Quantitative Characterization of Seismic Tremors in the Northern Cascadia Margin

By

Philip John Thompson
B.Sc., University of Southampton, 2003

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the School of Earth and Ocean Sciences

Quantitative Characterization of Seismic Tremors in the Northern Cascadia Margin

By

Philip John Thompson
B.Sc., University of Southampton, 2003

Supervisory Committee

Dr. G.D. Spence (School of Earth and Ocean Sciences)

Co-Supervisor

Dr. H. Kao (Geological Survey of Canada, Pacific Geoscience Centre)

Additional Member and Co-Supervisor

Dr. G. Rogers (Geological Survey of Canada, Pacific Geoscience Centre and School of Earth and Ocean Sciences)

Departmental Member

Dr. H. Dragert (Geological Survey of Canada, Pacific Geoscience Centre and School of Earth and Ocean Sciences)

Departmental Member

Supervisors: Dr Honn Kao and Dr George Spence

ABSTRACT

The episodic tremor-and-slip (ETS) events refer to the concurrence of westward crustal movements, as evident by GPS observations, and tremor-like seismic activity in the northern Cascadia margin. The regular occurrence of ETS events in the region has been remarkable. In addition to the 14-month period, secondary tremor activities, most of them lasting less than one or two days, are also found with no corresponding GPS signatures. However, the identification of tremor activity is mainly based on visual examination of regional/local seismic records. In this study, we attempt to develop an algorithm that can quantitatively characterize the level of tremors from a collection of seismic waveform data. For each hour of waveform at a given station, the process begins with the calculation of moving average and scintillation index with various time lengths. The scintillation index, essentially the "normalized variance of intensity of the signal", is adapted from the studies of pulses in radio waves and is an efficient tool to identify the pulse-like characteristics of tremor signals. Values of the indices are fed into a series of logic gates that use a combination of both parameters to determine if sufficient tremor activity exists. To demonstrate the effectiveness of our algorithm, seismic waveform data are collected for the known February/March 2003 ETS event. Our analysis gives results consistent with the work done manually. Implementation of our algorithm is straightforward and free from human intervention. Thus, it is potentially possible to automate the tremor monitoring process that may give early warning of the exact arrival time of ETS events.

TABLE OF CONTENTS

Abstract iii
Table of Contents iv
List of Figures vi
List of Tables viii
Acknowledgements ix
Dedication x

CHAPTER 1: INTRODUCTION 1

1.1 Purpose 2
1.2 Thesis Outline 2

CHAPTER 2: BACKGROUND SETTING 4

2.1 Tectonic Setting 4
2.2 Past Plate Motions 6
2.3 Present Plate Motions 8
2.4 Local Geology 9
2.5 Subduction Thrust Earthquakes 11
2.6 Episodic Tremor-and-Slip (ETS) Events 12
2.7 Thesis Objectives, Motivation and Importance 18

CHAPTER 3: WAVEFORM DATA 20

3.1 Canadian National Seismic Network 20
3.2 Region of Study 24
3.3 The February – March 2003 ETS sequence 26
3.4 Data Acquisition from the CNSN Database 26

CHAPTER 4: DATA PROCESSING 28

4.1 The Scintillation Index of a Waveform 28
4.2 The Moving Average of a Waveform 32
4.3 Waveform Processing Procedures 35
4.4 A representative example 37

CHAPTER 5: ANALYSIS & RESULTS 42

5.1 Characteristics of Background Noise 42
5.2 Characteristics of a Spike/Earthquake 45
5.3 Characteristics of Tremors 50

5.4 Typical Values for MAV and SIR 53
5.5 Testing the Theory 56
5.6 Application to the 2003 tremor sequence 61

CHAPTER 6: UNCERTAINTY & ERRORS 72

6.1 Examples of inconsistent identification and possible reasons 72
6.2 Final Remarks 82

CHAPTER 7: CONCLUSIONS 86

REFERENCES 88

LIST OF FIGURES

2.1 Tectonic setting of western North America 5
2.2 Cross-section cartoon of the northern Cascadia subduction zone 7
2.3 Regional geology of Vancouver Island 10
2.4 Map of the locations of GPS sites in southwestern British Columbia and northern Washington State 13
2.5 Cartoon of areas where ETS slip occurs 15
2.6 GPS time series along with tremor occurances 17

3.1 CNSN network of broadband instruments 22
3.2 Distribution of seismic stations in south-western British Columbia 23
3.3 Distribution of seismic stations used for this study 25

4.1 Synthetic waveform examples of background noise, spike/earthquake, and tremor patterns with corresponding SI traces 31
4.2 Synthetic waveform examples of background noise, spike/earthquake, and tremor patterns with corresponding MA traces 34
4.3 Hour-long original broadband waveform, vertical component taken from station TWBB for the hour beginning at 09:00:00 on March 17, 2003, in SAC format 39
4.4 Waveform from station TWBB for the hour beginning at 09:00:00 on March 17, 2003, after mean and linear trend removal and application of a high-pass Butterworth filter 40
4.5 Waveform from station TWBB for the hour beginning at 09:00:00 on March 17, 2003, fully processed with SI and MA traces shown 41

5.1 An example of a background noise pattern at station MGCB for the hour beginning at 23:00:00 on February 23, 2003 43
5.2 An example of a spike/earthquake pattern at station CBB for the hour beginning at 15:00:00 on March 03, 2003 46
5.3 An example of a teleseismic event 49
5.4 An example of a tremor pattern at station MGCB for the hour beginning at 10:00:00 on February 28, 2003 51
5.5 Graphs of MAV vs. SIR values for all data from the 2003 tremor sequence 57
5.6 Logic flow used for classification 60
5.7. Selected examples of waveforms (high pass filtered 1.5 Hz) for the hour beginning at 09:00:00 March 17, 2003 62
5.8 Classification results for five representative stations on two different days 64
5.9 Day-classification of stations in the lower BC area for 23 February 2003 66
5.10 Day-classification of stations in the lower BC area for 25 February 2003 67
5.11 Day-classification of stations in the lower BC area for 09 March 2003 69

5.12 Day-classification of stations in the lower BC area for 18 March 2003 70
5.13 Day-classification of stations in the lower BC area for 23 March 2003 71

6.1 High pass filtered waveform of the hour beginning at 04:00:00 from ..... station ANMB, 26 February 2003 74
6.2 High pass filtered waveform of the hour beginning at 09:00:00 from ..... station GOWB, 24 February 2003 75
6.3 High pass filtered waveform of the hour beginning 12:00:00 from ..... station ANMB, 02 March 2003 76
6.4 High pass filtered waveform of the hour beginning 00:00:00 from ..... station ENGB, 16 March 2003 78
6.5 High pass filtered waveform of the hour beginning 15:00:00 from ..... station GOWB, 7 March 2003 79
6.6 High pass filtered waveform of the hour beginning 04:00:00 from ..... station TXB, 24 March 2003 81
6.7 A graph of coherent tremors, classified by automatic and visual studies. 85

LIST OF TABLES

3.1 Stations used for this study 24

5.1 SIR and MAV values for the examples seen in Figures 5.1, 5.2, and 5.4. 54
5.2 Value ranges for background noise, spike/earthquake and tremor patterns (the ranges represent 98% of each type of pattern). 56
5.3 Various SIR values and the corresponding percentages of background noise (BG), tremor (TR) and spike/earthquake (SP) correctly classified. 58
5.4 Various MAV values and the corresponding percentages of background noise (BG), tremor (TR) and spike/earthquake (SP) correctly classified. 59

ACKNOWLEDGEMENTS

I wish to extend my greatest appreciation to my supervisor **Dr Honn Kao**, for giving me the opportunity to work with you on the exciting ETS phenomenon when I was a CASSIS project orphan. For your countless ideas, constant support, guidance, patience and inspiring, motivating chats. Thank you for your belief in me when things weren't going so well, without which, this would not have been possible. I have learnt so much from our discussions and your advice (even on the smallest programming problems).

To **Dr. George Spence**, thank you for giving me the opportunity to come to Canada, by offering me a place at the University of Victoria. It has been an amazing opportunity, both in terms of work and life experience. Thank you for sorting out financial support even when the CASSIS project was no more.

To **Dr. Herb Dragert** and **Dr. Garry Rogers**, thank you for your insight into the world of ETS and sharing with me your knowledge on the subject.

To everyone at the Pacific Geoscience Center, especially my numerous office mates as I moved around. Thank you for letting me share your offices and some great chats to make working fun. Thank you to **Ikuko Wada**, **Marc-Andre Chen** and **Lucinda Leonard** for helping me get to PGC on time with your ride sharing, and for entertaining me with conversation early in the morning. Thank you to **Steve Taylor** for your computer assistance on even the smallest matters.

I must obviously thank **my family** for supporting me when I decided to head off to a different country. Although I don't get to see them as much as I would like, their love, encouragement and support is always with me.

To **the Harrison's** who became a second family to me from the day I landed in Victoria. It's very comforting to know I can always come around, and your Sunday dinners are delicious.

I would like to thank **Michelle Zuccaro** for always being there for me, supporting and loving me, making me get my work done and punctuating it with numerous tea breaks. An Englishman cannot work without a daily tea intake! Last but not least I would like to thank the rest of the **Zuccaro family** for their support, for taking me into their family and for their numerous good conversations!

In Loving Memory of Marian Pearson (Granny)

I wish you were still with us to see my achievement, as where would I be if I didn't listen to our motto.

"*Fail to prepare and you prepare to fail*"

CHAPTER 1

INTRODUCTION

The Cascadia subduction zone extends from the northern terminus of the San Andreas Fault in southern California to the Queen Charlotte Fault in northern British Columbia, following the coastline of western North America. Here, the Juan de Fuca plate system converges with the continental North America plate.

The Cascadia subduction zone marks an active fault. Earthquakes of magnitudes ranging from 0.5 to approximately 7 have been experienced in recorded history. It is believed from native oral records and geological evidence such as buried peat layers, ocean turbidities, tree ring data and tsunami sands that the Cascadia subduction zone has had great earthquakes of magnitude 8 – 9 at regular intervals of 300 – 900 years [e.g. Atwater et al., 1990, Adams, 1990, Goldfinger et al., 2003].

More recently, a phenomenon termed episodic tremor-and-slip (ETS) has been identified on the northern part of the Cascadia subduction zone (in the northern Washington State, U.S.A, and southern British Columbia, Canada, region). The ETS events are defined by small magnitude tremors, which occur over a time period of several weeks, along with a westward crustal movement detected by Global Positioning System (GPS) [Rogers and Dragert, 2003].

1.1 Purpose

The objective of this thesis is to develop a computer procedure for the automated identification of tremor occurrence. Implicit in this objective is the need to define a quantitative description of a tremor waveform pattern and differentiate it from other waveform patterns encountered in seismic data.

A proficient automated procedure for tremor identification would provide a common tool for the near real-time detection of tremor within the Cascadia Subduction Zone as well as other similar subduction zones. It would also alleviate the need for time-consuming more subjective visual examination of seismic records. It is conceivable that an ETS episode can be a precursor of a great subduction thrust earthquake [Mazzotti and Adams, 2004]. From this perspective, early recognition of the ETS events could be very useful.

1.2 Thesis Outline

In Chapter 2, a review of the tectonic setting of the region of interest is presented, including past and present motions of the Juan de Fuca plate system, local geology of southern Vancouver Island, the seismicity of the region and an introduction to ETS and the work published thus far on ETS. Chapter 3 presents waveform data and how it may be acquired. Chapter 4 explains the processing procedures from a waveform in its original form to being fully processed for classification. Chapter 5 introduces the characteristics

that the automated procedure identifies and uses as a guide to classify each waveform. The application of the processing and classification procedure to a known ETS tremor sequence is also presented in this chapter. Chapter 6 describes the uncertainty and errors encountered in the automated procedure, and compares the advantages and disadvantages of the automated procedure with respect to a human visual procedure. Conclusions of this study are presented in Chapter 7.

CHAPTER 2

BACKGROUND AND SETTING

This chapter discusses the tectonic setting of the study area, the Cascadia subduction zone. It also describes past and present plate motions. The geology and seismicity of the area are reviewed. The chapter concludes with an introduction to the episodic tremor-and-slip phenomenon.

2.1 Tectonic Setting

The movement of three lithospheric plates dictates the modern plate tectonic regime of southwestern Canada and northwestern United States: the large Pacific and North American plates and the intervening Juan de Fuca plate system. The ongoing motions of these plates are responsible for most of the seismicity of western Canada.

The Cascadia subduction zone involves the northeast convergence of the Juan de Fuca plate system with the North American plate. The Juan de Fuca plate system comprises three plates, the Juan de Fuca plate being the largest, the smaller Gorda plate to the south, and the Explorer plate to the north. (Figure 2.1.)

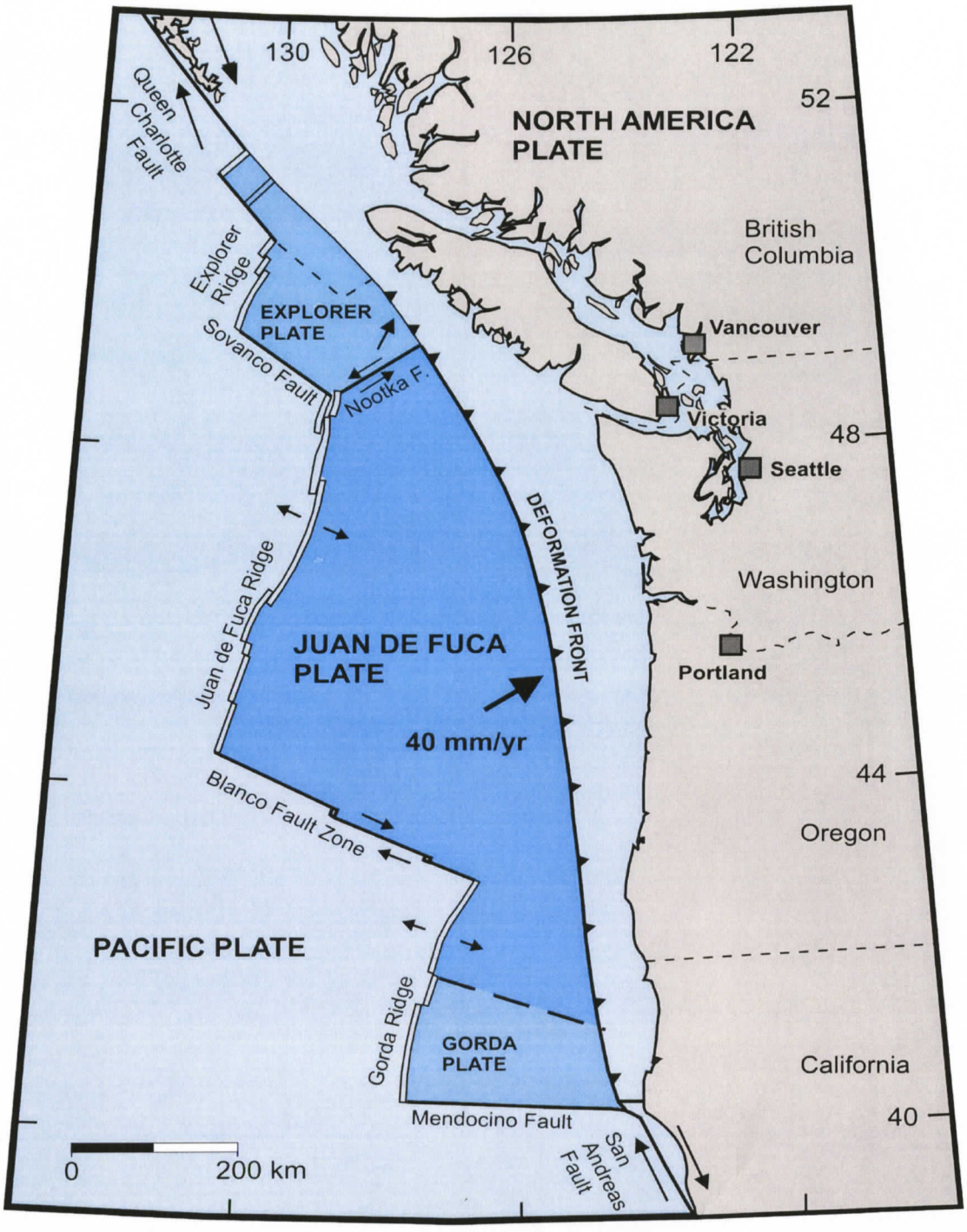

Figure 2.1
Tectonic setting of western North America. At the Cascadia subduction zone the Juan de Fuca plate system is subducting beneath the North American plate [adapted from *Hyndman, 1995*].

The boundary of the Pacific plate and the Juan de Fuca plate system is defined by several ridges and faults; from north to south they are: the Explorer ridge, Sovanco fault, Juan de Fuca ridge, Blanco fault zone, Gorda ridge, Mendocino fault. Active seismicity and bathymetrical features define these boundaries [Davis and Curie, 1993]. The eastern boundary of the system is the Cascadia subduction zone.

The Juan de Fuca plate system subducts beneath the North American plate (Figure 2.2). Seismicity in the subducting plate defines the Benioff zone. This reveals the dip at which the plate is being subducted. The slab is smoothly arched beneath the Puget Sound, subducting at a dip of 10-12° at the top of the arch. North and south of this crest the slab subducts at 15-20° [Crosson and Owens, 1987].

2.2 Past Plate Motions

Past plate motions within the study area can be reconstructed by examining the magnetic anomalies off western North America [Vine, 1966; Atwater, 1970; Riddihough, 1977]. A symmetrical pattern of magnetic anomalies can be seen on either side of the Juan de Fuca ridge. This pattern extends further to the west, under the Pacific Ocean, than to the east. This indicates convergence of an oceanic plate with the North American plate. This plate was the Farallon plate. The Farallon plate began convergence with the North American plate approximately 150 million years ago. The Juan de Fuca plate, the largest remaining fragment of the Farallon plate, became independent when the Pacific-Farallon ridge

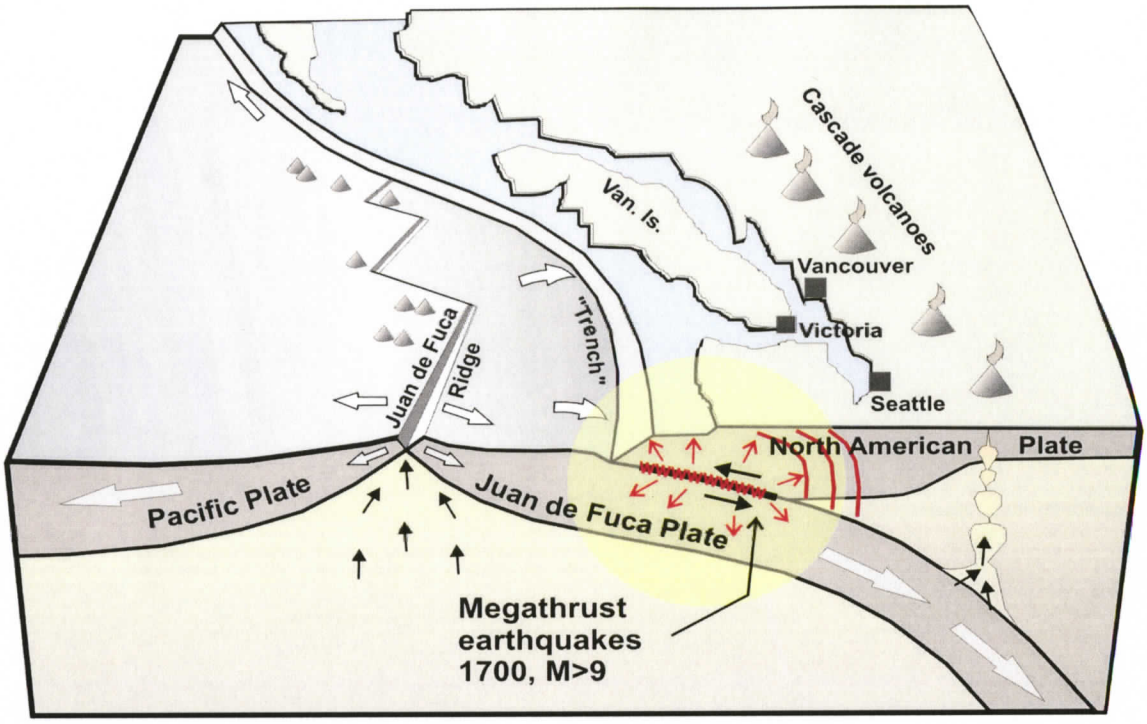

Figure 2.2
Cross-section cartoon of the northern Cascadia subduction zone [modified from *Hyndman et al., 1996*].

collided with the North American plate approximately 20-30 million years ago [Atwater, 1970; Riddihough, 1984]. The collision of the ridge at the north and south ends where formerly there was subduction eventually led to the formation the San Andreas transform fault south of Cascadia and the Queen Charlotte transform fault to the north.

2.3 Present Plate Motions

The contemporary motion of the Juan de Fuca plate can be inferred from present-day seismicity and marine magnetic anomalies.

There are variable rates of motion of the Juan de Fuca plate system. From north to south the Explorer plate is converging at approximately 2.0 cm/yr while the Juan de Fuca plate moves faster still at approximately 4.0 cm/yr [Riddihough, 1984]. The difference in convergence rates between the adjacent Explorer and Juan de Fuca plates is accounted for by the left lateral Nootka fault that separates the plates and exhibits motion of 2 cm/yr [Riddihough, 1984].

The Juan de Fuca plate is separating from the Pacific plate at the Juan de Fuca ridge at a half spreading rate of 2.9 cm/yr (full spreading rate 5.8 cm/yr) in a direction parallel to the Blanco fault zone [Atwater, 1970]. The Explorer ridge is not as well studied but is believed to be spreading more slowly than the Juan de Fuca ridge at a half spreading rate of 2.1 cm/yr (full spreading rate 4.2 cm/yr) [Riddihough, 1977].

2.4 Local Geology

The North American Cordillera consists of a collage of terranes accreted to one another and to the North American margin between Permian and Miocene time. Vancouver Island is located within the westernmost portion of the orogen, and is divisible into three terranes: Wrangellia, Pacific Rim and Crescent [Johnston and Acton, 2003] (Figure 2.3).

Wrangellia underlies 90% of Vancouver Island and comprises rocks of Devonian and Jurassic age. It consists largely of an oceanic plateau, a vast outpouring of basalt and more Mg-rich magma that erupted onto the seafloor and was subsequently accreted to the western margin of the North American plate [Greene et al., 2005]. Rocks of the Pacific Rim and Crescent terranes underlie the south and southwest coastal regions of the island. West and south of Wrangellia is the Pacific Rim terrane consisting of Triassic-Jurassic arc volcanics, marine sedimentary rocks, pillow basalts and amphibolites [Johnston and Acton, 2003]. Southwest of the Pacific Rim terrane lies the Crescent terrane comprising a Paleocene to early Eocene oceanic assemblage of basalt flows, breccia, tuff and volcanic sandstones cut by diabase intrusions. Some evidence suggests the most likely origin of the Crescent terrane was a series of seamounts formed above a hotspot [Duncan, 1982]. Other theories suggest that it is an ophiolite sequence and that the origin is a transform marginal basin analogous to the Andaman Sea [Massey, 1986].

