

# Improved earthquake detection as a probe for active fault structures in New Zealand's central Southern Alps

By

Calum J. Chamberlain

A thesis  
submitted to Victoria University of Wellington  
in fulfilment of the requirements for the degree of  
Doctor of Philosophy in Geophysics.

Victoria University Wellington  
2016

---

---

*“On and on, like dominoes. The more you experimented the less you really knew. ”*

Simon Ings, *The weight of numbers.*

*“Correlation is everything. Did you know it is an established fact that hemlines tend to rise in times of national crisis?”*

Terry Pratchett, *Making Money.*

*“The tremors underground are starting to surround us. . . ”*

Avalanche City (Dave Baxter), *Fault Lines.*

---

# Abstract

This thesis concerns the detection and analysis of micro-seismicity and low-frequency earthquakes in New Zealand’s central Southern Alps. We make use of the 6.5 year continuous seismic dataset collected using the Southern Alps Microearthquake Borehole Array (SAMBA), alongside other temporary and permanent seismic deployments nearby. The small station spacing of this deployment allows for high resolution seismic studies near the Alpine Fault, a dextral-transpressive plate boundary fault between the Pacific and Australian plates.

Using this dataset we have documented the first evidence of low-frequency earthquakes on or near the deep extent of the Alpine Fault. By using a network based cross-correlation detection method we have generated a 3 year catalogue of 14 low-frequency earthquake families. These low-frequency earthquake families locate close to other indicators and models of the deep extent of the Alpine Fault, and we interpret these low-frequency earthquakes to represent shear failure on or near the deep extent of the Alpine Fault. These low-frequency earthquakes highlight a near-continuous background rate of deformation, punctuated by short periods of tremor. We also observe higher rates of low-frequency earthquake generation after large regional earthquakes. The magnitudes of our low-frequency earthquakes range from  $M_L$ –0.8–1.8, and appear to follow an exponential distribution, implying that there might be a characteristic length-scale of failure.

We have extended the catalogue of low-frequency earthquake templates using the full 6.5 year dataset and an objective synthetic detection methodology. We developed a new methodology for template detection after other methods failed, or were not feasible. This method employs simple synthetic template events, which, rather than trying to capture all of the complexities of the body waves we try to detect, approximate a simple waveform that does not correlate well with background noise. To undertake this method we have developed a multi-parallel Python package, which is highly portable (we have run this on computers ranging from dual-core, 8 GB RAM laptops to a 393 node, 6349 CPU cluster computer) and distributed via an open-source model. This package was run through the 6.5 year dataset on the New Zealand E-Science PAN cluster to efficiently (<48 hours clock-time) generate a spatially and temporally continuous catalogue of low-frequency earthquake templates. Using this method to detect

---

an initial suite of over 25,000 detections grouped into 600 families we have generated 600 good quality, discrete stacked waveforms for use in further matched-filter detection routines. We have shown that, for templates with both P and S-phase picks, these templates locate near to our previously determined low-frequency earthquake family locations.

Using a network matched-filter detection technique we have generated a catalogue of micro-seismicity in a region of low-seismicity near the Whataroa Valley, motivated by the Deep-Fault Drilling Project; Phase-2. We detected 300 earthquakes that include a selection of near-repeating earthquakes. We find that most detected events are not similar enough to be termed repeating. For 106 earthquakes we are able to generate high-precision magnitudes calculated by singular-value decomposition of similar waveforms. We find a high b-value of 1.44 for these earthquakes, with no earthquakes above  $M_L 1.6$ . By generating high precision cross-correlation derived picks for individual detections and employing a double-difference location methodology we show that seismicity does not delineate a single structure; rather we interpret the detected seismicity as temporally-limited earthquake sequences on small asperities adjacent to the Alpine Fault. Focal mechanisms for the best recorded events show dominantly strike-slip mechanisms, with lesser reverse and normal components.

During the drilling of the Deep-Fault Drilling Project: Phase-2 borehole we operated a real-time earthquake detection system around the drill-site. This was a multi-national effort involving 16 seismologists in three countries monitoring the automatic detections in shifts. During the 5 month real-time monitoring period we detected and located 493 earthquakes, none of which occurred within 3 km of the drill-site, nor required changes to the drilling operations. We undertook this monitoring using open-source software, which employed a standard energy based detection scheme.

This thesis has contributed four complementary earthquake catalogues, a further three years of continuous seismic data from the central Southern Alps, and an open-source Python package for detection and analysis of earthquakes using cross-correlation techniques. The characteristics of these catalogues highlight deformation modes on and near one of the world's major strike-slip plate boundaries, both at depth, and at the upper extent of the seismogenic zone.

# Acknowledgements

I am incredibly grateful for the opportunities and support given to me by my supervisors, John Townsend and Tim Stern. Thanks also to David Shelly who provided essential advice, experience and enthusiasm for my little project. Thanks to Cliff Thurber and Steve Roecker for letting me assist in WIZARD work and for providing data and helpful advice. I am especially indebted to Carolin Boese for the thorough work she undertook in deploying the SAMBA network, and for her advice throughout my PhD.

Thanks to Huw Horgan and Simon Cox for including me in their work and allowing me to learn from them. Also thanks to Martha Savage and Euan Smith at VUW who have both provided invaluable advice at different stages of my PhD. My thanks also to the staff at Victoria University, particularly Aleksandr Beliaev for saving my skin when my hard-drive stopped working, and for his invaluable contribution to our real-time earthquake monitoring.

I'm very grateful for the support I have received financially from Victoria University of Wellington through the Doctoral Scholarship, the Earthquake Commission of New Zealand, and the GeoScience Society of New Zealand through the Jim Ansell Scholarship. Thanks also to the New Zealand eScience Infrastructure for providing me with time on their PAN cluster computer.

I am extremely grateful to those who helped me in the field, I hope that they enjoyed their time exploring the Southern Alps. Adrian Benson has been a key member of field trips, and has provided boundless knowledge, both technical and general, to aid this project. I certainly would not have learnt as much as I did about seismometers and Python without him. Thanks to our pilots in the Southern Alps for getting us to some tricky places and keeping us safe. Thanks to the people of the West Coast for the interest shown and their accommodating nature.

My family and friends have been a great support to me throughout my PhD. My parents enthusiasm for the outdoors and New Zealand got me into this, and their support has helped me get this far. Thanks also to my big sister Katy, who has been the ideal role-model and a great friend throughout my PhD. Thanks to Andy and Owen for being interested in my work over beers and on long bike rides.

My office-mates and fellow graduate students, Emily, Cécile, Jesse, Dom, Katrin,

---

Laura-May, Chet, Gabe, Rachel, Katrina, Adrian, Jenni, Shaun, Rich and Juliet (and others that I may have missed) have been great throughout the PhD, both for rock chats, and general social drinks. Thanks to Cécile (Cecil) specifically for teaching me about fractures, Katrina for helping with magnitudes and earthquake sequences, and Chet for being interested in EQcorrscan. Thanks to Laura-May Baratin for finding faults in my work, and doing things better than me; I hope your PhD goes swimmingly and that I can help you along the way.

Finally thanks Emily for making the days full of sun and the nights full of moreporks, thank you for questioning me, helping me and laughing with me. I can't wait for our next adventure.

# Contents

<b>Abstract</b>	<b>i</b>
<b>Acknowledgements</b>	<b>iii</b>
<b>Contents</b>	<b>v</b>
<b>List of Figures</b>	<b>xi</b>
<b>List of Tables</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation and objectives . . . . .	1
1.2 Tectonic setting . . . . .	6
1.3 Geology of the Alpine Fault and central Southern Alps . . . . .	10
1.4 Seismicity in the central Southern Alps . . . . .	11
1.5 Low-frequency earthquakes in a global context . . . . .	16
1.6 Thesis structure . . . . .	24
<b>2 Research methodology</b>	<b>25</b>
2.1 Data acquisition and processing . . . . .	25
2.1.1 SAMBA:The Southern Alps Micro-earthquake Borehole Array . . . . .	26
2.1.2 Data processing . . . . .	28
2.1.3 Data quality . . . . .	29
2.2 Event detection . . . . .	31
2.2.1 Earthquake detection . . . . .	31
2.2.2 Tremor detection . . . . .	40
2.2.3 Low-frequency earthquake Detection . . . . .	40
2.3 Location methods . . . . .	47
2.3.1 Linear methods . . . . .	48
2.3.2 Non-linear earthquake location . . . . .	50
2.3.3 Relative locations . . . . .	51
2.4 Magnitude estimation . . . . .	53

<b>3</b>	<b>A low-frequency earthquake catalogue for the central Southern Alps</b>	<b>55</b>
3.1	Abstract . . . . .	55
3.2	Introduction . . . . .	56
3.3	Methodology . . . . .	58
3.3.1	LFE identification and detection . . . . .	58
3.3.2	Magnitude calculation . . . . .	64
3.3.3	Event verification . . . . .	64
3.4	Observations . . . . .	66
3.5	Discussion . . . . .	72
3.5.1	LFE family locations . . . . .	72
3.5.2	Inter-event timing and magnitudes . . . . .	74
3.6	Conclusions . . . . .	75
3.7	Amendum . . . . .	76
<b>4</b>	<b>LFE template construction using synthetic waveforms</b>	<b>77</b>
4.1	Abstract . . . . .	77
4.2	Introduction . . . . .	78
4.3	Methods . . . . .	79
4.3.1	Template simulation . . . . .	79
4.3.2	Comparison with known waveforms . . . . .	80
4.3.3	A grid of templates . . . . .	87
4.3.4	Large-scale multi-parallel cross-correlation optimisation . . . . .	89
4.4	Results . . . . .	92
4.5	Discussion and conclusions . . . . .	97
4.5.1	Methodological extensions . . . . .	99
<b>5</b>	<b>High-precision micro-seismic detection adjacent to the Alpine Fault</b>	<b>101</b>
5.1	Abstract . . . . .	101
5.2	Introduction . . . . .	101
5.3	Data . . . . .	106
5.4	Methods . . . . .	107
5.4.1	Detection . . . . .	107
5.4.2	Correlation pick-determination . . . . .	109
5.4.3	Event location . . . . .	110
5.4.4	Repeater analysis . . . . .	110
5.4.5	SVD magnitude estimation . . . . .	111
5.4.6	Focal mechanism inversion . . . . .	114
5.5	Results . . . . .	114
5.6	Discussion . . . . .	118
5.6.1	Temporal evolution . . . . .	118

---

5.6.2	Locations . . . . .	122
5.6.3	Repeating earthquakes . . . . .	124
5.6.4	Secondary phases . . . . .	124
5.7	Conclusions . . . . .	128
5.8	Acknowledgements . . . . .	128
<b>6</b>	<b>DFDP real-time seismic monitoring</b>	<b>129</b>
6.1	Abstract . . . . .	129
6.2	Introduction . . . . .	130
6.2.1	Deep-Fault Drilling Project . . . . .	131
6.2.2	Seismicity around Whataroa . . . . .	134
6.3	Data and personnel . . . . .	135
6.4	Methods . . . . .	137
6.4.1	Earthquake detection . . . . .	137
6.4.2	Response to earthquakes . . . . .	138
6.4.3	Magnitude calculation . . . . .	140
6.4.4	Planned response to large or close earthquakes . . . . .	141
6.5	Results . . . . .	141
6.5.1	Malcolm Peak earthquake 22/10/2014 . . . . .	144
6.5.2	Drilling noise detected seismically . . . . .	146
6.5.3	Overall performance . . . . .	148
6.6	Conclusions . . . . .	148
6.6.1	Improvements . . . . .	149
<b>7</b>	<b>Discussion</b>	<b>151</b>
7.1	Summary . . . . .	151
7.1.1	Low-frequency earthquakes and slow deformation . . . . .	153
7.1.2	Micro-earthquake studies . . . . .	156
7.2	Opportunities for further investigation . . . . .	156
7.2.1	Low-frequency earthquakes and slow deformation . . . . .	156
7.2.2	Micro-earthquake studies . . . . .	161
7.2.3	Tasman glacier slip . . . . .	165
7.3	Concluding statement . . . . .	165
	<b>Appendices</b>	<b>167</b>
<b>A</b>	<b>Network set-up</b>	<b>169</b>
<b>B</b>	<b>Data continuity</b>	<b>173</b>
<b>C</b>	<b>Common noise characteristics of SAMBA data</b>	<b>175</b>

---

<b>D LFE event verification</b>	<b>179</b>
<b>E Seisan training manual</b>	<b>181</b>
E.1 An Introduction to Seisan . . . . .	181
E.2 Data download . . . . .	183
E.3 Pick data . . . . .	184
E.4 Further information . . . . .	187
<b>Bibliography</b>	<b>189</b>

# List of Figures

1.1	National seismograph network completeness in the central Southern Alps	2
1.2	Tremor and low-frequency earthquake publications per year . . . . .	3
1.3	New Zealand seismic hazard model . . . . .	4
1.4	New Zealand background seismicity . . . . .	5
1.5	Basement geology of New Zealand . . . . .	8
1.6	Alpine Fault schematic cross-section . . . . .	9
1.7	Background seismicity in the Southern Alps . . . . .	12
1.8	Along-strike changes in seismicity near the Alpine Fault . . . . .	15
1.9	LFEs compared to earthquakes . . . . .	17
1.10	Slow-earthquake overview . . . . .	17
1.11	Schematic model of stress transfer from depth to the seismogenic zone .	18
1.12	Earthquake magnitude-duration scales . . . . .	19
1.13	Low-frequency earthquake schematic cartoon for the Alpine Fault . . .	21
1.14	Cascade and preslip nucleation schematic cartoon . . . . .	23
2.1	Network map of passive seismic sites used . . . . .	26
2.2	Probabilistic power spectral densities of deployed seismic sites . . . . .	30
2.3	Earthquake detection using energy detection . . . . .	32
2.4	Single-channel cross-correlation earthquake doublet detection . . . . .	33
2.5	Network cross-correlation earthquake doublet detection . . . . .	35
2.6	Hidden extra earthquakes detected by network cross-correlation . . . . .	36
2.7	Waveform envelope cross-correlation tremor detection . . . . .	41
2.8	Low-frequency earthquake template and detected waveforms . . . . .	43
2.9	Cross-correlation for low-frequency earthquake detection . . . . .	44
2.10	Network response (brightness) for low-frequency earthquake detection .	46
2.11	Phase pick refinement by cross-correlation . . . . .	52
3.1	Station map for low-frequency earthquake detection . . . . .	57
3.2	Low-frequency earthquake template and detected waveforms . . . . .	60
3.3	Low-frequency earthquake family locations . . . . .	62
3.4	Tremor waveform with highlighted low-frequency earthquakes . . . . .	63
3.5	All low-frequency earthquakes detections and magnitudes with time . .	65

---

3.6	Low-frequency earthquake inter-event time distribution . . . . .	67
3.7	Low-frequency earthquake magnitudes within tremor . . . . .	68
3.8	Low-frequency earthquake magnitudes during the June 2011 Christchurch earthquake sequence . . . . .	69
3.9	Low-frequency earthquake detections with time by family . . . . .	70
3.10	Magnitude and amplitude distributions for low-frequency earthquakes .	71
4.1	Synthetic template generation . . . . .	79
4.2	Single-channel synthetic matching . . . . .	80
4.3	Multi-channel comparison of synthetic and real data . . . . .	82
4.4	Comparison of spectrograms for synthetic and real earthquakes . . . . .	83
4.5	Synthetic and real network cross-correlation sum . . . . .	85
4.6	Comparison of detections made by real and synthetic templates . . . . .	86
4.7	Synthetic compared to raw LFE waveform . . . . .	88
4.8	Comparison of spectrograms for a synthetic and real LFE . . . . .	89
4.9	Map of grid used for template generation . . . . .	90
4.10	Cumulative LFE detections using synthetic templates . . . . .	93
4.11	Example of detection . . . . .	94
4.12	Stack of detections made using a synthetic template . . . . .	96
4.13	Initial 1D LFE template location . . . . .	98
5.1	Background Seismicity in the central Southern Alps . . . . .	103
5.2	Data continuity of seismic stations used for detection . . . . .	104
5.3	Template locations for matched-filter detection . . . . .	106
5.4	Template waveform in continuous data . . . . .	108
5.5	Singular-value decomposition matching for magnitudes. . . . .	112
5.6	Temporal variations in detection frequency . . . . .	116
5.7	Template-detection inter-event times . . . . .	117
5.7	Template-detection inter-event times . . . . .	118
5.8	Relocated seismicity, Whataroa . . . . .	119
5.9	Focal mechanisms plotted with P-T axes and picks . . . . .	120
5.10	Whataroa focal mechanism map . . . . .	121
5.11	Magnitude-frequency relations calculated by singular-value decomposition	121
5.12	Similarity plot at GCSZ for near-repeating earthquakes . . . . .	122
5.13	Secondary phases on multiple detected events at GCSZ . . . . .	123
5.14	Secondary S phase detected on multiple sites. . . . .	125
5.15	Spectral plots of secondary phases . . . . .	126
5.16	Multiple S phases for multiple detected families . . . . .	127
6.1	Background seismicity around Whataroa . . . . .	133
6.2	Traffic-light response scheme . . . . .	134

---

---

6.3	Real-time processing flow . . . . .	139
6.4	Earthquake locations from real-time monitoring . . . . .	142
6.5	Magnitude-frequency plot for real-time seismicity . . . . .	143
6.6	Earthquakes detected in time and space . . . . .	145
6.7	Malcolm Peak earthquake sequence . . . . .	146
6.8	Drill drop recorded seismically . . . . .	147
7.1	Schematic cartoon of LFEs in the Southern Alps . . . . .	154
7.2	Triggered LFEs by a Christchurch earthquake . . . . .	157
7.3	Wanaka earthquake sequence . . . . .	163
7.4	Seismicity from a Tasman glacier speed-up event . . . . .	164
B.1	SAMBA continuity . . . . .	174
C.1	Lightning noise . . . . .	176
C.2	Acasual noise . . . . .	177
C.3	Solar induced spiking . . . . .	178
D.1	LFE threshold testing . . . . .	180



# List of Tables

2.1	SAMBA site locations . . . . .	27
3.1	LFE family information . . . . .	59
6.1	Real-time seismic site locations . . . . .	136
6.2	FilterPicker parameter . . . . .	138
6.3	Real-time velocity model . . . . .	138
A.1	SAMBA inventory . . . . .	171



# 1 Introduction

## 1.1 Motivation and objectives

Earthquakes are often the most obvious manifestation of dynamic crustal deformation. How we interpret seismogenic processes is, however, often limited by our ability to detect earthquakes across a range of magnitudes: large earthquakes ( $M > 7$ ) are poorly sampled due to large interseismic periods, and small earthquakes ( $M < 2.0$ , GeoNet magnitude of completeness ( $M_C$ ) in the central Southern Alps is  $\sim 2.3$ , Figure 1.1) are difficult to detect due to their low rupture energies, and the attenuating properties of the Earth resulting in low recorded amplitudes.

While little can be done to detect more large earthquakes, beyond recording for longer durations, understanding the significance of small earthquakes and their role in the seismic cycle can be improved. In the shallow crust, frequent small earthquakes contribute to the redistribution of elastic stress, and the triggering of subsequent events [Marsan, 2005]. In some cases, the spatial distribution of small earthquakes prior to large earthquakes has been observed to outline the region of maximum coseismic-slip [Haberman, 1981]. Temporal distributions of small earthquakes sometimes change with proximity to large earthquakes, however the significance of these observations is often questioned [e.g. Reasenber and Matthews, 1988, Shaw et al., 1992]. If we can understand how these more frequent and less destructive earthquakes influence, and are influenced by, the seismic cycle of large earthquakes, we may hope to find predictive models for large earthquakes.

In recent years the detection of slow earthquakes has offered insight into how faults are loaded, and provided a mechanism for the transferral of stress from deep-ductile regions to shallow seismogenic faults [e.g. Peng and Gomberg, 2010, Wech and Creager, 2011]. The implications of these discoveries are manifold; one such implication is that recent observations of slow-earthquakes prior to large earthquakes imply a preliminary phase of nucleation loading the fault patches that subsequently fail in large earthquakes [Kato et al., 2012, Ito et al., 2013, 2015, Ruiz et al., 2014]. These implications are discussed further in Section 1.5.

Interest in these slow-earthquakes has increased recently (Figure 1.2), and alongside

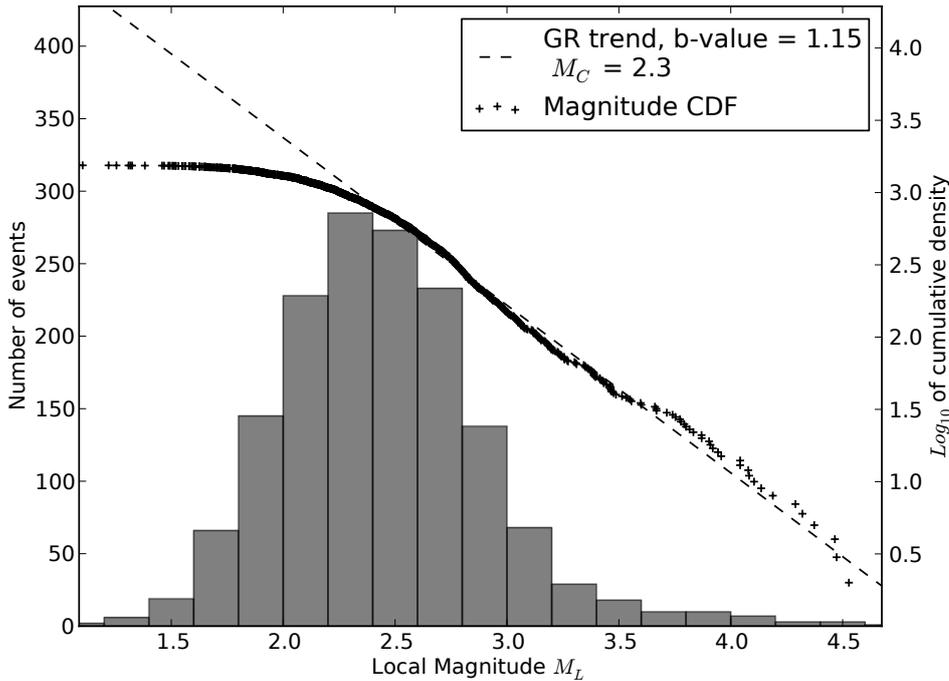


Figure 1.1: Magnitude histogram of GeoNet events detected and located between January 2000 and September 2015 in the central Southern Alps. An approximate Gutenberg-Richter fit is plotted alongside the cumulative density function (CDF) (crosses) for a  $M_C$  of 2.3 and b-value of 1.15 (fitted by linear least-square). Earthquakes plotted are located in a region bounded by the box:  $-44.13688$ – $43.11903$ ,  $169.70856$ – $170.91980$ .

this interest a range of specialised detection methods have been developed to better characterise these events [e.g. Shelly et al., 2009, Kao et al., 2009, Wech and Creager, 2008, Brown et al., 2013, Ghosh et al., 2009, Rubin and Armbruster, 2013, Frank and Shapiro, 2014]. These detection methods can often be applied to the detection of regular shallow micro-seismicity. The application of these detection techniques to a range of seismic signals on and near the Alpine Fault forms the core of this thesis.

Large earthquakes on New Zealand’s Alpine Fault have been identified by Stirling et al. [2012] as the primary seismic hazard faced in southern New Zealand (Figure 1.3), despite the current low rate of large earthquakes on the central portion of the Alpine Fault (Figure 1.4). This hazard assessment is based in part on the paleoseismic record of proposed large earthquakes on the Alpine Fault [Berryman et al., 2012]. This paleoseismic record demonstrates the regularity of likely large Alpine Fault earthquakes, with the last thought to have occurred in c. 1717 AD and to have exceeded M 8.0 [Wells et al., 1999]. This would put the Alpine Fault late in the seismic cycle at present, with the probability of a M 8 earthquake in the next 50 years of 27% [Biasi et al., 2015].

This thesis addresses small earthquakes and slow earthquakes on and near to New Zealand’s central Alpine Fault. To facilitate these investigations 6.5 years of continuous seismic data collected on the SAMBA (Southern Alps Microearthquake Borehole

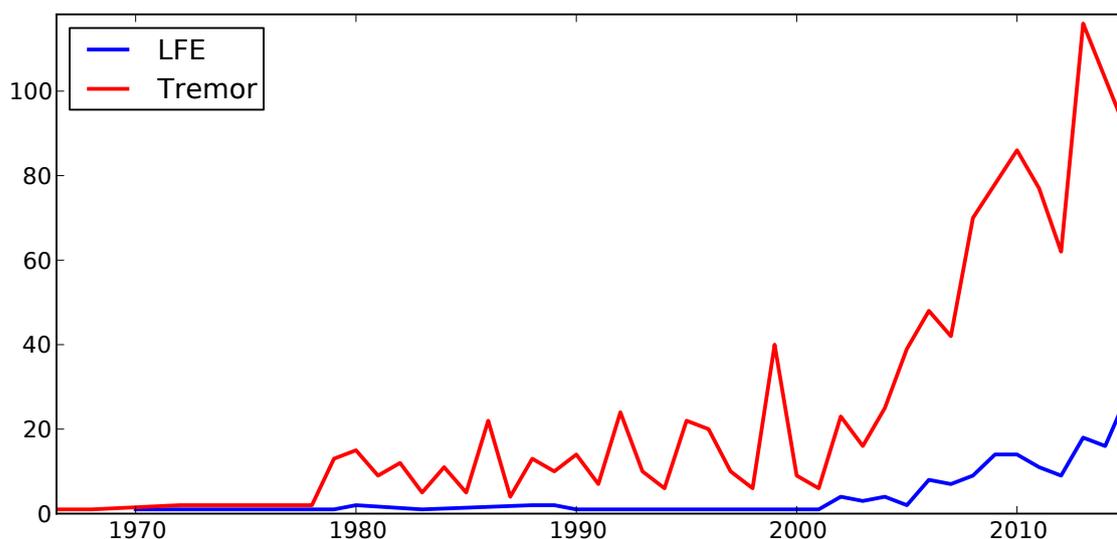


Figure 1.2: Publications per year with keywords, titles or abstracts including the terms tremor or slow slip or episodic tremor and slip, and LFE or low-frequency earthquake or low frequency earthquake in earth sciences respectively (source: <http://www.scopus.com>, accessed 21 September 2015<sup>1</sup>).

Array) network [Boese, 2012] have been analysed alongside data from other nearby temporary networks and the permanent national seismograph network, GeoNet. Information regarding the deployments used in this thesis, and the processing steps applied to the data can be found in Sections 2.1 and 2.1.1.

The overarching objectives of this thesis are as follows:

1. Determine how deformation occurs beneath the central Southern Alps near the brittle-ductile transition zone through the study of low-frequency earthquakes (LFEs);
2. Assess temporal variations in slow-deformation on and near the deep extent of the central Southern Alps in relation to other phenomena that might trigger slip, and in relation to how slow-slip may impact on the shallow seismogenic zone;
3. Determine levels of seismicity in the ‘seismic gap’ between Whataroa and Wanganui valleys prior to and during drilling of the Deep Fault Drilling Project, phase 2;
4. Reduce the magnitude of completeness for seismic catalogues in the central Southern Alps using match-filter detection techniques.

To achieve these objectives, the following tasks have been addressed:

<sup>1</sup>Full search strings are: ((TITLE-ABS-KEY("tremor") OR TITLE-ABS-KEY("slow slip") OR TITLE-ABS-KEY("episodic tremor and slip")) AND NOT (TITLE-ABS-KEY(volcanic)) AND (LIMIT-TO(SUBJAREA,"EART"))), and ((TITLE-ABS-KEY("LFE") OR TITLE-ABS-KEY("low-frequency earthquake") OR TITLE-ABS-KEY("low frequency earthquake")) AND NOT (TITLE-ABS-KEY(volcanic)) AND (LIMIT-TO(SUBJAREA,"EART"))).

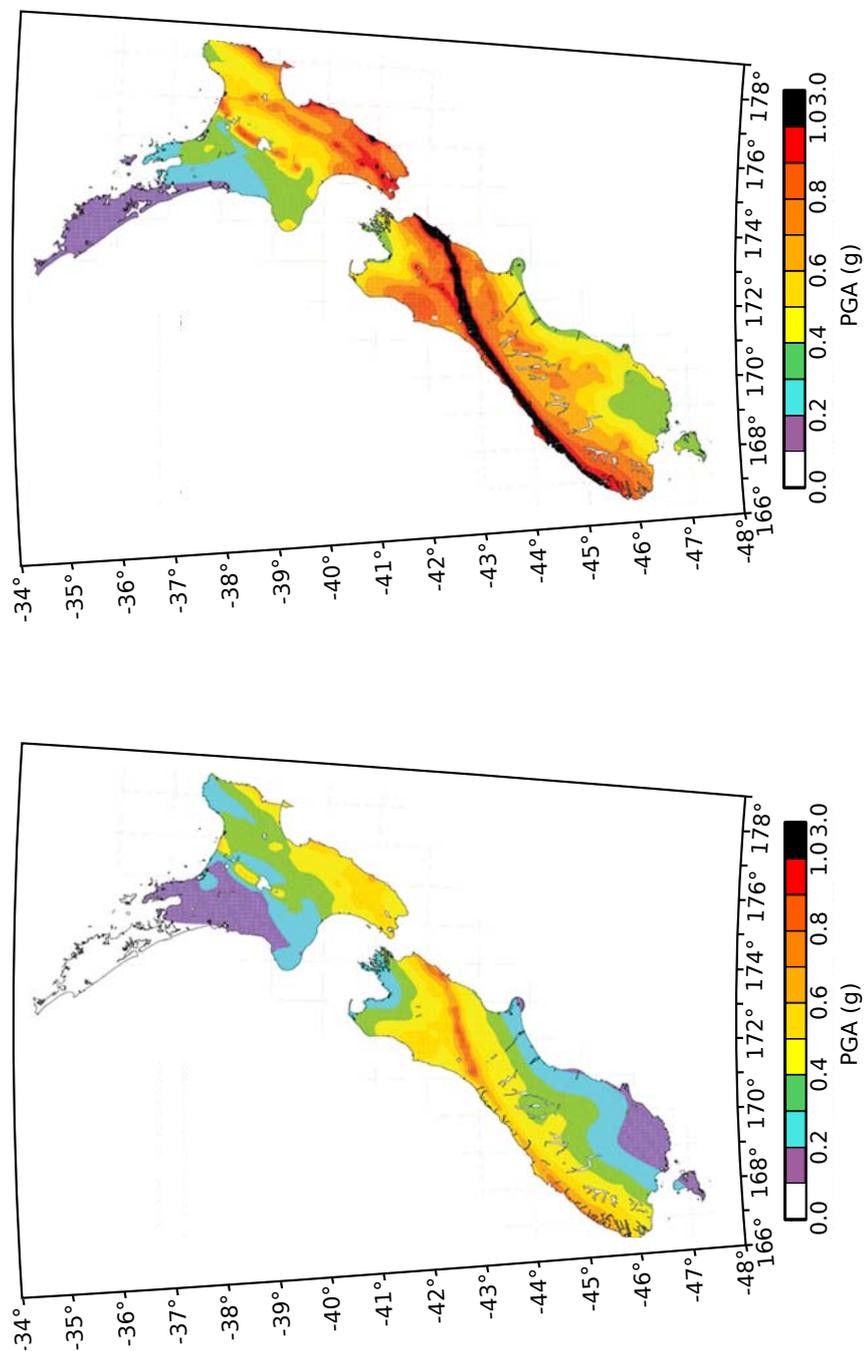


Figure 1.3: New Zealand seismic hazard model after Stirling et al. [2012]. Left panel shows the 475 year return period shallow soil PGA(g). Right panel shows the 2500 year shallow soil PGA(g). These models have probabilities of exceedence in 50 years of approximately 10 % and 2 % respectively. The most obvious feature in both maps is the zone of high PGA adjacent to the Alpine Fault.

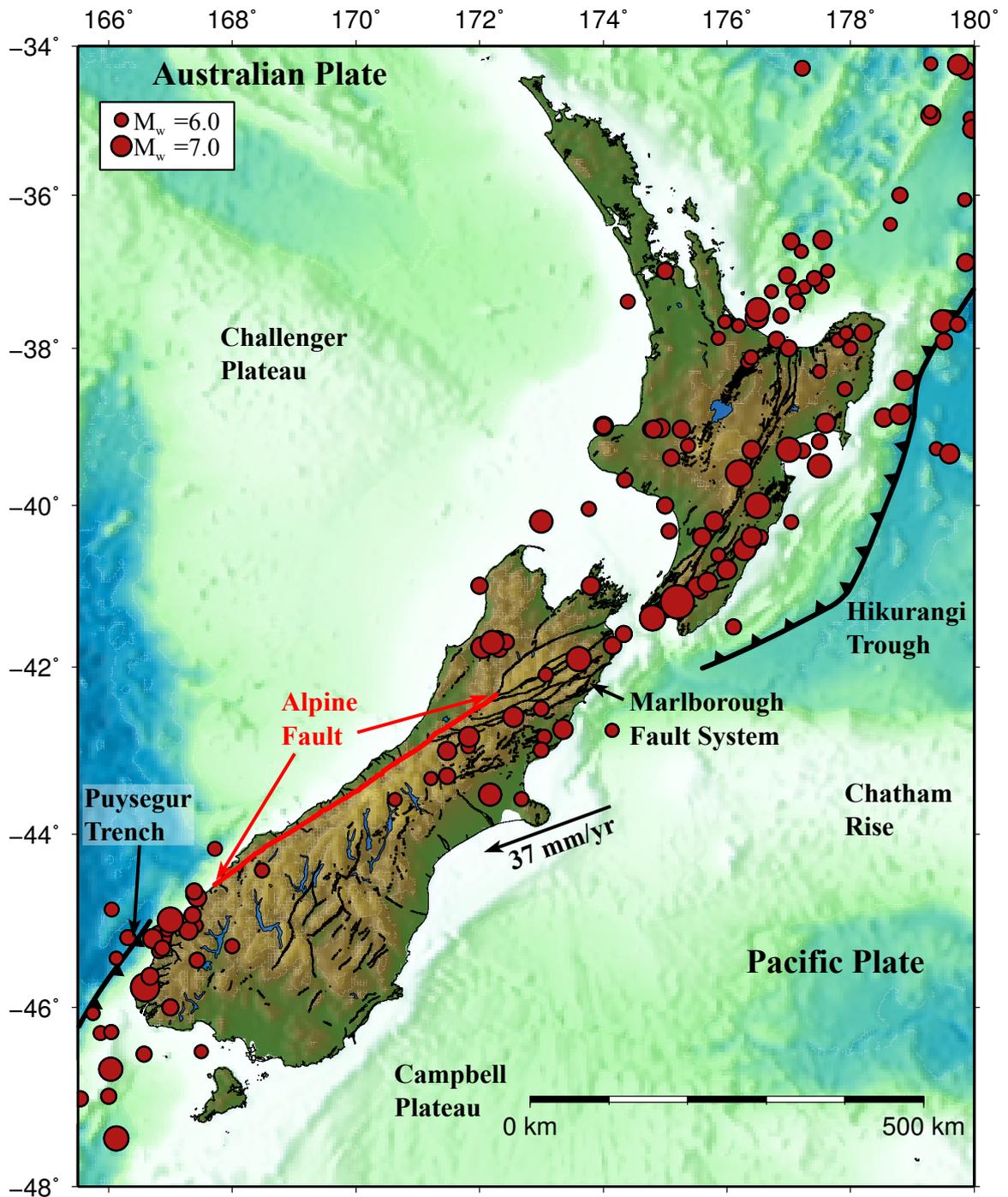


Figure 1.4: All events  $M_W > 6.0$  in the GeoNet national catalogue since 1815, as of 18 June 2015. The Alpine Fault is notable predominantly for the lack of large earthquakes along its length.

- Manually identify LFEs within tremor signal previously identified by Wech et al. [2012] (objective 1);
- Detect further LFEs using a match-filter detection technique to develop a continuous catalogue (objective 1);
- Generate a spatially and temporally continuous LFE catalogue using an automated template detection method coupled with match-filter detection and stacking (objective 2);
- Analyse temporal variations in the LFE catalogue and compare these to large regional earthquakes, and teleseismic events (objective 2);
- Generate magnitudes for LFEs and relate this to deformation rate and loading of the Alpine Fault (objective 2);
- Apply real-time detection methods to a small region around the Whataroa Valley during DFDP-2 drilling (objective 3);
- Generate templates and construct a catalogue of seismicity in the low-seismicity zone between the Whataroa and Wanganui valleys using match-filter detection (objective 4);
- Obtain relative locations of this matched-filter derived catalogue using high-precision cross-correlation derived picks (objective 4).

The work presented in this thesis has been conducted under the auspices of three Marsden funds awarded by the Royal Society of New Zealand. The initial Marsden fund (‘Putting a stethoscope on the Alpine Fault’) funded the development of the SAMBA seismic network and the continuation of detailed passive monitoring of Alpine Fault seismicity. The second of these Alpine Fault seismicity specific funds (‘Locked and loaded? Effects of deep seismic and aseismic deformation on Alpine Fault earthquakes’) has since enabled continued research into more esoteric seismic signals, specifically LFEs, with the goal of determining how slow tectonic deformation is loading the Alpine Fault towards an eventual large earthquake. Seismic monitoring associated with the Deep Fault Drilling Project was funded by a separate series of funds, including Marsden, the Earthquake Commission, Victoria University of Wellington, University of Otago, International Continental Scientific Drilling Project and GNS Science.

## 1.2 Tectonic setting

New Zealand’s present day structure is due to deformational and erosional processes dominated by deformation at and adjacent to the Australia-Pacific plate boundary

[Walcott, 1978]. Beneath the North Island the Pacific plate is subducted under the Australian plate at the Hikurangi subduction zone. To the south, interplate motion is transferred through the Marlborough Fault Zone to the Alpine Fault which runs for 650 km through South Island [Berryman et al., 1992] accommodating c. 75% of the total strike-slip inter-plate motion and up to 100% of the compression in the central portion of the fault [Norris and Cooper, 2000]. In the south of South Island, interplate motion occurs on the steeply dipping Puysegur subduction zone, where the Australian plate is subducted beneath the Pacific plate.

Deformation in the region of the Alpine Fault was dominated by extension to (what is now) the South between 40–38 Ma, before becoming dominantly strike-slip through the late Oligocene (29–24 Ma), with an increasing compressional component through time until the current transpressional regime [Sutherland, 1996]. Sutherland et al. [2000] favoured the formation of a transpressional Alpine Fault from around 25 Ma. These authors suggested that the location of the Alpine Fault is controlled by a previous, Eocene age, passive margin, that formed a separation between the continental lithosphere of the Challenger Plateau, and younger oceanic lithosphere, that has since been partially subducted beneath the Alpine Fault. The position and rheology of this material has key implications for the seismicity at depth beneath South Island [e.g. Boese et al., 2013, Lamb et al., 2015].

Total strike-slip displacement on the Alpine Fault is recorded by offset of the Dun Mountain Ophiolite Belt (Figure 1.5), giving c. 460 km dextral offset [Sutherland et al., 2000]. Plate reconstructions, however, estimate total dextral displacement between western North Island and eastern South Island of c. 850 km [Molnar et al., 1999]. The remaining c. 390 km of dextral offset is likely accommodated by shear on other, smaller structures [Norris and Cooper, 2000].

The present-day Alpine Fault has an average strike of  $055^\circ$ ,  $16^\circ$  anti-clockwise from the current Pacific-Australia plate motion vector ( $37 \pm 2$  mm/yr at  $071 \pm 2^\circ$ ) [DeMets et al., 1994]. This results in a transpressive margin accommodating both reverse and dextral strike-slip motion. At the northern extent of the Alpine Fault, where strain is partitioned across the Marlborough Fault Zone and compression is accommodated by the southern extent of the Hikurangi subduction zone, the Alpine Fault dips steeply to the East [Cox and Sutherland, 2007]. Through the central Southern Alps, the Alpine Fault becomes more listric in nature, with dips of  $45\text{--}60^\circ$  in the mid-crust [Stern et al., 2007]. The Alpine Fault becomes near vertical in its southern extent, with occasional instances of normal motion as compression is taken up by the Puysegur Trench [Sutherland et al., 2007].

The central Alpine Fault displays extensive surface segmentation, with strike-slip dominated sections striking at  $065\text{--}090^\circ$ , and thrust-dominated segments of strike  $020\text{--}$

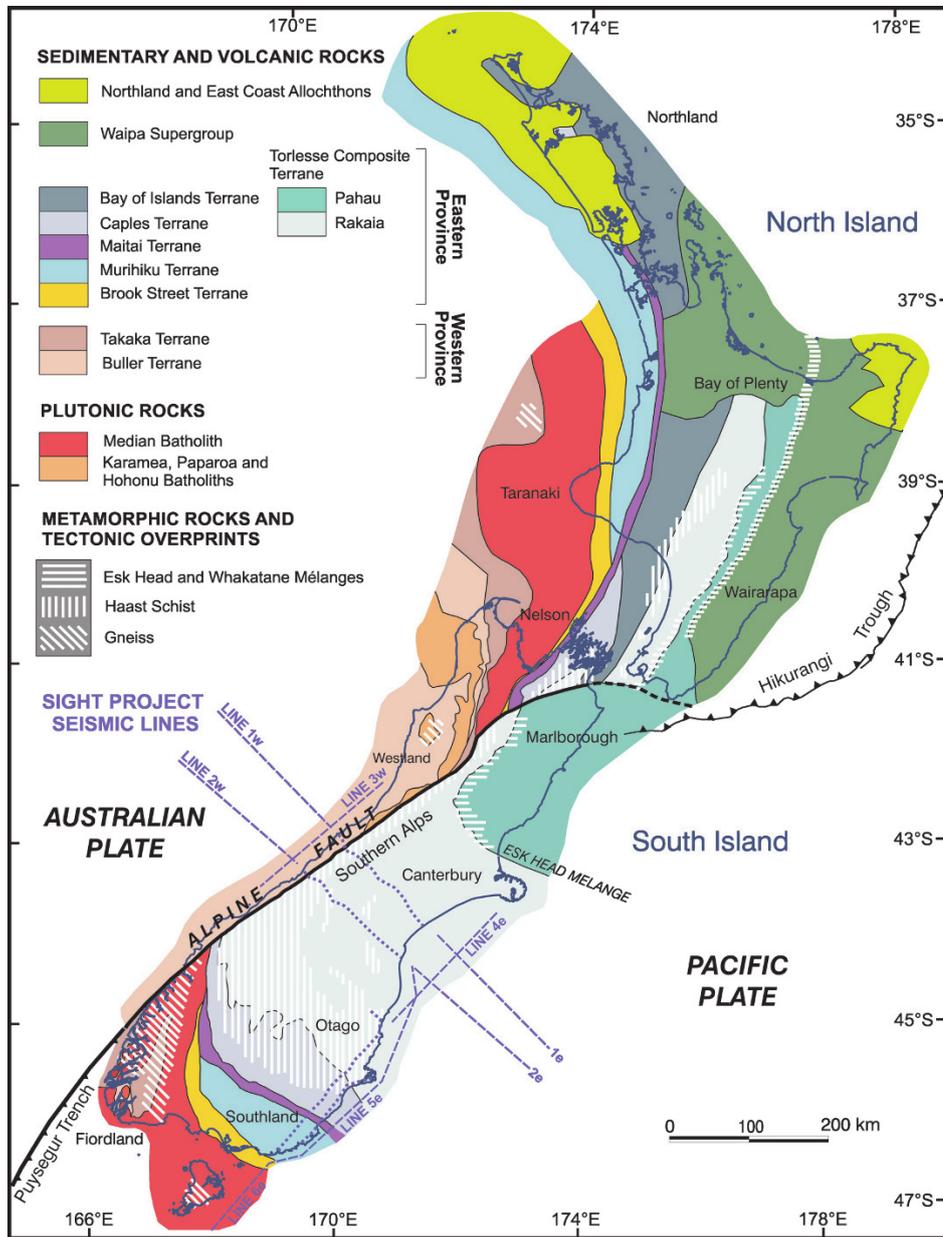


Figure 1.5: Basement terranes of New Zealand. From [Cox and Sutherland, 2007] after [Mortimer, 2004]. The Alpine Fault clearly offsets the Maitai, and other Terranes from Fiordland to northern South Island.

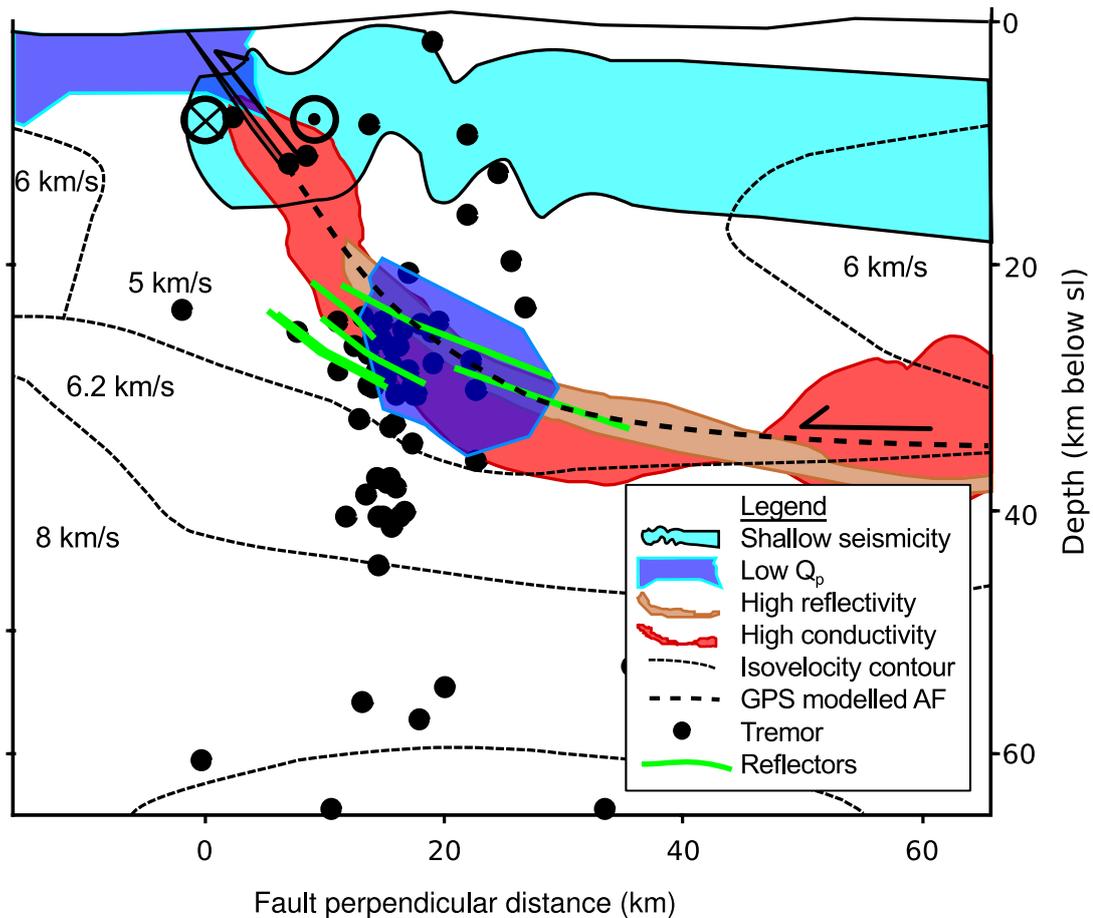


Figure 1.6: Summary cartoon of the possible deep structure of the Alpine Fault at approximately  $-43.6^\circ$  latitude. As plotted, the region of shallow seismicity is taken from the locations of Boese et al. [2012], the tremor locations are from Wech et al. [2012], zone of high reflectivity, reflectors and velocity model from Stern et al. [2007], zone of high conductivity from Wannamaker et al. [2002], regions of low  $Q_p$  from Eberhart-Phillips et al. [2008], GPS fault from Lamb and Smith [2013]. The listric fault foundering at 30–40 km depth is imaged well by the active source seismic studies [Stern et al., 2007].

$050^\circ$  [Norris and Cooper, 1995, Barth et al., 2013]. This segmentation is thought to be a near-surface effect of topography, however it is not known at what depth these segments merge into a continuous, supposedly through-going, crustal-scale fault [Houlié and Stern, 2012]. At depth the fault dip appears to become shallower (Figure 1.6), as imaged by active seismic studies [Stern et al., 2007], modelled by GPS inversion [Lamb and Smith, 2013], and indicated by geological exhumation profiles [Koons et al., 2003, Herman et al., 2009, Norris and Toy, 2014].

## 1.3 Geology of the Alpine Fault and central Southern Alps

New Zealand's basement geology is made up of a series of terranes and igneous suites [Cox and Sutherland, 2007]. On a large scale the basement terranes are grouped into Eastern and Western Provinces, separated by the Median Batholith (Figure 1.5) [Mortimer, 2004]. The rocks of the Western Province form the footwall of the present day Alpine Fault, and are juxtaposed with Eastern Province rocks at the fault. The Western province is dominated by quartzose sediments (Buller Terrane), which have been metamorphosed and later intruded by plutonic rocks [Cox and Sutherland, 2007]. Rocks of the Eastern Province are dominantly lithic and feldspathic greywackes, at differing levels of metamorphism [Cox and Sutherland, 2007]. Alongside the Alpine Fault, and throughout the Southern Alps, the rocks are grouped into the Rakaia Terrane, which transitions from greywacke to schist and on to mylonites and fault rocks near the Alpine Fault.

On a transect from Canterbury towards the Alpine Fault, the rocks of the Rakaia Terrane become progressively more metamorphosed and deformed, transitioning from greywacke into foliated schists. The foliation of the schists are thought to be due to distributed pure-shear strain from thickening and rotation [Little et al., 2002]. Within 2 km of the Alpine Fault, schists transition into mylonites (which can be further broken down into protomylonites, mylonites and ultramylonites [Norris and Cooper, 2007, Toy et al., 2015]). These mylonites have foliations with an average dip of  $\sim 45^\circ$ , which has been inferred to indicate the dip of the Alpine Fault at depth [Sibson et al., 1979], although this may not be the case due to a lack of rotation of fabrics [Toy et al., 2013].

Adjacent to the principal slip zone of the Alpine Fault cataclastic rocks separate the fault zone gouge from the mylonites in the hanging-wall. In the hanging wall the cataclastic units are predominantly formed from hanging-wall protoliths, however there is some indication of a foot-wall (Pacific plate) source for some of the cataclastics [Toy et al., 2015]. Recent drilling of the DFDP-1 boreholes [Townend et al., 2009, Sutherland et al., 2012] at Gaunt Creek revealed that the fault zone gouges have relatively high coefficients of friction ( $\mu \sim 0.61\text{--}0.76$  for  $T=70\text{--}350^\circ\text{C}$  and normal stress= $31\text{--}156\text{ MPa}$  [Boulton et al., 2014] and  $\mu \sim 0.67$  for  $T=160^\circ\text{C}$  and normal stress= $100\text{ MPa}$  [Ikari et al., 2015b]), with lower  $\mu$  occurring in gouges containing smectite clays [Schleicher et al., 2015]. These frictional values are in stark contrast to the low values of  $\mu$  found at the Nankai Trough of  $\mu < 0.4$  [Saffer et al., 2012], and Parkfield ( $\mu = 0.1\text{--}0.15$ ) [Lockner et al., 2011], or on the southern portion of the Alpine Fault ( $\mu = 0.12\text{--}0.37$ ) [Barth et al., 2012].

DFDP-1 drilling also revealed that, despite the relatively high permeability of the schists that make up the Southern Alps (highlighted by meteoric water intrusion up to the brittle-ductile transition [Menzies et al., 2014] and the response of hot-springs to regional earthquakes [Cox et al., 2015]), the Alpine Fault itself forms a barrier to fluid flow [Sutherland et al., 2012]. A high geothermal gradient was also found of  $\sim 63^\circ\text{Ckm}^{-1}$  [Sutherland et al., 2012]. A higher geotherm has also been seen in the more recent DFDP-2 borehole (Rupert Sutherland, pers. comm.).

Recent work by Bourguignon et al. [2015] revealed a step in the seismogenic thickness across the Alpine Fault, which they suggested indicated that changes in temperature across the fault are an important control on the seismogenic depth near the Alpine fault. Despite this body of work, little is known about the footwall beneath seismogenic depths and how the footwall interacts with the Alpine Fault. The footwall will likely play an important role in Alpine Fault earthquakes, not least as a possible source of fluids which may weaken the fault allowing for tremor to occur [Wech et al., 2012].

## 1.4 Seismicity in the central Southern Alps

Despite high deformation rates, the central Alpine Fault displays sparse seismicity of low magnitude [Evison, 1971, Scholz et al., 1973, Caldwell and Frolich, 1975, Walcott, 1978, Boese et al., 2012, Bourguignon et al., 2015]. Seismicity in the central Southern Alps has been studied by a number of authors. Initially only the New Zealand National Seismograph network was used [e.g. Haines, 1981, Anderson and Webb, 1994], before additional small-scale, temporary deployments were added.

Evison [1971] provided a focused overview of Alpine Fault seismicity, using a sparse catalogue of moderate magnitude ( $M_L \leq 4$ ) earthquakes. He noted the lack of seismicity in the central Southern Alps, and suggested a link between this observation and the idea that seismic gaps on the San Andreas denote a high probability of major earthquakes. Following from this work Scholz et al. [1973] conducted a more detailed study of Alpine Fault seismicity using a small transportable network of temporary seismic stations. They documented diffuse seismicity throughout the region, apparently not associated with major faults.

Reyners [1988] analysed seismicity associated with the impounding of Lake Pukaki using a small network of nine seismometers, recording for over 8 years from 1975 to 1983. Reyners [1988] noted that despite low seismicity in the region surrounding Lake Pukaki (a glacial lake within the central Southern Alps, Figure 1.7), many low magnitude ( $M_L < 2.0$ ) earthquakes were recorded. Reyners [1988] recorded 2825 earthquakes, of which 2777 (98%) were shallower than 15 km, with a completeness value of  $M_C=1.8$ .

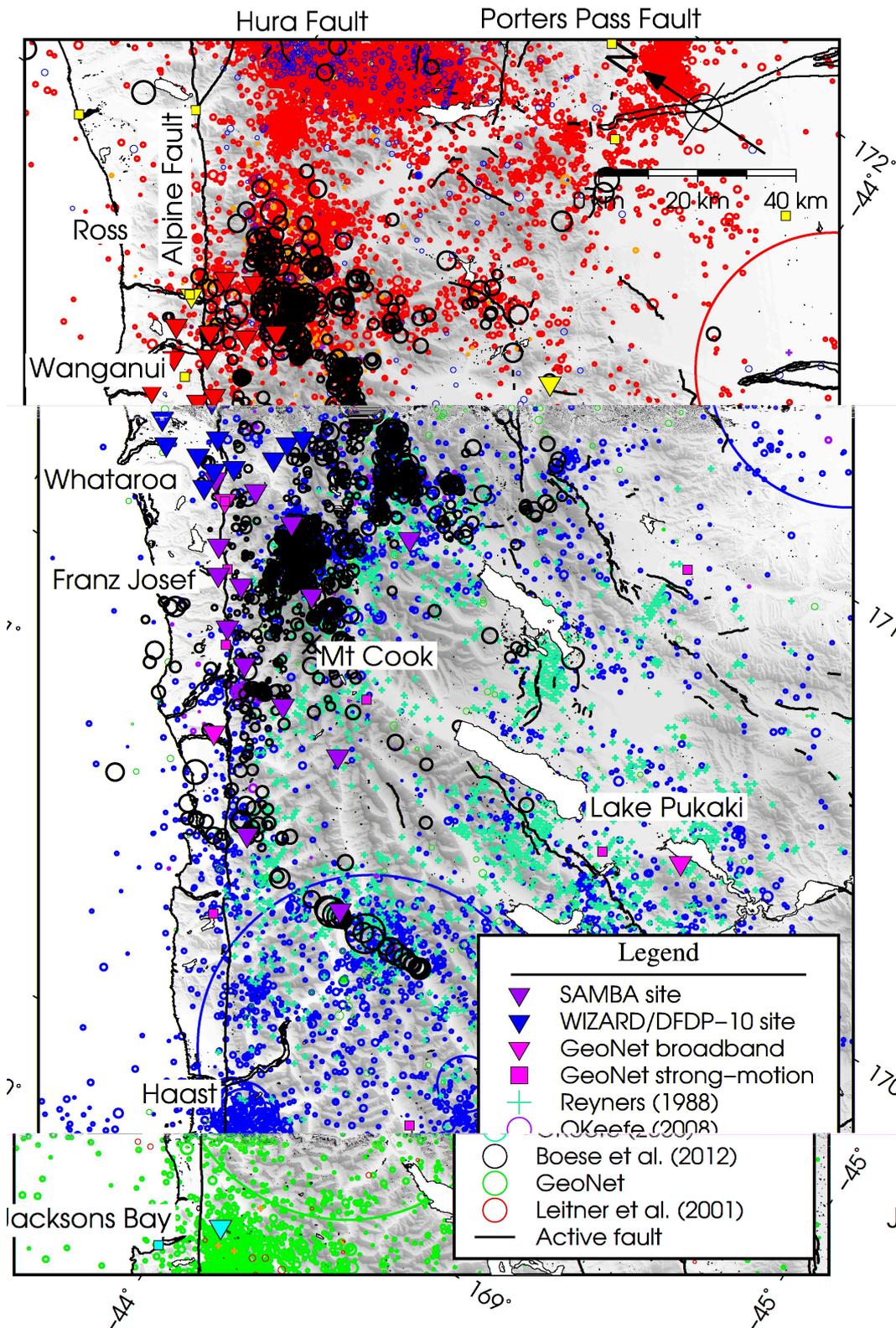


Figure 1.7: Seismicity in the Southern Alps from GeoNet (1990–2015), SAPSE [Leitner et al., 2001], Lake Pukaki seismic network [Reyners, 1988] and CIAF [O’Keefe, 2008]. All earthquakes are scaled by magnitude apart from the catalogue of Reyners [1988], for which magnitudes could not be extracted from the catalogue. The Hura Fault merges with other faults further north to form the Hope Fault (Faults and names from the New Zealand Active Fault Database (<http://data.gns.cri.nz/af/>)).

Leitner et al. [2001] studied the central Southern Alps using a more extensive seismic network of 40 temporary stations (Southern Alps Passive Seismic Experiment (SAPSE)) alongside 15 permanent, national network seismograph stations. They recorded 5,491 earthquakes of  $M_L$  2.0–4.2, of which 195 were selected to simultaneously invert for 1D velocity structure and hypocentre locations. From these earthquakes 53 well-constrained focal mechanisms were generated from first-motion picks and amplitude ratios. Leitner et al. [2001] found that the focal mechanisms were dominated by oblique strike-slip and thrust mechanisms.

Combining data from the short-term SAPSE deployment with other longer-term deployments (New Zealand national network, Lake Pukaki network and the small deployment of three seismometers following the 1997  $M_L$  5.0 Mount Cook earthquake) Leitner et al. [2001] found a consistent seismogenic cut-off depth throughout the Southern Alps of  $12 \pm 2$  km. Beneath the central Southern Alps, they noted a change in cut-off depth to 3–4 km shallower levels attributed to elevated temperatures at depth due to high uplift rates associated with the orogen. The width of the seismogenic zone on the east of the Alpine Fault becomes wider to the south, extending from c. 60 km-wide in the central Southern Alps to up to 200 km-wide in Central Otago.

Near the Alpine Fault between Franz Josef and the Wanganui River (Figure 1.7), Leitner et al. [2001] found a zone of low seismicity (within 10–20 km of the surface expression of the fault). They linked this apparent seismicity gap to a transition from deformation on the Alpine Fault to partitioned deformation across multiple fault structures in the Marlborough Fault Zone (the Hura/Hope fault splays off the Alpine Fault  $\sim$ 50 km north of the Wanganui river: Figure 1.7).

Leitner et al. [2001] observed 60 earthquakes that they suggested occurred on or near the Alpine Fault, yielding a dip of c.  $45^\circ$ . They suggested that the presence of seismicity on the Alpine Fault shows that the fault is capable of releasing elastic strain, and that the rates of seismicity on the Alpine Fault are comparable with those (for similar magnitudes) found on locked portions of the San Andreas Fault in California [Leitner et al., 2001].

The MSc study by O’Keefe [2008] addressed seismicity in the seismically quiet zone found by Leitner et al. [2001] between September 2006 and March 2007 using a temporary network of eight seismographs. This study recorded 411 events with a magnitude of completeness of 1.6. These earthquakes were used to compute a new 1D velocity model for the region. O’Keefe [2008] also generated an updated local magnitude scale showing that attenuation in the central Southern Alps is anomalously high compared to the national magnitude scale. O’Keefe [2008] again found low seismicity rates in the area between Franz Josef and the Wanganui River, but described this seismicity gap as a ‘horseshoe’ rather than of triangular shape.

More recently Boese [2012] conducted a detailed study on microseismicity in the central Southern Alps, resulting in the publications: Boese et al. [2012, 2013, 2014]. Boese [2012] deployed a temporary network of ten short-period seismographs between November 2008 and June 2009 forming the Southern Alps Micro-earthquake Borehole Array (SAMBA) network. This network continues to operate and is described in detail in Section 2.1.1.

Boese et al. [2012] recorded a detailed earthquake catalogue of 1,791 earthquakes between November 2008 and December 2009. They found that the earthquakes tend to occur in clusters along and to the east of the Main Divide Fault Zone, and appear distributed through the crust rather than confined to mapped faults. They again found similar seismogenic depths to Leitner et al. [2001], with seismogenic depths near the Alpine Fault trace of  $10\pm 2$  km, shallowing to  $8\pm 2$  km within 20 km of the Fault trace, before deepening to  $15\pm 2$  km. Boese et al. [2012] related this to the shape of the high resistivity zone [Wannamaker et al., 2002]. Earthquakes cluster on the boundaries between high and low resistivity zones, which Boese et al. [2012] suggested indicated zones of extensive fracturing, but limited fluid saturation.

Boese et al. [2012] reported magnitudes of  $M_L -0.3-4.2$ , with infrequent high magnitudes ( $M_L > 2.5$ ) in the region of highest uplift. These magnitudes were calculated using a magnitude calibration derived by including frequency-dependent attenuation. Boese et al. [2012] found that high frequencies ( $f > 10$  Hz) were attenuated more than twice as much as low frequencies ( $f < 5$  Hz). The lack of frequency dependent attenuation in the national network magnitude calculations resulted in magnitude estimations biased by hypocentral distance in the central Southern Alps. Boese et al. [2012] also computed focal mechanisms and stress inversions for their catalogue, giving an angle between  $S_{Hmax}$  and the average strike of the Alpine Fault of  $59^\circ$  which they interpreted to mean that the fault is poorly oriented and therefore weak. However, due to the segmented nature of the Alpine Faults surface trace, individual thrust and strike-slip segments are favourably oriented for faulting in this stress regime.

Boese et al. [2013] located 20 subcrustal earthquakes recorded between December 2008 and February 2012 at depths of 47–74 km. These events combined with previously documented subcrustal earthquakes [Calhaem et al., 1977, Haines et al., 1979, Reyners, 1987, Kohler and Eberhart-Phillips, 2003], and the non-volcanic tremor documented by Wech et al. [2012] (discussed in more detail in Section 1.5) led Boese et al. [2013] to identify three distinct zones of seismic character along the Alpine Fault (Figure 1.8). Zones at both north and south ends of the Alpine Fault are defined by subduction zone intermediate-depth and deep earthquakes. Directly beneath most of the SAMBA network few intermediate-depth earthquakes are seen and there is little crustal seismicity north of Mt. Cook. The final zone is marked by a sharp transition south of Mt. Cook

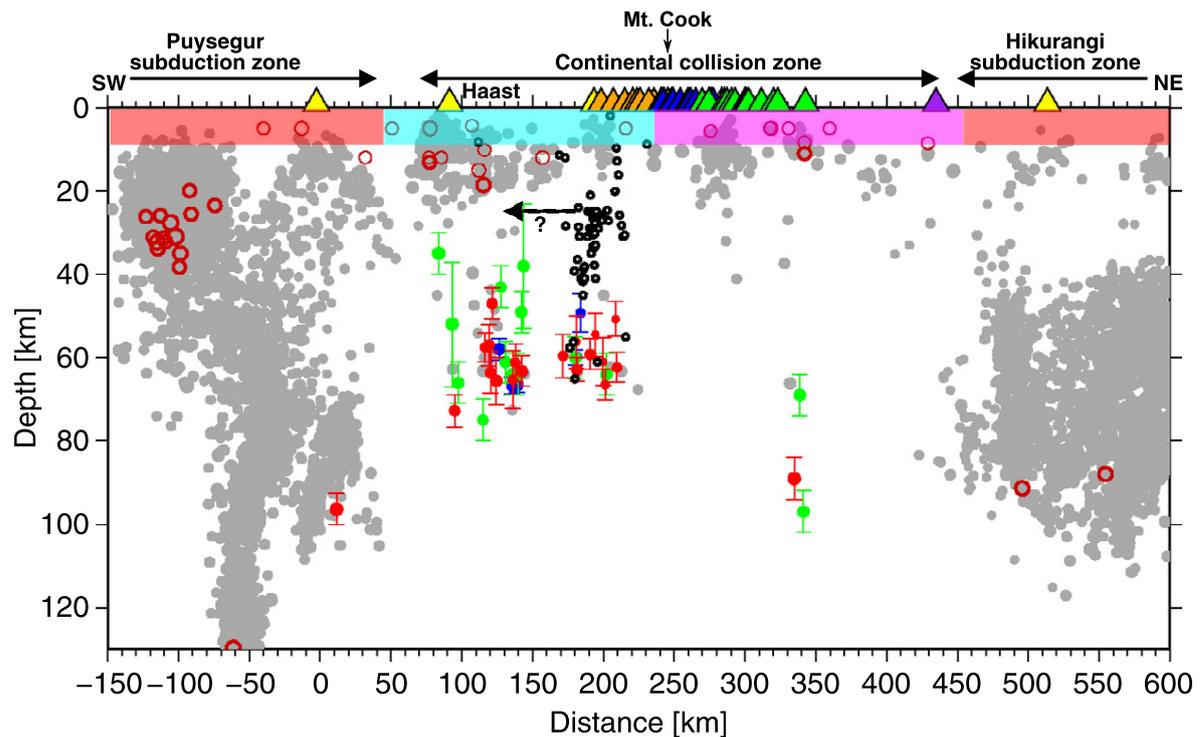


Figure 1.8: Adapted from Boese et al. [2013], highlighting the changes in seismic character along strike of the Alpine Fault. The red boxes signify regions characterised by subduction driven intermediate and deep earthquakes. The pink box indicate a region is seismically quiet, while the blue box indicates a region that displays more seismicity along with intermediate depth earthquakes and tremor. Tremor locations from Wech et al. [2012] are shown by the open black circles, with large depth uncertainty. Open red circles mark earthquakes of  $M > 5$ , grey circles mark well located seismicity in the GeoNet catalogue. Red and green closed circles with uncertainty are subcrustal earthquake locations.

and shows both tremor and intermediate depth earthquakes.

The zone of intermediate-depth earthquakes adjacent to the Alpine Fault appears to correlate spatially with the hypothesised extent of partially subducted Eocene-Miocene rifted crust [Sutherland et al., 2000]. This crust would likely be rich in volatiles, providing a fluid source capable of weakening the fault at depth and allowing for slow-slip episodes to occur as highlighted by tremor [Wech et al., 2012].

Boese et al. [2014] documented earthquake swarms within the central Southern Alps and provided evidence of both background and triggered swarms consisting of low magnitude ( $M_L \leq 2.8$ ) earthquakes. These earthquake swarms are distinct from mainshock-aftershock sequences by virtue of having no apparent mainshock. Boese et al. [2014] noted that some of the swarms are triggered by regional earthquakes. These swarms showed a delay between the passage of the surface waves of the triggering earthquake and the beginning of the swarm, indicating that pore fluid diffusion is the likely trigger rather than dynamic stresses due to surface waves. Following these swarms were periods of quiescence, which may be indicative of so-called ‘clock-advancing’, where

critically stressed faults are forced towards failure earlier than would be expected given background tectonic stressing [Gomberg, 2001].

Bourguignon et al. [2015] catalogued more than 1,300 earthquakes north of the Wanganui river and south of Ross. They computed high precision earthquake relocations, and generated a tomographic model, and calculated focal mechanisms for 148 events. Their earthquake locations showed strong variations in seismicity perpendicular to the Alpine Fault, with sparse seismicity of dominantly reverse mechanisms in the footwall, juxtaposed with sparse seismicity in the hangingwall within 9 km of the fault trace, before transitioning into elevated seismicity rates out to 20 km. The change from low seismicity rates in the hangingwall to higher rates of seismicity around 9 km away from the fault is coincident with region of interpreted peak horizontal strain rates [Lamb and Smith, 2013], and suggest that this peak strain rate implies slip on the deep extent of the Alpine Fault, thus linking deep slip with elevated rates of crustal seismicity.

Bourguignon et al. [2015] also found shallower seismogenic cut-off depths in the hangingwall of the fault and attributed this to higher geothermal gradients, which they inferred to be associated with increased permeability in the hangingwall, and the Alpine Fault forming a barrier to fluid flow (and thus a barrier to heat transport by convection).

## 1.5 Low-frequency earthquakes in a global context

Non-volcanic tremor (herein termed tremor) is thought to be, in part, composed of a superposition of low-frequency earthquakes (LFEs) [Shelly et al., 2007]. The current working hypothesis of how LFEs are generated holds that LFEs represent repeated shear failure of small asperities (‘sticky-points’) within an otherwise aseismic region. LFEs are characterised by dominant power in the low-frequency (1–10 Hz) band (Figure 1.9) compared to regular seismicity. The deficiency in high-frequency energy observed for LFEs and longer source-time functions has recently been used to infer LFE stress drops 2–3 orders of magnitude less than for regular seismicity (order of  $10^4$  Pa), and reduced rupture-velocities (0.7 km/s) compared to regular seismicity (often  $>2.0$  km/s) [Thomas et al., 2016].

To fully understand LFEs it is important to also discuss the two related observations of geodetically observed slow-slip and seismically observed tremor (Figure 1.10). The first observations of slow-slip in the Cascadia subduction margin [Dragert et al., 2001], and tremor in southwest Japan [Obara, 2002], and of subsequent studies [e.g. Shelly et al., 2006, Nadeau and Dolenc, 2005, Payero et al., 2008, Brown et al., 2009, Shelly et al., 2007, Ito et al., 2007, Kao et al., 2009, Delahaye et al., 2009, Wech et al., 2012,

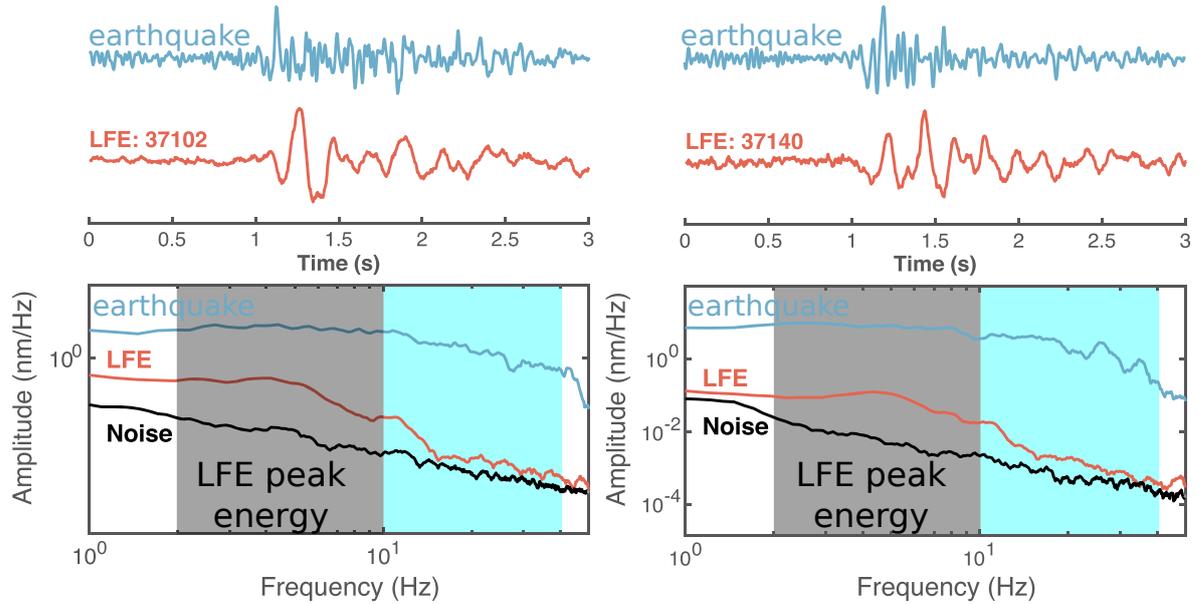


Figure 1.9: Comparison of LFE waveforms and spectrums (labelled LFE) with closely located regular earthquakes (labelled earthquake). The highlighted grey region of both spectral plots shows the 1–10 Hz band within which LFEs often have peak power. In contrast the regular earthquakes maintain power to much higher frequencies. Adapted from Thomas et al. [2016]

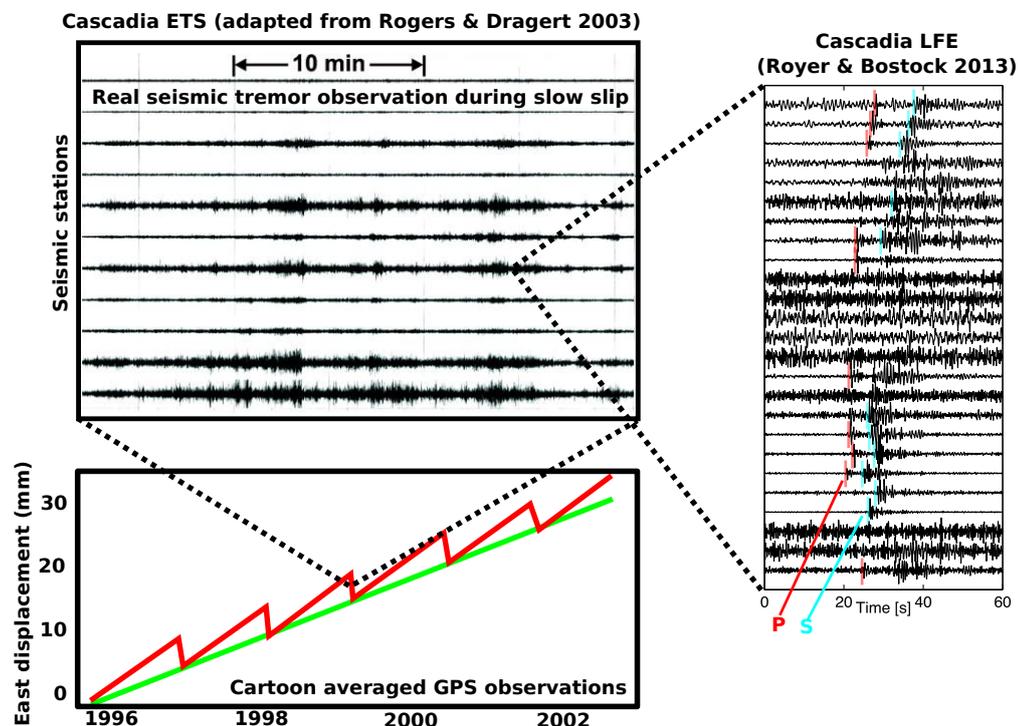


Figure 1.10: Overview of the range of scales slow-earthquakes on the Cascadia subduction margin (encompassing geodetically observed slow-slip, bottom panel, tremor, upper panel, and LFEs, right panel). Geodetically observed slow-slip often occurs on time-scales of days to weeks and is often accompanied by seismically observed tremor. Within the tremor are LFEs with often emergent seismic phase arrivals.

