Sensor Placement for Microseismic Event Location

A Thesis Submitted to the

College of Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the degree of Master of Science

in the Department of Electrical Engineering

University of Saskatchewan

Saskatoon

By

Angus Errington

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Electrical Engineering

57 Campus Drive

University of Saskatchewan

Saskatoon, Saskatchewan

Canada

S7N 5A9

Abstract

Mining operations can produce highly localized, low intensity earthquakes that are referred to as microseismic events. Monitoring of microseismic events is useful in predicting and comprehending hazards, and in evaluating the overall performance of a mine design.

A robust localization algorithm is used to estimate the source position of the microseismic event by selecting the hypothesized source location that maximizes an energy function generated from the sum of the time–aligned sensor signals. The accuracy of localization for the algorithm characterized by the variance depends in part upon the configuration of sensors. Two algorithms, MAXSRC and MINMAX, are presented that use the variance of localization error, in a particular direction, as a performance measure for a given sensor configuration.

The variance of localization error depends, in part, upon the energy spectral density of the microseismic event. The energy spectral density characterization of sensor signals received in two potash mines are presented and compared using two spectral estimation techniques: multitaper estimation and combined time and lag weighting. It is shown that the difference between the the two estimation techniques is negligible. However, the differences between the two mine characterizations, though not large, is significant. An example uses the characterized energy spectral densities to determine the variance of error for a single step localization algorithm.

The MAXSRC and MINMAX algorithms are explained. The MAXSRC sensor placement algorithm places a sensor as close as possible to the source position with the maximum variance. The MINMAX sensor placement algorithm minimizes the variance of the source position with the maximum variance after the sensor has been placed. The MAXSRC algorithm is simple and can be solved using an exhaustive search while the MINMAX algorithm uses a genetic algorithm to find a solution. These algorithms are then used in three examples, two of which are simple and synthetic. The other example is from Lanigan Potash Mine. The results show that both sensor placement algorithms produce similar results, with the MINMAX algorithm consistently doing better. The

MAXSRC algorithm places a single sensor approximately 100 times faster than the MINMAX algorithm. The example shows that the MAXSRC algorithm has the potential to be an efficient and intuitively simple sensor placement algorithm for mine microseismic event monitoring. The MINMAX algorithm provides, at an increase in computational time, a more robust placement criterion which can be solved adequately using a genetic algorithm.

Acknowledgements

I would like to express my sincere gratitude and appreciation to my supervisor, Dr. Brian Daku for his guidance and encouragement throughout the course of this work. I would also like to thank Prof. David Dodds for his encouragement and financial assistance.

This work would not have been completed without the encouragement and support from Siu-ching. I would also like to thank my father, John Errington, for reading the rough draft of this thesis.

Contents

List of Tables

List of Figures

List of Abbreviations

CTLW	Combined Time and Lag Weighting
dB	deciBels
ESD	Energy Spectral Density
FFT	Fast Fourier Transform
GA	Genetic Algorithm
PSD	Power Spectral Density
WOSA	Weighted Overlapped Segment Averaging

Chapter 1

INTRODUCTION

An increase in seismic activity surrounding open pit and deep underground mines has been observed at many sites. This kind of seismic activity is often called *induced seismicity* to indicate the triggering nature of mining activities in the release of preexisting stresses and stresses caused by mining [1]. Knowledge of this induced seismic activity is useful for evaluating mine design, mine planning and comprehending and predicting hazards. Also, as noted by Gibowicz and Kijko in [1]:

> Seismicity associated with underground mining is probably the most adverse phenomenon, among different types of triggered earthquakes, in relation to the safety and productivity of mining. Rockbursts are very often the major cause of fatalities in mines[1]. The problem becomes progressively more severe as the average depth and the extent of mining operations increase.

Thus, monitoring of seismic activity in the mine area is a very important topic. For an excellent overview of mining induced seismicity in Canadian mines please see [2].