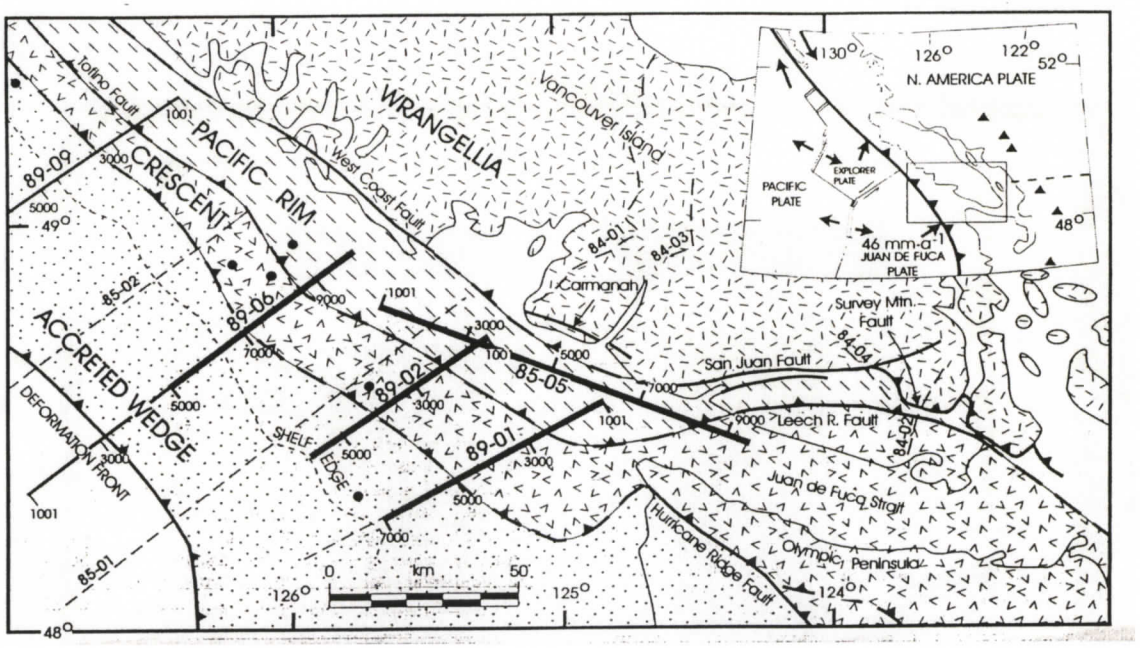

Figure 2.3
Regional geology of Vancouver Island. Adapted from Calvert, 1996

2.5 Subduction Thrust Earthquakes

The combined evidence for sudden subsidence, tsunamis, and shaking shows that earthquakes of magnitude 8 or larger have occurred on the boundary between the overriding North America plate and the downgoing Juan de Fuca and Gorda plates [Atwater et al, 1995]. There have been none in the region's documented history, which begins about 1790 A.D. However, evidence from buried peat layers, geodesy, tsunami records, and tree-ring data suggests that great earthquakes have occurred on the Cascadia subduction zone over the past thousands of years, most recently in 1700 A.D [e.g. Atwater et al., 1990, Satake et al., 1996].

Sediment cores from the Cascadia deep-sea channel contain sequences of turbidites. Pelagic intervals deposited between the turbidites suggest that at separate locations the turbidity currents occurred simultaneously and fairly regularly, every 590 ± 170 years on average. The best explanation for this pattern is that 13 great earthquakes on the Cascadia subduction zone triggered the turbidity currents [Adams, 1990; Goldfinger et al., 2003].

A magnitude 8 earthquake ruptures an area of about 10,000 km^2. This would correspond to a plate-boundary rupture 50km wide and 200km long at the Cascadia subduction zone [Atwater et al., 1995]. With convergence at the Cascadia Subduction zone averaging about 4 m per century, three centuries of convergence yields more than 10 m of potential seismic slip, five centuries about 20 m. By comparison, slip during the 1960 Chile and

1964 Alaska earthquakes (on similar type and size subduction zones) averaged about 20 m [Atwater et al., 1995].

The thrust fault of the Cascadia subduction zone has been well studied. Near Vancouver Island, it is believed to comprise a 60-km-wide locked, seismogenic zone underlying the continental slope offshore from Vancouver Island, with a 60-km-wide transition zone, which extends landward to just beneath the west coast (defined by transition from a completely locked to free slip). Further south beneath the northern Olympic Peninsula, the zones grow to 100-km-wide each [Dragert et al., 1994] (Figure 2.4).

2.6 Episodic Tremor-and-Slip (ETS) Events

Geodetic measurements over the past decade at sites on the Cascadia margin show that the shallow offshore part of the fault is locked and accumulating stress, which will ultimately be released in a great earthquake. Cooler temperatures and greater friction characterize the shallow part of the fault. On the deeper part of the fault, which is the hotter more plastic part, slip is assumed to be occurring steadily at the rate of plate convergence.

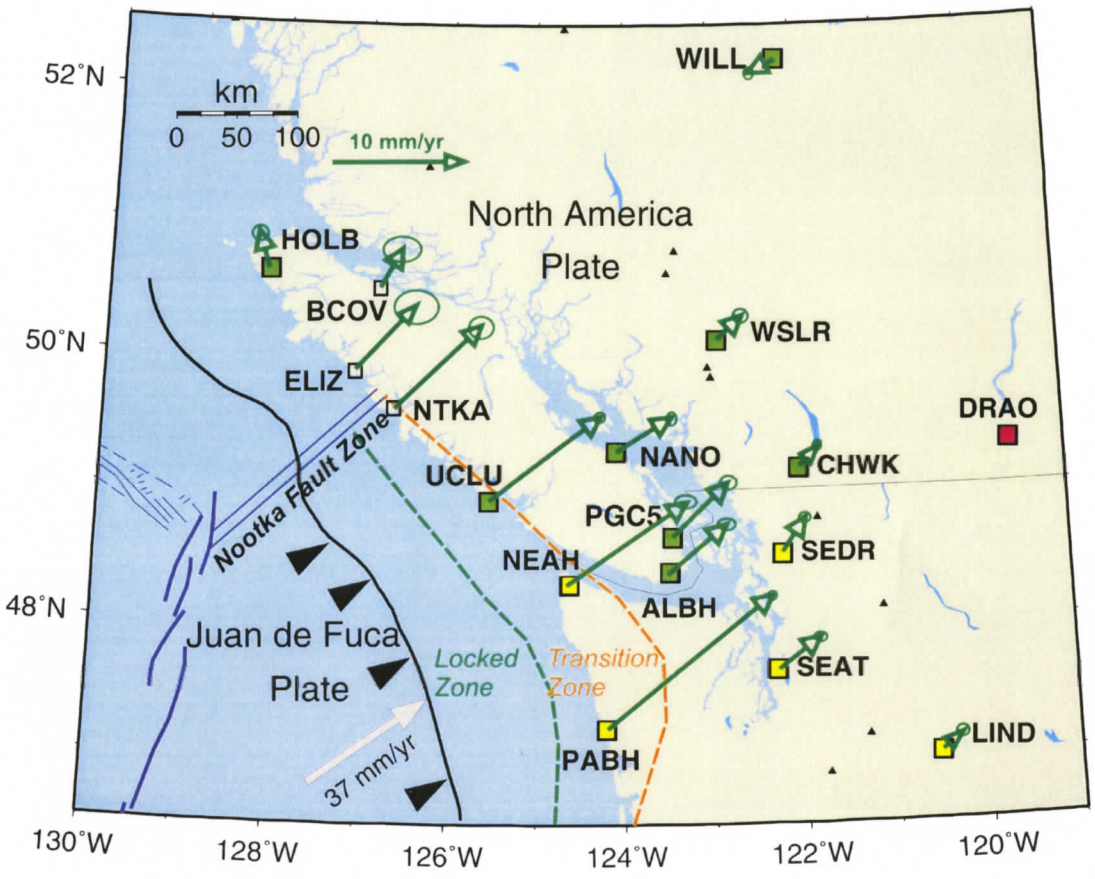

Figure 2.4
The location of GPS sites in southern British Columbia and northern Washington State. Their corresponding long-term horizontal displacements are indicated by green arrows [from Dragert et al., 2004].

Sites using the Global Positioning System (GPS) in southwestern British Columbia were established specifically for the study of Cascadia crustal motions. They can detect changes in latitude, longitude and height of sites relative to the GPS site at Penticton in the stable interior of North America. The observations indicate a long-term northeastward motion, largest at outer coastal sites (Figure 2.4). The motions are consistent with surface displacements predicted from finite element modeling of the convergent margin, with a 60 km region of the interplate boundary currently locked [Fluck et al., 1997].

Precise processing of the GPS data identified unexpected transient displacements at several contiguous sites in the latter half of August 1999. Total horizontal displacements ranged from approximately 2 to 4 mm, and occurred over a time span of one to two weeks. This transient motion was in the opposite direction to the usual northeast convergence motion [Dragert et al., 2001]. It can be accounted for by about 2 cm of aseismic slip on the subducting plate interface, down dip of the locked and transition zones, at depths between 25 km and 45 km (Figure 2.5).

The slip initiated near Seattle and then propagated northwestward along the plate boundary over a period of approximately 35 days and thus affected a 50 by 300 km fault area. To put it in context, if this slip had occurred as a seismic rupture, it would have produced an earthquake of magnitude 6.7, very similar in energy to the Nisqually earthquake near Seattle in 2001. There is no evidence for an obvious seismic trigger for these events. It is believed to be brought on by sudden changes in rheology and friction on the interface

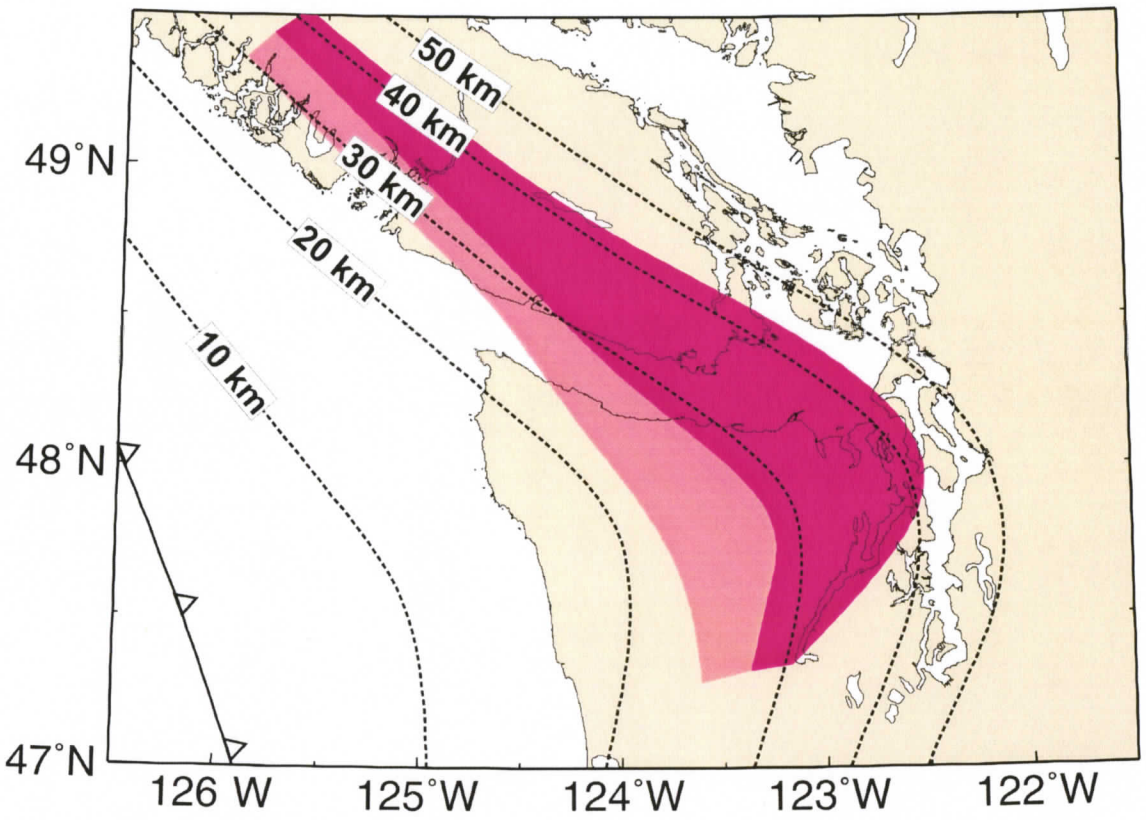

Figure 2.5
A model of a slow slip event from August 1999. The geometry of the subducting plate interface is adopted from structural studies (Fluck et al., 1997) and it is assumed that simple linear slip distribution directed up-dip. Dark shading indicates fault areas with full slip; light shading indicates fault areas where slip is tapered linearly from full to zero at the up-dip end. The equivalent moment-magnitude for this event (assuming rigidity of 40Gpa) is 6.7 [adapted from Dragert et al., 2004].

[Dragert et al., 2001]. Stress transfer from such deep slips could bring the locked zone of the fault closer to failure and so enhanced seismic hazard may well accompany slip events. [Mazzotti and Adams, 2004]

Transient slip events on the Cascadia margin have been studied in more depth in recent years. They appear to have a recurrence interval of 14±2 months. The slip was originally thought to be seismically silent as no earthquakes were observed. However, unique non-earthquake tremor-like seismic signals were subsequently observed to accompany the slip events. The tremor is similar to that observed in the forearc region of Japan, particularly in the Nankai subduction zone [Obara, 2002]. For six deep slip events between 1996 and 2003 the tremorlike signals in the Cascadia region were found to correlate spatially and temporally with the slip. Rogers and Dragert [2003] referred to the phenomenon as episodic tremor-and-slip (ETS) events (Figure 2.6).

The seismic tremors have a lower frequency content than regular small earthquakes. The frequency content for tremors is generally between 1 and 5 Hz, while much of the energy in a small earthquake is above 10 Hz. Unlike earthquakes, tremors emerge gradually and the energy pulses often last for several minutes. A series of pulses can last for several days. Tremors are most strongly seen on horizontal seismographs, and move across the seismic network at shear wave velocities [Rogers and Dragert, 2003]. A tremor on an individual seismogram is unremarkable and does not appear different from transient noise due to wind or cultural sources. It is only when a number of seismograph signals are

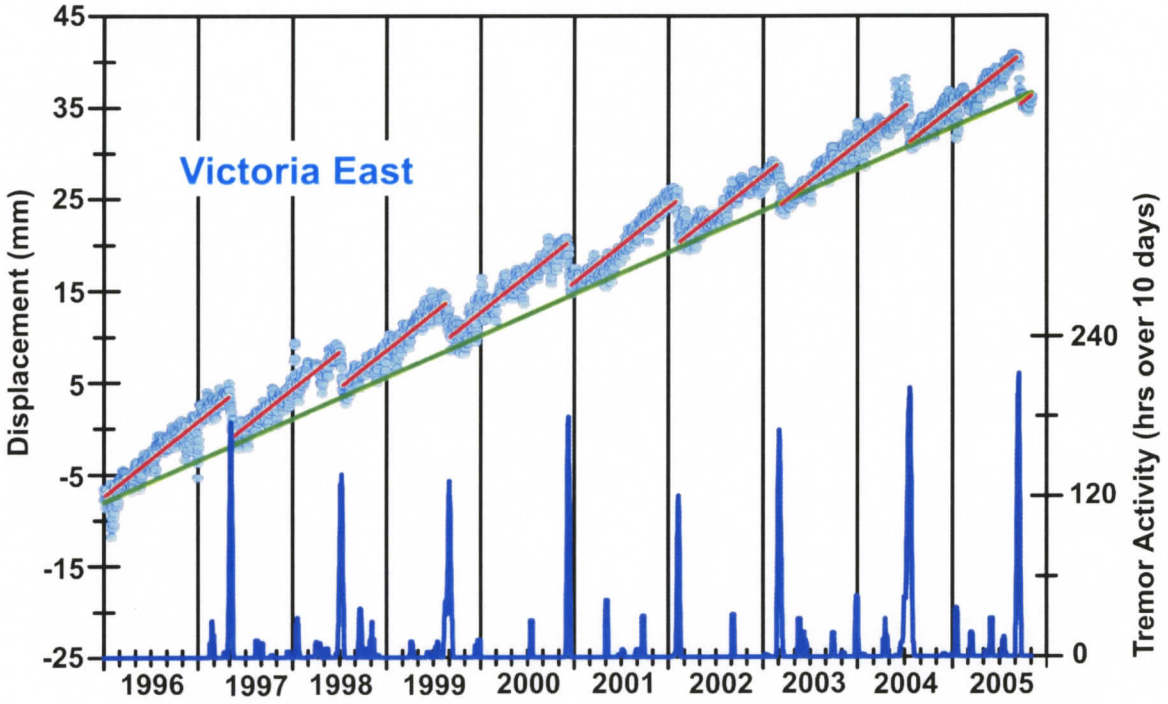

Figure 2.6
This shows the tremor and slip on Vancouver Island. The blue dots represent displacements in daily east-west position (of GPS instruments) with respect to Penticton. The green line represents the long-term (interseismic) linear motion observed. The red lines indicate the short-term (inter-slip) linear motion followed by reversals. The blue graph shows tremor activity on southern Vancouver Island [from Dragert et al., 2004, updated by G. Rogers and H. Dragert, pers. comm., 2006].

viewed together that the similarity of the seismic waveforms across the seismic array identifies ETS events [Rogers and Dragert, 2003].

Rogers and Dragert [2003] identified the ETS events by visual inspection of recorded seismograms. Every hour of data from regional seismic stations was classified manually as to containing or not containing tremor activity. There were no set criteria, but a tremor pattern is quite evident once one knows what to look for. This process, though time consuming and subjective, accurately identified many tremor hours for the 2003 ETS event.

2.7 Thesis Objectives, Motivation and Importance

The objective of this thesis is to create a logical, non-subjective, automated method for identifying tremors of data on seismograms. This computer-based technique is intended to replace the current time-consuming (and more subjective) methodology of visual examination of each hour of network data and to decide whether or not tremor has occurred during that hour. To do this, a program with logic gates to classify each hour and station was designed. The motivation behind this idea is to be able to eventually have real time processing of waveform data, which will be able to detect the onset of a tremor event automatically. It is also hoped that the method will be objective and that possibly we will be able to use it on other subduction zones where the same ETS phenomenon

may occur. Identification of ETS could be extremely important as ETS events may serve as a trigger for a megathrust earthquake [Mazzotti and Adams, 2004].

CHAPTER 3

WAVEFORM DATA

This chapter introduces the waveform data available for Canada and the specific region selected for this study. It then introduces the February – March 2003 ETS sequence, the events forming the basis of the study. The chapter concludes with a description of the procedure to acquire the waveforms from the Canadian National Seismic Network.

3.1 Canadian National Seismic Network

The primary source of seismic waveform data in Canada is the Canadian National Seismic Network (CNSN), operated by the Geological Survey of Canada (GSC). The network consists of more than 100 seismograph stations nationwide, with 35 stations in British Columbia. Stations are generally located in areas relatively free from artificial noise, such as traffic and mining. Signals from these instruments are sent to GSC offices near Sidney, British Columbia (on Vancouver Island) and in Ottawa, Ontario. The signals are transmitted in a variety of ways, including satellite transmission for the most remote sites, radio transmission over shorter distances, and in a few cases, telephone lines. Wherever possible, data transmission via Internet is also established. These signals are continuously monitored for seismic activity and stored on computer hard disks.

The seismograph network consists primarily of two types of instruments: three-component broadband instruments, and vertical-component short-period instruments. The broadband instruments record over a wide frequency band (typically from 0.03 Hz to 10 Hz) at a sample rate of 40 samples per second. Thus, they are sensitive to high-frequency energy associated with local earthquakes, as well as longer period energy that is associated with larger and more distant earthquakes. These instruments can easily record earthquakes larger than about magnitude 3.0 anywhere in Canada, and larger than about M=5.5 anywhere on Earth. The short-period instruments record at a higher sample rate (100 sample/second) and are useful for analyzing the high-frequency waves associated with local earthquakes.

In Canada, the broadband instruments form the nucleus of the CNSN. Figure 3.1 shows the CNSN network of broadband instruments across Canada. It can be seen that the southwestern region of British Columbia is very densely covered by CNSN. This is mainly due to the high rates of earthquake activity in this region. Data from this network are available to seismologists around the world within minutes via Internet. Short-period instruments are deployed in areas more susceptible to earthquakes and earthquake damage. For example, more than 20 of these instruments are located in the seismically active and more densely populated southwest corner of British Columbia. Figure 3.2 shows the distribution of seismic stations in the southwestern British Columbia region.

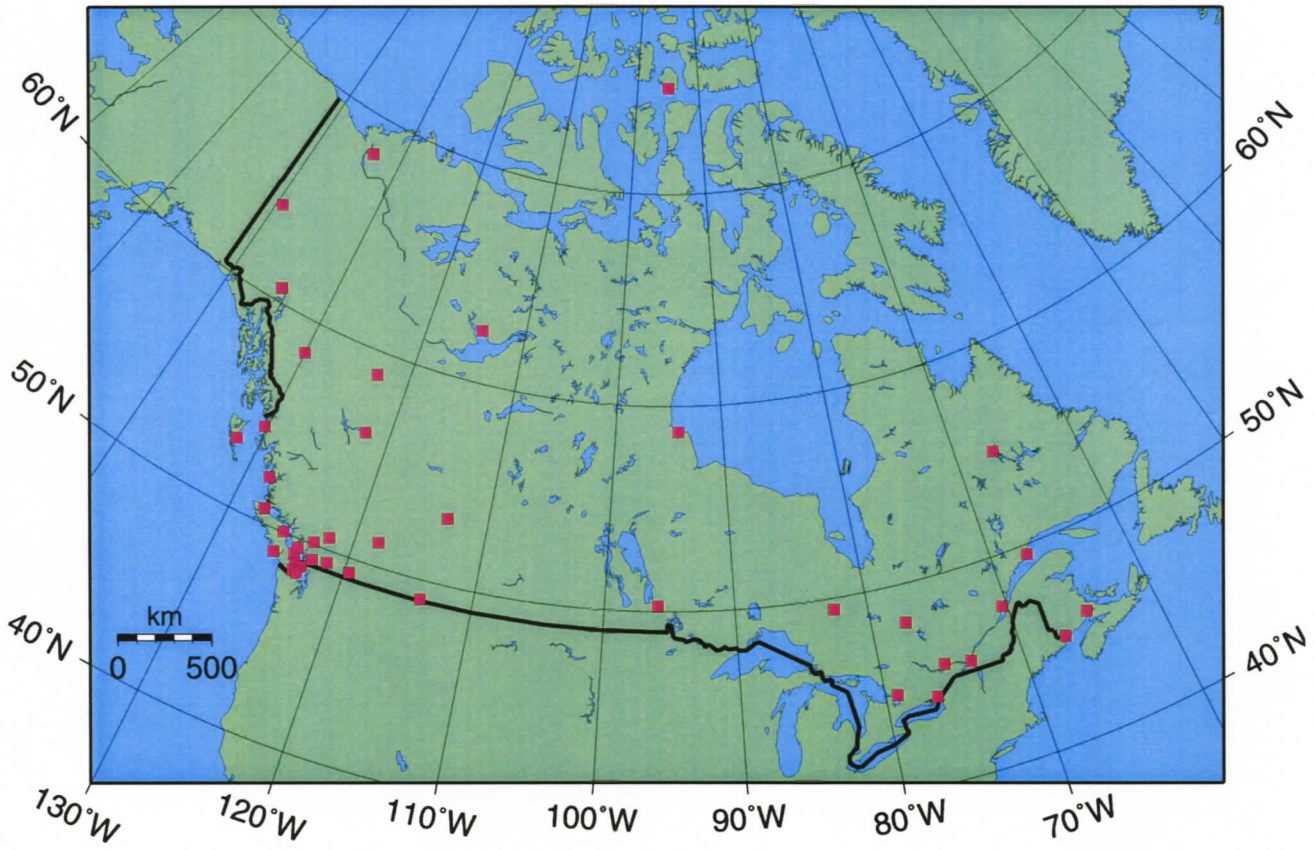

Figure 3.1
Canadian National Seismic Network (CNSN) network of broadband instruments.