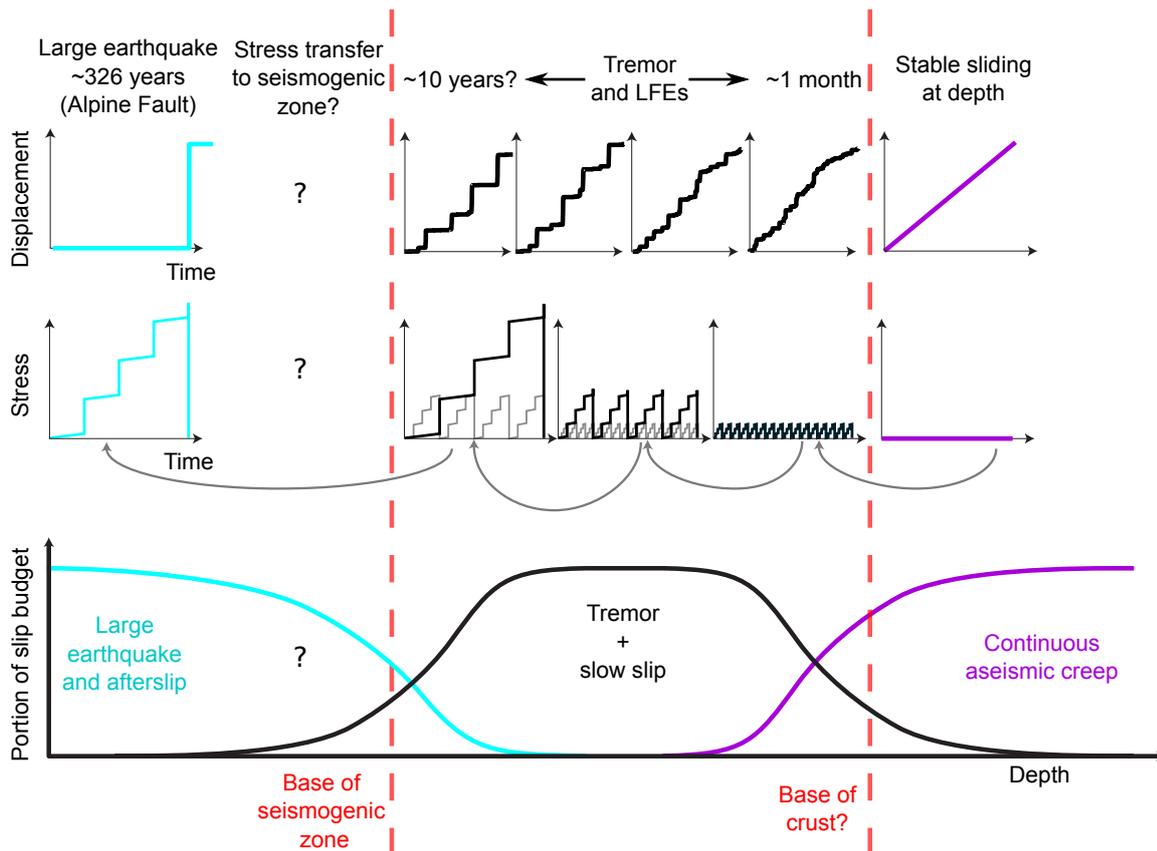


Figure 1.11: Adapted from Wech and Creager [2011]. Schematic model of how stress is transferred from deep stable sliding at the base of the crust to the seismogenic zone capable of generating large earthquakes.

Peng and Chao, 2008, Thomas et al., 2009] provided evidence that the transition of faults from locked (near the surface, on the seismogenic portion of faults) to stably sliding at depth is likely to be a gradational change, with a variety of slip mechanisms possible from large earthquakes through to slow, aseismic creep [Wech and Creager, 2011] (Figure 1.11).

These early observations, and reconciliation of tremor and slow-slip into Episodic Tremor and Slip [Dragert et al., 2004], showed that despite the slow moment release observed by Dragert et al. [2001], seismic energy was being radiated by these slow slip events, as non-volcanic tremor. The moment release of large slow-slip events is comparable to that of large earthquakes, with moment magnitudes of Cascadia slow-slip often exceeding of  $M_W$  6.4 (Figure 1.12) with relatively constant and predictable recurrence intervals of  $15 \pm 2$  months [Wech, 2010].

Although Cascadia Episodic Tremor and Slip events appear to be periodic in nature [Rogers and Dragert, 2003] and, on a broad scale, follow a relatively simple progression along strike of the subduction zone [Wech and Creager, 2008] (with some caveats [e.g. Wech and Bartlow, 2014] and short-duration rapid tremor reversals [Ghosh et al., 2010, Houston et al., 2011]), this is not the case in all subduction zones. For

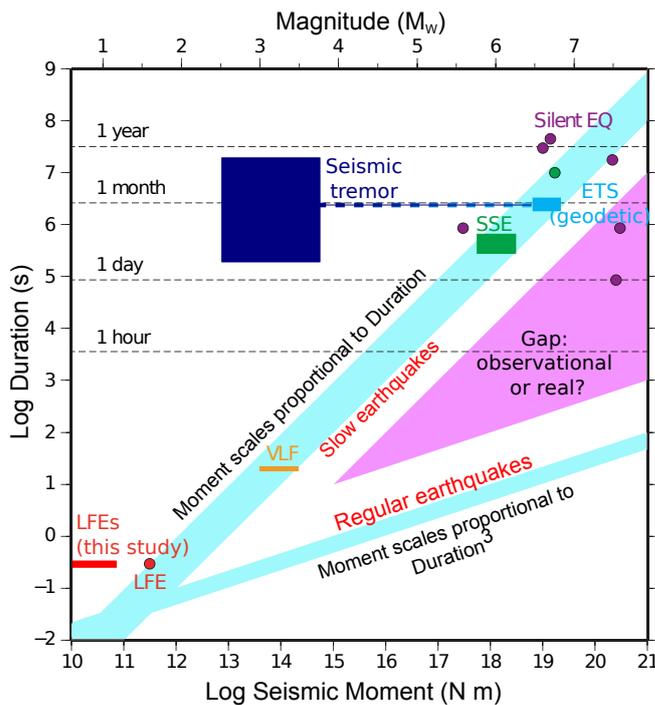


Figure 1.12: Adapted after Ide et al. [2007a]. Observational scaling of regular earthquakes and slow earthquake with duration. The gap in pink between regular and slow earthquakes reported by Ide et al. [2007a] may be due in part to a lack of available tools for observation of seismic phenomena at these moments and durations. The gap between seismically observed very low frequency earthquakes (VLF) and geodetically observed slow-slip in itself highlights a lack of observations within a band of durations from 100s to 10 days. This may be rectified through the recent development of constantly observing strain gauges.

example slow-slip in the Hikurangi margin of New Zealand is observed at a range of depths (possibly as shallow as the trench), and a variety of durations, from 6 months to 1.5 years [Douglas et al., 2005, Delahaye et al., 2009, Wallace and Beavan, 2010]. Even in the apparently simple Episodic Tremor and Slip zone of Cascadia a range of short-term transients are observable within tremor including rapid tremor reversals and streaks that propagate at speeds up to  $350\times$  faster than the main rupture front [e.g. Thomas et al., 2013, Houston et al., 2011, Ghosh et al., 2010]. The wide range of behaviour within tremor is however reproducible in relatively simple rate-and-state friction based models with frictional healing [e.g. Colella et al., 2012].

There are some caveats to the relationship between tremor and slow-slip. Notably in Cascadia, slow-slip can occur without tremor [Wech and Bartlow, 2014], where the authors observe a slow-slip event transitioning from tremor-generating to silent and back again as the slip front progresses along strike. Wech and Bartlow [2014] suggest this may be due to a combination of changing slip-speed and changes in rheology along strike. On the San Andreas tremor is documented without observable slow-slip [Shelly et al., 2009], however here the apparent disparity between tremor and slow-slip may be due to detection thresholds, with slow-slip on the near vertical San Andreas unlikely to be observable at the scales and depths proposed by the tremor locations and durations. On New Zealand’s Hikurangi margin Delahaye et al. [2009] documented increased reverse-faulting microseismicity, but no tremor accompanying a shallow slow-slip episode.

Following the early observations of tremor as a long duration, emergent, low-frequency signal, Shelly et al. [2007] described tremor in Japan as a swarm of discrete LFE events.

The discovery of LFEs and their relation to tremor quickly allowed for more accurate locations of tremor using traditional P and S-phase picking than the previous waveform envelope cross-correlation locations using a single-phase (S) lag times for location (see Section 2.2.2 for more details of the detection and location methods). The presence of distinct, albeit often emergent, phase arrivals also allowed for the generation of LFE focal mechanisms [e.g. Ide et al., 2007b, Shelly et al., 2007]. The combination of LFE locations clustering around fault interfaces and LFE focal mechanisms consistent with slip on the main subduction thrust fault in both Cascadia [Royer and Bostock, 2013] and Japan [Ide et al., 2007b, Shelly et al., 2007] provided further evidence that both LFEs and tremor are the seismic manifestation of slow-slip on the main plate interface fault.

The delicate balance between tremor and slow-slip is again highlighted through observations of triggered tremor, both by tidal stresses on the scale of 1–10 kPa [e.g. Thomas et al., 2012, 2013] and by passing seismic waves [e.g. Chao et al., 2013, Hill et al., 2013, Rubinstein et al., 2009, Gomberg and Prejean, 2013], again with very low dynamic stress perturbations on the scale of 2–10 kPa. Multiple authors appeal to high pore-pressures, close to lithostatic, resulting in low effective stresses to allow for triggering by such low external stresses [e.g. Thomas et al., 2012, Beeler et al., 2013]. The presence of fluid is also suggested by low  $Q_p$  regions close to many tremor generating regions [e.g. Wech et al., 2012].

The ability of tremor and LFEs to be triggered by such low stresses enabled Beeler et al. [2013] to make inferences about the rheological state of the crust in the region of LFE generation on the San Andreas fault. The study of the state of rheology relies on the currently pervasive simple model of LFE sources as small, persistent asperities within an otherwise stably creeping region (Figure 1.13). This model at the same time explains why tremor and slow-slip magnitudes do not match (not all of the sliding region is generating seismic energy), and why we observe LFEs as repeating events. It is also consistent with geological observations of presumed tremor-genetic regions [Fagereng et al., 2014], where the fault zone is characterised by lenses of brittle material within a more viscous region. These brittle lenses are postulated to be capable of generating seismic energy, while the matrix surrounding them may deform plastically around the lenses, accommodating the greater moment release observed geodetically than the pure seismic moment release would suggest.

Detailed studies of LFEs and tremor have suggested that the magnitudes of these events do not scale in the same way as traditional earthquakes (Figure 1.12) [Wech, 2010, Sweet et al., 2014, Bostock et al., 2015]. Tremor magnitudes have been calculated by various authors based on either a duration scaling related to the geodetically observed moment [e.g. Aguiar et al., 2009], or using duration-amplitude scaling [Watanabe et al.,

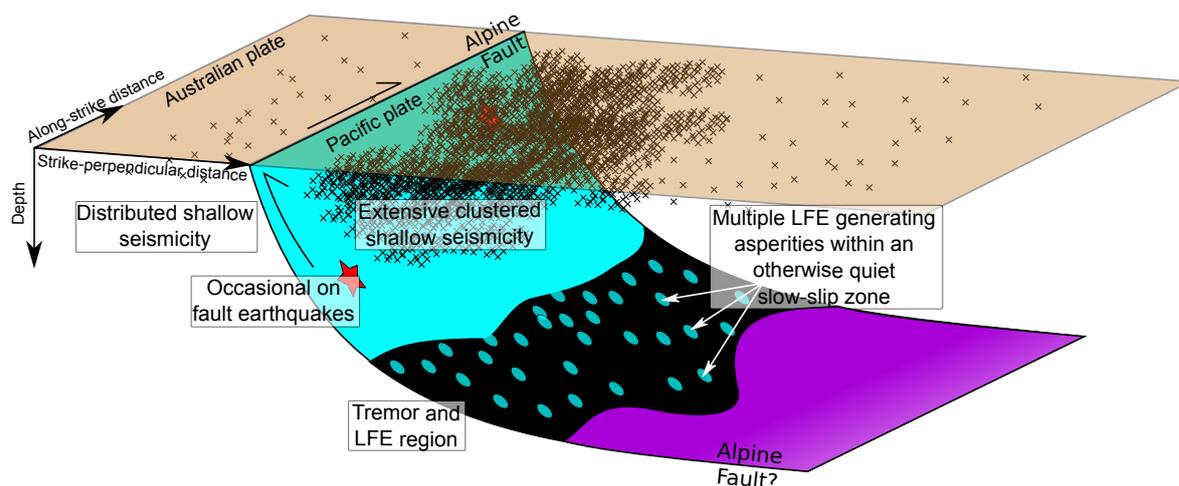


Figure 1.13: Schematic model for LFE generation in the Southern Alps. The transition from stable sliding at depth to the shallow seismogenic region in the upper crust is characterised by tremor and LFE generation [Wech et al., 2012, Chamberlain et al., 2014]. We suppose that LFEs (and therefore tremor) may be generated on small asperities within an otherwise stably sliding region. These asperities undergo brittle failure to generate seismic energy, while the stable sliding region surrounding them may provide a means for slowing the seismic slip of these brittle regions producing the slow rupture speeds required to generate true LFEs.

2007]. Ide et al. [2007a] suggested that slow earthquake moments scale in proportion to the duration-cubed. However as can be seen in Figure 1.12, there is a large gap in observations on time-scales of between 100s and 10 days (tremor magnitudes can be calculated for tremor durations within this band, but these do not show the true moment release as described above).

The gap in observations at moderate duration time-scales may be due to a lack of resolution at this level rather than a true lack of activity [Gomberg et al., 2014]. For instance geodetic instruments are not sensitive to slow-slip in Cascadia below  $M_W 6.1$  Aguiar et al. [2009], and most analysis of GPS time-series is done on daily GPS solutions. More work with 1 Hz GPS networks and novel techniques such as stacking GPS time-series during periods of elevated LFE generation [Frank et al., 2015a] may reveal smaller slow-slip events. There is, however, a gap in observations between Very Low Frequency earthquake (VLF) and regular earthquakes: we have resolution between these phenomena, but have not observed any seismicity of this moment-duration scale.

Understanding how slow-earthquake size relates to the number of events is important as this provides information on the source properties, which impact how we can expect these slow events to propagate. Studies into LFE amplitudes and magnitudes [Shelly and Hardebeck, 2010, Sweet et al., 2014] show that LFE magnitudes do not appear scale according to a power law relating to number of events, rather amplitudes scale according to an exponential law [Shelly and Hardebeck, 2010]. This is also observed in a detailed non-volcanic tremor study by Watanabe et al. [2007], who find that tremor

magnitudes scale exponentially. This exponential relationship suggests that LFE and tremor sources are scale invariant. The implications of this are discussed further in Section 3.5.2. However Bostock et al. [2015], who computed moment magnitudes for LFEs in Cascadia, indicated that magnitude calculations calibrated by local earthquakes have a different relationship than their moment magnitudes, which fit a power-law model with a high b-value ( $b > 5$ ). Despite this power-law fit, Bostock et al. [2015] conclude that LFEs do not scale self-similarly and rather their magnitudes depend more on slip than on slip-area.

The fine detail provided by LFE locations on the San Andreas has provided key information on how the fault behaves prior to [Shelly, 2009] and following a large earthquake (Parkfield) [Shelly, 2010] and the depth extent of the transition zone [Shelly et al., 2009]. The precise relocations of events by Shelly and Hardebeck [2010] provide constraints on the location of the fault deep in the crust. Precise locations of LFEs in the Southern Alps of New Zealand could provide strong constraints on the location of the Alpine Fault at depth and be a strong test of the models of Norris and Toy [2014].

Precise LFE locations can also provide constraints as to how faults are deforming at depth and the rheology of the fault zone [e.g. Beeler et al., 2013]. This is exemplified by the investigation of Frank and Shapiro [2014] in Mexico, where different zones of deformation are seen in the near flat-lying subduction zone. The continuous catalogue in Mexico highlighted small deviations in LFE generation rate and LFE locations, providing a basis for the model of Frank et al. [2015b] who suggested an almost fault-valve like modulation of LFEs, suggesting that fluids play an active role in tremor and LFE migration.

The importance of studying LFEs is highlighted most effectively by the observation of tremor increases prior to, and in the region of, large earthquakes. Prior to the  $M_W 9.0$  Tohoku (Japan) earthquake of 2011 Kato et al. [2012] observed a migration of small earthquakes they linked to the propagation of a slow-slip front towards the mainshock rupture point. This observation has since been corroborated by observations made on ocean bottom pressure sensors and an on-shore strainmeter [Ito et al., 2013], and observations of tremor on ocean bottom seismometers [Ito et al., 2015]. This type of observation has also been made prior to the  $M_W 8.1$  Iquique earthquake (Chile) in 2014, with coupled foreshock migration patterns and slow-slip transients towards and around the mainshock rupture area [Ruiz et al., 2014].

Observations of slow-slip prior to large earthquakes suggest that these large earthquakes may have been triggered by this slow deformation. For this to be the case, and for slow-slip and tremor to have the potential to be used as a diagnostic tool in earthquake prediction we can consider some variant of the pre-slip model of earthquake nucleation rather than a purely random cascade model [Beroza and Ellsworth,

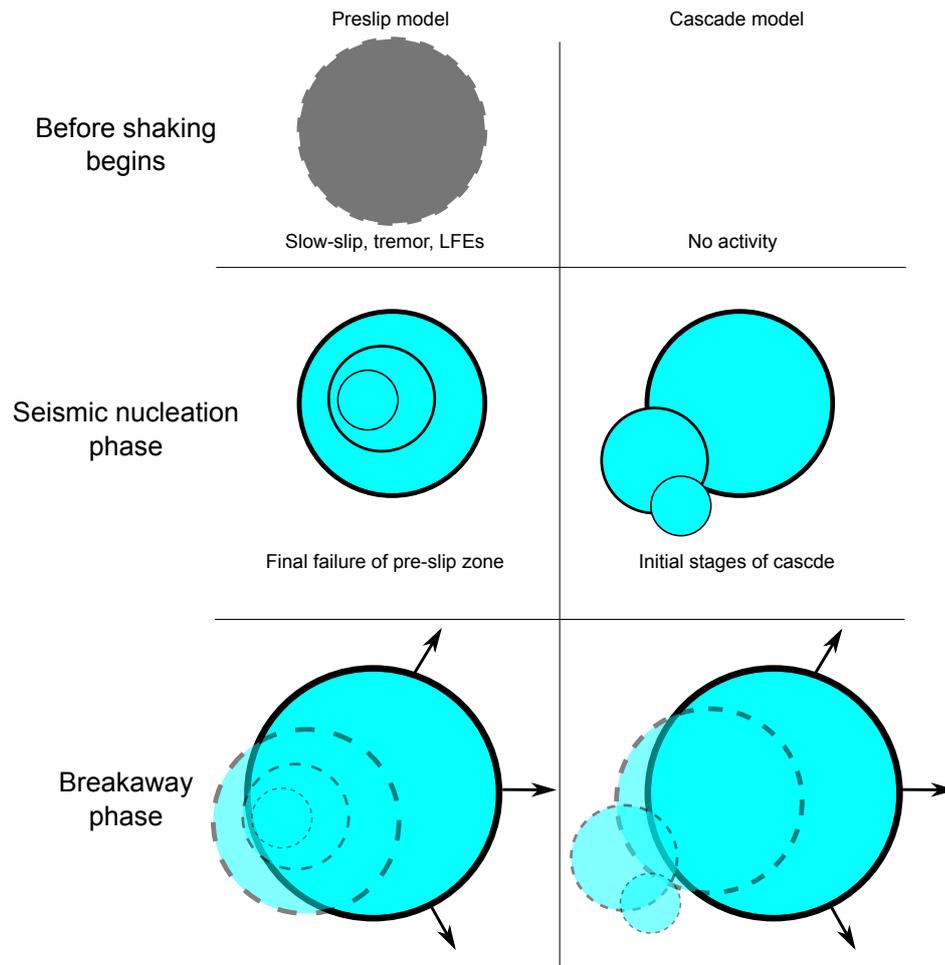


Figure 1.14: Cascade and preslip nucleation models adapted after Beroza and Ellsworth [1996]. The preslip model would see nucleation restricted to a zone undergoing pre-slip, or slow slip, before breaking away in the main earthquake rupture. In contrast the cascade model has no pre-earthquake aseismic creep, rather the nucleation phase begins by a small earthquake randomly triggering a larger event in a random walk scenario.

1996] (Figure 1.14). This appears to be the case in some laboratory experiments [e.g. McLaskey and Kilgore, 2013], where small foreshocks are promoted by aseismic creep in a block slider experiment.

If this indeed were the case then LFEs, tremor and other slow-slip associated phenomena may provide our most realistic chance of monitoring the state of stress of major faults in real-time prior to a major rupture.

## 1.6 Thesis structure

Following this introduction, the SAMBA network is described in detail along with data processing and research methodology in Chapter 2.

Chapter 3 discusses initial findings from 3 years of LFE detections including LFE family locations, inter-event times and frequency-magnitude characteristics.

Chapter 4 extends this work using an automated template detection method to generate an extended catalogue of LFE templates using 6.5 years of data.

Chapter 5 presents a catalogue of microseismicity detected using cross-correlation methods near the Deep Fault Drilling Project (DFDP) phase-2 borehole in Whataroa. In this study earthquakes are detected by correlation before being relocated using refined picks generated by single-channel correlation values. Earthquakes are then located precisely using double-difference methods.

Chapter 6 describes the real-time seismic monitoring associated with the DFDP phase-2 drilling, where microseismicity and moderate magnitude earthquakes were detected using standard energy based detection.

Finally Chapter 7 summarises the major findings of this thesis and suggests extensions of this work.

These chapters are largely stand-alone projects, with their own specific introductions and methodologies. These complement this general introduction and the following, more theoretical, methodology (Section 2). Chapters 3–6 have been written in the form of papers for publication. In the case of Chapter 3, this has already been published as [Chamberlain et al., 2014]. A shortened version of Chapter 4 is in preparation for submission to *Journal of Geophysical Research*, Chapter 5 has been written for submission to *Earth and Planetary Science Letters*, and Chapter 6 will be submitted to *Seismological Research Letters*.

Most of the work in these projects has been conducted by me (Calum Chamberlain), with all work presented in this thesis my own unless explicitly stated. Chapter 6 was a joint monitoring effort during the real-time monitoring, however I led this monitoring alongside Carolin Boese, and set-up and administered the software during real-time monitoring. All the analysis presented here of the dataset collected is my own work expect the relocations of the Malcolm Peak earthquake sequence which were conducted by Cliff Thurber. To be consistent with the format of Chapter 3 I have chosen to write the rest of this thesis in the first person plural, rather than first person singular. This also reflects the importance of my supervisor's guidance in this project.

## 2 Research methodology

To achieve the aims outlined in Section 1.1 a series of methodologies have been developed for constructing Low-Frequency Earthquake (LFE) and earthquake catalogues within the Southern Alps of New Zealand. To facilitate development of continuous catalogues, a multi-year suite of passive, continuous seismic data have been collected, the acquisition and processing of which are outlined in Section 2.1. To fully utilise these seismic data for the purpose of understanding the tectonics of the Southern Alps we consider fundamental methodological problems prior to analysis of earthquake catalogues.

The most basic of concerns when interpreting earthquake catalogues should be that of completeness, which directly relies on detection techniques. This chapter forms a discussion of some of the available detection methods and how the chosen methods have been implemented at different stages throughout this thesis (Section 2.2). Once the relevant seismic signals have been identified, we then locate the source of energy (Section 2.3) and obtain information regarding the source properties (Section 2.4).

### 2.1 Data acquisition and processing

Data from a range of deployments are presented within this thesis, however the core of the data presented have been collected using the Southern Alps Micro-earthquake Borehole Array (SAMBA) network. This network has been operational throughout this study, and the maintenance and extension of this network formed a large part of my work. The network is described in the following section (Section 2.1.1), with details of instrumentation, fieldwork protocols and operational parameters outlined in the associated appendix (Appendix A).

Further to this core network, data are also presented from select GeoNet stations (<http://geonet.org.nz/>), Wisconsin New Zealand Array Regional Deployment (WIZARD) sites [Thurber et al., 2012], and from the DFDP-2 local borehole array and associated short-term stations. All these data were collected from short-term passive deployments except for the GeoNet data, which are collected from sites that form part of the New Zealand national network operated by GeoNet, a group housed within Ge-



Table 2.1: SAMBA site locations, adapted from [Boese, 2012] and with the addition of three further sites deployed for this thesis, LARB, SOLU and MTBA.

Name	Locality name	Latitude [°]	Longitude [°]	Altitude [m]	Depth [m]	Site conditions
COSA	Cook Saddle	-43.448	170.060	377	1.40	Nr. Road, buried by gravel
LABE	De La Beche	-43.546	170.245	1590	2.15	Greywacke, surrounded by glaciers
COVA	Copland Valley (Sheils Peak)	-43.613	169.968	1477	c. 1 (surface)	Bedrock, strong schist fabric
EORO	End of Road (Franz Josef)	-43.426	170.169	233	1.35	Schist, low-light
FRAN	Franz Josef Dump (Canavan's Knob)	-43.380	170.160	124	98.00	Granite, bed-rock
GOVA	Godley Valley	-43.639	170.503	814	1.95	Greywacke bedrock
MTFO	Mount Fox	-43.512	170.003	1216	2.00	Schist, bedrock, but part of a sackungen
POCR	Potters Creek	-43.352	170.223	185	46.00	Sediments, poor signal quality
WHAT	Whataroa Quarry	-43.279	170.360	95	61	Granite, bedrock, North channel dead
REYN	Reynold's Creek	-43.353	170.391	1310	1.20	Semi-schist, short deployment
WHYM	Whymper Hut	-43.441	170.371	906	1.30	Well-cemented moraine-dam
LARB	Lansborough Rock Biv.	-43.749	169.936	1184	c. 1 (surface)	Well-cemented moraine-dam
SOLU	Solution Ranges	-43.908	169.606	1146	c. 1 (surface)	On schist bedrock, close to ridge peak
MTBA	Mt. Baird	-43.688	169.641	1232	c. 1 (surface)	Sediments

The findings from microseismicity studies using SAMBA have been published in [Boese et al., 2012, 2013, 2014]. Boese [2012] also contains a full description of the SAMBA network and site parameters for the main SAMBA deployment. Full site details will only be given here for those three sites (MTBA, LARB and SOLU) deployed as part of this PhD. Details of the network are given in Appendices A and B.

SAMBA was originally deployed as part of the Marsden project titled: ‘Putting a stethoscope on the Alpine Fault’, and has since found further use in the detection of more esoteric signals, such as glacier derived seismic signal [Horgan et al., 2015], tremor [Wech et al., 2013] and low-frequency earthquakes [Chamberlain et al., 2014, and this thesis]. Detection of such weak signals has been possible through careful deployment of sensitive seismometers. The choice to use borehole seismometers for the core of the network has allowed for significantly reduced noise levels compared to co-located surface sensors [Boese, 2012] due in part to lower temperature fluctuations beneath the surface and the decay of weak, surface generated signals with depth.

### 2.1.2 Data processing

Data have been collected from the SAMBA network biannually, usually in Austral autumn and spring, alongside network servicing to prepare for the winter (where snow-damage is likely) and rebuild instruments after winter respectively. Data are recorded by RefTek RT-130 dataloggers on compact flash disks of between 2–16 GB volume. Data are stored in a proprietary RefTek compression format based on miniseed Steim-2. After collecting the data from sites the data are then transferred from the flash disks to hard-drive and converted to miniseed format. Data files were originally stored as 15 min files, but have, during this project, been stored as hour-long files. Data conversion is done using the pascal tools *rt2ms* tool. Data are also converted to SAC format, which was used in the LFE and tremor detection routines.

Once data have been converted to miniseed, individual channel files were merged to generate merged hour-long miniseed files along with all other available data from WIZARD, DFDP2 and GeoNet stations, using either the Seisan tool *seisei* or the Obspy package in Python. These files are then queried by a simple energy based detector running a short-term average/long-term average filter (STA/LTA) for routine earthquake detection. For different elements of this project, different detection parameters and programming packages were used and as described in the following sections. Detections made by these routines were then visually inspected and picked for earthquake phases.

Separate to this, SAC data have been merged for individual channels into day-long files to be used for LFE and tremor detection. For the period between the start of SAMBA

and 2012/03/26, data have been downsampled to 100 Hz to speed processing for tremor and LFE detection. After this date, data have been left at their original sampling rate (100 Hz for GeoNet sites, 200 Hz for SAMBA and 250 or 200 Hz for DFDP2 stations dependent on time-period; DFDP2 sites were changed to 200 Hz sampling while telemetered to reduce data volumes) to allow for more accurate detection of higher frequency signals such as microseismicity.

### **2.1.3 Data quality**

The quality of data for the original SAMBA network have been discussed at length by Boese [2012]. We have included in Appendix C some key noise signals that disrupt earthquake detection. There has been, however, a significant change in data quality at station WHAT2, where one horizontal channel became irreparably damaged. The site is still recording on all channels and the remaining two channels have high quality signal.

Of the three sites installed during this PhD, only LARB has run continuously without issue (continuity of all sites is shown in Appendix B). Data from LARB are of reasonable quality, although, as the noise spectrum in Figure 2.2 shows, the noise levels often vary. We think this large variation in noise levels is likely due to landslides or other natural noise sources: LARB is situated in a steep valley, with the sensor deployed in an apparently stable land-slide. MTBA experienced sensor failures with the initially deployed Guralp 3-ESP-C sensor due to the vault flooding. After re-deploying with a Marks L4 short-period sensor we experienced write-issues for over 6 months. These write issues have since been remedied by using longer file-lengths. SOLU worked well until struck by lightning, which resulted in 6 months of data loss. SOLU has since been repaired. For both SOLU and MTBA the noise levels within the sensor working range (periods  $<0.5$  s) fall within the upper and lower noise level models (Figure 2.2).

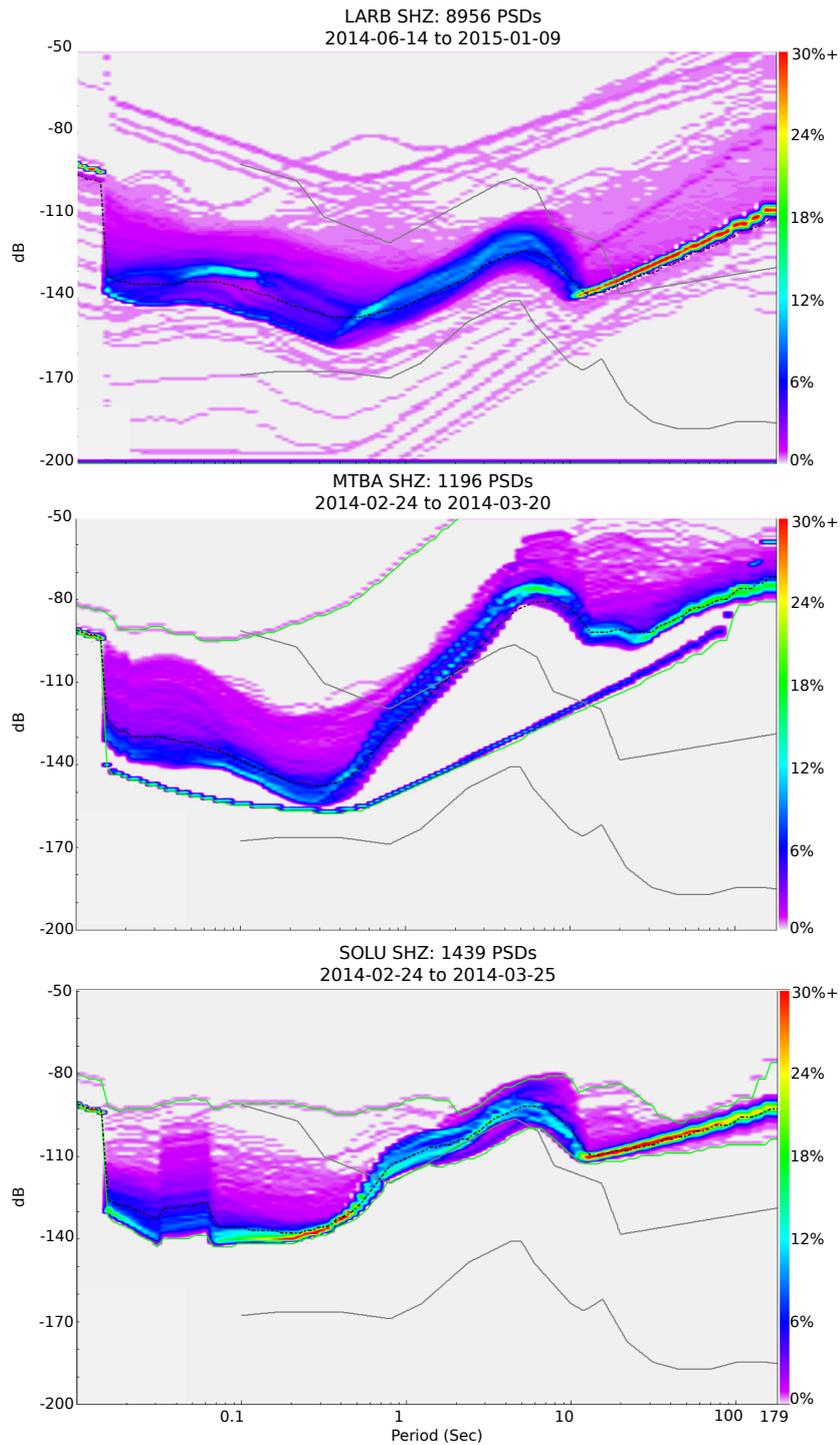


Figure 2.2: Probabilistic power spectral densities for the three SAMBA sites installed as a southern extension to SAMBA. Note power spectral densities do not cover the same time-period due to different installation dates, and problems with MTBA (sensor damage) and SOLU (lightening strike). Thin grey lines show the standard high (upper) and low noise models. LARB has a range of noise conditions, which appear to be from single days, likely when the site suffered temporary power failures.

## 2.2 Event detection

Of fundamental importance to our understanding of seismicity is our ability to detect earthquakes of a range of magnitudes. This section will discuss the detection techniques used in this thesis, compare them with each other, and discuss the limitations and benefits of the methods used here. A single perfect detection method that can reliably capture all of the earthquake related events in a seismic time-series has not been developed, hence understanding the limitations inherent to the methods deployed is key to the robust interpretation of this study's results.

### 2.2.1 Earthquake detection

The use of the term earthquake in this Chapter refers to regular earthquakes. That is, natural seismic sources with a broad frequency spectrum assumed to be the release of energy from the brittle failure of rock. Although this syntax prescribes a difference between LFEs and earthquakes, this is not necessarily true. Rather than LFEs and earthquakes being truly discrete, LFEs are here thought of as a specific class of earthquakes with limited power above 8 Hz, and are likely the result of a slightly different physical mechanism (the possible mechanisms are discussed in Section 1.5). The difference in frequency content, amplitudes and impulsivity between earthquakes and the subset of LFEs requires somewhat different event detection techniques.

#### Energy-based

The simplest of digital detection techniques for earthquakes within continuous data are energy detectors, based on changes in seismic amplitude [Withers et al., 1998]. This style of detector requires short increases in amplitude above the background noise on single seismic traces to declare detections. In principle this works by calculating the average amplitude within two moving windows; one short-term window, to represent a possible seismic signal, and one long-term window, to represent the background amplitude. These are interpreted as a ratio of short-term average/long-term average (STA/LTA). When this ratio reaches some threshold a trigger is declared. To run effectively and reduce false triggers associated with near-station noise a network detection system has been employed in this thesis, where final detections are declared when there are multiple triggers within a certain moveout window across a network of seismometers.

The main benefit of this system is its simplicity, allowing it to run faster than real-time for most networks. It is also simple to understand and determine physically realistic

parameters for the window lengths and trigger ratios without needing to conduct a full inversion. Under normal conditions short-term averages should be set to the length of a normal P-phase arrival, while the long-term average should be set to at least the period of the dominant background noise. A demonstration of STA/LTA triggering is shown in Figure 2.3 for an earthquake detected during DFDP real-time monitoring. This earthquake is of interest due to the obvious second, untriggered earthquake following the main earthquake in the record. The following sections will show how correlation based techniques can reduce detection thresholds for this type of repeating event.

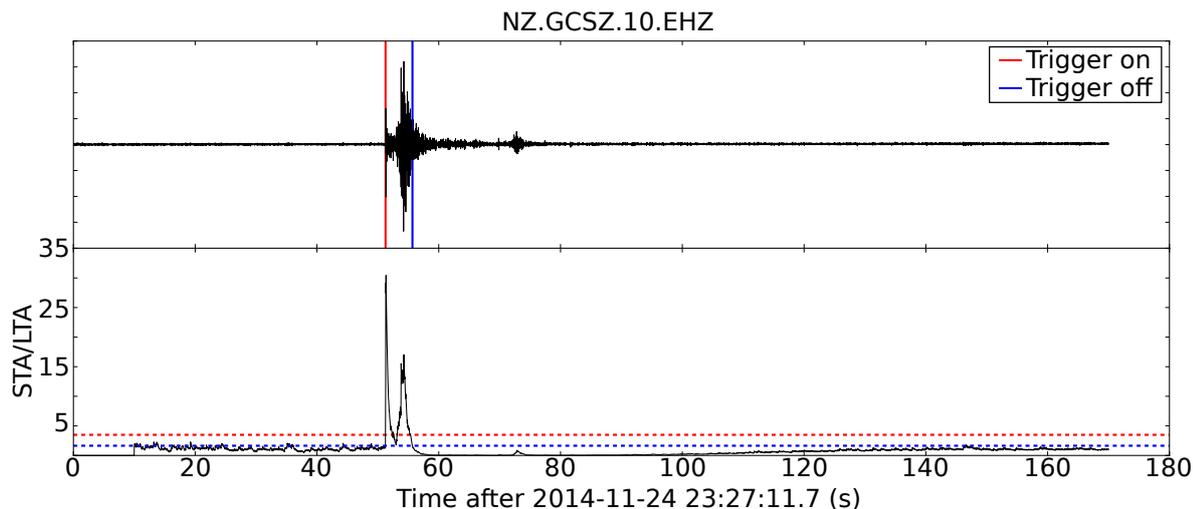


Figure 2.3: Doublet recorded on GCSZ vertical channel during the DFDP-2 real-time active seismic monitoring effort (see Chapter 6). Upper panel shows the data band-passed 5–25 Hz and sampled at 100 Hz, with marked trigger on and off points for both events. Lower panel shows a recursive STA/LTA characteristic function for the doublet, with STA length 0.3 seconds and LTA length of 10 seconds. The trigger-on ratio is set to 3.5 and the trigger-off ratio set to 1.5. These are the values used for triggering on GCSZ during drilling, however they do not trigger on the second event (see Figures 2.4 and 2.5).

### Correlation-based

Cross-correlation of known waveforms is one of the most accurate ways of detecting signals within continuous data [Gibbons and Ringdal, 2006]. In this way signals with amplitudes significantly below the SNR can be detected when the data matches the template. This type of detection routine is often termed as a matched-filter routine and is used in a range of disciplines for finding known signals within continuous data.

Initially matched-filter cross-correlation techniques in seismology were used solely to detect repeating earthquakes, relying on high single-channel correlation values (often  $cc > 0.75$ ) to detect repeating events [e.g. Schaff and Richards, 2011]. An example of single channel cross-correlation for detection is shown in Figure 2.4, with the same time-series as in Figure 2.3. Here the main earthquake is used as a template with

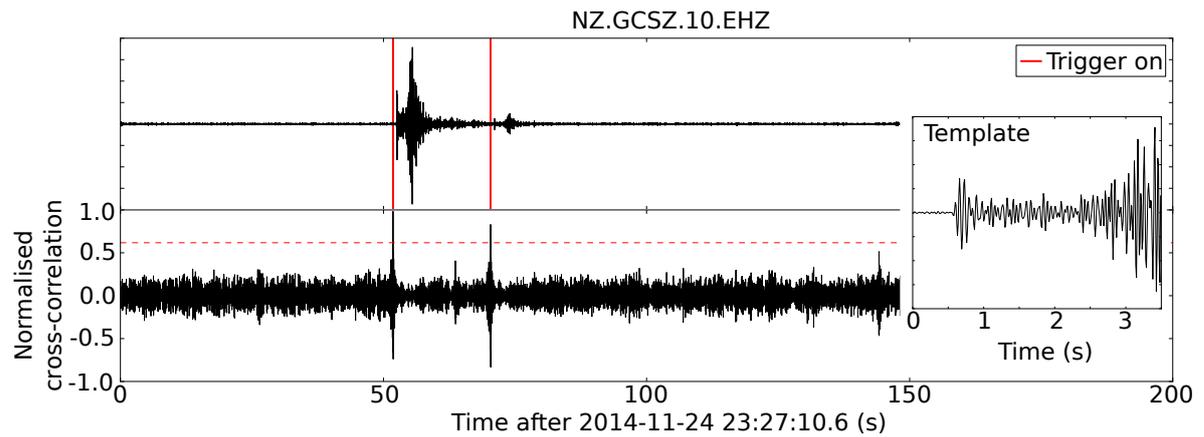


Figure 2.4: Doublet detection using single channel cross-correlation. In this example the data are the same as for Figure 2.3, with the same bandpass, and sampled at 100 Hz. The inset panel shows the template used, beginning just prior to the p-phase arrival of the first earthquake in the doublet and during the S-phase. Template length is 3.5 seconds. In this method the detection threshold is set to  $8 \times$  Median Absolute Deviation (MAD), which easily detects the second event with a normalized cross-correlation value of 0.84 (and corresponding negative peak, two samples or 0.02 s out of phase at -0.83).

which to detect itself and the second large event in record. This obviously performs better than the energy-based detector when looking for known signals, however it is not as adaptive as the energy-based detection routine. It is also important to consider the possibility of correlated noise, which appears to be occurring at around 140 s in Figure 2.4. Careful selection of thresholds is of great importance for all detection methods. Here the detection threshold is based on the Median Absolute Deviation (MAD) of the data itself, which we calculate daily.

To adapt the single-channel cross-correlation technique to allow for more varied sources Gibbons and Ringdal [2006] showed that the matched-filter technique could be extended across a network for the detection of nearly-repeating events using a network cross-correlation sum. This network method is able to detect events with relatively poor single-channel correlation values. This works by applying a delay-and-stack technique to the correlation functions, relying on the fact that although the seismic signal itself may not be coherent, the correlations should be.

This network cross-correlation method has much proven ability in the field of seismology, from the correlation of known explosions [Gibbons and Ringdal, 2006], detailed analysis of seismicity patterns [Schaff and Richards, 2011] to detection of global seismic waves [Dodge and Walter, 2015]. Recent studies of LFEs [e.g. Shelly et al., 2007] have also demonstrated the effectiveness of a network cross-correlation method for detecting LFEs, which are weak and emergent events often buried within noise.

An example of the network cross-correlation technique is shown in Figures 2.5 and 2.6,

for the same time-series as in the previous figures. This method is able to detect not only the main shock, which is the template event, and the second large-earthquake, but also two further earthquakes within the coda of the mainshock and below the noise level. The first of the events is close to the detection threshold, but is still clearly an outlier in the cross-correlation sum time-series. To estimate the confidence in this detection it is necessary to undertake thorough testing, which is outlined in Appendix D.

In this thesis correlation methods have been used both for the detection of nearly repeating events (LFEs and crustal earthquakes) by network cross-correlation, and for the generation of lag-times for the computation of relative locations of detected events. For detection the normalised cross-correlation values have been computed using the openCV<sup>1</sup> computer vision codes, specifically the *matchTemplate* function. This function is designed for 2D images with R,G,B values for each pixel, but is equally well suited to the 1D application of seismic data where every sample (pixel) has a single amplitude measurement. This function is highly efficient, computation of a day-long single-channel correlation can be achieved in under 1.5s for 100 Hz data (8640000 samples) with 6s templates on normal (Intel i5 processor) systems. The algorithm used to compute the normalized cross-correlation values in the time domain is:

$$R(x, y) = \frac{\sum_{x', y'} (T'(x', y') \cdot I'(x + x', y + y'))}{\sqrt{\sum_{x', y'} T'(x', y')^2 \cdot \sum_{x', y'} I'(x + x', y + y')^2}} \quad (2.1)$$

where  $I'$  is the normalised source image,  $T'$  is the normalised template image and  $R$  is the result.  $(x, y)$  is a sample in the image (where  $(x, y)$  range from  $0-(N_{(x,y)} - w_{(x,y)})$ , with  $N_{(x,y)}$  equal to the length of the image in the  $x$  and  $y$  dimensions and  $w_{(x,y)}$  equal to the length of the template in  $x$  and  $y$  dimensions.  $(x', y')$  is the sample position within the window being correlated. In our case the  $y$  values can be ignored simplifying Equation 2.1 to:

$$R(x) = \frac{\sum_{x'} (T'(x') \cdot I'(x + x'))}{\sqrt{\sum_{x'} T'(x')^2 \cdot \sum_{x'} I'(x + x')^2}} \quad (2.2)$$

here the normalised template and image are computed as:

$$T'(x') = T(x') - 1/w \cdot \sum_{x''} T(x'') \text{ and} \\ I'(x + x') = I(x + x') - 1/w \cdot \sum_{x''} I(x + x'')$$

with  $x''$  representing the position in the template or image being summed over for the normalisation.

In practice, computing cross-correlations in the time-domain is slow for large datasets. The implementation we use computes the cross-correlation in the frequency domain.

To compute the network cross-correlation sum for the day requires both the template

---

<sup>1</sup><http://code.opencv.org>

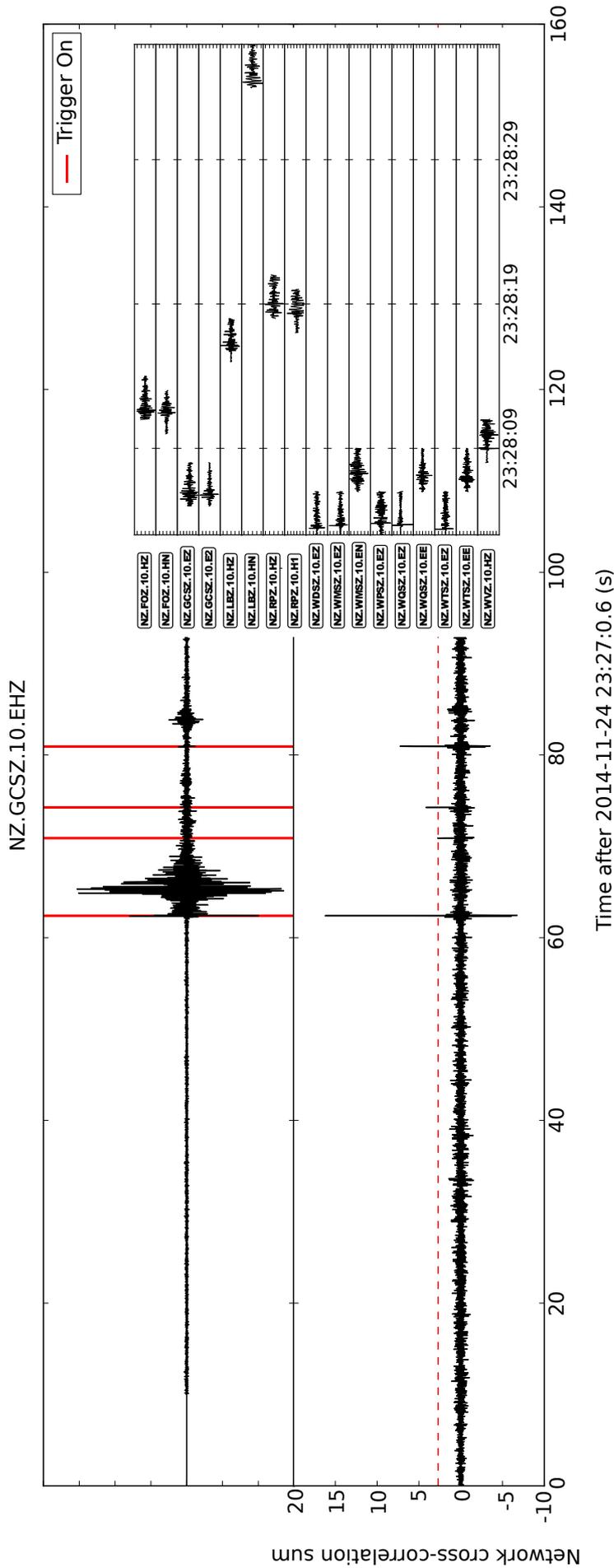


Figure 2.5: Doublet as recording during DFDP-2 real-time monitoring and as shown in Figures 2.3 and 2.4. Data are bandpassed 5–25 Hz and sampled at 50 Hz, to enable inclusion of the strong-motion site, WHFS. The inset shows the templates automatically generated from the duty seismologists picks of the larger event in the doublet. Templates are shown with their moveout applied and have length 3.0 seconds. Each template is picked to include only one phase. Due to the start of the template being at the pick time, the detections flagged are directly on the P arrival rather than before it as in Figure 2.4. The detection threshold is again set to  $8 \times$  Median Absolute Deviation (MAD). Here we see an increase in detections due to the coherence of the correlation functions through the network. Instead of just identifying the obvious doublet we are now able to identify two smaller events between them. The correlation values are lower due to the interference of noise, but still significant. Figure 2.6 shows a zoom of the detected repeats.

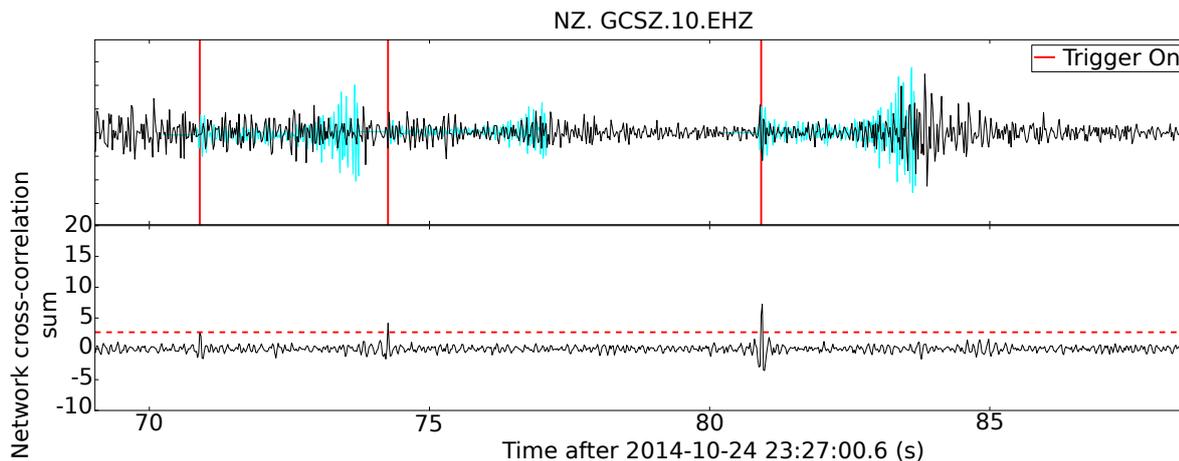


Figure 2.6: Zoom of Figure 2.5, with the template for GCSZ vertical channel superimposed (cyan) on the detected events for comparison. The final event in the sequence displays clear correlation with the template event at a high level (previously identified as a single channel normalized cross-correlation value of 0.84), with most phases of the template matching well. The two smaller events in the decaying coda of the mainshock are less obvious, however major phases correlate by eye well, and, although individual station correlation values are low, the stack of these is significantly above the noise level to be declared a detection.

waveform and the delay of the template waveform for each channel. These delays can either be applied to the correlation vectors, or the day-long data vectors. These delays should not be used to pad the template data with zeros as this will reduce the cross-correlation values. For this study the data vectors are padded with zeros at the start of the day of data according to the lag for that channel in the template. Data are then re-cut to day-long length. The template waveforms are then cross-correlated with their respective day of data for the same station and channel. The output correlation vectors are then summed to give the network cross-correlation sum.

To detect correlated events within the network cross-correlation sum a threshold is set based on the MAD value. MAD is a robust statistic to outliers in otherwise normally distributed data. For random noise the values of the network cross-correlation sum are normally distributed, however if there are correlated events with the same moveout (lags) as the template then outliers to this normal distribution are introduced. By careful selection of threshold values (see Appendix D), events below the SNRs of 1 can be retrieved.

Template generation for earthquakes can be achieved by taking events from a pre-formed catalogue and extracting the waveform and lag information. However for LFE detection, without a pre-formed catalogue template generation becomes non-trivial. The details on template generation for LFE detection are outlined in Section 2.2.3.

This thesis uses the terms *template*, *detection* and *family* throughout to describe the different waveforms used and created through the matched-filter process. These terms

are defined as follows:

- **template:** Short waveform (either single or multiple channels or traces of seismic data) containing information (usually earthquake phase arrivals and coda); used to scan for similar events in continuous data using a matched-filter methodology;
- **detection or event:** Single detection of a similar event to the template (when used in terms of matched-filter detections);
- **family:** All detections generated from a single template, this should represent all the detections that closely resemble the template. If used in a matched-filter sense then we assume all detections within a family represent closely spaced events with similar focal mechanisms (see Section 5.5 for validation in the shallow crust).

All LFE locations in this thesis have been computed for family. These family locations are computed using phase-picks on stacked data, whereby all events within the family are stacked together. In this thesis we use a linear stacking method. Individual events within a family may have slightly different locations to the family as a whole, however, we expect these differences to be small ( $< 500$  m, Section 5.5); if these location differences were not small, we would expect significant differences in the waveforms between the template and the event, which would result in the event not being detected.

## Computational considerations

To use the above cross-correlation methodology a series of codes have been written as part of this thesis. These were first written in Matlab before being re-written in Python to allow more portability, scalability and further parallel workflows. These Python codes are hosted on github<sup>2</sup> at <https://github.com/calum-chamberlain/EQcorrscan> and are licensed under the LGPL v.3<sup>3</sup>. The codes are included as the latest release prior to thesis submission in the electronic appendices. These codes have been developed, up to now, solely by me. They have been tested on OSX, Ubuntu and RedHat operating systems, and allow scalable continuous cross-correlation on machines from 4 GB RAM, single-core laptops to large cluster computers. The latest release can be accessed via the github page, or on pypi<sup>4</sup>.

Because the computation behind cross-correlation detections does not rely on other computations it is possible to design software in a massively parallel fashion. On a low level, the cross-correlations of discrete time-steps can be paralleled, which is handled in the EQcorrscan package by the openCV matchTemplate function. This can either

---

<sup>2</sup>github is an online software repository which uses the git protocol for version control.

<sup>3</sup>GNU Lesser General Public License, a free software license that allows for open-source distribution and contributing - this licence was chosen as the EQcorrscan package relies on the Obspy package, also licensed under the LGPL.

<sup>4</sup><https://pypi.python.org/pypi/EQcorrscan>

be run on threads within a multi-threaded CPU or on multiple cores within a GPU. Further to this, templates can be run in parallel through one day of data, in EQcorrscan this is handled by the multiprocessing module in Python, which allows multiple child processes to be spawned on multiple CPUs. On an even higher level still, multiple days of data can be analysed in parallel by running multiple instances of the matched-filter routine in EQcorrscan, this is handled by overarching Python scripts.

To what extent the codes are paralleled is limited mostly by resources, meaning that processing around 100 templates through the 6.5 year continuous SAMBA dataset presented in Chapter 4, on a desktop with a dual-core intel processor and 32 GB of RAM would take approximately 1 month. For the same processing, but using the School of Geography, Environment and Earth Sciences, Victoria University of Wellington (VUW) (SGEES) cluster machine with 64 cores and 256 GB of RAM takes approximately 5 days, however this is not achievable as other users require time on the cluster.

To enable the processing of large continuous datasets, with large numbers of templates ( $>100$ ), this project has leveraged time on the New Zealand eScience Infrastructure (NeSI) high performance computer cluster, PAN. This machine is a cluster computer with over 390 nodes. Each node has at-least 12 cores, with the majority of nodes (Intel Sandy Bridge architecture, 236 nodes) having 16 cores ( $2\times$ Intel E5-2680 CPUs), and 128 GB of RAM. This system allows computation of the full 6.5 year dataset, which corresponds to over 4,000 hours of computation time, in 2.5 hours when distributed across 100 nodes with 100 templates.

### Real-time detection

To assist the DFDP-2 project during the Phase-2 attempt to drill into the Alpine Fault, a local borehole seismic network was deployed in early 2013. This network was subsequently telemetered in mid 2014 to allow for real-time monitoring of seismicity during the drilling of DFDP-2. Earthquake detection and automated hypocentre determination was computed using the *Rtquake* software package [Utheim et al., 2014].

*Rtquake* downloads data in as near to real-time as is available via the *Seedlink* protocol. Data are then scanned as they arrive using a simple STA/LTA routine on single channels to detect increased energy for short periods. These single channel detections are then compared in time to other single channel detections, if there are more than a set value of detections within the network moveout window then a trigger is declared.

Following this trigger a short segment of data is downloaded for each station and channel defined in the set-up files. These are multiplexed and saved as a miniseed file with an empty S-file associated with it (*Seisan*, Nordic format pick file). An automated picker routine is then run over the data to pick both P and S phases with the S

phases downweighted (assigned weight of 3 by default). The picker routine used is the *FilterPicker* of Lomax et al. [2012].

In this picking routine a range of filters are used to generate filtered time-series. The envelope of these filtered time-series' is then calculated and converted to a characteristic function following equations 2.3.

$$F_n^C(i) = \frac{E_n(i) - \langle E_n \rangle (i-1)}{\langle \sigma(E_n) \rangle (i-1)}, \quad (2.3)$$

where  $E_n$  is the envelope for a given pass-band  $n$ ,  $\langle E_n \rangle (i-1)$  and  $\langle \sigma(E_n) \rangle (i-1)$  are the time averages up to sample  $i-1$  of  $E_n$  and the standard deviation of  $E_n$  respectively. Accordingly the characteristic function  $F_n^C$  characterises the change in the envelopes with respect to the previous background level. A final characteristic function is defined by the maximum sample for each of the bandpassed characteristic functions:  $F^C(i) = \max\{F_n^C(i); n = 0, N_{band} - 1\}$ . When this summary function exceeds a set trigger threshold a trigger time is declared. This is only declared as a pick if the integral of the summary characteristic function exceeds a further predefined threshold. For further details on the method see Lomax et al. [2012]. P picks are made on the vertical channel and the S picks made on the horizontal channels.

The *FilterPicker* routine has a number of parameters which should be optimized for greatest picking accuracy. This routine was used during the DFDP drilling period without optimized parameters (due to time and data constraints), which resulted in mostly good P-picks, but mostly poor S-picks. Picks from the *FilterPicker* routine are then written to the S-file which is then loaded by *Hypo71* [Lee et al., 1972] for location. This routine iteratively reduces residuals by removing picks with the largest residuals one at a time until the RMS residual falls below the maximum value assigned by the user. *Seisan's automag* program is then run on the S-file (if a location has been possible). This program generates automated amplitudes for local magnitudes and spectral corner frequencies for moment magnitudes. Finally the *Rtquake* routine sends a mail to a pre-defined alert email using the *Sendmail* and *Mutt* programs within Linux. This mail contains a plot of the waveforms to allow the seismologist on duty to check the picks by eye before further analysis, alongside the S-file and a location map.

Although our implementation of the *FilterPicker* technique was not perfect, it did provide a fast and simple picking routine which allowed for automated checking of events followed by manual final picks and locations by an expert seismologist. For further information as to the implementation within the DFDP drilling period and for results of this work see Chapter 6.

### 2.2.2 Tremor detection

Tremor detection has been carried out in a systematic, manual fashion for the period from 2009 to mid 2012 by Wech et al. [2012]. Manual detection was utilised as high noise levels in the tremor frequency band, and short tremor duration relative to subduction tremor [Wech and Creager, 2008], made automated detection un-viable.

Initial detection is done within the CORAL [Creager, 1996] toolbox extended by Aaron Wech for tremor analysis. Waveforms are first read in, filtered between 1-6 Hz, converted into envelopes, smoothed using a 20 s low pass filter and decimated to 1 Hz. The conversion to envelopes does not detract from the ability to detect tremor, and allows for faster loading of data for reading. Envelopes are then manually scanned for signal of tremor characteristics determined for SAMBA:

1. Emergent arrival;
2. Appears to correlate by eye over entire network;
3. Low frequencies dominate;
4. Not appearing on northerly control station (THZ).

The WECC method of tremor detection is based on waveform envelope cross-correlation (Figure 2.7). In the automatic method every 5 minutes of data are correlated to find the optimal lag times for maximal cross-correlations (in the manual method only user selected windows are correlated). A threshold for cross-correlation values is set by the user to include envelopes in the detection. If a robust set of correlations is found a location is then generated by performing a grid-search through a series of travel-time grids for each station for the optimal lag-times found.

The locations generated are based on the S-wave velocity of the medium and are only ever single phase locations. Due to this (and the location of the tremor outside of the SAMBA network), depths are poorly resolved. However this method does constrain the epicentral location well. Wech et al. [2012] found that tremor in the Southern Alps located in an almost linear trend vertically approximately 20 km south of the extent of SAMBA. The smearing of depths is evident when comparing the depth extent of the tremor locations with the depths of the LFEs reported here and in Chamberlain et al. [2014].

### 2.2.3 Low-frequency earthquake Detection

LFE catalogues are generated here using a cross-correlation matched-filter routine as described in 2.1. This method relies on the cross-correlations of template events

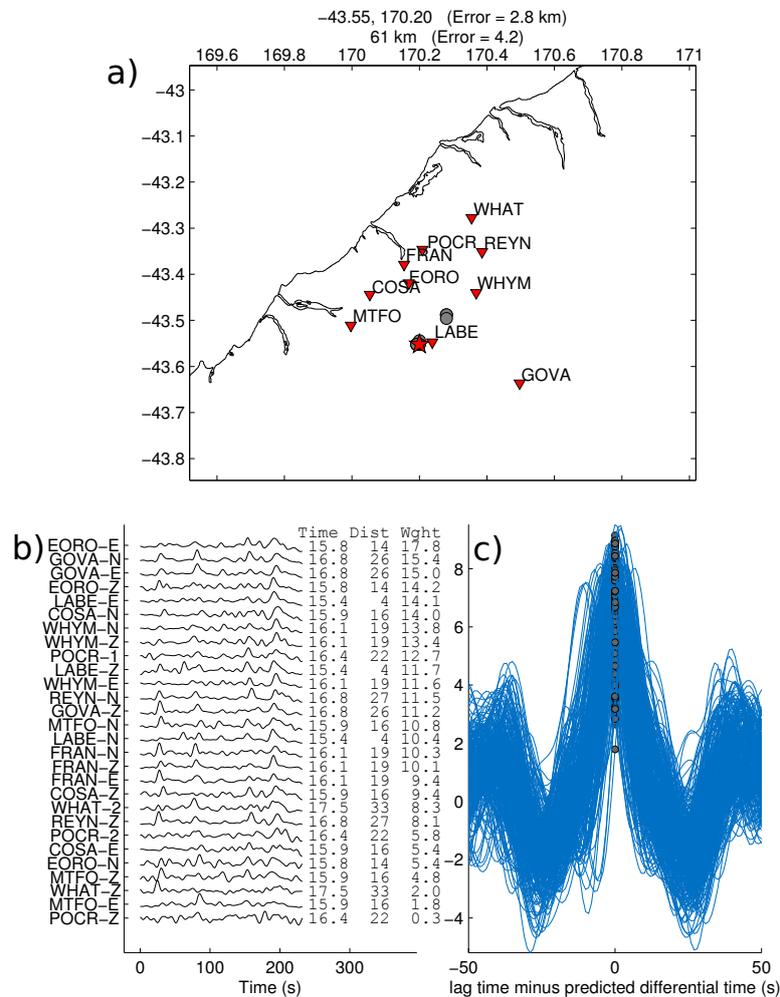


Figure 2.7: Tremor detected on 2009/07/14 at 17:28 UTC. a) Tremor location as a red star, with other bootstrap locations plotted as grey circles. b) Envelope waveforms sorted by weight (a function of correlation), alongside the waveforms respective times, distances and weights. c) Correlation functions, which show a good alignment of peaks indicative of a good location. The location obtained for this burst of tremor is  $-43.55, 170.20$ , with an approximate depth of 61 km. Uncertainties in depth are high for tremor, and LFEs within this tremor burst locate considerable shallower (15-30 km deep).

throughout continuous data. This section discusses how templates have been generated for cross-correlation, whereby a template is simply a series of channels of seismic data containing the event we want to identify, with each channel having a specific delay. Templates should be filtered to emphasise the signal above the noise, and contain as little pre-event noise as possible.

### Manual detection

Initial LFE identifications were made by manual inspection of tremor waveforms. These LFEs were characterised by impulsive phase arrivals within the first few seconds of tremor, prior to complications of superposition within the tremor wavetrain (Figure 2.8). These events were identified as likely LFEs by comparison of their frequency spectra with regular micro-earthquakes and with the background tremor (Figure 2.9), and their similar initial locations to the tremor they were found within.

Despite these characteristics, our model of LFEs as being the essential component of tremor and occurring as repeating earthquakes (see Section 1.5) effectively nullifies the need for accurate determination of templates to take through the cross-correlation stages of the LFE detection routine providing computer time is not limited and that the cross-correlation codes are written efficiently. If the template is shown to repeat then these repeats can be stacked allowing for improved SNR and more accurate picking (Figure 2.8). The locations for LFEs given are locations for the family of detections (the stack) rather than the locations of the initial template events.

Detection of further events from this initial, sparse catalogue was undertaken using the iterative-stacking and cross-correlation method of Shelly and Hardebeck [2010], whereby the network cross-correlation method described in Section 2.2.1 is run through the data multiple times, with a redefinition of the template used after each iteration given by the stack of detected events. This method was initially coded in Matlab, however to allow for more paralleled code and increased portability code have been converted to Python language, and can be found at github: EQcorrscan<sup>5</sup>

A full description of the specific parameters used for LFE detection by matched-filter, and the testing routines behind these parameters, can be found in Section 3.3.1.

### Beamforming: brightness

This method of detection applies a delay-and-stack protocol to seismic data from multiple stations within an array. If a signal is coherent across the array the signal will sum coherently while the noise (assumed to be Gaussian on each station) will stack to near

---

<sup>5</sup><https://github.com/calum-chamberlain/EQcorrscan>

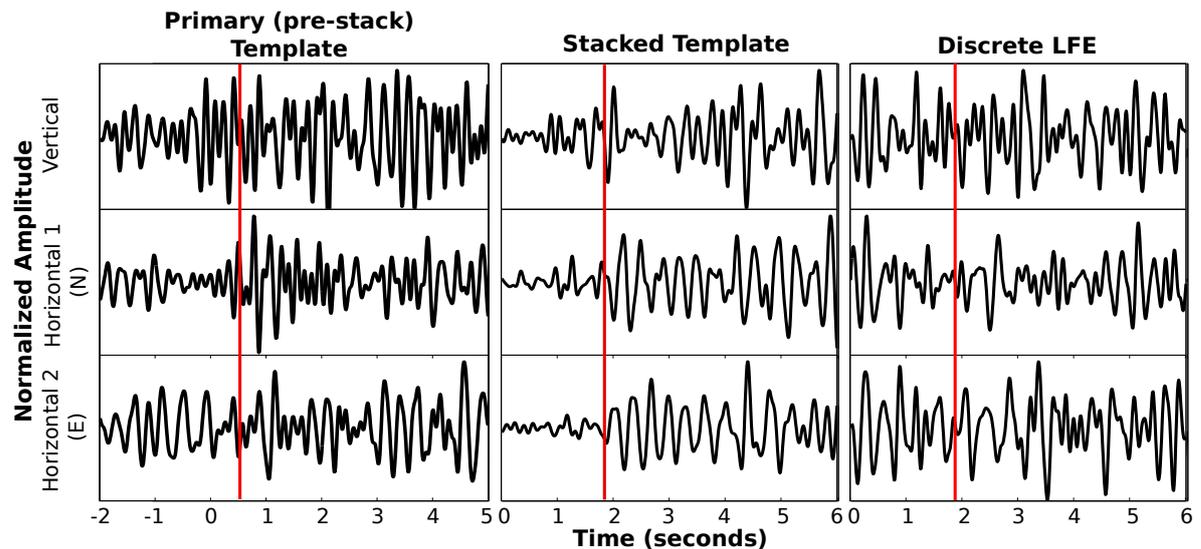


Figure 2.8: S-wave window for a LFE detected on the SAMBA network and plotted here on all three components of SAMBA station LABE. Times are relative to the start of the window used in correlation for detection of similar events, this begins between 0.5 and 1 s before the S-arrival (red line). From left to right are plotted the primary (pre-stack) template; the same template after as a stack of approximately 1000 waveforms (middle); and a final detection of a discrete (outside of tremor)  $M_L$  0.5 LFE (right). Note that the primary template is padded by two seconds prior to the beginning of the search window to highlight the change from background noise to LFE. A second LFE with S-arrival between 5–6 s is also cut from this plot but included in the primary template. All plots are bandpassed 2–8 Hz as this frequency band contains the highest amplitudes above background noise for our LFE detections. The red line marks the S-phase arrival in all plots. SNR improves after stacking, allowing for more accurate phase picks of both S and P (not shown here as P arrivals are not correlated for at any stage in our method). Note the change in S-pick before and after stacking. This is likely due to noise in the initial manually detected primary template (left) obscuring the true phase arrival and highlights the importance of stacking before locating LFEs. This figure is repeated as Figure 3.2.

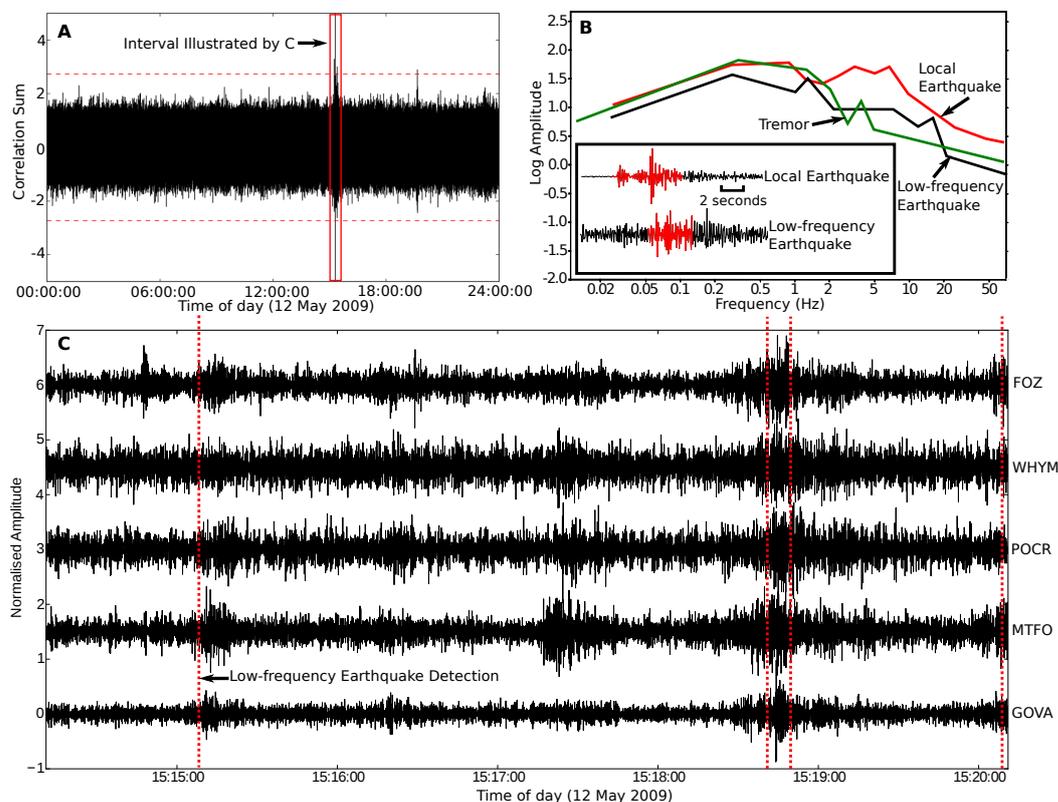


Figure 2.9: LFEs detected by matched-filter cross-correlation. Panel A shows the day-long cross-channel correlation sum in black, with the threshold of  $8 \times \text{MAD}$  (Median Absolute Deviation) plotted as a red dashed line. This is plotted both positive and negative to allow for both strongly positively and negatively correlated detections, thereby including the possibility of different focal mechanisms. In practice, for LFEs, positive correlations are often preceded and followed by weaker negative peaks indicative of the template being out of phase by half a phase. Panel B shows the spectra obtained on SAMBA stations for LFEs (green), tremor (black) and local earthquakes (red), as documented by Shelly et al. [2007], tremor and LFEs are deficient in high frequency energy. Panel C shows the waveforms for the first horizontal component (approximately North oriented) for five sites used in the detection of LFEs. Red dashed lines indicate an LFE detection for a single template.

zero. This only works well on true seismic data (not absolute values of data) if all sites in an array are in one focal sphere, otherwise the situation where one site may be in a compressional quadrant and another site in a dilatational quadrant is likely. In this situation the signal would sum incoherently and cancel itself out. Hence beamforming is often only used for teleseismic events or regional events rather than local events.

It is also important to consider array deployment when conducting beamforming as some of the issues around incoherent stacking can be overcome through careful array deployment. Rost [2002] also notes that inter-station spacing is important when considering what frequency of waves will be coherent across an array.

For SAMBA, where we have an array with large station spacing and analyse local earthquakes, the delay-and-sum beamforming procedure has been performed on the square of the seismic trace (proportional to seismic energy), which will always stack coherently. This method does not reduce the Gaussian (noise) portion of the beam unfortunately and rather relies on peaks in energy across the beam for given points in the earth.

The sole use of beamforming here is for the determination of LFE templates according to the method of Frank and Shapiro [2014]. As discussed above, LFE templates do not have to be a perfect characterisation of an LFE waveform, rather they should be the bulk energy within an LFE. In this way beamforming over the energy of the data in our network effectively discriminates between energy coming from the LFE source region and background noise. The seismic data from this period of elevated energy can then be extracted as a template and run through an efficient cross-correlation routine before stacking the detected events. The stack can then be located using normal seismic methods (phase-picking) rather than relying on the energy source location implicit in the beamforming generation.