Seismic events, caused by mining, can be separated into three broad categories [3]

1. acoustic emissions
2. microseismicity
3. large–scale seismicity (earthquakes)

Acoustic emissions are caused by microscopic failure in rock and are often precursors to larger failures. Quite often monitoring of these emissions is the basis of roof–fall warning systems [3]. In potash mines, for example, a crackling sound is often heard, in freshly mined rooms, as salt crystals adjust to the change in stress [2]. Large–scale seismicity, or earthquakes, involve rock failure over

[1] There is a distinction made in mining seismology between rockbursts and mine tremors. Rockbursts are violent failure of rock which cause actual visible damage to underground voids and mine workings within a mine while mine tremors are seismic events that do not cause damage [1].

hundreds of meters or more [3]. The most prevalent and monitored class of seismic activity in mines is microseismicity. The following section discusses microseismicity in detail, since this is by far the most common form of induced seismicity and the greatest indicator of stress release in mines.

1.1 Microseismicity

Microseismicity involves failure over a few tens of meters and is often part of normal mine room closure in plastic rock mines, such as coal and potash. Room closure often involves roof–delamination and/or floor heave. Microseismic events produce mechanical waves which propagate outward from the source position. There are P–waves and S–waves; the P, or primary, waves are compressional while the S, or secondary, are shear waves [4]. The P–waves travel faster than S–waves, as can be seen in the following equations,

$$V_p = \left(\frac{\lambda + 2\mu}{\rho}\right)^{1/2} \tag{1.1}$$

and

$$V_s = \left(\frac{\mu}{\rho}\right)^{1/2}, \tag{1.2}$$

where V_p and V_s are the velocity of the P and S waves, respectively [4]. In equations (1.1) and (1.2), ρ is the density of the propagation medium and λ and μ are the two Lamè constants. The Lamè constants are both positive and depend on the medium in question [1]. Typical values for the constants in potash are: $\rho = 2100$ kg/m^3, $\mu = 13$ GPa, and $\lambda = \mu$. This gives an S wave velocity of about 2500 m/s and a P wave velocity of about 4300 m/s. P waves are rarely observed on the monitoring systems used in Saskatchewan potash mines.

Quantitatively seismic events can be classified relative to magnitude as seen in Table 1.1 [4]. The convention used throughout this thesis is to refer to microseismic events as those events with magnitude less than 3 and refer to earthquakes as those events with magnitude greater than 3 [4]. Also, the terms microseismic event and microearthquake are used interchangeably.

Magnitude (M)	Classification
$M \geq 7$	Major Earthquake
$5 \leq M < 7$	Moderate Earthquake
$3 \leq M < 5$	Small Earthquake
$1 \leq M < 3$	Microearthquake
$M < 1$	Ultra–microearthquake

Table 1.1: Earthquake classification based on Richter Magnitude scale

1.1.1 Microseismic Monitoring

The most common type of seismic monitoring system deployed in mines are microseismic monitoring systems. These systems are designed to monitor microseismic events in the vicinity of a mine. Though the main target group are microearthquakes these systems are capable of recording large–scale seismic events as well.

The purpose of monitoring is to obtain three types of information regarding seismic events in mines:

1. **When** rock is breaking,
2. **Where** rock is breaking,
3. **How much** rock is breaking (the size of the failure).

This type of information allows one to, "accurately quantify the geomechanical response of the 'host rock' to the opening of rooms during mining [3]." Here the word 'rooms' refers to tunnels in underground mines. Seismic monitoring of the geomechanical response to mining is useful in evaluating the day–to–day performance of a mine design [3].

Even though rockbursts are a serious issue relating to mine safety, there is currently no system which can reliably predict when and where they will occur. Seismic monitoring systems, however, aim to contribute a portion of knowledge which hopefully can be analyzed in the context of other

similar observations [3].