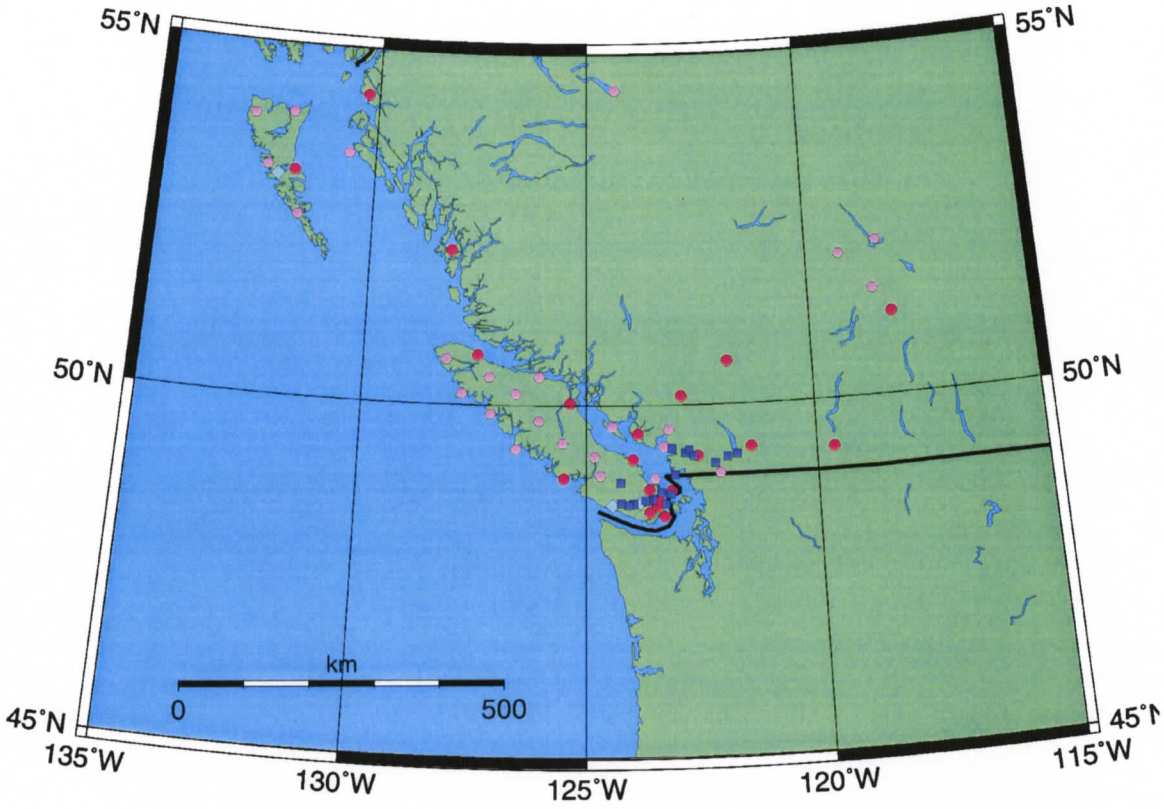

Figure 3.2
Distribution of seismic stations in lower British Columbia. Red circles and pale blue diamonds represent broadband stations with low (<40 samples per second) and high (100 samples per second) sampling rates respectively. Pink hexagons and royal blue squares are temporary stations with sampling rates of at least 40 samples per second.

3.2 Region of Study

In previous studies that identified and located ETS events [e.g. Kao et al. 2005], the tremors began south of Vancouver Island in Washington State, USA. The tremors then migrated northwestwards through southern Vancouver Island at a rate of 5-10 km per day, sometimes with a smooth transition, other times with sudden jumps. The current study concentrates on the Canadian stations in the region bounded by 122°-128° W and 48° - 50° N (Figure 3.3). These stations are situated mainly on southern Vancouver Island, with a few in the Lower Mainland of British Columbia near Vancouver (Table 3.1). Fewer stations were used in the study than appear in Figure 3.2. This is due to the fact that data were not available from the CNSN database for all stations in the area. The southern region of Vancouver Island remains well covered. However, the northern region is more sparsely covered.

Table 3.1 *Stations used for this study*

Station Name	Latitude	Longitude
ALB	49.2718	-124.8277
ANMB	49.319	-122.859
BIB	49.4094	-123.31
BTB	49.4683	-125.5214
CBB	50.0315	-125.3653
CLVB	49.3817	-123.1447
COQB	49.3542	-122.7747
ENGB	49.008	-123.0891
GOWB	48.737	-123.1844
HNB	49.2745	-122.5792
KELB	48.6611	-123.5702
LZB	48.6117	-123.8236
MGB	49	-124.6975
MGCB	48.6318	-123.6811

Station Name	Latitude	Longitude
PFB	48.575	-124.4444
PGC	48.65	-123.4505
PIMB	49.2745	-122.666
SHB	49.593	-123.8805
SNB	48.776	-123.1712
SSIB	48.7558	-123.3875
TWBB	48.5846	-124.092
TWGB	48.6076	-124.2559
TWKB	48.6449	-123.7332
TXB	49.6969	-124.436
WPB	49.6481	-123.2083
VDB	49.0261	-122.1028
VGZ	48.4139	-123.3244

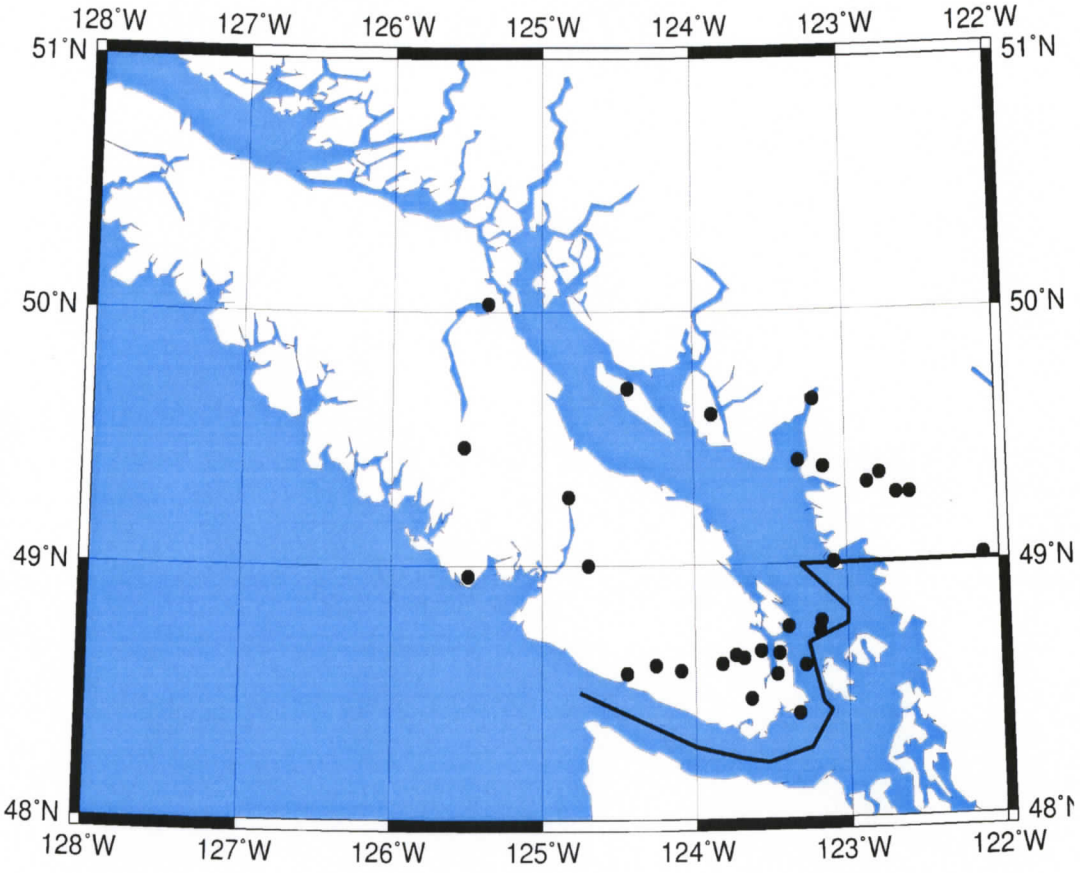

Figure 3.3
Distribution of seismograph stations available on Vancouver Island and the Lower Mainland of British Columbia used in this study. There is a concentration of stations on the southern part of Vancouver Island, with much sparser station coverage further north (above 49 degrees North).

3.3 The February – March 2003 ETS sequence

Extended tremor sequences have been identified in northern Cascadia dating back to 1996. The most recent event at the time this study began was the 2003 February-March sequence. The 2003 tremor sequence began on 25 February 2003 and ended on 22 March 2003. Tremors in this sequence occurred over a large fraction of the total period. This sequence was thus chosen for automated identification of tremors, due to its recent occurrence and previous work available for comparison.

Data from all available seismic stations were selected over the tremor sequence, including three days before the start (25 February 2003) and three days after the end (22 March 2003).

3.4 Data Acquisition from the CNSN Database

To acquire the relevant waveform data for the selected stations in the region of study, an email request must be sent to the Data Management Centre of CNSN (autodrm@seismo.nrcan.gc.ca). The request states the station names, start time of waveform data, duration of data wanted (usually one hour), channel list (just the z component), and the format of the data (usually the Standard for the Exchange of Earthquake Data, SEED format [http://www.iris.washington.edu/manuals]). The CNSN data centre puts the requested waveforms onto their ftp site, and an email is returned with

the location of the data and retrieval instructions. The following is an example of CNSN data request email.

```
BEGIN
EMAIL        pjt@uvic.ca
START_TIME       2003/03/04.00:00:00
DURATION  3600
STA_LIST     PGC, TWBB, GOWB, TWGB, ALB, BIB
CHAN_LIST  *
FORMAT       SEED
WAVEFORM
STOP
```

The first line begins the requesting script. The next line lists the email address where the retrieval details are to be sent. The start time is in the format year/month/day.hour:minutes:seconds. The fourth line gives the requested time duration in seconds, in my case 3600 s (one hour). The fifth line lists the stations from which data are requested. The sixth line gives a list of select channels. An asterisk means that I request all available channels. The format of the data requested is given on the seventh line, in this case SEED. The eighth line gives the type of requested data, in my case waveform data. The ninth and final line ends the requesting script.

CHAPTER 4

DATA PROCESSING

This chapter introduces the two calculated parameters used to process the waveforms for tremor identification, the scintillation index (SI) and the moving average (MA). Generic examples of the types of waveform encountered, and their respective MA and SI characteristics, are discussed. A step-by-step guide to processing a waveform is presented. This chapter concludes with the automated processing procedures, from requesting a waveform to its final classification.

4.1 The Scintillation Index of a Waveform

Scintillation is a flash of light produced in a phosphor by absorption of an ionizing particle or photon. The scintillation index (SI) is defined as the normalized variance of intensity of the signal [Yeh and Liu, 1982]. It was originally used to measure the intensity of flashes of light or bursts of energy (scintillations) in the ionosphere.

In this study we use the scintillation index on seismic traces, likening the bursts of energy found in a tremor to those in radio waves. The formula for calculating the SI is defined as:

$$SI(i) = \sqrt{\left[\left(\left\langle A^2 \right\rangle_i - \left\langle A \right\rangle_i^2\right) / \left(\left\langle A \right\rangle_i^2\right)\right]} \qquad (4.1)$$

where A represents the absolute amplitude of a sample within a time window containing N samples. <A> is the mean value of A for all sample points within the time window. Thus, $<A^2>$ is the mean of the square of the absolute amplitudes of all points in the window, whereas $<A>^2$ is the square of the mean value of the absolute amplitudes of all points. That is:

$$\left\langle A^2 \right\rangle_i = \sum_{j=i-(N-1)/2}^{j=i+(N-1)/2} \left[|y(j)|^2 \right] / N \qquad (4.2)$$

$$\left\langle A \right\rangle_i = \sum_{j=i-(N-1)/2}^{j=i+(N-1)/2} |y(j)| / N \qquad (4.3)$$

In theory, N can be any positive odd integer. In practice, however, the length of the time window should be comparable to the duration of tremor bursts to manifest the characteristics of tremors. For example, for a 10 s window with a sampling rate of 10 Hz, N is 101.

Synthetic examples of the three types of events we expect to appear on seismograms in the 2003 tremor sequence (i.e. background noise, spike/earthquake, and tremor) have been created and shown in Figures 4.1a, 4.1b and 4.1c. The synthetic waveforms shown here are normalized to a maximum nominal value of 10. The normalization will be explained further in section 4.3 (waveform processing procedures). It must be pointed out that the spike shown in Figure 4.1b does not exhibit typical earthquake pattern (i.e. a P-

wave arrival, followed by an S-wave arrival then by a coda). However, an earthquake occurs over a short enough time period (much shorter than a tremor) that it can be modeled as one or more spikes with a trailing coda that tends to be low amplitude. Therefore I simply use a spike to mimic the signal generated by an earthquake. The SI values for these events are also calculated and shown as the red lines above the waveform. The background noise pattern, seen in Figure 4.1a, is made up of perfectly random amplitudes. It fluctuates between highs and lows, with no particular phases in the waveform. The SI value in this case remains almost constant, at the value of 0.57.

The spike/earthquake pattern, seen in Figure 4.1b, exhibits a constant low value (approximately zero) with a single spike at 1800 s having a maximum value of 10.8, followed by a return to the low value. The SI detects this large change. For this example a 300 s time window is used, so as soon as the time window reaches the spike/earthquake event (half a time window length before), the SI value increases dramatically from almost zero to a maximum of approximately 0.9. The SI value remains high until the time window is completely beyond the event, and then falls back to the low value again.

A seismogram containing a tremor is shown in Figure 4.1c, and shows a low level (approximately 1.0) of background noise for the first 300 s of the trace, followed by a gradual increase of amplitude to a maximum value of approximately 9.0. The tremor signal lasts for about 600 s, from approximately 1700 s to 2100 s, followed by a similar gradual decline back to the starting amplitude of approximately 1.0. The SI, which is very useful for detecting changes of waveform patterns, exhibits two peaks. One peak is at the

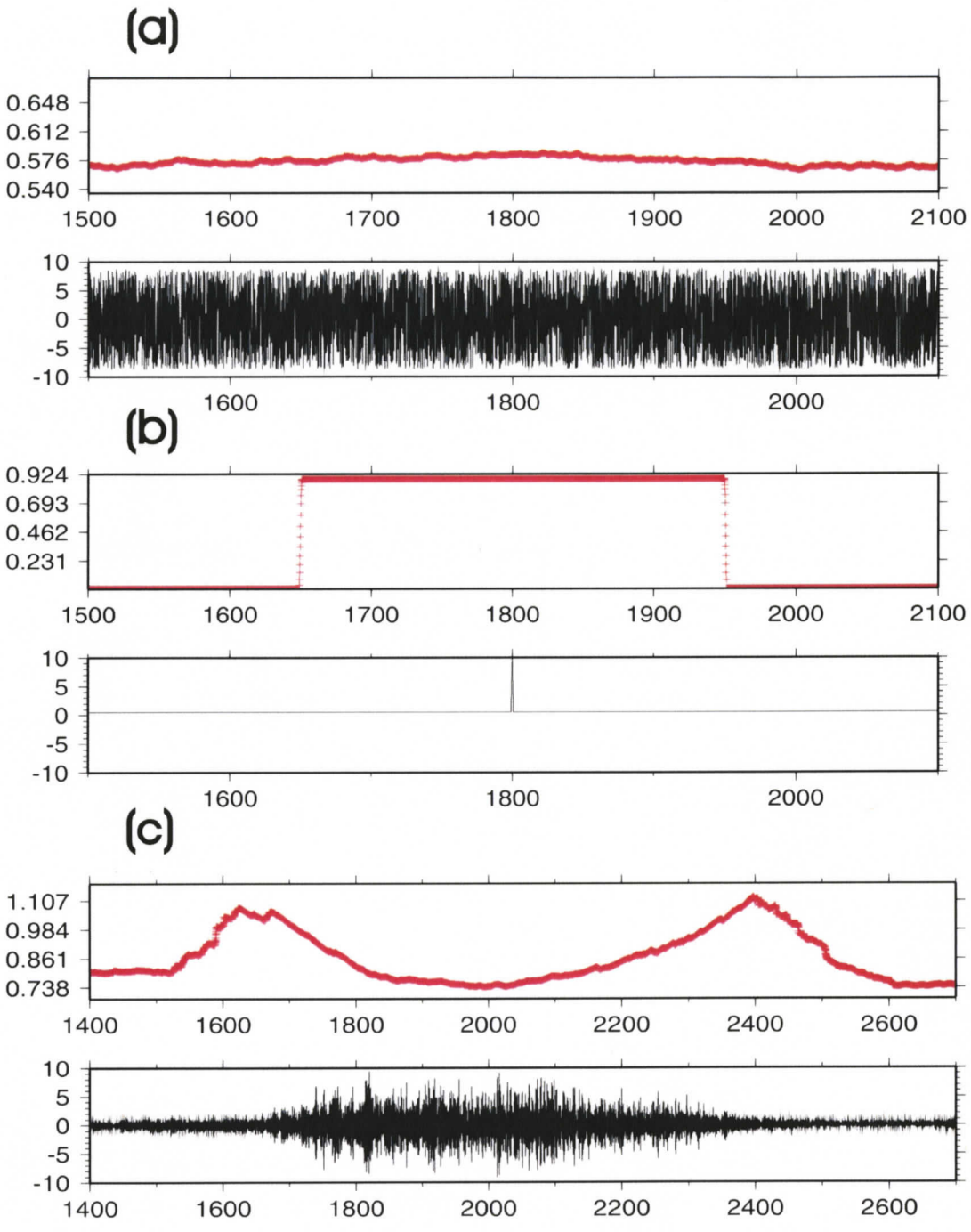

Figure 4.1
Synthetic examples of (a) background noise, (b) spike/earthquake and (c) tremor patterns. For each pattern, the normalized waveform is shown in black with the corresponding SI trace above in red. The x-axis is time in seconds, and the y-axis gives the waveform amplitude and SI values for each case.

onset of the tremor pattern and one is at the end of the tremor pattern; both have maximum amplitudes of approximately 1.1. Between these two peaks, the SI value drops back to low values of approximately 0.75.

It is evident from analysis of these three waveform patterns that the ratio of the maximum SI value to the average SI value varies significantly from one case to another. For a pattern of background noise, we would expect a ratio of around 1.0 due to the relatively constant SI values. For a pattern of tremor, the maximum SI value is approximately 0.35 greater than the average SI value so it may yield a ratio of 1.0 – 1.5. The spike/earthquake pattern on the other hand generates a very large ratio, as the peak SI value is far greater than the corresponding average. This suggests a good way of separating the three different patterns of waveforms.

4.2 The Moving Average of a Waveform

The moving average (MA) of a normalized seismic waveform is also used as a parameter for identifying tremors. The moving average is a simple calculation, which effectively smoothes the data and makes it easier to spot trends. It averages the absolute values of all samples within a specific time window and assigns the centre point that average value. The calculation of MA is defined as:

$$MA(i) = \sum_{j=i-(N-1)/2}^{j=i+(N-1)/2} |y(j)| / N \qquad (4.4)$$

where N is the number of data points within a prescribed time window (various lengths are tested in this study to examine the waveform trends that can best characterize tremors). The moving average is the same as <A>, in equation 4.1, the mean value of the absolute amplitude of all points in a time window.

The same synthetic examples as those used for the SI calculation (Figure 4.1) are used to demonstrate the corresponding MA characteristics and the blue lines in Figure 4.2 show the results. For a pattern of background noise (Figure 4.2a), the MA trace exhibits similar characteristics to the SI trace in Figure 4.1a. Both remain relatively constant, although the MA values are much higher at 4.4 with a small fluctuation.

Figure 4.2b shows the synthetic spike/earthquake pattern. Again the MA exhibits a similar pattern to the SI in Figure 4.1b, except that the MA values are much lower (approximately 0.5 with a maximum value of 0.54). Because the time window used is also 300 s, the half time window lag is evident by the plateau extending 150 s on both sides of the spike.

The tremor pattern seen in Figure 4.2c shows a much different MA trace than the corresponding SI trace in Figure 4.1c. Rather than a peak at the beginning and end of the tremor, the MA trace closely follows the envelope of the tremor waveform. Low values

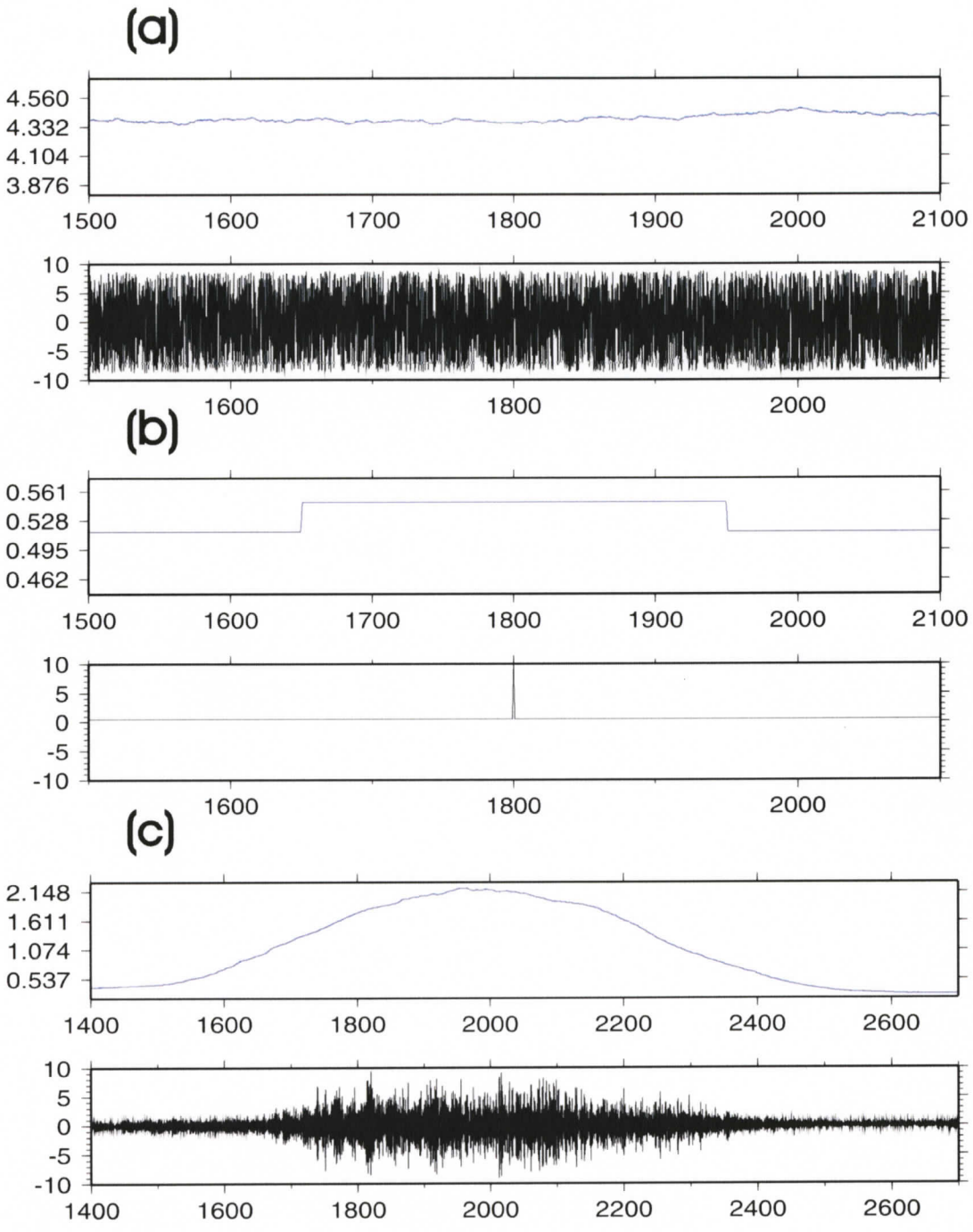

Figure 4.2
Synthetic examples of (a) background noise, (b) spike/earthquake and (c) tremor patterns. For each pattern, the normalized waveform is shown in black with the corresponding MA trace above in blue. The x-axis is time in seconds, and the y-axis gives the waveform amplitude and MA values for each case.