Frank and Shapiro [2014] presented a simple, but effective method of generating a continuous catalogue of templates which we have also employed here. This method relies on an array of seismic stations and a known LFE source region in a similar methodology to the brightness based source-scanning algorithm of Kao and Shan [2004]. Seismic data from all stations are squared and normalised to provide a proxy for seismic energy. Data from all stations are then lagged according to a set of lag-times (moveouts) required for a hypothetical LFE source node before being stacked. This delay-and-stack procedure is repeated for all the possible LFE source nodes within the grid. Careful selection of grid-space defines the possible sources, therefore by selecting only node points beneath the seismogenic zone energy from crustal earthquakes is stacked incoherently.

For each node this generates an energy stack. The energy stacks combined into a beam by taking only the maximum values in the stacks for each time interval (or sample).

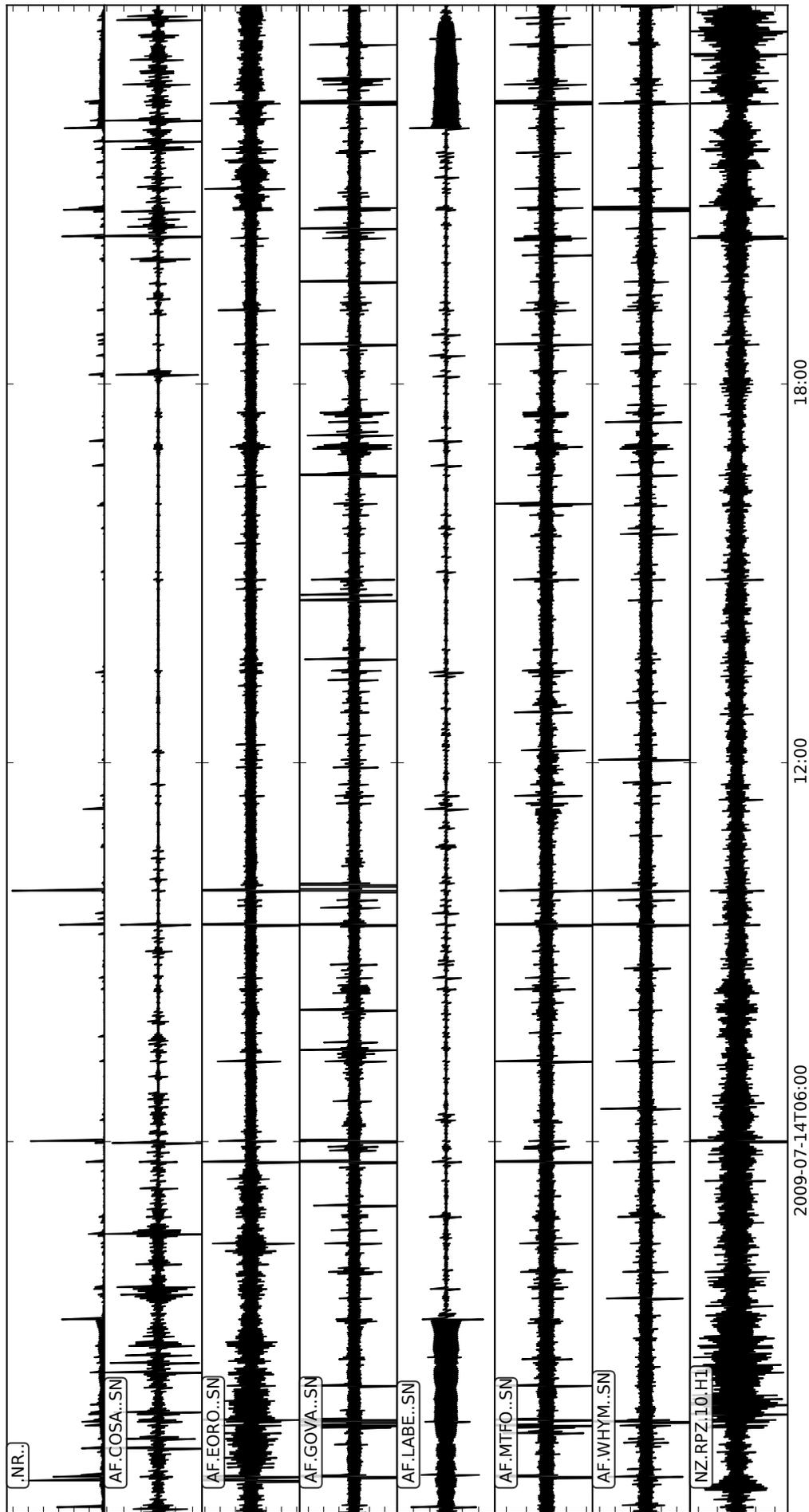


Figure 2.10: Network response (upper panel) for a test set of three nodes within the LFE source region: (169.511, -43.796, 19 km), (169.511, -43.778, 19 km), (169.511, -43.76, 19 km). On this day (2009-07-14) Wech et al. [2012] detected tremor at the start and end of the day, coincident with peaks in the network response from which templates can be generated.

The beam is termed the network response ( $NR$ ) and is given by

$$NR(\tau) = \max_{\phi, \lambda, z} \sum_i^N \frac{s_i(\tau, \phi, \lambda, z)^2}{\text{rms}[s_i(\tau, \phi, \lambda, z)^2]} \quad (2.4)$$

where  $NR(\tau)$  is the network response at time  $\tau$ .  $\phi$ ,  $\lambda$  and  $z$  are geographic co-ordinates for a source node,  $s_i(\tau, \phi, \lambda, s)$  is the seismic data at station  $i$  in  $N$  stations at time  $\tau$  lagged to correspond to the moveout from node  $(\phi, \lambda, z)$ .

Alongside  $NR$ , we also maintain a record of the node providing the data for the network response at each sample (e.g., a vector of the indexes related to the maximum values extracted for the  $NR$ ). These nodes must be maintained to maintain the moveout information required for template generation.

Peaks in the network response (Figure 2.10) represent periods of elevated energy release from a single, or multiple source locations. Data corresponding to peaks in the network response are extracted as possible templates, with the appropriate moveout for the hypothetical source location. False template detections are then removed by computing the network coherence ( $C$ ) of each extracted template and rejecting those with coherences less than some threshold ( $C_{thresh}$ )

$$C = \frac{\sum_{i=1}^N \sum_{j=i+1}^N |R(i, j)|}{N} \quad (2.5)$$

where  $R(i, j)$  is the cross-correlation of data from channel  $i$  with data from channel  $j$ , and  $N$  is the total number of traces. This is effectively the sum of the strict (e.g. zeros along the diagonal) upper triangular matrix of possible cross-correlation pairs divided by the number of channels. This is maximised as  $N$ .

After extensive testing this method was found to be unreliable at detecting LFEs in the central Southern Alps due to high noise levels resulting in many false initial detections. Low inter-station coherence due to high levels of scattering (likely due to strong foliations in the schist and many moderate scale faults in the crust) precludes the ability to discriminate between signal and noise using the coherence threshold outlined above.

## 2.3 Location methods

A critical component of any earthquake catalogue are the locations of the detected earthquakes. These locations can provide constraints on the structural controls of seismicity. Hypocentre location and origin time are generally modelled together either by forward modelling from estimated source parameters, or as an inverse problem from

information gleaned from seismograms. Here ‘earthquake location’ is a combination of hypocentre location and origin time where hypocentre location is a three dimensional spatial location. For local events a Cartesian co-ordinate system will often be used in processing, but results will be quoted in global spherical co-ordinates (Latitude, Longitude and Depth).

### 2.3.1 Linear methods

The following equations (2.6, 2.7 and 2.8) and explanations relate to the local problem in Cartesian space for simplicity, but could equally be converted to the spherical co-ordinate system.

For a simple homogeneous half-space the relation between arrival time ( $t$ ) and hypocentre parameters in Cartesian co-ordinates ( $X, Y, Z$ ) and origin time ( $T$ ) follows the equation:

$$t_i = T + \frac{1}{v} \sqrt{(X - x_i)^2 + (Y - y_i)^2 + (Z - z_i)^2} \quad (2.6)$$

where  $i$  denotes the characteristic of the  $i^{th}$  recording station and  $v$  is the velocity of the homogeneous half-space.

Utilising equation 2.6 allows calculation of travel times for a given station by forward modelling from a chosen hypocentre and origin time. This can be extended to calculate the travel times at all stations for many hypocentre locations and origin times; this is often done by gridding the subsurface and calculating the traveltimes for all stations at each grid node. From these locations and subsequent traveltimes, the best location is chosen by minimising the residuals ( $r_i$ ) where each station residual is given by:

$$r_i = t_i^{calc} - t_i^{obs} \quad (2.7)$$

The inverse approach to the above forward problem is utilised in Geiger’s location method for local earthquakes [Geiger, 1910] and more recently summarised and adapted by Thurber [1985]. Geiger’s method relies on the expansion of the first order Taylor series about a hypocentre of co-ordinates as above and solving in a least squares fashion giving the following:

$$r_i = \frac{\delta t_i}{\delta X} \Delta X + \frac{\delta t_i}{\delta Y} \Delta Y + \frac{\delta t_i}{\delta Z} \Delta Z + \Delta t_0 \quad (2.8)$$

where  $t_0$  is the origin time and terms preceded by  $\Delta$  are adjustments to be made to the model (hypocentre), and other symbols are as previously defined. The first order

partial derivatives are expressed below:

$$\frac{\delta t}{\delta X} = \frac{X - x_i}{vS} \quad (2.9)$$

$$\frac{\delta t}{\delta Y} = \frac{Y - y_i}{vS} \quad (2.10)$$

$$\frac{\delta t}{\delta Z} = \frac{Z - z_i}{vS} \quad (2.11)$$

where all terms are as before, with the addition of a path length term ( $S$ ).

The multivariate nature of this system requires at-least four travel-time recordings to constrain all earthquake location parameters, with the normal result that over-determination of the system will yield increased accuracy in earthquake locations.

In real Earth situations we are rarely dealing with a constant velocity system; and we are required to expand the relation in equation 2.6 to include a known velocity structure. For the Southern Alps a number of three and one dimensional velocity structures have been estimated; this study uses the one-dimensional velocity structure of O’Keefe [2008] as used by Boese et al. [2012] for initial linear earthquake location.

In locating events by non-linear methods both the one-dimensional model, and the 3D velocity models of the WIZARD group for DFDP events, and the national 3D model of Eberhart-Phillips and Bannister [2002] for LFE detection and location have been used. The inclusion of a layered structure yields a more complex time-series with a greater range of seismic wave arrivals including: refracted and reflected arrivals, phase conversions, and scattered energy. For a discussion of the associated complexities of layered media see Thurber [1985].

The approach outlined above; Geiger’s method, relies on the approximation of an inherently non-linear system (equation 2.6) to a linear system of equations 2.8. This is reasonable for local events in a simple medium (e.g. slowly varying seismic velocities between source and receiver), where the problem is weakly non-linear, with all variables displaying little co-dependence. However for more complex media or for distant hypocenters, attempting to solve in a linear fashion breaks-down. In this instance all parameters display a large co-dependence and the linear inversion approach becomes unstable (“a small change in the inputs of the problem produces a physically unacceptable large change in the outputs”: Tarantola and Valette [1982]).

In the use of the *Seisan* software package [Havskov and Ottemoller, 1999] locations have been obtained using a modified version of the *hypocenter* program of Lienert et al. [1986]. This software utilises an adaptive damping procedure to iteratively improve residuals in a least-squared linear inversion style. *Hypocenter* allows a full range of seismic phases to be input and generates an approximate initial location through the

use of a starting location algorithm, using azimuth and slowness of a plane wave within the network, before applying the iterative least-squares linear inversion method with adaptive damping.

### 2.3.2 Non-linear earthquake location

To provide a more stable earthquake location for complex media or/and a disparate or shallow source we must include the second order terms of the Taylor series presented in equation 2.8 which take the form of second order partial differentials [Thurber, 1985], which can be calculated analytically as in equations 2.12, 2.13 and 2.14, with similar solutions for the remaining x and y terms.

$$\frac{\delta^2 t}{\delta Z^2} = \frac{1}{vS} \left[ 1 - \frac{(Z - z_i)^2}{S^2} \right] \quad (2.12)$$

$$\frac{\delta^2 t}{\delta Y \delta Z} = \frac{(Y - y_i)(Z - z_i)}{vS^3} \quad (2.13)$$

$$\frac{\delta^2 t}{\delta X \delta Z} = \frac{(X - x_i)(Z - z_i)}{vS^3} \quad (2.14)$$

Thurber [1985] notes that for the case of  $Z = z_i$  the first-order partial derivative for  $\frac{\delta t}{\delta Z}$  (2.11) vanishes whilst the second-order partial  $\frac{\delta^2 t}{\delta Z^2}$  (2.12) is maximised. This represents a case where the source depth is the same as the receiver depth: a shallow source. This is exacerbated for short path lengths (such as those for shallow events whose hypocentre lies within a small aperture array, as is often the case for microearthquakes detected by SAMBA).

A further case for which non-linear earthquake location is useful occurs for hypocentres lying outside of an array. The increased stability in the non-linear location method again relies on the greater sensitivity of the second order partial derivatives at shallow depths ( $Z \approx z_i$ ) compared to the first order terms.

In both these cases (shallow-local and shallow-disparate) Geiger's method (undamped) does not achieve convergence; to combat this *Hypocentre in Seisan* [Havskov and Ottemoller, 1999] fixes hypocentre depths to zero for both these cases.

Further to greater stability, this non-linear method can be extended to include a robust, probabilistic error assessment. This error assessment can be represented by a spatial Posterior Density Function (PDF), which takes the form of a volume in space. This volume is often markedly disparate to the idealised error ellipse produced by linear error estimation [Lomax et al., 2000].

For relocation the *NonLinLoc* [Lomax et al., 2000] software package has been used,

which provides a more realistic estimation of error through the provision of a PDF. This PDF can then be plotted as a cloud of possible locations around the best fitting hypocentre.

The NonLinLoc package first generates a traveltime grid for a every source in a chosen grid at grid-nodes defined by a grid-node spacing, to each receiver. The picks are then compared to these grid traveltimes by non-linear inversion in a Metropolis-Gibbs fashion [Lomax et al., 2000].

### 2.3.3 Relative locations

Relative locations of regular earthquakes are determined using double-difference methods (Section 5). Locations are computed using hypoDD [Waldhauser and Ellsworth, 2000], from picks generated by calculation of the optimal correlation lag-times for events detected by network cross-correlations (Section 2.2.1).

#### Double-difference location

The method of double difference location requires a catalogue of earthquakes with absolute travel-time measurements and/or differential travel-time measurements determined by cross-correlation. These travel times are used to generate a series of double difference equations of the form:

$$dr_k^{ij} = (t_k^i - t_k^j)^{obs} - (t_k^i - t_k^j)^{cal} \quad (2.15)$$

where  $dr_k^{ij}$  is the double difference residual for a pair of earthquakes,  $i$  and  $j$ . Waldhauser and Ellsworth [2000] note that the travel-times,  $t$  are absolute travel times, but the bracketed  $(t_k^i - t_k^j)$  could be replaced with the differential travel-time generated from cross-correlation.

This equation relies on the assumption of a constant slowness vector for both events between the source and receiver. This breaks down for events that are not sufficiently close together in space (and possibly time if the velocity structure is time varying). Expansion of Equation 2.3.3 with the appropriate origin times and slowness vector results in the following equation (from Waldhauser and Ellsworth [2000]):

$$\frac{\partial t_k^i}{\partial x} \Delta x^i + \frac{\partial t_k^i}{\partial y} \Delta y^i + \frac{\partial t_k^i}{\partial z} \Delta z^i + \Delta \tau^i - \frac{\partial t_k^j}{\partial x} \Delta x^j - \frac{\partial t_k^j}{\partial y} \Delta y^j - \frac{\partial t_k^j}{\partial z} \Delta z^j - \Delta \tau^j = dr_k^{ij} \quad (2.16)$$

with terms as the partial derivatives of the travel-times  $t$  with respect to the locations  $(x, y, z)$  and origin times  $\tau$ .  $\Delta x$ ,  $\Delta y$ ,  $\Delta z$  and  $\Delta \tau$  are the changes required to make the model better fit the data. The computation of these changes is done by linear,

least-squares inversion for all stations, and all hypocentral pairs. *HypoDD* solves this problem in a series of iterations whereby different weighting schemes are applied on each iteration, and those earthquakes deemed to be unclustered are removed at each iteration.

The practical result of this inversion technique is to allow for variations in velocity structure not apparent in the original velocity model. Other relative location methods often invoke station correction terms to account for differences in calculated and observed travel-times at individual stations. *HypoDD* however can be extended to the program *TomoDD* [Zhang and Thurber, 2006], which inverts for both earthquake location and velocity structure using the double-difference approach.

Double-difference techniques often improve location accuracy, providing locations accurate to within  $1/4$  wavelength according to correlation limits. This often results in a tightening of earthquake clusters, highlighting structures on which earthquakes occur.

### Lag-time determination by cross-correlation

In this study lag-times between a master event and a detected event have been generated using time-domain correlations (Figure 2.11). This method uses the cross-correlation function in *Obspy* [Beyreuther et al., 2010] which computes normalised cross-correlations for signals beyond their length (e.g. with some overlapping allowing for two time-series of equal length to find the optimum offset without padding data). In practice lag generation has been implemented after the detection stage, but could be implemented while detecting events.

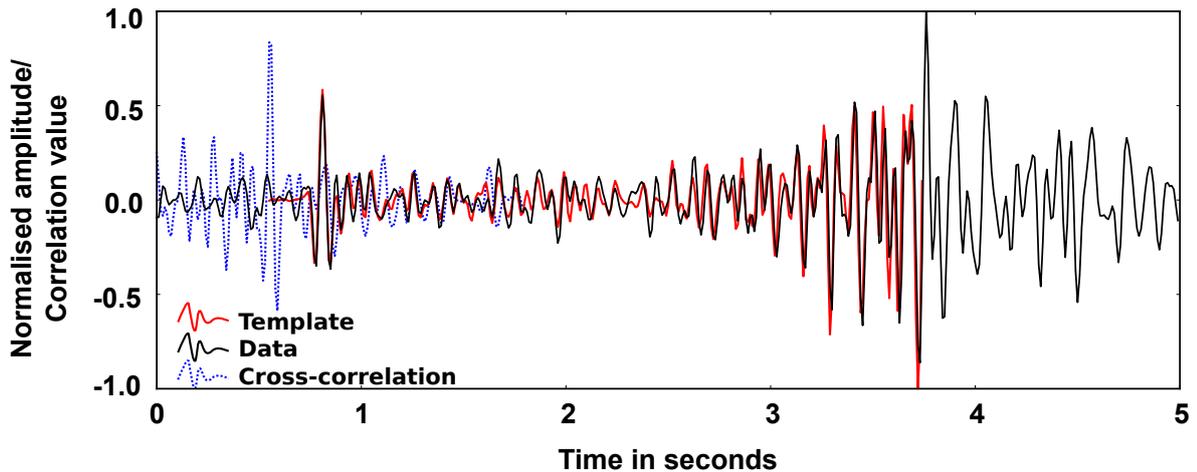


Figure 2.11: Lag-generation for the detected event in Figure 2.4, for channel Z of station GCSZ. Note that the peak in the correlation function (blue dashed) is aligned with the start of the template rather than the true P-arrival. Peak normalised cross-correlation value for this detection is 0.835. In this instance the window is 1.8s longer than the template, allowing 1.8s range in the lag-time.

For each detected event and each channel used in the detection, a small window around the detection time is cut. For each of these cut windows the cross-correlation function is computed with the appropriate template channel of data. The resultant cross-correlation function will be of the length  $length(window) - length(template)$ . The peak of this correlation function is taken to be the optimal lag-time.

Lag-times can then be converted to absolute pick-times by the sum:

$$T^i = t_{window}^i + \delta t^i + \tau^i \quad (2.17)$$

where  $T$  is the resultant absolute pick,  $t_{window}$  is the start of the window in real-time (e.g. UTC),  $\delta t$  is the lagtime generated by cross-correlation, and  $\tau$  is the delay between the start of the template window and the known phase arrival time. This is repeated for all channels  $i$ . These absolute times are then saved in a *Seisan* s-file, while the delays and correlation values are saved in text files. The s-files can be located in an absolute fashion, while the lag-times and correlation values can be fed into *HypoDD*.

## 2.4 Magnitude estimation

Seismic magnitude is an important measure in any earthquake catalogue; comparing earthquakes of different magnitudes is important when considering how seismic budget is distributed. In this thesis local magnitudes ( $M_L$ ) have been calculated based on the local magnitude scale developed for SAMBA and the Southern Alps by Boese et al. [2012]. In this magnitude scheme amplitudes of local events (within 100 km of the network centre) are converted to seismic magnitude based on the following relation:

$$M_L = \log_{10} A(\Delta) + \alpha \log_{10} \Delta + 0.4343\gamma\Delta + S \quad (2.18)$$

where  $A$  is the peak amplitude measured on a simulated Wood Anderson seismograph,  $\Delta$  is the hypocentral distance,  $\gamma$  is the anelastic attenuation parameter and  $S$  is the station specific correction term.

Boese et al. [2012] found that for the Southern Alps,  $\gamma$  is significantly frequency dependent, and  $\gamma$  should be represented as  $\gamma = \gamma_0 f$  where  $f$  is frequency. Boese et al. [2012] found a frequency dependent anelastic attenuation parameter of  $\gamma_0 = 1.89 \pm 0.02 \times 10^{-3}$  s/km.

In the calculation of magnitudes we used the amplitude measured on a Wood-Anderson simulated seismograph, on the vertical channels of seismographs. This is in contrast to the original definition of Richter, who measured amplitudes on the horizontal components. The practice of measuring on the vertical channel was adopted according to

the advice in Havskov and Ottemoller [2010], who state that the vertical channel is less susceptible to near-surface amplification affects, giving a more reliable estimate of the true seismic magnitude. When calculating amplitudes we also adopted the practice of using half the peak-to-trough amplitude rather than the true peak amplitude as the zero level on our seismographs for small earthquakes was often difficult to find. In practice this is more likely to underestimate the magnitude slightly, but consistently.

In Chapter 5 we also calculate relative moments using the singular-value decomposition method described in Rubinstein and Ellsworth [2010]. We use this method for near repeating seismicity to obtain precise magnitudes. This method is outlined in detail in Section 5.4.5.

All magnitude scales used in this thesis are adjusted to be comparable with the GeoNet national magnitude scale, which is rooted in the local magnitude scale. It is of note that this work was undertaken prior to that of Ristau et al. [2016], which may result in a systematic shift of the GeoNet national magnitude scale, which may render comparison between this work and the national scale meaningless.

# 3 Low-frequency earthquakes reveal punctuated slow slip on the deep extent of the Alpine Fault, New Zealand

This work was originally published as Chamberlain et al. [2014] in *Geochemistry, Geophysics, Geosystems*. This work has been modified to fit the formatting of this thesis, and amended slightly from the published version (Section 3.7). The full catalogue is included in the electronic appendices.

## 3.1 Abstract

We present the first evidence of low-frequency earthquakes (LFEs) associated with the deep extension of the transpressional Alpine Fault beneath the central Southern Alps of New Zealand. Our data base comprises a temporally continuous 36 month-long catalogue of 8760 LFEs within 14 families. To generate this catalogue, we first identify 14 primary template LFEs within known periods of seismic tremor and use these templates to detect similar events in an iterative stacking and cross-correlation routine. The hypocentres of 12 of the 14 LFE families lie within 10 km of the inferred location of the Alpine Fault at depths of approximately 20–30 km, in a zone of high P-wave attenuation, low P-wave speeds and high seismic reflectivity. The LFE catalogue consists of persistent, discrete events punctuated by swarm-like bursts of activity associated with previously and newly identified tremor periods. The magnitudes of the LFEs range between  $M_L - 0.8$  and  $M_L 1.8$ , with an average of  $M_L 0.5$ . We find that the frequency–magnitude distribution of the LFE catalogue both as a whole and within individual families is not consistent with a power law, but that individual families’ frequency–amplitude distributions approximate an exponential relationship, suggestive of a characteristic length-scale of failure. We interpret this LFE activity to represent quasi-continuous slip on the deep extent of the Alpine Fault, with LFEs highlighting asperities within an otherwise steadily creeping region of the fault.

## 3.2 Introduction

The original recognition of non-volcanic tremor in southwest Japan [Obara, 2002], and subsequent studies in several locations worldwide [e.g. Rogers and Dragert, 2003, Nadeau and Dolenc, 2005, Wech et al., 2012] have revealed seismic deformation occurring under temperature and pressure (depth) conditions previously thought to preclude seismogenesis. Tremor has previously been demonstrated to be a superposition of multiple low-frequency earthquakes (LFEs) in a swarm-like manner [Shelly et al., 2007, Ide et al., 2007b]. The spatio-temporal persistence of LFEs and tremor [e.g. Rogers and Dragert, 2003, Frank et al., 2013] and their response to small stresses [Peng et al., 2009, Thomas et al., 2009, Fry et al., 2011] provide a means of examining the state of stress at depth and mechanisms of stress transfer from the deep crustal roots of faults to their shallow seismogenic zones [Wech and Creager, 2011].

The constituent LFEs of tremor bursts have occasionally distinct, often emergent, P and S phases. Waveform cross-correlation techniques [e.g. Ide et al., 2007b, Shelly et al., 2007, Brown et al., 2008, 2013] for identification of repeating LFEs provide a means of detecting multiple similar events. These similar events can then be stacked to improve signal to noise ratios (SNRs). Picking both P and S phases of these stacked LFEs allows LFE families to be located more accurately than the tremor signals themselves. Tremor is often otherwise located by cross-correlating a long (2–5 minute) waveform packet [e.g. Wech and Creager, 2008]. Not only does locating the constituent LFEs of tremor allow for more robust hypocentre locations, but it also allows for tremor fronts to be tracked in high spatio-temporal definition by identifying individual LFE family activity during tremor [e.g. Shelly, 2010]. Emergent phase arrivals of LFEs have also been shown to provide sufficient constraints for the calculation of LFE focal mechanisms [Ide et al., 2007b, Bostock et al., 2012, Royer and Bostock, 2013].

Tremor has been documented at both subduction and strike-slip plate boundaries, including the transpressive Alpine Fault in South Island, New Zealand. Using data recorded by the Southern Alps Micro-earthquake Borehole Array (SAMBA) [Boese et al., 2012], Wech et al. [2012, 2013] documented tremor occurring south of Mt. Cook and attributed this phenomenon to deep slow slip on the Alpine Fault. However, due to large depth uncertainties, ambiguity remains over the relationship of the tremor to the Alpine Fault, with the possibility that tremor occurs not only on the deep extent of the Alpine Fault but also on the northern extent of the subducted Australian plate [Sutherland et al., 2006, Cox and Sutherland, 2007] or on other undocumented structures at depth. This ambiguity may be overcome using refined source locations, made possible through accurate hypocentre locations of LFEs embedded within the tremor.

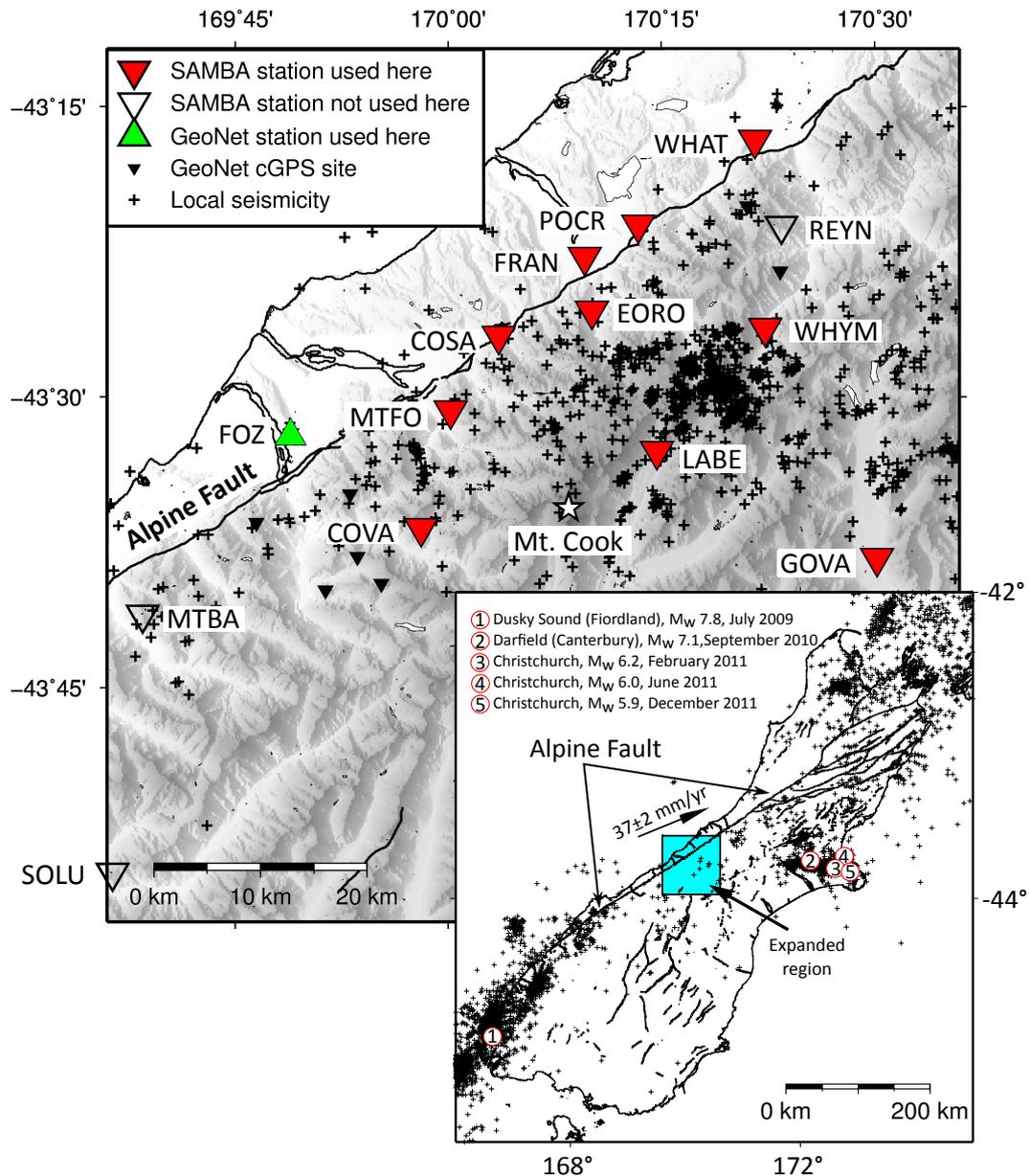


Figure 3.1: SAMBA network used in this study to detect LFEs with background microseismicity of [Boese et al., 2012] plotted as black crosses. The two southernmost SAMBA stations (MTBA and SOLU) were not active during the period analysed but may provide future constraints on LFEs south of the current LFE family hypocentres. Inset: Tectonic setting of South Island New Zealand, with earthquakes seen to trigger increased LFE generation indicated by numbered circles. Background seismicity  $M_w \geq 3.0$  from the Geonet catalogue for the period 26 March 2009–02 April 2012 plotted as black crosses.

The Alpine Fault is a transpressive margin of dextral-reverse sense at the boundary of the Australian and Pacific Plates (Figure 3.1). Geological investigations [e.g. Norris and Cooper, 2000, Sutherland et al., 2006] find a c. 23 mm/yr long-term fault-parallel Quaternary slip rate, approximately half the contemporary fault-parallel component of Australia-Pacific motion of 35–40 mm/yr [Beavan et al., 2002]. Uplift rates on the Alpine Fault are greatest (5–8 mm/yr [Beavan et al., 2002, Houlié and Stern, 2012]) in

the area of highest topographic relief surrounding the Mt. Cook massif [Little et al., 2007].

Despite high deformation rates, seismicity rates for moderate magnitude ( $M_w$  3–5) earthquakes are lower around Mt. Cook than on the northern extent of the Alpine Fault or the Fiordland Seismic Zone. Seismicity in the central Southern Alps is concentrated in areas of high resistivity gradients [Boese et al., 2012]. Boese et al. [2012] suggested that outside the seismically active shallow regions the material is either too hot, too fluid-saturated or too weak to produce detectable seismicity.

The well-documented quasi-periodic record of ground-breaking Alpine Fault earthquakes with a mean recurrence interval of c. 330 years [Berryman et al., 2012] and the inferred year of the last great earthquake on the Alpine Fault ( $1717 \pm 5$  C.E.) Sutherland et al. [2007] suggest that the Alpine Fault is late in its average seismic cycle of  $M_w$  7–8 earthquakes, and as such poses substantial hazard to southern New Zealand [Sutherland et al., 2012]. By studying LFEs we may gain some understanding of how stress transfer on the deep extent of the Alpine Fault behaves prior to a major earthquake.

Here we provide the first documentation of LFEs occurring beneath the Southern Alps. Using manually identified LFE templates and the match-filter detection technique employed by Shelly et al. [2007], we detect similar LFEs within known tremor periods and throughout the 36 month seismic record of 26 March 2009–2 April 2012. We analyse the inter-event times and frequency–magnitude characteristics of the resulting LFE catalogue and demonstrate a near-constant background rate of LFE generation punctuated by distinct rate increases on scales of minutes to hours corresponding with tremor and of days to weeks following large ( $M_w \geq 5.9$ ) regional earthquakes.

## 3.3 Methodology

### 3.3.1 LFE identification and detection

In this study LFE candidates are first identified by manual inspection of tremor waveforms detected by Wech et al. [2012]. We recognise LFEs by their distinctive earthquake-like, near-impulsive arrivals within otherwise emergent tremor waveforms. We require an LFE to have been recorded on at least five stations in order to undergo further analysis. All candidates considered here have S-picks (estimates of S-wave arrival times) at all stations used. The frequency content of LFE candidates is then checked to confirm their difference from ‘typical’ earthquakes: as in other studies [e.g. Shelly et al., 2007] our LFEs have peak SNR at frequencies of 2–8 Hz.

We compute preliminary locations using the 1D velocity model described by Boese et al. [2012] (modified after O’Keefe [2008]) and used subsequently by Boese et al. [2013] to locate deep seismicity in the region. We compute hypocentre locations and uncertainties for all candidate events using *NonLinLoc* [Lomax et al., 2000] and model parameters calculated by Boese et al. [2012] for local seismicity, namely a typical model error of 0.075 s in travel-time residuals, a correlation length of 8.0 km (equal to the average station spacing), a travel-time uncertainty of 5 % in the velocity model, a minimum picking travel-time error of 0 s and a maximum of 0.2 s. We do not recalculate velocity model errors as we do not have enough hypocentres within our dataset. As noted by [Wech et al., 2012] for the tremor signals within which these LFEs are detected, depths are anti-correlated to both deep (c. 50–100 km) [Boese et al., 2013] and shallow (0–12 km) seismicity [Boese et al., 2012]. Due to the use of an overly simplistic 1D velocity model, and uncertainty parameters more appropriate for shallow seismicity, we consider the uncertainties calculated here to represent minimum uncertainty estimates.

Table 3.1: Locations of LFE families, picked on stacked waveforms. Mean  $M_L$  values for each LFE family. Overall mean  $M_L$  for all detected events is 0.5 over the 8760 catalogued events. Uncertainties quoted represent minimum uncertainties and do not include uncertainties arising from velocity model errors, which are likely to result in much larger hypocentral uncertainties. Note, table differs from that published, numbers in brackets in the detections column are the published, erroneous numbers, with those un-bracketed being the corrected numbers. This affects Figure 3.9.

Family ID	Latitude	Longitude	Depth	Mean $M_L$	Detections
55115	-43.67	169.91	$25 \pm 2$	0.3	(419) 329
55200	-43.56	169.93	$27 \pm 3$	0.3	(258) 223
55432	-43.64	170.06	$25 \pm 3$	0.4	(486) 545
61100	-43.64	169.93	$24 \pm 2$	0.4	(362) 476
30441	-43.68	170.14	$27 \pm 2$	0.4	(426) 590
60905	-43.64	170.00	$21 \pm 3$	0.3	(308) 410
61044	-43.77	169.98	$22 \pm 3$	0.9	(1726) 1847
61220	-43.71	170.02	$20_{-4}^{+3}$	0.5	(346) 412
37575	-43.67	170.01	$23 \pm 2$	0.4	(1456) 1562
63877	-43.67	169.94	$26_{-3}^{+5}$	0.4	(269) 366
59966	-43.67	169.98	$23 \pm 3$	0.4	(372) 483
60070	-43.64	170.00	$18_{-3}^{+2}$	0.3	(324) 432
54905	-43.71	170.04	$20_{-5}^{+3}$	0.5	(509) 430
17208	-43.52	170.06	$28 \pm 2$	0.4	(577) 655

Once LFE candidates are confirmed, they are then used as primary templates in the detection of similar events by correlation. We scan with fourteen primary template LFEs (Table 3.1) through the SAMBA waveforms recorded between 26 March 2009 and 2 April 2012. To construct templates, we use a 6 s-long window beginning approximately 0.5 s before the S-pick to ensure that we span both the S-phase arrival and

its coda for each channel at each station (Figure 3.2). Templates are correlated on all three channels at each station, yielding a minimum of 15 channels per template upon which detection is based. We do not correlate P arrivals as these are less evident at most stations than the S arrivals for our primary template LFEs.

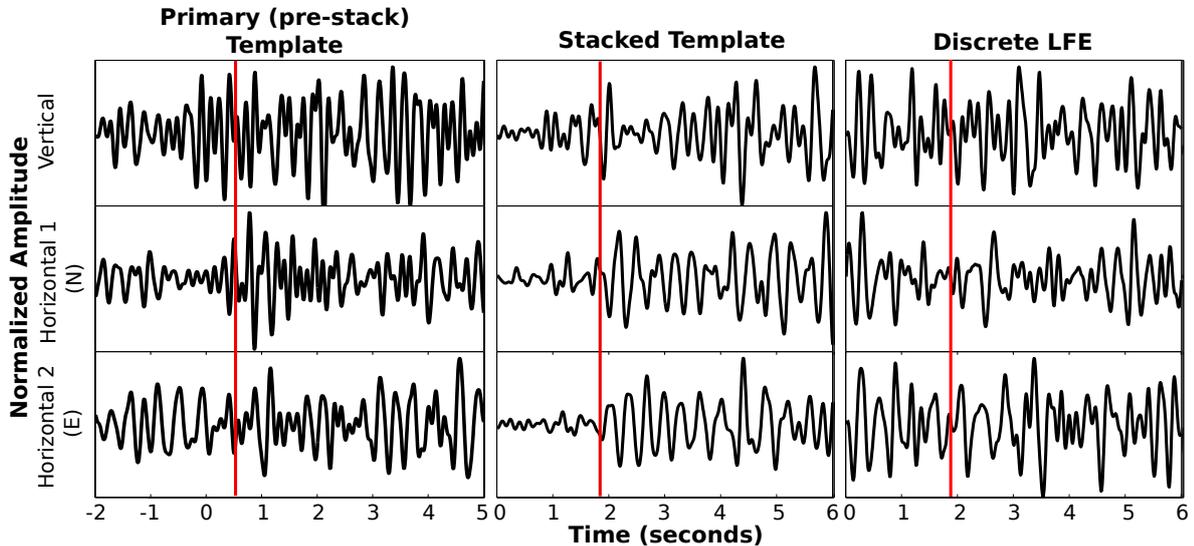


Figure 3.2: S-wave window for a LFE detected on the SAMBA network and plotted here on all three components of SAMBA station LABE. Times are relative to the start of the window used in correlation for detection of similar events, this begins between 0.5 and 1 s before the S-arrival (red line). From left to right are plotted the primary (pre-stack) template; the same template after as a stack of approximately 1000 waveforms (middle); and a final detection of a discrete (outside of tremor)  $M_L$  0.5 LFE (right). Note that the primary template is padded by two seconds prior to the beginning of the search window to highlight the change from background noise to LFE. A second LFE with S-arrival between 5–6 s is also cut from this plot but included in the primary template. All plots are bandpassed 2–8 Hz as this frequency band contains the highest amplitudes above background noise for our LFE detections. The red line marks the S-phase arrival in all plots. SNR improves after stacking, allowing for more accurate phase picks of both S and P (not shown here as P arrivals are not correlated for at any stage in our method). Note the change in S-pick before and after stacking. This is likely due to noise in the initial manually detected primary template (left) obscuring the true phase arrival and highlights the importance of stacking before locating LFEs.

In constructing the primary templates, we impose an implicit *a priori* constraint on LFE hypocentre locations by incorporating station-specific delay times. Delay terms are required to cope with large ranges in source–receiver distances throughout the network (4–60 km epicentrally), which in some cases exceed the 6 s template window length. We have found longer windows to increase the risk of correlations being dominated by non LFE-derived noise and also increase the risk of cycle-skipping when LFEs occur close together in time.

To detect similar events, we cross-correlate each of the primary templates through the dataset at 0.01 s intervals. Both the template and dataset are bandpassed at 2–8 Hz. We

detect new LFEs when the network correlation sum exceeds a threshold based on the Median Absolute Deviation (MAD) value, a statistic that is robust to outliers expected to arise from real event detections [Shelly et al., 2007]. We set our detection threshold at 8 times MAD based on synthetic testing outlined in Section 3.3.3. Detections for a single template are constrained to have an 8 s gap between them to reduce the risk of cycle-skipping.

Following an initial detection run, we then stack the best 10% of primary detections to generate a higher-SNR secondary template (Figure 3.2) in a similar fashion to the iterative stacking and cross-correlation routine of Shelly and Hardebeck [2010]. During stacking, the template length is extended to 15 s to include both P and S arrivals. At this stage the stacked waveforms are retained for location analysis. After stacking, P phases become clearer, in many cases allowing for hypocentre refinement using more phase picks than we are able to achieve with the corresponding primary templates. The secondary template is then re-cut to 6 s surrounding the S arrival and used in a further cross-correlation detection routine.

In general, the process of iteratively stacking and cross-correlating increases the number of detections. We repeat this iterative routine until there is less than a 10% change in detections (either greater or fewer detections), or until the fifth iteration. Once the final stack has been generated, we re-evaluate the P and S picks and re-compute the hypocentre locations (Figure 3.3). All final template stacks are located with P-picks at two or more stations and S-picks at every station.

Some SAMBA stations have encountered problems with noise introduced by solar controllers (LABE, Figure 3.4) [Boese, 2012, Boese et al., 2012]. Although the correlation detection routine does not initially detect this noise, any spikes present within a detection will be incorporated in the following, stacked, template. This results in the next correlation routine correlating the spike rather than the LFE. To combat, this we have manually scanned each record and removed channels for time periods showing spiking (< 6 months data removed at each of LABE and POCR). As we require at least five stations for a detection, weaker templates with initial picks on only five station do exhibit fewer detections when we have data loss. De-spiking routines were not used as they were found to reduce initial detections and introduced other noise into the stacked templates.

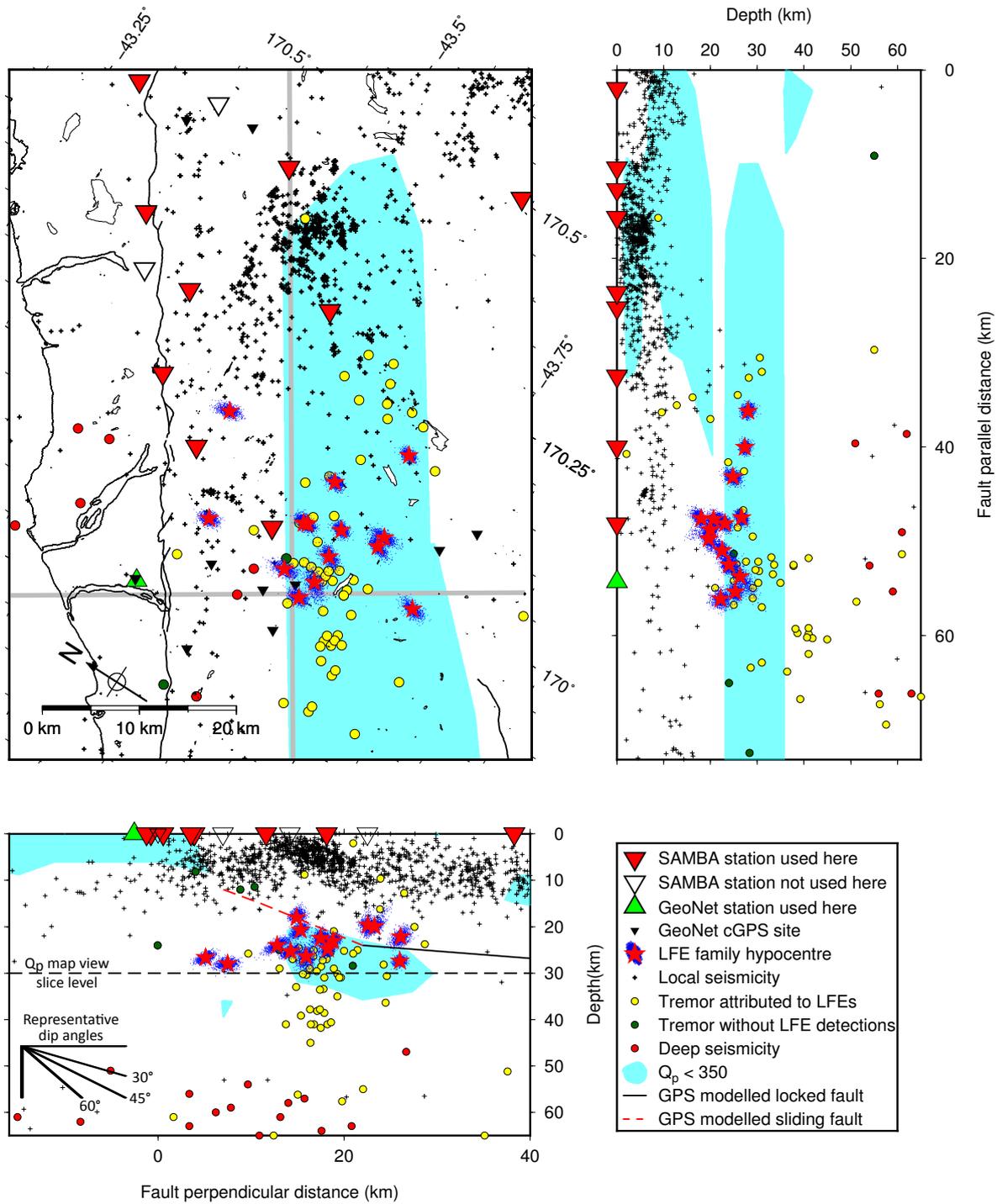


Figure 3.3: Hypocentre locations of 14 LFE families identified in this study, with uncertainties plotted as blue cloud of probability density function. Note that uncertainties are systematically underestimated due to unquantifiable uncertainties in velocity model, therefore these uncertainties should be considered minimum uncertainties. Local seismicity [Boese et al., 2012, 2014], deep seismicity [Boese et al., 2013] and previously identified tremor [Wech et al., 2012, 2013] are plotted as in the legend. Cross-sections projected onto gray lines in map view. GPS modelled fault shown on lower cross section after Lamb and Smith [2013]. Low  $Q_p$  region (at 30 km depth in map plot as shown by dashed line in fault perpendicular cross-section) shaded in cyan after Eberhart-Phillips et al. [2008]. All coloured triangles indicate stations on which LFEs have been detected. Note the vertical contraction in both cross-sections.

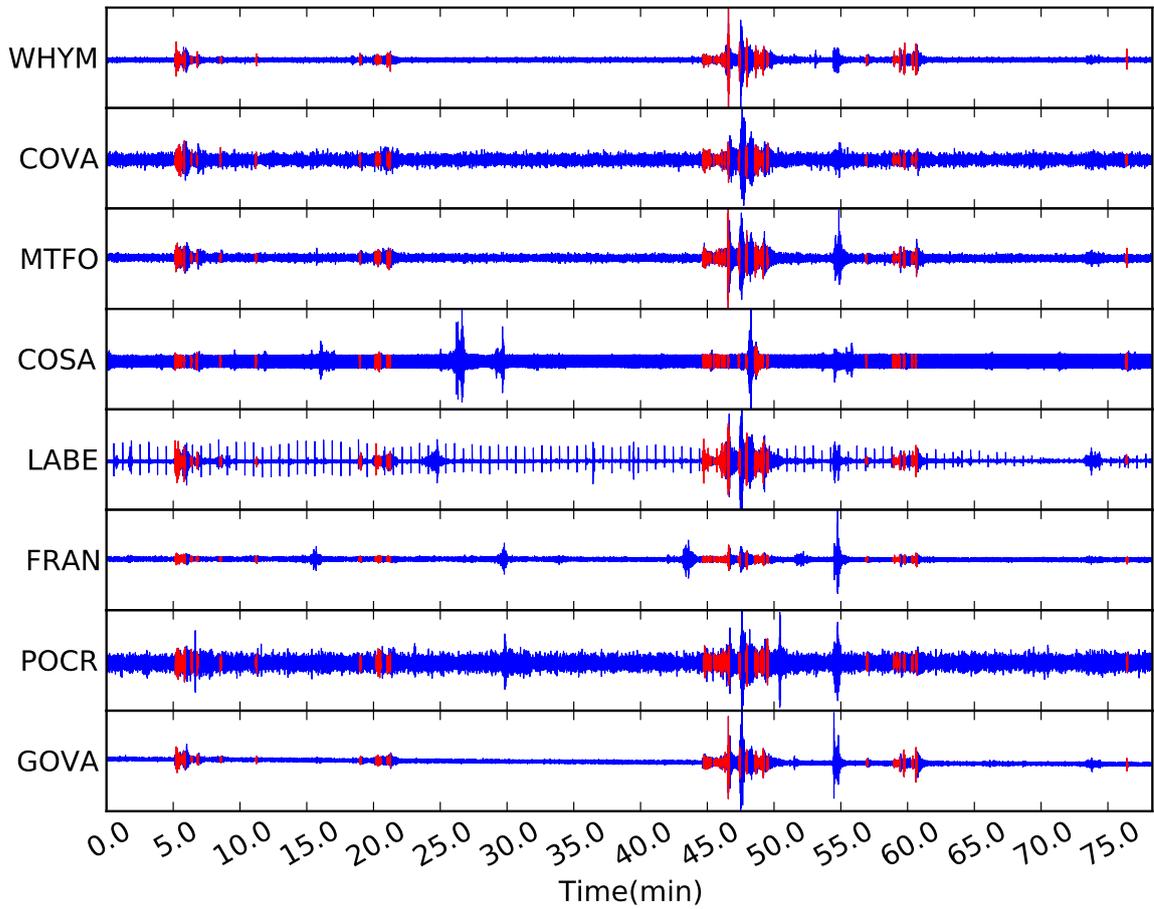


Figure 3.4: LFE detections (red) within tremor identified by Wech et al. [2012] on eight SAMBA stations. Plot start time is 2010 August 20 04:41:40.00 UTC. Amplitudes are normalised and data bandpassed 2–8 Hz. Tremor is discontinuous within the plot window, with the main tremor burst highlighted by Wech et al. [2012] between 43 and 50 minutes in this plot containing many LFE detections. Obvious spikes in time series recorded on LABE are due to solar-controller defects, the effect of these spikes is discussed in the text.

### 3.3.2 Magnitude calculation

For each detected event at each station and channel we bandpass at 2–8 Hz the 6 s window in which the detection was made. We then remove the instrument response to estimate ground velocity and find the minimum and maximum amplitudes and the time between them to give the peak ground velocity (PGV) and frequency. The PGV is then converted to a Wood-Anderson amplitude at the specific frequency determined. Bandpassing and using a single PGV at a single frequency mitigates the effects of low-frequency amplification of microseismic noise when simulating Wood-Anderson responses for an unfiltered time-series [Havskov and Ottemoller, 2010]. This amplitude is finally converted to  $M_L$  using the coefficients derived for local earthquakes by [Boese et al., 2012] recorded with SAMBA. We found this routine to give magnitudes within  $\pm 0.2$  of the published magnitudes of local seismicity [Boese et al., 2012]. For template LFE events, the automated routine gives the same PGVs and magnitudes as manual picking.

We remove spuriously large and small single-channel magnitudes ( $M_L > 2.0, M_L < -4.0$ ), which often correspond, respectively, to detections within the coda of large earthquakes or periods of solar controller-induced spiking [Boese et al., 2012], and magnitudes determined on the vertical component. We are always left with more than three magnitude estimates for each event (average 16 channels used). A final event magnitude is given by the mean of all single-channel magnitude estimates for each event. Single-event PGVs are given by the median of the single-channel PGVs for that event. We do not attempt to correct for source-station distance in this calculation.

### 3.3.3 Event verification

To determine an appropriate MAD multiplier for our detection threshold, we have undertaken synthetic testing following the method of Rawlinson [2011]. This involves seeding a day of random noise with copies of real LFE templates at random locations and with random SNRs before applying the same cross-correlation algorithm used in detection.

This testing shows that using a threshold of  $8 \times \text{MAD}$  will give between 0 and 3 false detections per day per template (mean of 0.2 for days of high noise, such as after large regional earthquakes, and mean of 0.01 for quiet days without large earthquakes) and detect 70 to 90% of all actual LFEs above a SNR of 0.2, depending on the template and noise conditions. Testing reveals that a complete (100% positive detections) catalogue is achieved at approximately SNR=0.8 throughout the time period analysed. By empirically determining the daily noise levels, we determine an average magnitude of

completeness ( $M_c$ ) of 0.6. This  $M_c$  is strongly time varying, however, with fluctuating noise levels (Figure 3.5) resulting in completeness values varying between  $-0.5$ – $2.0$ . The implications of this time varying  $M_c$  are discussed later.

We do expect a higher false detection rate during the aftershock sequences of large regional earthquakes. However, of the 1400 tests run at  $8 \times \text{MAD}$  throughout a day of phase-randomised data containing the amplitude information from the Dusky Sound earthquake and its aftershocks, in 1176 (84%) of tests we made 0 false detections and 196 (14%) of tests contained 1 false detection. The maximum false detections per template per day is 3, at a threshold of  $8 \times \text{MAD}$ . This maximum false detection rate was found in 2 (0.1%) test runs through high noise conditions. For further details of the synthetic testing see Appendix D.

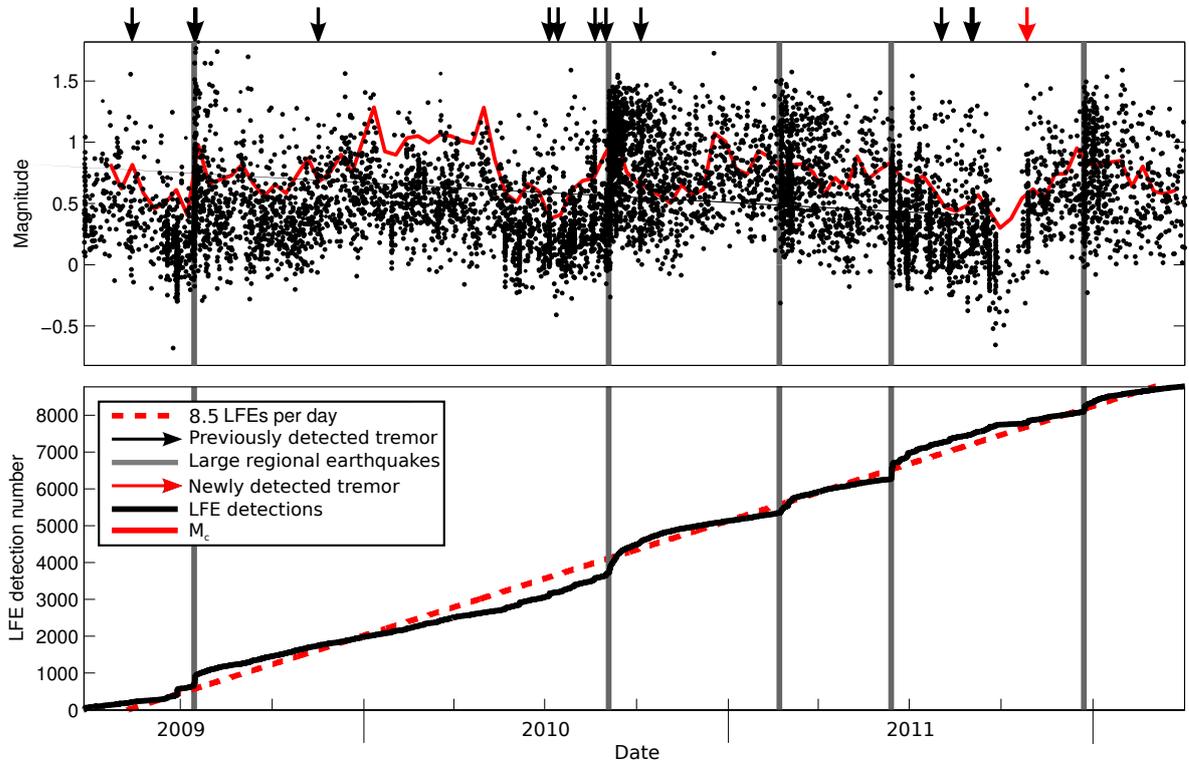


Figure 3.5: LFE detection rates (bottom panel) and magnitudes (top panel) with time. Magnitude of completeness variations with time (10 day running average) plotted as a red line in the upper panel. The apparent gap in detections between September and October 2011 is due to data corruption due to spikes as discussed in the text. Gray bars indicate the date of certain major regional earthquakes as discussed in the text (section 3.5.2). Black arrows indicate tremor periods identified by Wech et al. [2012]. A simple linear fit corresponds to 8.5 LFE detections per day is plotted as a red dashed line. We identify a possible new tremor period outside of the time period analysed by Wech et al. [2012] as a spike in LFE generation (highlighted by a red arrow); this has been confirmed as tremor by applying standard tremor waveform envelope cross correlation techniques and observing similar tremor locations to the LFE source region. We also see rate increases after the regional events indicated here, suggesting LFE and tremor triggering or modulation: tremor was active both before and after the Dusky Sound  $M_w$  7.8 [Beavan et al., 2010] (15 July 2009) earthquake.

## 3.4 Observations

The fourteen LFE families in our catalogue occur within a cross-sectional area of  $c. 21 \times 36$  km and depth extent of 15 km, taking into account the uncertainties in locations in *NonLinLoc* [Lomax et al., 2000]. The hypocentres appear anti-correlated in depth with ‘typical’ seismicity, lying beneath the seismogenic zone (0–12 km, [Boese et al., 2012]). Twelve of the fourteen families lie on or near the Alpine Fault inferred from GPS measurements [Lamb and Smith, 2013], coincident with deep reflectors [Okaya et al., 2007] within a zone of low seismic velocities [Smith et al., 1995, Stern et al., 2001] and high P-wave attenuation (low  $Q_p$ ) [Eberhart-Phillips et al., 2008] (Figure 3.3). The other two families appear to be separated from the main cluster of families and lie to the south-west of the inferred Alpine Fault. LFEs also occur at markedly shallower depths than deep subcrustal earthquakes [Boese et al., 2013], suggesting a different process of seismogenesis.

LFEs are ubiquitous throughout the record, occurring at a relatively constant long term rate of 8.5 per day (Figure 3.5), rather than being confined to tremor periods. They are nevertheless clustered, with 50% of detections having inter-event times of less than 1000s, and 11% having inter-event times less than 8 s (Figure 3.6). The clustering of inter-event times close to 0s is typical of LFE swarms coinciding with tremor periods identified by Wech et al. [2012] and with newly identified tremor periods verified here by eye and found to be located near the LFE source region using the WECC method [Wech, 2010] (Figure 3.4).

Tremor periods previously identified by Wech et al. [2012] correspond to abrupt increases in LFE detection rates (Figures 3.4 and 3.7). We are able to identify all but three of the nineteen tremor bursts [Wech et al., 2012] recognised, reinforcing the idea of tremor being composed of LFE swarms. LFEs account for roughly 12–24% of the previously detected tremor, depending on whether phase-delays due to traveltimes are included or not. This estimate does not include the effects of scattering or multiples and should be considered a minimum value. Two of the three tremor periods not identified here were located at least 15 km from our template locations and at least 10 km from the nearest detecting station. One location for tremor not detected by LFEs however, lies near to the region of LFE detections. For the tremor episode on 3 August 2011, two distinct tremor locations were calculated by Wech et al. [2012] within 5 min of each other, one within the LFE source region and one  $c. 15$  km southwest of the LFEs. We suggest that the tremor location within the LFE source region may be an artefact and the second location (to the southwest) is more likely. Attempts to pick LFEs within these three bursts of tremor have not yet proven fruitful, with fewer than five stations showing detections. An ongoing southern extension of the SAMBA may yield greater

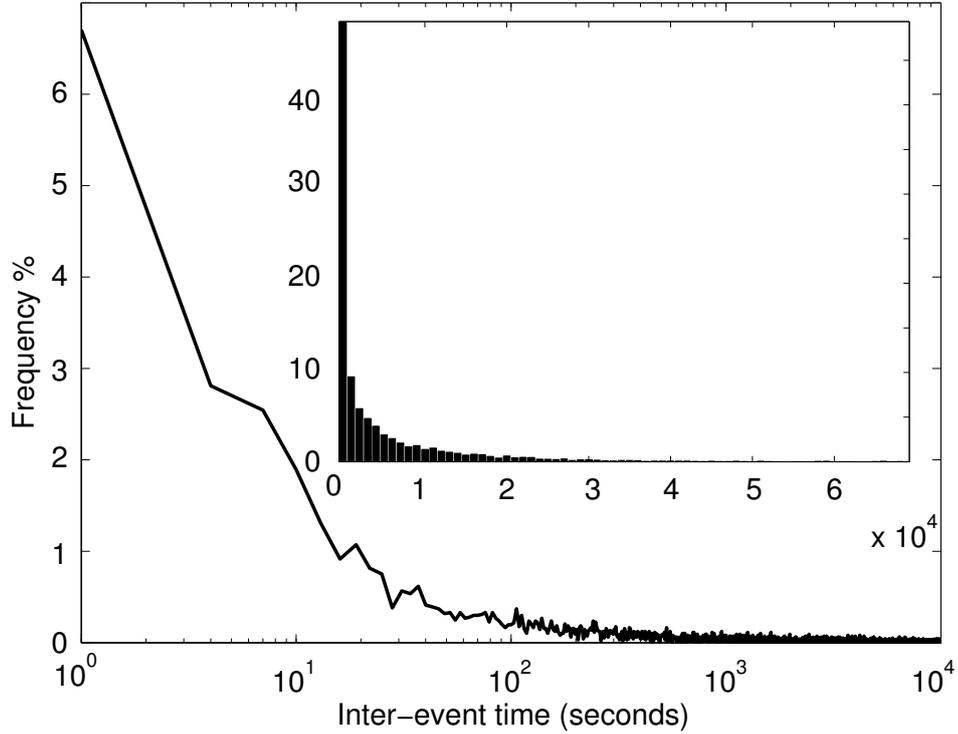


Figure 3.6: Inter-event times for all detected events. Main panel shows events within 3 s bins on a logarithmic scale, inset has a bin size of 1000 s and uses a linear scale. Inter-event times peak between 0 and 3 s, with 27% of all detections having inter-event times less than 100 s.

template coverage for future studies.

One of the strengths of the iterative cross-correlation and stacking method is our ability to identify LFEs within noisy time-series, and in particular within aftershock sequences and during the mainshock coda of large regional earthquakes (Figure 3.8). Following five large regional earthquakes during the study period (Dusky Sound 15 July 2009,  $M_w$  7.8 [Beavan et al., 2010]; Darfield (Canterbury) 3 September 2010,  $M_w$  7.1 [Quigley et al., 2012]; Christchurch 21 February 2011,  $M_w$  6.2 [Holden, 2011]; Christchurch 13 June 2011,  $M_w$  6.0 [Holden, 2011], and Christchurch 23 December 2011  $M_w$  5.9 [Bannister and Gledhill, 2012], dates in UTC) we observe distinct increases in LFE rate above the expected increase in false detections (see Section 3.3.3). After each earthquake, we see a decay towards the background rate, but the duration of the decay after each earthquake is different. Reliable magnitude estimation within aftershock sequences remains problematic, with high LFE magnitudes within the aftershock sequences assessed here often corresponding to earthquake contamination.

Each LFE family analysed shows a slightly different temporal evolution during known tremor periods and in response to large regional earthquakes (Figure 3.9). Nevertheless, the range of magnitudes and average magnitudes of each family are quite similar. Family ID:61044 has the largest mean magnitude ( $M_L$  0.9) and highest number of detections (1,847; Table 3.1) suggesting that detection rate and magnitude are correlated;

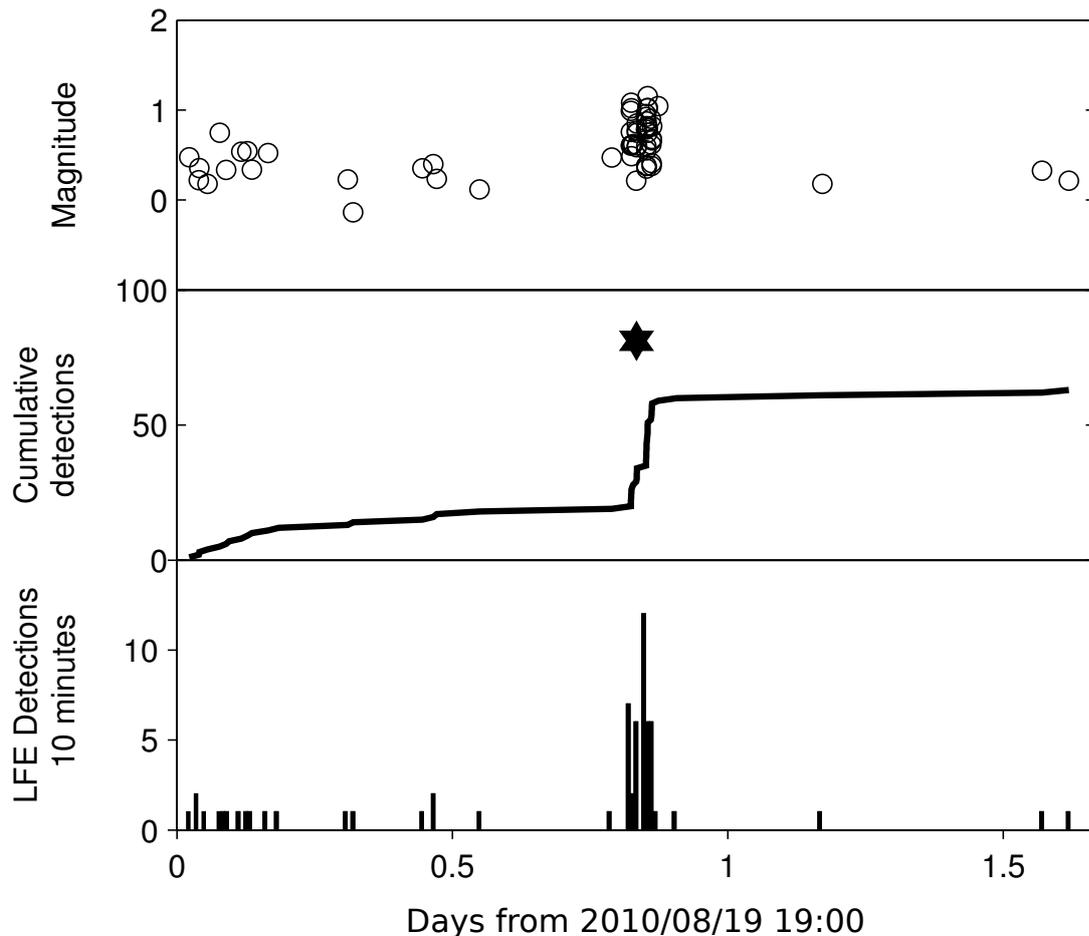


Figure 3.7: LFE detections for a representative tremor period (onset marked by black star) previously detected by Wech et al. [2012] at UTC 2010 August 20, 04:46–05:57. Upper plot shows LFE magnitudes with time, the middle plot shows the cumulative detections, and the bottom plot shows the detections per hour. Tremor periods observed through LFE detection display a sharp spike in LFE detections at tremor onset (5–15 detections per 10 minutes) without major variation in magnitudes or change in background detection rate.

however, this is not the case because, for instance, family ID:37575 has a similarly high number of detections (1,562), but a much lower mean magnitude ( $M_L$  0.4). Basic analysis shows no trend in our small dataset between family mean  $M_L$  and number of detections. In general, the cumulative detections of each family is quite consistent, with only two families containing more than 1,000 events.

The LFEs we have detected do not exhibit a simple Gutenberg-Richter power law scaling, as plots of the logarithm of the cumulative frequency against magnitude (Figure 3.10) are not well fit by a linear gradient above the  $M_c$ . The implications of this are discussed below. However, PGV amplitudes do approximately follow an exponential law for individual LFE families as reported previously by Shelly and Hardebeck [2010] for the San Andreas fault. We see evidence for two main forms of this exponential rela-

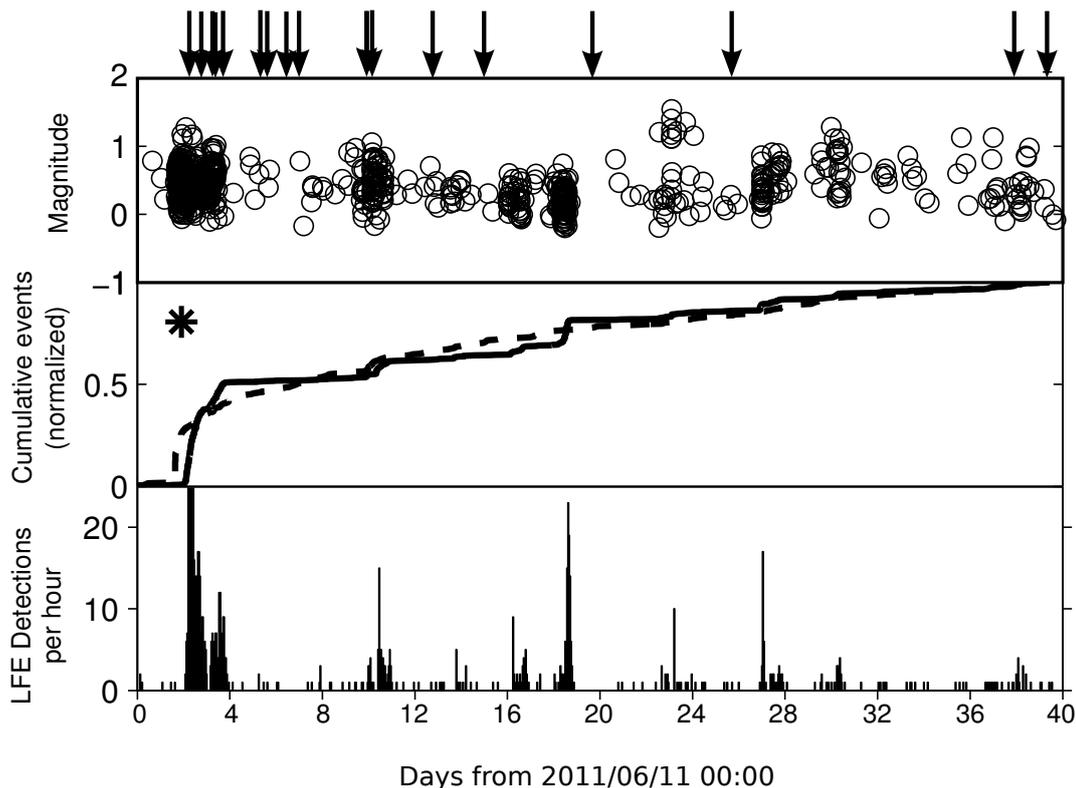


Figure 3.8: LFEs detected after the June Christchurch earthquake ( $M_w$  6.0 [Holden, 2011], UTC 2011 June 13 02:20 origin time plotted as black asterisk). Upper plot shows LFE magnitudes with time, the middle plot shows the normalized cumulative detections for both LFEs (solid line) and ‘typical’  $M_w > 3.0$  earthquakes (dashed line), and the bottom plot shows the detections per hour. LFE detections after large regional earthquakes show more gradual response with detection rates increasing to 5–15 per hour, a larger range of magnitudes, and a steady decrease in rates towards background levels. This decay towards background levels is more staccato than that seen in ‘typical’ earthquake aftershocks (dashed line upper panel), with spikes in detection rate associated with larger aftershocks suggesting triggering by earthquakes of lower magnitudes than the main  $M_w$  6.0 earthquake. Earthquakes  $M_w > 4.0$  in the GeoNet catalogue covering the whole of South Island, highlighted by black arrows.

tion, with five families showing similar gradients to family ID:61220 (Figure 3.10) and the other nine families exhibit a shallower gradient similar to family ID:37575.

For the two LFE families with much higher numbers of detections than average, the majority of events occur closely spaced in time (Figure 3.9). For family ID:61044, 30 % of the 1,847 detections were made following the Darfield earthquake. This family showed a muted response to the Dusky Sound earthquake in 2009 but more pronounced responses to the Christchurch earthquakes following the Darfield earthquake. Family ID:37575 produced nearly 20 % of its constituent LFEs within a month of the 13 June 2011 Christchurch earthquake, and shows a marked increase in activity after this earthquake. This family showed a decreased rate of detection after the Darfield earthquake, in contrast to family ID:61044.

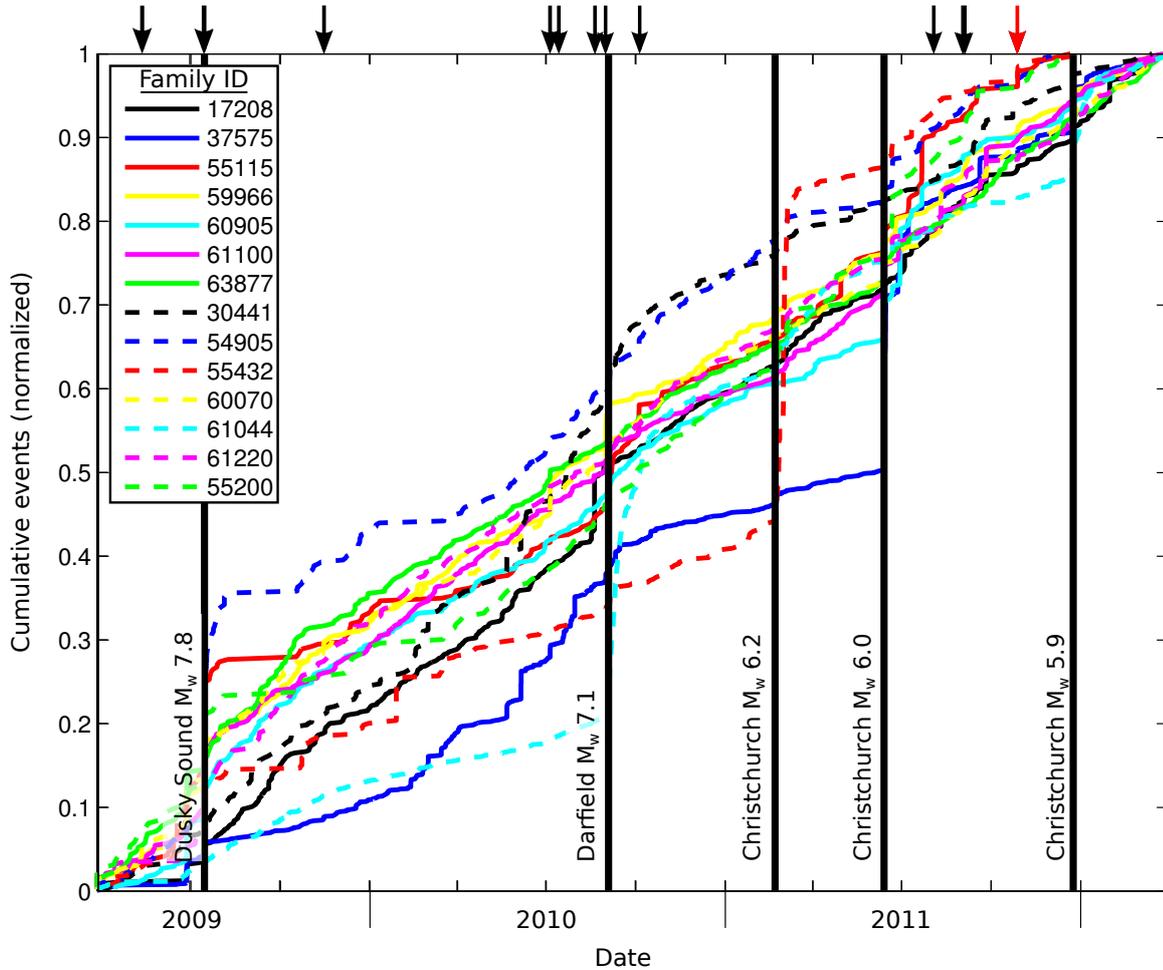


Figure 3.9: Normalized detections for each family with previously detected tremor events [Wech et al., 2012] indicated by arrows and large regional earthquakes seen to elicit an LFE response indicated by solid black vertical lines. Red arrow indicates a newly identified tremor period defined by a spike in LFE generation. Individual families respond differently to different tremor events and earthquakes. Notably, the family with the most detections (Family ID:61044, 1,847 detections) responds most strongly to the Darfield earthquake and also shows a strong response to the two Christchurch earthquakes. Note this figure differs from that published by slight changes in event detections to match the full catalogue.