A monitoring system consists of a set of sensors distributed throughout the mine. This set of sensors is often called a sensor array or sensor network. The sensors used are vibration sensors which measure ground motion in the form of either velocity or acceleration. Velocity sensors are called geophones and acceleration sensors are called accelerometers. Although both types of sensor are available, geophones are much more prevalent in mine monitoring systems. An example of a set of geophone outputs for a microseismic event is given in Figure 1.1. The traces shown in Figure

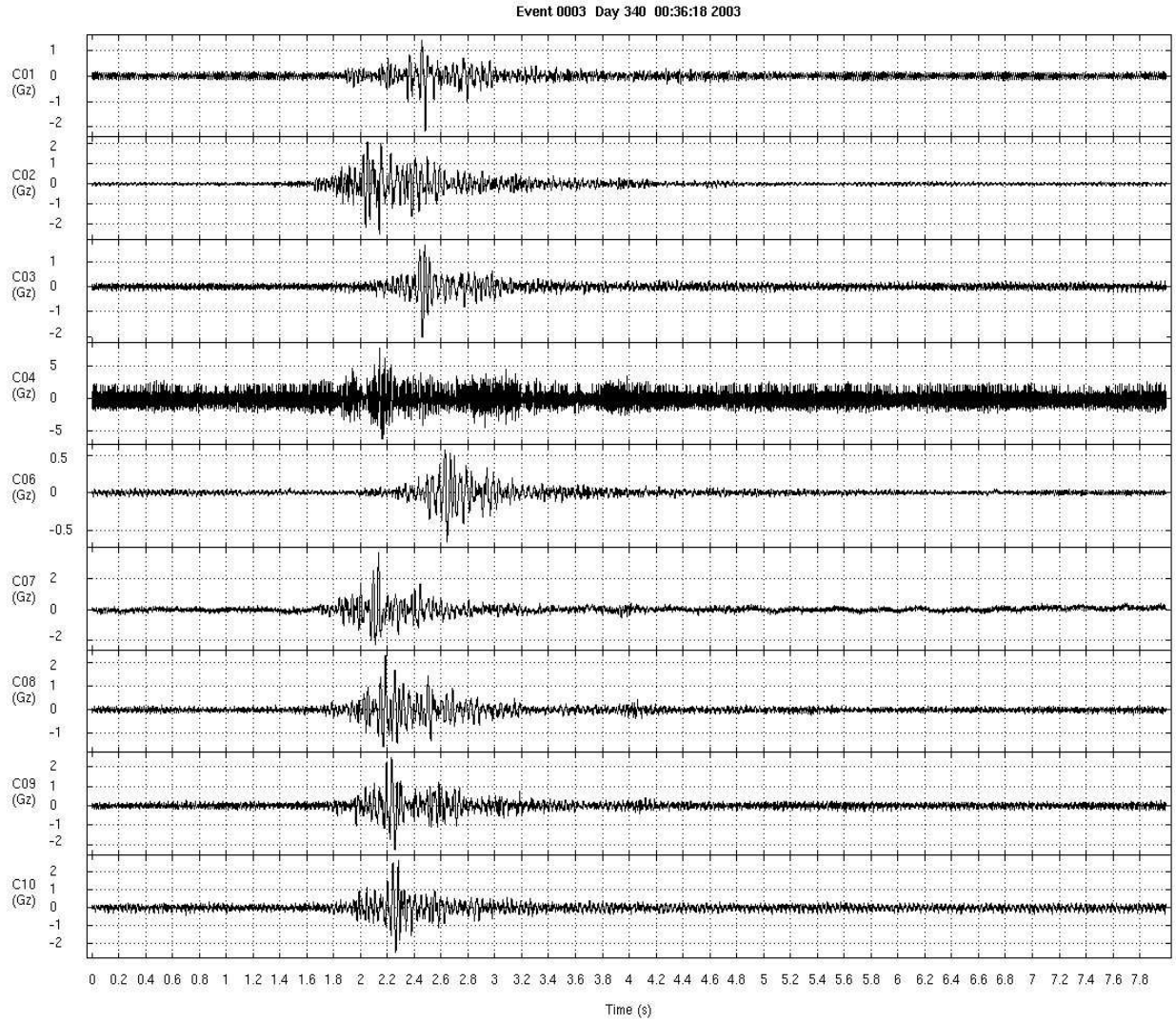

Figure 1.1: An example of a seismic trace recorded in December 2003 at Lanigan Potash Mine near Saskatoon, SK