(<0.5) are seen before the gradual rise to the peak amplitude of approximately 2.1 at 2000 s. As the tremor signal gradually decreases in its amplitude, the corresponding MA values also show similar reduction.

The exact values of MA depend critically on how the waveform is normalized. Given the same normalization scheme (e.g. normalized against the maximum peak-to-peak amplitude), we may notice that the majority of the MA values are small if the waveform contains spike/earthquake signals. On the other hand, MA values of a background noise pattern are generally larger. Therefore, the mean value of the MA trace can also be used to separate different patterns of waveforms. The next section and Chapter 5 will have more detailed discussion on this subject.

4.3 Waveform Processing Procedures

After requesting and retrieving the data from the CNSN data centre, a waveform data file in SEED format is ready for processing. This section explains the step-by-step processing procedures used in this study.

Step 1: Convert the waveforms from SEED to Seismic Analysis Code (SAC) [http://www.iris.edu/manuals/sac/SAC_Home_Main.html] format. The program "rdseed", provided by the Incorporated Research Institutions of Seismology (IRIS), is used for this purpose.

Step 2: Use the SAC program to remove the mean and linear trend of the original waveform. This step is necessary to prevent artificial spikes in the later high-pass filtering process.

Step3: Apply a high-pass Butterworth filter with a corner frequency at 1.5 Hz. This is applied because tremors are known to have a rich frequency content of 1-5 Hz. Using the high-pass filter can effectively isolate any possible tremor signal from the background noise of the waveform.

Step 4: Normalize the waveform to further suppress background noise, and enhance the tremor signal. The normalization scheme used in this study normalizes the entire waveform against the mean of the eight maximum amplitudes. The normalization value is then set at 10 (not 1 as with most normalizations). The decision of using the eight largest amplitudes is to prevent the normalization being dominated by an erroneous outlier in the waveform due to unknown/unwanted reasons (e.g. power or transmission instability). We have also tried several different normalization factors, choosing from the six largest to the fourteen largest. Eight appeared to give the most satisfactory results. A big advantage of this normalization scheme is that after normalization, traces with different noise levels will have very different mean MA values. For example, a pattern of background noise should give a much larger average MA because the overall normalized amplitude is larger. In contrast, the mean MA value of a spike/earthquake waveform is very small because most data points before and after the spike/earthquake have very small

amplitudes after normalization. The tremor signal will give a result between the two end members.

Step 5: In the final step of processing, apply equations 4.1 and 4.4 to the normalized waveform with different lengths of time windows (300 s, 100 s, 50 s, 20 s, 10 s) to obtain the SI and MA traces. The results are then used for analysis and classification.

4.4 A representative example

This section shows a representative example of the processing procedures. Figure 4.3 shows the original broadband, vertical-component waveform taken from station TWBB for the hour beginning at 09:00:00 on March 17, 2003 after conversion from the SEED format to SAC. The waveform does not clearly exhibit any particular tremor pattern at this point. It appears to fluctuate between ±150 with the amplitudes increasing slightly towards the end of the time segment, around 3000 s.

Figure 4.4 shows the same waveform after trend removal and high-pass filtering (steps 2 and 3). Now the waveform has small amplitude fluctuations between ±15 for the first 2000 s. At approximately 2200 s the amplitudes begin to increase for about 1000 s before falling back to the smaller amplitudes by 3400 s. It appears that a tremor event occurs between 2200 s and 3200 s. This was not clear before the processing had been applied.

Figure 4.5 (bottom) shows results after normalization (step 4). The normalization can be seen to have decreased the amplitudes of the noise in the segment before the tremor. The tremor is still clear and appears to have two peaks, one at 2650 s and one at 3200 s. The MA and SI have been calculated for the five different lengths of time windows (step 5), and can be seen in blue and red lines in the plots, respectively.

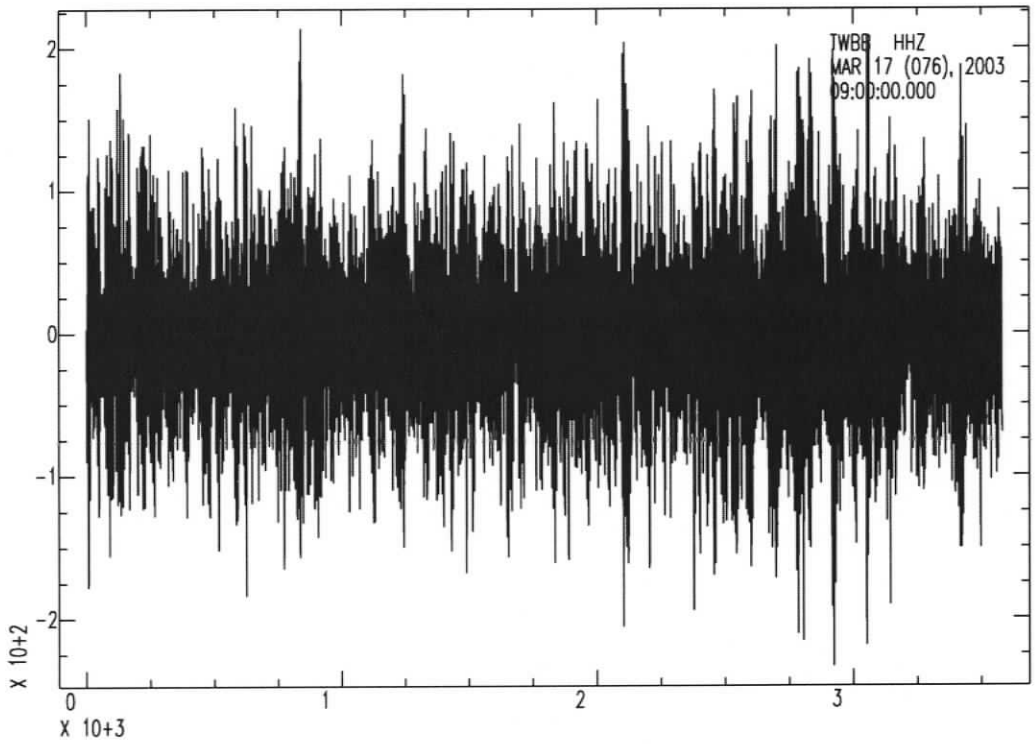

Figure 4.3
Hour long original broadband waveform, vertical component taken from station TWBB for the hour beginning at 09:00:00 on March 17, 2003 in SAC format. The x-axis represents time in seconds, the y-axis amplitude.

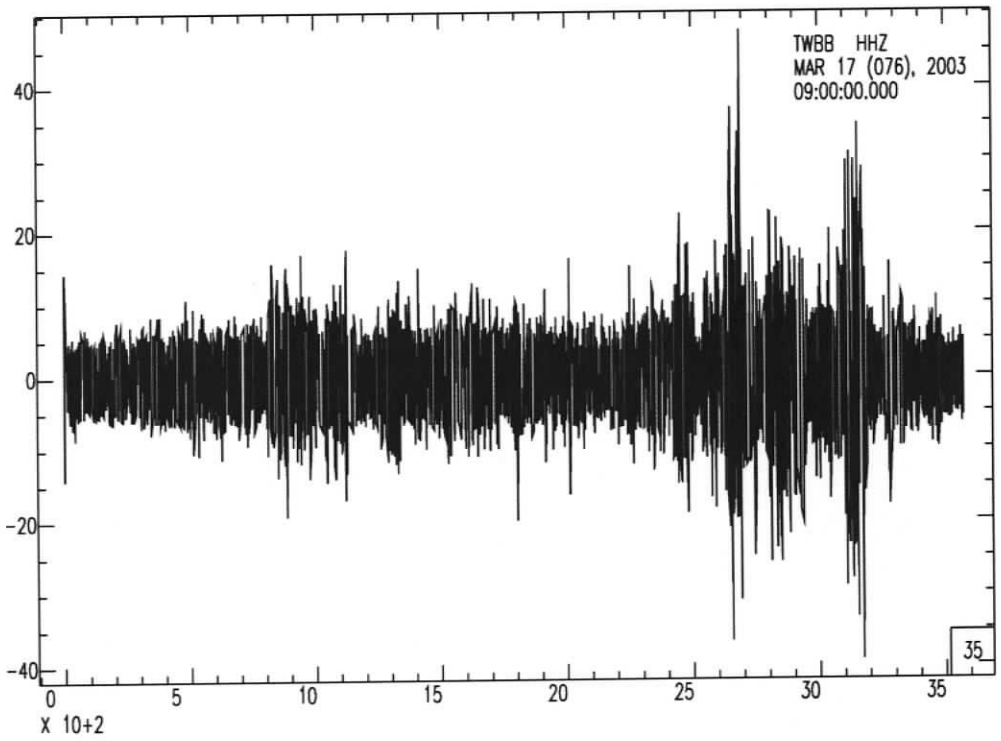

Figure 4.4
The same waveform as in Figure 4.3, after mean and linear trend removal and application of a high-pass Butterworth filter with corner frequency 1.5 Hz. The x-axis represents time in seconds, the y-axis amplitude.

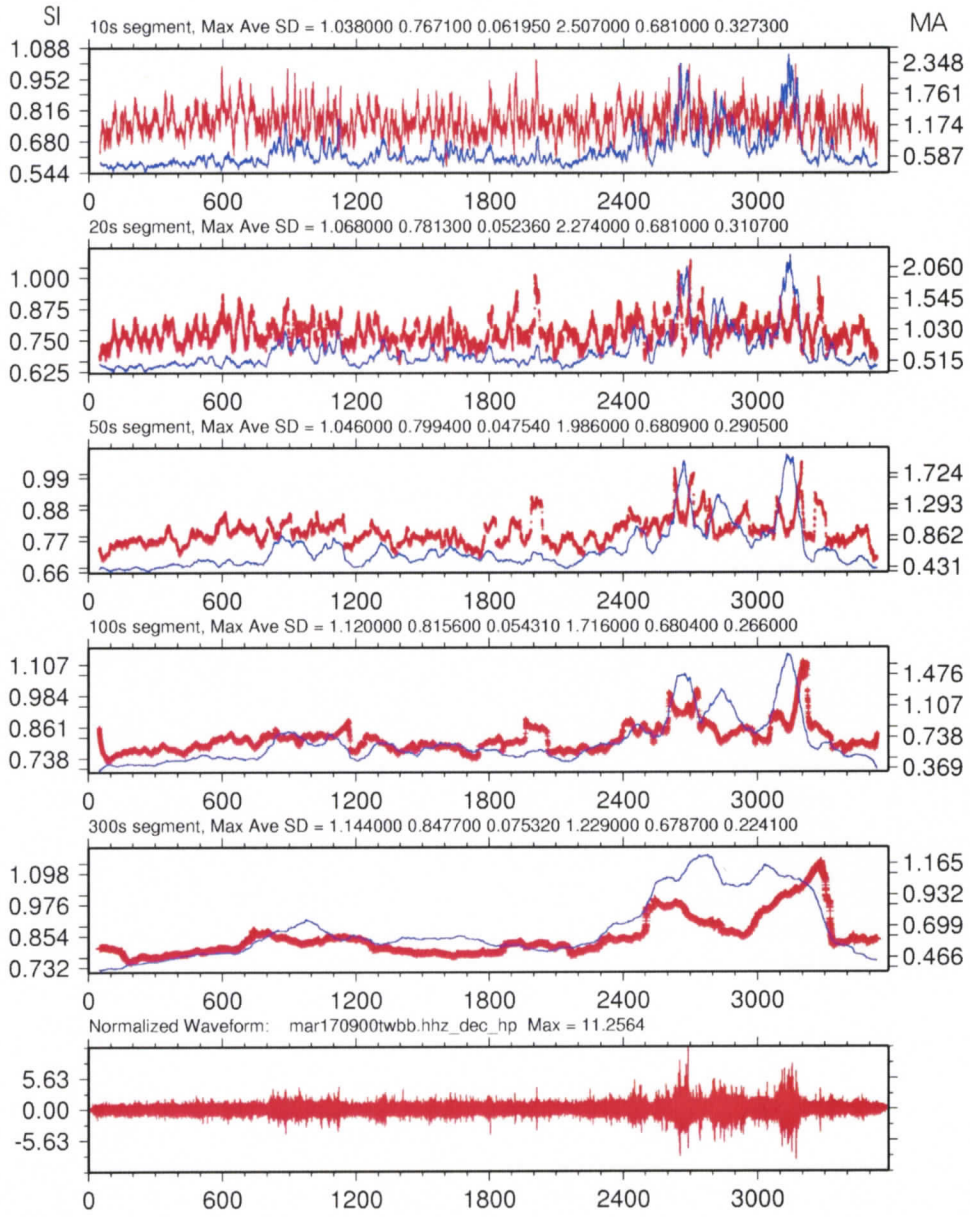

Figure 4.5

The same hour segment as in Figures 4.3 and 4.4, which has now been normalized, and processed using equations 4.1 and 4.4 to calculate SI and MA values, respectively. The bottom graph shows the original waveform high-passed at 1.5 Hz, and normalized. The x-axis runs from 0 - 3600 s (0 s is the start at 09:00:00). In the graph directly above, the red line represents the scintillation Index (SI) and the blue line the Moving Average (MA), calculated over a 300 s time window. The left and right vertical axes correspond to the SI and MA values, respectively. The six numbers listed immediately above the plot are the maximum, average and standard deviation of the red and blue lines. The four graphs above show the same quantities but for time windows of decreasing size (100 s, 50 s, 20 s, 10 s), the smallest time window being the top graph.

CHAPTER 5

ANALYSIS & RESULTS

In this chapter, characteristic patterns for typical examples of background noise, an earthquake/spike and tremors are described. Based on the theory and data processing procedures, the ranges of typical scintillation index ratio (SIR) and moving average values (MAV) for each example are then discussed. A series of tests are performed to select the optimal values of SIR and MAV that can best separate the three different groups. The classification algorithm is then applied to the March 2003 ETS tremor sequence, and a comparison of the results from automated classification and those from a visual examination of data are presented at the end of this chapter.

5.1 Characteristics of Background Noise

Figure 5.1 shows a typical pattern of background noise. The waveform fluctuates randomly about the zero value with maximum amplitude of approximately 4.88. No large peaks or troughs are observed.

The moving average (MA) values of this background noise pattern are shown as blue lines in Figure 5.1. In the 300 s window, the MA shows slight fluctuations around the value of 1.78. In general, the range of fluctuations is small (within 0.2) and the duration

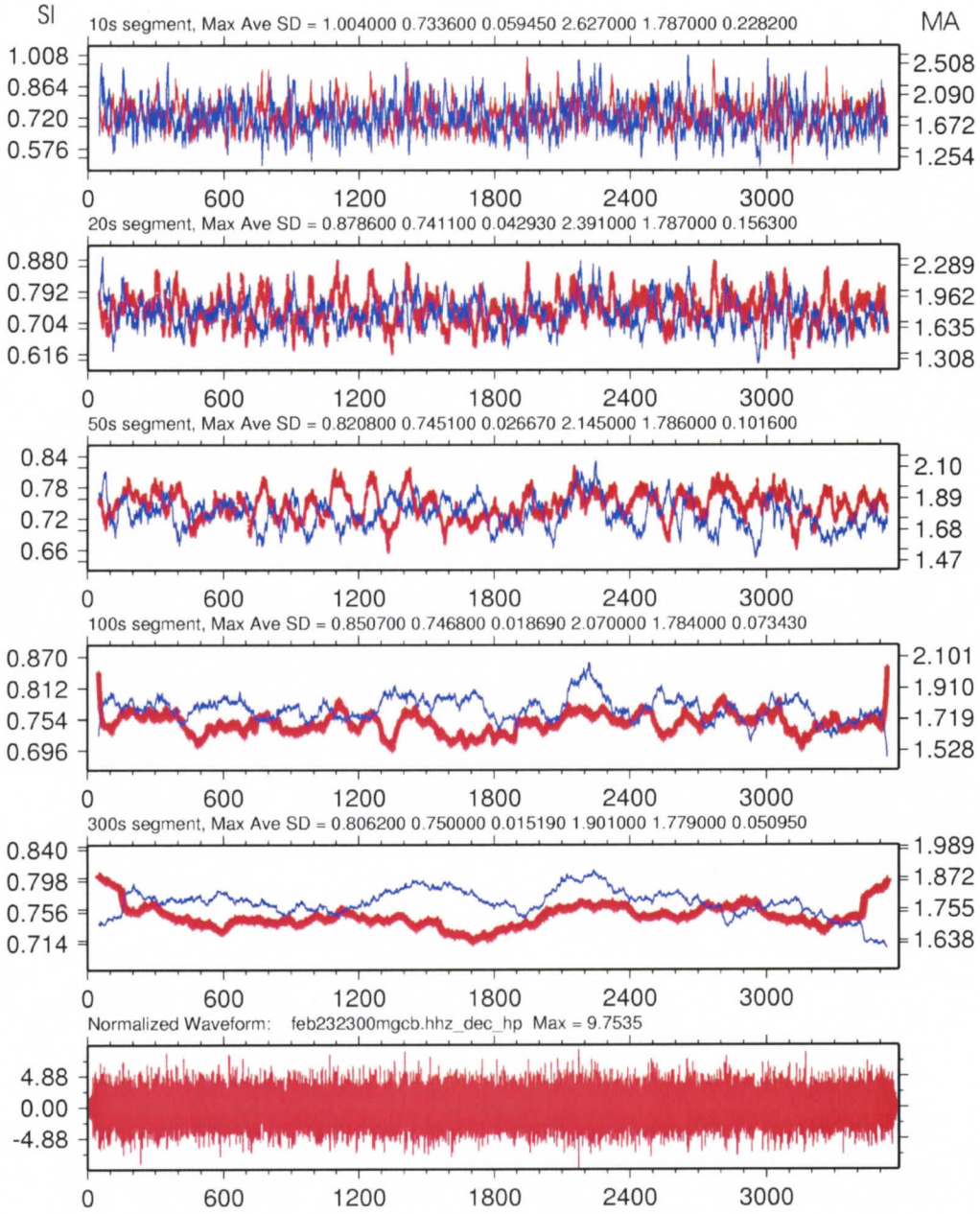

Figure 5.1
An example of a background noise pattern at station MGCB for the hour beginning at 23:00:00 on February 23, 2003. The bottom graph shows the original waveform high-passed at 1.5 Hz, and normalized. The x-axis runs from 0 - 3600 s (0 s is the start at 23:00:00). The layout is the same as that of Figure 4.5.

of the fluctuations is generally longer than the 300 s length of the time window (e.g. from the peak at 1450 s to the trough at 1920 s). As the length of time window decreases, the fluctuations increase in both amplitude and frequency. This is evident from the 100 s, 50 s, 20 s, and 10 s windows shown in Figure 5.1. Notice that no matter how the pattern of MA fluctuation changes with the length of time window, the MAV remains practically unchanged at approximately 1.78.

The corresponding SI values are shown as red lines in Figure 5.1. The SI is most sensitive to a change of waveform pattern. In the case of the background noise, the longest change of pattern occurs as the trace begins and ends. This is especially evident from the 300 s time window in which a sudden drop and rise exist in the first and last 200 s of the trace, respectively. These are computational artefacts because for a time equal to half the time window at the beginning and end of each trace, the calculation cannot be fully made. The average SI value is 0.750 with the peak being 0.806. Similarly the 100 s time window shows a drop in the first 50 s followed by a relatively flat pattern with a sharp rise in the last 50 s. The average and peak SI values are 0.747 and 0.851 respectively. As the length of window gets shorter, the SI values show more fluctuations. This is mainly due to the fact that the change of waveform pattern toward the two ends becomes comparable to the rest of the trace as the length of time window decreases. As a result, SI values show many peaks indicating local pattern changes within the trace (e.g. red lines within the 50 s, 20 s, 10 s windows in Figure 5.1). Nevertheless, the fluctuations are still small in amplitude. For example, the average value for the 10 s window is 0.73 with a peak value of 1.00.

In summary a pattern of background noise can be characterized by relatively flat SI and MA curves with exceptions towards the two ends of the trace. As the time windows decrease in size, both SI and MA show more fluctuations. However, the fluctuations remain relatively small (± 0.2) with a virtually constant average value.

5.2 Characteristics of a Spike/Earthquake

Figure 5.2 shows a typical example of a spike/earthquake pattern. The waveform has a pattern similar to background noise, but with very low amplitude (< 0.5) until the first arrival of the P-wave. In the case presented, the P-wave arrives at approximately 2250s. This is marked by a sharp increase in the amplitude to a maximum of 11.24. The S-wave arrival is picked up at 2300 s. Its amplitude is approximately 6.00. During the next 150 s the coda wave of the earthquake is observed as a gradual decrease in amplitude. The whole event lasts approximately 200 s.

The MA values for the spike/earthquake are shown as blue lines in Figure 5.2. In the 300 s time window, the pattern shows a very low value (< 0.15) from the start of the trace to 2100 s when the contribution of P phase starts to influence the MA value. In other words, the rise of MA values precedes the actual beginning of the event by a time equal to half the size of the time windows. The value steadily increases to a maximum value of 0.655 at 2400 s, and then drops rapidly in the next 100 s to approximately 0.2. At 2800 s the MA value decreases to its initial state (<0.15) and remains small to the end of the trace.