LFE detections made within aftershock sequences do not appear to be ‘typical’ aftershocks. Extensive detection testing within synthetic aftershock sequences has shown that although our false detection rate increases, this increase is unable to account for the overall detection rate increases observed. We have also confirmed by manual assessment for a day of data within the aftershocks of the Darfield earthquake that LFE detections are not made at the same time, or crucially, with the same station-delays as aftershocks. Furthermore LFE decay patterns after triggering events do not follow the same aftershock decay pattern as the typical earthquake aftershocks of these events: rather, we observe spikes in LFE detections concurrent with the mainshock and subsequent large events with periods of relative quiescence between (Figure 3.8).

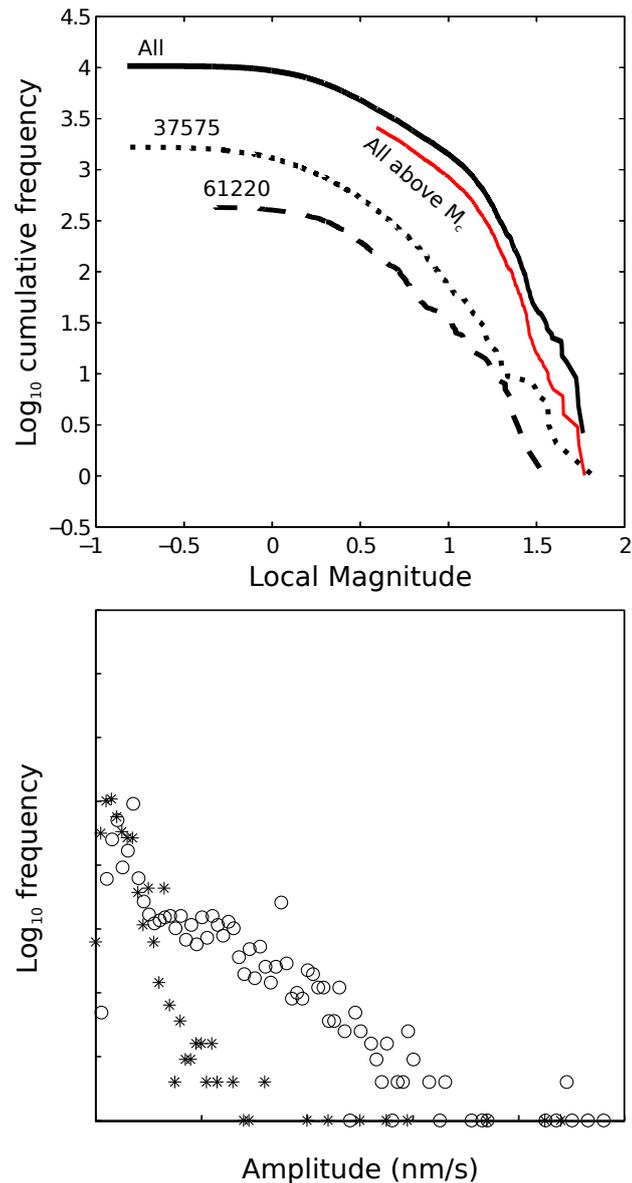


Figure 3.10: Cumulative frequency–magnitude (upper) and frequency–amplitude (PGV) (lower) relations for all LFEs (solid line and crosses), family ID:37575 (dotted line and stars) and family ID:61220 (dashed line and circles). Red line in upper panel show the frequency–magnitude relation for the complete part of the catalogue (LFEs with  $M_L \geq M_c$  for days with  $M_c \leq 0.6$ ) which follows a similar relationship to the full catalogue (black). PGV amplitudes scale approximately linearly with the logarithm of frequency, with each family having a different scaling parameter. Individual family scaling parameters have two dominant forms exemplified by families ID:37575 and ID:61220. The sum of these two forms is evident as a spike in occurrence at low PGV amplitudes in the plot of all LFE detections due to the amplitude relation of families similar to family ID:37575, followed by a change in gradient due to families more similar to family 61220.

## 3.5 Discussion

### 3.5.1 LFE family locations

Tremor and LFEs detected at plate boundaries have been interpreted to be the seismic manifestation of slow shear slip [e.g. Ide et al., 2007b, Wech and Creager, 2007, Shelly, 2010, Rubin and Armbruster, 2013]. In this case, the consistency in the locations of LFEs with the location of the deep extent of the Alpine Fault inferred by other means (GPS inversion [Lamb and Smith, 2013], seismic reflection [Okaya et al., 2007] and velocity models [Stern et al., 2001, Smith et al., 1995]), and the contemporaneous detection of tremor and LFEs suggest that the LFEs represent slow shear slip on the deep extent of the Alpine Fault. However, without reliable focal mechanisms, and given the small number of LFE families detected to date, we cannot yet unequivocally relate the LFEs detected in this study to shear slip. We are hopeful that the improved azimuthal coverage provided by the recent southern SAMBA extension (stations SOLU and MTBA in Figure 3.1) will enable focal mechanism parameters to be evaluated in the future. We note also that the locations presented here have under-estimated uncertainties due to insufficient information to constrain velocity model uncertainties; therefore these uncertainties should be considered minimum uncertainties.

We hypothesise that these LFEs are generated on asperities within an otherwise aseismic, creeping region of the Alpine Fault, following the model of Ide [2008]. Due to low SNR values of the LFEs, and corresponding low single-channel correlations (average 0.5), we have not yet been able to reliably relocate the LFEs within individual families. However, because we do not detect the same event with more than one template, we infer that the scatter of hypocentres within an individual LFE family is smaller than the average distance between different families.

Twelve of the fourteen LFE locations in this region are clustered in space, near the inferred change in dip from a steep ( $45\text{--}60^\circ$ ) stably sliding fault updip of the LFE generation region, to a shallower-dip ( $15\text{--}20^\circ$ ) partially locked zone [Lamb and Smith, 2013]. We hypothesise that this transition allows for near-constant slip while having enough asperities to generate LFEs. The remaining two locations may be indicative of a more steeply dipping strand of the Alpine Fault as discussed by Norris and Toy [2014], but further events are required to substantiate this.

Some tremor depths estimated by Wech et al. [2012] exceed those of the LFEs reported here. It is possible our analysis misses deep tremor due to low SNR values, but during all deep tremor periods identified by Wech et al. [2012] we see active LFE generation at shallower levels. We thus suggest that the discrepancies in estimates of depths reflect the large uncertainties due to tremor-source geometry and network limitations Wech

et al. [2012] pointed out.

Tremor and LFEs are often interpreted to be associated with fluids, either in relation to fluid migration or high fluid pressure resulting in reduced normal stress and allowing for shear slip [e.g. Shelly et al., 2006, Thomas et al., 2009, Bostock et al., 2012, Wech et al., 2012]. If the LFEs documented here do indicate deep slip on the Alpine fault as we hypothesise, then the low  $Q_p$  and low seismic velocity that characterise the source region (Figure 3.3) may in some manner be associated with fault weakening. Eberhart-Phillips et al. [2008] interpreted the region of low  $Q_p$  here to be the down-dip extension of the Alpine Fault beneath South Island, and the high attenuation itself to mark the presence of metamorphic fluids. This is compatible with the interpretation of Stern et al. [2001] who concluded that the low seismic velocities correspond to increased fluid pressures and associated fault weakness.

The fourteen LFE templates and their subsequent detections are unlikely to represent all of the tremor on the Alpine Fault in this region; instead they represent the strongest LFE signals detected with a sub-optimal seismic network. The recent extension of SAMBA southwards should allow future studies of LFEs and tremor to be made with lower magnitudes of completeness and greater spatial coverage.

As noted by other authors [e.g. Boese et al., 2012, Wech et al., 2012, Boese et al., 2013] there is a distinct change in the pattern of seismicity along strike on the Alpine Fault. Of particular relevance to this study is the lack of LFE detections north-east of Mt. Cook. This change is coincident with the abrupt change in shallow seismicity noted by Boese et al. [2012] and in deep seismicity [Boese et al., 2013] (Figure 3.3). However, as we have only looked for LFE templates within known tremor periods, also seen to locate south-west of Mt. Cook, the lack of LFE locations away from tremor locations is unsurprising. We therefore cannot rule out LFE generation north-east of the current LFE source region, but suggest that if LFEs are generated to the north-east they do not generate tremor as demonstrated by Wech et al. [2012].

Seismic [e.g. Okaya et al., 2007] and geological [e.g. Little et al., 2005] evidence suggests that the geometry of the Alpine Fault changes in the region beneath Mt. Cook from a listric fault south of Mt. Cook to a near-vertical structure in the north. This change in geometry may be associated with the termination of a partially subducted remnant passive margin attached to the Australian plate and extending south-westward beneath the plate boundary [Sutherland et al., 2000]. [Boese et al., 2012] considered that the partially subducted remnant passive margin may explain the transition in deep seismicity beneath South Island. The similarity in termination points for deep earthquakes, tremor generating LFEs and the proposed remnant passive margin suggests a common control.

### 3.5.2 Inter-event timing and magnitudes

We observe tremor beneath the central Southern Alps to be a superposition of multiple LFEs occurring with small inter-event times in a swarm-like manner, as demonstrated by Shelly et al. [2007]. However, a large portion of the catalogue presented here exhibits discrete behaviour, with large inter-event times that are not tremor-like. As our catalogue is largely composed of events near or at the limit of our detection and outside our network (Figure 3.3) we cannot be certain that those discrete events within the catalogue are not related to smaller events, or events outside our detection region.

Our observations of near-constant LFE generation rate suggest that LFEs represent quasi-continuous punctuated creep on the deep extent of the Alpine Fault, and provide the first evidence of such slip on the deep extent of the Alpine Fault. With magnitudes apparently at or near our detection threshold ( $M_c$  0.6) and deep hypocentres, the lack of geodetically detectable slow slip is unsurprising. Detection of slow slip is also hindered by the constancy of LFE generation and hence slow slip, with brief swarms (tremor events) lasting 5–120 mins suggesting similarly brief slip rate increases. Abrupt LFE detection rate increases coincident with large regional earthquakes may be indicative of triggering by transient or static stress shifts [Boese et al., 2014]. A full investigation of LFE triggering characteristics in the Southern Alps is the subject of ongoing research and beyond the scope of this paper. However, initial analysis of the apparently triggered LFE rate increases shows that LFE decay rates following regional earthquakes do not follow the same decay pattern as ‘typical’ aftershocks generated in the brittle crust by the same events.

We are not aware of LFE catalogue completeness being computed before. Shelly and Hardebeck [2010] provided some indication of completeness by demonstrating the level at which LFE size stops following an exponential law, but this approach requires the assumption that LFEs follow an exponential law. Rather than imposing this condition (or indeed assuming a G-R law), we have estimated the completeness empirically based on the SNR values required for 100% detections in synthetic testing (Section 3.3.3). Analysis of the magnitude-frequency characteristics above this level shows no relationship between magnitude and recurrence interval. This is again likely to be a characteristic of our swarm-dominated catalogue.

The fact that  $M_c$  varies from day to day does not preclude analysis of the frequency-magnitude characteristics of the LFE dataset as a whole. Events of  $M_L \geq 0.6$  occurring on the subset of days for days on which  $M_c \leq 0.6$  exhibit the same frequency-magnitude characteristics as the full catalogue (Figure 3.10). If LFEs scaled in a similar fashion to ‘typical’ repeating earthquakes, we might expect a linear relationship between the logarithm of cumulative frequency and magnitude, but this is not the case. Rather,

individual families are well fit by an exponential model above  $M_c$  and below a maximum magnitude value. This maximum magnitude may be an indication of the limit of fault (asperity) size as identified on a global scale for ‘typical’ earthquakes by Kagan [2002].

The approximately exponential distribution of LFE amplitudes seen here is consistent with the San Andreas LFE amplitude distributions shown by Shelly and Hardebeck [2010], albeit here at much higher amplitudes, and with the observations of Watanabe et al. [2007] of amplitude-duration distributions in non-volcanic tremor. This exponential distribution is also consistent with tremor amplitude-duration characteristics in volcanic settings [e.g. Benoit et al., 2003]. Such an exponential relationship in tremor implies that the source has a characteristic scale. In volcanic settings this can be attributed to constraints imparted by the magmatic plumbing system [e.g. Aki and Koyanagi, 1981, Benoit et al., 2003]. Watanabe et al. [2007] argued that a characteristic length scale for the non-volcanic tremor source process, possibly the length of fluid-filled cracks, is required to describe the exponential relationship found in tremor properties in southwest Japan.

The source process for non-volcanic tremor and LFEs remains poorly understood. The exponential frequency-magnitude relationships seen here may be related to a characteristic length scale for individual LFE families. This characteristic length-scale is likely to be related to the asperity size, the limit of which may be inferred from the cut-off magnitude. If this is the case then LFE magnitude could be used to directly monitor the amount of slip at the deep extent of faults. However, the frequency-magnitude characteristics could also be related to some process required to slow earthquake rupture in the generation of these ‘slow’ earthquakes [Kaproth and Marone, 2013], related to the damping term described by Ide [2008].

### 3.6 Conclusions

We have presented the first evidence of LFEs associated with the Alpine Fault in New Zealand’s South Island. Our results demonstrate that tremor occurring on or near the Alpine Fault is composed of LFE swarms, but that LFEs also occur quasi-continuously without generating extended-duration tremor. LFEs in our catalogue exhibit a range of inter-event times dominated by short inter-event times coincident with swarms (tremor) while also containing many temporally isolated events. We interpret that these discrete events indicate that the deep extent of the Alpine Fault is undergoing quasi-continuous creep punctuated by rate increases manifesting as tremor.

We have located fourteen LFE families and found that twelve of these fourteen families lie within c. 10 km of the Alpine Fault inferred from GPS measurements [Lamb and

Smith, 2013], within an area of high seismic reflectivity [Okaya et al., 2007], anti-correlated in depth with both shallow [Boese et al., 2012] and deep [Boese et al., 2013] seismicity in the region. We conclude that LFEs (and therefore tremor) in the Southern Alps are the manifestation of slow shear slip on the plate interface. Furthermore, twelve of fourteen LFE families lie within a zone of low  $Q_p$  which we interpret to be a sign of fluid-weakened rock, thereby allowing for ongoing slow slip on an otherwise locked region of the Alpine Fault. We suggest that these fluids may be generated as the products of metamorphism of a partially subducted passive margin which terminates beneath Mt. Cook. This also may explain the lack of LFEs and tremor north-east along strike from our LFE families.

The LFE magnitudes we compute do not scale according to a power law, but instead appear to scale according to an exponential law suggestive of a characteristic length scale in LFE source generation. We interpret this to be indicative of constant slip patch sizes for individual LFE families associated with the size of asperities upon which LFEs nucleate in an otherwise stably sliding region of the Alpine Fault.

Further study using data from the forthcoming SAMBA extension south of the current LFE source region will further constrain the extent of LFE generation and lead to a more representative view of slip occurring at depth on the Alpine Fault.

## 3.7 Amendum

In the published version of this chapter the number of detections in Table 3.1 was erroneous, this affected Figure 3.9 and is corrected here. This occurred as numbers were taken from the previous iteration in the iterative stacking routine. This error does not affect the conclusions made here, or the discussion points. Numbers of detections in the text quoted from Table 3.1 have been corrected also. We have also added statements in the text and in captions to highlight the under-estimated uncertainties that arise from velocity model uncertainties. These uncertainties are not quantifiable with such a small location set.

# 4 Earthquake detection using synthetic templates: developing a continuous catalogue of low-frequency earthquake templates for New Zealand's central Southern Alps

## 4.1 Abstract

We develop a method of detecting earthquakes employing simple synthetic waveforms as templates in a large-scale, network-based matched-filter routine. To enable us to compute these detections efficiently we have also developed an open-source, multi-parallel Python package. We first demonstrate that the synthetic templates are capable of detecting both discrete and swarm-like shallow seismicity, and subcrustal earthquakes. By constructing synthetic S-phase waveforms to represent possible low-frequency earthquake sources, we are able to detect a continuous (6.5 year-long) suite of events, which we then stack to generate real low-frequency earthquake templates. We use only the S-phase for low-frequency earthquake templates as we do not know the focal-mechanism *a-priori*. Using only the S-phase results in more false detections than using the P and S phases. We overcome false detections when generating real template waveforms by stacking, for which only the true detections stack coherently. In this manner we generate a spatially and temporally continuous catalogue comprising 600 LFE templates. The preliminary hypocentres of these low-frequency earthquakes are poorly constrained in depth due to a lack of P-phases in most template stacks caused by false detections in the initial synthetic detection stage, which reduce the coherency of the stack outside of the S-phase. This approach yields an objective catalogue of real templates for use in further low-frequency earthquake detection routines.

## 4.2 Introduction

Detection of low-frequency earthquakes (LFEs) is often undertaken using matched-filter techniques [e.g. Shelly et al., 2007]. A major weakness of these methods is their reliance on *a-priori* information (a template). In noisy environments such as the Southern Alps, or when signals are of low-amplitude or sparse, finding useful LFE templates manually is both time-consuming and unreliable due to inherent subjectivity. In our previous work [Chamberlain et al., 2014] we were able to detect only fourteen LFE families, in contrast, Frank et al. [2014] were able to detect and analyse 1,120 LFE families in Guerrero, Mexico using a systematic template detection method. The results of this catalogue have exemplified the importance of continuous LFE catalogues, Their catalogue has allowed for the detection of small slow-slip events [Frank et al., 2015a], analysis of temporal variations in LFE rates and studies of how these LFEs relate to physical models of the slow-slip region [Frank et al., 2015b]. These studies would not have been possible without a continuous catalogue of low-frequency earthquakes. To develop a similar catalogue in the central Southern Alps, while making use of the matched-filter technique to extract weak signals, we require a method of objectively identifying appropriate templates.

A number of methods are available for template detection, including auto-correlation [Brown et al., 2008], cross-station correlation [Rubin and Armbruster, 2013, Thomas and Bostock, 2015], linkage methods [Aguiar and Beroza, 2014] and brightness [Frank and Shapiro, 2014]. All these techniques have peculiarities which either preclude their usage, or reduce their effectiveness in the central Southern Alps. After extensive testing of a brightness approach [Frank and Shapiro, 2014], we found that we could not usefully discriminate between noise and signal; signal amplitudes are low and noise is often detected rather than signal. Cross-station coherences are also often lower for known signals than they are for noise, due to complex scattering and strong wind noise field [Petersen et al., 2011]. The poor cross-station coherence in the 2–8 Hz tremor band means that the cross-station correlation technique is also poorly suited to the current network configuration. The autocorrelation technique, while useful when analysing discrete tremor events, is so computationally expensive [Frank et al., 2014] that attempting to use this technique to develop a continuous set of LFE templates would be unreasonably slow: autocorrelation scales non-linearly with catalogue duration. Further comparison of the methods is given in Section 4.5.

To develop a suitable computationally viable method of detecting LFEs in the central Southern Alps we suggest using synthetic waveforms as preliminary templates in a large-scale cross-correlation detector. These detections can then be clustered and stacked to generate a uniform set of real template events. The term ‘real’ is used to re-

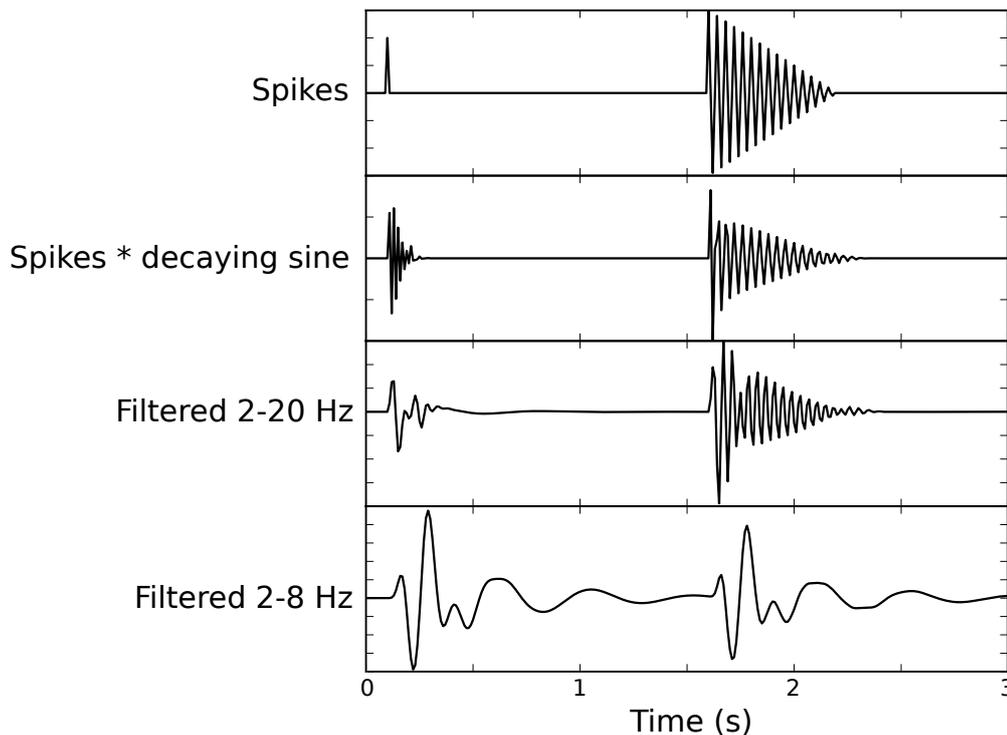


Figure 4.1: Generation of a synthetic template. Top to bottom: A time-series of spikes is constructed, which is then convolved with a 50 Hz decaying sine wave. This synthetic can then be filtered to better approximate real waveforms. Here the synthetic template is constructed with an S–P time of 1.5 s sampled at 100 Hz and with an S/P amplitude ratio of 1.5.

fer to recorded data, and ‘synthetic’ to refer to manually and automatically constructed waveforms.

In principle, a seismogram can be simulated by convolving known source properties with an appropriate Greens function [Aki and Richards, 2002]. In practice, however, accurate simulation of earthquake body waves is made difficult by the dominant high-frequencies and strong scattering resulting from path effects [Carrington et al., 2008]. These factors mean that estimating Green’s functions for all possible ray-paths is often not viable. Nevertheless, as we show below, dominant periods of 0.05–0.5 s features of real earthquake waveforms can be reasonably matched with simple approximations.

## 4.3 Methods

### 4.3.1 Template simulation

We represent the seismic waveform initially as a single spike representing the P arrival, and a series of spikes of diminishing amplitudes and alternating polarities to represent the S arrival and coda (Figure 4.1). The S–P time is set using a pre-determined

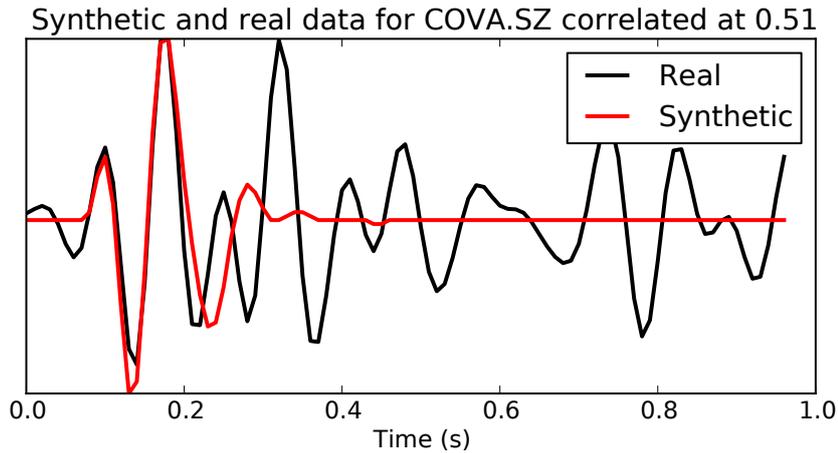


Figure 4.2: Synthetic (red) and real (black) P-waves for a crustal earthquakes located by Boese et al. [2014] as part of a swarm. The vertical channel seismograms for station COVA are shown (c. 15 km hypocentral distance, directly above the LFE source region identified by Chamberlain et al. [2014]). The P coda is not well matched by the synthetic, and the two waveforms diverge after the first cycle. Nevertheless the clear correlation of the early part of the waveform is strong enough to distinguish this from noise.

3D velocity model (Section 4.3.3). The difference in representations between P and S arrivals is chosen to simulate the typically more complicated coda of the S-phase. The spiked time-series is then convolved with a 50 Hz decaying sine wave, chosen as this is the Nyquist frequency of the lowest sampling-rate site used here. We generate a high-frequency signal rather than restricting waveforms solely to the 2–8 Hz band (where LFE signals are dominant in the central Southern Alps [Chamberlain et al., 2014]) to generalize the approach; filtering the waveform later allows a specific dominant frequency to be generated.

Single channel synthetic templates aligned with seismic phases correlate with real earthquakes between a 0.0–0.8 level (normalised cross-correlation computed within a 1 s, 100 sample window, Figure 4.2). Correlations for S and P phases are similar, meaning that the synthetics do not match P phases better than S or vice-versa. The most important characteristic of the templates is the lack of similarity between the synthetic and the background noise. This appears to be key: rather than trying extensively to match the signal, we seek to *not* match the noise. This method can be extended to generate a multi-component template for use in a network cross-correlation detection routine.

### 4.3.2 Comparison with known waveforms

To test our methodology we first compare the synthetic templates with real earthquakes, and then examine the detections made by real and synthetic templates. We

undertake the following four tests:

1. Comparison with shallow micro-earthquakes;
2. Comparison with a subcrustal earthquake beneath the LFE source region;
3. Shallow earthquake swarm detection and;
4. Comparison with a raw LFE template.

The first, second and fourth tests compare discrete real waveforms with synthetic waveforms, whereas in the third test we generate a grid of synthetic templates and attempt to detect a known swarm of shallow earthquakes in the central Southern Alps.

To compare with previously picked events synthetic templates are generated with arrivals at the time of the manually picked real P and S arrivals. For the regular earthquake tests, we filter between 2–10 Hz and use a 100 Hz sampling rate, whereas for the LFE test we use a 2–8 Hz filter and 20 Hz sampling rate to more closely represent the parameters used in large-scale cross-correlation (Section 4.3.4). Following this step, we cross-correlate the real template (generated by filtering the real seismic waveforms in the same frequency band as the synthetic, and cutting 0.1 s prior to the manual phase pick and 0.9 s after the pick) with the synthetic template to determine the absolute maximum of the cross-correlation and corresponding lag. We allow lags of as much as  $\pm 20$  samples or 0.2 s to allow for differences between the manual phase-pick and the theoretical arrival in the synthetic template. If the cross-correlation measured is negative we reverse the polarity of the synthetic. The synthetic template is shifted to achieve the optimum alignment with the real event (Figure 4.3). We also compare the synthetic template and real earthquake spectra (Figure 4.4), with obvious discrepancies in the timing of the arrival, which we correct for when aligning the templates.

For the first two tests, we use both the real waveforms and the synthetic waveforms as templates in a network cross-correlation detection scheme for the day of the real earthquake. For both the shallow and deep earthquakes, we are able to detect the target earthquake in the continuous data using the synthetic template (Figure 4.5). No other events are detected during the days analysed using the synthetics, nor using the real waveforms. For both cases the synthetic has a lower network cross-correlation sum than the real template by approximately a half. The synthetic network cross-correlation sum however is still greater than the  $8 \times \text{MAD}$  threshold used for that day of data. The shallow earthquake synthetic detection has a cross-correlation sum of  $12.8 \times \text{MAD}$ , and the deep equivalent has a cross-correlation sum of  $15.3 \times \text{MAD}$ . In both cases the MAD level for the synthetic template is lower than for the real template, indicating that our synthetic templates do not match the background (noise) as well as the real templates.

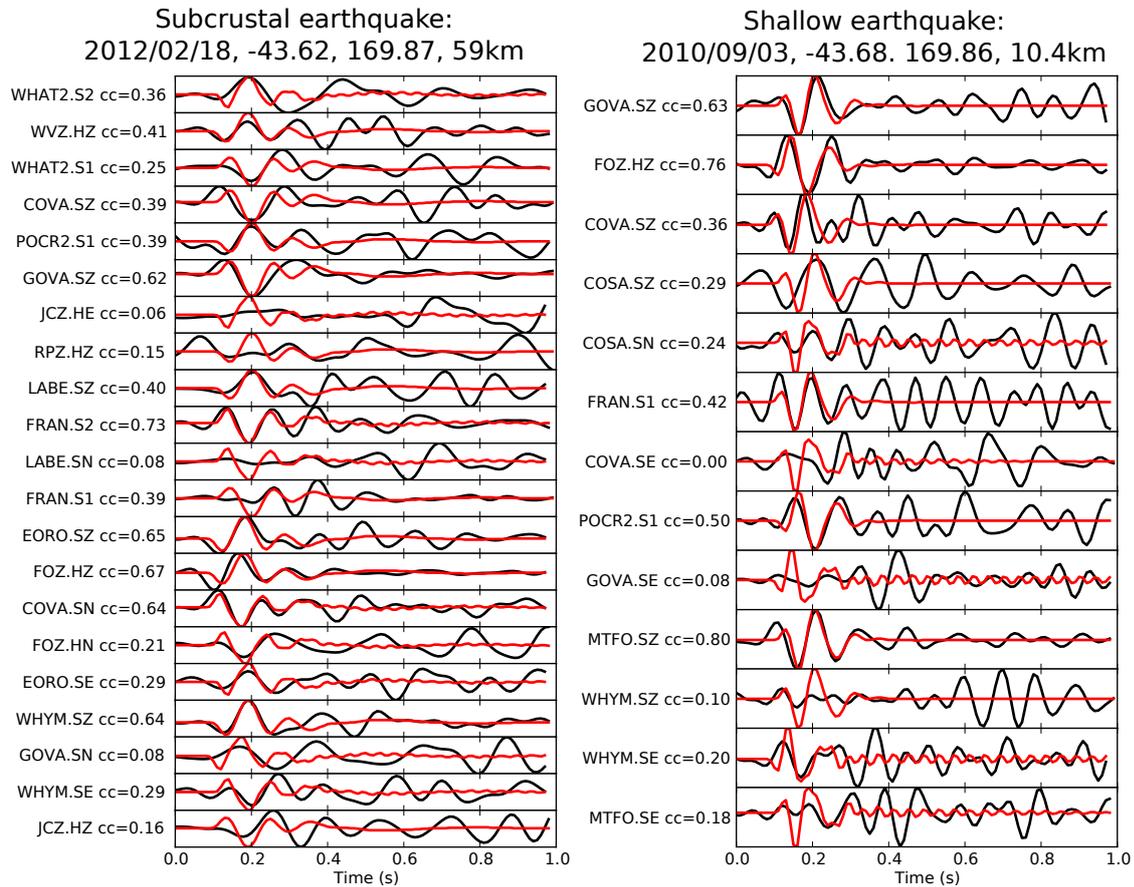


Figure 4.3: Real (black) and synthetic (red), 1 s-long templates for the shallow and subcrustal earthquakes examined. Waveforms are sampled at 100 Hz, bandpassed 2–10 Hz and aligned 0.1 s before the manual phase-pick. Channels are only plotted if a phase-pick (either P or S) was made on that channel. P-phase templates are displayed on the vertical (Z) channels and S-phase on the horizontal (1,2,N,E) channels. We obtain some high correlation values (0.8 for the P-wave of the shallow earthquake on station MTFO), with occasional near-zero correlations (0.0–0.1 for three of 13 channels used for the shallow event, and for three channels of the 21 used for the subcrustal event).

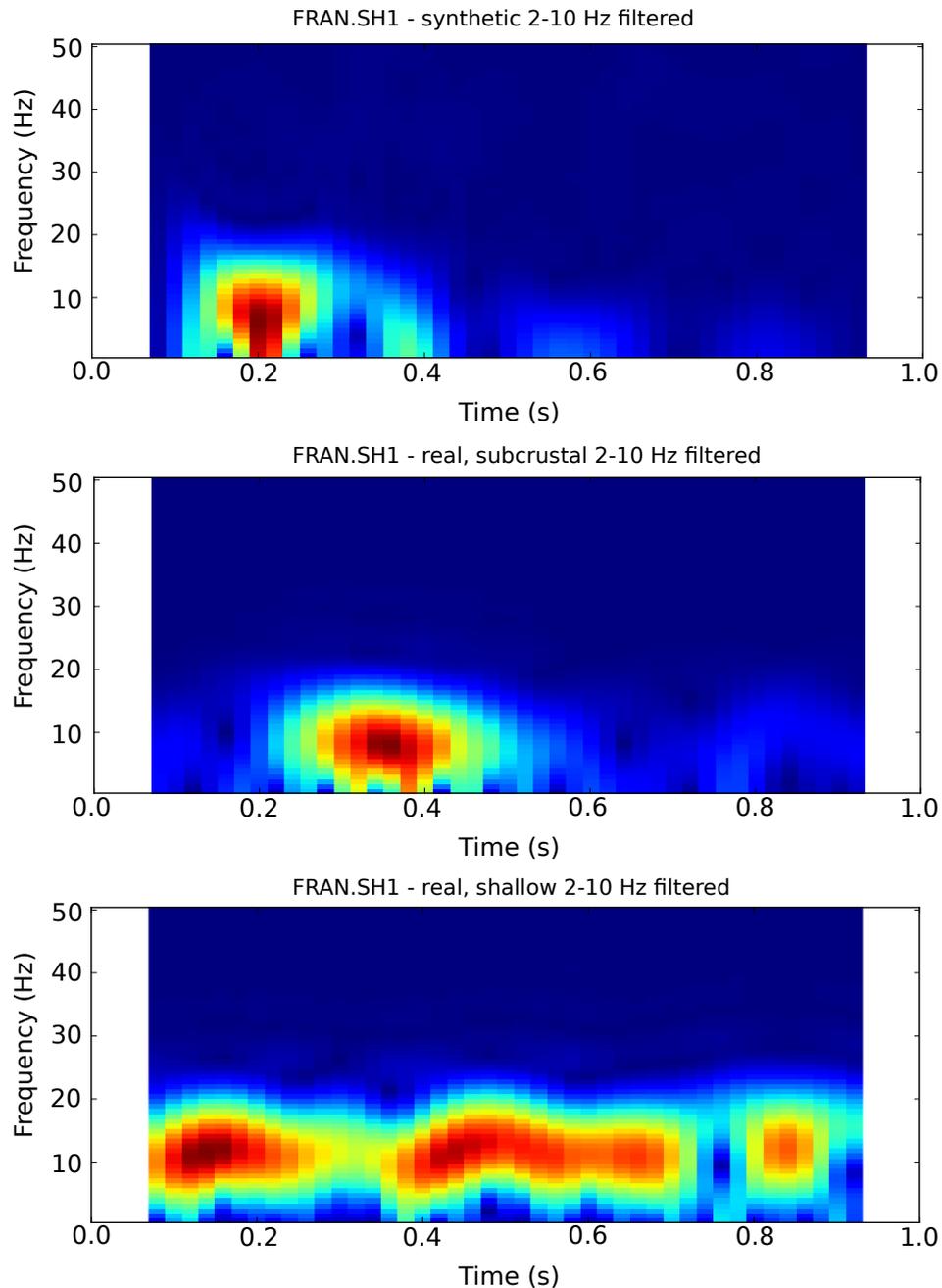


Figure 4.4: Spectra for synthetic (top) and two real earthquakes described in the text. Seismograms are filtered between 2–10 Hz before their spectra are calculated. All plots show the vertical channel of station FRAN. The poor alignment between energy peaks is likely indicative of pick uncertainties as all seismograms are extracted with a 0.1 s pad prior to the P-phase (note, these plots are made prior to alignment to illustrate this effect). The simplicity of the subcrustal earthquake spectrogram highlights the ease of correlation. The shallow earthquake has a more complicated spectrogram due to multiple phases being present in the window, which are not modelled in the synthetic waveform.

To determine whether we can detect near-repeating earthquakes, we attempt to detect a background (un-triggered) swarm. Of the background swarms catalogued by Boese et al. [2014], we focus on the only swarm that occurred before the  $M_W$ 7.8 Dusky Sound earthquake. This swarm occurred on 24 May 2009, and comprised 19 events. We use the same approach that we applied to the detection of discrete earthquakes, whereby a synthetic is generated that best matched one earthquake within the swarm and used this synthetic and the real earthquake as templates for that day.

Of the 19 earthquakes documented by Boese et al. [2014], the real template detected 13. Two of the missed events had smaller inter-event times than the 6 s minimum we allowed (Figure 4.6). The real template also detected a further nine previously uncatalogued events. Under inspection, eight of these additional events are unambiguously earthquakes, while the final event has a clear S-wave, but the P-phase is not visible.

The synthetic template detected ten of the 19 earthquakes previously catalogued, and as with the real template, two earthquakes were missed due to the 6 s minimum inter-event time. The synthetic template made two additional detections compared to the real template which both had negative cross-correlation sums. These events were not associated with catalogued events. One of these additional detections is a real earthquake with low signal-to-noise ratio that had not been detected previously. The final additional detection is not associated with any visible events and we classify it as a false detection.

We then generate a grid of templates representing a finite source volume to further test our ability to detect the swarm using synthetic templates without *a-priori* knowledge of the source. Using a coarse grid of 210 synthetic template sources and inter-node spacing of 2 km we obtain no detections at  $8 \times \text{MAD}$ . In light of the previous test this suggests that our grid spacing is too coarse for this swarm, with no node sources matching the swarm source location. Lowering the detection threshold results in an increase in false detections with little to no increase in real detections indicating that this issue cannot be solved by changing the threshold.

The last result indicates that synthetic templates in the near surface are source-specific, which is likely due to the impulsive nature of the waveforms from shallow earthquakes in the central Southern Alps. This is in contrast to the more emergent LFE waveforms found in the central Southern Alps. We have attempted using only P-phases, only S-phases and both body phases together with the same result. Of importance here is the lack of false detections: if we have the correct template source location we can detect these events, and if we have the wrong source location we neither detect the events, nor make any false detections.

We also compare our synthetic templates to previously identified LFE templates. We compare with the raw, un-stacked templates rather than those generated by stacking

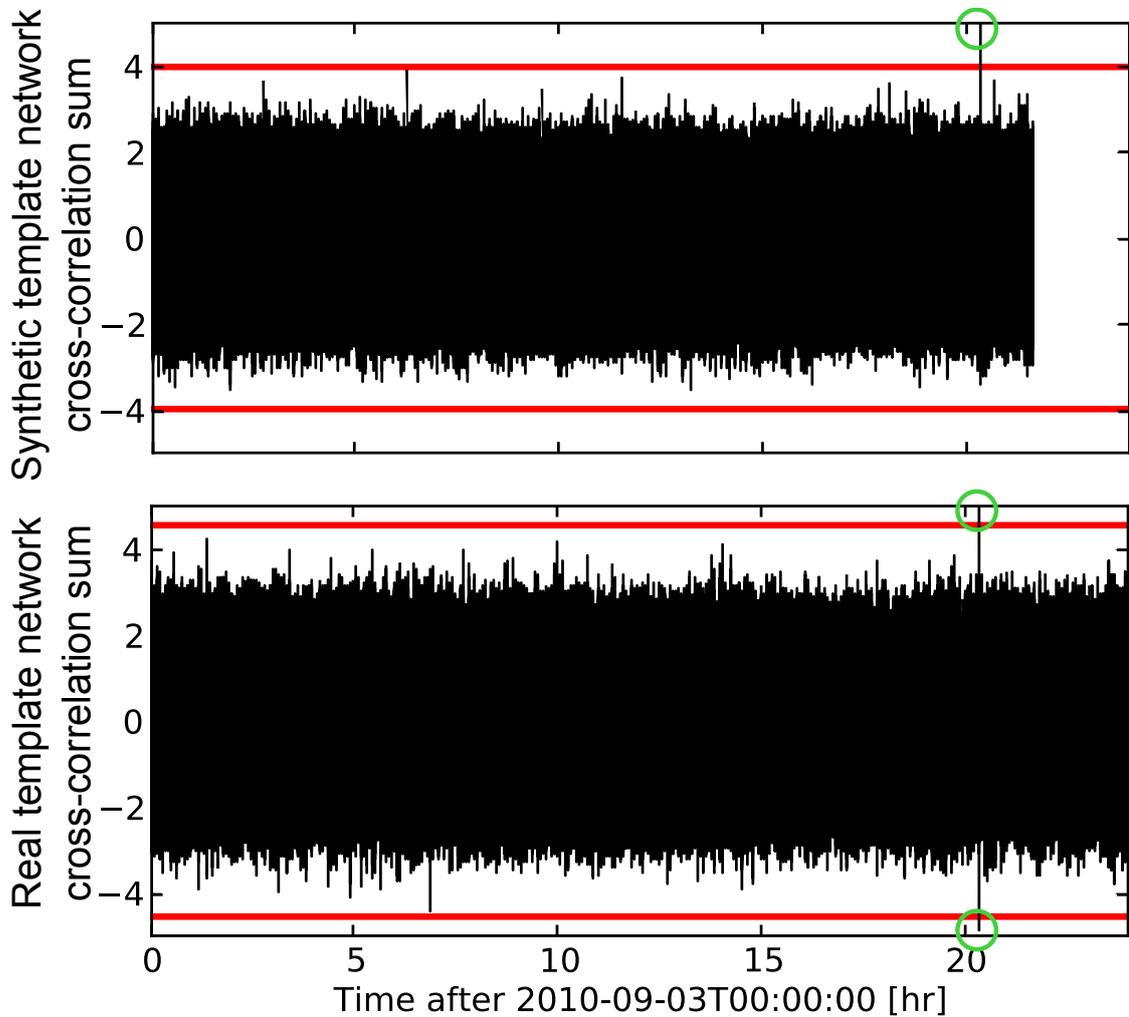


Figure 4.5: Day-long network cross-correlation sums for the day of the shallow event tested, for both the synthetic template (upper) and real template (lower). Note, we do not show the subcrustal earthquake as this was on a different day. Red bars mark both the positive and negative  $8\times\text{MAD}$  threshold levels. Threshold exceedances are highlighted by green circles. Only one sample in both time-series' exceeds the positive threshold, while only the real template has a negative peak greater than the threshold, which likely corresponds to the template one period out of phase (cycle-skipping). Note: maximum network cross-correlation sums are clipped to better show peaks near the threshold.

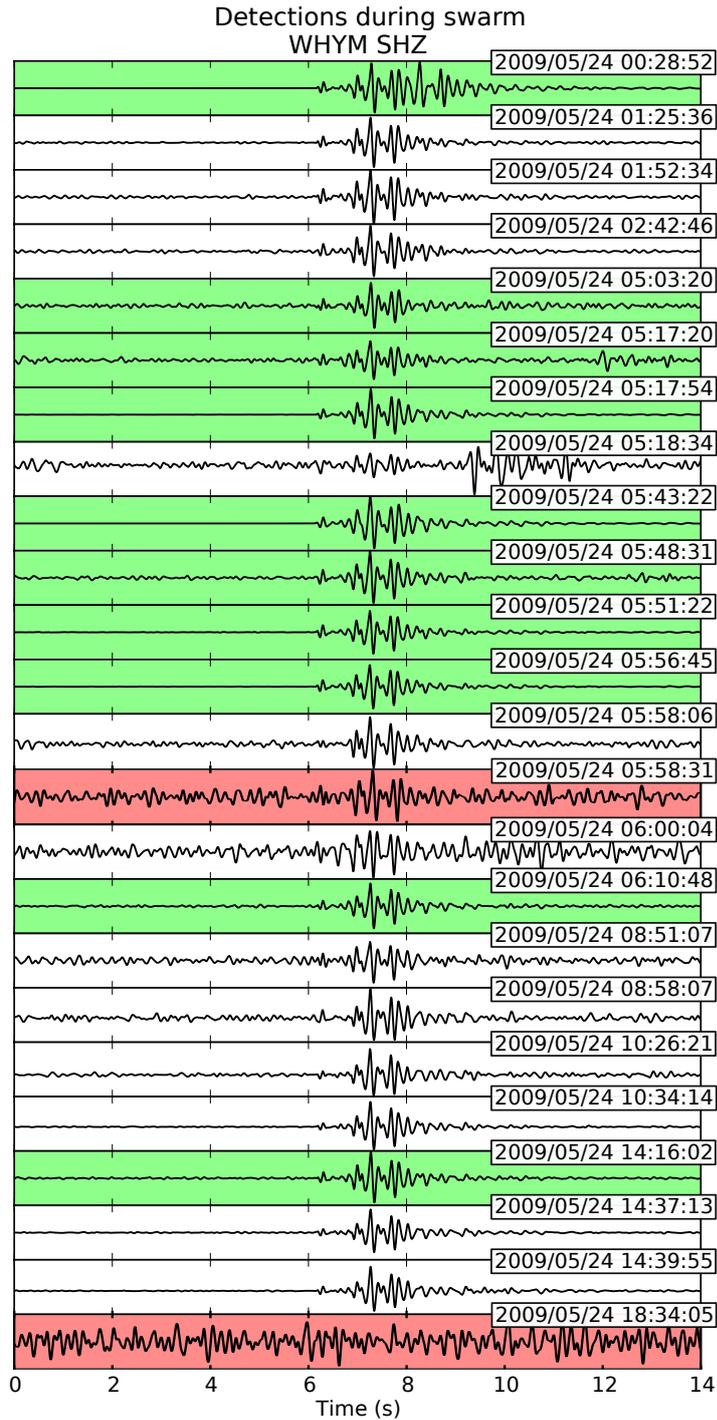


Figure 4.6: Detections made during one day of swarm activity (24 May 2009) using synthetic and real templates. Red backgrounds to traces indicate detections made only by the synthetic, which also correspond to negative cross-correlation sums. The final event appears to be a false detection. Green backgrounds indicate shared detections between the real and synthetic templates. Detections with white backgrounds were only detected by the real template. The synthetic template detects fewer events and makes one false detection for this day. Events are ordered by time, with their detection times given in the top right of every trace. Amplitudes are not equal between traces. Note that the red-shaded event at 05:58:31 is obviously an earthquake, but shifted half a cycle out of phase, resulting in the negative network cross-correlation sum.

multiple detections as this is more representative of the initial detection phase applied to the continuous data. Our original templates were constructed using S-phases only as this is the most dominant energy in the tremor waveforms [Chamberlain et al., 2014]. We obtain single channel cross-correlation values of between 0.3 and 0.6. The cross-correlation sum of the LFE event plotted (Figure 4.7) at 3.19 would be close to the threshold for the swarm period ( $8 \times \text{MAD} = 3.54$ ) However, this threshold is significantly higher than the threshold expected for the LFE due to the use of more channels: we used ten channels for the synthetic swarm detections, whereas the LFE template has only seven channels, we might therefore expect the MAD level to be lower, 70% of the swarm threshold is 2.5; at this level the LFE would trigger a detection.

Comparison of the spectrograms of the real and synthetic events shows clear differences (Figure 4.8), although the peaks in energy occur in similar spectral bounds. The synthetic template shows the interference in the coda of the S-phase with multiple low-frequency energy peaks following the main arrival. This is due to the alternating spikes used in the synthetic template generation procedure. These multiple phases are also observed in the real LFE spectrogram, however it is unclear whether these are subsequent LFEs in the tremor wavelet or multiple phases.

### 4.3.3 A grid of templates

To detect earthquakes with different source locations we require a suite of templates. To achieve this either all possible moveouts and all possible S–P times can be modelled, or templates can be generated for a set of known sources (e.g. within a region of interest). By specifying a grid the computational load can be reduced (by undersampling the grid-space and selecting specific boundaries) while simultaneously ignoring detections for events we are not interested in (e.g. teleseisms, or regional earthquakes). In the case of LFEs in the central Southern Alps, the source region is relatively well defined as being anti-correlated in depth with shallow seismicity [Wech et al., 2012, Chamberlain et al., 2014]. This grid selection is similar to that used by Frank and Shapiro [2014] for work in Mexico. Alongside reducing computational load, using a set of pre-defined sources has the potential to constrain the location in tandem with detection.

Here we use a grid extending between depths of 15–40 km with corners as shown in Figure 4.9 and a coarse node spacing of 2 km. Our initial grid has been chosen to be deeper than regular seismicity, while epicentrally encompassing some standard seismicity, to test whether we can distinguish between LFEs and regular seismicity. It has also been set to encompass the LFE and tremor source region. For each node in the grid we calculate travel-times to all stations using a 3D velocity model [Eberhart-Phillips et al., 2008], and the ray-tracing algorithm in *NonLinLoc* [Lomax et al., 2000].

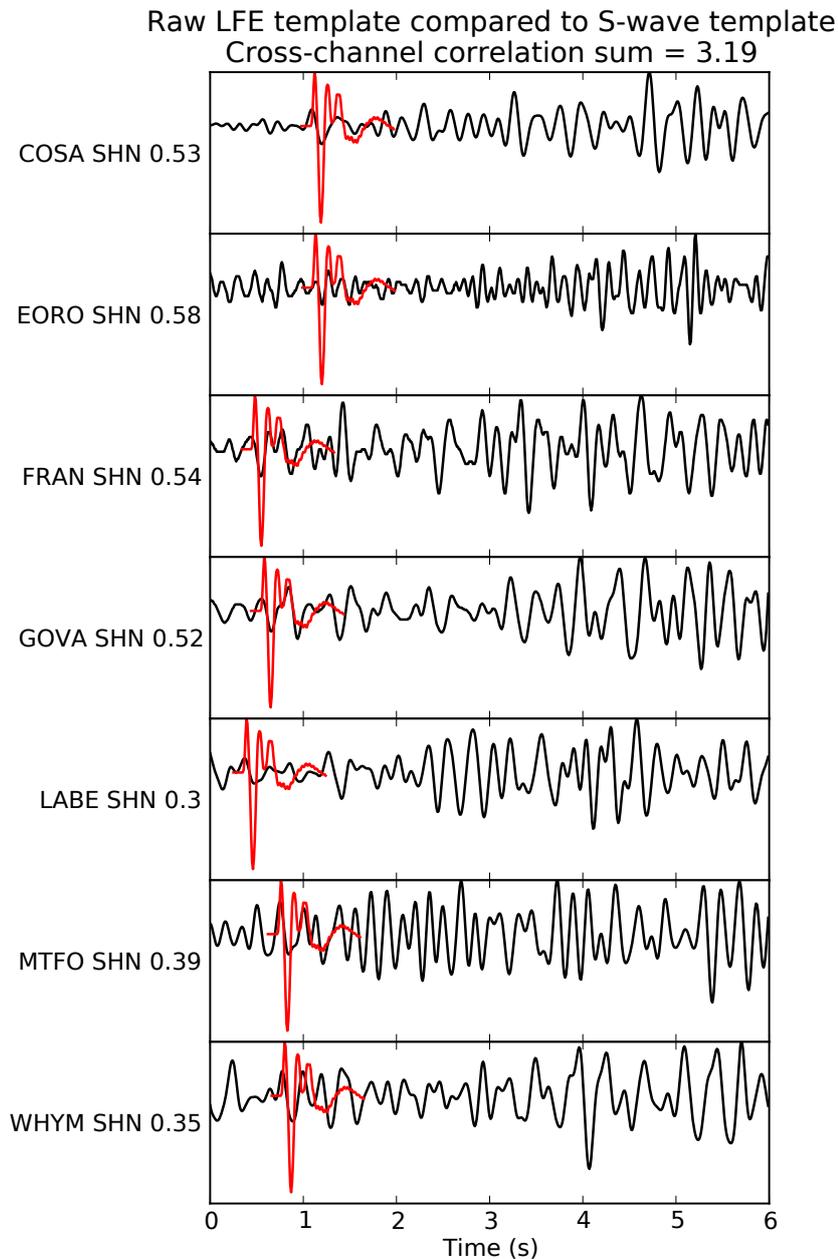


Figure 4.7: Synthetic S-phase template (red) overlaid on the real, un-stacked LFE template (37575, Table 3.1) at the maximum, positive cross-correlation within the first two seconds. We designed the initial (real) templates to begin approximately 0.5 s before the preliminary S-pick and thus only look for the maximum cross-correlation early in the trace. We can obtain higher single-channel cross-correlation values by scanning the entire six second window, but this would likely correspond to later arrivals of other LFEs. The cross-correlation sum is given in the title. We only show the north horizontal channel here as this is what we use for the large-scale cross-correlation detection routine. Individual single-channel cross-correlation values are given after the station and channel name label.

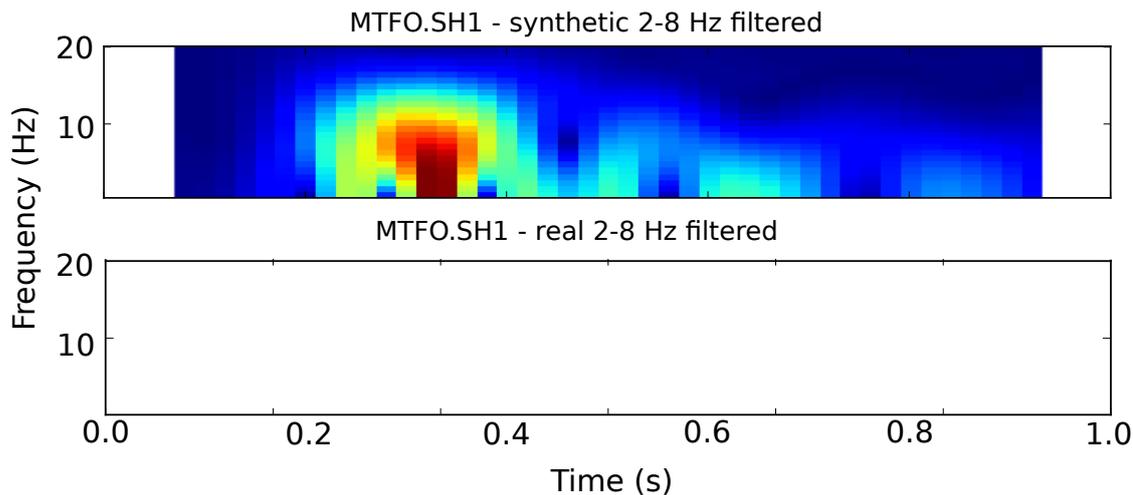


Figure 4.8: Comparison between synthetic (top) and real (bottom) S-phases for LFE templates. Seismograms are filtered 2–8 Hz prior to calculation of spectrograms. The multiple energy peaks are roughly met by the synthetic.

When generating templates for continuous detection of LFEs, we use only the S-phases of our synthetic templates. We do this in an attempt to cope with unknown focal mechanisms; reversed P-wave polarities at some sites due to different positions on the focal sphere would result in negative cross-correlations for P-wave arrivals if we did not model this. As we do not know the focal mechanisms for our LFEs yet, we cannot account for this without modelling many other sources. Despite S-phase polarities being dependent on take-off angle and focal mechanism, this does not detract overly from our detection ability, possibly due to inherent cycle-skipping that can occur when using a grid of only single phase templates. All the S-phase templates generated have positive first-motions and are generated and correlated with the north horizontal component only.

As eluded to above, using only S-phase templates reduces our location accuracy significantly. For example, our S-phase template may correlate well with the first arrival of the LFE S-phase at one site, but correlate better one or more periods later at another site. This results in a detection for an incorrectly located template event.

#### 4.3.4 Large-scale multi-parallel cross-correlation optimisation

Our problem of 6.5 years of data (<2 terabytes) is small compared to the scale of so-called ‘Big Data’ datasets currently being analysed in computer science [Snijders et al., 2013]. Nevertheless, earthquake detection can be thought of as a data-mining problem (methods encompassing pattern extraction from large (‘Big Data’) datasets) whereby we search for specific features within the dataset. We can therefore make use of some of the algorithms and computational architectures designed for this type of problem: for example Aguiar and Beroza [2014] applied the Google PageRank algorithm to the

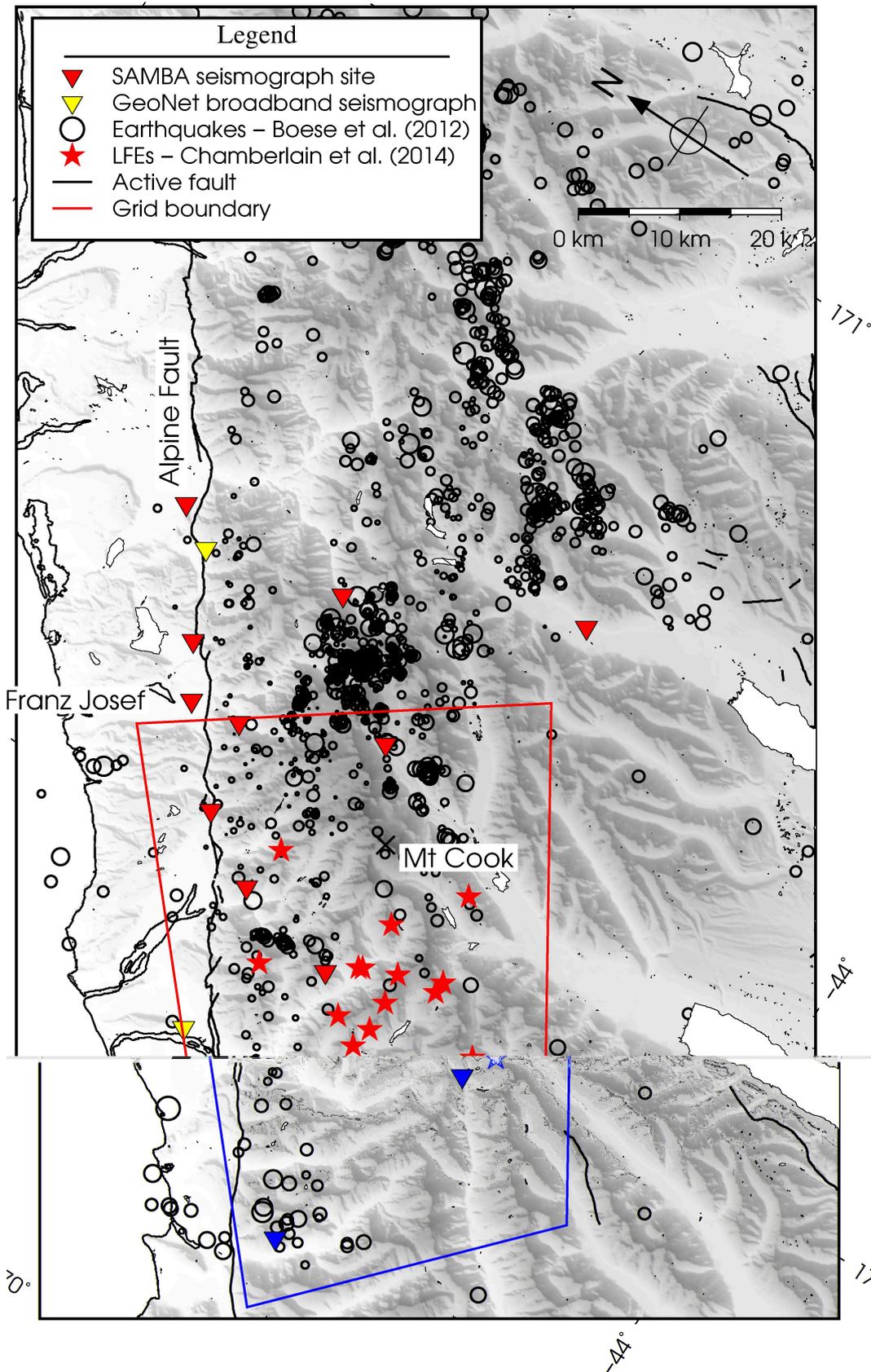


Figure 4.9: Target LFE source region in the central Southern Alps, with grid outline (red) used in this study. Seismicity is from Boese et al. [2012]. We select our grid to be deeper than most of the documented microseismicity in the region. Note that map is rotated anticlockwise from north.

problem of earthquake detection. Here we make use of the possibility to deploy the matched-filter routine in a multi-parallel architecture. The matched-filter routine we use is written in Python as part of the open-source EQcorrscan package<sup>1</sup>. This section outlines how this routine is optimised for large-scale, multi-parallel earthquake detection.

For a single day of seismic data and for each template, every correlation at every time-step can be computed in parallel. EQcorrscan uses a normalised cross-correlation routine from the openCV<sup>2</sup> image processing library, which handles multi-threading of correlations at a CPU (Central Processing Unit) level. At the level above this, daily cross-correlations for individual templates are computed in parallel, which are paralleled using multiple CPUs accessing seismic data in shared memory. At our top-level of parallel processing, each day is computed separately. To exploit this, we separate the 6.5 year dataset into chunks of days and compute the detections for this sub-set of days on individual computers or nodes in a High Performance Computing cluster architecture.

Given enough compute resources, we could in principle generate detections for as many templates as we have CPUs in the same time that it takes to compute the detections for a single template in a non-parallel architecture. We can simultaneously compute detections for as many days as we have nodes: this is simple to achieve in EQcorrscan. In practice, however, compute resources are often limited. In this study we make use of the NeSI<sup>3</sup> infrastructure and the PAN cluster computer. This cluster has 393 nodes, each node having between 12 and 40 CPUs. Some nodes have access to GPUs (Graphics Processing Units) which could, in the future, be used for this problem. Using a normal availability of resources (20–25 nodes), we are able to complete correlation of more than 600 templates with the full 6.5 year dataset in less than 48 hours clock-time.

The main limitation in this methodology is memory consumption, and steps have been taken to reduce the size of arrays stored in memory. In particular, data are down-sampled to 20 Hz. This resampling reduces computational load, while still capturing the information in the 2–8 Hz tremor band we are interested in. Seismic data are forced to be of integer type; for the Stein2 compressed miniseed data used in this study which is stored as 32-bit integers this has no effect.

Filtering the input integer data yields floating-point outputs. These are then amplified to the limits of 32-bit integer expression before being converted back to integers. Data are amplified to reduce the loss of fidelity when converting from floating point numbers to integers. This step also normalises across the channels. When we compute day-long

---

<sup>1</sup><https://github.com/calum-chamberlain/EQcorrscan>, version 0.0.9 as used in this work, is included in the electronic appendices

<sup>2</sup><http://opencv.org/>

<sup>3</sup>New Zealand eScience Infrastructure: <https://www.nesi.org.nz/>

cross-correlation vectors, we convert these from 32-bit floating-point numbers to 16-bit floating-point numbers. As 16-bit floating-point numbers are not a native C-type this conversion has some compute overhead, but effectively reduces memory consumption by half. Using 16-bit floating-point numbers also has implications for detections near the threshold due to rounding errors.

The 128 GB memory limit of PAN nodes means that we are limited to approximately 100 templates running in parallel on a single node. This is because we compute detections for each day only after we have completed the parallel run through all templates. That is, we compute cross-correlations for templates in parallel for each channel individually, and sum each channels correlation array; the result of this is that we store multiple single-channel cross-correlation vectors in memory. EQcorrscan is written to allow these arrays to be stored on disk, which is especially useful for low-memory computers, however the extra read and write time is significantly more than the time penalty for limiting the number of simultaneously processed templates to 100.

## 4.4 Results

In our initial run using 638 synthetic templates we made 347,111 detections at  $8\times\text{MAD}$ . Of the 638 templates, 38 did not detect any events during the 6.5 year period, leaving us with 600 families (groups of detections associated with a single template) consisting of between 388 and 862 individual detections. We remove detections associated with data gaps in the GeoNet national data (where day-long files are partially filled), which reduces the number of detections to 331,489. We make considerably more detections in September and October of 2011 than at any other point in the 6.5 year timespan. This corresponds to a reduction in the number of functioning stations. If we increase the minimum number of stations used for detection we remove this period of increased detection while maintaining other trends. We see a similar but lesser, effect by increasing the detection threshold to  $10\times\text{MAD}$ . To remove this artefact we only use detections with at least nine stations. This leaves us with 233,853 detections in 600 families. For each of these detections we extract a short (90 second) waveform for all channels and stations available at the time of detection from the continuous data. We group these waveforms into families by template and stack them. All stacked waveforms are included in the electronic appendices.

Initial analysis of two days of detections reveals several types of detected event, rather than just tremor and LFEs. Our use of S-phase templates only results in a significant number of false detections (usually detections of unwanted earthquakes), including local microseismic events and the surface waves of large earthquakes. The duration of

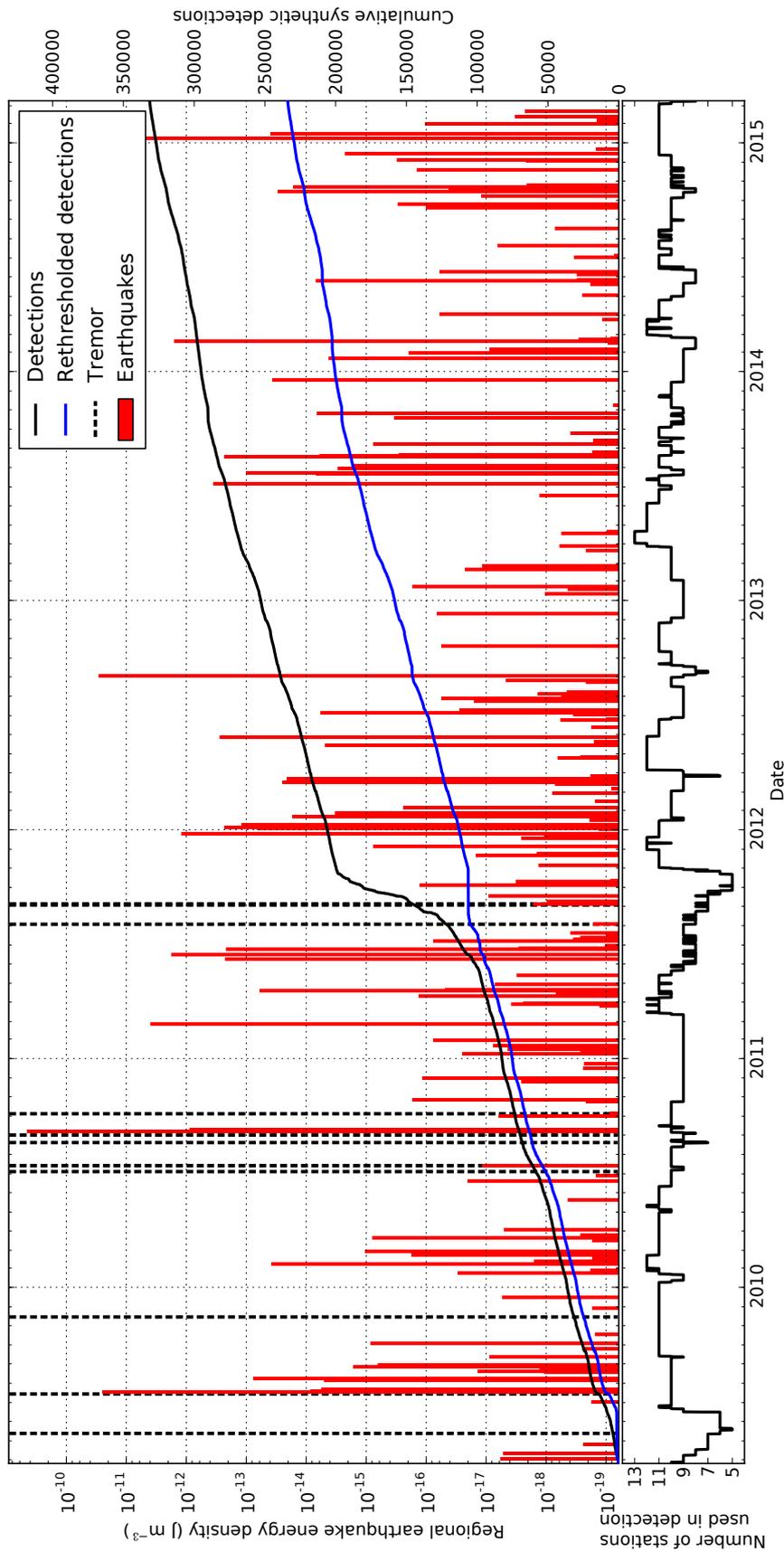


Figure 4.10: Cumulative detections from synthetic templates at  $8 \times \text{MAD}$ , with the raw detections ('Detections', black), and those with at-least nine stations ('Rethresholded detections', blue) plotted. Plotted alongside the cumulative detections are tremor events identified by Wech et al. [2012] as black dashed lines, and the energy density calculated for all earthquakes in the GeoNet catalogue in the previously determined LFE source region (red). The lower plot shows the running average of the number of stations used for detection.

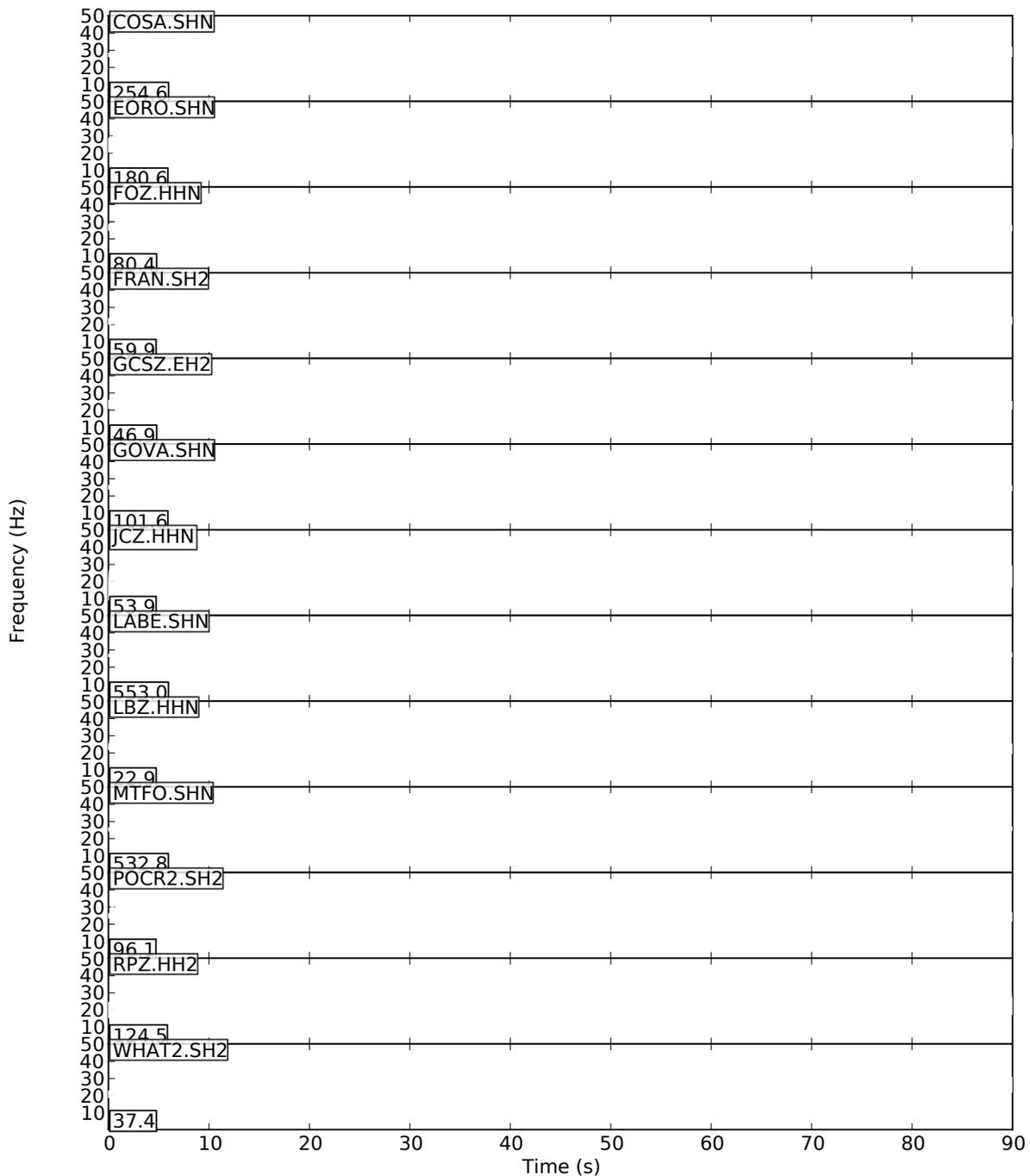


Figure 4.11: Example waveforms and spectrograms at the time of a single LFE detection. Waveforms are resampled to 100 Hz (from either 200, or left at 100 Hz), and filtered between 2 and 49 Hz to remove long-period noise on the broad-band sites, while retaining signals outside the tremor band to demonstrate the consistent dominant power in the 2–8 Hz band. There are clear spectral peaks between 2 (lowcut bound) and 10 Hz. The waveform is clearly emergent, and does not correlate with any earthquakes either globally above magnitude 5.0 (USGS catalogue), or within New Zealand above magnitude 2.0 (GeoNet catalogue). This waveform is interpreted to represent a short burst of tremor.

our dataset allows us to mitigate against the effect of false detections during stacking, providing we are detecting mostly real events, which is confirmed by visual inspection of a subset of days. We find that we are able to detect events within tremor (Figure 4.11), and that we have detection rate increases coincident with known tremor events. Therefore, when detections are stacked, positive detections stack coherently and false detections do not. This coherency is exemplified in the stacks which have clear S-phases and occasional P-phases (Figure 4.12), which allow for phase picks to be made and hypocentre estimates generated. These distinct phases are the topic of on-going research to generate focal mechanisms.

Comparison of the stacked waveforms generated by different templates shows low similarity between stacks, suggesting that each stack represents a distinct family of events. When computing the similarity between stacks we extract a short 20 s waveform from the 90 s stack beginning 5 s before the detection time (which should be the first S-arrival). This 5 s pre-pad is designed to include any P-arrivals present above the noise. The 15 s after the first S-arrival is used to include later arrivals on other channels. We do not apply a moveout to different channels to allow comparison of channels from different stacks (families) and therefore moveouts. We compute cross-correlations between all stacks and generate a channel average cross-correlation value. No stack pairs have average single-channel correlations above 0.4.

We automatically pick all 600 stacks using a simple STA/LTA routine, tuned for picking the emergent arrivals in the waveforms. We independently pick and compare 75 events of the 600, and find that the automatic picker makes slightly fewer picks than our analyst picks, but the picks it does make are usually good (within analyst pick uncertainty, which is often high for these emergent waveforms). Most stacks (579) have only S-picks due to the higher signal-to-noise ratio of this phase, which leads to large depth uncertainties in locations. This uncertainty in depth is demonstrated when P-picks are added (which are available for only 21 stacks), which reduces the depth estimates from 40–60 km to 14–30 km (Figure 4.13). We note that if we locate the 21 events for which P and S-phase picks are available using only the S-phase information, then only seven stacks locate shallow compared to the deeper events. We think this indicates the importance of P and S-phase information, especially in this case with an uncertain velocity model. We therefore only present the locations for families which have both P and S-phase picks in Figure 4.13, all other families locate in a broad cloud at depth due to their high uncertainty.

Template locations are very uncertain (more than their Gaussian estimates of  $\pm 5$  km suggest) and require further iterative detection and stacking to enhance signal-to-noise ratios to ensure that spurious noise phases are not incorrectly picked within the stack, and to enhance weak P-phases. These locations should not be treated as LFE fam-

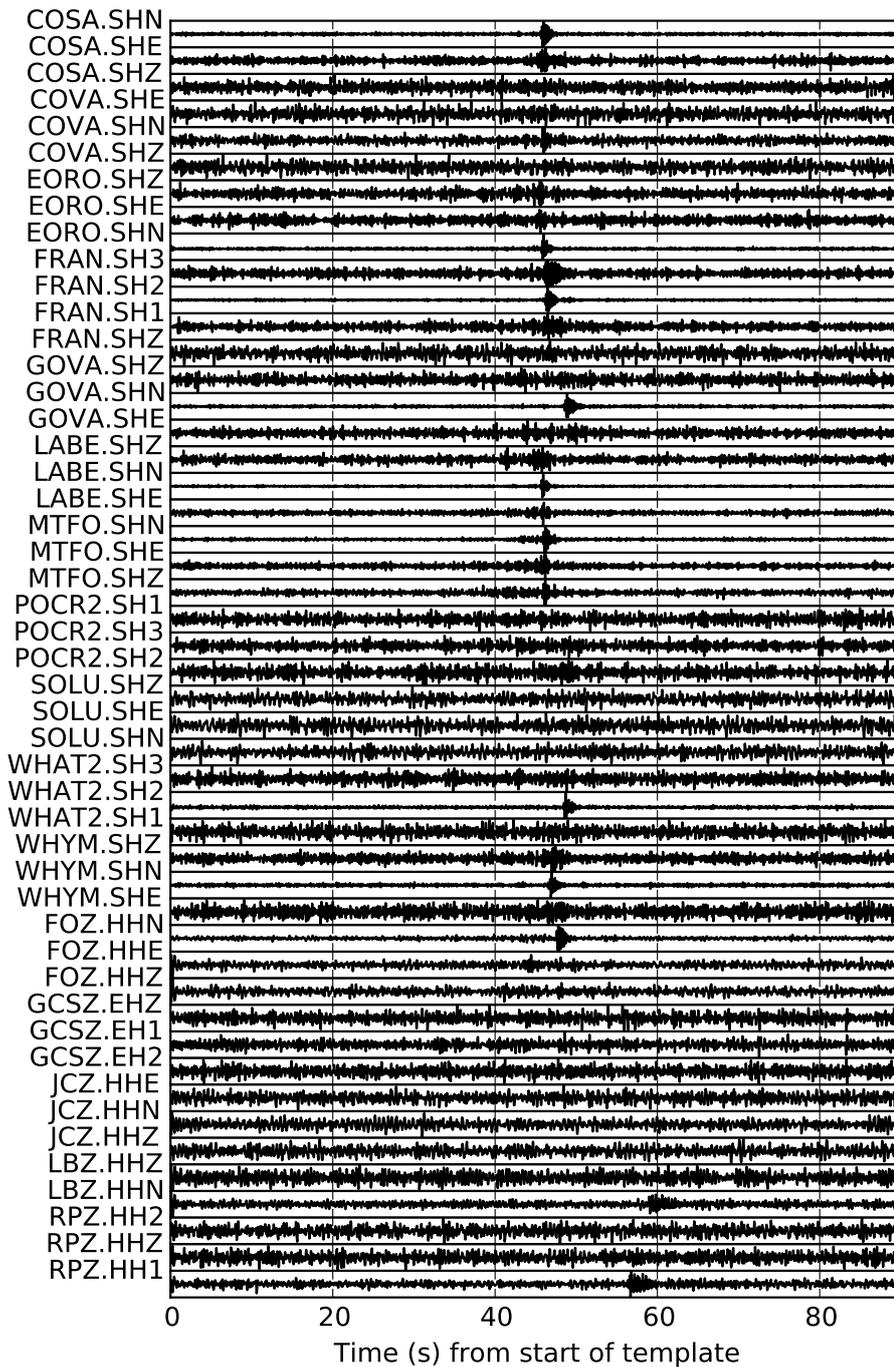


Figure 4.12: Initial linear stack of 549 detections from node 170.149, -43.5293, 18 km. The synthetic template for this node made the most detections of all synthetics. For many stations there are clear S arrivals, with some also showing P-phases enabling hypocentres to be determined. For sites not operational throughout the whole period (notably SOLU and GCSZ), fewer detections have been stacked resulting in poorer-quality waveforms. Waveforms are downsampled to 20 Hz and bandpassed 2–8 Hz prior to stacking.

ily locations, and rather are a validation that our detected events do not co-locate with regular seismicity, and that we are able to achieve sufficient signal-to-noise ratios through stacking to develop seismic phase arrivals.

## 4.5 Discussion and conclusions

We have presented a new method for detecting earthquakes using simple synthetic templates in a matched-filter routine, and have proven the method's effectiveness for detecting both shallow and deep earthquakes in New Zealand's central Southern Alps. We have also demonstrated that this method can detect low-frequency earthquakes and used it to develop a spatially and temporally continuous catalogue of 600 LFE templates spanning 6.5 years beneath New Zealand's central Southern Alps.

The synthetic template detection method presented here works well for detecting LFEs within New Zealand's central Southern Alps, however, detection rates for the synthetic templates are notably lower than for a stack of real waveforms: we obtain between 251 and 549 detections per template in a 6.5 year dataset, whereas the previous study (Chapter 3) found between 223 and 1847 detections per template in a 3 year period. Of these detections, the synthetic templates make a significant number of false detections (due to only using the synthetic S-phase, which also correlates well with shallow seismicity and surface waves). However, stacking the detections produces useful waveforms for further matched-filter detection by coherently stacking real detections to produce higher quality, real waveforms. As such this method should be used to complement standard matched-filter methods, and may be thought of as a template construction methodology.

Compared to other template construction/detection methodologies, this method works better for the central Southern Alps dataset due to a number of key features. Firstly, to exploit the duration of the dataset, a method needs to be computationally efficient. This fact alone excludes 'many-to-many' type methods such as autocorrelation [Brown et al., 2008]. For the example of autocorrelation using all the data available rather than only analysing periods of known tremor, we would effectively be cross-correlating approximately  $4 \times 10^9$  templates, rather than 600. This increase in correlations would yield a  $6 \times 10^6$  increase in computational time, given the same resources. For the case presented here, with 638 templates and using 20–25, nodes with 16 CPUs each, takes less than 48 hours. For the autocorrelation method we predict that we would require 37,000 years. This would however be a more exhaustive search for possible LFE templates, in contrast to our bounded method.

The second feature of the central Southern Alps seismic dataset that precludes the use

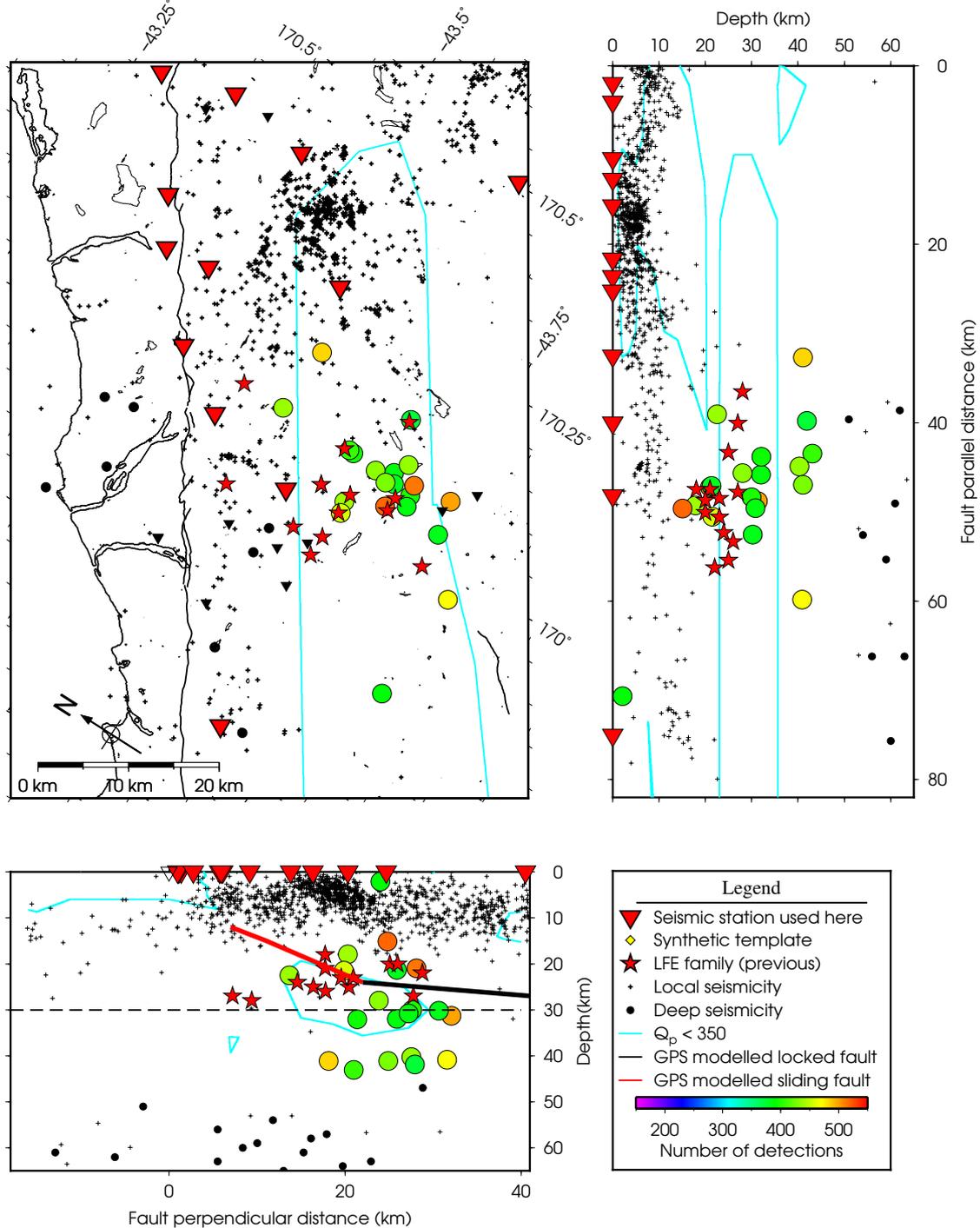


Figure 4.13: Initial locations of LFE templates computed using *hypocentre* and the same 1D velocity model used by Chamberlain et al. [2014]. Most events are poorly constrained by only S-picks (marked by diamonds), however, those with P-picks (marked as circles) tend to locate shallower than those without. LFE stack locations are coloured by number of detections used to build the stack for that template. Note that locations should not be considered final LFE family locations, but indicate the importance of using both P and S-phases for location in this uncertain velocity structure. Out of the 21 stacks for which P and S-phase picks are available, if we locate these events using only the S-picks, 7 of them locate shallower than the large cloud of stacks for which only S-phase picks are available.

of other methods, such as the cross-station [Rubin and Armbruster, 2013, Thomas and Bostock, 2015] and the brightness methods [Frank and Shapiro, 2014], is the low signal-to-noise ratio, and complex signal characteristics. The low signal-to-noise ratios result in noise stacking to greater amplitudes than signal in the brightness method. Frank and Shapiro [2014] suggest the use of an inter-station coherence threshold to remove false detections generated by noise. However, the complex LFE signals observed in the central Southern Alps result in very low inter-station coherences, often lower than the inter-station coherence of noise signals. This lack of inter-station coherence also precludes the use of the cross-station method, which looks for sources that provide the highest inter-station coherency between three sites. In testing we have found that the noise signal has higher inter-station coherence than LFE signals. We think this low inter-station coherency is due to the short duration of tremor in the central Southern Alps, the strong noise field [Petersen et al., 2011], and the complex scattering apparent due to highly foliated rocks.

The templates generated by this method currently have poor depth constraints due to a lack of P-phases in most waveforms. The S-phase only synthetic templates also make many false detections, which precludes an analysis of events detected by the synthetic templates. The stacks generated using this method should therefore be used in further, possibly iterative, matched-filter and stacking routines until the number of detections, and template, stabilises. This should reduce the number of false detections, which would enhance the stack, possibly allowing for more P-picks to be made: as we have demonstrated, with the inclusion of P-picks, depths are more well constrained (Figure 4.13).

Currently, rather than the phase picks presented here being able to give robust locations, they should be used to generate further templates around the phase picks for a second stage of matched-filter detection. The detections from this run could be stacked to enhance signal-to-noise ratio which would allow for more robust phase picks, including P-picks on some channels for more templates. This would provide more robust locations, as well as reducing the number of false detections to provide a detailed, continuous catalogue of LFEs.

### 4.5.1 Methodological extensions

By extending this method with subspace detection methods, we may be able to use both P and S phases for a range of unknown focal mechanisms. However, it is not possible to generate a single subspace template that captures all possible P amplitudes due to the reversed polarity at focal-planes. Detecting a range of focal mechanisms and separating these into discrete families remains an ongoing problem in matched-filter

methodologies. Currently low signal-to-noise ratios preclude a standard clustering of detections by cross-correlation: for individual detections the noise dominates inter-detection cross-correlations.

When including P-phases initial detection location accuracy is increased, resulting in a need for more densely sampled template grids. This is simply limited by computation time: increasing the density of the grid from 2 km to 0.5 km increases the number of templates, and hence the computational duration, by a factor of 16. This is still a computationally viable option on cluster computers. Including P-phases also reduces the number of false detections (of other unwanted earthquakes), which would yield a better stack, and a catalogue that could be analysed without a need for further stacking and detection routines. However, as eluded to above, we have yet to overcome the difficulties inherent in unknown focal-mechanisms when using P-phases. This may be overcome by exploration of the minimum number of stations required to give reliable detections, and generating P-wave (and S-wave) synthetic templates for these stations with negative and positive polarities. For example, rather than having just a positive P-wave at three stations, each station would have both a positive (compressional) and negative (dilatational) P-wave template. For three stations this would result in 8 templates for a single source location: the number of templates for each template source scales as  $2^n$  where  $n$  is the number of stations.

The method presented here is an efficient way of detecting unknown signals in continuous data, and could be well suited to other noisy environments such as drill-sites, geothermal areas or waste-water injection areas. In these settings a small volume of crust could be targeted with a high density grid of templates and a range of focal mechanisms; small earthquakes could then be detected in real-time, with each detection having a specific location attached to it according to the template used. This is discussed further in Section 7.2.2.

# 5 High-precision detection and relocation of highly correlated micro-earthquakes adjacent to the central Alpine Fault

## 5.1 Abstract

We detect microseismicity adjacent to New Zealand’s central Alpine Fault in a region of low seismicity, based on cross-correlation of pre-determined template events, for the period 2012/08/17–2014/12/31. A suite of 300 earthquakes and 195 explosions are detected, from which high-precision relocations for 157 earthquakes are generated. Relocations are generated by double-difference relocation using automatic, correlation derived picks, with relative pick precision of less than 0.01 s. We find clustered seismicity dominated by events near the site of the DFDP-1 borehole in Gaunt Creek. We extract a subset of 106 highly correlated events and calculate high-precision magnitudes by singular-value decomposition. These magnitudes range from  $-0.5 < M_L < 1.6$ ; other events are excluded from this calculation due to low signal-to-noise ratios likely due to very small magnitudes. We generate probabilistic focal mechanisms for 24 earthquakes in the catalogue using manual picks. These exhibit predominantly strike-slip faulting, with subordinate normal and reverse components. We interpret our catalogue to represent short-duration earthquake sequences on small asperities adjacent to the Alpine Fault, rather than repetitive failure on the through-going Alpine Fault.

## 5.2 Introduction

Deformation in New Zealand’s South Island is dominated by oblique convergence between the Australian and Pacific plates which is accommodated by the  $\sim 800$  km-long Alpine Fault. Between 5–10 mm/yr reverse motion has been estimated on the central Alpine Fault [Norris and Cooper, 2000, Sutherland et al., 2006, Barnes, 2009, Beavan et al., 2010] corresponding to up to 100% of interplate compression, and strike-slip motion of 23–27 mm/yr equivalent to  $\sim 75\%$  of interplate dextral motion [Norris and

Cooper, 2000]. The remaining deformation is thought to be taken up by distributed deformation on other structures such as the faults of the main divide fault zone [Cox and Findlay, 1995]. Despite these high deformation rates there is little documented on-fault seismicity on the central section of the Alpine Fault, with what on-fault seismicity there is being confined to low magnitudes (mostly  $M_L < 5$ ) [Boese et al., 2012, Leitner et al., 2001, Bourguignon et al., 2015, O’Keefe, 2008].

The Alpine Fault is thought to fail in major (M7–8) earthquakes with paleoseismic records indicating a likely recurrence timescale for the southern Alpine Fault of  $329 \pm 68$  years [Berryman et al., 2012]. The last Alpine Fault earthquake occurred before written records in New Zealand, but is dated at c.1717AD [Wells et al., 1999]. This gives a probability of a large ground-rupturing Alpine Fault earthquake in the next 50 years of  $\sim 27\%$  [Biasi et al., 2015]. Recent evidence from the central section of the Alpine Fault suggests that four of the five most recent Alpine Fault earthquakes ruptured large portions of the central section of the fault, with the 1717AD earthquake thought to rupture throughout the central section [Howarth et al., 2014]. Howarth et al. [2014] computed a recurrence interval of  $271 \pm 73$  years for the central section of the Alpine Fault, within uncertainty for the longer record presented by Berryman et al. [2012] for the southern Alpine Fault, but notably shorter. If this recurrence interval is correct, this would increase the probability of rupture in the next 50 years.

The region bounded by Franz Josef Glacier and the Wanganui river displays low rates of off-fault seismicity (15–30 km southeast from the surface trace) as documented by a number of authors [Boese et al., 2012, Leitner et al., 2001, Bourguignon et al., 2015, O’Keefe, 2008, Evison, 1971]. This region, hereafter termed the ‘seismic gap’, is bounded to the northeast by a zone of higher rates of moderate magnitude (M3–5) seismicity. On the southwestern edge of the gap, Boese et al. [2012] reported magnitudes on average one unit lower than at the northeastern edge.

This zone of low-seismicity is characterised by fewer mapped faults [Cox and Barrell, 2007] and low attenuation in the shallow crust ( $< 5$  km) [Eberhart-Phillips et al., 2008]. Boese et al. [2012] proposed this implies that the seismic gap region is a strong, unfractured block. Despite this paucity in off-fault seismicity, Boese et al. [2012] report no significant difference in seismicity within 10 km of the fault (near and on-fault). Rather, this region adjacent to the fault has low magnitude seismicity consistent with similarly near-fault areas both north-east and south-west along strike. Bourguignon et al. [2015] recently interpreted that the low seismicity zone was predominantly controlled by temperature, with low thermochronometric ages between the Hokitika and Karangarua rivers [Little et al., 2005].

Leitner et al. [2001] interpreted the region of low seismicity as being the initial transition of deformation from the Alpine Fault to the southern extent of the, more favourably

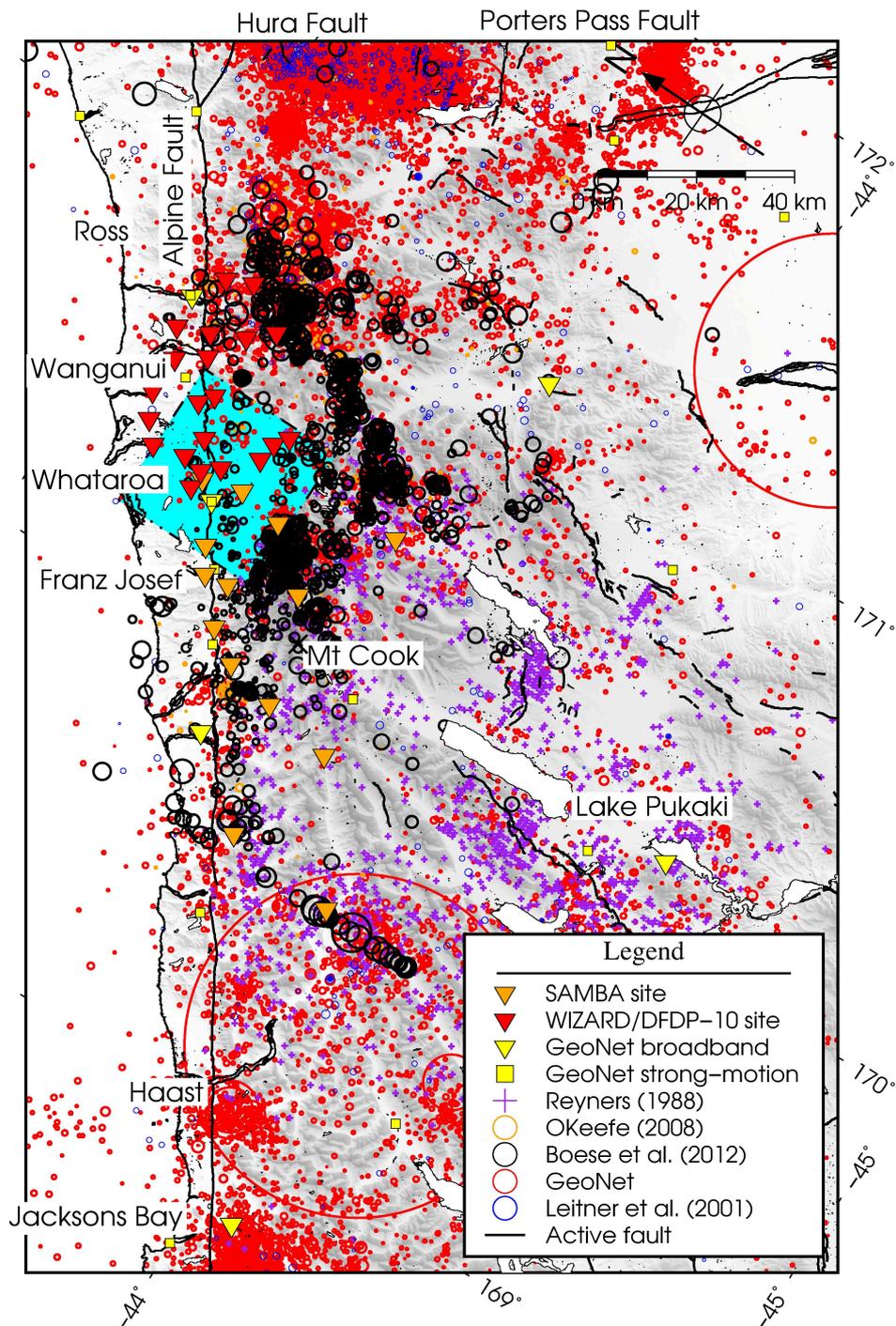


Figure 5.1: Background seismicity in the central Southern Alps as documented by several groups. Note that the catalogues do not span the same time-periods, nor do they have the same magnitude of completeness. The GeoNet catalogue has the longest duration (plotted between 29/07/1990–28/07/2015). The sites labelled were all active for at least part of the period analysed here, although only those sites with clear phases were used for template generation and subsequent detection. The cyan box shows the study region illustrated in later figures.

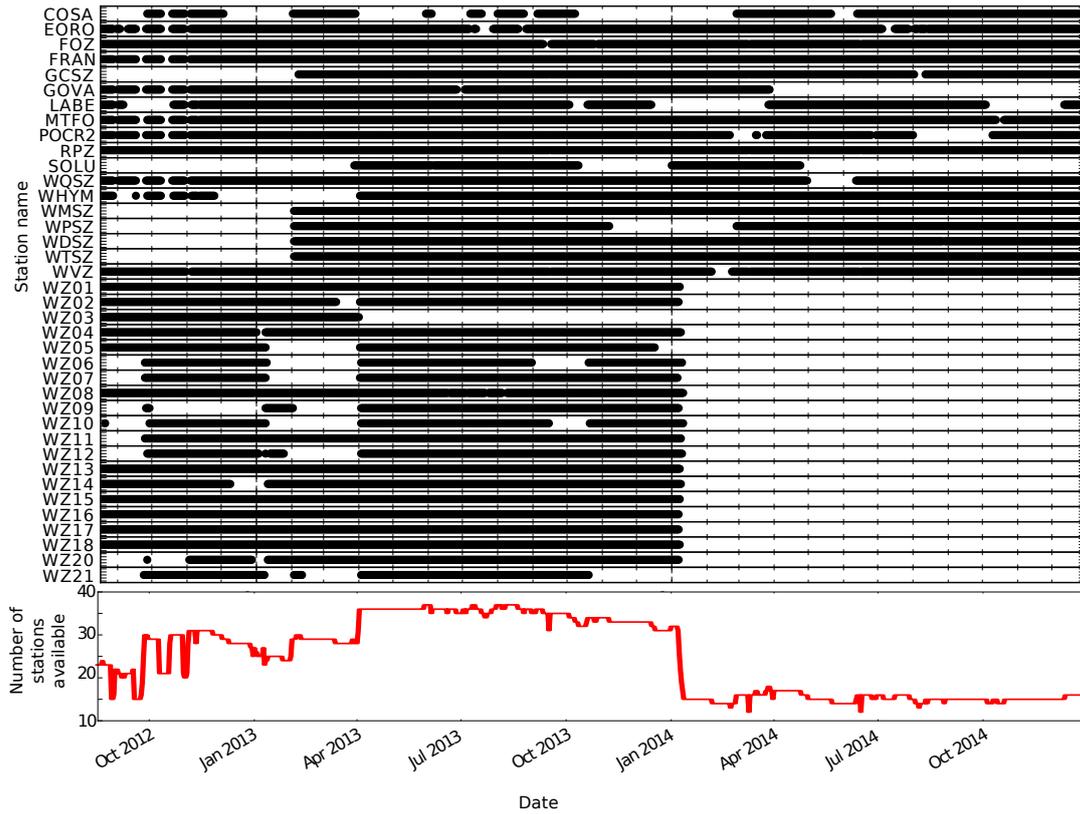


Figure 5.2: Data continuity of the stations used for matched-filter detection. In the upper panel site up-times are shown by black lines, the lower panel shows the overall number of stations available. Not all sites were used for all templates, as such single station outages can adversely affect different templates. Sites in the WIZARD network were removed in January 2014, resulting in a large reduction of possible stations for use in detection. Due to difficult high Alpine conditions some stations suffer data-loss due to lightning strikes (SOLU), extreme temperatures resulting in data corruption during writing (LABE, MTFO) and power failures (GOVA, COSA).

oriented (for strike-slip) faults of the Marlborough Fault Zone. The southwest extension of the Porters Pass fault marks the southwestern edge of the seismic gap [Leitner et al., 2001].

The Deep Fault Drilling Project (DFDP) addresses the *in-situ* conditions of the Alpine Fault using a range of techniques [Townend et al., 2009]. The first stage of the project, DFDP-1 [Sutherland et al., 2012] was undertaken in 2011 at Gaunt Creek (Figure 5.1). The results of this study provided detailed information on the geological [Toy et al., 2015], petrophysical [Townend et al., 2013], frictional [Boulton et al., 2014, Ikari et al., 2015b, Schleicher et al., 2015] and hydraulic [Sutherland et al., 2012, Carpenter et al., 2014] conditions in the near-surface (150 m depth). Notably this study showed that the Alpine Fault forms a hydrological barrier between the hanging and foot-walls of the Alpine Fault.

In 2014 a second phase of drilling, DFDP-2, was undertaken with the aim of intersecting the fault at a depth of 1–1.3 km, deep enough that stress conditions and pore-fluid

temperatures could be measured without the effect of topography. The Whataroa Valley was chosen to be the most suitable site along the central Alpine Fault for the DFDP-2 borehole. This decision was in part informed by the lack of seismicity in the region, with a well known background seismic pattern.

To enable a more precise understanding of the seismic conditions surrounding the borehole, and to monitor seismicity around the drill-site during drilling (Chapter 6) a small aperture array of four borehole seismometers was installed in February 2012 (Table 6.1) to supplement the SAMBA [Boese et al., 2012] and WIZARD [Thurber et al., 2012, Feenstra et al., In Review] networks in place at the time. Here we apply a matched-filter detection method using earthquakes detected between 2009 and late-2013 as templates to generate a more complete catalogue of seismicity in the region.

The application of cross-correlation methods to seismic data has proven to be one of the most effective methods of extracting known signals from continuous seismic data [e.g. Gibbons and Ringdal, 2006, Shelly et al., 2007, Schaff and Richards, 2011]. The extension of single-channel cross-correlation detections to detections based on multi-channel cross-correlation beams has allowed for many small signals to be uncovered from background seismic noise, with more variability than would be allowed by single-channel cross-correlation detections. Network based cross-correlation methods have recently been used to detect low-frequency earthquakes (LFEs) [Shelly et al., 2007], earthquakes buried in the coda of mainshocks [e.g. Peng et al., 2009], and within earthquake swarms [e.g. Shelly et al., 2013]. Using single-channel cross-correlation of respective template and detected events we compute high-precision relative pick times to quantify minor differences in source location between events within a family (the selection of events detected by a single template).

Notably cross-correlation methods have been used to study repeating earthquakes [e.g. Nadeau and McEvilly, 2004, Uchida et al., 2003]. Repeating earthquakes on faults are often modelled as asperities within otherwise creeping segments of faults [Zechar and Nadeau, 2012]. These asperities may be thought of as having different frictional properties to the surrounding fault, possibly they are more rate-weakening than the surrounds, which would result in more likelihood of frictional instability (e.g. earthquake nucleation). These asperities are usually constant in time, resulting in the repeating nature of these earthquakes.

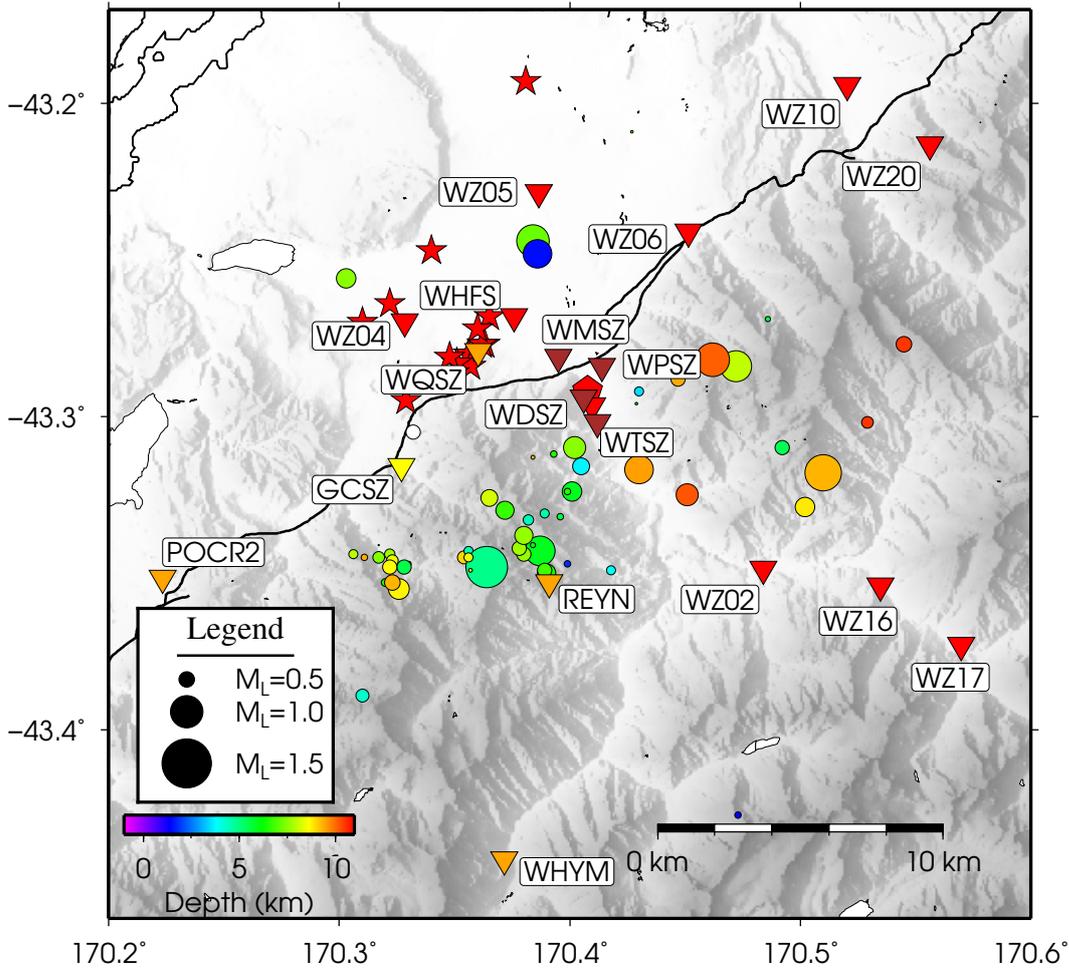


Figure 5.3: Locations of templates used in this study. Circles mark earthquakes, coloured by depth and stars mark explosion locations. Explosions occur regularly at the Whataroa quarry (next to site WQSZ), and depths are fixed to zero. Locations are often poorly constrained for explosions.

## 5.3 Data

In this study, we use data from three temporary passive seismic deployments alongside data obtained from the GeoNet national network<sup>1</sup> (Figure 5.1). The SAMBA network has been running almost continuously since late 2008 [Boese et al., 2012, 2014], notwithstanding intermittent data losses (Figure 5.2). All sites of the SAMBA network are short-period 2 or 4.5 Hz sensors recording at sample rates of 200 Hz. The WIZARD network [Thurber et al., 2012] was active between January 2012 and January 2014, with a combination of high alpine sites deployed with short-period sensors, and lower altitude sites with broadband sensors, both recording at 100 Hz.

In February 2013, a further four shallow borehole seismograph sites were deployed around the proposed DFDP-2 drill-site. These sites are short-period (4.5 Hz) sensors

<sup>1</sup><http://info.geonet.org.nz/display/appdata/Applications+and+Data>

deployed in 28 m-deep holes. These sites recorded at 250 Hz initially, but since March 2014 have been recording at 200 Hz to reduce data volumes for real-time telemetry during drilling. All sites record in continuous format and are digitized on-site with timing information obtained from GPS receivers. GPS issues rendered data from one DFDP site and occasionally from discrete SAMBA sites unreliable for cross-correlation as we require highly accurate timing to compute network cross-correlations at specific lag-times relative to all other stations.

When computing cross-correlations, we downsample all the channels used to a consistent 100 Hz sample frequency to match the sampling frequency of the GeoNet sites. A key site in this study has been the GeoNet site GCSZ, deployed as part of the DFDP-1 observatory in Gaunt Creek. This sensor is deployed in the fault-zone of the Alpine Fault and has been used to detect fault-zone guided waves [Eccles et al., 2015]; it provides some of the highest quality waveforms for the events studied here.

## 5.4 Methods

### 5.4.1 Detection

We first conduct standard earthquake detection using an energy-based STA/LTA (ratio of short-term average to long-term average) detector applied to the continuous data collected between the installation of the DFDP-2 borehole sites (February 2013) and October 2013, in order to detect a suite of template events. We specifically target events within a radius of c. 20 km of the Whataroa Valley (Figure 5.3). Detected events are then picked manually to obtain accurate body-wave first-arrival picks at as many stations as possible. These picks are used to compute initial locations using *hypocentre* [Lienert et al., 1986]. In this way we extract 69 possible template events (Figure 5.3) to carry forward to the match-filter detection stages. In addition to these newly detected events we also include eight high-quality earthquakes catalogued by Boese et al. [2012]. Templates from earlier periods were not included due to changes in station distribution, and a lack of earthquakes within the region. We include 14 explosions in our study from the nearby Whataroa quarry to test our detection method, but these explosions are not included in our final relocation.

We subsequently use the catalogued and verified earthquake waveforms as templates for a further network cross-correlation detection routine between 17 August 2012 to 31 December 2014. Template events are extracted from continuous waveform data, where the data have been bandpass filtered between 5–15 Hz. We use this frequency band because, despite dominant frequencies of these events extending to above 40 Hz, this

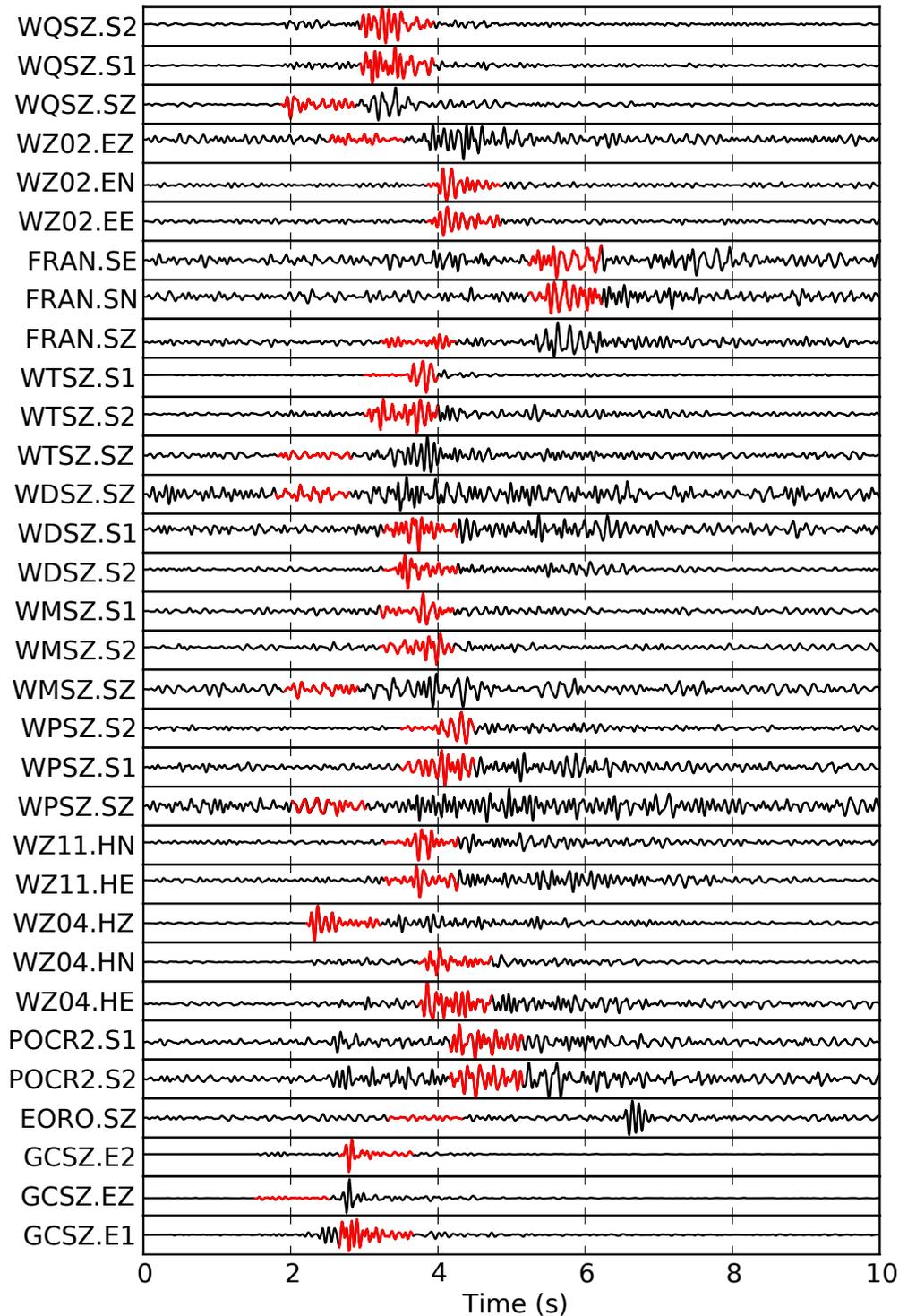


Figure 5.4: Waveforms extracted for use as a single template event from an earthquake detected on 2013/04/17 at 13:36 UTC. Red waveforms mark the 1 s template plotted within the background (black) waveform. Waveforms are plotted with a 5–15 Hz band-pass applied, as used in the detection routine.

lower-frequency band is more coherent and allows for more detections, while retaining sufficient individual event information. Template windows for each channel are 1 s long and begin 0.1 s before the manual pick for that channel (Figure 5.4). We construct P-phase templates on the vertical channels, and S-phase templates on the horizontal channels.

The network cross-correlation detection routine is similar to that developed for use in the central South Island with low-frequency earthquakes by Chamberlain et al. [2014], and on the matched-filter routine of Shelly et al. [2007]. In these methods an initial template event is correlated with continuous (day-long) waveforms. High correlation values indicate likely repeats or near-repeats of the signal.

When detecting new events, we use all channels of data represented in the template, correlated with the day-long data from the same station and channel, bandpassed at 5–15 Hz. Continuous data were lagged according to the appropriate delay (channel pick time—first pick-time) for that station-channel pair. The normalised cross-correlation values for each day of data are summed across the channels to give a network cross-correlation sum time-series. Events are then detected from this network cross-correlation sum when it exceeds a threshold value of  $10 \times$  Median Absolute Deviation (MAD), a statistic robust to outliers in noisy data [Shelly et al., 2007]. We choose to use a high threshold ( $10 \times$  MAD gives a theoretical rate of 1 false detection per 100 days across all 77 templates) as we aimed to obtain relocations for as many events as possible. With the approach taken here, this requires high single-channel normalised cross-correlation values.

### 5.4.2 Correlation pick-determination

Precise pick-times for individual channels of individual detected events are calculated by extracting a window of 1.5 s around the detection time (start of window is 0.25 s before the detection time) and cross-correlating this with the original template event. The corrected pick-time is assigned to the time of the maximum (positive) cross-correlation value within the calculated 0.5 s correlation function. The absolute times of these picks are then stored in *Seisan* S-files and located initially using *hypocenter* [Lienert et al., 1986] to generate origin times for events.

Post-processing of the detected events prior to final relocation is undertaken to regularise the location results by only including events with medium to high correlation values. In this way we remove events from the catalogue if their average single-channel correlation value is below 0.5, or if they have picks from fewer than five sites (P or S). We also remove picks with correlation values below 0.2 as these low correlation values within otherwise strong detections have been found to be associated with data

artefacts (gaps padded with ‘nan’ values, or electrical noise).

Our final step in post-processing is to remove duplicate events, defined by a detected event with the same origin time ( $\pm 0.5$ s) as another event detected by a different template. We define this based on origin time rather than pick time due to the range of data qualities within the templates (e.g. the first pick is often not made on the same site for multiple templates), and the small geographic region examined. In the case of multiple detections with similar origin times, we only retain the detection with the highest average cross-correlation value.

### 5.4.3 Event location

We compute double-difference relocations for all detected events using a 1D velocity model. We solve the double difference equations by least-squares in *hypoDD* [Waldhauser and Ellsworth, 2000], using the cross-correlation-derived picks as catalogue pick times, with individual absolute pick-times weighted with the same values as the manually assigned template weights for that detection–template pairing. We use a secondary correlation routine in *Seisan* [Havskov and Ottemoller, 1999] to compute relative cross-correlation-derived pick adjustments for all detected events relative to all others. We only use cross-correlation values exceeding 0.5, and weight the cross-correlation corrections by the correlation value. We restrict our corrections to be within  $\pm 0.1$ s of the pick-time, on the assumption that we already have very good picks from our cross-correlation determined absolute picks.

Due to the cross-correlations taking into account all of the waveform information, rather than purely the first arrival, locations are centroid rather than hypocentre locations. However, for the small events we detect here, differences between centroid and hypocentre should be small due to small rupture areas. We also note that due to likely large velocity changes between the hangingwall (schist) and footwall (granite overlain with gravels in this valley), which are not included in our 1D velocity model, events may be clustered incorrectly by *hypoDD* if they are hypocentrally close, but on opposite sides of the Alpine Fault. This would lead to significant variation in ray-paths for earthquakes located on opposite sides of the fault; given that a key assumption of the double difference method is that clustered events follow the same raypath, this would lead to higher uncertainty in relocations.

### 5.4.4 Repeater analysis

Following detection and pick-correction we extract a subset of repeating earthquakes from our catalogue defined by high cross-correlation values. Our threshold for repeating

events is based on that used by Rubinstein and Ellsworth [2010], and we required cross-correlations values of at least 0.95 on 5 or more stations. These cross-correlations are computed using 1.28 s of 5–40 Hz bandpassed, 100 Hz sampled seismic data beginning 0.28 s before the pick time. We only compare events within families, as part of the declustering routine compares cross-correlation values for events at similar times and assigns events to the most similar family.

### 5.4.5 SVD magnitude estimation

Due to high levels of similarity between events within families we are able to calculate magnitudes for a subset of the most similar events using singular value decomposition (SVD). We attempted to compute automated magnitudes for all events using a simple automated amplitude picker, but found the uncertainties in these magnitudes to be very high. In contrast, magnitudes derived by SVD have been shown to have lower uncertainties than catalogue magnitudes [Rubinstein and Ellsworth, 2010] when considering highly correlated events.

Before computing magnitudes for events, we select events with high single channel cross-correlations, relative to the template, and pre-process them. We use only the vertical channels of seismographs to reduce the computational requirements. We do not remove the instrument responses as the SVD method computes relative amplitudes at a single site. We then compare these relative amplitudes across the network, rather than comparing the true amplitudes. None of the sites used to compute SVD magnitudes had any known changes in response during the study period (e.g. the sensor was constant and the digitizer type was constant).

When selecting traces to use for SVD analysis, we require highly similar waveforms that can be adequately represented by the first singular vector. As we have already aligned our traces in the cross-correlation picking stage (relative to the template) we should not need to re-align them; however, we found that a further re-alignment by cross-correlation relative to the stack of the waveforms increased the similarity of waveforms. These cross-correlations were computed using 5.0 s-long windowed waveforms beginning 1 s before the cross-correlation derived P-pick, resampled to 100 Hz and bandpassed between 2–10 Hz. Calculated shifts were on the order of 1 sample (0.01 s), and likely arise due to the higher signal-to-noise ratio of the stack compared to the raw template.

We then extract a smaller, 4.0 s, window for SVD analysis, beginning 0.5 s before the re-aligned P-phase. Of these re-aligned traces we only use those with a correlation-coefficient greater than 0.85 relative to the stack. We compare with the stack as this is similar to the first singular vector [Barrett and Beroza, 2014], and therefore should provide an estimate of how well the first basis vector describes the waveforms. We also

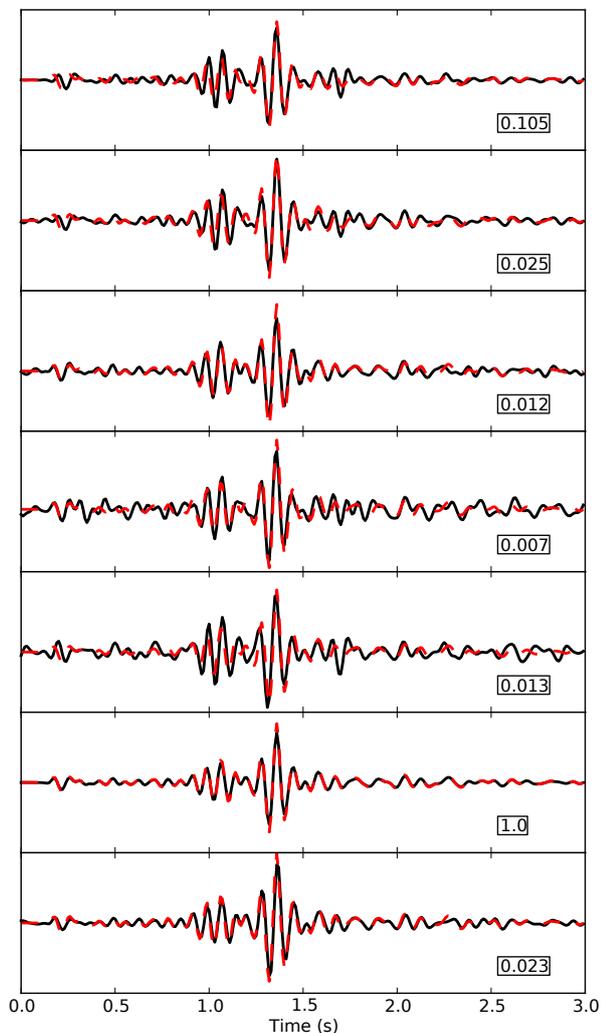


Figure 5.5: First basis vector (red, dashed) scaled to match the maximum amplitude of a family of earthquakes. Seismograms plotted are the filtered vertical channel of GCSZ (black). The number on the right of each waveform indicates the relative amplitude of the event scaled to the largest event.

require at least three events per channel for that channel to be included in the SVD analysis.

The SVD method is used to compute the basis-vectors of a set of horizontal vectors (in this case seismograms)  $\mathbf{M}$ , by solving the equation:

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V} \quad (5.1)$$

where  $\mathbf{U}$  is the set of input basis-vectors,  $\mathbf{\Sigma}$  is the array of singular values, and  $\mathbf{V}$  is the matrix of output basis-vectors. The output basis-vectors describe the input matrix of seismograms, with the first basis-vector ( $\mathbf{V}_1$ ) having the most power, and those with higher order having lower descriptive power. The relative descriptive power of a single output basis-vector is given by the appropriate singular value (e.g.  $\mathbf{V}_i$  has the descriptive power of  $\mathbf{\Sigma}_i$ ).

The relative moments of the output basis-vectors are given by their complementary input basis-vectors ( $\mathbf{U}$ ). We use the relative amplitudes given by the first input basis-vector to describe the relative moments of events within a family, for a single channel.

We assume that higher-order basis-vectors describe mostly small-scale variations in either noise, or seismogram mis-alignment (Figure 5.5).

To combine relative moments from single channels to a total event relative moment we follow the method of Rubinstein and Ellsworth [2010]. As not all stations record every event, a mean across sites may unfairly bias events with few readings. To circumvent this we pose the calculation as an inverse problem with the input as a series of randomly paired relative moments, and the model as the actual relative moments. This is posed such that the weighted sum of paired relative moments equals zero. For example, as Rubinstein and Ellsworth [2010] explain, if a first event was half the size of a second event the following would arise:

$$2.0m_1 - m_2 = 0 \quad (5.2)$$

where  $m$  are the event relative moments. In all cases the relative moments are normalised by the smaller of the two relative moments.

Incorporating more events for a single station generates an increasingly large square matrix in which each row has only two non-zero elements. To incorporate more stations into the system further matrices are vertically concatenated. Any empty rows (corresponding to a station not being used for an event) are removed. The width of the resulting matrix corresponds to the total number of events and the length is the sum of the number of events recorded at each station. This results in the system:

$$\mathbf{KM} = \mathbf{D} \quad (5.3)$$

where  $\mathbf{K}$  is the matrix describing the relative amplitudes of all events,  $\mathbf{M}$  is the unknown vertical vector of relative event moments, and  $\mathbf{D}$  is a column of zeros. Following the convention of Rubinstein and Ellsworth [2010] we append a final row to  $\mathbf{K}$ , whose elements are given by  $1/N$ , with  $N$  as the number of events, to normalise the relative moments  $\mathbf{M}$  to 1. We weight this final row by the number of events, and solve by least squares.

To enable comparison of events between different families we convert from relative moment to local magnitude, using the local magnitude of the template to normalise all the moments. This template magnitude is converted to moment magnitude using the relationship obtained by Ristau [2009] for shallow earthquakes in New Zealand;

$$M_L = 0.88M_W + 0.73 \quad (5.4)$$

where  $M_L$  is the local magnitude and  $M_W$  is the moment magnitude. This scale has been derived for earthquakes of larger magnitudes than we are studying here,

however, we are using this as an interim step, and remove this calibration later. Our template events (and therefore our final magnitudes) are calibrated to approximate the local magnitude scale of Boese et al. [2012], without using the frequency dependent attenuation term found by these authors.

This moment magnitude is then converted to seismic moment using the standard relation [Hanks and Kanamori, 1979]:

$$M_W = 2/3(\log M_0 - 9.0). \quad (5.5)$$

We then normalise by multiplying all relative moments by the ratio of the converted template moment and the calculated template relative moment. For this to work we must have calculated the relative moment of the template by the previous SVD analysis. We then convert all the moments back to local magnitude using the previous relationships. We have not compared our final magnitudes to other event magnitudes, and therefore do not present magnitudes that can be directly compared to other magnitude scales (such as the local magnitude scale of [Boese et al., 2012])

#### 5.4.6 Focal mechanism inversion

For the subset of events with clear, impulsive, first-motions (24 earthquakes), we calculate focal mechanisms using the Bayesian method of Walsh et al. [2009]. With this approach, earthquakes are first located probabilistically using the NonLinLoc [Lomax et al., 2000] location program. This generates a posterior density function of probable earthquake locations and take-off-angles for each station. These take-off angles and polarity picks are used to constrain a least-squares inversion for focal mechanism strike, dip and rake. We allow the inversion to span all strikes, dips and rakes. We compared our results to those from a constrained inversion for only positive rakes [c.f. Walsh et al., 2009], but found that some solutions were less robust.

## 5.5 Results

We detect 495 events in the search region, of which 195 are quarry blasts (Figure 5.6, full catalogue included in the electronic appendices). In the same time period the national seismic network recorded two earthquakes in the study area. Earthquake detections are concentrated close to their template event in time; 54% occur within 10 days of their template event (Figure 5.7). In contrast, 11% of explosions detected occur within 10 days of their respective templates. This temporal relationship is also evidences at

other levels: 90% of earthquake detections occur within 53 days of their template, and 50% occur within 9 days, whereas explosions detections are more distributed in time (90% at 376 days and 50% at 171 days).

Of the 63 earthquake templates used, 42 detected events other than themselves. The numbers of events detected per family varied from one to thirty (discounting self-detections). In contrast, of the 14 explosion templates, nine detected events other than themselves, with up to 49 explosions detected by one template.

Of the 300 earthquakes recorded, 11 events within 5 families have similarities exceeding our threshold for repeating earthquakes. These highly similar events have S-P time differences of less than 1 sample (1 sample=0.01 s at the 100 Hz sampling frequency used) when aligned on the dominant phase. This would put their separation at less than 100 m assuming a constant P/S ratio and P-wave velocity of 4.7 km/s (upper layer velocity used in our 1D velocity model). Of these repeaters, 4 of the 5 families have only two events (one repeat), while the other has three events (two repeats). One family has an inter-event time of less than 1 hour (13.5 minutes). Within families not all detected events are classified as repeaters, due to larger differences in waveforms.

We are able to obtain relocations for 157 of our 300 earthquakes (Figure 5.8). Events cluster around the template used to detect them, rarely deviating more than 500 m from the template source. Our earthquakes do not appear to define a single fault-plane, rather they are distributed through the crust close to where we expect the Alpine Fault to be, with apparent dips of between 45–60°.

Despite low signal-to-noise ratios, we are able to obtain focal-mechanisms for 24 events in our catalogue (Figure 5.10). These events nevertheless have poorly constrained focal mechanisms due to a paucity of available picks resulting from low signal-to-noise ratios (Figure 5.9). They are dominated by strike-slip solutions with one focal plane close to alignment with the Alpine Fault. We also see some normal motion (e.g. event five in Figure 5.9) which may be due to mis-location; the scatter from the 1D non-linear locations is wrapped around both sides of the focal-sphere. This uncertainty is illustrated by the poorly constrained P and T axes in Figure 5.9.

We obtain high-precision SVD-derived magnitudes for 106 earthquakes (Figure 5.11) ranging between  $M_L - 0.46$ –1.60. The magnitude-frequency relationship of these earthquakes above a completeness value of  $M_C = 0.6$  are fit by a power-law relation with a high b-value of 1.44 indicating a large quantity of small magnitude events. We are unable to generate magnitudes for the majority of events due to low signal-to-noise ratios. These low signal-to-noise ratios result in low cross-correlation values which exclude them from our SVD magnitude calculation. They also result in unreliable manual and automatic amplitude picks, precluding the use of standard local magnitude calculations.

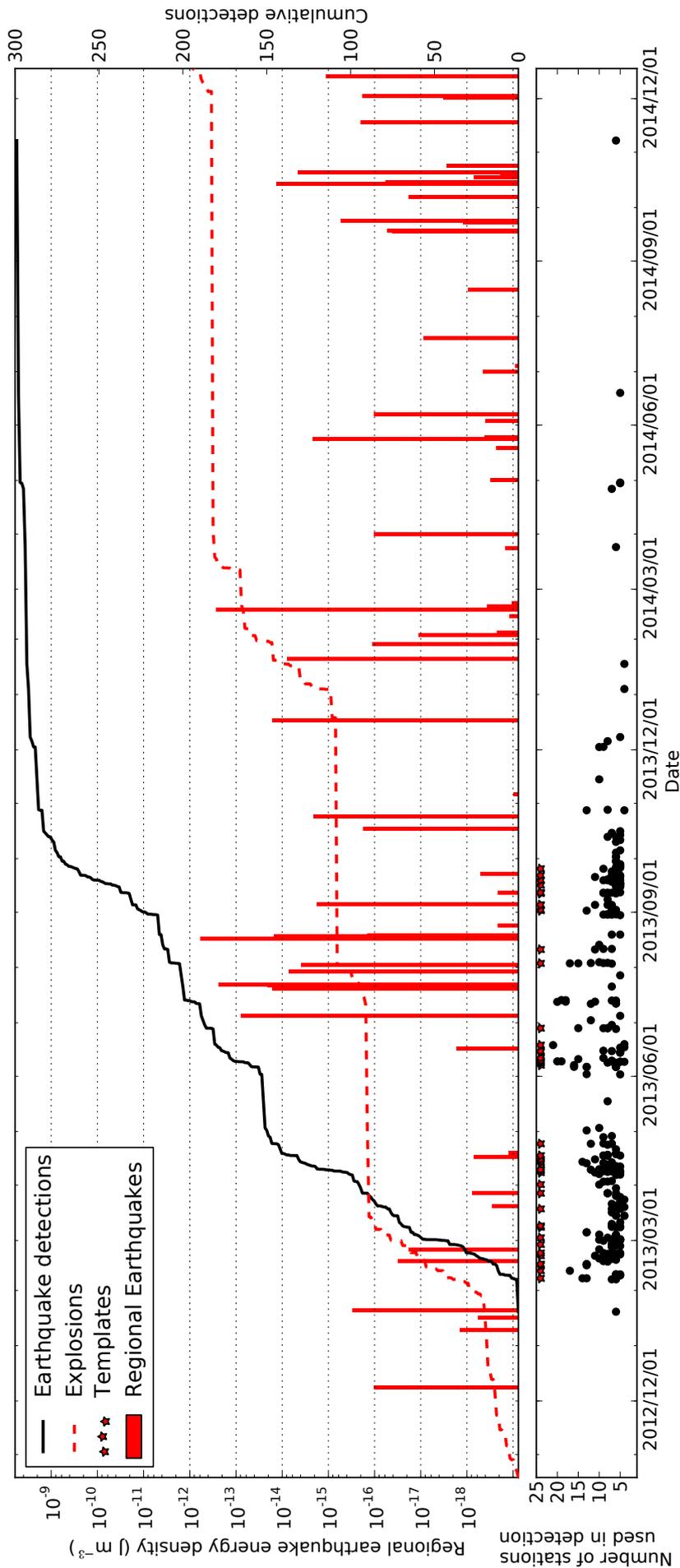


Figure 5.6: Main panel: Cumulative densities for earthquake (solid black line) and explosion (dashed red line) templates. Also plotted as red bars are the energy densities from regional earthquakes of the GeoNet catalogue (calculated for the DFDP-2 drill-site). The lower panel shows the number of stations used in the earthquake detection. Template detection times are marked by red stars in the lower panel, the number of stations used is fixed rather than being indicative of how many stations are used in the template.

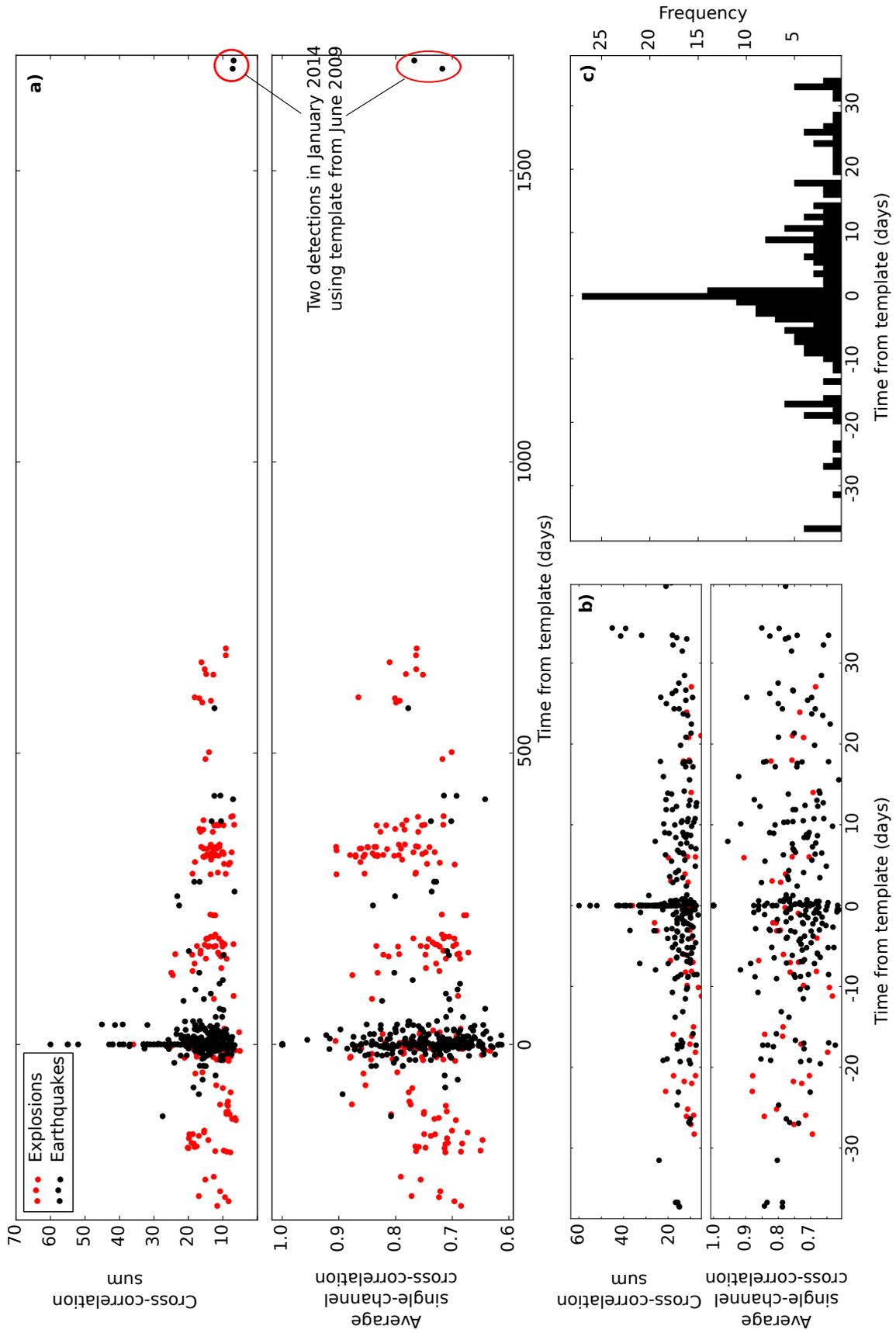


Figure 5.7: Caption overleaf.

Figure 5.7: Previous page: a) Cross-correlation sums (upper) and averaged single-channel cross-correlation values for individual detections plotted against the time difference between the detection and its respective template. Negative time indicates a detection before the template event, while positive time represents a detection after the template. Detections at time=0 are self detections and have the strongest cross-correlation values. Earthquakes are plotted as black circles and explosions from the Whataroa quarry are plotted as red circles. Explosions tend to have more detections at larger relative times than earthquakes. There is no obvious decay in cross-correlation values with time: in fact, the lowest average cross-correlation values are found close to time=0. Two outliers at large positive times are the sole two detections in January 2014, made by a template from June 2009. b) Zoom of 79 days around time=0 to highlight the majority of earthquake detections. c) Histogram of detections per day (self-detections removed) relative to the template time (same time scale as panel b). In the five to ten days prior to template times, there are more detections than in the same days after the template, and the detections increase with proximity to the template.

## 5.6 Discussion

### 5.6.1 Temporal evolution

Our earthquake detections are predominantly limited to the period directly surrounding the template events, despite continuous high-quality data spanning the study-period. After October 2013, we obtain only 13 earthquake detections from 9 families (Figure 5.6). In 2014 there was a reduction in station number (Figure 5.2), but this change in network does not correspond to a change in detection rates. In general, the number of stations used to detect earthquakes correlates poorly to detection rate.

Using a set of 14 explosion templates, we are able to detect 195 explosions which, despite being clustered in time during periods of quarry work (although we have been unable to validate the explosion times, or shot-sizes with the quarry owners), can have large inter-event times for individual templates (Figure 5.7). This is illustrated by individual templates generating detections during multiple periods of activity, rather than just being capable of detections close to the template time. Of note is that this is despite the explosions being a destructive source, therefore the source properties must be different between separate explosions (different shot-size, coupling, location and rock-mass).

Correlation values for both earthquake and explosion templates do not deteriorate with time (Figure 5.7). From this, and the more long-lived detection capability of our explosion templates, we infer that the paucity of earthquake detections after October 2013 is real. This likely represents a reduction in activity for our discrete templates, rather than a reduction of seismicity rates at a larger scale. We infer that our families

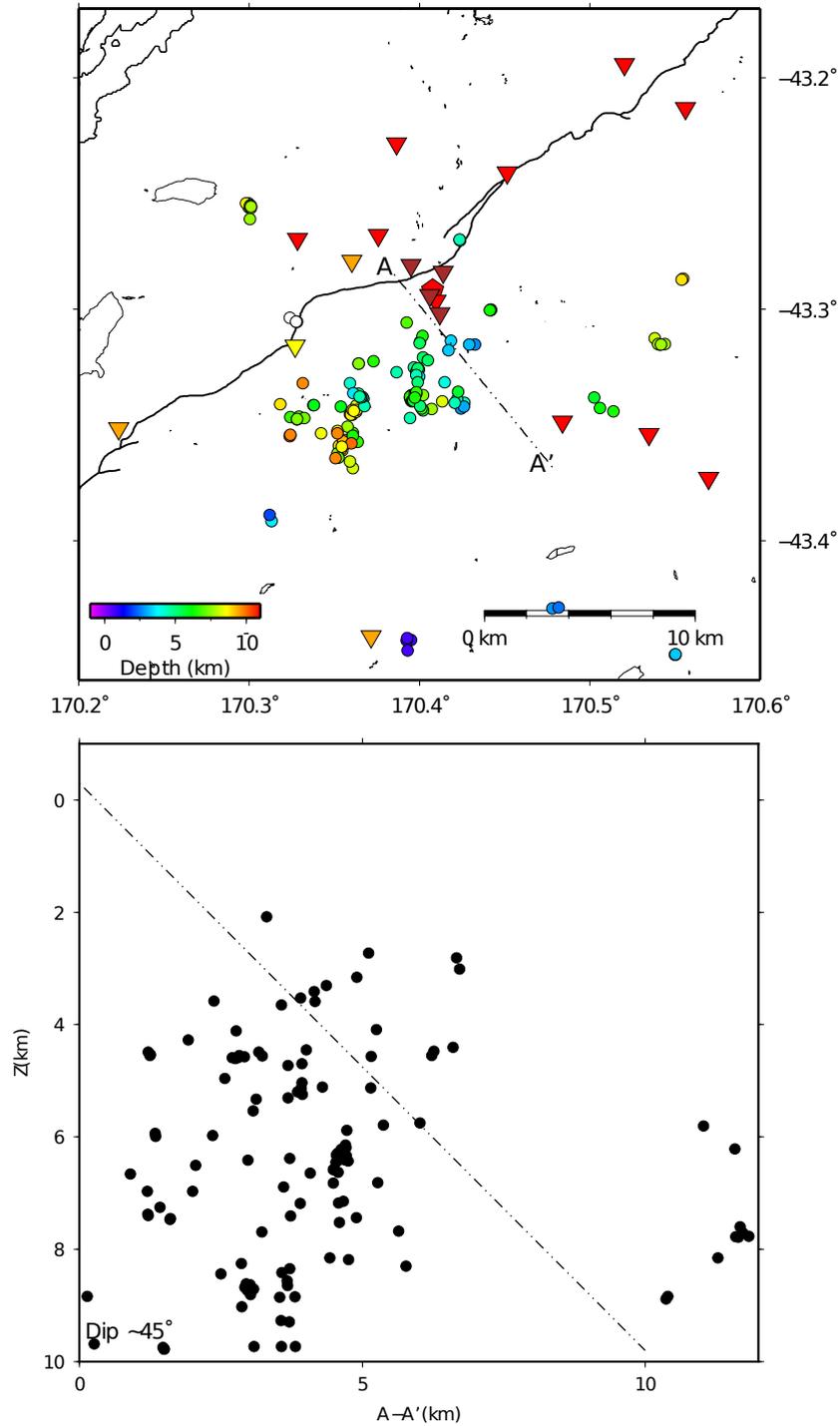


Figure 5.8: Earthquake relocations as obtained for 157 of the detected events. Events cluster around template locations, with variations of less than 500 m. Left panel shows a cross-section plotted along the line A–A'. Dashed line on cross-section is plotted at a representative dip of 45°.

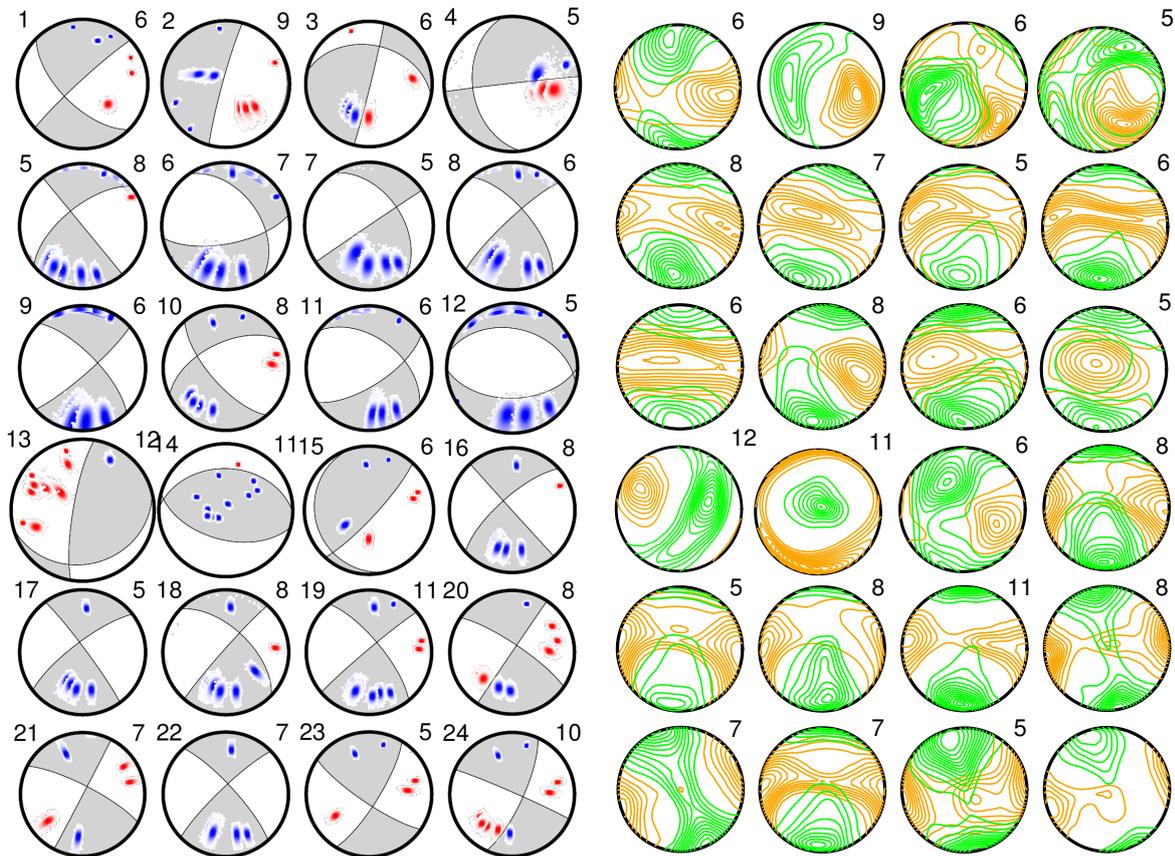


Figure 5.9: Focal mechanisms for all earthquakes that we are able to constrain. Each earthquake is plotted twice, once with picks and shaded best fit focal mechanism (left) and once with pressure (orange)-tension (green) axes with their respective uncertainty contours. On the left plots, stations are shown with their uncertainty scatter obtained from non-linear location of the earthquakes. Red stations indicate a downward first motion of the waveform at that site, while blue indicates upward. Due to the low magnitudes of events studied here, we were unable to obtain accurate picks for most sites, which results in poor coverage of the focal sphere and thus poorly constrained mechanisms. Nevertheless, of the 24 mechanisms obtained there is a propensity for strike-slip mechanisms, with one plane subparallel to the Alpine Fault, and giving the expected dextral faulting. Numbers on the right of beach-balls indicate the number of polarities used to constrain the mechanism, while the number on the left is the earthquake id-number ordered by date.