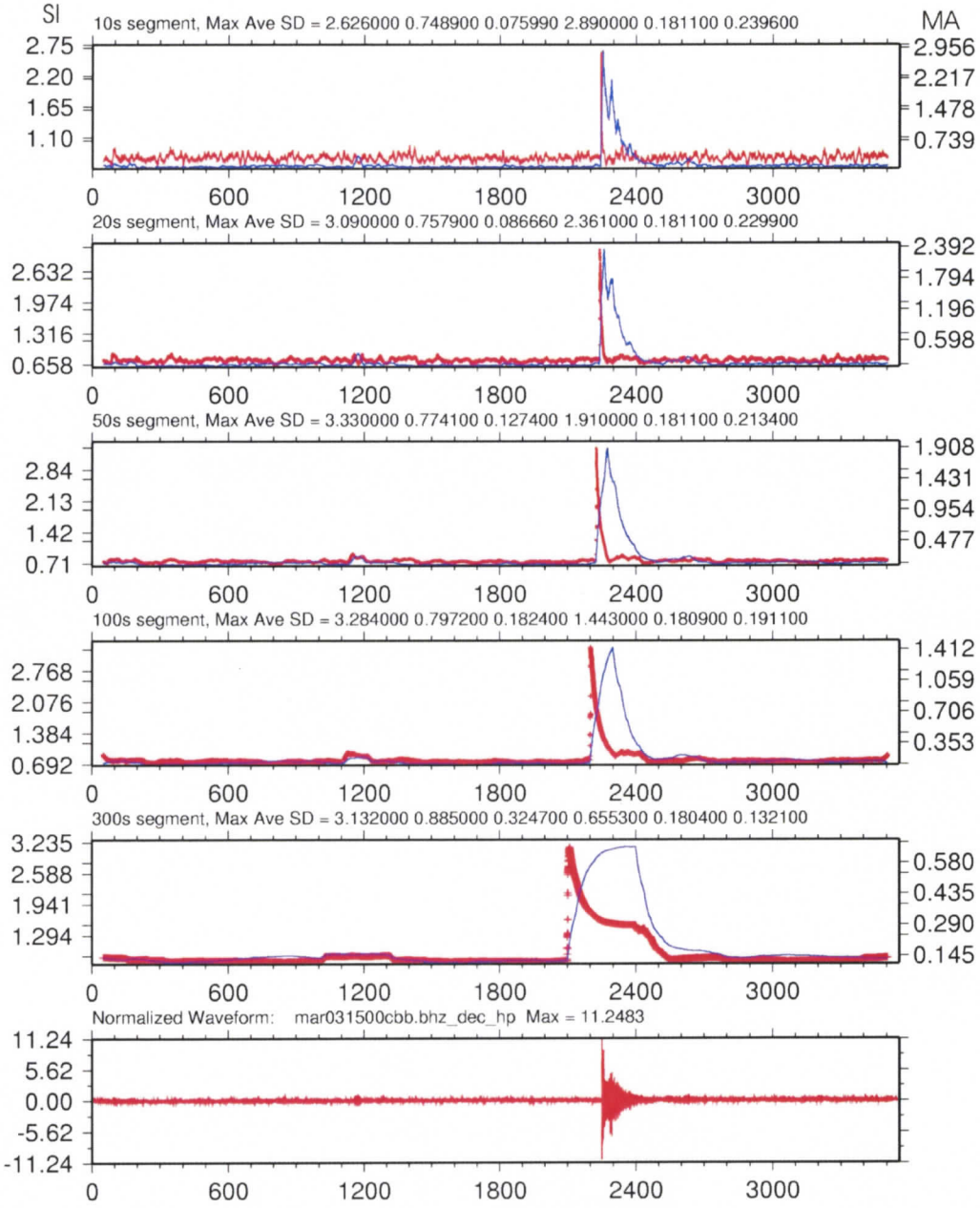

Figure 5.2
An example of a spike/earthquake pattern at station CBB for the hour beginning at 15:00:00 on March 03, 2003. The bottom graph shows the original waveform high-passed at 1.5 Hz, and normalized. The x-axis runs from 0 - 3600 s (0 s is the start at 15:00:00). The spike event begins at 2250 s and ends at approximately 2400 s. The layout is the same as that of Figure 4.5.

The MAV for this time window is 0.18. The 100 s time window shows a similar pattern but with a narrower peak due to the smaller time window. The first arrival is detected at 2200 s (now a 50 s difference relative to the P phase). The MA shows a rapid increase to a maximum value of 1.44 at 2300 s followed by a similar rapid decrease to approximately 0.2 by 2400 s. The MAV for the trace remains at 0.18. As the size of the time window decreases, the width of the peak becomes narrower and aligns better with the real time arrival time of the P-wave (as can be seen in the 50 s, 20 s, 10 s time windows in Figure 5.2). The S-wave arrival can also be seen to create a second peak of smaller amplitude approximately 25 s behind the P-wave peak as the time windows decrease in size (this is evident in the 20 s and 10 s windows in Figure 5.2). For all the time windows the MAV remains practically unchanged at 0.18.

The corresponding SI values are shown as the red lines in Figure 5.2. They rise dramatically when the first arrival is detected, but otherwise the values remain low at approximately 0.8. The 300 s time window demonstrates this very well. As the P-wave arrival (a large pattern change which the SI picks up) is detected at 2100 s, the SI increases instantly to a maximum value of 3.13. By 2200 s the SI has decreased sharply to approximately 1.8. The SI remains relatively constant at this value until 2400 s and by 2500 s it decreases back to the low value of approximately 0.8. The average value for the SI is low at 0.88. As the length of the time windows decreases, the similar pattern of rapid rise and fall are still visible, but over shorter periods of time. The peaks also begin to align more directly with the arrival of the P-phase (e.g. red lines in 100 s, 50 s, 20 s, 10

s). The 10 s time window simply exhibits a very narrow sharp peak with a maximum value of 2.63.

Similar SI and MA characteristics can also be seen for teleseismic events. Figure 5.3 shows an example from an earthquake (magnitude 7.7) that occurred at Koryakia, in the eastern region of Siberia, Russia on April 20, 2006. The event was a shallow earthquake at a depth of 43 km. It can be seen in Figure 5.3 that the high-pass filtered waveform resembles a local earthquake without a clear S-phase or surface waves. This is mainly due to the attenuation of high-frequency energy as shear waves propagate across teleseismic distances to the recording stations.

In summary, a relatively flat line can characterize the pattern of an earthquake/spike, with a very large peak immediately ahead of the arrival of P-phase for both MA and SI values. The longer the time window, the longer the time lags. The 10 s and 20 s windows for both MA and SI show the peak much more accurately. In addition, the S-phase can be detected as the second peak of MA values in the 10 s and 20 s time windows. Both the MA and SI show low average values that remain relatively constant, with significantly higher maximum values.

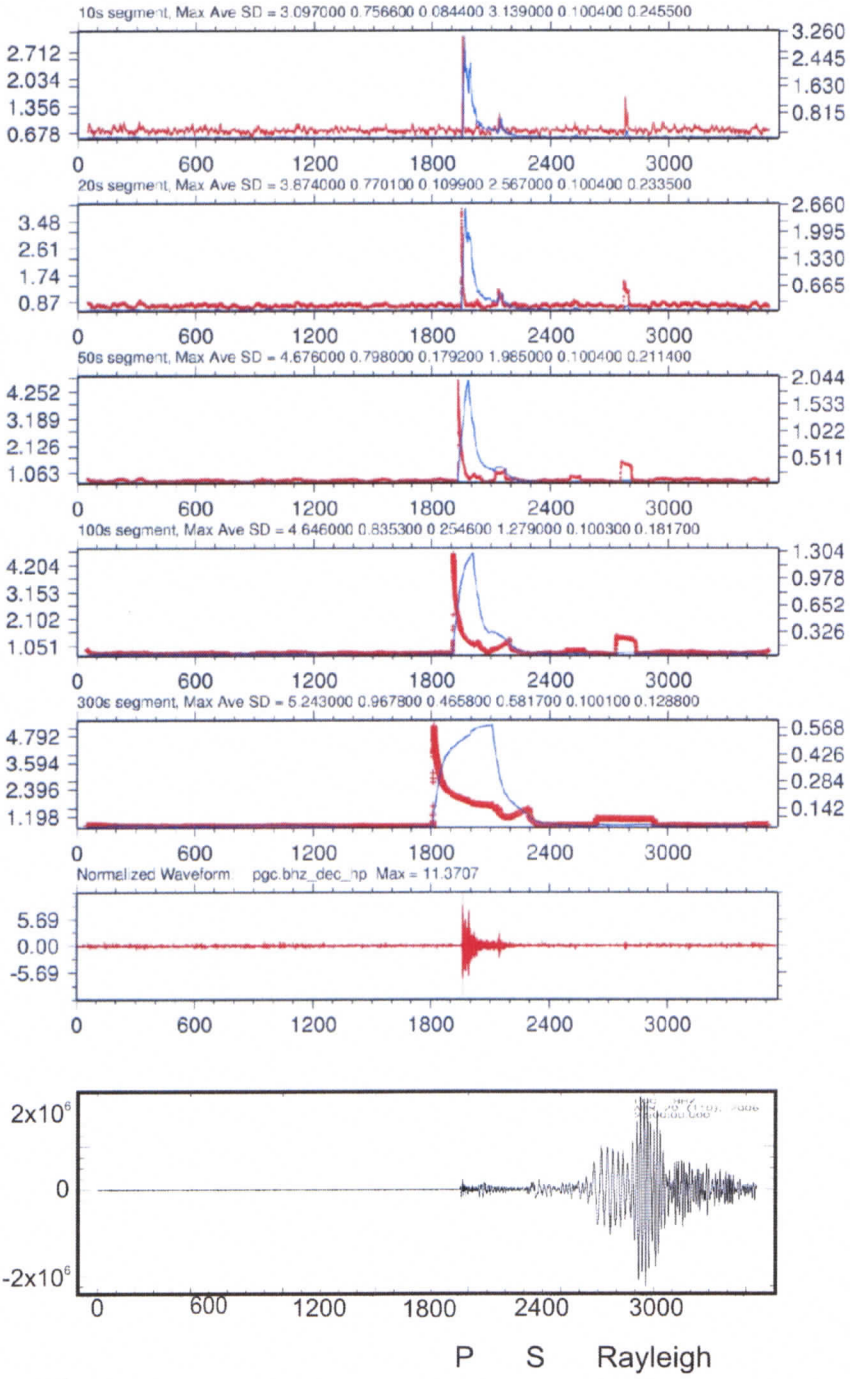

Figure 5.3

An example of a teleseismic earthquake at station PGC for the hour beginning at 23:00:00 on April 20, 2006. The lower graph shows the original broadband waveform. The labels P, S and Rayleigh identify the P-wave, S-wave, and Rayleigh wave arrivals respectively. The graph above takes the same layout as Figure 4.5. The x-axis runs from 0 - 3600 s (0 s is the start at 23:00:00). The first motion for this event begins at 1850 s and ends at approximately 2200 s.

5.3 Characteristics of Tremors

Figure 5.4 shows a typical example of tremor patterns. The waveform begins with small amplitude fluctuations averaging approximately 2.3. At 600 s the amplitudes start to increase. By 900 s the average amplitude is approximately 5.0. By 1300 s, (i.e. the peak of the tremor) the maximum amplitude is greater than 9.0. The amplitude then slowly decreases. By 1600 s the average amplitude is approximately 4.0-5.0 again. By 2000 s amplitudes are back to the starting level of approximately 2.3. There is no clear definition for the start of the tremors. Unlike the earthquake/spike pattern, tremors do not have a clear P-wave first arrival, followed by the S-wave and then the earthquake coda. The tremor can also last a lot longer than the earthquake/spike pattern.

The MA values of this tremor example are shown as blue lines in Figure 5.4. Using a 300 s time window, the MA values approximately follow the envelope of the tremor event. The values rise from 0.10 at 0 s gradually to a peak value of 1.72 at 1200 s. The MA stays at this peak until 1400 s, and then gradually declines back down to 0.20 at 2600 s. A small rise to 0.70 is observed at 3000 s where it remains to the end of the trace. The 100 s window begins at a value of approximately 0.44 and does not begin to rise until 600 s. It rises to a small peak of 1.70 at 1000 s and then falls to a small trough with a value of 1.35 at 1200 s. The MA value then rises to the largest peak (2.08) at 1300 s before falling to approximately 0.50 by 1900 s where it fluctuates by ±0.20 for the rest of the trace. The smaller time windows show very similar patterns, but with a greater correlation to the waveform. The largest peak is always observed at 1300 s with the rise starting at 600 s

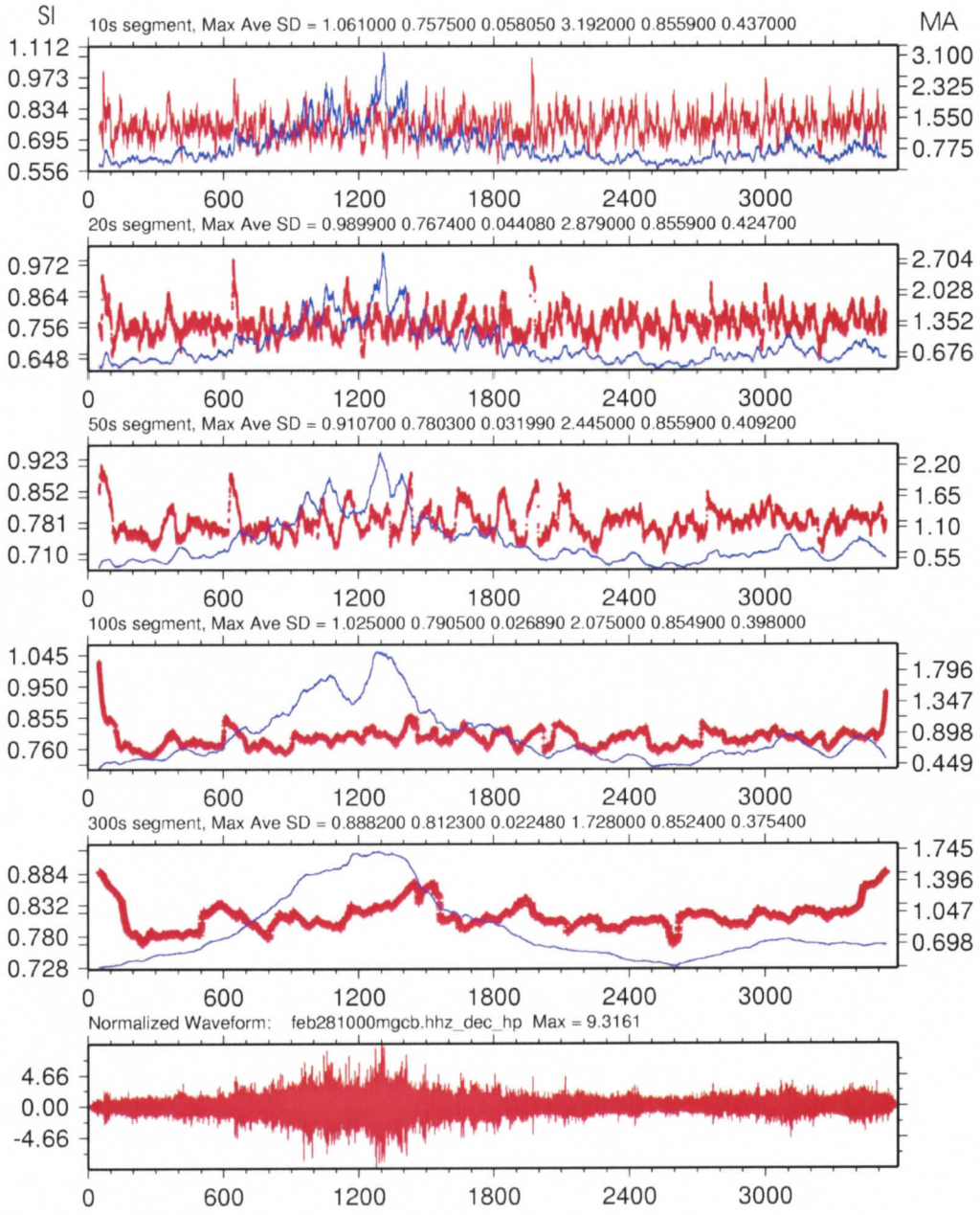

Figure 5.4
An example of a tremor pattern at station MGCB for the hour beginning at 10:00:00 on February 28, 2003. The bottom graph shows the original waveform high-passed at 1.5 Hz, and normalized. The x-axis runs from 0 - 3600 s (0 s is the start at 10:00:00).The tremor pattern begins at 600 s, peaks at 1300 s and ends by 2000 s. The layout is the same as that of Figure 4.5.

and the decline finished by 1800 s (see blue lines in the 50 s, 20 s, and 10 s time windows in Figure 5.4). For all window lengths, the MAV for the tremor pattern remains constant at approximately 0.85.

The SI values are shown as red lines in Figure 5.4. The SI detects changes of patterns. Similar to the example of background noise, there are large changes at the beginning and end of the waveform that are artefacts. This is very evident in the 300 s time window where a sudden drop and rise are visible in the first and last 200 s, respectively. The SI, exhibiting a slightly broad peak in both cases, detects the beginning of the tremor at 600 s and partial end at 1500 s. The rest of the pattern fluctuates about the average value of 0.812. The 100 s time window shows the same artefacts in the first and last 50 s. The rest of the trace fluctuates about the average of 0.79 with no apparent peaks. As the length of the time windows decreases, the fluctuations can be seen to increase in amplitude about the average value of approximately 0.76. The peaks at 600 s and 2000 s signifying the beginning and end of the tremor signal, respectively, have maximum values in the region of 0.95 – 1.06 as shown in Figure 5.4.

To summarize, the MA tends to mimic the envelope of the tremor waveform and the SI shows a fluctuating pattern with peaks at the beginning and end of both the tremor signal. The MAV remains relatively constant at 0.85.

5.4 Typical Values for MAV and SIR

The MAV (the average of all the MA values across the entire segment) and the SIR (the ratio of the maximum value of the SI to the average value of the SI across entire segment) are the two values of most interest to this study. They are defined as:

$$MAV = \left[\sum_{i=1}^{K} MA(i) \right] / K \qquad (5.1)$$

$$SIR = SI(\max) / \left(\left[\sum_{i=1}^{K} SI(i) \right] / K \right) \qquad (5.2)$$

where K is the total number of samples.

In the previous three sections, ideal examples of each group were presented. The five different time windows were analyzed and the 10 s time window was chosen as the most effective for distinguishing the three events. The 10 s time window reduces the extent of artificial edge effects, which are clearly seen in the 100 s and 300 s windows. It also has the most contrasting MAV and SIR values. The corresponding MAV and SIR values for the 10 s time windows in figure 5.1, 5.2 and 5.4 are listed in Table 5.1.

Table 5.1: *SIR and MAV values for the examples seen in Figures 5.1, 5.2, and 5.4.*

	SIR	**MAV**
Background Noise	1.39	1.79
Spike/Earthquake	3.51	0.18
Tremor	1.40	0.86

For background noise, the SIR value is similar to that of the tremor, but much lower than that of the spike/earthquake. The MAV, on the other hand, is far higher than that of the spike/earthquake and about twice as large as that of the tremor. The higher value of MAV for background noise results from the normalization scheme, as the whole trace has no large peaks.

The spike/earthquake has a much higher value for the SIR and an extremely low value for the MAV. The spike/earthquake should always have a large SIR as it creates a very large peak in the SI values due to the dramatic change of waveform pattern from very low background noise to the large amplitude of the event. The low MAV is also due to the normalization scheme.

The tremor example has a SIR value similar to that of background noise (i.e. both of which are much lower than that of the spike/earthquake). Meanwhile, the MAV is

relatively low compared to that of the background noise, but much higher than that of the spike/earthquake. This is again due to the normalization scheme, as the normalization for the tremor decreases the values of the background noise in the waveform.

Figures 5.1, 5.2, 5.4 are three very ideal examples with ideal values. To obtain the full range of these values, the whole 2003 tremor sequence was processed and then classified visually. The visual classification was a time consuming and subjective process. Of the waveform patterns, the spike/earthquake hours are the hardest to classify visually as often a tremor can look like signal with one or more spikes. In other cases, a pattern of background noise can be misidentified as being a spike/earthquake hour due to the existence of one extremely high reading.

To make the comparison meaningful, each station-hour waveform segment is examined independently from the others. This approach is different from that of previous studies in which tremors are recognized only when they are coherent across a number of nearby stations [Rogers and Dragert, 2003; Kao et al., 2006]. Approximately 744 hours of data was visually examined from over 22 seismic stations. The entire visual classification took over 200 human hours. In total, the 2003 tremor sequence between 22 February 2003 and March 27, 2003 yielded 5894 hours classified as having a pattern of background noise, 2477 hours having a pattern of spikes/earthquakes and 6788 hours having a pattern of tremors.

From this visual classification result, the 98% ranges of SIR and MAV values for each category are listed in Table 5.2.

Table 5.2 *Value ranges for background noise, spike/earthquake and tremor patterns (the ranges represent 98% of each type of pattern).*

	SIR range	**MAV range**
Background Noise	1.2 – 1.5	1.4 – 2.5
Spike/Earthquake	1.2 – 6.0	0.0 – 1.4
Tremor	1.2 – 1.7	0.0 – 1.6

5.5 Testing the Theory

Figure 5.5 shows the distribution of SIR and MAV values of all data samples. To better display how the distribution of tremor samples overlaps with that of background noise, the tremor samples are plotted on top of the background noise samples in Figure 5.5a and vice versa in Figure 5.5b.

It appears waveforms with an earthquake/spike pattern have high SIR values. Therefore in theory, it is possible to set an SIR threshold to identify waveforms showing a spike/earthquake pattern. Using a lower SIR threshold will increase the percentage of correct identification for spikes/earthquakes, but will also decrease the accuracy for patterns of tremors and background noise. On the other hand, if we increase the threshold

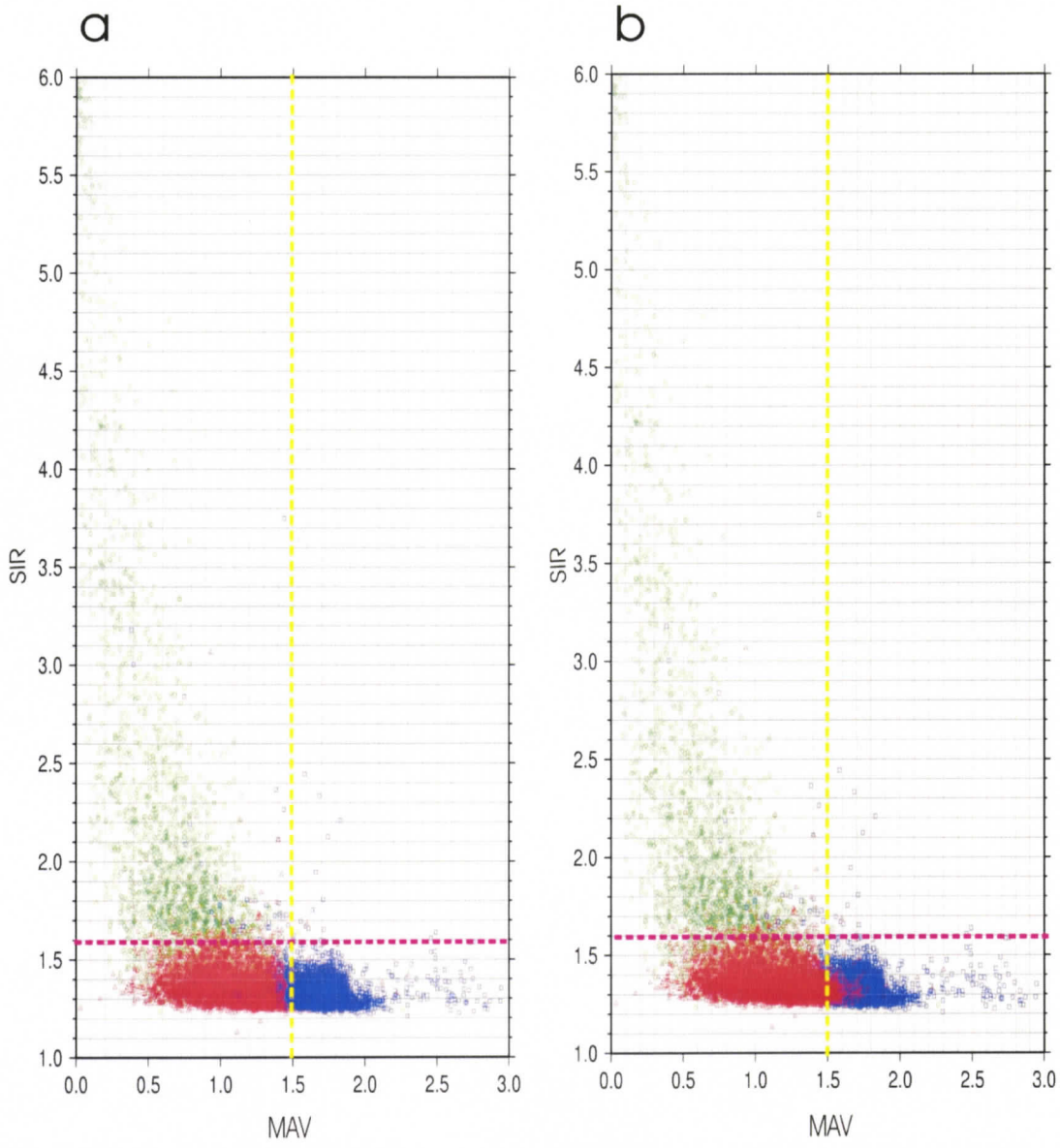

Figure 5.5
Graphs of MAV vs. SIR values for all data from the 2003 tremor sequence. Tremor, background noise and spike/earthquake are represented by red triangles, blue squares and green circles, respectively. The yellow line represents the optimum MAV cut-off value of 1.5 to separate tremors from background noise. The purple line represents the optimum SIR cut-off value of 1.6 to separate spike/earthquake pattern from background noise and tremor patterns. In (a) background noise pattern is the top layer, while in (b) the tremor pattern is the top layer.

of SIR value we can easily identify 100% of background noise and tremor hours correctly. But the drawback is that the number of spikes/earthquakes classified in the wrong categories increases dramatically. Table 5.3 summarizes the percentages of correctly identified events with different SIR thresholds. It appears that a threshold value of 1.6 can produce the optimum compromise, with most tremor and background noise samples below this threshold and most spike/earthquake samples above.

Table 5.3 Various SIR values and the corresponding percentages of background noise (BG), tremor (TR) and spike/earthquake (SP) correctly classified.

PERCENTAGES						
BG > SIR	**TR > SIR**	**SP > SIR**	**SIR Value**	**BG < SIR**	**TR < SIR**	**SP <SIR**
11.8	29.6	95.1	***1.4***	88.2	70.4	4.9
4.7	18.5	92.7	***1.45***	95.3	81.5	7.3
2.1	11.3	89.1	***1.5***	97.9	88.7	10.9
1.3	6.5	85.2	***1.55***	98.7	93.5	14.8
1.0	3.5	**82.2**	***1.6***	**99.0**	**96.5**	17.8
0.8	1.8	77.9	***1.65***	99.2	98.2	22.1
0.6	1.2	70.9	***1.7***	99.4	98.8	29.1

With the SIR value set at 1.6, 99% of background noise hours and 96.5% of tremor hours are available to be classified. The MAV threshold must now be examined. It turns out that the lower the MAV, the more background noise hours will be classified correctly. For example, setting the MAV at 1.4 will identify almost 99% of background noise hours, but only 81.9% of tremors. In contrast a higher MAV value of 1.6 can correctly identify 95.4% of tremor hours, but at the expense of wrongly identifying many background noise hours (79.2%). By trying several values in-between, it appears that a

MAV of 1.5 gives the optimum identification result, with 94.9% of background noise and 92.3% of tremor hours correctly identified (see Table 5.4).

Table 5.4 Various MAV values and the corresponding percentages of background noise (BG), tremor (TR) and spike/earthquake (SP) correctly classified.

Percentages (if SIR < 1.6)						
BG <MAV	**TR < MAV**	**SP < MAV**	**MAV Value**	**BG > MAV**	**TR > MAV**	**SP > MAV**
0.3	81.9	16.7	***1.4***	98.7	14.6	1.2
0.9	88.3	17.2	***1.45***	98.2	8.2	0.6
4.2	**92.3**	17.4	***1.5***	**94.9**	4.2	0.4
10.9	94.2	17.6	***1.55***	88.1	2.3	0.2
19.8	95.4	17.7	***1.6***	79.2	1.1	0.2

Based on these results, a logic flow can now be constructed with the optimum values to classify all three patterns. Figure 5.6 shows this flow chart. The SIR value is first examined; if it is greater than 1.6 then the pattern is classified as a spike/earthquake; if not, the MAV is examined. If the MAV is less than 1.5 the classification is a tremor pattern; if the MAV is greater the classification result is background noise.

Using this flow chart, the spike category will yield the least percentage of correct identification (82%) but the spike is also the most ambiguous of all three to classify. Using these values, 92% of all tremors and 95% of all background noise hours are correctly identified. Figure 5.6 also shows the percentage of each event classified by the logic flow.

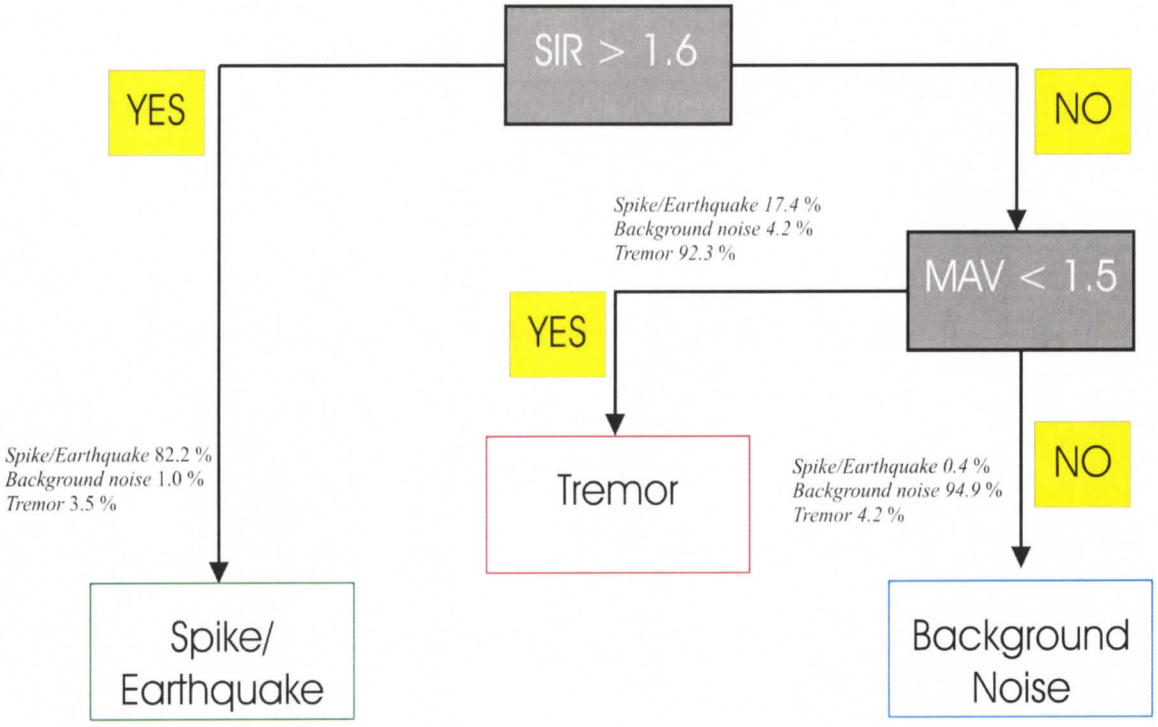

Figure 5.6
Logic used for classification. Each hour segment enters at the top of the chart. If the segment has a SIR value greater than 1.6 it is classified as having a spike/earthquake pattern. If it is not greater than 1.6, it moves to the right and the second logic gate is encountered. If the MAV is less than 1.5, the hour is classified as having a tremor pattern; if greater than 1.5 it is classified as being background noise. The corresponding percentages of events are also included.

5.6 Application to the 2003 tremor sequence

In this section, the method shown in Figure 5.6 is applied to the 2003 tremor sequence. By visual inspection, the tremor sequence was determined to begin on 25 February 2003 and end approximately on 22 March 2003. To have a comparison with non-tremor days, 3 days before and after this time period were also analyzed.

The processing procedures and logic flow previously described were used to classify all the data. Available waveforms of all stations are processed chronologically using hour-long segments for the entire time period. Approximately 19,000 station-hour segments were processed and classified using the method.

Figure 5.7 shows the distribution of stations and examples of some corresponding waveforms for the hour beginning at 09:00:00 on March 17, 2003 (i.e. the 21st day of the tremor sequence). The waveforms shown are high-passed at 1.5 Hz to better illustrate the tremor characteristics. The classification results are color-coded, with blue corresponding to background noise, green to spikes/earthquakes and red to tremors. The visual identification of waveforms and corresponding classifications all agree. There are five waveforms with tremor patterns (three on the bottom row, one on the lower-middle left and one top middle), four with background noise (2 top left and 2 top right) and one that exhibits a spike/earthquake pattern (lower-middle right).

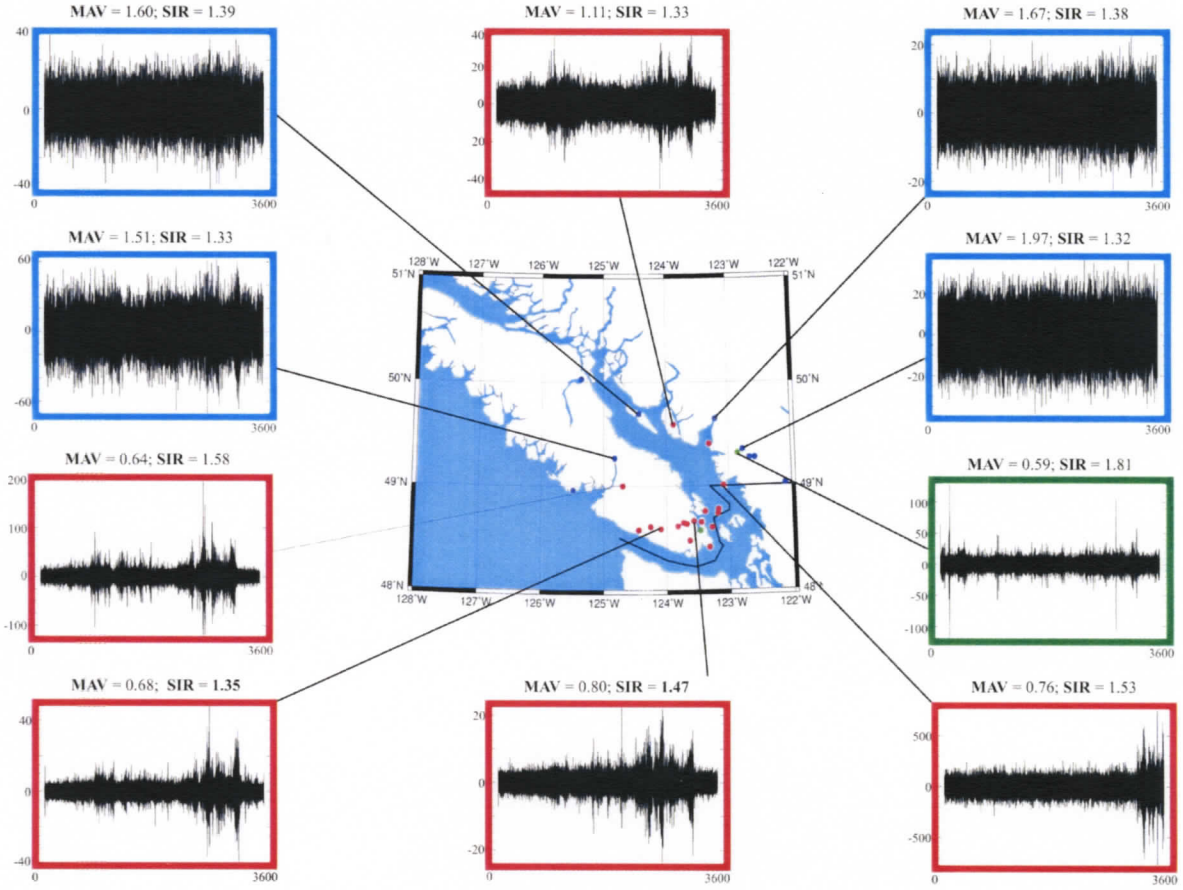

Figure 5.7

Selected examples of waveforms (high pass filtered 1.5 Hz) for the hour beginning at 09:00:00 on March 17, 2003. The stations are colour coded. Red dots represent a station-hour classified as having a tremor pattern, blue dots represent background noise, green dots represent a spike/earthquake. From the top central waveform clockwise, the stations with example waveforms are; BIB, WPB, COQB, ANMB, ENGB, KELB, TWBB, MGB, ALB, TXB. The waveforms have red, blue or green outlines indicating tremor, background noise or spike/earthquake patterns. The x-axis of the waveforms goes from 0 s to 3600 s; the y-axis represents the amplitude and the scale is given. The associated MAV and SIR values for each waveform are directly above each waveform panel.

This same procedure was repeated for all 24 hours segments within that day. Figure 5.8 shows classification results of 5 representative stations for two selected days, one from before the tremor sequence (23 February 2003, two days before the tremor sequence began) and one during the sequence (4 March 2003, the 8th day of the tremor sequence). Red, blue and green squares represent the waveform characteristics of station-hour segments for tremors, background noise and spikes/earthquakes, respectively. One evident feature is the abundance of blue squares from the pre-tremor sequence day and the abundance of red squares in the day during the tremor sequence. These abundances need to be quantified to be able to classify each station for the whole 24 hours into one category, i.e., station-day classification. Twelve hours of the same class per day would be a minimum requirement to mark a station as having a consistent pattern throughout that day. Trials of several different thresholds were conducted (12, 13, 14, 15, 16, 17, 18 hours per day), and the most diagnostic value appeared to be 14. Requiring 18 hours of one category for the station-day classification appears to be too restrictive. On the other hand, requiring only 12 or 13 hours of the same category appears to be too loose and an ambiguity of station-day classification can happen in some cases. The 14-hour threshold classified most stations accurately, while also allowing some flexibility for a few stations to be dominated by local signals of other types.

At station TXB (station 2 in Figure 5.8a) before the tremor sequence there are 19 hours of background noise, 4 hours of tremor and 1 hour of spike/earthquake patterns. Therefore 23 February 2003 would be classified as a background noise day at station TXB. In contrast during the tremor sequence on 4 March 2003, (Figure 5.8b) station TXB exhibits

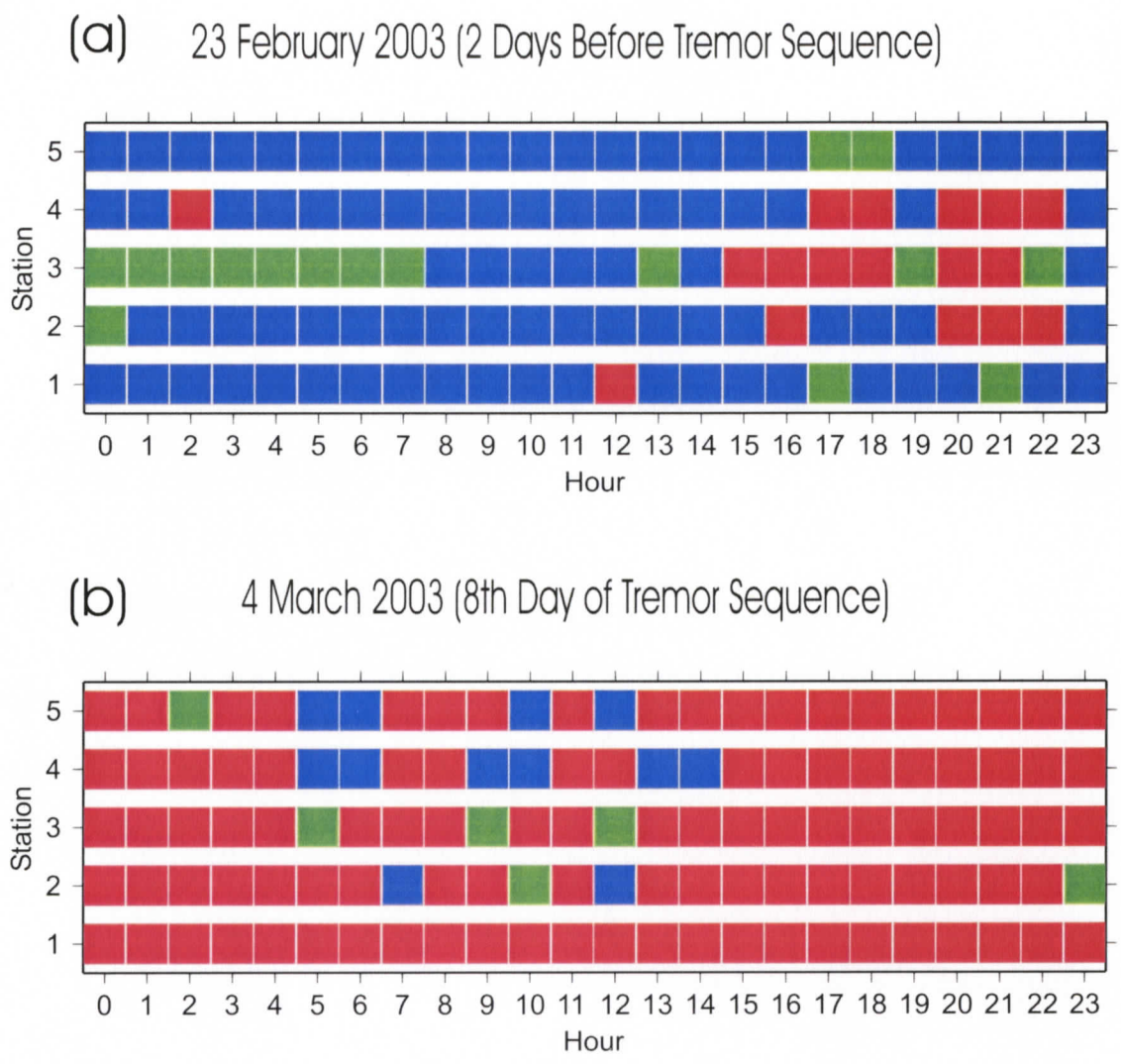

Figure 5.8
Classification results for five representative stations on two different days. Stations 1 - 5 are, PGC, TXB, TWBB, CBB, SNB, respectively, with their hourly classification for two days (24 hour long periods). Classification of tremor, background noise and spike/earthquake are shown by red, blue and green squares, respectively. (a) shows 23 February 2003, 2 days before the tremor sequence and is dominated by background noise classification. (B) shows 4 March 2003, the 8th day of the tremor sequence and is dominated by tremor classifications.

2 hours of background noise, 20 hours of tremor and 2 hours of spike/earthquake patterns. As a result 4 March 2003 would be classified as a tremor day at station TXB. Conversely, if there are no more than 14 hours of one category at one station in a day it will go unclassified. An example of this is station TWBB on 23 February 2003 (station 3 in Figure 5.8a), exhibiting 7 hours of background noise, 6 hours of tremor and 11 hours of earthquake/spike patterns.

The whole 2003 tremor sequence was examined on a daily basis using the station-day classification approach. The pattern expected to be seen should be mainly background noise before 25 February 2003, a high percentage of stations classified as having tremor patterns between 25 February and 22 March 2003, followed by a return to a pattern of background noise for most stations afterwards. The following figures illustrate the daily classification results for 5 selected days of the 2003 tremor sequence.

Figure 5.9 corresponds to 23 February 2003, 2 days before the onset of the tremor sequence. The distribution of blue circles demonstrates that all stations are associated with more than 14 hours of background noise. This is what we expected for a pre-tremor day.

On the first day of the tremor sequence (25 February 2003, Figure 5.10), the pattern was clearly changed and the onset of tremor was clear. The red circles in southeast Vancouver Island represent stations in the class of tremor-day pattern. The stations further north continue to detect background noise. Also fewer stations are classified than in Figure 5.9.

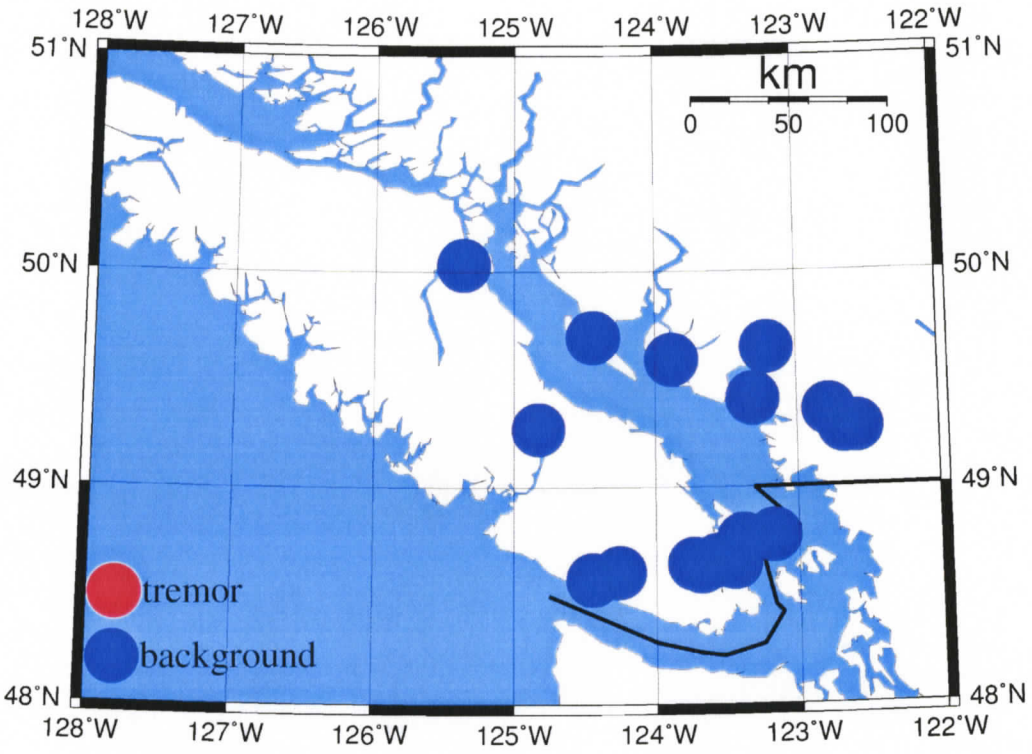

Figure 5.9
Day-classification of stations in the lower BC area for 23 February 2003, two days before the 2003 tremor sequence started. Background noise and tremor classification are shown in blue and red circles respectively if more than 14 hours of the same class can be observed. In this case the pattern is dominated by background noise.

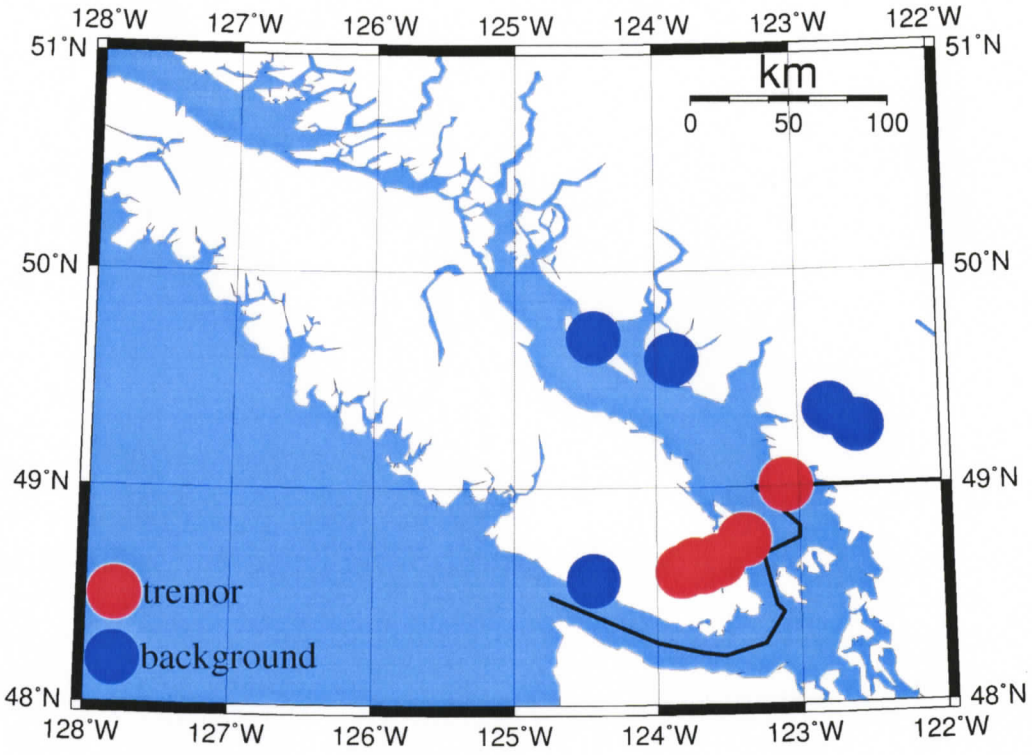

Figure 5.10
Day-classification of stations in the lower BC area for 25 February 2003, the starting day of the 2003 tremor sequence. The layout is the same as that of Figure 5.9. In this case, clear tremor patterns are visible on the south eastern part of Vancouver Island and south western lower mainland, BC. Background noise is still evident north, east and west of the tremors.