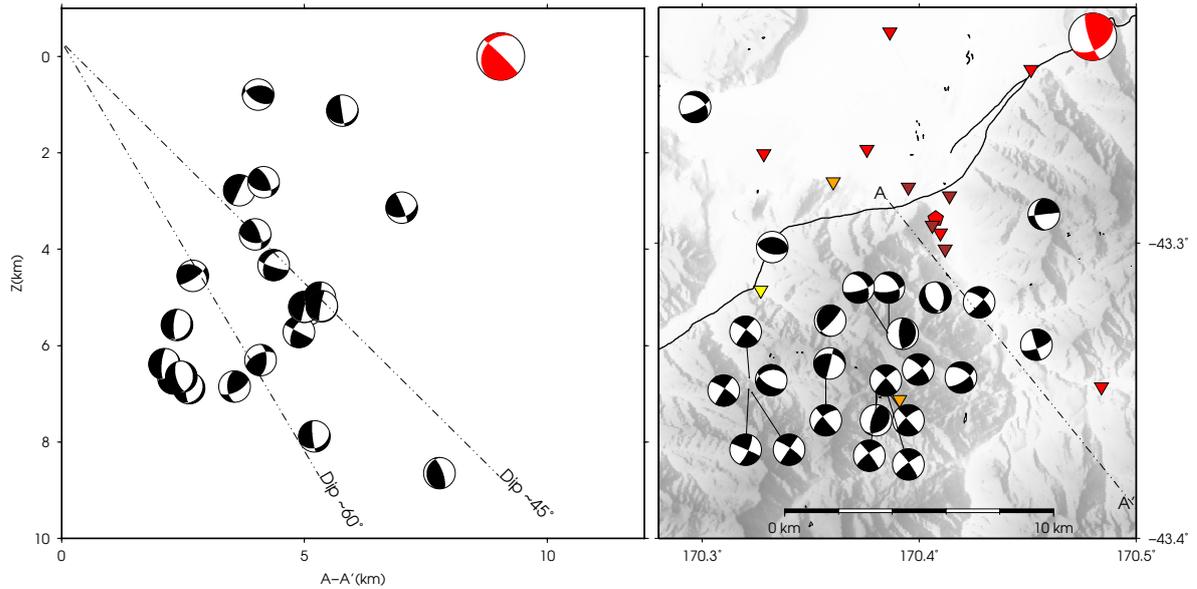


Figure 5.10: Focal mechanisms computed using the Bayesian routine of Walsh et al. [2009]. The locations plotted differ from those in Figure 5.8 due to the different location methods. Red focal mechanism indicates an expected oblique dextral-reverse Alpine Fault earthquake, with strike  $057^\circ$ , dip  $45^\circ$  and rake of  $160^\circ$ .

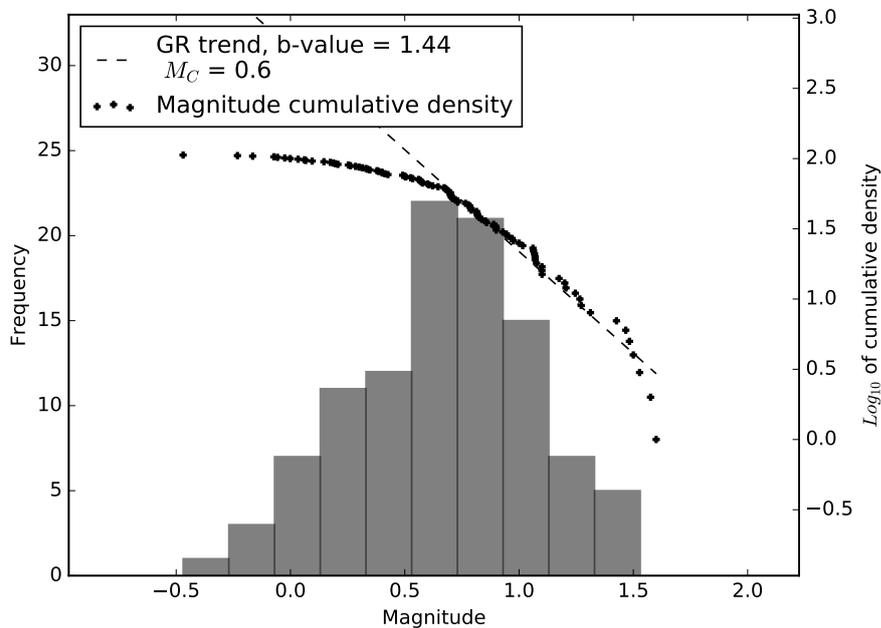


Figure 5.11: Magnitude histogram for the 106 earthquakes similar enough to calculate high-precision singular-value decomposition-derived magnitudes. Magnitudes are computed relative to manually-determined template local magnitudes (using the scale of [Boese, 2012]). We fit the cumulative density function with a power-law above a completeness of  $M_C = 0.6$ , and below  $M_L 1.4$ .

therefore represent local, usually non-repetitive, failure of small patches or asperities.

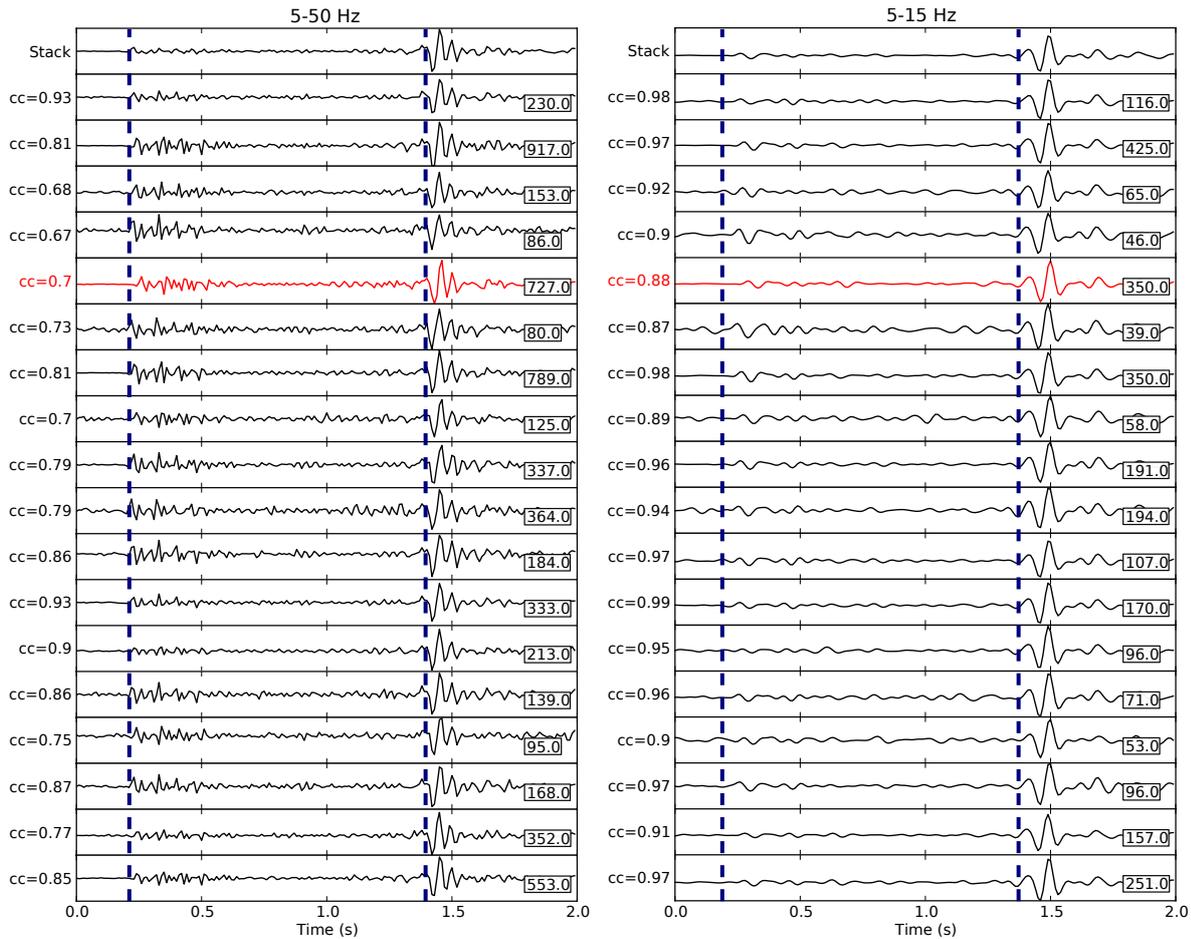


Figure 5.12: Near-repeats for template event 2013-03-02-45331 at site GCSZ, vertical channel, for correlation values above 0.8. Left panel shows the 5–50 Hz filtered seismograms, with high correlation values, right panel shows the 5–15 Hz filtered seismograms. Correlation values are denoted by ‘cc=’ and the number in the box of each axis shows the maximum amplitude in that plot in counts. Blue dashed lines indicate approximate arrivals of body phases. The red trace shows a clear delay in P arrival.

## 5.6.2 Locations

Rather than the earthquakes detected here having hypocenters delineating some structure (e.g. the Alpine Fault), they are distributed in a damage zone throughout the crust, near the Alpine Fault (Figure 5.8). The lack of seismicity on the fault would suggest that the fault is either locked and not able to fail in small earthquakes at this point in the seismic cycle (prior to an expected large magnitude earthquake), or is deforming aseismically. As there is no observed creep at the surface the second scenario seems unlikely, and the Alpine Fault is likely locked in this region and accumulating strain.

We observe earthquakes near the Alpine Fault, as also seen by Boese et al. [2012],

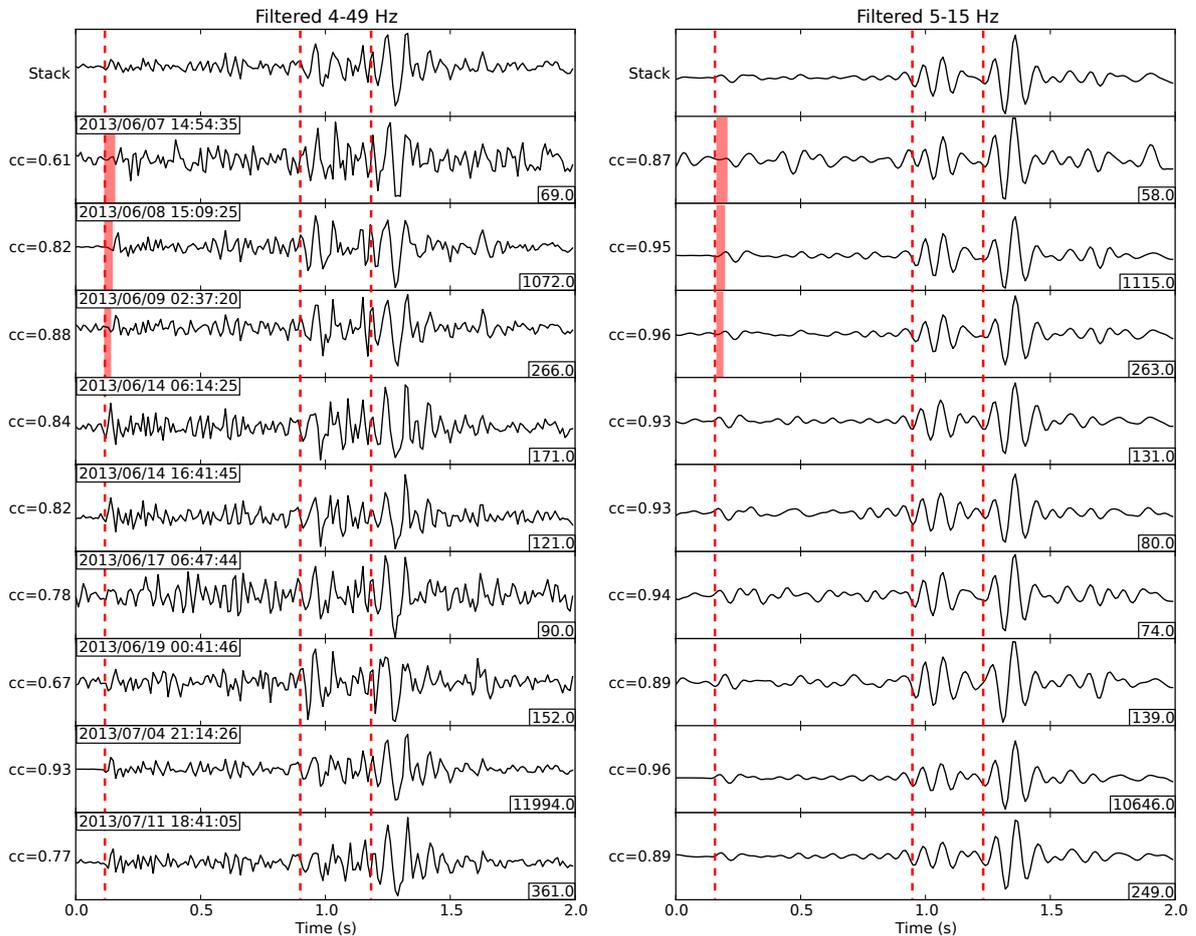


Figure 5.13: Vertical channel seismograms from events detected within one family on site GCSZ. Phase arrivals are marked by red dashed lines, with red shading showing P-delays for three earthquakes. All earthquakes occurred within a month of each other (plot start time is given in the top left corner). Maximum amplitudes in counts are given in the bottom right corner, and span the range 69–11994 (the largest amplitude is approximately 174 times bigger than the smallest), suggesting a difference in magnitudes of  $c. 2.2$ . The deviation in S–P times of up to 0.04 s indicates either location differences between events of up to 500 m, or small velocity changes. We observe multiple phases for many event documented here (tertiary phase marked by the third red-dashed line)

O’Keefe [2008], with less seismicity away from the fault (34 earthquakes are located more than 5 km from the fault, and 106 earthquakes locate within the first 5 km of the hanging-wall). Other authors have suggested that the low-seismicity zone between Franz Josef and the Wanganui river represents a strong, unfractured block [e.g. Boese et al., 2012]. While we can say little about the presence of seismicity away from the fault, near the fault, seismicity is dominated by low-magnitude events, likely on multiple small asperities, rather than on one large structure. These earthquakes likely represent failure at the margins of the presumed strong block, with strike-slip focal mechanisms consistent with the background patterns of seismicity in the central Southern Alps.

We are unable to compute robust uncertainty estimates by solving the double-difference equations by singular-value decomposition due to the size of the problem. Instead, we estimate uncertainties by comparison of our double-difference relocations with the locations obtained by non-linear probabilistic locations for the subset of events for which we generate focal mechanisms. These are shown in Figures 5.8 and 5.10. Epicentral changes in location are on the order of 2 km, whereas depth uncertainties are smaller, on the order of 1 km.

### 5.6.3 Repeating earthquakes

A small number of events within our catalogue have high enough similarity on enough stations to be considered as repeating earthquakes [Rubinstein and Ellsworth, 2010]. However on the highest signal to noise ratio site, GCSZ, we observe consistently high similarities for earthquakes within given families (Figure 5.12), indicative of small location changes and only minor changes in focal mechanisms. We consider that the scarcity of events with high enough correlation values to be termed repeating may be due to the low signal-to-noise ratios found for these small magnitude events. To better assess the possibility of repeating earthquakes on and near the Alpine Fault, more high-quality seismic sites are required. This would likely require more borehole sites (in bedrock, the sites deployed for the Deep-Fault Drilling Project have lower signal-to-noise due to their emplacement in sediments) close to the fault.

### 5.6.4 Secondary phases

For multiple events in multiple families we observe a further seismic phase following the S-phase, and occasionally a phase arrival following the P-phase (Figure 5.13). We first observed these phases on the borehole site GCSZ, however on further inspection, we also found these secondary phases other sites near the fault, dominantly also appearing on WTSZ (Figure 5.14). This may mean that the phase is not a fault-zone guided wave,

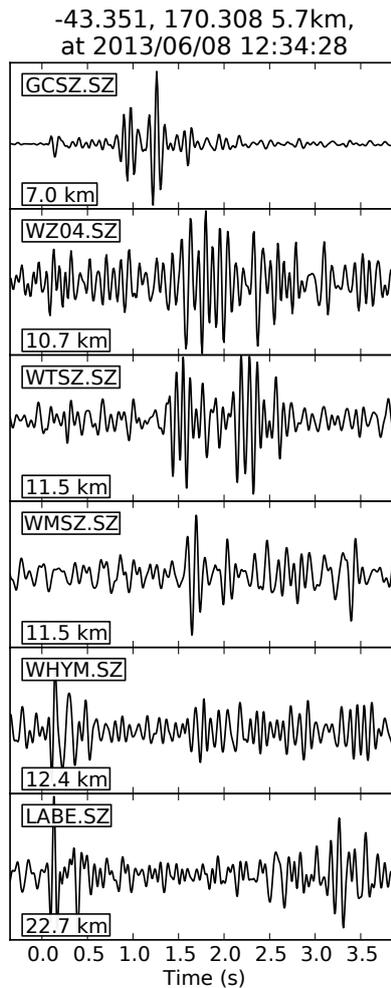


Figure 5.14: Single event aligned on the P-phase at six stations. Station names are given in the top left of every plot window, with hypocentral distance in the bottom left. Vertical channels for these stations are plotted which, for nearby stations, have clear S-phase energy. Waveforms are filtered between 5–15 Hz. At both GCSZ and WTSZ there are clearly two phases following the P-arrival.

as they have previously only been documented on the on-fault seismometer at GCSZ. However, it could be that either the damage zone (the wave-guide) encompasses more distant sites, or that the trapped wave interacts with the free-surface at the Earth's surface, allowing it to travel to other near-fault sites. We also note that the phase is occasionally, but not always dispersive (Figure 5.15).

Another explanation for the secondary phase may be anisotropy. The appearance of these secondary phases varies with source location, but does not appear to vary consistently with either back azimuth relative to GCSZ, nor distance (Figure 5.16). Delays between the main S-phase and the secondary phase of c. 0.3 s at 7 km hypocentral distance, Figure 5.14 implies  $V_p/V_s$  ratios of 1.48 and 1.68 for a fixed  $V_p$  of  $4.7 \text{ km s}^{-1}$ , or changes of c. 12%. Bourguignon et al. [2015] observed a similar secondary phase to the north of our study region which they interpreted to be due to anisotropy, with changes in  $V_p/V_s$  ratio of 9% using an assumed  $V_p$  of  $4.5 \text{ km s}^{-1}$ .

Due to the proximity of GCSZ to the fault-zone it seems unlikely that this phase would be a converted phase resulting from waves passing through, or reflecting from, the Alpine Fault, however this phase could be a conversion or reflection from some other structure. The secondary phases documented here appear similar to both the secondary

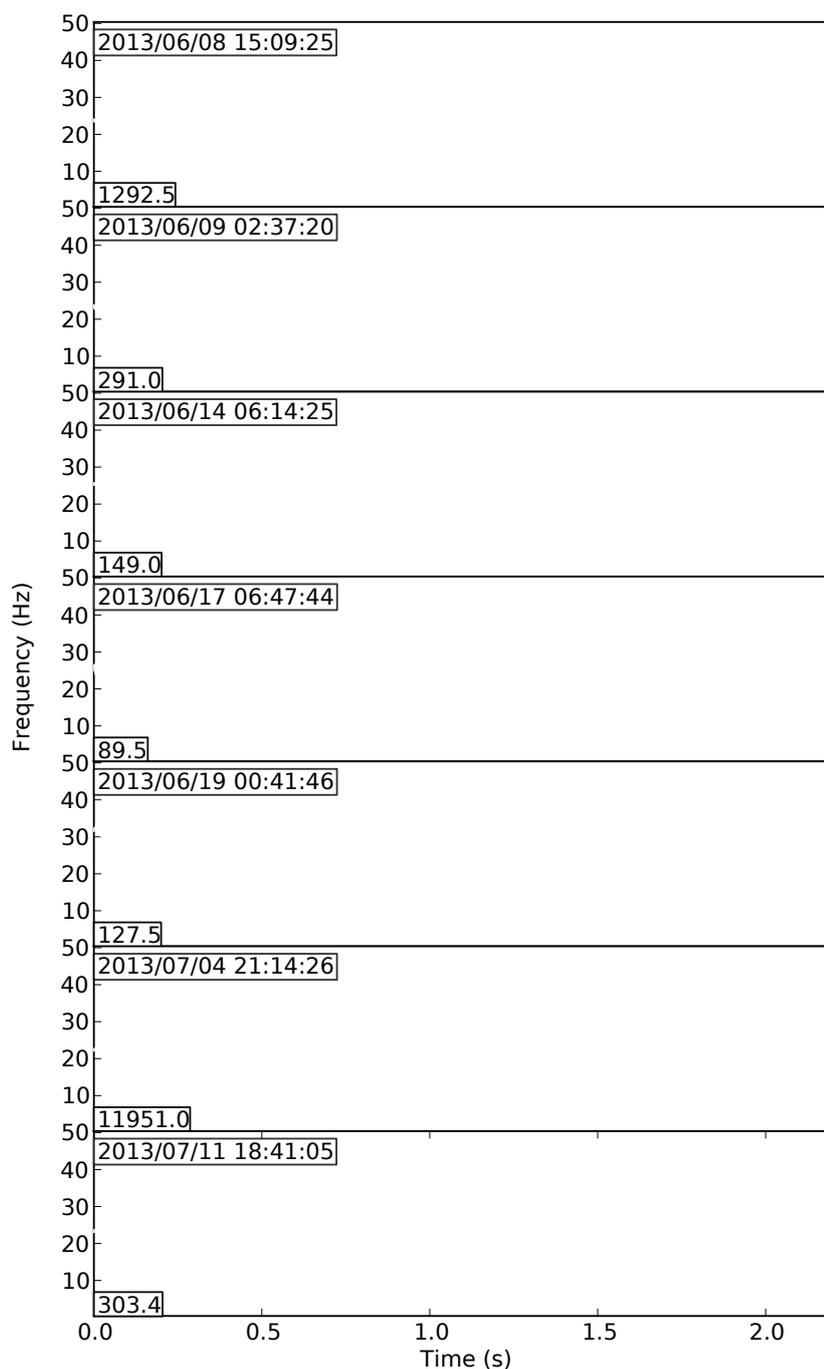


Figure 5.15: Unfiltered seismograms overlain on spectrgrams for the vertical channel of GCSZ showing tertiary phases after the S-phase. Spectral colours are normalised, with red indicating high power and blue indicating low. All events were detected by the same template. Detection times are given in the top left of every sub-plot, and the maximum amplitude within the waveforms are given in the bottom left. The tertiary phase has the most power in these plots, with peak power around 12 Hz. In some cases the tertiary phase appears dispersive, but this is not pervasive across all plotted events.

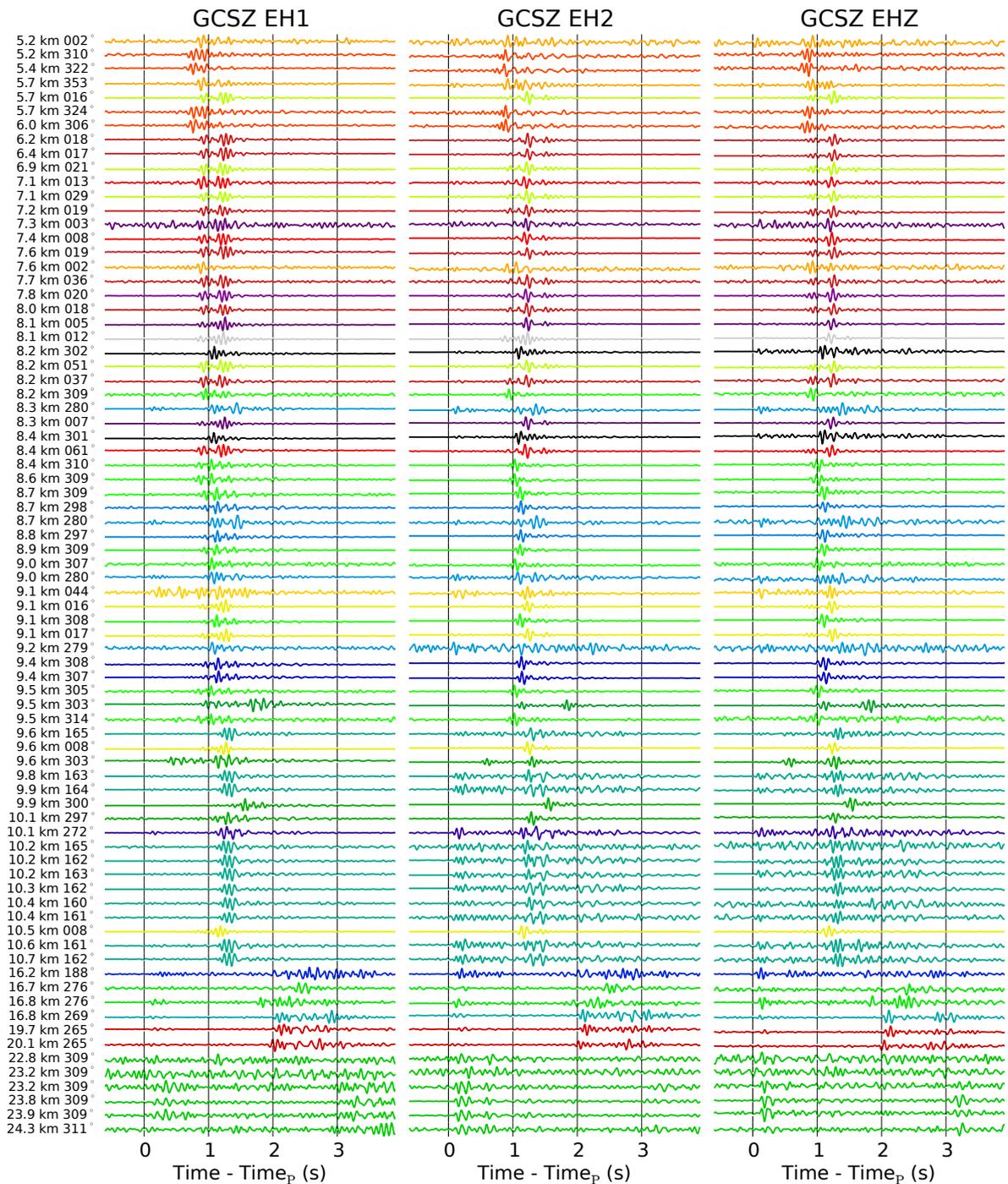


Figure 5.16: 78 detected events from 24 different families between June and October 2013 of above  $M_L 0.7$  plotted on all three channels of GCSZ. Waveforms are band-passed between 5–15 Hz, coloured by family and aligned on the P-phase. Numbers on the left show hypocentral distance and back-azimuth from GCSZ to the initial location (computed using *Hypocenter* of the event detected). Events are sorted by hypocentral distance. On multiple events less than 8 km away there are clearly two arrivals around the 1 s relative time: on more distant events, this pair of phases becomes less pronounced, with events often showing only one of the two arrivals.

S-phases observed by Bourguignon et al. [2015], and the fault-zone guided wave documented by Eccles et al. [2015]. The changes in  $V_p/V_s$  required for the anisotropy interpretation fall well below laboratory measurements of S-wave anisotropy for the Alpine Schist [Christensen and Okaya, 2007] of up to 26%. However the dispersive nature of some of the secondary phases (Figure 5.15) is suggestive of trapped waves. Further examination of these phases is required.

## 5.7 Conclusions

We have documented 300 earthquakes of small magnitude ( $-0.5 < M_L < 1.6$ ) adjacent to the Alpine Fault in a seismically quiet zone. Our detection method, and the scarcity of repeating events, means that our catalogue is limited mostly to periods surrounding template events. We therefore think that rather than our catalogue being composed of repeating events in the classical sense, we have captured small-scale repeating earthquakes on small, localised asperities at the edge of a strong block in the hanging-wall of the Alpine Fault.

Our catalogue is dominated by small earthquakes (b-value=1.44), which likely indicates a scale-limit for the asperities in this region - rather than the events occurring on large-scale faults capable of accommodating large earthquakes (either large slip, or large slip-patches), these events are constrained to small fractures in the crust.

Although this is a seismically quiet region, by applying the matched-filter methodology, we are able to extract a detailed catalogue of near-fault microseismicity. This method could be extended along strike to provide comparisons between this region and more active region. Particularly, the high b-value obtained here is in contrast to that obtained for the central Southern Alps as a whole by Boese et al. [2012] of 0.86. Whether b-values are consistently higher near the Alpine Fault, or if this section is unique could indicate stress concentrations near the fault [Nakaya, 2006].

## 5.8 Acknowledgements

We acknowledge the New Zealand GeoNet project and its sponsors EQC, GNS Science and LINZ, for providing data used in this study. We used obspy, matplotlib and eqcorrscan for our waveform plots, and gmt for our map plots. We are grateful to Adrian Benson for all his assistance in the field when collecting the data used in this study. We are also grateful to Cliff Thurber and Steve Roecker for sharing the WIZARD data used in this study.

# 6 Real-time seismic monitoring at the Deep Fault Drilling Project, Alpine Fault, New Zealand

This work is written to be submitted to Seismological Research Letters. This work has been modified to fit the formatting of this thesis only. This chapter is written by me, with assistance and editorial input from co-authors. The work was conducted as a team effort during the period of real-time monitoring, however prior to monitoring I set up the real-time network alongside Carolin Boese, Adrian Benson and the GeoNet telemetry team. I also set up the procedures and software, led training sessions, acted as one of the senior seismologists and conducted most of the post-drilling analysis that is contained in this chapter.

## 6.1 Abstract

The Deep Fault Drilling Project (DFDP) is a multinational scientific drilling effort to study the evolution, structure, and seismogenesis of the Alpine Fault, New Zealand, via in-situ measurements of fault-rock properties. The second phase of drilling (DFDP-2), undertaken in the Whataroa Valley in late 2014, was intended to intersect the Alpine Fault at a depth of 1 km. In conjunction with the drilling and on-site science activities, a real-time seismic monitoring scheme and “traffic-light” response protocol were established to detect, locate and if necessary respond to seismicity within 30 km of the drill-site. This network was operated around the clock between late August 2014 and early January 2015 and detected and located 494 earthquakes of  $M_L$ 0.6–4.2. None of these earthquakes occurred within 3 km of the drill-site, nor did any of the seismicity detected require changes to drilling operations. The monitoring was undertaken using free-to-research software operated by an international team of 16 seismologists (including eight postgraduate students) working in seven institutions and three countries to provide rapid on- and off-site manual checking and relocating of

events. Our standard response time between detection and final location was less than 30 minutes under normal background seismicity conditions and up to 1 hr during swarm activity and for low-priority, distant ( $\geq 30$  km epicentrally from the drill-site) earthquakes.

## 6.2 Introduction

The Deep Fault Drilling Project (DFDP) is a scientific drilling project aiming to sample and monitor the *in-situ* conditions of New Zealand's Alpine Fault prior to an expected major (M7–8) earthquake [Townend et al., 2009, Berryman et al., 2012, Sutherland et al., 2012]. The Alpine Fault is New Zealand's major on-land plate boundary structure, undergoing transpressional deformation due to the oblique convergence of the Australian and Pacific plates. The Alpine Fault accommodates  $\sim 75\%$  of the dextral strike-slip inter-plate motion, and up to 100% of the compressional motion [Norris and Cooper, 2000] in the central section of its 650 km length. Slip rates range from 23–27 mm/yr dextral [Norris and Cooper, 2000, Sutherland et al., 2006, Barnes, 2009] and 6–10 mm/yr reverse motion [Norris and Cooper, 2000], with the highest uplift corresponding to the highest topography around Mt. Cook in the central Southern Alps.

The Alpine Fault separates the east-dipping Fiordland subduction zone to the southwest from the west-dipping Hikurangi subduction zone to the northeast. In northern South Island, crustal deformation is accommodated on a distributed network of faults (Marlborough Fault Zone) as deformation is transferred from the Hikurangi subduction zone to the transpressional Alpine Fault [Berryman et al., 1992]. The faults of the Marlborough Fault Zone converge to the main Alpine Fault structure around 100 km north of the Whataroa Valley, the site of the DFDP-2 drill-hole. In the central portion of the Alpine Fault, the fault appears to be a singular structure, however at the surface there is segmentation into individual strands of oblique-reverse and dextral sense [Norris and Cooper, 1995, Barth et al., 2012]. South of Haast, on the southern section of the fault, reverse motion decreases as compression is taken up by offshore thrust faults and the Fiordland subduction zone [Sutherland and Norris, 1995, Barth et al., 2013].

Seismic [Scherwath et al., 2003, Stern et al., 2007, Okaya et al., 2007, Wech et al., 2012, Chamberlain et al., 2014, Bourguignon et al., 2015], magnetotelluric [Wannamaker et al., 2002], geodetic [Lamb and Smith, 2013] and geological [Little et al., 2007, Herman et al., 2009] observations of the Alpine Fault indicate that the central part of the fault has a listric geometry, with its dip increasing from approximately  $60^\circ$  at the surface, to less than  $15^\circ$  at 30 km depth [Scherwath et al., 2003, Stern et al., 2007]. These observations do not rule out a second steeply dipping structure at depth as modelled

by Koons et al. [2003], who show that a mechanical model with strain partitioned onto two structures, one listric, dominantly reverse structure, and the other near-vertical with mostly strike-slip motion, could explain the strain histories recorded geologically. In the near-surface thrust segments of the central Alpine Fault often appear to have shallow dips which has been attributed to near-surface topographic effects [Norris and Cooper, 1995].

The central portion of the Alpine Fault is thought, on paleoseismological grounds to have last ruptured c. 1717 AD, in a  $M \sim 8$  earthquake [Wells et al., 1999, Sutherland et al., 2007], prior to written records in New Zealand. Paleoseismic work by Berryman et al. [2012] on the southern section of the Alpine Fault has generated a record of probable Alpine Fault ruptures from the last 8000 years, with these ruptures occurring regularly with a recurrence time of  $329 \pm 68$  years. The maximum likelihood probability of a large earthquake occurring in the next 50 years on the southern Alpine Fault has been calculated to be 27% [Biasi et al., 2015]. This implies that the Alpine Fault is late in its typical seismic cycle. The Alpine Fault therefore provides the opportunity to study the conditions of a continental boundary fault prior to a major rupture, and the opportunity to deploy observatories to record the processes leading up to a major earthquake.

### 6.2.1 Deep-Fault Drilling Project

Much of our current understanding of the Alpine Fault and the processes controlling its seismogenesis is derived from surface geological observations and inherently non-unique and often low resolution geophysical observations. Although these observations have provided valuable large-scale models describing the Alpine Fault and the processes occurring at depth, these models require real observations of the frictional and rheological properties, stress state and permeability within the fault zone to test their suitability and improve on them. To this end DFDP aims to drill into the Alpine Fault to determine the temperature, fluid pressure and chemistry, rock properties and the prevailing stress state through direct observations of the fault zone at depth [Townend et al., 2009].

The first phase of the project, DFDP-1, was completed in early 2011 at Gaunt Creek (Figure 6.1). Two shallow boreholes were drilled through the hanging wall of the Alpine Fault and into the footwall at depths of 128 m and 98 m [Toy et al., 2015]. The fault core was observed to be of low permeability, effectively forming a barrier to fluids across the fault [Sutherland et al., 2012, Carpenter et al., 2014] and to exhibit distinctive electrical properties related to the presence of phyllosilicate alteration products [Townend et al., 2013].

Examination of the fault rocks revealed that smectite clays only formed in the narrow band of fault gouge zones [Schleicher et al., 2015]. These smectitic clays contribute to lower coefficients of friction ( $\mu$ ) at low sliding velocities ( $\mu = 0.49\text{--}0.74$ , increasing with increasing P-T conditions) for the smectite bearing gouge relative to the chlorite/white mica-bearing gouge material ( $\mu = 0.61\text{--}0.76$ ) also found at DFDP-1 [Boulton et al., 2014]. Ikari et al. [2015b] also found high coefficients of friction for gouge from DFDP-1, with  $\mu$  for gouge at P-T conditions appropriate for seismogenic depths of around 0.67; high compared to those found at other active fault drilling projects, e.g.  $\mu \leq 0.4$  at the Nankai Trough [Saffer et al., 2012] and  $\mu = 0.1\text{--}0.15$  at Parkfield [Lockner et al., 2011].

A second stage of drilling, DFDP-2, was undertaken in 2014 in the Whataroa Valley (Figure 6.1) approximately 7 km east of the DFDP-1 drill-site. This phase of drilling was intended to characterize temperature, pressure, stress and structure of the hanging- and footwalls of the Alpine Fault, and to intersect the principle slip zone at a depth of  $\sim 1$  km. Due to the perceived risk of triggering earthquakes by drilling into the Alpine Fault, a formal safety review was undertaken in March 2014, involving an international panel of experts in earthquake seismology, fault mechanics, New Zealand geology, and drilling engineering. The panel was chaired by Dr. William Ellsworth (USGS), to assess the drilling plan. The aims of this safety review were to:

1. Determine if the drilling plan would have a significant effect on Alpine Fault seismic hazard;
2. Decide if the geophysical data to be collected, and responses to these data were appropriate;
3. Modify the operational procedures to reduce the impact of drilling on the hazard of the Alpine Fault if deemed necessary.

The safety review concluded that drilling as planned would have no significant effect on the seismic hazard of the Alpine Fault. Nevertheless, given the planned depth and scope of the drilling activities, and especially the unknown conditions that would be encountered below the principal slip zone, the science team implemented procedures for monitoring seismicity near the borehole throughout DFDP-2 operations and for responding to any observed seismicity in a prudent and effective manner. The response plan was set out using a traffic-light system (Figure 6.2, [e.g. Zoback, 2012, Walters et al., 2015]) that had been discussed at length during the review panel's meeting and subsequently revised based on available infrastructure, personnel, and drilling plans. The real-time seismic monitoring was planned to allow the DFDP-2 team to effectively and proactively communicate the locations of nearby seismicity with high precision to the general public, while also monitoring the seismicity around the borehole and providing operational advice based on this.

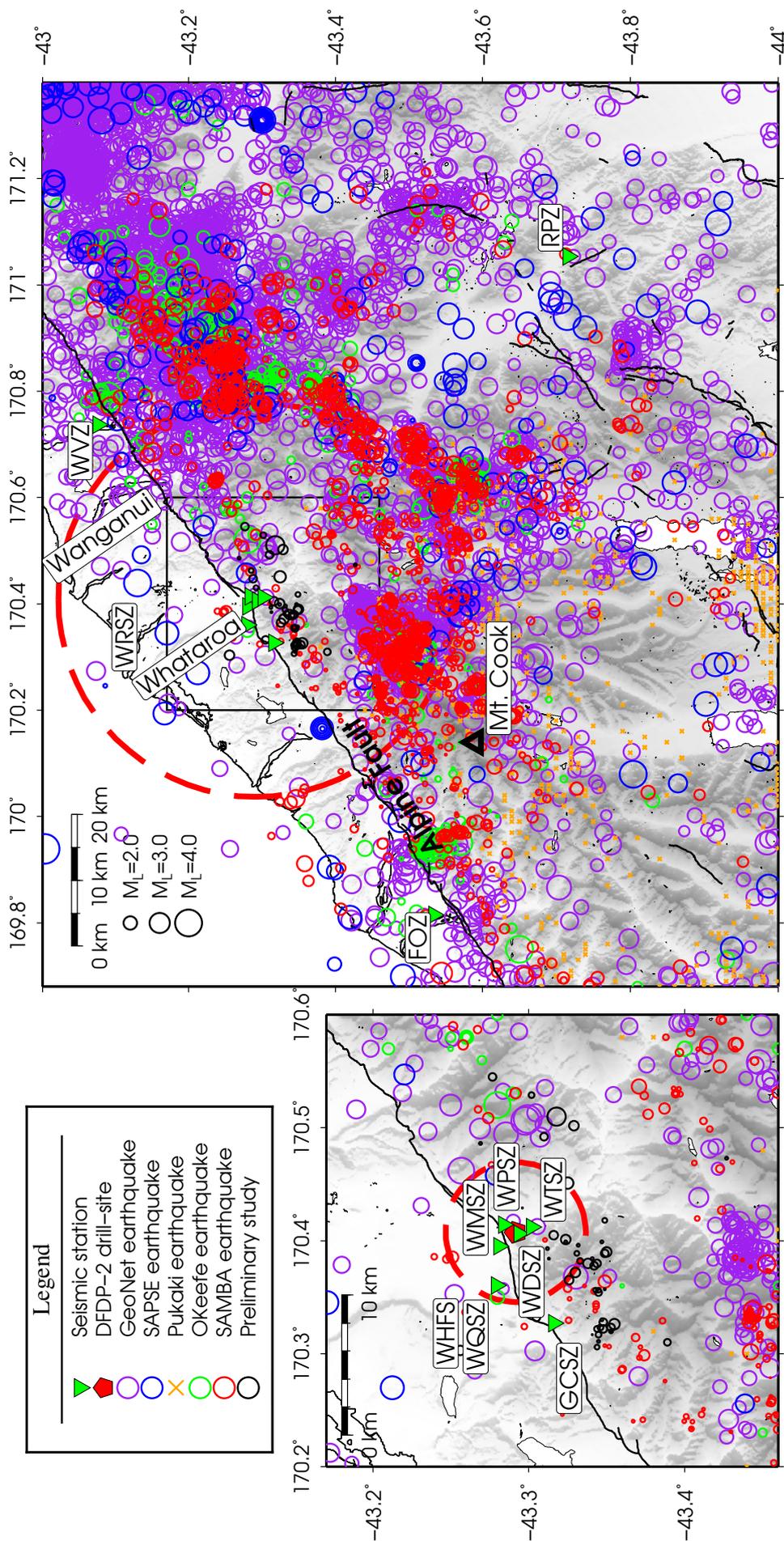


Figure 6.1: Background seismicity for the central Southern Alps region recorded by GeoNet (<http://info.geonet.org.nz/display/appdata/Earthquake+Catalogue>), Boese et al. [2012] (SAMBA), Reyners [1987] (Pukaki), O’Keefe [2008], Leitner et al. [2001] (SAPSE) and the preliminary study undertaken prior to drilling. This seismicity covers the period 1984–2013, none of the catalogues span the duration. Seismic stations used in this study are shown as green triangles Red dashed circle on the main map shows a 30 km radius around the drill-site. Inset: zoom around Whataroa with the DFDP-2 drill-site shown as a red pentagon, the DFDP-1 drill-sites are located at the site of the GCSZ seismic station. Dashed red circle is radius 3 km around the drill-site as described in the traffic-light scheme. There is little seismicity within 10 km of the drill-site (4 events over M 3 since 2004), with seismicity in the region bounded by the Wanganui and Whataroa river valleys concentrated in clusters around the edge of the region of low-seismicity.

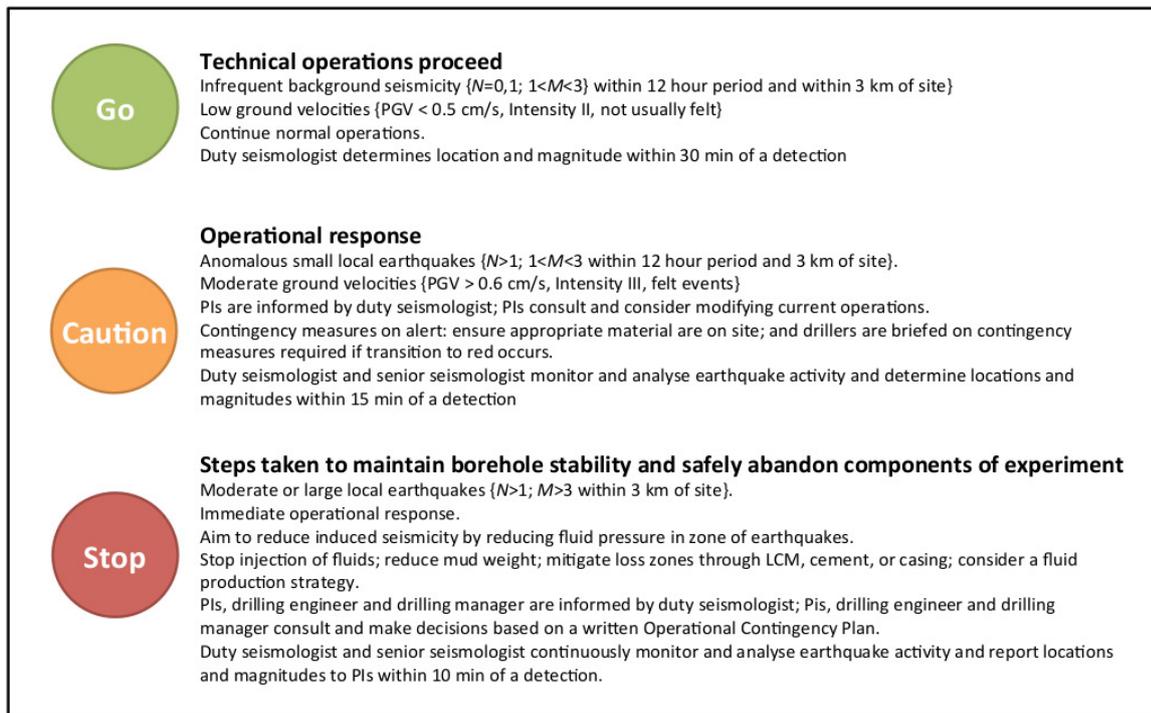


Figure 6.2: Traffic light scheme adopted during real-time seismic monitoring of the DFDP-2 drilling project.

### 6.2.2 Seismicity around Whataroa

Despite high deformation rates, the seismicity on and surrounding the central part of the Alpine Fault is sparse, with a notable seismic gap extending between the Whataroa and Wanganui Valleys as documented by Leitner et al. [2001], O’Keefe [2008], Boese et al. [2012] and Bourguignon et al. [2015] (Figure 6.1). Boese et al. [2012] presented the most recent detailed earthquake catalogue of 1791 earthquakes between November 2008 and December 2009 in the central Alpine Fault region, using data from the SAMBA, DFDP-10 and GeoNet networks. They found that earthquakes tend to occur in clusters along and to the east of the Main Divide Fault Zone, and are distributed throughout the crust rather than confined to mapped faults. They found seismogenic depths near the Alpine Fault trace of  $10 \pm 2 \text{ km}$ , shallowing to  $8 \pm 2 \text{ km}$  within 20 km of the Fault trace, before deepening to  $15 \pm 2 \text{ km}$ . Boese et al. [2012] related the geometry of the seismogenic zone to the shape of an underlying zone of low resistivity [Wanamaker et al., 2002], with highly clustered seismicity occurring in areas of strong resistivity gradients. These areas are interpreted to be associated with extensive fracturing allowing for fluid transport. In general, Boese et al. [2012] found little seismicity in the upper 2 km of the crust.

Work is currently underway to document seismicity patterns and velocity structures in the low seismicity region between the Whataroa and Wanganui rivers using the WIZ-

ARD seismic network [Thurber et al., 2012] which was operated from January 2012 to January 2014, and reoccupied many of the DFDP-10 sites. To the north of this region, Bourguignon et al. [2015] conducted a microseismicity and seismic tomography study between 2008 and 2010 using the ALFA08 and ALFA09 seismic networks. They again confirmed reduced seismicity between the Whataroa and Wanganui valleys. Eccles et al. [2015] showed that, despite this low seismicity, there is seismicity capable of generating fault-zone guided waves near or on the Alpine Fault near Gaunt Creek.

Ahead of DFDP-2 drilling we conducted a preliminary study using the borehole stations used during real-time monitoring, SAMBA sites and GeoNet sites to produce a catalogue of seismicity in the region around the DFDP-2 drill-site between January and October 2013 (Figure 6.1). This targeted analysis revealed low rates of seismicity and in general only low-magnitude ( $M_L < 3.0$ ) seismicity in the region. Further, this analysis allowed us to test triggering parameters and observe the distribution of background seismicity before any drilling operations had started, using a similar network. A cluster of earthquakes near the site of the DFDP-1 borehole c. 7 km west of the DFDP-2 drill-site was found to be active as part of the background seismicity near the drill-site. We discuss the results of this analysis in more detail in the following section.

### 6.3 Data and personnel

To obtain a representative catalogue of seismicity before and during drilling we installed four shallow (28 m) borehole seismometers in the Whataroa Valley to complement the SAMBA [Boese et al., 2012], WIZARD and GeoNet ([info.geonet.org.nz](http://info.geonet.org.nz)) networks. The installation of these four seismometers enabled the detection of small magnitude ( $M_L > -0.8$ ) earthquakes near the Whataroa Valley. These four sites, along with one SAMBA site (WQSZ, previously known as WHAT2) and one broadband surface seismometer (WRSZ) were telemetered in addition to the three GeoNet stations in the vicinity of the drillsite (Figure 6.1). The resulting real-time network is outlined in Table 6.1.

Data were telemetered via cellular 3G wireless modem to the GeoNet national data centre where they were then archived. Data were pulled from the GeoNet data centre via seedlink protocol in real-time using the *Rtquake* software [Utheim et al., 2014]. The *Rtquake* software package was used throughout the detection process to collect data, compute triggers, generate automated picks and send alerts to the duty seismologists. This software package enabled us to use *SEISAN* [Havskov and Ottemoller, 1999] for post-processing the detected events and for database management. The standard data flows and processing steps are outlined in Figure 6.3.

Table 6.1: Site locations for all seismic stations utilised in this study. Tabulated elevations are to the surface, with depths of borehole sites given in brackets

Station	Latitude (°)	Longitude (°)	Elevation (m)	Instrument	Install Date
WTSZ	-43.302	170.412	100 (-29 borehole)	IESE HS-1-LT 4.5 Hz	2013/02/03
WMSZ	-43.284	170.394	78 (-29 borehole)	IESE HS-1-LT 4.5 Hz	2013/02/03
WPSZ	-43.284	170.412	77 (-29 borehole)	IESE HS-1-LT 4.5 Hz	2013/02/03
WDSZ	-43.294	170.406	91 (-29 borehole)	IESE HS-1-LT 4.5 Hz	2013/02/03
WRSZ	-43.135	170.416	73	Guralp CMG-3ESP	2014/03/18
WQSZ	-43.279	170.360	96 (-61 borehole)	IESE HS-1-LT 2 Hz	2009/03/22
WHFS	-43.261	170.359	60	Kinematics BASALT	2014/06/16
GCSZ	-43.316	170.327	210 (-81 borehole)	IESE HS-2-shunted 2 Hz	2012/04/01
FOZ	-43.566	169.687	10	Guralp CMG-3ESP	2004/10/13
RPZ	-43.715	171.054	453	Guralp CMG-3TB	2001/06/06
WVZ	-43.074	170.736	91	Guralp CMG-3ESP	2008/02/15

To manually check detections and compute final hypocentre locations and magnitudes, a team of 16 seismologists worked both on- and off-site. Of these, nine seismologists covered the on-site monitoring in 2–4 week rotations. On-site, there was always at least one-of-five senior seismologists partnered with a junior seismologist or/and off-site seismologists. Prior to the drilling we trained six students to work with the system, in total our team included eight students. Our off-site seismologists were located in New Zealand, USA and Japan across six institutions. To train our seismologists and provide a reference manual we compiled documentation and reviewed this document as part of our training preparation to ensure that all members of the seismic monitoring team understood and complied with our observatory practices (Appendix E).

## 6.4 Methods

### 6.4.1 Earthquake detection

Detections were computed in near real-time using a simple STA/LTA filter routine within *Rtquake*. We used differing short-term and long term window lengths and triggering parameters for different stations, optimized for the individual station signal and noise conditions. Short-term windows ranged from 0.3 to 1.0 seconds, and long-term windows from 6.0 to 10.0 seconds. Sites closer to the drill-site had higher trigger thresholds ( $STA/LTA > 4.5$ ) determined due to high sensitivity to events local to the drill-site (and therefore of interest to this study), and high noise levels during drilling, whereas quieter sites had trigger-on thresholds of 3.0–3.5. Detections were raised when at least four sites were triggered. We did not incorporate data from sites WMSZ and WPSZ in this triggering routine to avoid detections becoming too reliant on the four sites closest to the drill-pad, and to avoid false detections associated with drilling noise.

Initial picks were made automatically using the *FilterPicker* routine [Lomax et al., 2012] built into *Rtquake*. The default parameters for the *FilterPicker* routine [Vassallo et al., 2012, their Table 3] (Table 6.2) worked well for P-phases, but we were unable to optimize the parameters for S-picks in time for drilling. As a result, S-picks were systematically downweighted in the automatic locations, resulting in reasonable epicentral locations, but large depth uncertainties.

Automatic locations were computed using *Rtquake*'s inbuilt automatic location program, which runs *SEISAN*'s modified version of *Hypocenter* [Lienert et al., 1986] iteratively. At each iteration picks with high time residuals are removed until either too few stations remain, or the RMS time residual is below a set threshold. For all locations we used a 1D velocity model (Table 6.3, based on that used by Boese et al. [2012], with an

Table 6.2: Parameters used for the *FilterPicker* routine during real-time monitoring. Parameters used are the suggested default values, where  $\Delta T$  is the sample interval, which was 0.01 s for all sites.

Parameter	Description	Value
$T_{filter}$	Longest period to consider picking	$300\Delta T = 3\text{ s}$
$T_{long}$	Time averaging scale	$500\Delta T = 5\text{ s}$
$S_1$	Trigger threshold for detection	10 s
$S_2$	Pick declaration threshold within $T_{up}$	10 s
$T_{up}$	Time window for pick validation	$20\Delta T = 0.2\text{ s}$

Table 6.3: 1D velocity model used in this study, depths are to to top of the velocity layer.

Depth (km)	Velocity (km/s)
-2.0	4.721
1.0	5.670
8.0	5.790
18.0	6.280
43.0	7.350
48.0	8.000

upper layer accounting for sediment velocities derived from averaging previous velocity models [O’Keefe, 2008, Feenstra et al., In Review]. Final checked locations were also generated using *Hypocenter*.

The automatic locations, combined with waveform plots were sent by email to a central mailing address accessed by the monitoring team. These emails provided detection alerts alongside useful information allowing on-duty seismologists to prioritize events. As these emails could be pushed to seismologists smartphones it allowed seismologists to continue other duties while maintaining response times.

### 6.4.2 Response to earthquakes

All events detected by *Rtquake* were checked by the on-site duty seismologist. This check had two tasks. The first of these was an initial check of the emailed information containing an initial location, pick-file in Nordic format, and a PNG image of the waveforms, with automatic picks in place, and hypocentre and magnitude estimation for relevant events. As we were operating a local network with a focus on earthquakes within a small region (main focus on earthquakes within 30 km of the drill-site), we did not conduct further checks on earthquakes with large ( $>3\text{ s}$ ) S–P times. Due to unreliable automatic S-picks, S–P times were checked by eye on the waveform images emailed.

If the on-duty seismologist deemed the earthquake detected to be of relevance to the

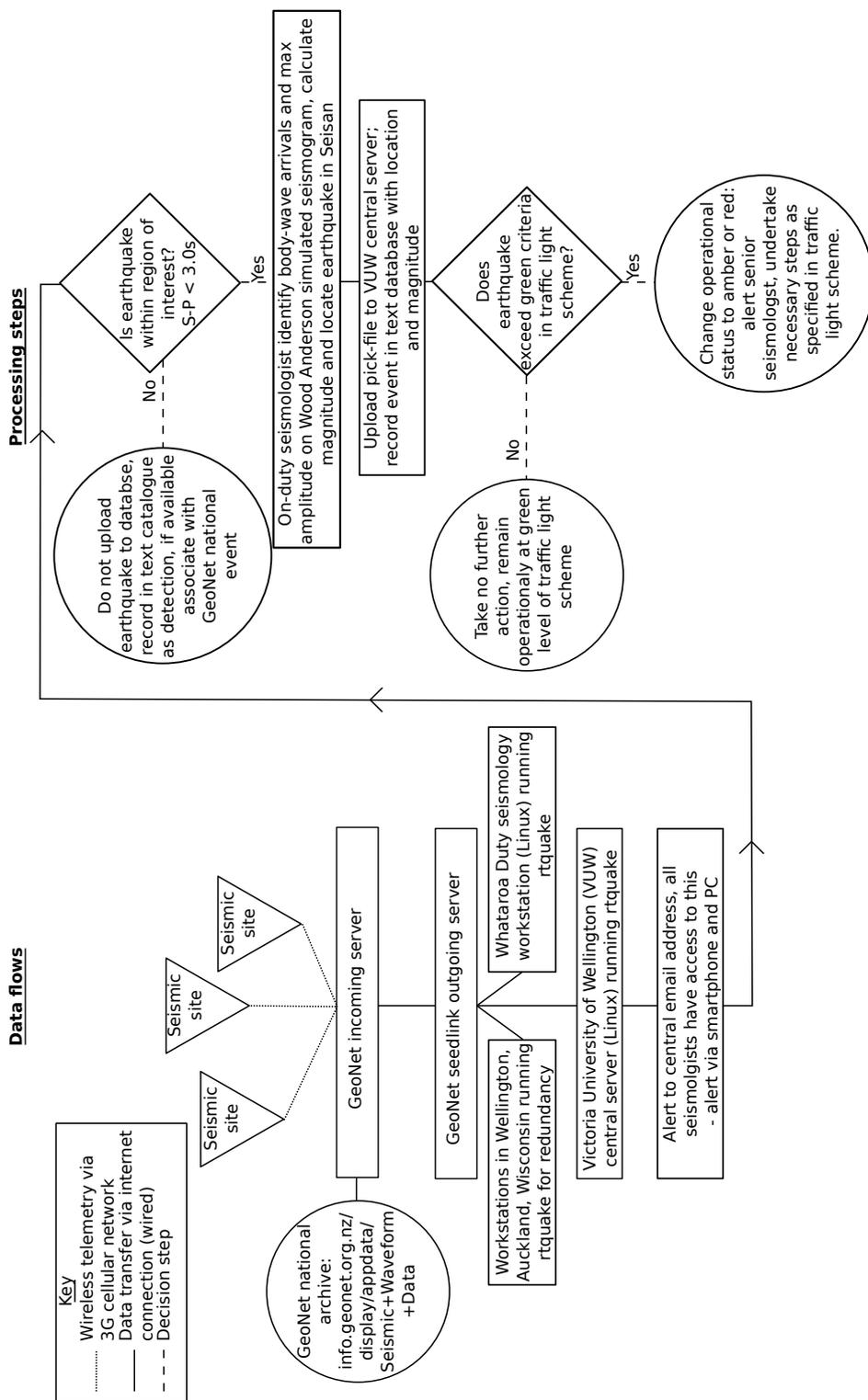


Figure 6.3: Flow diagram of the data and processing steps actively undertaken during monitoring. Data from the sites were telemetered in real-time and archived in the GeoNet national seismic database. Data were downloaded to multiple concurrent systems for redundancy via seedlink protocol. The central Victoria University server was used to store final locations and as the mailserv client to send alerts to the on-duty seismologists via a central email address which all users had access to.

drilling operations (small S–P time at stations close to the drill-site, <3 s, or phases first arriving on sites close to the drill-site), or of general interest to the duty seismologist, e.g. part of an earthquake sequence, or a subcrustal earthquake), then a second task was undertaken. Although in general our catalogue contains mostly earthquakes within our region of interest, some earthquakes outside of this region were picked based on subjective qualities, which affects the overall catalogue completeness.

The second stage of checking by the duty seismologist involved obtaining the waveform and pick-file data, either from the central server, or locally if the duty seismologist was running a local copy of the *Rtquake* software, then re-picking the event in *SEISAN*. The first step in the picking process was to correct S-picks to better constrain the depth. P-picks were often good, and needed little to no adjustment. The on-duty seismologist undertook normal observation practices [Havskov and Ottemoller, 2010] when picking the event before computing the hypocentral location and magnitude using *Hypocenter* in *SEISAN*.

This second stage (hypocentre and magnitude estimation) was usually conducted by the on-site duty seismologist, but during times of above-average seismicity (e.g. during the Malcolm Peak earthquake sequence, Section 6.5.1) we outsourced some of this re-picking to off-site seismologists in different time zones to enable our 24 hour monitoring operation to continue with little disruption. During normal background seismicity, this re-picking stage was completed within 15 minutes of detection. Earthquakes were then uploaded to the central server, and recorded in a simple text database containing time, magnitude and seismologist notes. This text record was available to all members of the drilling project.

### 6.4.3 Magnitude calculation

During the drilling preliminary local magnitudes ( $M_L$ ) were calculated automatically in *SEISAN* using an average station correction term. These magnitudes followed the relation:

$$M_L = \log_{10} A + \alpha \log_{10} \Delta + c \Delta + S \quad (6.1)$$

where  $A$  is half the measured peak-to-peak amplitude for a given site at distance  $\Delta$  from the hypocentre.  $\alpha$ ,  $c$  and  $S$  are constants, with  $\alpha$  the geometrical spreading factor,  $c$  the anelastic attenuation parameter, and  $S$  the station correction term. Initial magnitudes were calculated using  $\alpha=1.0$ ,  $c=0.0041$  and a station correction term of  $-1.68$ . This provided a good estimate of magnitude without taking into account the frequency dependent anelastic attenuation and station-specific site correction terms inverted for by Boese et al. [2012].

After drilling, re-calculated magnitude estimates for events recorded by at-least three

sites within 100 km of the epicentre site were made. In this case we use the site dependent site-correction terms derived by comparison with earthquakes also recorded by GeoNet. We use the frequency dependent anelastic attenuation parameter for the central Southern Alps found by Boese et al. [2012], giving the relation below:

$$M_L = \log_{10} A + \log_{10} \Delta + 0.00189f\Delta + S \quad (6.2)$$

where  $c$  has been replaced by the frequency dependent term and  $f$  as the frequency in Hz.

#### 6.4.4 Planned response to large or close earthquakes

As part of the seismic monitoring, and in accordance with the safety review, a response plan was put in place for earthquakes deemed to necessitate changes to drilling operations. This planned response employed a traffic light scheme tailored to the range of background seismicity encountered prior to drilling. The traffic light scheme is outlined in Figure 6.2 and was developed for the Whataroa region taking into account the previously recorded background seismicity near the drill-site (Figure 6.1).

The originally proposed traffic light scheme was based solely on number of events within a certain magnitude range. An earthquake of  $M$  2.9 occurred one month prior to drilling and real-time monitoring, at c. 8.5 km epicentral distance from the drill-site. This event provided a test for our traffic light scheme due to its larger than average magnitude in this region and relative proximity to the drill site. Only 4 events of comparable or larger magnitude have occurred since installation of the three permanent GeoNet sites in the low seismicity area surrounding the drill site. This event was considered part of background seismicity, and of little risk to the drilling operations (having a peak ground velocity (PGV) value  $\leq 0.01$  cm/s at the drill-site). However, it showed that additional traffic light criteria were needed, that specify thresholds on distance ranges, event rates and PGV values for nearby earthquakes. We implemented criteria based on felt intensities and peak ground motions from traffic light schemes developed in response to induced seismicity (for examples see Fig. 6 of Van Eck et al. [2006], Fig. 10 of Majer et al. [2007], and Fig. 8 of Grünthal [2014]).

## 6.5 Results

During the active monitoring we detected 1,121 events, of which we manually picked and located 493 events of interest. A further event was later picked as part of the Malcolm Peak earthquake sequence, approximately 30 km east of the drill-site: this

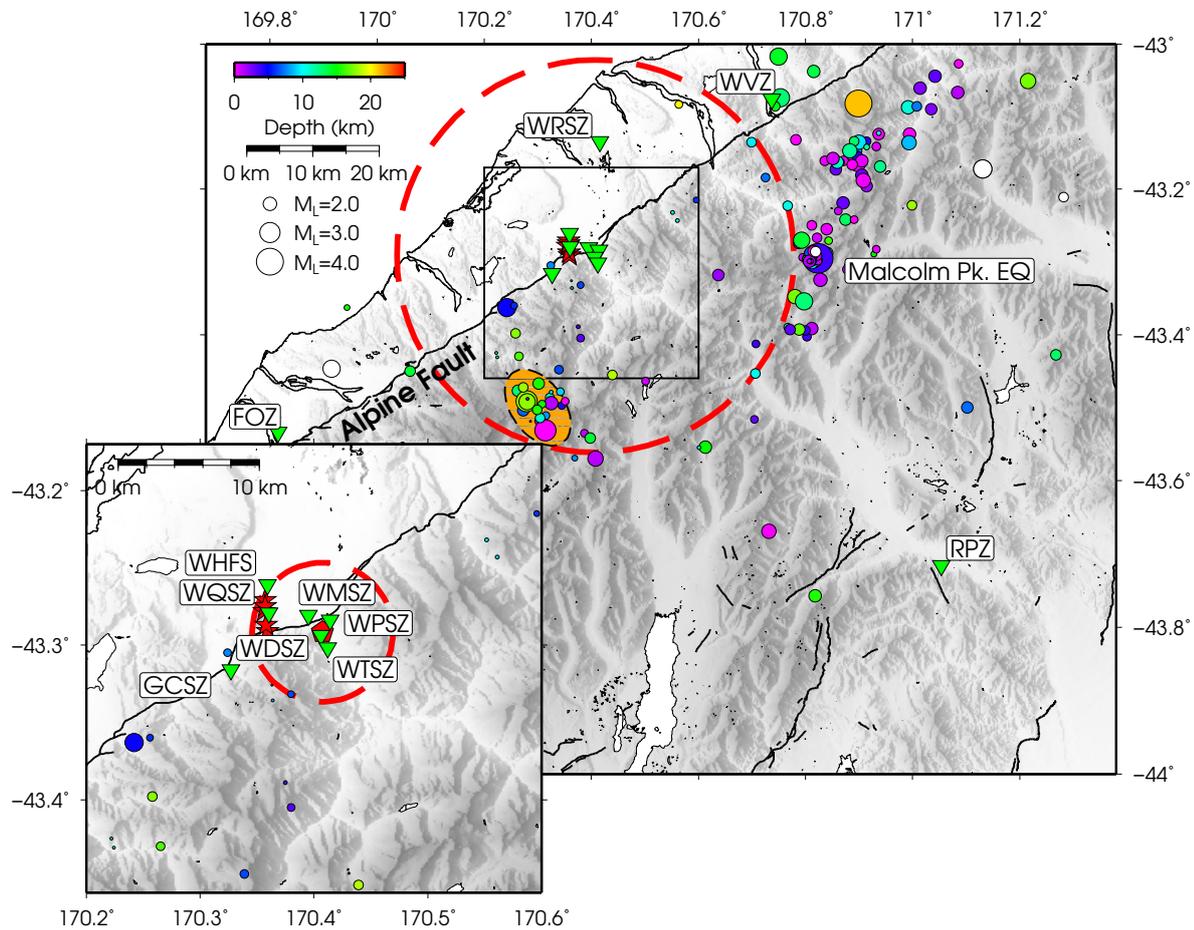


Figure 6.4: Earthquakes detected and located during real-time seismic monitoring. Total events located number 494. Red dashed circle on the main map shows a 30 km radius around the drill-site. The orange ellipse marks the southern cluster described in the text. Earthquakes plotted without colour are deeper than 25 km. Inset: enlarged view around drill-site with drill-site shown as a red pentagon. Dashed red circle is radius 3 km around the drill-site as described in the traffic-light scheme. Red stars denote explosions at the nearby Whataroa Quarry. Green triangles mark seismic sites used in this study. Magnitudes plotted are those derived during the real-time monitoring, without frequency dependent attenuation or site-specific station correction terms. These magnitudes are systematically over-estimated.

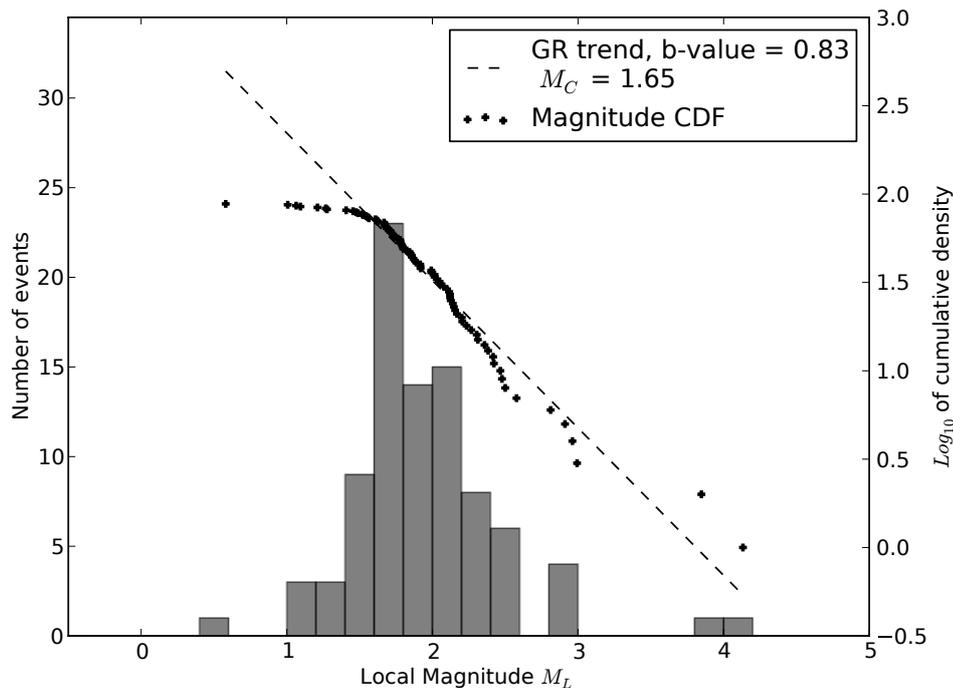


Figure 6.5: Histogram of magnitude distribution observed during real-time monitoring for earthquakes within 40 km of the drill-site ( $n=88$ ). Magnitudes plotted are recalculated after drilling to correspond to the magnitude scale of Boese et al. [2012], using a frequency dependent attenuation model and individually determined station correction terms. The GR fit for this dataset is not robust due to a lack of variation in magnitudes, however this small range appears to fit a b-value of 0.83 with an  $M_C$  of 1.65.

event had been missed during real-time monitoring. Of these 494 picked events we generated local magnitudes for 140 events which had at least three seismic stations recording the event within 100 km. We report magnitudes of between  $M_L$  0.6 and 4.2, the frequency distribution of the magnitudes of earthquakes within 40 km ( $n=88$ ) appears to follow a Gutenberg-Richter law with a high magnitude of completeness ( $M_C=1.65$ ) and b-value of 0.83 (Figure 6.5). However this fit is derived over less than a unit of magnitude, so is likely unreliable. The high  $M_C$  is likely due to a real lack of seismicity within 30 km of the drill-site during the monitoring period combined with a skew towards the detection of larger events and incomplete reporting further from the network (most of the events considered in the  $M_C$  calculation occur between 30–40 km from the drill-site). Thus the reported  $M_C$  is unlikely to reflect the networks ability to detect small magnitude earthquakes close to the drill site (see Section 6.5.2). The full catalogue and magnitudes is included in the electronic appendices.

The closest earthquake to the drill-site, with  $M_L$  1.4 ( $M_L$  before recalculation=1.8), occurred on the 3rd of September 2014, approximately 5.3 km south-southwest of the drill-site at a depth of 6.4 km. Three further events occurred at similar distance to the drill-site, with magnitudes of  $M_L$  0.5, 0.7 and 1.4. All of these earthquakes were outside of the actionable range defined in the traffic light scheme (Figure 6.2). Therefore

throughout the drilling period no actions were taken due to seismicity outside of the standard operational plan.

During monitoring we also recorded five local quarry blasts near site WQSZ (Whataroa Quarry). The locations obtained for these blasts are plotted as stars in Figure 6.4. Locations for these events are automatically fixed at 0 km (rather than the surface at  $-0.096$  km in our model) depth and have larger uncertainties due to high travel-time residuals at nearby stations. These high residuals are the result of poorly constrained shallow velocities in our velocity model, especially around the quarry where footwall granite outcrops rather than the slower fill beneath most of our sites.

We detected more earthquakes within 30 km of the drill-site than GeoNet detected, with our catalogue including all the GeoNet events up to a radius of 40 km from the drill-site after 05/09/2014 (when all sites went online, prior to this date only WTSZ was telemetered in addition to GeoNet sites Figure 6.6). Of those events within 30 km of the drill-site, GeoNet were able to detect and locate three events, occurring in December 2014. We were able to detect and locate 19 events within 30 km of the drill-site.

During real-time monitoring we detected a subcrustal earthquake on 11/10/2014, at a depth of 64 km and  $M_L$  2.6. This earthquake locates on the north-east side of the surface trace of the Alpine Fault, co-located with a cluster of four subcrustal depth (55–65 km) earthquakes documented by Boese et al. [2013].

### 6.5.1 Malcolm Peak earthquake 22/10/2014

During the active-monitoring period we detected an earthquake sequence of 26 events between 22/10/2014 04:17–23/10/2014 21:12 UTC. With magnitudes ranging from  $M_L$  1.0 to 4.2. Two earthquakes of  $M_L$  1.9 occurred 13 and 6 minutes before the largest ( $M_L$  4.2) earthquake in the sequence. Immediately (20 seconds) prior to the mainshock, a  $M_L$  3.8 earthquake in approximately the same location as the  $M_L$  4.2 mainshock occurred. This event was missed during real-time monitoring due to its proximity in time to the larger mainshock, and as such is not included in the relocations (Figure 6.7). Due to our incomplete catalogue we cannot determine whether this sequence is a swarm, or a mainshock-aftershock sequence with foreshocks. The sequence took place approximately 30 km due East of the drill-site and produced a PGV of  $< 0.2$  cm s $^{-1}$  and PGA of 0.13% g at the strong motion site WHFS.

The region around the here-named Malcolm Peak sequence was one of the most seismically active regions studied during the real-time monitoring. This cluster of earthquakes locate  $\sim 20$  km epicentrally east-northeast from the Alpine Fault surface trace and likely occurs on faults associated with the main divide fault-zone. Precise double-difference relocations computed on-site using *TomoDD* [Zhang and Thurber, 2006]

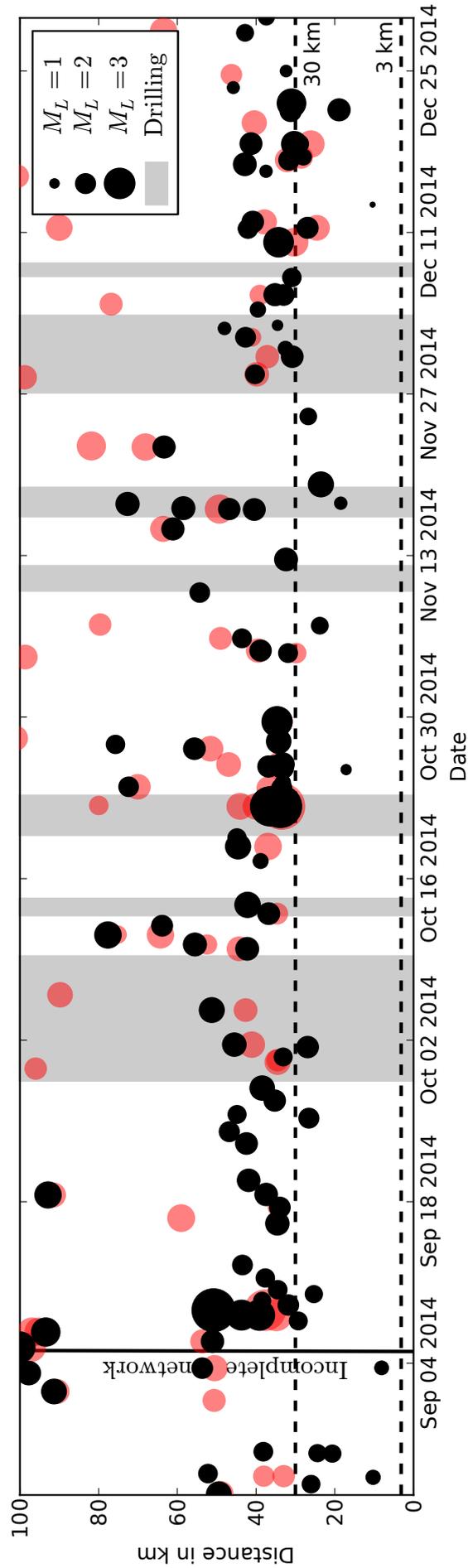


Figure 6.6: Time series plot of detected and located earthquakes, sized according to magnitude, and plotted according to distance from the drill-site. Black circles are earthquakes detected and located as part of this project, red circles are catalogued events from GeoNet ([info.geonet.org.nz](http://info.geonet.org.nz)). Prior to 05/09/2014 we were operating an incomplete network, with only site WTSZ telemetered in addition to the GeoNet sites. Dashed lines at 30 km and 3 km correspond to regions of interest plotted on Figure 6.4.