This is the pattern one would expect with tremors beginning in the south of the region, background noise to the north and with the onset of tremors being picked up by stations in between. As a result, fewer stations can be classified due to the 14-hour threshold of station-day classification.

Figures 5.11 and 5.12 correspond to the classification results of two selected days during the tremor sequence (9 and 18 March 2003). The number of stations showing a pattern of tremor signal becomes more abundant across the southern Vancouver Island. The main difference between Figures 5.11 and 5.12 is that the region showing tremor patterns on 18 March 2003 appears to migrate more toward the northeast than that on 9 March 2003. This northeastward migration pattern was also confirmed by a detailed tremor relocation analysis [Kao et al, 2005]. The furthest north station (at ~50°N, 125.3°W, Figures 5.11 and 5.12) showed a pattern of background noise for both 9 and 18 March 2003, indicating that the 2003 tremor activity was primarily limited to the southern part of Vancouver Island. Note that for the given station coverage, a gradual increase in the intensity of tremors in a fixed location can give the appearance of tremor migration as the area of detection becomes larger. Expanded network coverage is required to resolve these two possibilities.

Finally on 23 March 2003 (1 day after the tremor sequence stopped, Figure 5.13) all stations showed blue circles, meaning that the tremor sequence was practically over as only the pattern of background noise was observed. This pattern is very similar to the pre-tremor sequence day shown in Figure 5.9.

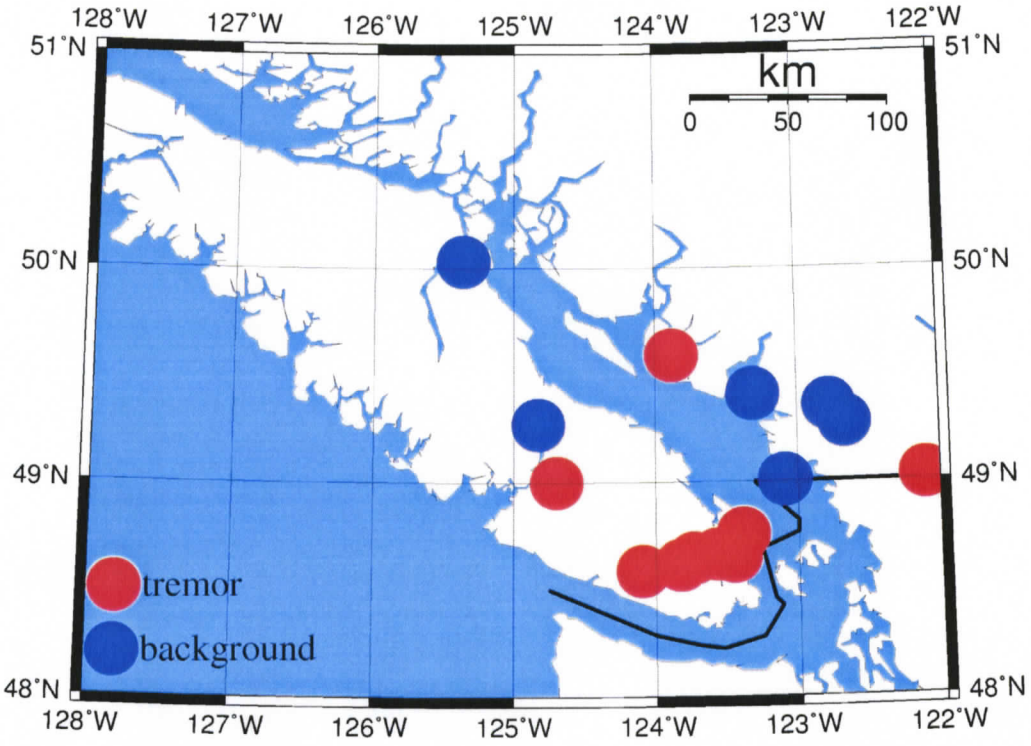

Figure 5.11
Day-classification of stations in the lower BC area for 09 March 2003, the thirteenth day of the 2003 tremor sequence. The layout is the same as that of Figure 5.9. In this case, clear tremor patterns are picked up on the southern part of Vancouver Island and further north to mid-island and across to the parts of the mainland north of Vancouver. Background noise is still evident in the northern area of Vancouver Island and around Vancouver on the mainland.

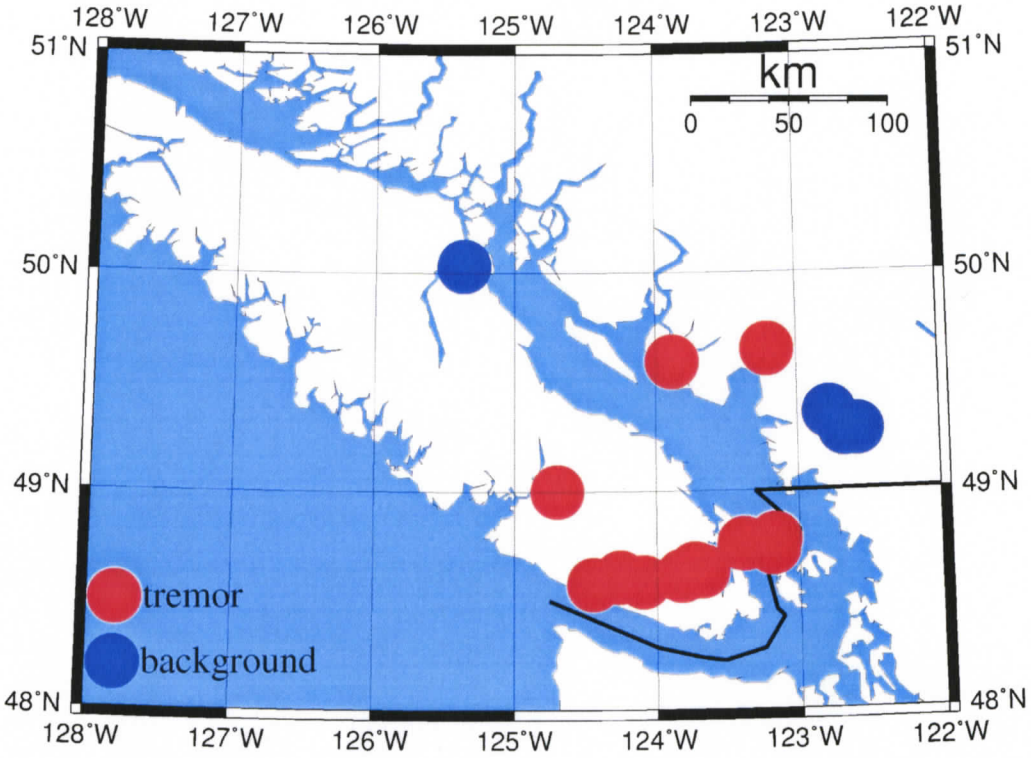

Figure 5.12
Day-classification of stations in the lower BC area for 18 March 2003, the twenty-second day of the 2003 tremor sequence. The layout is the same as that of Figure 5.9. In this case, clear tremor pattern is picked up on the southern and central part of Vancouver Island and across to the mainland, north of Vancouver. Background noise is still evident in the Vancouver area and the northern part of the island.

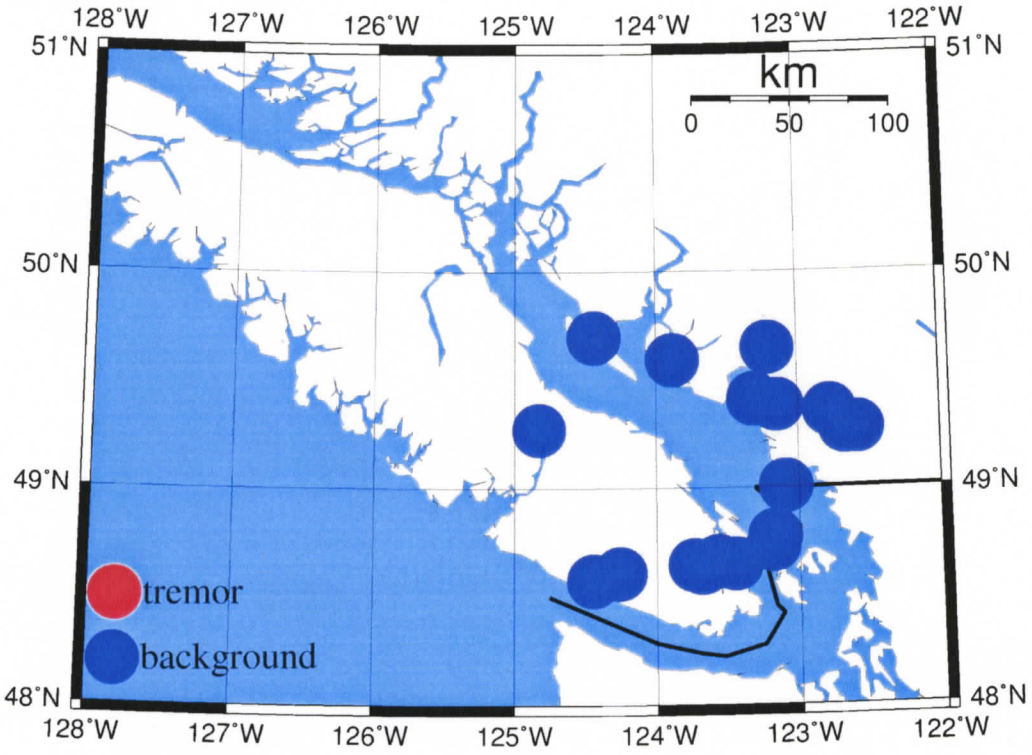

Figure 5.13
Day-classification of stations in the lower BC area for 23 March 2003, one day after the end of the 2003 tremor sequence. The layout is the same as for that of Figure 5.9. In this case there is no longer tremor evident and the map is dominated by background noise throughout.

CHAPTER 6

UNCERTAINTY AND ERRORS

The success of an automatic tremor identification algorithm depends on, at least partially, whether the results are consistent with that of visual inspection. In this chapter, I will revisit the misclassified waveforms, which were evident from the flow chart in Figure 5.6 and discuss the possible reasons. I will also present a direct comparison between the results of the visual inspection in the previous study to demonstrate that the two are consistent to each other.

6.1 Examples of inconsistent identification and possible reasons

As can be seen from the flow chart, not all hours are identified correctly. 17% of all spikes and 4% of background noise are incorrectly classified as tremors, 4% of all tremors and 1% of all background noise are classified as spikes, and 4% of all tremors are classified as background noise. In this section an investigation of possible reasons why these hours are not correctly classified is presented. Some of these may legitimately be given the correct visual classification, but they do not fall into the correct category using the computerised method and vice versa. Some are simply very ambiguous and the results are therefore open to individual judgements.

Example 1: An hour-long seismogram beginning at *04:00:00 February 26, 2003 at station ANMB; a pattern of spike/earthquake by visual classification but classified as a pattern of tremors by the automatic process (Figure 6.1).*

This is an example of one of the 17% of spike/earthquake hours that fall into the tremor category with an SIR of 1.49 and an MAV of 0.89. When carefully inspected, the trace can be arguably re-classified as a pattern of tremors if the two very clear spikes of much greater amplitude at 750s and 3100s are ignored. In this particular case, the tremor-like signal throughout the trace obviously overwhelms the two spikes and dominates the SIR and MAV calculations.

Example 2: An hour-long seismogram beginning at *09:00:00 February 24, 2003 at station GOWB; a pattern of background noise by visual classification but classified as a pattern of tremors by the automatic process (Figure 6.2)*

This example is one of the 4% of background noise hours that fall into the tremor category with an SIR of 1.38 and an MAV of 1.32. There seems to be small tremor events at 500 s and 1500 – 2000 s, but the corresponding amplitudes are small. Obviously, these small tremor amplitudes are sufficient to be picked up by the automatic process. To say the least, this case should be considered ambiguous.

Example 3: An hour-long seismogram beginning at *12:00:00 March 2, 2003 at station ANMB; a pattern of tremors by visual classification but classified as a pattern of spikes/earthquakes by the automatic process (Figure 6.3).*

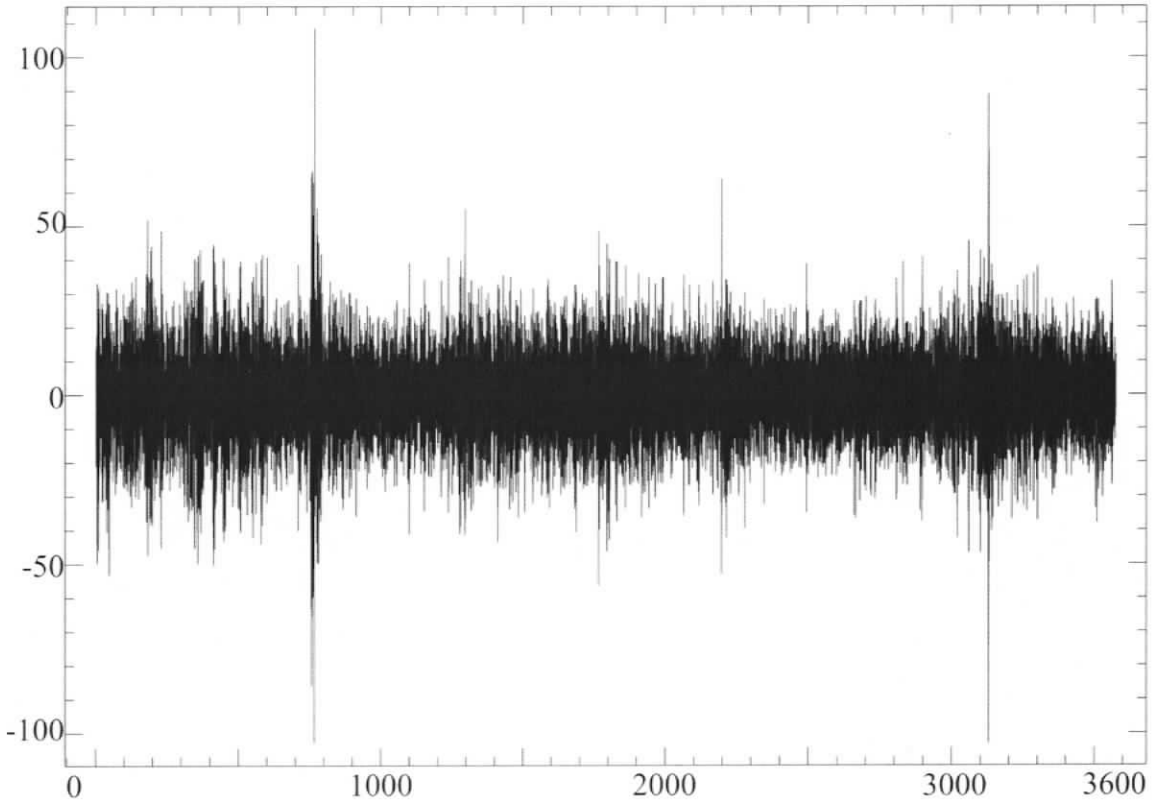

Figure 6.1
High pass filtered waveform of the hour beginning at 04:00:00 from station ANMB, 26 February 2003. The x-axis represents time in seconds from 0 - 3600. The y-axis represents the amplitude of the waveform.

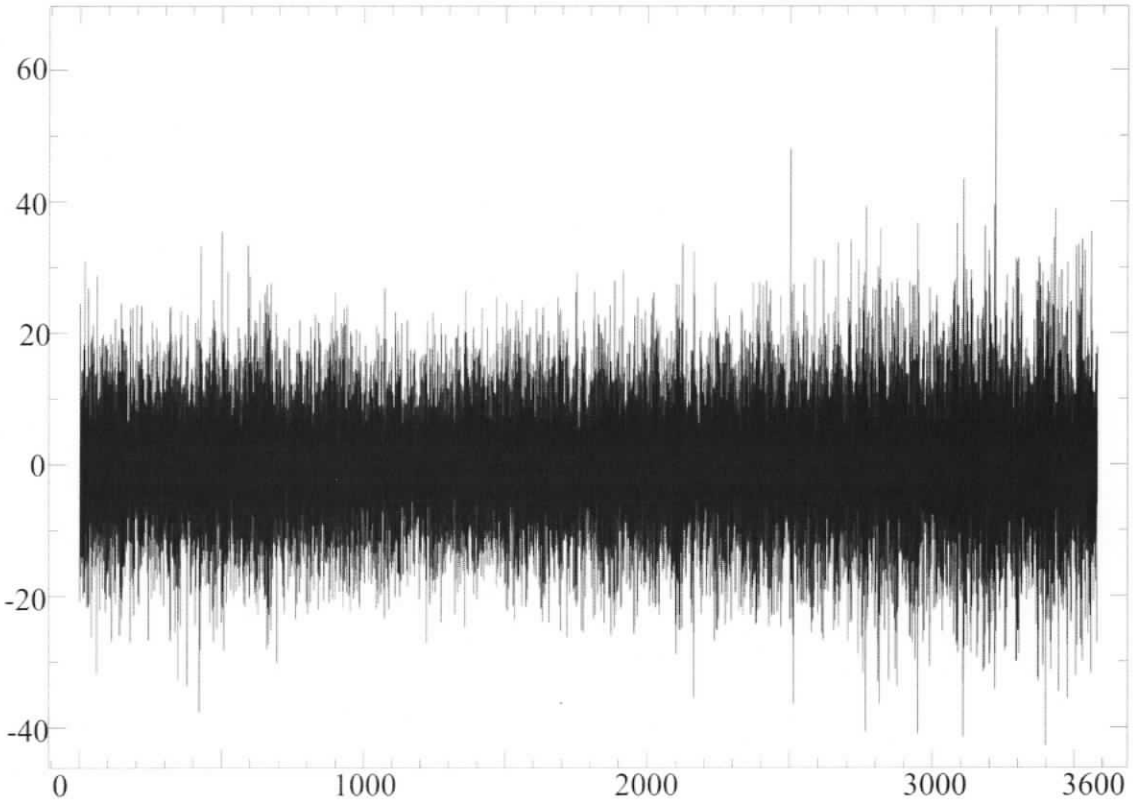

Figure 6.2
High pass filtered waveform of the hour beginning at 09:00:00 from station GOWB, 24 February 2003. The x-axis represents time in seconds from 0 - 3600. The y-axis represents the amplitude of the waveform.

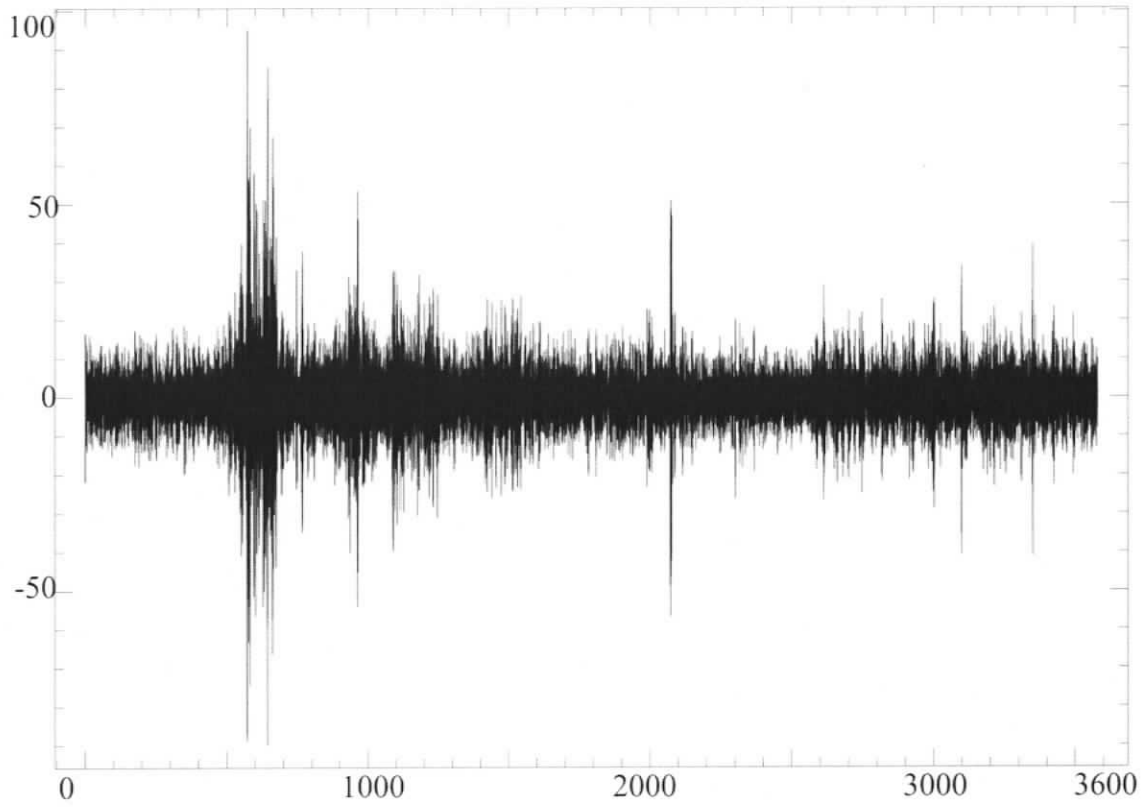

Figure 6.3
High pass filtered waveform of the hour beginning at 12:00:00 from station ANMB, 02 March 2003. The x-axis represents time in seconds from 0 - 3600. The y-axis represents the amplitude of the waveform.

This example is one of the 4% of tremor hours that falls into the spike/earthquake category with an SIR of 1.92 and an MAV of 0.58. When inspected closely, the trace appears to have several tremor events between at 500 s and 1600 s. Several large amplitude spikes are present in the waveform. These high amplitudes obviously dominate the calculations of MAV and SIR values, and result in the classification of a spike/earthquake pattern.

Example 4: An hour-long seismogram beginning at *00:00:00 March 16, 2003 at station ENGB; a pattern of background noise by visual classification but classified as a pattern of spikes/earthquakes by the automatic process (Figure 6.4).*

This example is one of the 1% of the background noise hours that falls into the spike/earthquake category with an SIR of 1.69 and an MAV of 1.55. The trace appears to exhibit clear background noise properties, with the exception of several peak values between 600 s and 700 s. These few high amplitudes dominate the calculations of SIR and cause misidentification.

Example 5: An hour-long seismogram beginning at *15:00:00 March 7, 2003 at station GOWB; a pattern of tremors by visual classification but classified as a pattern of background noise by the automatic process (Figure 6.5).*

This example is one of the 4% of tremor hours, which falls into the background noise category with an SIR of 1.40 and an MAV of 1.61. The trace appears to have a possible tremor pattern from 0 s to 1900 s with relatively small amplitudes. This time, in contrast to the example shown in Figure 6.3, the amplitudes are not significant enough to be recognized as a tremor pattern by the automatic process.

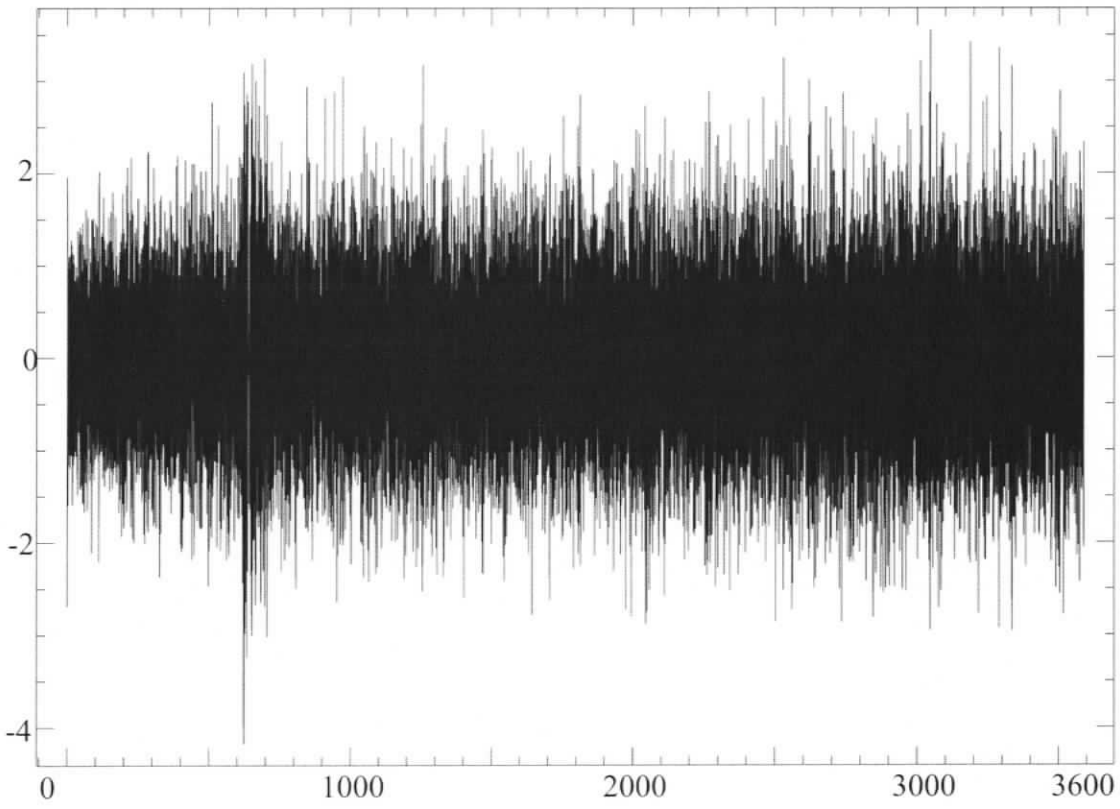

Figure 6.4
High pass filtered waveform of the hour beginning at 00:00:00 from station ENGB, 16 March 2003. The x-axis represents time in seconds from 0 - 3600. The y-axis represents the amplitude of the waveform.

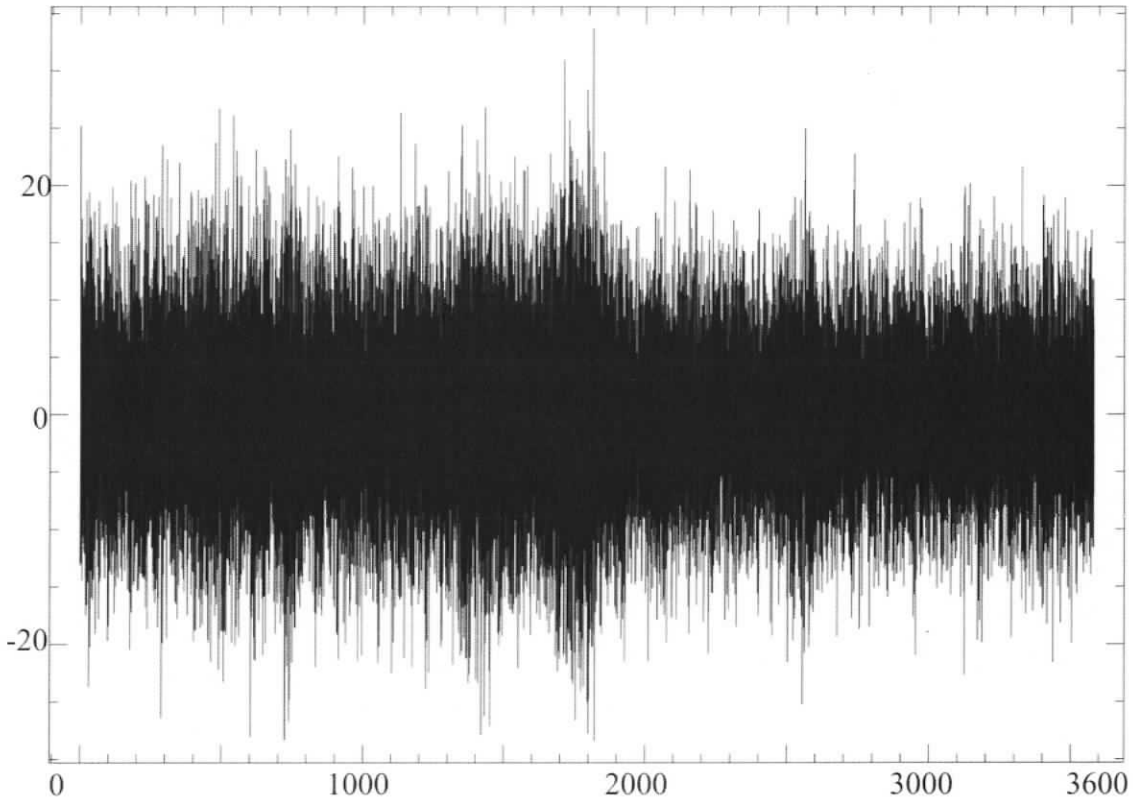

Figure 6.5
High pass filtered waveform of the hour beginning at 15:00:00 from station GOWB, 7 March 2003. The x-axis represents time in seconds from 0 - 3600. The y-axis represents the amplitude of the waveform.

Example 6: An hour-long seismogram beginning at *04:00:00 March 24, 2003 at station TXB; a pattern of spikes/earthquakes by visual classification but classified as a pattern of background noise by the automatic process (Figure 6.6).*

This example is one of the 0.4% of spike/earthquake hours, which falls into the background noise category with an SIR of 1.47 and an MAV of 1.55. When the waveform is visually examined, the trace has a clear spike at 700 s. However the rest of the trace is dominated by background noise. The single spike at 700 s does not influence the SIR value enough for the spike pattern to be recognised, and the trace appears most closely to a pattern of background noise if the single spike is ignored.

From the examples presented in this section, it can be seen that stand-alone 'spikes' can sometimes cause the waveforms to be wrongly classified. After some careful examination of the original waveforms, I found that many of these spikes could be associated with artefacts produced during data processing. For example, if there is a gap in the data acquisition (maybe only for a very brief period), then the gap will be filled with zero values, effectively putting a square waveform into the trace. If this square waveform is undetected and processed with a high-pass filter, then the edges of the square waveform will generate artificial spikes. These will also be detected very strongly by the SI, as they correspond to a dramatic change in the waveform pattern.

To reduce the number of the ambiguous events in the classification, a procedure to remove artificial spikes in the waveforms is needed. To do this, the zero values (square waveforms) must be identified first and then replaced with interpolated values. This would still give a change of waveform pattern for the SI to detect and give a spike in SI

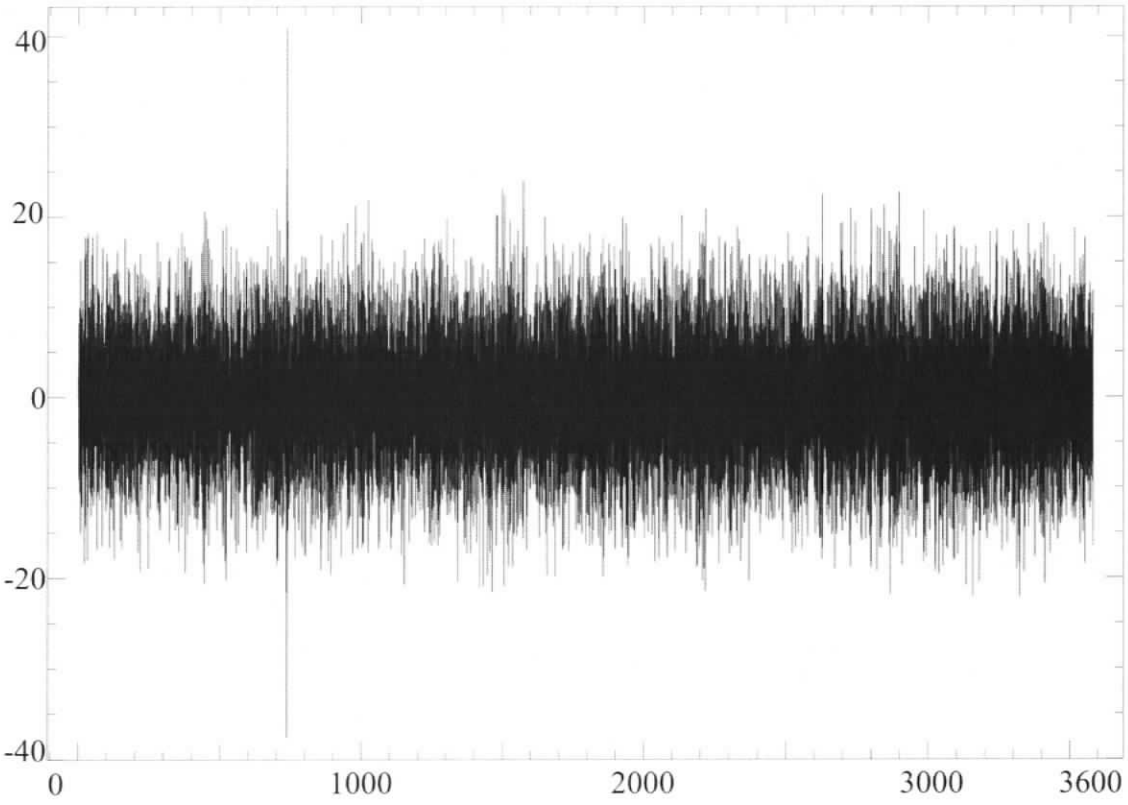

Figure 6.6
High pass filtered waveform of the hour beginning at 04:00:00 from station TXB, 24 March 2003. The x-axis represents time in seconds from 0 - 3600. The y-axis represents the amplitude of the waveform.

values. However if after interpolation a segment of the waveform from just before the gap is used to replace the gap, then the pattern change can be greatly reduced and the artificial spikes would not be created. Unfortunately due to time constraints, this procedure could not be implemented in this study, but should be included in future studies.

6.2 Final Remarks

Current studies of Cascadia tremors have all relied on visual classification. This study is the first to automate the identification of tremors. There are discrepancies between the two approaches when compared (as presented in the previous section).

The question then arises as to which method is more robust, reliable, accurate, and consistent. Visual classification by one person could be said to be more accurate, but it is also very time consuming, labour intensive, and certainly much less efficient. The second problem of visual classification is the subjective difference of opinions. Two people classifying the same event may come to different conclusions, especially when looking at the ambiguous examples previously presented. Some hours could have fallen in any of the three categories depending on the observer's point of view. In other words, classification itself inherits a source of uncertainty from human judgements.

The automated classification that I have developed can be used through both tremor and non-tremor sequences. It can classify accurately (over 92% for tremor patterns), efficiently (requiring much less human labour), and is more reliant when classifying the ambiguous hours. It also eradicates the possibility of different personal opinions.

The previous method of visual identification [Rogers and Dragert, 2003] looks for coherent tremors within the same hour across multiple stations. If coherent tremor signals are observed at three or more nearby stations, it is considered a tremor hour. The automated procedure presented in this study, on the other hand, looks at the hour-long waveform of individual stations and determines its classification based on its overall waveform characteristics. So technically, the two approaches are conceptually different.

To do a truly meaningful comparison to the visual inspection results of previous studies, the waveform of one station must be examined against the waveforms at the neighbouring stations. For example, our automated classification may state that there is a tremor pattern within the hour long-segment. It does not note the time of the tremor event. So two neighbouring stations may both exhibit tremor patterns, but one may have the tremor at the beginning of the hour segment and one may have it at the end of the hour segment. Obviously, the two stations should not be considered as having coherent tremor signals. To reliably distinguish if the classified tremor patterns are coherent, the procedure would need another step to identify the timing of tremor and then compare with other stations.

Nevertheless, I tried to mimic the visual tremor hour counts by examining the automatic classification results of nearby stations. In this case, a tremor hour is counted if three or

more nearby stations are all classified as having tremor patterns. Figure 6.7 shows the daily tremor counts by both approaches. Although the exact numbers are slightly different, the overall trends are very similar. Therefore it is reasonable to conclude that the automatic process can yield classification results consistent to visual inspections.

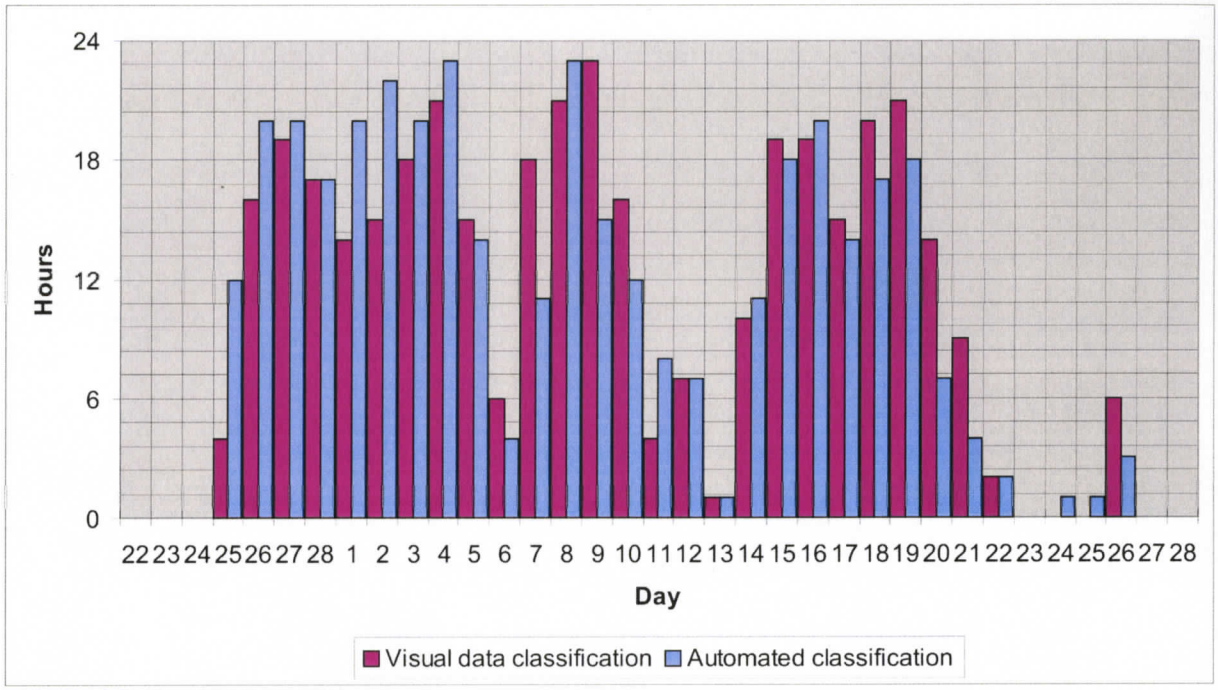

Figure 6.7
A graph of coherent tremors. The purple graph is created by a previous visual study of the 2003 event. Tremors are considered coherent if they appear on the seismograms of at least 3 stations in close proximity to each other geographically. The automated classification (shown in mauve) took the hourly classifications for each station and then searched for other stations within 100 km. If tremor was identified on more than 3 stations within this radius, and hour of coherent tremor was recorded.

CHAPTER 7

CONCLUSIONS

In this thesis, the pattern of ETS tremors in a waveform has been identified using calculations of moving average and scintillation index. A fully automated procedure for identifying tremor patterns in seismic waveforms is presented for the Cascadia subduction zone. The merits over previous visual studies have been summarized and the accuracy of the automated procedure has been quantified.

Using the SIR and MAV as defining values, over 92% of tremor patterns can be accurately identified for the February-March 2003 tremor sequence. It has been shown that some of the remaining 8% of tremors, which are incorrectly classified, are somewhat ambiguous. It is possible that if two people were to visually classify the waveforms of this particular group, different results may arise subject to individual's bias. As far as the majority of clear tremors are concerned, our automated identification algorithm is very reliable, efficient and non-subjective.

This comprehensive investigation of ETS and creation of an automated procedure to identify tremors with 92% accuracy can hopefully take ETS studies to a new level. Now greater volumes of data can be analyzed in a faster, less subjective way. It may also be applied to other subduction zones, (maybe with a slight change to the defining values of SIR and MAV) to further the search for the ETS phenomenon. To discover the values which best work at other subduction zones, a comparison similar to that presented in

Figure 5.5 would have to be done, with at least a few hundred samples. The exact threshold values of SIR and MAV may need to change depending on the level of background noise and network instrument configuration. In any case, the percentage of correct identification should be kept at >90% to be deemed successful.

Future uses of this automated procedure would be to create a real-time identification procedure to identify the next tremor sequence, and constantly monitor seismicity as it can be shown to also identify earthquakes, and background noise.

REFERENCES

Adams, J., 1990, Paleoseismicity of the Cascadia Subduction zone: Evidence from Turbidites off the Oregon-Washington Margin, *Tectonics*, 9, 569-583.

Atwater, T., 1970, Implications of Plate Tectonics for the Cenozoic Tectonic Evolution of Western North America, *Geological Society of America Bulletin*, 81, 3513-3536.

Atwater, B., Nelson, A., Clague, J., Carver, G., Yamaguchi, D., Bobrowsky, T., Bourgeois, J., Darienzo, M., Grant, W., Hemphill-Haley, E., Kelsey, H., Jacoby, G., Nishenko, S., Palmer, S., Peterson, C., and Reinhart, M., 1990, Summary of Coastal Geological Evidence for Past Great Earthquakes at the Cascadia Subduction Zone, *Earthquake Spectra*, 11, 1-18.

Calvert, A.J., 1996, Seismic reflection constraints on imbrication and underplating of the northern Cascadia convergent margin, *Canadian Journal of Earth Sciences*, 33, 1294-1307.

Crosson, R. and Owens, T., 1987, Slab Geometry of the Cascadia Subduction Zone Beneath Washington from Earthquake Hypocenters and Teleseismic Converted Waves, *Geophysical Research Letters*, 14, 824-827.

Davis, E., and Currie, R., 1993, Geophysical Observations of the Northern Juan de Fuca ridge system: lessons in sea-floor spreading, *Canadian Journal of Earth Sciences*, 30, 278-300.

Dragert, H., Hyndman, R., Rogers, G., Wang, K., 1994, Current deformation and the width of the seismogenic zone of the northern Cascadia subduction zone, *Journal of Geophysical Research*, 99, 653-668.

Dragert, H., Wang, K., and James, T.S., 2001, a silent slip event on the deeper Cascadia subduction interface, *Science*, 292, 1525-1528.

Dragert, H., Wang K., and G. Rogers, 2004, Geodetic and seismic signatures of episodic tremor and slip in the northern Cascadia subduction zone, *Earth Planets Space*, 56, 1143-1150.

Duncan, R.A., 1982, A captured island chain in the Coast Range of Oregon and Washington, *Journal of Geophysical Research*, 84, 10827-10837.

Fluck, P., Hyndman, R.D., and Wang, K., 1997, Three-dimensional dislocation model for great earthquakes of the Cascadia subduction zone, *Journal of Geophysical Research*, 102, B9, 20539-20550.

Goldfinger, C., Nelson, C.H., and Johnson, J.E., 2003, Holocene earthquake records from the Cascadia subduction zone and northern San Andreas fault based on precise dating of offshore turbidites, *Annual Review of Earth and Planetary Sciences*, 31, 555-577.

Greene, A.R., Scoates, J.S., and Weis, D., 2005, Wrangellia Terrane on Vancouver Island: distribution of flood basalts with implications for potential Ni-Cu-PGE mineralization in southwestern British Columbia, *B.C. Geological Fieldwork 2004*, Paper 2005-1, 209-220.

Hyndman, R.D., 1995, Giant Earthquakes of the Pacific Northwest, *Scientific American*, 273, 50-57.

Hyndman, R.D., Rogers, G.C., Dragert, H., Wang, K., Clague, J.J., Adams, J., and Bobrowsky, R.T., 1996, Giant earthquakes beneath Canada's west coast, *Geoscience Canada,* 23, 63-76.

Johnston, S., and Acton, S., 2003, The Eocene Southern Vancouver Island Orocline – a response to seamount accretion and the cause of fold-and-thrust belt and extensional basin formation, *Tectonophysics*, 365, 165-183.

Kao, H., Shan, S-J., Dragert, H., Rogers, G., Cassidy, J., and Ramachandran, K., 2005, A wide depth distribution of seismic tremors along the northern Cascadia margin, *Nature*, 436, 841-844.

Kao, H., S. Shan, H. Dragert, G. Rogers, J. F. Cassidy, K. Wang, T. S. James, and K. Ramachandran (2006), Spatial-temporal patterns of seismic tremors in northern Cascadia, *Journal of Geophysical Research*, 111, B03309, doi:10.1029/2005JB003727.

Massey, N.W.D., 1986, Metchosin Igneous Complex, southern Vancouver Island: Ophiolite stratigraphy developed in an emergent island setting, Geology, 14, 602-605.

Mazzotti, S., and Adams J., 2004, Variability of near-term probability for the next great earthquake on the Cascadia subduction zone, *Bulletin of the Seismological Society of America, 94,* 1954-1959.

Obara, K., 2002, Nonvolcanic Deep Tremor Associated with Subduction in Southwest Japan, *Science*, 296, 1679-1681.

Riddihough, R., 1977, A model of recent plate interactions off Canada's west coast, *Canadian Journal of Earth Science*, 14, 384-396.

Riddihough, R., 1984, Recent movements of the Juan de Fuca Plate system, *Journal of Geophysical Research*, 89, 6890-6994.

Rogers, G., and Dragert, H., 2003, Episodic Tremor and Slip on the Cascadia Subduction Zone: The Chatter of Silent Slip, *Science*, 300, 1942-1943.

Satake, K., Shimazaki, K., Tsuji, T., and Ueda, K., 1996, Time and size of a giant earthquake in Cascadia inferred from Japanese tsunami records of January 1700, *Nature*, 379, 246-249.

Vine, F.J., 1966, Spreading of the ocean floor – New evidence, *Science*, 154, 1405-1415.

Wells, R.E., Engebreston, D.C., Snavely, P.D., and Coe, R.S., 1984, Cenozoic plate motions and volcano-tectonic evolution of western Oregon and Washington, *Tectonics*, 3, 275-294.

Yeh, K., and Liu C-H., 1982, Radio Wave Scintillations in the Ionosphere, *Proceedings of the IEEE*, 70, 324-360.