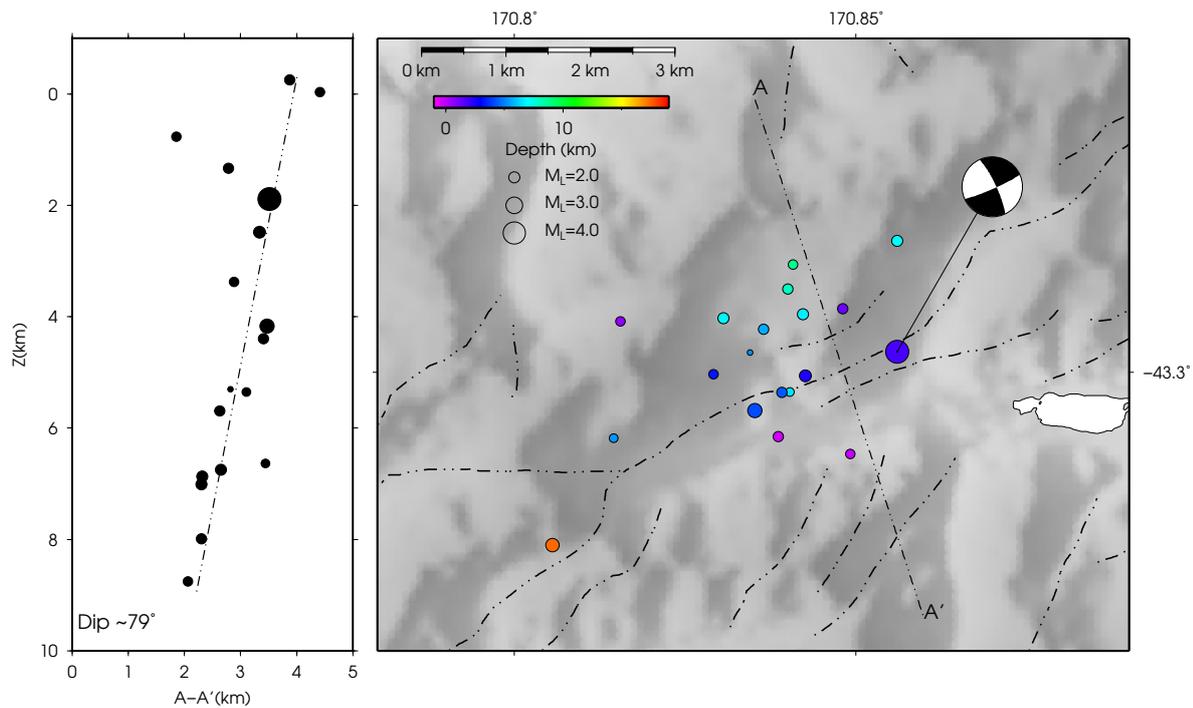


Figure 6.7: Malcolm Peak earthquakes relocated using the velocity model of Feenstra et al. [In Review]. Main map includes a lower hemisphere projection of focal mechanism obtained from GeoNets moment tensor inversion of the largest event in the sequence (<http://info.geonet.org.nz/display/appdata/Earthquake+Catalogue>). Cross-section runs from A–A', a line perpendicular to the strike ( $246^\circ$ ) of one of the nodal planes of the focal mechanism. Fitting a line to these aftershocks gives an approximate dip of  $79^\circ$  NW, similar to the  $84^\circ$  NW obtained from the moment tensor solution ( $66^\circ$  strike,  $-170^\circ$  rake). These plots do not include the earthquake 20s before the largest earthquake in sequence as this was not relocated.

indicate a steeply dipping ( $\sim 80^\circ$ ) structure striking  $246^\circ$  (Figure 6.7), with antithetic dip to the Alpine Fault. This is likely a back-thrust within the main-divide Fault Zone [Cox and Findlay, 1995]. The focal mechanism obtained from moment tensor inversion [Ristau, 2008] of the largest event in the sequence is that of a steeply-dipping dextral strike-slip event striking sub-parallel to the plate convergence (obtained from <http://info.geonet.org.nz/display/appdata/Earthquake+Catalogue>, last accessed 19 October 2015), consistent with the hypocentre relocations.

### 6.5.2 Drilling noise detected seismically

On 24 October 2014, a winch cable on the drill rig failed as the bottom hole assembly (BHA) was being assembled in preparation for renewed drilling after a wireline logging run. This resulted in the BHA falling approximately 70 m to the bottom of the hole at this time. This impact was recorded on the nearest stations (WTSZ, WMSZ, WPSZ and WDSZ, located 1.2, 1.6, 1.0 and 0.2 km from the drill-site respectively) with dominant frequencies in the 2–10 Hz range (Figure 6.8). By picking similar re-

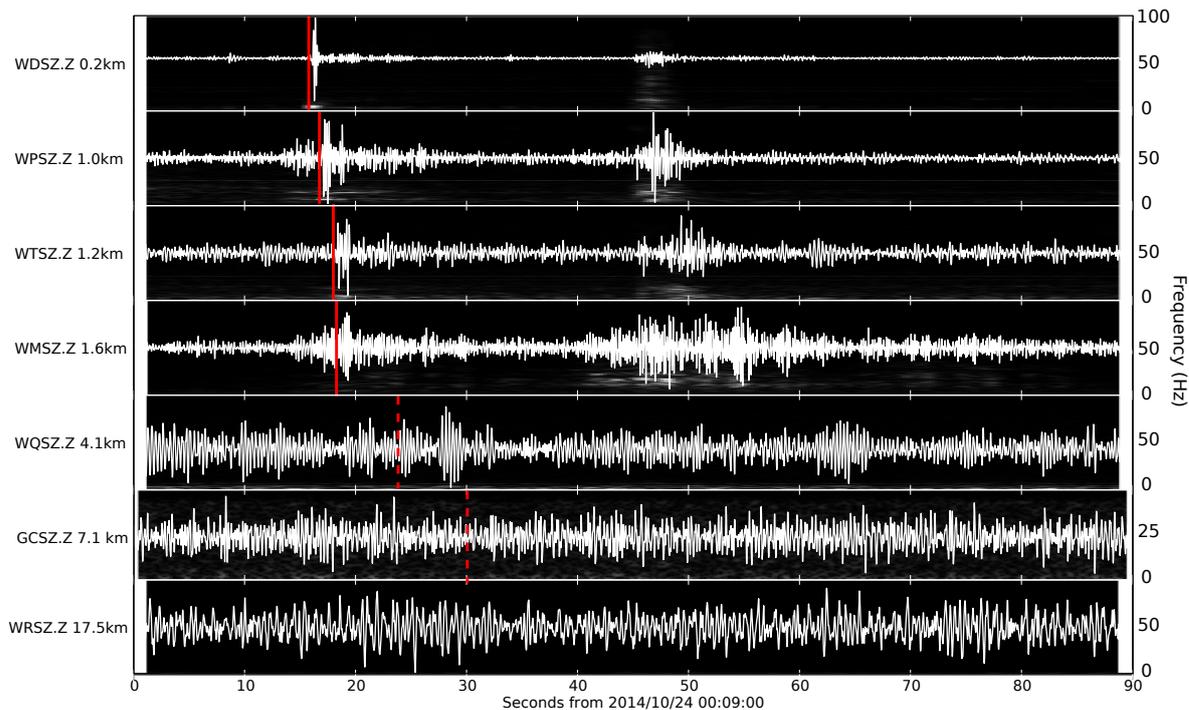


Figure 6.8: Waveforms (unfiltered) and spectra showing the seismic signal resulting from dropping the  $\sim 8$  T bottom hole assembly  $\sim 70$  m to the bottom of the hole. There is a clear peak in the seismic amplitudes on the four closest sites (marked in red), which indicate a moveout of  $\sim 2.1 \text{ km s}^{-1}$ . The dashed red lines on WQSZ and GCSZ indicate when the seismic energy is expected to arrive at these sites given this moveout. Seismograms are plotted overlaying the spectrograms for that site. Note the difference in frequency scale for the 100 Hz sampled GCSZ. On nearby sites a subtle peak in power can be seen in the 2–10 Hz frequency band, which cannot be seen on more distant sites.

gions on the waveforms (there are no distinctive phase arrivals) we obtain a moveout of  $\sim 2.0 \text{ km s}^{-1}$ . We are only able to see this signal on the stations closest to the borehole, with no discernible signal observed on the next nearest site, WQSZ (4.1 km).

Using the dry weight of the BHA prior to cable failure (8560 kg), we obtain a gravitational potential energy of 587 kJ, which, if all this energy was released in this impact with the base of the borehole, would equate to a moment magnitude ( $M_W$ ) of  $-1.5$ . This is a maximum bound on the moment magnitude, as drag forces of a piston-like BHA entering a partially fluid-filled hole of a similar diameter to the BHA would be significant. Also some energy was likely lost as the BHA bounced on the walls of the hole during descent; these bounces appear to be manifest seismically as smaller arrivals prior to the main arrival on Figure 6.8. Notably, we did not detect this signal with our real-time detection scheme as we required STA/LTA triggers on four sites, exclusive of WMSZ and WPSZ. Therefore this BHA drop shows that we would be unable to detect very small events ( $M_W < -1.5$ ) using this triggering arrangement. However, if triggers from drilling noise could be excluded in another way, the network is capable of recording events to this level close to the drill-site.

### 6.5.3 Overall performance

Throughout the monitoring period the network performed well, allowing us to systematically and efficiently check and relocate earthquakes of a range of magnitudes. We suffered two telemetry failures during monitoring at site WRSZ, where birds had damaged cables attaching the modem and antenna. We were able to fix these issues quickly (within a week), and did not suffer a noticeable deterioration in detections due to the loss of a single site. We also suffered a modem failure at WQSZ for two days due to high rainfall flooding the modem.

During the monitoring we suffered a server failure at the main Victoria University of Wellington server used for *Rtquake* and database storage. This was due to a failed concurrent real-time system designed to detect earthquakes by cross-correlation which wrote a large (>2 TB) log-file of errors. During the down-time of the server (20 hours) we reverted to using the on-site workstation as our main *Rtquake* system and mailserve machine. During this period we also stored detections and picked waveforms on this system and in a shared *Dropbox* folder before uploading them to the server once this was back online.

During the monitoring we relied heavily on reliable internet connections, both to the servers and for the 3G cellular telemetry of the seismic sites. We had a plan and Python scripts in place to enable continued monitoring in the event of a failure of the 3G cellular network. This involved the on-site seismologist driving to key sites every two hours and downloading data before running these data through automatic detection scripts and the *FilterPicker* routine. The key sites determined were WQSZ and WMSZ, chosen for their proximity to the drill-site (therefore good signal to noise for events close to the drill-site) and ease of access. These stations would be analysed alongside the GeoNet sites FOZ, WVZ and RPZ which have back-up radio-link telemetry. We did not operate this system at any point during our monitoring, but did undertake tests of the codes during drilling.

After the drilling, security issues were raised with our method of external access to the Victoria University network. These security concerns have resulted in the closure of this access method. To allow us to operate in a similar fashion in future we would either need to use a more secure method of access, or a different, external server.

## 6.6 Conclusions

Complications during drilling led to the premature cessation of drilling at a depth of 893 m, prior to intersecting the principal slip zone. Nevertheless, throughout the drilling

process, and for two months before and one month after, we successfully operated a low-cost system of real-time seismic monitoring, focused on the local area around the drill-site. We detected few earthquakes ( $n=4$ ) within 10 km of the drill-site, with three of these earthquakes located c. 5 km away near the site of the DFDP-1 borehole at Gaunt Creek.

Prior to drilling we undertook testing to determine what level of seismicity we could detect with our real-time network, and used this to tailor our traffic-light scheme to the background seismicity. During real-time monitoring we observed no obvious deviations from the background seismicity either during periods of drilling, or when we were not drilling, nor did we detect any seismicity within 3 km of the drill-site. This was as expected since drilling did not reach the principal slip-zone, nor should drilling as planned have induced seismicity at any point. No seismicity occurred during drilling that required a change to operational procedures, with all events meeting the parameters of the green/go traffic-light level.

We documented sparse seismicity in the previously noted seismic gap between Whataroa and Wanganui valleys, but found strongly clustered seismicity on the margins of this region, highlighted by the Malcolm Peak earthquake sequence. We also recorded one subcrustal earthquake of  $M_L$  2.9 at a depth of 64 km within 7 km epicentrally of the Alpine Fault surface trace. This intermediate depth earthquake occurred in a similar location to a cluster of four events at intermediate depths (55–65 km) documented by Boese et al. [2013].

### 6.6.1 Improvements

Overall, the DFDP-2 seismic monitoring program highlights the feasibility of near-real-time earthquake surveillance in conjunction with drilling or other scientific activities. Using cost-effective technical components and open-source software, a team working in seven locations in three time zones was able to deliver continuous monitoring coverage for more than three months.

Nevertheless, several improvements could be made to the real-time monitoring in case of continued drilling operations in the Whataroa Valley. These include tuning of the parameters used for: STA/LTA detection, automatic picking and magnitude determination. With the acquired and manually checked dataset, we can now train the *FilterPicker* routine for reliable S-picks following the method of Vassallo et al. [2012]. This should allow for better initial hypocentre locations, especially regarding depths, reducing the manual labour of the duty seismologist. We would still require the duty seismologist to check and confirm locations and magnitudes at all times, especially for those events close to the drill-site.

The real-time magnitude calculation was based on station correction terms derived after 1.5 months of real-time monitoring and were overestimating event size compared with GeoNet magnitude estimates. Given the amount of intermediate size events that were recorded simultaneously by the real-time monitoring and the GeoNet network, and can be used for magnitude calibration, we can now tune magnitude station correction terms.

In preparation for higher event numbers in the direct vicinity of the drill site, as may be the case when intersecting the principal slip zone of the Alpine Fault, the traffic light system needs clearly specified conditions that allow the return to a lower level once precautionary measures have been taken.

Lastly, the number of on-site duty seismologists needs to be increased for drilling periods exceeding four months. The majority of New-Zealand based seismologists were on site for 14–42 days. Therefore, longer monitoring periods would require international collaborators to be involved in on-site operations.

# 7 Discussion

This thesis is underpinned by the development and analysis of multiple earthquake catalogues derived from continuous seismic data in New Zealand’s central Southern Alps. These catalogues represent the seismic manifestation of deformation occurring on and near to the Alpine Fault, one of the world’s major strike-slip plate boundary faults and thought to have a high probability of rupture in a M 7–8 earthquake in the next 50 years [Biasi et al., 2015].

## 7.1 Summary

We have compiled and examined four complementary earthquake catalogues, each with specific characteristics. In Chapter 3 we described the first low-frequency earthquake (LFE) catalogue for the central Southern Alps. In Chapter 4 we developed a new, objective, method of LFE template construction based on synthetic templates, leading to a more spatially and temporally continuous catalogue of LFE templates, which can be used to build a detailed, continuous catalogue of LFEs on the Alpine Fault. The methods employed and expertise gained in analysing LFEs have also been used to compile a detailed catalogue of shallow microseismicity adjacent to the Alpine Fault (Chapter 5). This catalogue highlights some of the limitations of the matched-filter methodology, while demonstrating a lack of classical repeating earthquakes on or near the Alpine Fault beneath Whataroa. We also conducted real-time monitoring of seismicity in the vicinity of the DFDP-2 drill-site during the second stage of the Deep Fault Drilling Project (Chapter 6), and we provide detailed documentation of the procedures and practices in this thesis and in the appendices which will be of use for further drilling operations.

The main findings and contributions of this thesis are as follows:

Chapter 3:

- Low-Frequency Earthquakes (LFEs) occur on and/or near the deep extent of New Zealand’s Alpine Fault;
- These LFEs respond to stresses imposed by large ( $M_W > 5.7$  observed) regional

earthquakes;

- The magnitudes of catalogued LFEs scale according to an exponential law, suggesting an underlying scale-bound process: in other words, the source has a characteristic length-scale, possibly corresponding to the length of asperities that fail as LFEs.

Chapter 4:

- Earthquakes and LFEs can be detected using simple synthetic templates with a network-based matched-filter detection methodology;
- To efficiently scan through 6.5 years of continuous seismic data, we have developed a multi-paralleled, scalable, portable and open-source Python package for the detection and analysis of repeating and near-repeating seismicity which includes this methodology;
- Using this method, we have extended the catalogue of LFE templates to include 600 possible templates;
- Events in this catalogue are poorly constrained in depth due to a lack of P-phases in the stacked waveforms (resulting from high noise), but when P-picks are available the LFEs' hypocentres are close to those of the much smaller set of events described in Chapter 3.

Chapter 5:

- We have used the matched-filter network-based cross-correlation technique to detect non-repeating, shallow microseismicity within a seismically quiet region adjacent to the Alpine Fault near Whataroa;
- The earthquakes detected exhibit clustering into temporally discrete sequences;
- In the study region, where the Alpine Fault is currently locked near the surface and has not generated moderate magnitude earthquakes during the SAMBA deployment, earthquakes cluster on small-scale asperities;
- Asperities near the Alpine Fault in Whataroa generate small earthquake sequences that are then not re-activated within 1 year.

Chapter 6:

- We applied open-source real-time earthquake detection software, operated by a multi-national team of scientists, for around-the-clock monitoring of seismicity during an active fault drilling project (DFDP-2);
- As previously documented, seismicity around the Whataroa Valley is found to be scarce (four earthquakes within 10 km of the DFDP-2 drill-site, of magnitudes

( $M_L$ ) between 0.5 and 1.4);

- As anticipated following the pre-drilling safety review, DFDP-2 drilling did not produce detectable changes in seismicity.

Below we expand and discuss some of the implications of these points, separated into low-frequency earthquakes and slow deformation (Chapters 3 and 4), and micro-earthquake studies (Chapters 5 and 6). These results would not have been possible without the extensive data from the seismic networks maintained and extended as part of this PhD work. These networks (SAMBA, DFDP-2 boreholes and WIZARD) are all deployed in challenging Alpine environments, often requiring helicopter access. Maintenance of these networks has been a large component of this PhD study, to ensure they continue to provide high quality data which yield such interesting results.

### 7.1.1 Low-frequency earthquakes and slow deformation

Our initial LFE catalogue (Chapter 3) extended the work of Wech et al. [2012] who documented sparse tremor beneath the central Southern Alps. By employing a network-based matched-filter methodology with manually determined templates, we were able to construct a 3 year-long continuous catalogue encompassing 14 LFE families, which we infer to be located on or near to the deep extent of the Alpine Fault. By extracting LFEs with discrete seismic phases we were able to reduce the depth uncertainty compared to previous tremor locations, with the resulting LFE locations clustering within a zone of high attenuation, in the vicinity of other structures imaged and modelled to be the deep extent of the Alpine Fault (Figure 7.1). This LFE catalogue is the first documentation of such events in the central Southern Alps.

Our continuous record of LFE detections show a quasi-continuous rate of seismicity, with increased rates associated with previously reported tremor and following large regional earthquakes ( $M_W > 5.7$ ). Increased LFE detection rates are likely to be coincident with increased deformation rates, which have implications for the state of stress on the deep extent of the Alpine Fault. If slip on the deep extent of the fault can be triggered by small stress changes, then we would expect that the fault is in a state near to failure, either as a weak fault, or critically stressed [cf. Beeler et al., 2013].

If slip on the deep extent of the Alpine Fault directly loads the seismogenic portion, then we might expect to see an increase in LFE generation prior to a large Alpine Fault earthquake as observed prior to large subduction earthquakes [Kato et al., 2012, Ito et al., 2015, Ruiz et al., 2014] (see Section 1.5 for further discussion). If this is the case, then studying how LFEs are generated, and how their rate, location and magnitude change in time through active monitoring may provide insight into how the seismogenic zone of the Alpine Fault is loaded.

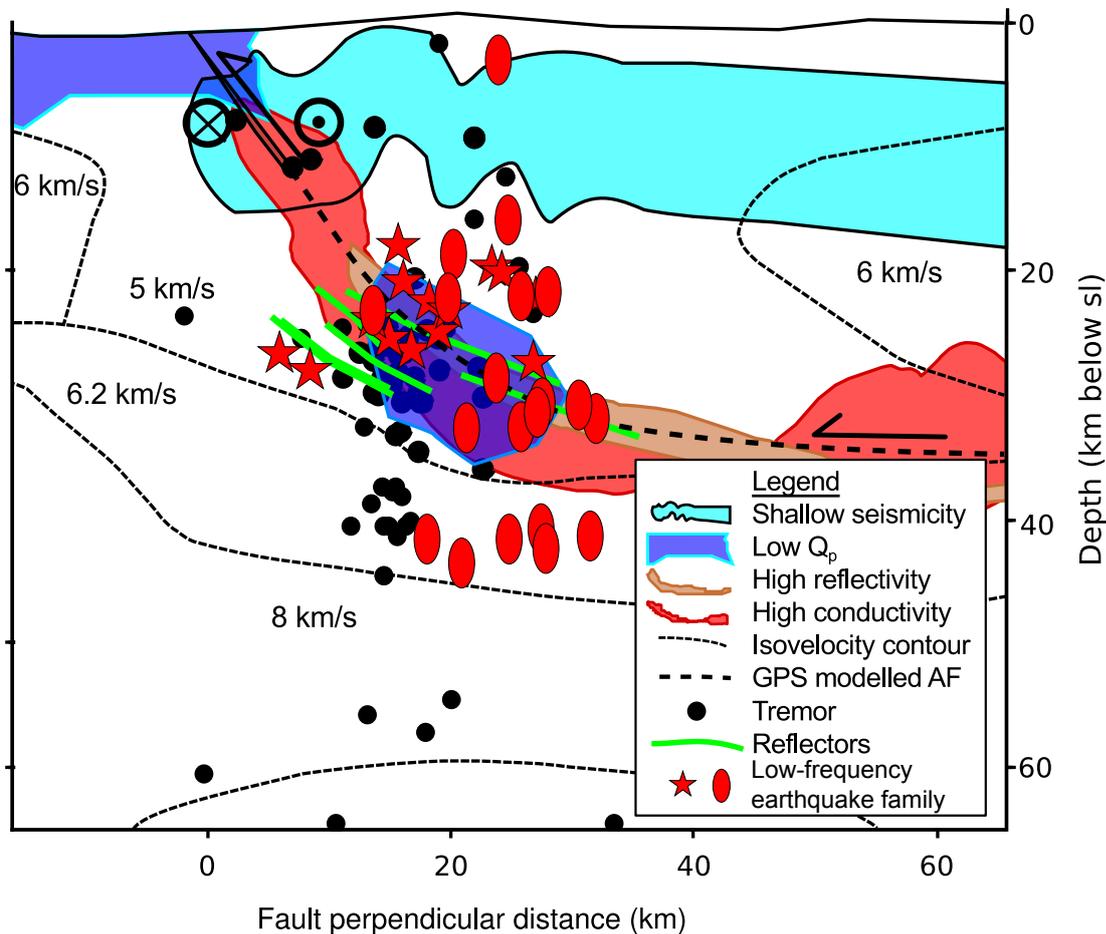


Figure 7.1: Low-frequency earthquakes documented here are associated with a number of other indicators of the Alpine Fault at depth. We interpret the low-frequency earthquakes to be associated with slow-slip on the deep extent of the Alpine Fault, which is complemented by the GPS model of Lamb and Smith [2013] who inferred the presence of a slipping region adjacent to the low-frequency earthquakes. Two different groups of LFE family are plotted here: those marked by stars are from Chapter 3, and those marked by ovals are the templates with P-picks from Chapter 4. As plotted, the region of shallow seismicity is taken from the locations of Boese et al. [2012], the tremor locations are from Wech et al. [2012], zone of high reflectivity, reflectors and velocity model from Stern et al. [2007], zone of high conductivity from Wannamaker et al. [2002], regions of low  $Q_P$  from Eberhart-Phillips et al. [2008], GPS fault from Lamb and Smith [2013]. The listric fault foundering at 30–40 km depth is imaged well by the active source seismic studies [Stern et al., 2007].

---

In Chapter 3 we also provided the first catalogue magnitudes for LFEs. LFE local magnitudes follow an exponential magnitude-frequency relationship which we interpret to indicate an intrinsic scale-bound process in the generation of LFEs in the central Southern Alps. Two other groups have since produced LFE catalogue magnitudes. Sweet et al. [2014] found a similar exponential relationship, using a methodology similar to that employed here. In contrast, Bostock et al. [2015] observed a power-law relationship using a different methodology where they directly computed moment magnitudes. Bostock et al. [2015] nevertheless interpreted the high  $b$ -values obtained, and the invariant signal-duration, to be indicative of a scale-bound process. It seems, therefore, that moment magnitudes and local magnitudes, despite giving different results, both show that LFE generation is scale-bound. This scale-bound process may allow for simple conversion between magnitude and slip, which could allow for the quantification of slip rates on the deep extent of faults: if asperity rupture size is fixed in some way, then magnitudes recorded seismically could be directly related to slip length.

To provide a thorough quantification of slip on the deep extent of the Alpine Fault requires a more objective and complete catalogue of LFEs. We developed a new method of LFE template detection to address some shortcomings of other methods in this region (Chapter 4). Using this approach we developed 600 possible LFE templates beneath the central Southern Alps. We intend these templates to be used to generate a complete 6.5 year catalogue of LFEs. We have also deployed and made use of additional seismometers which extend the SAMBA network southwards, to surround the tremor-generating zone found by Wech et al. [2012]. Using these new templates with the longer dataset and the objective template set, we aim to obtain further, additional and more accurate LFE locations.

LFE locations may have implications for the structure of the Alpine Fault at depth. For example, our initial locations in Chapter 3 showed clustering near an inferred listric Alpine Fault, but two families were located west of this, possibly on a secondary, subvertical strand of the fault as modelled by Koons et al. [2003]. LFE locations also have implications for the rupture extent of Alpine Fault earthquakes. Laboratory studies show that at moderate sliding velocities that would be achieved or exceeded during coseismic slip, fault gouge materials can transition from velocity strengthening to velocity weakening [Ito et al., 2015]. In other words, during an earthquake, coseismic slip can penetrate into slow-slip zones.

Coseismic rupture through a slow-slip region was physically observed for the  $M_W$  9 Tohoku earthquake, which ruptured through a shallow slow-slip zone [Ikari et al., 2015a]. The geometry proposed by other authors of a listric Alpine Fault [e.g. Koons et al., 2003, Herman et al., 2009, Stern et al., 2007, Lamb and Smith, 2013, Lamb et al., 2015], and highlighted here by the locations of LFEs (Figure 7.1) therefore has implications

for seismic hazard east of the Southern Alps. If the Alpine Fault is listric at depth and coseismic rupture can continue through the LFE region then shaking intensities south-east of the Alpine Fault have likely been underestimated. The hazard due to this has to date not been considered in models of coseismic shaking in New Zealand: for instance, Stirling et al. [2012] considered an Alpine Fault source extending to a depth of 12 km.

### 7.1.2 Micro-earthquake studies

In Chapter 5 we showed that near-fault seismicity near Whataroa is dominated by small earthquake sequences, which we interpret as non-repetitive failure of small asperities close to the fault. These small earthquakes may be redistributing strain near the main strand of the Alpine Fault. The near-fault seismicity we document is notably depleted in earthquakes larger than  $M_L$  3.0 and has many small earthquakes. This is evident in the high b-value found for near-repeating seismicity near Whataroa. We interpret this high b-value ( $b=1.44$ ) to be due to a lack of large faults available for slip other than the Alpine Fault in this region.

The b-value we document is higher than that found for the central Southern Alps region in general ( $b=0.86$  [Boese, 2012]). However, without detailed catalogues of near-fault seismicity along-strike of the Alpine Fault it is difficult to understand the significance of this high b-value. To understand how seismicity changes along-strike of the Alpine Fault, further detailed studies of this type would be beneficial.

In Chapter 6 we outlined and made use of real-time earthquake detection methods during the DFDP-2 drilling, with the seismic monitoring focussed on the region analysed in Chapter 5. We observed limited seismicity within 10 km of the drill-site, and did not detect any changes in seismicity during drilling outside of background seismicity. We have provided documentation and advice for further monitoring (Appendix E), which should be useful for future Alpine Fault drilling projects.

## 7.2 Opportunities for further investigation

### 7.2.1 Low-frequency earthquakes and slow deformation

To take full advantage of LFE catalogues and other seismological observations in the central Southern Alps we need a greater understanding of the state of stress on both the shallow and deep portions of the Alpine Fault. There are a number of properties that detailed studies of the LFE catalogues generated in this thesis could elucidate.

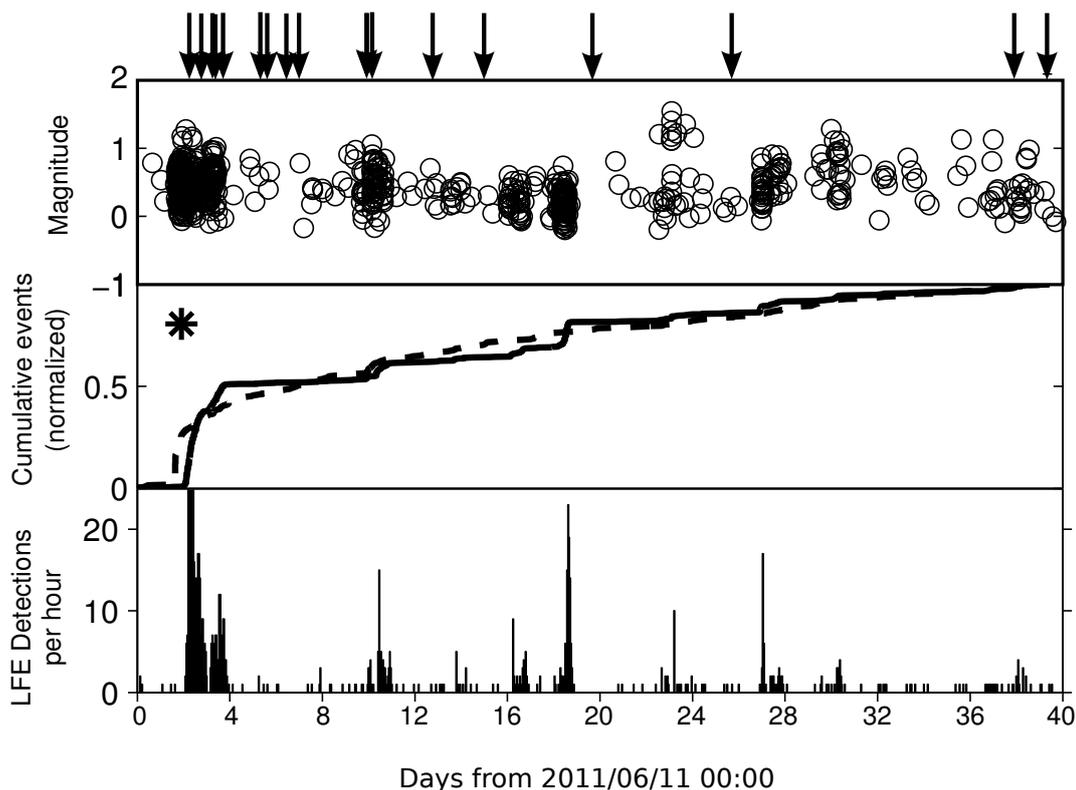


Figure 7.2: Low-frequency earthquake detection rates increase directly following a large ( $M_W$  6) earthquake in Christchurch marked by the star, likely indicating triggering. By quantifying of the stresses exerted by earthquakes that appear to trigger LFEs on the deep extent of the Alpine Fault, we may be able to quantify the strength of the Fault at depth. Arrows indicate the occurrence of earthquakes of  $M_W > 4$  in the GeoNet, nationwide catalogue.

### Temporal variations

How the deep extent of the Alpine Fault responds to large regional earthquakes might be well quantified by variations in LFE detection rate (Figure 7.2). To study the triggering relationship between regional earthquakes and LFEs, the dynamic and static stresses experienced due to these regional earthquakes could be calculated, and the methods of either Hill et al. [2013] or Beeler et al. [2013] applied to estimate the state of stress on the Alpine Fault at depth. This information could be useful for relating LFE stress-drops to rupture scales, making use of the scale-invariance inferred herein and by other authors [Sweet et al., 2014, Bostock et al., 2015]. Of note is the apparent lack of triggering by teleseismic events in the LFE catalogue. We have studied the wavetrain of the Tohoku M9 earthquake and found no evidence of triggered tremor or LFEs on the Alpine Fault.

Understanding the difference between background (ambient) and triggered tremor or LFE swarms, and the patterns inherent in both, could be key to differentiating between tremor that will not trigger an earthquake, and tremor that may trigger an earthquake.

This might be achieved by modelling both ambient and triggered tremor to understand the mechanism that stops slow-slip from propagating into the seismogenic region on the Alpine Fault.

Of note is that we have not observed a characteristic return period of LFEs and tremor in the central Southern Alps, unlike Cascadia [e.g. Rogers and Dragert, 2003] or other subduction zones. Whether this is real or due to the limited catalogue in Chapter 3 is unclear. By extending the catalogue through the full 6.5 year dataset using the templates generated in Chapter 4 we may be able to determine whether there is an episodic nature to tremor and LFEs in the central Southern Alps.

### **Magnitudes**

Magnitude distributions of LFEs may elucidate certain properties of the asperities that fail during LFEs, specifically rupture length and rupture area. By understanding these properties, and the stress drops associated with LFEs, we may be able to quantify the amount of coseismic slip associated with LFEs and tremor events. These estimates would likely be significantly less than any geodetically measurable slip as LFEs likely do not represent the whole of the region slipping [Aguiar et al., 2009].

### **Focal mechanisms**

High quality stacked LFE waveforms offer the opportunity to constrain LFE focal mechanisms, either by inversion of first motions and body wave amplitude ratios, or by modelling waveforms produced by possible focal mechanisms and comparing these with the recorded waveforms (e.g. in a moment tensor inversion methodology [cf. Ide and Yabe, 2014]). LFE focal mechanisms would provide constraints on the slip-direction and stress orientations at depth on the Alpine Fault. These stress orientations coupled with the highly detailed catalogues generated here would be useful in generating and improving realistic models of the Alpine Fault. Slip styles (likely either reverse or strike-slip or oblique) would also be key for determining if and how strain is partitioned on structures beneath the central Southern Alps, particularly testing if there is a single listric fault at depth, or multiple structures similar to the model of Koons et al. [2003]. Work is currently underway as part of Laura-May Baratin's PhD work at Victoria University of Wellington to calculate focal mechanisms of LFE families.

### **Crustal anisotropy**

High quality LFE stacks have also been used to investigate crustal anisotropy resulting in delayed S-phases on orthogonal components [e.g. Matharu et al., 2014]. In our initial

template stacks in Chapter 4 we are beginning to see S-phase delays between north and east components of some stations. Measurements of crustal anisotropy from LFEs would provide useful extra ray-paths for analysis of splitting directions in the crust.

### **Scattered phases and improved locations**

Further analysis of phases other than the main P and S-phases in high quality LFE stacks (scattered and reflected phases) is being used in Cascadia for more precise location of LFE families, and for analysis of the low-velocity zone that the LFEs in Cascadia lie within [Bostock, 2015]. Bostock [2015] suggested that these scattered phases represented phase conversions and reflections at the boundary of the low-velocity (and low  $V_p/V_s$ ) zone. LFEs under the central Southern Alps also locate near a zone imaged to have low seismic velocities [Smith et al., 1995, Stern et al., 2001] and low  $V_p/V_s$  (low  $Q_P$ ) [Eberhart-Phillips et al., 2008], and we observe scattered phases within high-quality family stacks. We may, therefore, be able to analyse the scattered phases to provide greater spatial resolution of both the LFEs and the low-velocity zone. As yet it is unclear whether LFEs occur on a discrete structure, or throughout the crust adjacent to the Alpine Fault.

### **Geodetic observations**

Frank et al. [2015a] demonstrated how GPS time-series from different tremor periods can be stacked together to increase the signal-to-noise ratio of the geodetic data and extract the motion associated with slow-slip. As there is a GPS transect running across the Southern Alps, above the region documented to generate tremor and LFEs, it may be possible to implement this method here. However, the GPS data are not all of high quality, which may result in low quality stacks (Hreinsdottir 2015, pers. comm.). This is the topic of further research led by Laura-May Baratin at Victoria University of Wellington.

### **Catalogue extension**

Alongside detailed studies based on the catalogues generated here, it would be logical and methodologically simple to extend the LFE catalogue north-east along-strike using data collected on the WIZARD network. We have, however, attempted to look for tremor using a short (one month) period of the WIZARD data using the WECC methodology [Wech and Creager, 2008], and have been unable to identify tremor. This shouldn't preclude a more thorough LFE search using the methodology outlined here:

as we have demonstrated, tremor in the central Southern Alps is of short duration and low energy, making it difficult to detect with most methods.

By telemetering key sites of the SAMBA network, and by implementing a real-time or near real-time matched-filter detector we could use the LFE templates generated here to construct real-time LFE catalogues. The knowledge gained from studying the properties of LFEs in the central Southern Alps could be applied to these real-time catalogues to forecast LFE rates, and the stresses imposed on the seismogenic zone of the Alpine Fault by them.

### EQcorrscan development

We have developed an extensible, portable and scalable Python package (EQcorrscan) based on Obspy [Beyreuther et al., 2010] for the detection of LFEs, earthquakes and explosions. Because EQcorrscan is based on Obspy, there is scope to extend the codes to incorporate real-time data using Obspy's real-time module. Work by Chet Hopp of Victoria University of Wellington is currently underway to implement standardised data handling within EQcorrscan which would be necessary to make use of real-time functionality. This data handling will be completed using Seishub<sup>1</sup>. Later we hope to integrate the Adaptable Seismic Data Format<sup>2</sup>, which is currently in development. The Adaptable Seismic Data Format will allow more reproducible handling and processing of data (through improved provenance handling), while also implementing parallel input/output operations, which would provide significant speed-ups to EQcorrscan. The Adaptable Seismic Data Format project also includes ties to SpecFEM3DGlobe<sup>3</sup>, which models earthquake waveforms in a 3D finite element mesh from realistic earthquake sources. This could provide more realistic synthetic templates for use in the synthetic template matched-filter detection methodology presented here.

### Modelling

Ideally any LFE catalogue would feed into a model of the Alpine Fault that uses physically realistic fault properties to probabilistically model LFE generation. Properties that could be included in such a model include, but are not limited to:

- strength of the LFE region gained from studies of triggering;
- fault geometry from a range of studies including precise LFE hypocentre locations;

---

<sup>1</sup><https://github.com/barsch/seishub.core>

<sup>2</sup><http://seismic-data.org>

<sup>3</sup>[https://geodynamics.org/cig/software/specfem3d\\_globe](https://geodynamics.org/cig/software/specfem3d_globe)

- frictional strength of the seismogenic portion of the Alpine Fault gained from laboratory studies of fault gouge [e.g. Boulton et al., 2014, Ikari et al., 2015b, Schleicher et al., 2015].

Modelling software available to do this include, but are not limited to, RSQSim [Richards-Dinger and Dieterich, 2012] (a rate-and-state based model which has been applied to both earthquakes and slow-slip [Colella et al., 2012]) and PyLith<sup>4</sup> (which models quasi-static and dynamic deformation).

## 7.2.2 Micro-earthquake studies

### Synthetic templates for earthquake detection

The synthetic template matching method is well-suited to detecting microseismicity in the central Southern Alps. However, we have shown that to detect such earthquakes templates need to be precisely co-located with the seismicity. This is also true for real earthquake templates (real, detected waveforms rather than synthetic waveforms). In Chapter 5 we showed that our earthquake templates were only appropriate over a short time-scale, and that detected earthquakes were close to the template used to detect them (<500 m). From this it seems that to detect a spatially complete earthquake catalogue using either synthetic or real templates we would need a finely spaced grid of templates. If using a finely spaced grid of templates with known locations, earthquakes could be detected and located in tandem, with fine relocations possible using relative picks as described in Chapter 5.

In Chapter 4 we suggested extending the synthetic matched-filter methodology with a subspace detector [Harris, 2006, Barrett and Beroza, 2014]. This may allow for earthquakes of a range of focal mechanisms to be detected with one (subspace) template. This could also be designed to allow greater variation in location for detections made by one (subspace) template.

### Along-strike variation in near-fault seismicity

If these methodological extensions proved fruitful, then a systematic along-strike study of microseismicity in the Southern Alps using the combination of SAMBA, WIZARD, GeoNet and the COSA (Central Otago Seismic Array [Warren-Smith et al., 2015]) data could be undertaken. This might elucidate changes in fault properties along-strike, which could inform rupture simulations, e.g. by mapping on-fault seismicity we could map both the fault at depth, and where there is no seismicity on the fault, to

<sup>4</sup><https://geodynamics.org/cig/software/pylith/>

look for locked patches which may be more likely to nucleate large earthquakes [e.g. Haberman, 1981].

A systematic along-strike matched-filter analysis would also be worthwhile to scan for repeating events on the Alpine Fault. Currently the Alpine Fault has no documented repeating earthquakes along its length. In comparison the San Andreas [Nadeau and Johnson, 1998, Nadeau and McEvilly, 2004, Zechar and Nadeau, 2012] and Japanese subduction zones [e.g. Uchida et al., 2003, Igarashi et al., 2003] have well documented and studied repeating earthquakes. These repeaters have been shown to respond to large earthquakes [e.g. Schaff et al., 1998, Nadeau and McEvilly, 2004] with responses consistent with post-seismic afterslip. In the case of the Tohoku M9 earthquake, small repeating earthquakes show accelerated recurrence prior to the mega-thrust earthquake [Mavrommatis et al., 2015]. If we could detect and document changes in repeating seismicity on the Alpine Fault, alongside the long-duration LFE catalogues developed here, we may be able to extract similar information regarding strain-rate on both the shallow seismogenic region of the fault, and the deep transitional region.

### **Seismic monitoring during drilling**

We would also recommend making use of the matched-filter methodology during future drilling projects, both on the Alpine Fault, and in other regions, such as geothermal drilling, or in oil regions where injection and fracking are used. The matched-filter method could be used in an adaptive framework, whereby the matched-filter routine runs in tandem with standard observatory practices and newly detected earthquakes are used as templates in real-time and hindcast matched-filter routines [Ringdal et al., 2009]. The synthetic matched-filter routine outlined here could also be made use of to generate a fine-grid of sources around the drill-site to provide real-time, high precision detection and location of small earthquakes within drilling noise.

### **Studies near the Southern Alpine Fault**

We have also proved the effectiveness of the matched-filter technique applied to COSA data, with which, alongside fellow Victoria University of Wellington graduate student, Emily Warren-Smith, we generated a detailed aftershock catalogue of 2545 earthquakes following the 04/05/2015  $M_L$ 6.0 Wanaka earthquake. We were able to generate this catalogue rapidly (within two months of the earthquake, using one month's data) by taking advantage of the matched-filter detection routines developed in this thesis. We also computed high-precision double-difference locations for 80% of events, allowing us to define the fault plane (Figure 7.3). This catalogue also highlights how different template events have different detection capabilities. For example, the  $M_L$ 6.0 main-

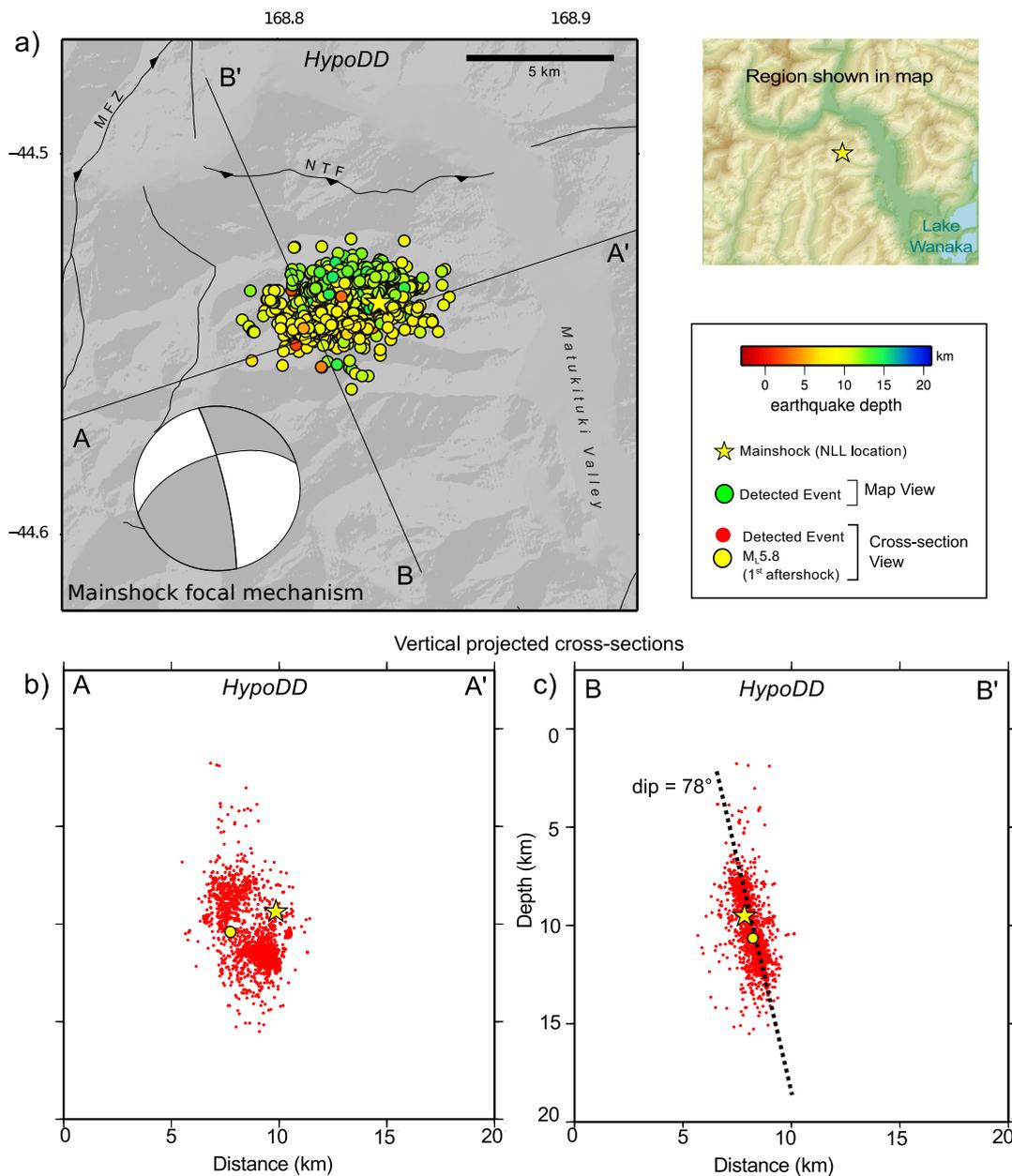


Figure 7.3: Earthquakes detected following the  $M_L$ 6.0 Wanaka earthquake on 04/05/2015 using matched-filter detection. Locations have been generated using hypoDD with cross-correlation derived picks. Figure from Emily Warren-Smith (pers. comm. 2015).

shock, when used as a template, did not detect any other events. This may be due to the inherent increased waveform complexity of this large earthquake. Further analysis of this catalogue relating to what qualities in a template allow it to detect the largest range of magnitudes could provide important constraints on the expected completeness of matched-filter catalogues.

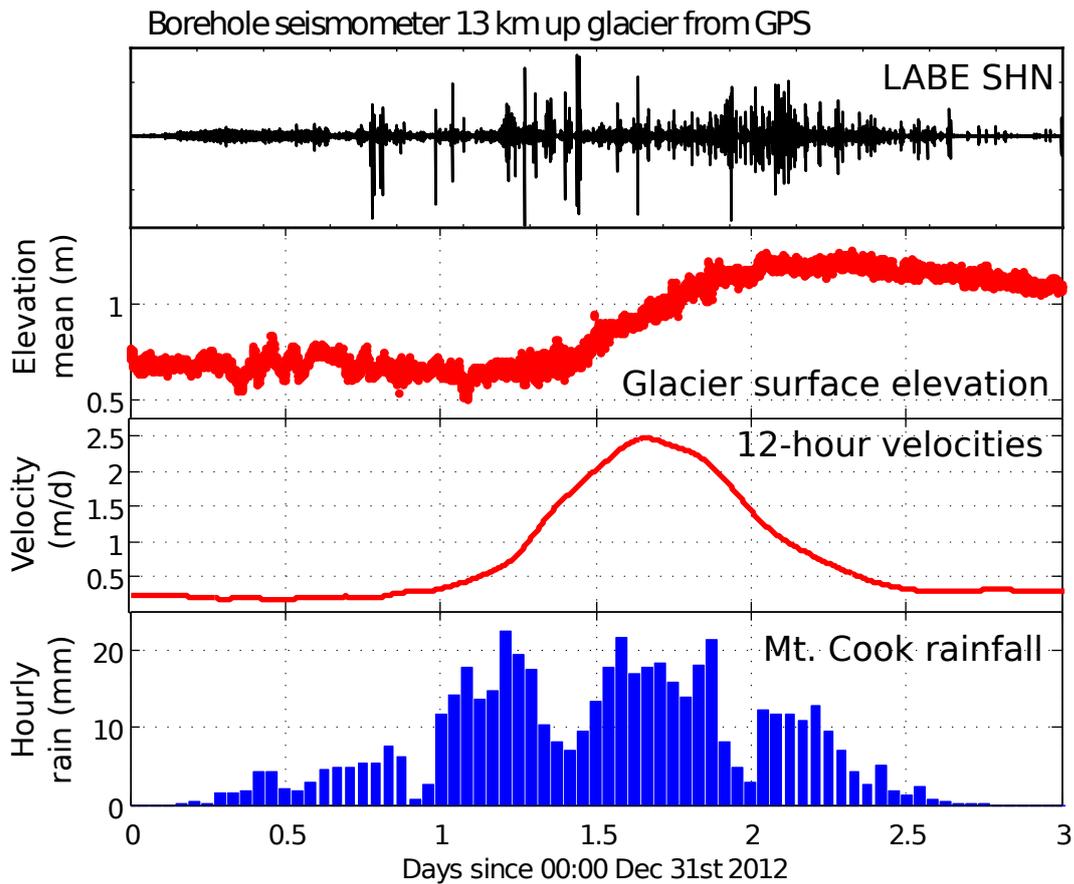


Figure 7.4: Data collected on and nearby the Tasman glacier in the central Southern Alps. Rainfall data collected from Mt. Cook village, approximately 5 km from the glacier. Velocities and elevation data are collected from GPS sites deployed on the upper surface of the glacier, which is covered in debris. Seismic data are taken from SAMBA site LABE, which is within 200 m horizontally of the glacier. We observe increased seismic amplitudes associated with speed-up events. Figure from Huw Horgan, GPS data are published in Horgan et al. [2015].

---

### 7.2.3 Tasman glacier slip

SAMBA has also recorded seismicity from the Tasman Glacier, including the seismic signal of large speed-up events (Figure 7.4) which occur during periods of high rainfall. Initial analysis of the seismic signal at SAMBA site LABE (above and c. 300 m horizontally of the Tasman glacier) shows directivity of shear waves within the signal, with optimum orientations up and down the glacier. In contrast, the signal at GOVA, the next nearest seismic station and experiencing similar rainfall shows random orientations, suggesting that the dominant signal is due to rain noise.

These speed-up events (and the seismic signal associated with them as recorded on LABE) have been attributed to buoyant sliding of the Tasman Glacier associated with large rainfall events and basal cavitation [Horgan et al., 2015]. This signal, although currently difficult to locate, provides an interesting possible proxy for seismic, non-volcanic tremor, with similarly emergent waveforms. Further study of this signal coupled with the detailed geodetic, near-source data may provide constraints on the mechanism of tremor on faults, which could be fed into realistic models.

A group within Victoria University of Wellington's Antarctic Research Centre and School of Geography, Environment and Earth Sciences led by Huw Horgan and Sam Taylor-Offord plans to deploy a small seismic and geodetic array on the glacier in 2016. This array should provide better constraints on the source of the seismic signal, and may allow mapping of rupture fronts during speed-up events.

## 7.3 Concluding statement

This study identifies and characterises the first LFE events observed along the Alpine Fault. Our LFE catalogues provide insight into contemporary deformation on the deep extent of the Alpine Fault, which is likely to have implications for seismic hazard assessment, earthquake modelling and potentially earthquake forecasting. Our work provides a framework for further study of LFEs and slow deformation in the central Southern Alps. This study has also highlighted near-fault micro-seismicity in a region of relative quiescence, which will form the basis for further along-strike catalogues of near-fault seismicity, and form the basis of continued searches for repeating earthquakes on the Alpine Fault.



# Appendices



## A Network set-up

SAMBA is principally a network of short-period borehole seismometers deployed within the Central Southern Alps of New Zealand’s South Island. The locations of all SAMBA sites are shown in table 2.1. The backbone of the network has recently (in March 2013 and again in March 2014) been supplemented by three additional short-period surface sensors to complement the single short-period surface sensor deployed above the Copland Valley in early 2010. The nuances of installing the borehole seismometers are covered in some detail by [Boese, 2012].

The three seismic sites deployed to the south of the main SAMBA deployment were installed to assist in the identification and location of LFEs after [Chamberlain et al., 2014] identified sparse LFE activity beneath this region. It was hoped that these sites alongside the Haast site of Emily Warren-Smith’s Central Otago Seismic Array (COSA, not to be confused with the Southern Alps Micro-earthquake Borehole Array (SAMBA) site of the same name) network would allow constraints to be placed on the southern extent of LFE and tremor activity. This however was not the case as the Haast site encountered multiple failures, never providing enough data to hope to detect LFEs.

All sites of the SAMBA network are deployed with RefTek (Refraction Technologies) RT-130 dataloggers recording at 200 Hz. All borehole sites use GeoSpace Technologies HS-1-LT short-period sensors. All borehole sites use a 2 Hz version of this sensor apart from FRAN, which uses a 4.5 Hz sensor due to sizing constraints imposed by the drilling procedures when drilling this, the deepest of our holes. All surface sites are, at the time of writing, deployed using 1 Hz Marks Products L4 sensors. MTBA was initially deployed with a Guralp 3-ESP-C broadband sensor between 30/03/2013–21/03/2014 before being replaced with the more sturdy L4 after the Guralp failed.

Sites are powered by 12 V deep-cycle batteries, which themselves are kept charged by solar panels wired through solar controllers. [Boese, 2012] found that certain models of solar controllers resulted in frequent noise spikes, hence care was taken to use predominantly SunSaver and ProStar solar controllers. Solar panel power is site dependent with sites with less sun exposure requiring 80 W panels, and those with more consistent sunlight needing only 60 or 40 W panels.

Timing for the RT-130 dataloggers requires the deployment of a GPS sensor at the site.

---

Most sites in the SAMBA network have a RefTek model GPS clock, but some have been used with Garmin model GPS clocks. These have been found to be less reliable than their RefTek branded counterparts. At sites with consistent power the clocks run continuously, however for sites with more problematic powering (e.g. EORO, COSA with the previous solar panel deployment) clocks are run in cycled mode where they update the timing every 5 minutes. Phase differences encountered in the mode are not significantly increased when considering LFE signals, but for higher frequencies can be an issue. In processing data for LFE and earthquake detections the timing uncertainty was always considered, and if uncertainties were too high for the frequencies being investigated the data were discarded. The pros of running the GPS clocks in cycled mode are a large reduction in power demand, which can result in more continuous data when solar power is limited.

Details of the deployment at each site are reported in table A.1.

Station	DAS	Firmware	Sensor	GPS	S. Panel	S. Mount	S. Contoller	Disk 1	Disk 2	Status
GOVA	9160	3.4.3	HS-1-LT 102	130-101 2684	2x60W	Waratah	Sansaver	<b>CF-RT16-12</b>	R7	Good
LABE	915C	3.4.5	HS-1-LT 107	130-102 5873	60W	A-Frame	Flexcharge	<b>CF-SD08-02</b>	Unknown	Good
EORO	A970	3.4.3	HS-1-LT 104	130-102 5842	60W	A-frame	Prostar	<b>2G</b>	CF-SD04-04	Good
FRAN	A852	3.4.3	HS-1-LT 132 (?)	130-102 5569	60W	Pole Mount	Sansaver	<b>CF-SD16-01</b>	R9	Good
COSA	915E	3.4.3	HS-1-LT 103	130-102	80W	A-frame	Prostar	<b>CF-SD16-08</b>	<b>8B</b>	Good
POCR	A979	3.4.3	HS-1-LT 132 (?)	130-102 3664	80W	A-frame	Prostar	<b>CF-SD16-17</b>	17B	Good
WHAT	A895	3.4.3	HS-1-LT 106	130-102 2467	2x60W	Waratah	Sansaver	<b>CF-SD08-10</b>	R3	Good - telemetered
COVA	BAD5	3.4.5	Marks L4 733	130-102 3690	60W	Pole Mount	Prostar	<b>CF-RT08-05</b>	CF-RT04-01	Good
MTFO	915F	3.4.3	HS-1-LT 105	130-102 5867	60W	A-frame	Prostar	<b>2710</b>	15B	Good
WHYM	AC14	3.4.3	HS-1-LT 101	130-102 5906	60W	A-frame	Prostar	<b>10G</b>	CF-SD04-01	Good
SOLU	<b>9F99</b>	2.9.9	Marks L4 602	Garmin	40W	Pole Mount	Prostar	<b>CF-SD08-04</b>	<b>CF-SD04-07</b>	Major service
MTBA	<b>AC13</b>	3.4.5	L4 734	AB3	60W	A-frame	Prostar	<b>CF-SD08-15</b>	<b>3E</b>	DAS issues
LARB	9C99	3.4.5	Marks L4 586	130-102 3145	80W	A-frame	Prostar	<b>CF-RT08-13</b>	17B	Good

Table A.1: Inventory of deployed SAMBA equipment and status as of March 2015 (last service led by Calum Chamberlain). Components in bold type indicate a change from the previous set up.

---

## B Data continuity

The climate and morphology of the Southern Alps makes maintaining a seismic network problematic. Under standard operational conditions servicing of the network is undertaken bi-annually in the Austral autumn and spring. This is to ensure that the stations are ready to be left for the winter, and to fix problems that have occurred over the winter. Some of the physical problems encountered during these service trips include, but are not limited to:

- Snow-cover resulting in lose of power (solar panel receives no light);
- Damage caused by Kea (cables chewed);
- High winds blowing over sites;
- Lightning damage;
- Flooding of sites.

While every effort has been made to ensure continuity of data, there are occasionally small to large periods of data loss. These have occasionally been exacerbated by a lack of opportunity to get to sites (high cost of helicopters and lack or a replacement sensor prevented fixing SOLU after a lightning strike in which every electrical component of the site apart from the battery underwent failure). The resultant data continuity is shown in Figure B.1.

Data Continuity

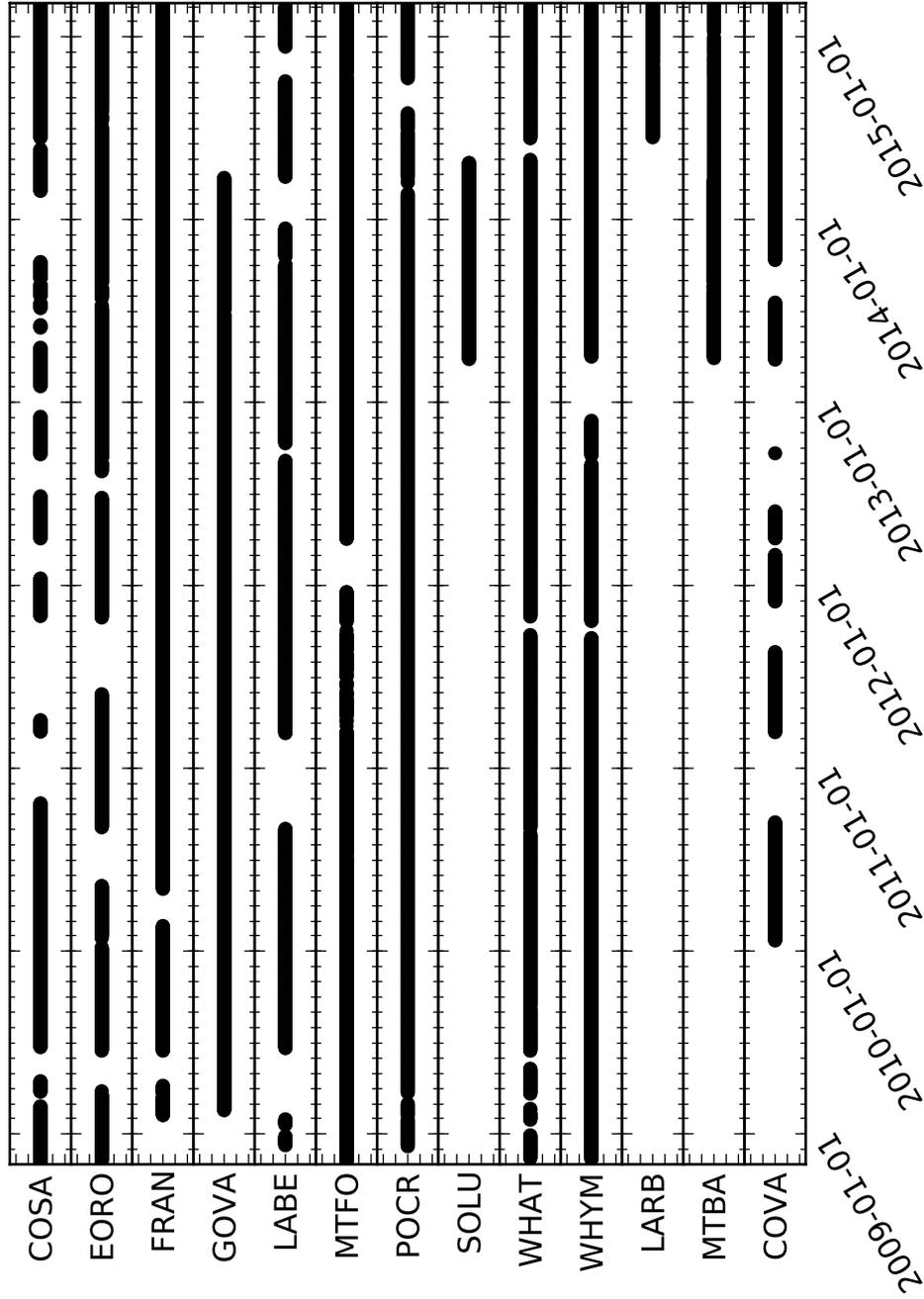


Figure B.1: Continuity of the SAMBA stations used here. Black marks indicate a station is online, but do not provide any indication of quality.

## C Common noise characteristics of SAMBA data

In this study non-earthquake derived seismic amplitudes are considered noise. This covers a broad spectrum from electrical noise, and analogue to digital filtering effect to weather events and wave noise. All of these factors need to be considered before drawing conclusions from seismic data. In this appendix some of the more common noise is shown and discussed.

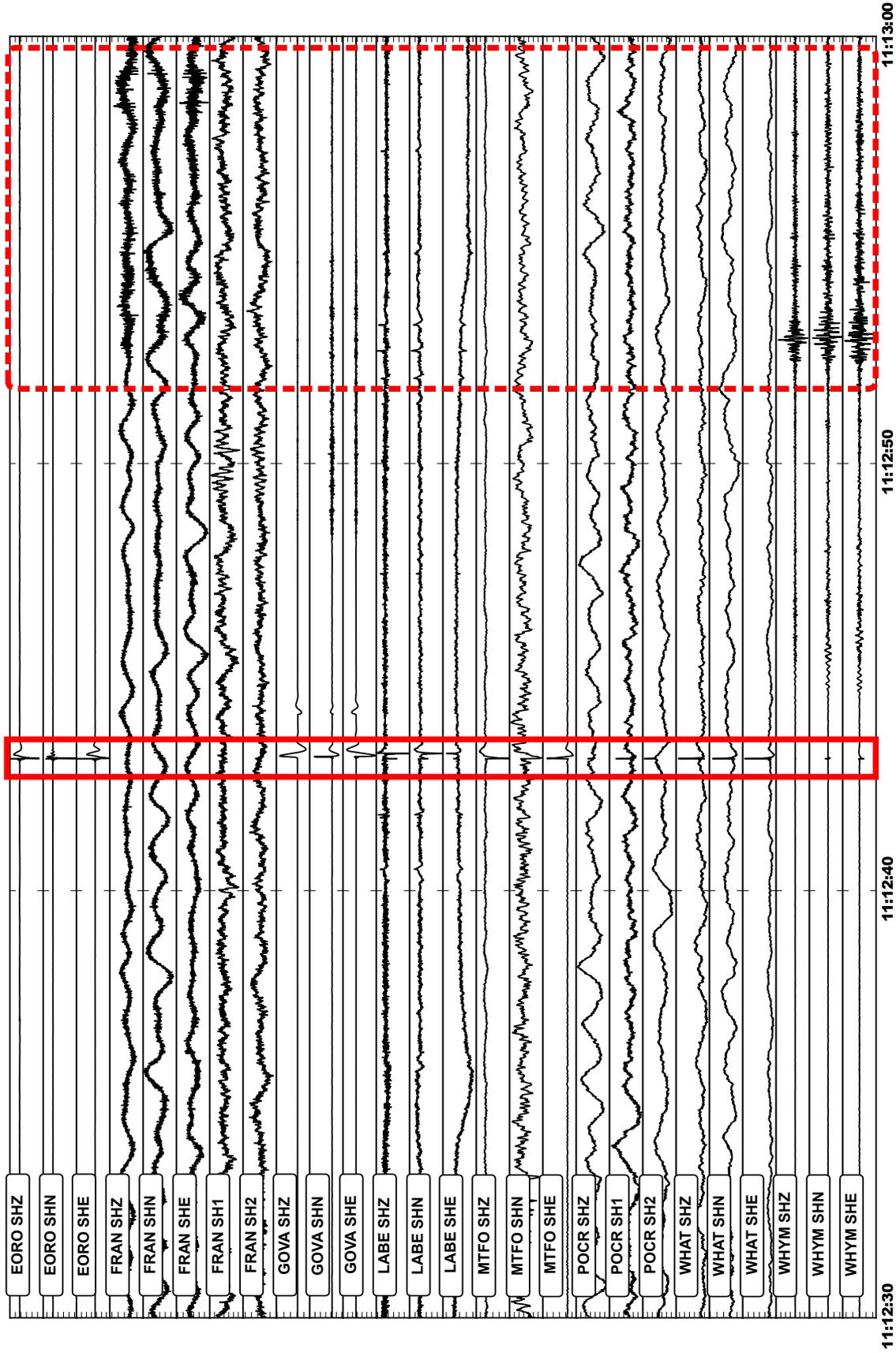
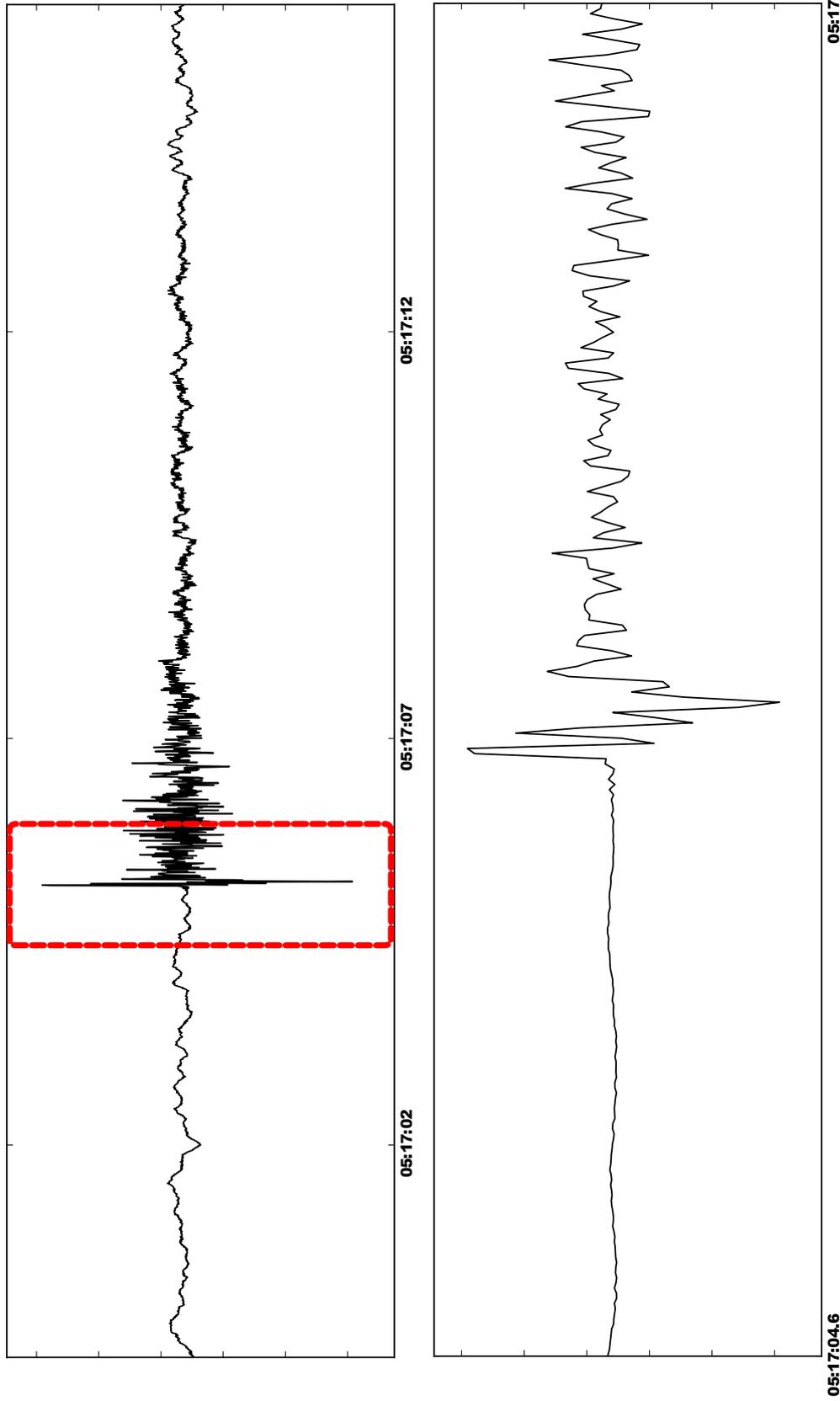


Figure C.1: Lightning strike on the SAMBA network on 2012/08/26. The red box shows the timing of the lightning seen as an impulse arriving with zero movement across the network. Not all sites see the lightning, notably both sensors at FRAN do not see the lightning, nor does the sensor at WHYM, in both of these records though the thunder that follows is clear (dashed red box). These lightning strikes often result in false detections, and automated pickers often pick the lightning as a P-phase and the thunder as an S-phase.

P-onset of ML 1.0 Local earthquake on vertical channel of WHAT on 2014/11/30



05:17:04.6

05:17:05.9

Figure C.2: A-causal ringing prior to the P-phase onset thought to be introduced in the digitization by the application of an anti-aliasing Finite Impulse Response (FIR) filter in the RefTek datalogger. The upper panel shows the vertical component of station WHAT, with a zoom into the p-phase (shown by the red box) arrival in the lower panel. This acausal ringing can lead to the P-phase being obscured. Care must be taken when analysing the early onset of seismic waves to check that the data are due to real nucleation process affects rather than due to filtering effects. This filter effect occurs as the FIR filter is maximum-phase rather than zero-phase to be able to have a very sharp cut-off at the Nyquist frequency.

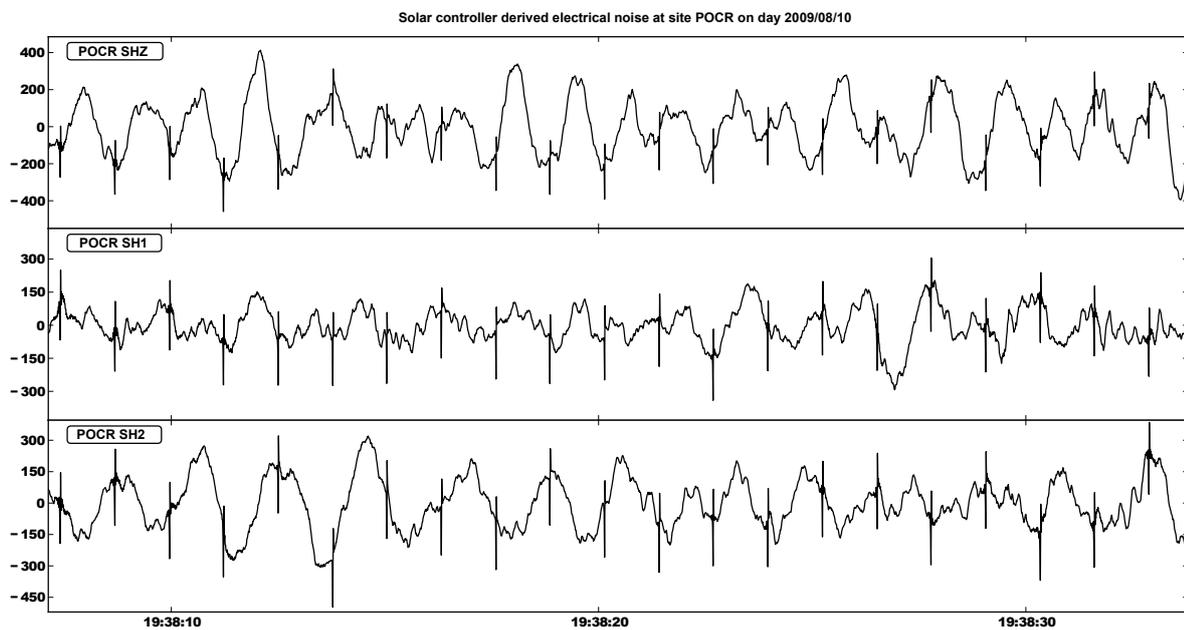


Figure C.3: Solar controller induced spiking on station POCR on date 2009/08/10. While obvious to the eye, these can be problematic for automated detection. While generating the LFE catalogue we found that filtering these spikes in the LFE frequency band (2–8 Hz) resulted in relatively good (below 0.2, but non-zero) correlations. To overcome this data from stations with this spiking were removed as de-spiking algorithms also introduced correlated noise.

## D LFE event verification

In testing threshold levels for cross correlation we developed the following methodological flow:

1. Extract 1 day of raw seismic data per channel per station used in detection of a specific template, here we use the day of the Dusky Sound earthquake as it has both earthquake and LFE signals;
2. Fast Fourier Transform the day of data and extract the amplitude and phase information;
3. Generate a random phase spectra for the same duration, before replaciong the first sample and sample at the half the length of the time series with those in the original phase spectra;
4. Combine the randomised phase spectra and the original amplitude spectra and compute the inverse Fourier transform of these;
5. Confirm that the amplitude spectra of both the phase randomised data and the original data are the same;
6. Generate 100 different random day long time series with 50 random impulses (both random in position and amplitude) on a day of otherwise zeroed time series;
7. Convolve the impulses trace with the 6 s long LFE template, repeated for each channel and station with the appropriate template to generate a clean time series;
8. Normalise the clean (according to the maximum amplitude) and the phase randomised (according to the mean amplitude, used as the spike in amplitude from the Dusky earthquake has been maintained and so skews the SNR values) time series;
9. Sum the normalised clean and normalised phase randomised data to generate a seeded time series;
10. Apply the correlation routine across all traces and sum the correlation values at each time to generate a time series of correlation sums (ccsum);

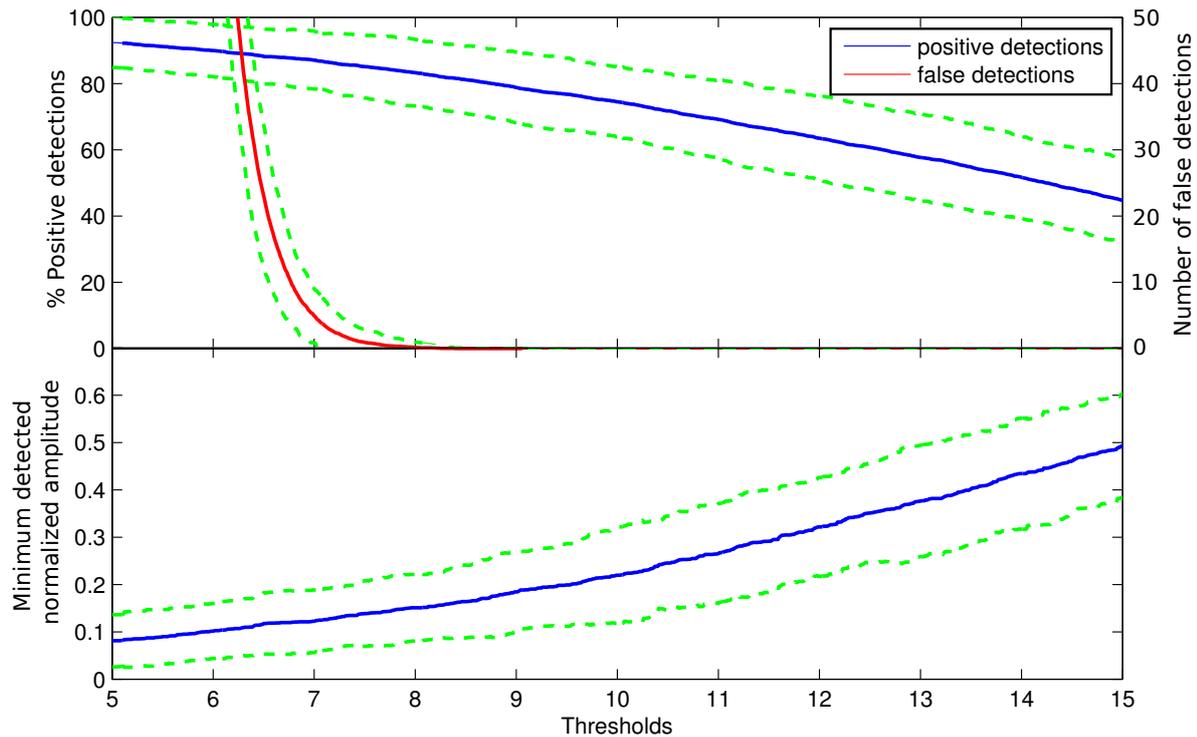


Figure D.1: Testing thresholds (MAD multipliers) for template 17208.1.2 for testing during a high noise day. Green dashed lines show standard deviation from the means plotted as solid lines. % positive detections corresponds to the percentage of seeded templates recovered by the routine whereas the number of false detections is the actual number of false detections made per day for this template.

11. Iteratively scan through the cccsum time series at a range of MAD multiplier thresholds from 5–15 at steps of 0.01 and output the detections;
12. Finally compare the detected events with the impulses and calculate the percentages of positive and false detections and the minimum SNR.

Rather than using purely random noise we chose to use noise with the same amplitude spectra as the data we intend to analyse. In generating this random noise we take a day of actual seismic data containing LFE template detections and convert to the Fourier domain. To remove the signal from the data we randomise the phase component, but retain the amplitude information. We then combine the old amplitude component and the newly randomised phase component and compute the inverse Fourier transform. We see that prior to phase randomisation we achieve template matches, but after phase randomisation there are no matches (prior to seeding). We also see that the amplitude spectra for both the phase randomised and the actual data are the same. For each template we check a range (1001 thresholds between 5 and 15 spaced at 0.01) of Median Absolute Deviation (MAD) multipliers to determine the best (near 95% confidence in detections) threshold. For each MAD value 100 tests are run each with a separate randomly seeded time series, resulting in 100,100 tests for each master template (over 14 million full tests).

# E Seisan training manual

To pick and confirm earthquakes for the DFDP-2 drilling procedures we will be using the Seisan software package. You will be alerted to events through the `dfdp.seismology@gmail.com` (Password \*\*\*\*<sup>1</sup> this is the same password as your server login) from the `rtquake` realtime analysis software. After receiving an email check the `All.png` file, which is a `.png` image of the traces for which detection occurred, if picks have been made these will be displayed. You should also check the location (if given) by clicking the link. If the event appears local to the drill-site it should be considered urgent. Once you have received an alert you should pick the event in Seisan on your machine. The procedure for this and a brief intro to Seisan is provided below.

## E.1 An Introduction to Seisan

Seisan is a group of programs written in C and Fortran for routine seismic processing and analysis in passive source seismology. Seisan is written to be interactive both through the command line and within x-windows. Seisan can be installed on most standard Linux distributions, OS X and Windows. Running on Windows requires the use of Seisan explorer, an interactive GUI.

On all distributions the directory structure is the same, underneath the top directory where Seisan is installed, there are a group of directories as follows:

- PRO: Programs, source codes and executables
- LIB: Libraries and subroutines
- INC: Include files for programs and subroutines
- COM: Command procedures
- DAT: Default and parameter files
- CAL: System calibration files
- INF: Documentation and information

---

<sup>1</sup>Password removed from public thesis

- ISO: Macroseismic information
- SUP: Supplementary files and programs
- TMP: Temporary files
- REA: Earthquake recording information , S-files (nordic format pick files)
- WAV: Waveform database

The three directories that we care most about are the REA, WAV and CAL directories as these contain the information that we will edit or install information into. For a detailed description of the nordic format, which Seisan pick files (S-files) are written in can be found in Appendix A of the Seisan manual (Seisan.pdf in the INF directory), pages 341–350. Waveform data will be stored in multiplexed miniseed format by default, in the WAV directory. The CAL directory contains the response information to be extracted from the DFDP seismology dropbox.

Seisan has a lot more potential than we will use it for, as for our routine purposes we only need to locate earthquakes and generate magnitudes. This will be done using the Seisan default location program initially (for which Seisan uses HYPO). For more reading the SEIAN manual contains much information and a detailed training guide for most of the functionality.

As my experience is with OS X and Linux, I will not cover Windows distributions here. Installation: Full installation information can be found in the Seisan manual (chapter 3, page 17), problems with installation should be directed to calum.chamberlain@vuw.ac.nz. Specific Set-up: Once everything is installed and the distribution is activated (by setting your top directory in COM/Seisan.bash or COM/Seisan.csh) and editing your .bashrc or .cshrc file as described on p.19 of the Seisan manual in the INF directory, we need to set the database structure up for use with our data. This is a simple procedure you should:

1. Change to the working directory by typing 'wo' in the command line;
2. Run 'makerea' from the command line and follow the prompts, you should call the database RTAF as we have done on the server the start date should be the 20140101 and the end date 20150201 (beyond the end of drilling by quite a bit). You should ask for both REA and WAV when it prompts about them. This will generate directories in the WAV and REA directories topped by your network name (RTAF) of the structure: NET\_/\_YYYY/MM (e.g., RTAF\_/2014/09). This will be where you download data to.
3. You should repeat the above but with the database name RTAFC, in which you will register your picked and checked waveforms and s-files this is where you will

upload data from.

## E.2 Data download

Next you should download data from the VUW servers to test picking. You will need to use a command based on the following instructions (replace YYYY and MM with date values and S-file with the name of the s-file, wav file with the name of the waveform file - \$Seisan\_TOP should be a set environment variable on your system when you have completed Seisan installation): You will need to copy both the S-file (pick file) and waveform file for the appropriate time and date. The S-file name is in the email on the first line just after the link to the map (however do not copy from the full path, only the year, month and filename are of use here for permissions reasons the data is hosted in multiple places). Get the S-file first and check it to find the name of the waveform file, which is on the 2nd line of the S-file. The S-files are named according to the the following convention: dd-hhmm-ssL.Syyyymm

For pick files:

```
scp -P10 dfdpseismology@evison.geo.vuw.ac.nz:/Volumes/GeoPhysics_09/users-data/-  
chambeca/DFDP_Seisan/REA/RTAF_/YYYY/MM/S-file $Seisan_TOP/REA  
/RTAF_/YYYY/MM/.
```

Read the S-file to get the name of the wav-file For waveform files:

```
scp -P10 dfdpseismology@evison.geo.vuw.ac.nz:/Volumes/GeoPhysics_09/users-data/  
chambeca/DFDP_Seisan/WAV/RTAF_/YYYY/MM/wav-file $Seisan_TOP/WAV  
/RTAF_/YYYY/MM/.
```

You will also require the response files for the sites which can be copied as follows if you have not already done it:

```
scp -P10 dfdpseismology@evison.geo.vuw.ac.nz:/Volumes/GeoPhysics_09/users-data/  
chambeca/DFDP_Seisan/CAL/* $Seisan_TOP/CAL/.
```

Finally you will need the appropriate STATION0.HYP file in your working directory. (This will overwrite your existing STATION0.HYP file. I would recommend moving your current file to something like STATION0.HYP\_old before using the command below):

```
scp -P10 dfdpseismology@evison.geo.vuw.ac.nz:/Volumes/GeoPhysics_09/users-data/  
chambeca/DFDP_Seisan/STATION0.HYP $Seisan_TOP/WOR/.
```

## E.3 Pick data

Once the data are in place you should be sure to move to the working directory. From here type 'eev yyyyymmdd net' where yyyy is year, mm is month, dd is date and net is the name you have given to your database (RTAF). This will launch a command line interface, listing the events and waiting for your response. Typing 'p' will launch mulplt, the interactive, multi-channel picking interface. A full description of the possible commands run from mulplt can be found in section 6.2.5, P. 90 Seisan.pdf. Mulplt should plot all the channels and show the picks made in the s-file as red lines with phase names labelled.

It is worth checking s-p time before going ahead and picking events, if the s-p time is larger than 2 seconds then the event is likely to be outside of our sphere of interest and picking is less urgent. If the event does not appear to be a true earthquake you may delete the event by pressing q within the mulplt window then pressing d at the end of the eev line when asked for an input. You will then be asked to confirm that you want to delete the event and for your user ID (see later for explanation).

Picking data is done by a combination of mouse and keystroke. Holding the mouse over where you wish to place the pick, typing 1 will register an impulsive P-phase, 2 = an emergent P-phase, 7 an impulsive S-phase and 8 an emergent S-phase.

Zooming is done by mouse clicks. Window the data you wish to zoom over with two left mouse-clicks, on at the start of the window, and the second at the end. Zooming out can be done by clicking at a position before the green line from the first click.

If you wish to plot fewer channels these can be selected by clicking the channel names on the left and typing 'r' for replot once you have selected all channels, or you can type 'o' and select the traces that way. Typing 'o' followed by 'o' will plot all traces again.

Filtering can be done by default filters (z through m, the bottom row on an English keyboard) which are described if you click the menu button at the top right of the mulplt window. Typing 'r' will replot with the selected filter. You can assign high and low cut frequencies by typing '.' and entering the frequencies (separated by a space) then typing enter/return and 'r'.

Picks can be weighted by hovering over the pick and typing 'shift [1,2,3,4,9]', the equivalent of [!,@,#,\$,(] (important for mac's and foreign keyboards where the hash-key may be accessed in another way, e.g. on my mac, alt-3 is # which is required to give a weight of 3). Weights of 04 decrease in certainty as the number increases, a weight of 9 can be registered when timing is uncertain, this should be registered on both the P and S phases for the station, then the location code will take into account the s-p time, but not the absolute pick times. Picks can be removed by hovering over

the pick and pressing d. To locate the event type 'l', the mulplt window will disappear and text will appear in the terminal with the stations, distance from event, azimuth, ain (angle of incidence), weight, phase name, calculated phase, pick timing, residual, weight after location and di(?). Large residual picks should be re-evaluated. Typing enter after this output will re-open the mulplt window, this time with both the red manual picks and the blue predicted arrival times. This can be repeated until you are satisfied.

Note: if you find an error in location describing an IASP error then you should quit eev (type 'q' in both the mulplt window and the eev command line), then run 'remodl' followed by 'setbrn', this should rebuild the travel-time tables required for location. The error is usually: At line 1645 of file tau.for (unit = 22, file = /ISAP91\_linux.HED) Fortran runtime error: End of file

Seisan can calculate magnitudes in a variety of ways. Our preferred method is the local magnitude generated from the amplitude measured on a simulated Wood Anderson instrument. RTQUAKE has the potential to generate this type of magnitude, but I have so far failed to do this. However it often generates coda magnitudes, which are invariably higher than the local magnitudes. We therefore require you to pick the amplitudes of as many stations as possible. These should be picked on the largest body-wave phase of the Z-channel. The procedure is as follows:

1. Click on the channel you want to measure the amplitude on and type 't' for toggle (I would suggest having located the event before this stage to generate the predicted travel times to give you something to window around);
2. In the toggle window a new menu will appear, either click the WA button or type 'w' and select a window around the S-phase;
3. This should bring up another time-series beneath the previous one as a simulated Wood Anderson trace if the trace has the words no response info for ... beneath it then you do not have the appropriate response information for this station in your CAL directory.
4. Typing upper-case a 'A' will generate an automatic amplitude pick, this should be manually checked. If this is not appropriate then you can pick the amplitude manually by hovering over the trough and typing lower-case a 'a' then hovering over the peak associated and typing the same this should remove the automatic pick and place you manual pick there. In the Wood Anderson view the phase will automatically be registered as IAML.
5. Typing 't' for toggle will return you to the main mulplt window typing 'l' to locate will also generate a magnitude which will be shown in the text.

We will also be using PGA for assessment of hazard to the drill-rig and hole. These can be calculated routinely in Seisan in a similar fashion to magnitude picking. These picks are then also stored in the s-file. The procedure for this is as follows (do this on all channels of the sites WTSZ, WMSZ, WPSZ, WDSZ and WHFS):

1. Click on the channel you want to measure the amplitude on and type 't' for toggle (I would suggest having located the event before this stage to generate the predicted travel times to give you something to window around);
2. In the toggle window a new menu will appear, either click the Groun button or type 'g' and select a window encompassing the P- and S-phases;
3. You will be asked to select displacement, velocity or acceleration, you will need to pick acceleration (PGA);
4. This should bring up another time-series beneath the previous one as a simulated ground motion trace if the trace has the words no response info for ... beneath it then you do not have the appropriate response information for this station in your CAL directory.
5. Typing upper-case a 'A' will generate an automatic amplitude pick, this should be manually checked, if this is not appropriate the you can pick the amplitude manually by hovering over the trough and typing lower-case a 'a' then hovering over the peak associated and typing the same this should remove the automatic pick and place you manual pick there.
6. You need to select the type of amplitude pick, type 1 for a simple amplitude, or if you have identified the phase, click the appropriate amplitude;
7. Typing 't' for toggle will return you to the main mulplt window.

While these do not strictly give PGA (you will note that you have picked a peak-to-peak amplitude and PGA requires you to have simply the maximum amplitude), this is the most useful way to store amplitudes in the s-files as it stands. When interpreting these values it is sensible to halve the amplitude given in the s-file to give an indication of PGA.

Finally you should register the event (locally) by typing 'p' in the mulplt window to put the event in the database. You will be asked for information:

1. Do you want to register (y/n)? - usually y
2. Give 2-5 letter data base, return for same base your RTAFC
3. Change event type to L,R or D we care about local events mostly, L is local, R is regional and D is Distant

4. Give operator code, max 4 characters mine is CALU, Carolin's is CARO, use an abbreviation of your name.

You can then copy the s-file and waveforms back to the server to the RTAFC database (RTAF-checked). I will update the database daily. Note: you should only upload data once you are satisfied and only if you are on duty, e.g. while practising DO NOT upload your picks. The following commands can be used to copy your files (again replacing YYYY and MM with the appropriate information and the wav-file and s-file with the files you want to move):

```
scp -P10 $Seisan_TOP/WAV/YYYY/MM/wav-file dfdpseismology@evison.geo.-  
vuw.ac.nz:/Volumes/GeoPhysics_09/users-data/chambeca/DFDP_Seisan/WAV/  
RTAFC/YYYY/MM/.
```

```
scp -P10 $Seisan_TOP/REA/YYYY/MM/s-file dfdpseismology@evison.geo.vuw.-  
ac.nz:/Volumes/GeoPhysics_09/users-data/chambeca/DFDP_Seisan/REA/  
RTAFC/YYYY/MM/.
```

## E.4 Further information

Daily heli-plots can be viewed in png format in a dropbox account associated with the dfdp.seismology@gmail.com address (same address and password). We will also have continuously updated seismicity maps generated from the RTAFC database with 24 Hr rolling seismicity and the total seismicity detected by the network alongside hour long spectral plots. While picking you can view your location in google earth by opening the file \$Seisan\_TOP/WOR/gmap\_cur.kml (information on p.171 of the Seisan manual). This will display your most recent located event. To save you from reloading this file I advise that you open the file \$Seisan\_TOP/DAT/gmap-automatic.kml in google earth and save this to 'my places'. This file checks and reloads the gmap\_cur.kml file every few seconds so you will always have the most recent event plotted. This view is a nice graphical way of checking your residuals when picking as the view also shows the stations used to pick the event, coloured according to their residuals (green is low, red high) and the error ellipse around the hypocentre.



# Bibliography

- A. C. Aguiar and G. C. Beroza. PageRank for Earthquakes. *Seismological Research Letters*, 85(2):344–350, 2014. doi: 10.1785/0220130162.
- A. C. Aguiar, T. I. Melbourne, and C. W. Scrivner. Moment release rate of Cascadia tremor constrained by GPS. *Journal of Geophysical Research*, 114(7):1–11, 2009. doi: 10.1029/2008JB005909.
- K. Aki and R. Koyanagi. Deep volcanic tremor and magma ascent mechanism under Kilauea, Hawaii. *Journal of Geophysical Research*, 86(B8):7095, 1981. doi: 10.1029/JB086iB08p07095.
- K. Aki and P. G. Richards. *Quantitative Seismology*. University Science Books, New York, 2 edition, 2002.
- H. Anderson and T. Webb. New Zealand seismicity: Patterns revealed by the upgraded National Seismograph Network. *New Zealand Journal of Geology and Geophysics*, 37:477–493, 1994.
- S. Bannister and K. Gledhill. Evolution of the 2010-2012 Canterbury earthquake sequence. *New Zealand Journal of Geology and Geophysics*, 55(3):295–304, 2012. doi: 10.1080/00288306.2012.680475.
- P. M. Barnes. Postglacial (after 20 ka) dextral slip rate of the offshore Alpine fault, New Zealand. *Geology*, 37(1):3–6, 2009. doi: 10.1130/G24764A.1.
- S. A. Barrett and G. C. Beroza. An Empirical Approach to Subspace Detection. *Seismological Research Letters*, 85(3):594–600, 2014. doi: 10.1785/0220130152.
- N. C. Barth, V. G. Toy, R. M. Langridge, and R. J. Norris. Scale dependence of oblique plate-boundary partitioning: New insights from LiDAR, central Alpine fault, New Zealand. *Lithosphere*, 4(5):435–448, 2012. doi: 10.1130/L201.1.
- N. C. Barth, C. Boulton, B. M. Carpenter, G. E. Batt, and V. G. Toy. Slip localization on the southern Alpine Fault, New Zealand. *Tectonics*, 32, 2013. doi: 10.1002/tect.20041.

- J. Beavan, P. Tregoning, M. Bevis, T. Kato, and C. Meertens. Motion and rigidity of the Pacific Plate and implications for plate boundary deformation. *Journal of Geophysical Research*, 107(B10):2261, 2002. doi: 10.1029/2001JB000282.
- J. Beavan, S. Samsonov, P. Denys, R. Sutherland, N. Palmer, and M. Denham. Oblique slip on the Puysegur subduction interface in the 2009 July MW 7.8 Dusky Sound earthquake from GPS and InSAR observations: implications for the tectonics of southwestern New Zealand. *Geophysical Journal International*, 183(3):1265–1286, 2010. doi: 10.1111/j.1365-246X.2010.04798.x.
- N. M. Beeler, A. Thomas, R. Bürgmann, and D. Shelly. Inferring fault rheology from low-frequency earthquakes on the San Andreas. *Journal of Geophysical Research: Solid Earth*, 118(11):5976–5990, 2013. doi: 10.1002/2013JB010118.
- J. P. Benoit, S. R. McNutt, and V. Barboza. Duration-amplitude distribution of volcanic tremor. *Journal of Geophysical Research*, 108(B3):2146, 2003. doi: 10.1029/2001JB001520.
- G. C. Beroza and W. L. Ellsworth. Properties of the seismic nucleation phase. *Tectonophysics*, 261(1-3):209–227, 1996. doi: 10.1016/0040-1951(96)00067-4.
- K. R. Berryman, S. Beanland, and A. F. Cooper. The Alpine Fault, New Zealand: variation in Quaternary structural style and geomorphic expression. *Annales tectonicae*, 6, 1992.
- K. R. Berryman, U. A. Cochran, K. J. Clark, G. P. Biasi, R. M. Langridge, and P. Villamor. Major earthquakes occur regularly on an isolated plate boundary fault. *Science (New York, N.Y.)*, 336(6089):1690–3, 2012. doi: 10.1126/science.1218959.
- M. Beyreuther, R. Barsch, L. Krischer, T. Megies, Y. Behr, and J. Wassermann. ObsPy: A Python Toolbox for Seismology. *Seismological Research Letters*, 81(3):530–533, 2010. doi: 10.1785/gssrl.81.3.530.
- G. P. Biasi, R. M. Langridge, K. R. Berryman, K. J. Clark, and U. A. Cochran. Maximum-Likelihood Recurrence Parameters and Conditional Probability of a Ground-Rupturing Earthquake on the Southern Alpine Fault, South Island, New Zealand. *Bulletin of the Seismological Society of America*, 105(1), 2015. doi: 10.1785/0120130259.
- C. M. Boese. *Microseismicity in the central Southern Alps, Westland, New Zealand*. PhD thesis, Victoria University of Wellington, 2012.
- C. M. Boese, J. Townend, E. Smith, and T. Stern. Microseismicity and stress in the vicinity of the Alpine Fault, central Southern Alps, New Zealand. *Journal of Geophysical Research*, 117(B2):B02302, 2012. doi: 10.1029/2011JB008460.

- 
- C. M. Boese, T. A. Stern, J. Townend, S. Bourguignon, A. Sheehan, and E. G. C. Smith. Sub-crustal earthquakes within the Australia-Pacific plate boundary zone beneath the Southern Alps, New Zealand. *Earth and Planetary Science Letters*, 376:212–219, 2013. doi: 10.1016/j.epsl.2013.03.030.
- C. M. Boese, K. M. Jacobs, E. G. C. Smith, T. A. Stern, and J. Townend. Background and delayed-triggered swarms in the central Southern Alps, South Island, New Zealand. *Geochemistry, Geophysics, Geosystems*, 15, 2014. doi: 10.1002/2013GC005171.
- M. G. Bostock. Constraints on Subduction Zone Processes from Low Frequency Earthquakes. *AGU Fall Meeting Abstracts*, 2015.
- M. G. Bostock, A. A. Royer, E. H. Hearn, and S. M. Peacock. Low frequency earthquakes below southern Vancouver Island. *Geochemistry, Geophysics, Geosystems*, 13(11):Q11007, 2012. doi: 10.1029/2012GC004391.
- M. G. Bostock, A. M. Thomas, G. Savard, L. Chuang, and A. M. Rubin. Magnitudes and moment-duration scaling of low-frequency earthquakes beneath southern Vancouver Island. *Journal of Geophysical Research: Solid Earth*, 120, 2015. doi: 10.1002/2015JB012195.
- C. Boulton, D. E. Moore, D. A. Lockner, V. G. Toy, J. Townend, and R. Sutherland. Frictional properties of exhumed fault gouges in DFDP-1 cores, Alpine Fault, New Zealand. *Geophysical Research Letters*, 41(2):356–362, 2014. doi: 10.1002/2013GL058236.
- S. Bourguignon, S. Bannister, C. M. Henderson, J. Townend, and H. Zhang. Structural heterogeneity of the midcrust adjacent to the central Alpine Fault, New Zealand: Inferences from seismic tomography and seismicity between Harihari and Ross. *Geochemistry, Geophysics, Geosystems*, 16(4):1017–1043, 2015. doi: 10.1002/2014GC005702.
- J. R. Brown, G. C. Beroza, and D. R. Shelly. An autocorrelation method to detect low frequency earthquakes within tremor. *Geophysical Research Letters*, 35(16):L16305, 2008. doi: 10.1029/2008GL034560.
- J. R. Brown, G. C. Beroza, S. Ide, K. Ohta, D. R. Shelly, S. Y. Schwartz, W. Rabbel, M. Thorwart, and H. Kao. Deep low-frequency earthquakes in tremor localize to the plate interface in multiple subduction zones. *Geophysical Research Letters*, 36(19):L19306, 2009. doi: 10.1029/2009GL040027.
- J. R. Brown, S. G. Prejean, G. C. Beroza, J. S. Gombert, and P. J. Haeussler. Deep low-frequency earthquakes in tectonic tremor along the Alaska-Aleutian subduction
-

- zone. *Journal of Geophysical Research Solid Earth*, 118(3):1079–1090, 2013. doi: 10.1029/2012JB009459.
- J. G. Caldwell and C. Frolich. Microearthquake study of the Alpine Fault near Haast, South Island, New Zealand. *Bulletin of the Seismological Society of America*, 65(5): 1097–1104, 1975.
- I. M. Calhaem, A. J. Haines, and M. A. Lowry. An intermediate-depth earthquake in the central region of the South Island used to determine a local crustal thickness. *New Zealand Journal of Geology and Geophysics*, 20(2):353–361, 1977.
- B. M. Carpenter, H. Kitajima, R. Sutherland, J. Townend, V. G. Toy, and D. M. Saffer. Hydraulic and acoustic properties of the active Alpine Fault, New Zealand: Laboratory measurements on DFDP-1 drill core. *Earth and Planetary Science Letters*, 390: 45–51, 2014. doi: 10.1016/j.epsl.2013.12.023.
- L. Carrington, D. Komatitsch, M. Laurenzano, M. M. Tikir, D. Michéa, N. Le Goff, A. Snavely, and J. Tromp. High-frequency Simulations of Global Seismic Wave Propagation Using SPECFEM3D\_GLOBE on 62K Processors. In *Proceedings of the 2008 ACM/IEEE Conference on Supercomputing, SC '08*, pages 60:1–60:11, Piscataway, NJ, USA, 2008. IEEE Press. ISBN 978-1-4244-2835-9.
- C. J. Chamberlain, D. R. Shelly, J. Townend, and T. A. Stern. Low-frequency earthquakes reveal punctuated slow slip on the deep extent of the Alpine Fault, New Zealand. *Geochemistry, Geophysics, Geosystems*, pages 2984—2999, 2014. doi: 10.1002/2014GC005436.
- K. Chao, Z. Peng, H. Gonzalez-Huizar, C. Aiken, B. Enescu, H. Kao, A. A. Velasco, K. Obara, and T. Matsuzawa. A Global search for triggered tremor following the 2011 Mw 9.0 Tohoku earthquake. *Bulletin of the Seismological Society of America*, 103(2 B):1551–1571, 2013. doi: 10.1785/0120120171.
- N. I. Christensen and D. A. Okaya. Compressional and Shear Wave Velocities in South Island, New Zealand Rocks and Their Application to the Interpretation of Seismological Models of the New Zealand Crust. In D. Okaya, T. Stern, and F. Davey, editors, *A Continental Plate Boundary: Tectonics at South Island, New Zealand. Geophys. Mono. Ser.*, volume 175 of *Geophysical Monograph Series*, pages 123–155. American Geophysical Union, Washington, D. C., 2007. ISBN 978-0-87590-440-5. doi: 10.1029/GM175.
- H. V. Colella, J. H. Dieterich, K. Richards-Dinger, and A. M. Rubin. Complex characteristics of slow slip events in subduction zones reproduced in multi-cycle simulations. *Geophysical Research Letters*, 39(20):1–5, 2012. doi: 10.1029/2012GL053276.

- 
- S. C. Cox and D. J. A. Barrell. Geology of the Aoraki area. Technical report, Geological and Nuclear Sciences, 2007.
- S. C. Cox and R. H. Findlay. The Main Divide Fault Zone and its role in formation of the Southern Alps, New Zealand. *New Zealand Journal of Geology and Geophysics*, 38(4):489–499, 1995. doi: 10.1080/00288306.1995.9514675.
- S. C. Cox and R. Sutherland. Regional geological framework of South Island, New Zealand, and its significance for understanding the active plate boundary. In D. Okaya, T. Stern, and F. Davey, editors, *A Continental Plate Boundary: Tectonics at South Island, New Zealand. Geophys. Mono. Ser.*, volume 175 of *Geophysical Monograph Series*, pages 271–288. American Geophysical Union, Washington, D. C., 2007. ISBN 978-0-87590-440-5. doi: 10.1029/GM175.
- S. C. Cox, C. D. Menzies, R. Sutherland, P. H. Denys, C. Chamberlain, and D. A. H. Teagle. Changes in hot spring temperature and hydrogeology of the Alpine Fault hanging wall, New Zealand, induced by distal South Island earthquakes. *Geofluids*, 15(1-2):216–239, 2015. doi: 10.1111/gfl.12093.
- K. C. Creager. Coral 2.1, 1996.
- E. J. Delahaye, J. Townend, M. E. Reyners, and G. Rogers. Microseismicity but no tremor accompanying slow slip in the Hikurangi subduction zone, New Zealand. *Earth and Planetary Science Letters*, 277(1-2):21–28, 2009. doi: 10.1016/j.epsl.2008.09.038.
- C. DeMets, R. G. Gordon, D. F. Angus, and S. Stein. Effect of Recent Revisions to the Geomagnetic Reversal Time Scale on Estimates of Current Plate Motions. *Geophysical Research Letters*, 21(20):2191–2194, 1994.
- D. A. Dodge and W. R. Walter. Initial Global Seismic Cross-Correlation Results: Implications for Empirical Signal Detectors. *Bulletin of the Seismological Society of America*, 105(1):240–256, 2015. doi: 10.1785/0120140166.
- A. Douglas, J. Beavan, L. Wallace, and J. Townend. Slow slip on the northern hikurangi subduction interface, new zealand. *Geophysical Research Letters*, 32(16), 2005. doi: 10.1029/2005GL023607.
- H. Dragert, K. Wang, and T. S. James. A silent slip event on the deeper Cascadia subduction interface. *Science (New York, N.Y.)*, 292(5521):1525–1528, 2001. doi: 10.1126/science.1060152.
- H. Dragert, K. Wang, and G. Rogers. Geodetic and seismic signatures of episodic tremor and slip in the northern Cascadia subduction zone. *Earth, Planets and Space*, 56(12):1143–1150, 2004. ISSN 13438832.
-

- D. Eberhart-Phillips and S. Bannister. Three-dimensional crustal structure in the Southern Alps region of New Zealand from inversion of local earthquake and active source data. *Journal of Geophysical Research*, 107(B10), 2002. doi: 10.1029/2001JB000567.
- D. Eberhart-Phillips, M. Chadwick, and S. Bannister. Three-dimensional attenuation structure of central and southern South Island, New Zealand, from local earthquakes. *Journal of Geophysical Research*, 113(B5), 2008. doi: 10.1029/2007JB005359.
- J. D. Eccles, A. K. Gulley, P. E. Malin, C. M. Boese, J. Townend, and R. Sutherland. Fault Zone Guided Wave generation on the locked, late interseismic Alpine Fault, New Zealand. *Geophysical Research Letters*, pages 5736–5743, 2015. doi: 10.1002/2015GL064208.
- F. F. Evison. Seismicity of the Alpine Fault, New Zealand. *Bulletin of the Royal Society of NZ*, 9:161–165, 1971.
- Å. Fagereng, G. W. B. Hillary, and J. F. A. Diener. Brittle-viscous deformation, slow slip, and tremor. *Geophysical Research Letters*, 41:4159–4167, 2014. doi: 10.1002/2014GL060433.1.
- J. Feenstra, C. H. Thurber, J. Townend, S. Roecker, S. Bannister, C. Boese, N. Lord, S. Bourguignon, and D. Eberhart-Phillips. Microseismicity and P-wave tomography of the central Alpine Fault, New Zealand. *New Zealand Journal of Geology and Geophysics*, In Review.
- W. B. Frank and N. M. Shapiro. Automatic detection of low-frequency earthquakes (LFEs) based on a beamformed network response. *Geophysical Journal International*, 197(2):1215–1223, 2014. doi: 10.1093/gji/ggu058.
- W. B. Frank, N. M. Shapiro, V. Kostoglodov, A. L. Husker, M. Campillo, J. S. Payero, and G. A. Prieto. Low-frequency earthquakes in the Mexican Sweet Spot. *Geophysical Research Letters*, 40(11):2661–2666, 2013. doi: 10.1002/grl.50561.
- W. B. Frank, N. M. Shapiro, A. L. Husker, V. Kostoglodov, A. Romanenko, and M. Campillo. Using systematically characterized low-frequency earthquakes as a fault probe in Guerrero, Mexico. *Journal of Geophysical Research: Solid Earth*, pages 1–15, 2014. doi: 10.1002/2014JB011457.
- W. B. Frank, M. Radiguet, B. Rousset, N. M. Shapiro, A. L. Husker, V. Kostoglodov, N. Cotte, and M. Campillo. Uncovering the geodetic signature of silent slip through repeating earthquakes. *Geophysical Research Letters*, 42, 2015a. doi: 10.1002/2015GL063685.

- 
- W. B. Frank, N. M. Shapiro, A. L. Husker, V. Kostoglodov, H. S. Bhat, and M. Campillo. Along-fault pore-pressure evolution during a slow-slip event in Guerrero, Mexico. *Earth and Planetary Science Letters*, 413:135–143, 2015b. doi: 10.1016/j.epsl.2014.12.051.
- B. Fry, K. Chao, S. Bannister, Z. Peng, and L. Wallace. Deep tremor in New Zealand triggered by the 2010 Mw8.8 Chile earthquake. *Geophysical Research Letters*, 38(15):1–5, 2011. doi: 10.1029/2011GL048319.
- L. Geiger. Determination of seismic centres. *Nachrichten von der Koniglicher Gesellschaft der Wissenschaften zu Gottingen, Mathematisch-Physikalische Klasse*, 4:331–349, 1910.
- A. Ghosh, J. E. Vidale, Z. Peng, K. C. Creager, and H. Houston. Complex nonvolcanic tremor near Parkfield, California, triggered by the great 2004 Sumatra earthquake. *Journal of Geophysical Research*, 114:1–12, 2009. doi: 10.1029/2008JB006062.
- A. Ghosh, J. E. Vidale, J. R. Sweet, K. C. Creager, A. G. Wech, H. Houston, and E. E. Brodsky. Rapid, continuous streaking of tremor in Cascadia. *Geochemistry, Geophysics, Geosystems*, 11(12):1–10, 2010. doi: 10.1029/2010GC003305.
- S. J. Gibbons and F. Ringdal. The detection of low magnitude seismic events using array-based waveform correlation. *Geophysical Journal International*, 165(1):149–166, 2006. doi: 10.1111/j.1365-246X.2006.02865.x.
- J. Gomberg. The failure of earthquake failure models. *Journal of Geophysical Research*, 106:16,253–16,263, 2001.
- J. Gomberg and S. Prejean. Triggered tremor sweet spots in Alaska. *Journal of Geophysical Research: Solid Earth*, 118:6203–6218, 2013. doi: 10.1002/2013JB010273.
- J. Gomberg, E. Roeloffs, K. Hodgkinson, D. Schmidt, P. Bodin, J. McGuire, B. Crowell, and S. Minson. Learning about modes and scaling of fault slip from in situ, fault-scale observations. Talk at 2014 Cargese, Corsica workshop on: Earthquakes: nucleation, triggering and relationship with aseismic processes, 2014.
- G. Grünthal. Induced seismicity related to geothermal projects versus natural tectonic earthquakes and other types of induced seismic events in Central Europe. *Geothermics*, 52:22–35, 2014. doi: 10.1016/j.geothermics.2013.09.009.
- R. E. Haberman. Precursory seismicity patterns: stalking the mature seismic gap. *Earthquake Prediction: An International Review*, 4:29–42, 1981.
- A. J. Haines. A local magnitude scale for New Zealand earthquakes. *Bulletin of the Seismological Society of America*, 71(1):275–294, 1981.
-

- A. J. Haines, I. M. Calhaem, and D. E. Ware. Crustal seismicity near Lake Pukaki, South Island. *The origin of the Southern Alps. The Royal Society of New Zealand bulletin*, 18:87–94, 1979.
- T. C. Hanks and H. Kanamori. A moment magnitude scale. *Journal of Geophysical Research*, 84(B5):2348–2350, 1979. doi: 10.1029/JB084iB05p02348.
- D. Harris. Subspace detectors: Theory. *Lawrence Livermore Natl. Lab. Rep. UCRL-TR-222758*, 2006.
- J. Havskov and L. Ottemoller. SEISAN earthquake analysis software. *Seismological Research Letters*, 70:532–534, 1999.
- J. Havskov and L. Ottemoller. *Routine Data Processing in Earthquake Seismology*. Springer Netherlands, Dordrecht, 2010. ISBN 978-90-481-8696-9. doi: 10.1007/978-90-481-8697-6.
- F. Herman, S. C. Cox, and P. J. J. Kamp. Low-temperature thermochronology and thermokinematic modeling of deformation, exhumation, and development of topography in the central Southern Alps, New Zealand. *Tectonics*, 28(5), 2009. doi: 10.1029/2008TC002367.
- D. P. Hill, Z. Peng, D. R. Shelly, and C. Aiken. S-wave triggering of tremor beneath the Parkfield, California, section of the San Andreas fault by the 2011 Tohoku, Japan, earthquake: Observations and theory. *Bulletin of the Seismological Society of America*, 103(2 B):1541–1550, 2013. doi: 10.1785/0120120114.
- C. Holden. Kinematic Source Model of the 22 February 2011 Mw 6.2 Christchurch Earthquake Using Strong Motion Data. *Seismological Research Letters*, 82(6):783–788, 2011. doi: 10.1785/gssrl.82.6.783.
- H. J. Horgan, B. Anderson, R. B. Alley, C. J. Chamberlain, R. Dykes, L. M. Kehrl, and J. Townend. Glacier velocity variability due to rain-induced sliding and cavity formation. *Earth and Planetary Science Letters*, 432:273–282, 2015. doi: 10.1016/j.epsl.2015.10.016.
- N. Houlié and T. Stern. A comparison of GPS solutions for strain and SKS fast directions: Implications for modes of shear in the mantle of a plate boundary zone. *Earth and Planetary Science Letters*, 345-348:117–125, 2012. doi: 10.1016/j.epsl.2012.06.029.
- H. Houston, B. G. Delbridge, A. G. Wech, and K. C. Creager. Rapid tremor reversals in Cascadia generated by a weakened plate interface. *Nature Geoscience*, 4(6):404–409, 2011. doi: 10.1038/ngeo1157.

- J. D. Howarth, S. J. Fitzsimons, R. J. Norris, and G. E. Jacobsen. Lake sediments record high intensity shaking that provides insight into the location and rupture length of large earthquakes on the Alpine Fault, New Zealand. *Earth and Planetary Science Letters*, 403:340–351, 2014. doi: 10.1016/j.epsl.2014.07.008.
- S. Ide. A Brownian walk model for slow earthquakes. *Geophysical Research Letters*, 35(17):3–7, 2008. doi: 10.1029/2008GL034821.
- S. Ide and S. Yabe. Universality of slow earthquakes in the very low frequency band. *Geophysical Research Letters*, 41, 2014. doi: 10.1002/2014GL059712.
- S. Ide, G. C. Beroza, D. R. Shelly, and T. Uchide. A scaling law for slow earthquakes. *Nature*, 447, 2007a. doi: 10.1038/nature05780.
- S. Ide, D. R. Shelly, and G. C. Beroza. Mechanism of deep low frequency earthquakes: Further evidence that deep non-volcanic tremor is generated by shear slip on the plate interface. *Geophysical Research Letters*, 34(3), 2007b. doi: 10.1029/2006GL028890.
- T. Igarashi, T. Matsuzawa, and A. Hasegawa. Repeating earthquakes and interplate aseismic slip in the northeastern japan subduction zone. *Journal of Geophysical Research: Solid Earth*, 108(B5), 2003. doi: 10.1029/2002JB001920. 2249.
- M. J. Ikari, Y. Ito, K. Ujiie, and A. J. Kopf. Spectrum of slip behaviour in tohoku fault zone samples at plate tectonic slip rates. *Nature Geoscience*, 8, 2015a. doi: 10.1038/NGEO2547.
- M. J. Ikari, S. Trütner, B. M. Carpenter, and A. J. Kopf. Shear behavior of DFDP-1 borehole samples from the Alpine Fault, New Zealand, under a wide range of experimental conditions. *International Journal of Earth Science*, 2015b. doi: 10.1007/s00531-014-1115-5.
- Y. Ito, K. Obara, K. Shiomi, S. Sekine, and H. Hirose. Slow Earthquakes Coincident with Episodic Tremors and Slow Slip Events. *Science*, 315(5811):503–506, 2007. doi: 10.1126/science.1134454.
- Y. Ito, R. Hino, M. Kido, H. Fujimoto, Y. Osada, D. Inazu, Y. Ohta, T. Iinuma, M. Ohzono, S. Miura, M. Mishina, K. Suzuki, T. Tsuji, and J. Ashi. Episodic slow slip events in the Japan subduction zone before the 2011 Tohoku-Oki earthquake. *Tectonophysics*, 600:14–26, 2013. doi: 10.1016/j.tecto.2012.08.022.
- Y. Ito, R. Hino, S. Suzuki, and Y. Kaneda. Episodic tremor and slip near the Japan Trench prior to the 2011 Tohoku-Oki earthquake. *Geophysical Research Letters*, pages 1725–1731, 2015. doi: 10.1002/2014GL062986.

- Y. Y. Kagan. Seismic moment distribution revisited: I. Statistical results. *Geophysical Journal International*, 148(3):520–541, 2002. doi: 10.1046/j.1365-246x.2002.01594.x.
- H. Kao and S. J. Shan. The Source-Scanning Algorithm: Mapping the distribution of seismic sources in time and space. *Geophysical Journal International*, 157(2): 589–594, 2004. doi: 10.1111/j.1365-246X.2004.02276.x.
- H. Kao, S. J. Shan, H. Dragert, and G. Rogers. Northern Cascadia episodic tremor and slip: A decade of tremor observations from 1997 to 2007. *Journal of Geophysical Research*, 114(June):B00A12, 2009. doi: 10.1029/2008JB006046.
- B. M. Kaproth and C. Marone. Slow earthquakes, preseismic velocity changes, and the origin of slow frictional stick-slip. *Science*, 341(6151):1229–32, 2013. doi: 10.1126/science.1239577.
- A. Kato, K. Obara, T. Igarashi, H. Tsuruoka, S. Nakagawa, and N. Hirata. Propagation of Slow Slip Leading Up to the 2011 Mw 9.0 Tohoku-Oki Earthquake. *Science*, 335 (6069):705–708, 2012. ISSN 0036-8075. doi: 10.1126/science.1215141.
- M. D. Kohler and D. Eberhart-Phillips. Intermediate-Depth Earthquakes in a Region of Continental Convergence: South Island, New Zealand. *Bulletin of the Seismological Society of America*, 93(1):85–93, 2003. doi: 10.1785/0120020043.
- P. O. Koons, R. J. Norris, D. Craw, and A. F. Cooper. Influence of exhumation on the structural evolution of transpressional plate boundaries: An example from the Southern Alps, New Zealand. *Geology*, 31(1):3–6, 2003. doi: 10.1130/0091-7613(2003)031<0003.
- S. Lamb and E. Smith. The nature of the plate interface and driving force of interseismic deformation in the New Zealand plate-boundary zone, revealed by the continuous GPS velocity field. *Journal of Geophysical Research: Solid Earth*, 118, 2013. doi: 10.1002/jgrb.50221.
- S. Lamb, E. Smith, T. Stern, and E. Warren-Smith. Continent-scale strike-slip on a low-angle fault beneath New Zealand’s Southern Alps: Implications for crustal thickening in oblique collision zones. *Geochemistry, Geophysics, Geosystems*, 2015. doi: 10.1002/2015GC005990.
- W. H. K. Lee, R. E. Bennett, and L. Meagher. A method for estimating magnitude of local earthquakes from signal duration. Technical report, Open file report, USGS, 1972.
- B. Leitner, D. Eberhart-Phillips, H. Anderson, and J. L. Nabelek. A focused look at the Alpine fault, New Zealand: Seismicity, focal mechanisms, and stress observations. *Journal of Geophysical Research*, 106(B2):2193, 2001. doi: 10.1029/2000JB900303.

- 
- B. R. Lienert, E. Berg, and L. N. Frazer. Hypocenter: an earthquake location method using centered, scaled, and adaptively damped least squares. *Bulletin of the Seismological Society of America*, 76(3):771–783, 1986.
- T. A. Little, M. K. Savage, and B. Tikoff. Relationship between crustal finite strain and seismic anisotropy in the mantle, Pacific-Australia plate boundary zone, South Island, New Zealand. *Geophysical Journal International*, 151(1):106–116, 2002. doi: 10.1046/j.1365-246X.2002.01730.x.
- T. A. Little, S. Cox, J. K. Vry, and G. Batt. Variations in exhumation level and uplift rate along the obliqu-slip Alpine fault, central Southern Alps, New Zealand. *Geological Society of America Bulletin*, 117(5):707, 2005. doi: 10.1130/B25500.1.
- T. A. Little, R. H. Wightman, R. J. Holcombe, and M. Hill. Transpression models and ductile deformation of the lower crust of the pacific plate in the central Southern Alps, a perspective from structural geology. In D. Okaya, T. Stern, and F. Davey, editors, *A Continental Plate Boundary: Tectonics at South Island, New Zealand. Geophys. Mono. Ser.*, volume 175 of *Geophysical Monograph Series*, pages 271–288. American Geophysical Union, Washington, D. C., 2007. ISBN 978-0-87590-440-5. doi: 10.1029/GM175.
- D. A. Lockner, C. Morrow, D. Moore, and S. Hickman. Low strength of deep San Andreas fault gouge from SAFOD core. *Nature*, 472(7341):82–5, 2011. doi: 10.1038/nature09927.
- A. Lomax, J. Virieux, P. Volant, and C. Berge. Probabilistic earthquake location in 3D and layered models: Introduction of a Metropolis-Gibbs method and comparison with linear locations. In C. Thurber and N. Rabinowitz, editors, *Advances in Seismic Event Location*, pages 101–134. Kluwer, Amsterdam, 2000.
- A. Lomax, C. Satriano, and M. Vassallo. Automatic Picker Developments and Optimization: FilterPicker—a Robust, Broadband Picker for Real-Time Seismic Monitoring and Earthquake Early Warning. *Seismological Research Letters*, 83(3):531–540, 2012. doi: 10.1785/gssrl.83.3.531.
- E. L. Majer, R. Baria, M. Stark, S. Oates, J. Bommer, B. Smith, and H. Asanuma. Induced seismicity associated with Enhanced Geothermal Systems. *Geothermics*, 36(3):185–222, 2007. doi: 10.1016/j.geothermics.2007.03.003.
- D. Marsan. The role of small earthquakes in redistributing crustal elastic stress. *Geophysical Journal International*, 163(1):141–151, 2005. doi: 10.1111/j.1365-246X.2005.02700.x.
-

- G. Matharu, M. G. Bostock, N. I. Christensen, and J. Tromp. Crustal anisotropy in a subduction zone forearc: Northern cascadia. *Journal of Geophysical Research: Solid Earth*, 119:7058–7078, 2014. doi: 10.1002/2014JB011321.
- A. P. Mavrommatis, P. Segall, N. Uchida, and K. M. Johnson. Long-term acceleration of aseismic slip preceding the Mw 9 Tohoku-oki earthquake: Constraints from repeating earthquakes. *Geophysical Reserach Letters*, 42:9717–9725, 2015. doi: 10.1002/2015GL066069.
- G. C. McLaskey and B. D. Kilgore. Foreshocks during the nucleation of stick-slip instability. *Journal of Geophysical Research: Solid Earth*, 118(6):2982–2997, 2013. doi: 10.1002/jgrb.50232.
- C. D. Menzies, D. A. H. Teagle, D. Craw, S. C. Cox, A. J. Boyce, C. D. Barrie, and S. Roberts. Incursion of meteoric waters into the ductile regime in an active orogen. *Earth and Planetary Science Letters*, 399:1–13, 2014. doi: 10.1016/j.epsl.2014.04.046.
- P. Molnar, H. J. Anderson, E. Audoiné, D. Eberhart-Phillips, K. R. Gledhill, E. R. Klosko, T. V. McEvelly, D. Okaya, M. K. Savage, T. Stern, and F. T. Wu. Continuous deformation versus faulting through the continental lithosphere of New Zealand. *Science*, 286(5439):516–9, 1999. ISSN 1095-9203.
- N. Mortimer. New Zealand’s Geological Foundations. *Gondwana Research*, 1:261–272, 2004.
- R. M. Nadeau and D. Dolenc. Nonvolcanic tremors deep beneath the San Andreas Fault. *Science*, 307:389, 2005. doi: 10.1126/science.1107142.
- R. M. Nadeau and L. R. Johnson. Seismological studies at parkfield vi: Moment release rates and estimates of source parameters for small repeating earthquakes. *Bulletin of the Seismological Society of America*, 88(3):790–814, 1998.
- R. M. Nadeau and T. V. McEvelly. Periodic pulsing of characteristic microearthquakes on the San Andreas fault. *Science*, 303(5655):220–2, 2004. doi: 10.1126/science.1090353.
- S. Nakaya. Spatiotemporal variation in b value within the subducting slab prior to the 2003 Tokachi-oki earthquake (M 8.0), Japan. *Journal of Geophysical Research-Solid Earth*, 111(B3), 2006. doi: 10.1029/2005JB003658.
- R. J. Norris and A. F. Cooper. Origin of small-scale segmentation and transpressional thrusting along the Alpine Fault, New Zealand. *Geological Society of America Bulletin*, 107(2):231–240, 1995. doi: 10.1130/0016-7606(1995)107<0231:OOSSSA>2.3.CO;2.

- R. J. Norris and A. F. Cooper. Late Quaternary slip rates and slip partitioning on the Alpine Fault, New Zealand. *Journal of Structural Geology*, 23(2000):507–520, 2000.
- R. J. Norris and A. F. Cooper. The Alpine Fault, New Zealand: Surface Geology and Field Relationships. In D. Okaya, T. Stern, and F. Davey, editors, *A Continental Plate Boundary: Tectonics at South Island, New Zealand. Geophys. Mono. Ser.*, volume 175 of *Geophysical Monograph Series*, pages 157–175. American Geophysical Union, Washington, D. C., 2007. ISBN 978-0-87590-440-5. doi: 10.1029/GM175.
- R. J. Norris and V. G. Toy. Continental transforms: A view from the Alpine Fault. *Journal of Structural Geology*, 64:3–31, 2014. doi: 10.1016/j.jsg.2014.03.003.
- K. Obara. Nonvolcanic Deep Tremor Associated with Subduction in Southwest Japan. *Science*, 296(5573):1679–1681, 2002.
- D. Okaya, T. Stern, and F. Davey, editors. *A continental Plate Boundary: Tectonics at South Island, New Zealand.*, volume 175 of *Geophysical Monograph Series*. American Geophysical Union, Washington, D. C., 2007. ISBN 978-0-87590-440-5. doi: 10.1029/GM175.
- B. C. O’Keefe. *Microseismicity of the Central Alpine Fault Region, New Zealand*. Msc thesis, Victoria Univ. of Wellington, Wellington, New Zealand, 2008.
- J. S. Payero, V. Kostoglodov, N. Shapiro, T. Mikumo, A. Iglesias, X. Pérez-Campos, and R. W. Clayton. Nonvolcanic tremor observed in the Mexican subduction zone. *Geophysical Research Letters*, 35(7), 2008. doi: 10.1029/2007GL032877.
- Z. Peng and K. Chao. Non-volcanic tremor beneath the Central Range in Taiwan triggered by the 2001 Mw 7.8 Kunlun earthquake. *Geophysical Journal International*, 175(2):825–829, 2008. doi: 10.1111/j.1365-246X.2008.03886.x.
- Z. Peng and J. Gomberg. An integrated perspective of the continuum between earthquakes and slow-slip phenomena. *Nature Geoscience*, 3(9):599–607, 2010. doi: 10.1038/ngeo940.
- Z. Peng, J. E. Vidale, A. G. Wech, R. M. Nadeau, and K. C. Creager. Remote triggering of tremor along the San Andreas Fault in central California. *Journal of Geophysical Research*, 114:1–18, 2009. doi: 10.1029/2008JB006049.
- T. Petersen, K. Gledhill, M. Chadwick, N. H. Gale, and J. Ristau. The New Zealand national seismograph network. *Seismological Research Letters*, 82(1):9–20, 2011. doi: 10.1785/gssrl.82.1.9.
- M. Quigley, R. Van Dissen, N. Litchfield, P. Villamor, B. Dufft, D. Barrell, K. Furlong, T. Stahl, E. Bilderback, and D. Noble. Surface rupture during the 2010 Mw

- 7.1 Darfield (Canterbury) earthquake: Implications for fault rupture dynamics and seismic-hazard analysis. *Geology*, 2012. doi: 10.1130/G32528.1.
- Z. Rawlinson. *Microseismicity associated with actively exploited geothermal systems: earthquake detection and probabilistic location at Rotokawa and statistical seismic network design at Kawarau*. Msc thesis, Victoria University of Wellington, Wellington, New Zealand, 2011.
- P. A. Reasenber and M. V. Matthews. Precursory seismic quiescence: A preliminary assessment of the hypothesis. *Pure and Applied Geophysics*, 126(2-4):373–406, 1988. doi: 10.1007/BF00879004.
- M. Reyners. Subcrustal earthquakes in the central South Island, New Zealand, and the root of the Southern Alps. *Geology*, pages 1168–1171, 1987.
- M. Reyners. Reservoir-induced seismicity at Lake Pukaki, New Zealand. *Geophysical Journal International*, pages 127–135, 1988.
- K. B. Richards-Dinger and J. H. Dieterich. RSQSim Earthquake Simulator. *Seismological Research Letters*, 83(6):983–990, 2012. doi: 10.1785/0220120105.
- F. Ringdal, D. B. Harris, D. Dodge, and S. J. Gibbons. Adaptive waveform correlation detectors for arrays: algorithms for autonomous calibration. *Lawrence Livermore National Laboratory Report UCRL-TR-222758*, 2009.
- J. Ristau. Implementation of Routine Regional Moment Tensor Analysis in New Zealand. *Seismological Research Letters*, 79(3):400–415, 2008. doi: 10.1785/gssrl.79.3.400.
- J. Ristau. Comparison of Magnitude Estimates for New Zealand Earthquakes: Moment Magnitude, Local Magnitude, and Teleseismic Body-Wave Magnitude. *Bulletin of the Seismological Society of America*, 99(3):1841–1852, 2009. doi: 10.1785/0120080237.
- J. Ristau, D. Harte, and J. Salichon. A revised local magnitude (ml) scale for new zealand earthquakes, 2016.
- G. Rogers and H. Dragert. Episodic tremor and slip on the Cascadia subduction zone: The chatter of silent slip. *Science*, 300(5627):1942–1943, 2003.
- S. Rost. Array seismology: Methods and applications. *Reviews of Geophysics*, 40(3):1008, 2002. ISSN 8755-1209. doi: 10.1029/2000RG000100.
- A. A. Royer and M. G. Bostock. A comparative study of low frequency earthquake templates in northern Cascadia. *Earth and Planetary Science Letters*, 1:1–10, 2013. doi: 10.1016/j.epsl.2013.08.040.

- 
- A. M. Rubin and J. G. Armbruster. Imaging slow slip fronts in Cascadia with high precision crossstation tremor locations. *Geochemistry, Geophysics, Geosystems*, 14(12):5371–5392, 2013. doi: 10.1002/2013GC005031.
- J. L. Rubinstein and W. L. Ellsworth. Precise Estimation of Repeating Earthquake Moment: Example from Parkfield, California. *Bulletin of the Seismological Society of America*, 100(5A):1952–1961, 2010. doi: 10.1785/0120100007.
- J. L. Rubinstein, J. Gomberg, J. E. Vidale, A. G. Wech, H. Kao, K. C. Creager, and G. Rogers. Seismic wave triggering of nonvolcanic tremor, episodic tremor and slip, and earthquakes on Vancouver Island. *Journal of Geophysical Research*, 114:B00A01, 2009. doi: 10.1029/2008JB005875.
- S. Ruiz, M. Metois, A. Fuenzalida, J. Ruiz, F. Leyton, R. Grandin, C. Vigny, R. Madariaga, and J. Campos. Intense foreshocks and a slow slip event preceded the 2014 Iquique Mw 8.1 earthquake. *Science*, 345(6201), 2014. doi: 10.1126/science.1256074.
- D. M. Saffer, D. A. Lockner, and A. McKiernan. Effects of smectite to illite transformation on the frictional strength and sliding stability of intact marine mudstones. *Geophysical Research Letters*, 39(11):2–7, 2012. doi: 10.1029/2012GL051761.
- D. P. Schaff and P. G. Richards. On finding and using repeating seismic events in and near China. *Journal of Geophysical Research: Solid Earth*, 116(3):1–20, 2011. doi: 10.1029/2010JB007895.
- D. P. Schaff, G. C. Beroza, and B. E. Shaw. Postseismic response of repeating aftershocks. *Geophysical Research Letters*, 25(24):4549–4552, 1998. doi: 10.1029/1998GL900192.
- M. Scherwath, T. Stern, F. Davey, D. Okaya, W. Holbrook, R. Davies, and S. Kleffmann. Lithospheric structure across oblique continental collision in New Zealand from wide-angle P wave modeling. *Journal of Geophysical Research*, 108(B12):2566, 2003. doi: 10.1029/2002JB002286.
- A. M. Schleicher, R. Sutherland, J. Townend, V. G. Toy, and B. van der Pluijm. Clay mineral formation and fabric development in the DFDP-1B borehole, central Alpine Fault, New Zealand. *New Zealand Journal of Geology and Geophysics*, 58(1):13–21, 2015. doi: 10.1080/00288306.2014.979841.
- C. H. Scholz, J. M. W. Rynn, R. W. Weed, and C. Frohlich. Detailed seismicity of the Alpine fault zone and Fiordland region, New Zealand. *Geological Society of America Bulletin*, 84(October):3297–3316, 1973.
-

- B. E. Shaw, J. M. Carlson, and J. S. Langer. Patterns of seismic activity preceding large earthquakes. *Journal of Geophysical Research: Solid Earth*, 97(B1):479–488, 1992. doi: 10.1029/91JB01796.
- D. R. Shelly. Possible deep fault slip preceding the 2004 Parkfield earthquake, inferred from detailed observations of tectonic tremor. *Geophysical Research Letters*, 36(17): 1–6, 2009. doi: 10.1029/2009GL039589.
- D. R. Shelly. Migrating tremors illuminate complex deformation beneath the seismogenic San Andreas fault. *Nature*, 463(7281):648–652, 2010. doi: 10.1038/nature08755.
- D. R. Shelly and J. L. Hardebeck. Precise tremor source locations and amplitude variations along the lower-crustal central San Andreas Fault. *Geophysical Research Letters*, 37(14), 2010. doi: 10.1029/2010GL043672.
- D. R. Shelly, G. C. Beroza, S. Ide, and S. Nakamura. Low-frequency earthquakes in Shikoku, Japan, and their relationship to episodic tremor and slip. *Nature*, 442(7099):188–91, 2006. doi: 10.1038/nature04931.
- D. R. Shelly, G. C. Beroza, and S. Ide. Non-volcanic tremor and low-frequency earthquake swarms. *Nature*, 446(7133):305–7, 2007. doi: 10.1038/nature05666.
- D. R. Shelly, W. L. Ellsworth, T. Ryber, C. Haberland, G. S. Fuis, J. Murphy, R. M. Nadeau, and R. Bürgmann. Precise location of San Andreas Fault tremors near Cholame, California using seismometer clusters: Slip on the deep extension of the fault? *Geophysical Research Letters*, 36(1), 2009. doi: 10.1029/2008GL036367.
- D. R. Shelly, D. P. Hill, F. Massin, J. Farrell, R. B. Smith, and T. Taira. A fluid-driven earthquake swarm on the margin of the Yellowstone caldera. *Journal of Geophysical Research: Solid Earth*, 118(9):4872–4886, 2013. doi: 10.1002/jgrb.50362.
- R. H. Sibson, S. H. White, and B. K. Atkinson. Fault rock distribution and structure within the Alpine Fault zone: a preliminary account. *Royal Society of New Zealand Bulletin*, 18:55–65, 1979.
- E. G. C. Smith, T. Stern, and B. O’Brien. A seismic velocity profile across the central South Island, New Zealand, from explosion data. *New Zealand Journal of Geology and Geophysics*, pages 37–41, 1995.
- C. Snijders, U. Matzat, and U. Reips. ”Big Data”: Big Gaps of Knowledge in the Field of Internet Science. *International Journal of Internet Science*, 7(1):1–5, 2013. ISSN 1662-5544.

- 
- T. Stern, S. Kleffmann, D. Okaya, M. Scherwath, and S. Bannister. Low seismic-wave speeds and enhanced fluid pressure beneath the Southern Alps of New Zealand. *Geology*, 2001. doi: 10.1130/0091-7613(2001)029<0679.
- T. Stern, D. Okaya, S. Kleffmann, M. Scherwath, S. Henrys, and F. Davey. Geophysical exploration and dynamics of the Alpine Fault zone. In D. Okaya, T. Stern, and F. Davey, editors, *A Continental Plate Boundary: Tectonics at South Island, New Zealand*, in *Geophys. Mono. Ser.*, volume 175 of *Geophysical Monograph Series*, pages 207–233. American Geophysical Union, Washington, D. C., 2007. ISBN 978-0-87590-440-5. doi: 10.1029/GM175.
- M. Stirling, G. McVerry, M. Gerstenberger, N. Litchfield, R. Van Dissen, K. Berryman, P. Barnes, L. Wallace, P. Villamor, R. Langridge, G. Lamarche, S. Nodder, M. Reyners, B. Bradley, D. Rhoades, W. Smith, A. Nicol, J. Pettinga, K. Clark, and K. Jacobs. National seismic hazard model for New Zealand: 2010 update. *Bulletin of the Seismological Society of America*, 102(4):1514–1542, 2012. doi: 10.1785/0120110170.
- R. Sutherland. Transpressional development of the Australia-Pacific boundary through southern South Island, New Zealand: Constraints from Miocene-Pliocene sediments, Waiho-1 borehole, South Westland. *New Zealand Journal of Geology and Geophysics*, pages 37–41, 1996.
- R. Sutherland and R. J. Norris. Late Quaternary displacement rate, paleoseismicity, and geomorphic evolution of the Alpine Fault: Evidence from Hokuri Creek, South Westland, New Zealand. *New Zealand Journal of Geology and Geophysics*, 38 (November 2014):419–430, 1995. doi: 10.1080/00288306.1995.9514669.
- R. Sutherland, F. Davey, and J. Beavan. Plate boundary deformation in South Island, New Zealand, is related to inherited lithospheric structure. *Earth and Planetary Science Letters*, 177(3-4):141–151, 2000. doi: 10.1016/S0012-821X(00)00043-1.
- R. Sutherland, K. Berryman, and R. Norris. Quaternary slip rate and geomorphology of the Alpine fault: Implications for kinematics and seismic hazard in southwest New Zealand. *Geological Society of America Bulletin*, 118(3-4):464–474, 2006. ISSN 0016-7606. doi: 10.1130/B25627.1.
- R. Sutherland, D. Eberhart-Phillips, R. Harris, T. Stern, J. Beavan, S. Ellis, S. Henrys, S. Cox, R. Norris, K. Berryman, J. Townend, S. Bannister, J. Pettinga, B. Leitner, L. Wallace, T. Little, R. Cooper, M. Yetton, and M. Stirling. Do great earthquakes occur on the Alpine fault in central South Island, New Zealand? In D. Okaya, T. Stern, and F. Davey, editors, *A Continental Plate Boundary: Tectonics at South Island, New Zealand*. *Geophys. Mono. Ser.*, volume 175 of *Geophysical Monograph*
-

- Series*, pages 235–251. American Geophysical Union, Washington, D. C., 2007. ISBN 978-0-87590-440-5. doi: 10.1029/GM175.
- R. Sutherland, V. G. Toy, J. Townend, S. C. Cox, J. D. Eccles, D. R. Faulkner, D. J. Prior, R. J. Norris, E. Mariani, C. Boulton, B. M. Carpenter, C. D. Menzies, T. A. Little, M. Hasting, G. P. De Pascale, R. M. Langridge, H. R. Scott, Z. Reid Lindroos, B. Fleming, and A. J. Kopf. Drilling reveals fluid control on architecture and rupture of the Alpine fault, New Zealand. *Geology*, 40(12):1143–1146, 2012. doi: 10.1130/G33614.1.
- J. R. Sweet, K. C. Creager, and H. Houston. A family of repeating low-frequency earthquakes at the downdip edge of tremor and slip. *Geochemistry, Geophysics, Geosystems*, 15(9):3713–3721, 2014. doi: 10.1002/2014GC005449.
- A. Tarantola and B. Valette. Generalized nonlinear inverse problems solved using the least squares criterion. *Reviews of Geophysics and Space Physics*, 20(2):219–232, 1982.
- A. M. Thomas and M. G. Bostock. Identifying low-frequency earthquakes in central Cascadia using cross-station correlation. *Tectonophysics*, 658:111–116, 2015. doi: 10.1016/j.tecto.2015.07.013.
- A. M. Thomas, R. M. Nadeau, and R. Bürgmann. Tremor-tide correlations and near-lithostatic pore pressure on the deep San Andreas fault. *Nature*, 462(7276):1048–51, 2009. doi: 10.1038/nature08654.
- A. M. Thomas, R. Bürgmann, D. R. Shelly, N. M. Beeler, and M. L. Rudolph. Tidal triggering of low frequency earthquakes near Parkfield, California: Implications for fault mechanics within the brittle-ductile transition. *Journal of Geophysical Research*, 117(B5):1–24, 2012. doi: 10.1029/2011JB009036.
- A. M. Thomas, G. C. Beroza, and D. R. Shelly. Constraints on the source parameters of low-frequency earthquakes on the San Andreas Fault. *Geophysical Research Letters*, 43, 2016. doi: 10.1002/2015GL067173.
- T. W. Thomas, J. E. Vidale, H. Houston, K. C. Creager, J. R. Sweet, and A. Ghosh. Evidence for tidal triggering of high-amplitude rapid tremor reversals and tremor streaks in northern Cascadia. *Geophysical Research Letters*, 40(16):4254–4259, 2013. doi: 10.1002/grl.50832.
- C. H. Thurber. Nonlinear earthquake location: theory and examples. *Bulletin of the Seismological Society of America*, 75(3):779–790, 1985.

- C. H. Thurber, S. W. Roecker, J. Feenstra, N. Lord, G. O'Brien, J. D. Pesicek, S. C. Bannister, and J. Townend. Seismic site characterization for the Deep Fault Drilling Project (DFDP), Alpine Fault, New Zealand: Preliminary results from the WIZARD array. In *AGU Fall Meeting Abstracts*, 2012.
- J. Townend, R. Sutherland, and V. Toy. Deep Fault Drilling Project–Alpine Fault, New Zealand. *Scientific Drilling*, pages 75–82, 2009. ISSN 1816-3459. doi: 10.2204/iodp.sd.8.12.2009.
- J. Townend, R. Sutherland, V. G. Toy, J. D. Eccles, C. Boulton, S. C. Cox, and D. McNamara. Late-interseismic state of a continental plate-bounding fault: Petrophysical results from DFDP-1 wireline logging and core analysis, Alpine Fault, New Zealand. *Geochemistry, Geophysics, Geosystems*, 14(9):3801–3820, 2013. doi: 10.1002/ggge.20236.
- V. G. Toy, R. J. Norris, D. J. Prior, M. Walrond, and A. F. Cooper. How do lineations reflect the strain history of transpressive shear zones? The example of the active Alpine Fault zone, New Zealand. *Journal of Structural Geology*, 50:187–198, 2013. doi: 10.1016/j.jsg.2012.06.006.
- V. G. Toy, C. J. Boulton, R. Sutherland, J. Townend, R. J. Norris, T. A. Little, D. J. Prior, E. Mariani, D. Faulkner, C. D. Menzies, H. Scott, and B. M. Carpenter. Fault rock lithologies and architecture of the central Alpine fault, New Zealand, revealed by DFDP-1 drilling. *Lithosphere*, pages 155–173, 2015. doi: 10.1130/L395.1.
- N. Uchida, T. Matsuzawa, A. Hasegawa, and T. Igarashi. Interplate quasi-static slip off sanriku, ne japan, estimated from repeating earthquakes. *Geophysical Research Letters*, 30(15), 2003. doi: 10.1029/2003GL017452. 1801.
- T. Utheim, J. Havskov, M. Ozyazicioglu, J. Rodriguez, and E. Talavera. RTQUAKE, A Real-Time Earthquake Detection System Integrated with SEISAN. *Seismological Research Letters*, 85(3):735–742, 2014. doi: 10.1785/0220130175.
- T. Van Eck, F. Goutbeek, H. Haak, and B. Dost. Seismic hazard due to small-magnitude, shallow-source, induced earthquakes in The Netherlands. *Engineering Geology*, 87(1-2):105–121, 2006. doi: 10.1016/j.enggeo.2006.06.005.
- M. Vassallo, C. Satriano, and A. Lomax. Automatic Picker Developments and Optimization: A Strategy for Improving the Performances of Automatic Phase Pickers. *Seismological Research Letters*, 83(3):541–554, 2012. doi: 10.1785/gssrl.83.3.541.
- R. I. Walcott. Present tectonics and Late Cenozoic evolution of New Zealand. *Geophysical Journal International*, 52(1):137–164, 1978. doi: 10.1111/j.1365-246X.1978.tb04225.x.

- F. Waldhauser and W. L. Ellsworth. A double-difference earthquake location algorithm: method and application to the northern Hayward fault, California. *Bulletin of the Seismological Society of America*, pages 1353–1368, 2000.
- L. M. Wallace and J. Beavan. Diverse slow slip behavior at the Hikurangi subduction margin, New Zealand. *Journal of Geophysical Research: Solid Earth*, 115(12):1–20, 2010. doi: 10.1029/2010JB007717.
- D. Walsh, R. Arnold, and J. Townend. A Bayesian approach to determining and parametrizing earthquake focal mechanisms. *Geophysical Journal International*, 176(1):235–255, 2009. doi: 10.1111/j.1365-246X.2008.03979.x.
- R. J. Walters, M. D. Zoback, J. W. Baker, and G. C. Beroza. Characterizing and Responding to Seismic Risk Associated with Earthquakes Potentially Triggered by Fluid Disposal and Hydraulic Fracturing. *Seismological Research Letters*, 86(4):1–9, 2015. doi: 10.1785/0220150048.
- P. E. Wannamaker, G. J. Jiracek, J. A. Stodt, T. G. Caldwell, W. M. Gonzales, J. D. McKnight, and A. D. Porter. Fluid generation and pathways beneath an active compressional orogen, the New Zealand Southern Alps, inferred from magnetotelluric data. *Journal of Geophysical Research*, 107(B6):2117, 2002. doi: 10.1029/2001JB000186.
- E. Warren-Smith, S. H. Lamb, T. Stern, and C. J. Chamberlain. What Controls Slip Directions of Diffuse Microseismicity in a Zone of Continental Transpression, South Island, New Zealand? *AGU Fall Meeting Abstracts*, 2015.
- T. Watanabe, Y. Hiramatsu, and K. Obara. Scaling relationship between the duration and the amplitude of non-volcanic deep low-frequency tremors. *Geophysical Research Letters*, 34(7):L07305, 2007. doi: 10.1029/2007GL029391.
- A. G. Wech. Interactive tremor monitoring. *Seismological Research Letters*, 81(4):664–669, 2010. doi: 10.1785/gssrl.
- A. G. Wech and N. M. Bartlow. Slip rate and tremor genesis in Cascadia. *Geophysical Research Letters*, 41(2):392–398, 2014. doi: 10.1002/2013GL058607.
- A. G. Wech and K. C. Creager. Cascadia tremor polarization evidence for plate interface slip. *Geophysical Research Letters*, 34(22):L22306, 2007. doi: 10.1029/2007GL031167.
- A. G. Wech and K. C. Creager. Automated detection and location of Cascadia tremor. *Geophysical Research Letters*, 35(20):1–5, 2008. doi: 10.1029/2008GL035458.

- A. G. Wech and K. C. Creager. A continuum of stress, strength and slip in the Cascadia subduction zone. *Nature Geoscience*, 4(9):624–628, 2011. doi: 10.1038/ngeo1215.
- A. G. Wech, C. M. Boese, T. A. Stern, and J. Townend. Tectonic tremor and deep slow slip on the Alpine Fault. *Geophysical Research Letters*, 39(10):1–6, 2012. doi: 10.1029/2012GL051751.
- A. G. Wech, A. F. Sheehan, C. M. Boese, J. Townend, T. A. Stern, and J. A. Collins. Tectonic Tremor Recorded by Ocean Bottom Seismometers. *Seismological Research Letters*, 84(5):752–758, 2013.
- A. Wells, M. D. Yetton, R. P. Duncan, and G. H. Stewart. Prehistoric dates of the most recent Alpine fault earthquakes, New Zealand. *Geology*, 27(11):995–998, 1999. doi: 10.1130/0091-7613.
- M. Withers, R. Aster, C. Young, J. Beiriger, M. Harris, S. Moore, and J. Trujillo. A comparison of select trigger algorithms for automated global seismic phase and event detection. *Bulletin of the Seismological Society of America*, 88(1):95–106, 1998. ISSN 00371106.
- J. D. Zechar and R. M. Nadeau. Predictability of repeating earthquakes near parkfield, california. *Geophysical Journal International*, 190(1):457–462, 2012. doi: 10.1111/j.1365-246X.2012.05481.x.
- H. Zhang and C. H. Thurber. Development and Applications of Double-difference Seismic Tomography. *Pure and Applied Geophysics*, 163(2-3):373–403, 2006. doi: 10.1007/s00024-005-0021-y.
- M. D. Zoback. Managing the seismic risk posed by wastewater disposal. *Earth*, 1 (April):38–43, 2012. ISSN 1943345X.