1.1 can be used to estimate the size of the event as well as the location of the event. The time

the event took place is inherent in the monitoring system since it records the time the traces were captured. The accuracy of the location estimated from the received sensor signals is very important and can depend on a number of factors. Of concern in this thesis, are the sensor positions relative to an event's location, since this has a considerable effect upon the estimate's accuracy. This topic is introduced briefly below.

1.1.2 Locating Seismic Events

This subsection gives a brief overview of the problem of locating a seismic event. A detailed discussion of the various methods used in seismic event localization is given in Chapter 2. Locating a source using a sensor network relies on the fact that the signals received at sensors at different distances from the source will arrive at different times. That is, sensors closer to the source will receive the signals before those farther away. The first step in locating a given seismic event, assuming the received sensor signals as seen in Figure 1.1 are known, is to choose a velocity model for wave propagation. The simplest model is that of uniform earth with either a known or unknown velocity. Alternatively a layered velocity structure is often used to account for the geological stratification of the earth [1, 4]. Using the chosen model an event location can be selected and the model can be used to predict what the received signals should look like; specifically how each signal is shifted relative to the other signals. The estimated signals, or some parameters of them, can then be compared with the observed signals. In this way one can minimize the difference between the estimated and observed parameters to estimate the location of an event. Though the approach outlined above is used in most of the seismic and microseismic event location procedures, it is not the only approach. Different approaches do exist, but they will not be discussed here.

1.1.3 CANSEIS

The CANSEIS microseismic monitoring system was developed in–house by the Potash Corporation of Saskatchewan. It is deployed at a number of their potash mines, which are the focus of this thesis. The system can be used either in manual or automatic mode to calculate the three pieces of

information useful for microseismic monitoring: size of the event, time of the event and location of the event. In manual mode the 'picking' of the arrival times can be done by eye and then refined by the software. A 'misfit' function determines the differences between the observed and calculated arrival times, based on the model. If this value is large, then the arrival times can be 're–picked.'

The CANSEIS system has been in use since 1997 and has proved invaluable to mine engineers and planners. Currently, there is no procedure used to place sensors. Typically, they are placed every 600 meters along the main mine rooms. The following section introduces the topic of algorithmic sensor placement to improve the accuracy of the location estimate.

1.1.4 Sensor Placement

The accuracy of a location estimate depends upon a number of factors including: the precision of the signal recording equipment, the accuracy of the assumed velocity model, the suitability of the location estimation technique and the spatial distribution of the seismic sensors. The focus of this thesis is to improve the accuracy of the event location estimate by adding additional sensors to the sensor network to modify the spatial distribution of seismic stations. Related to this is the location estimator, itself. To improve the accuracy of an estimate one needs a measure of how good one spatial configuration of sensors is compared to another. A statistic of the estimator itself implicitly depends upon the positions of sensors and is therefore an excellent measure of how well a given spatial configuration performs.

The statistic and estimator chosen in this thesis is the variance of error, in a particular direction, of a one–step estimator. The reason for these choices will be explained in the following chapter. It is not necessary to understand either of these to see how the variance of error can be used to evaluate the performance of a sensor configuration. If two sensor configurations are used to evaluate the variance of error, it should be clear that the configuration that produces the smaller variance of error would be superior. Typically, sensor configurations that surround an area to be monitored produce better results than those that are clustered to one side.

The mines of interest in this thesis are potash mines in Saskatchewan and they are constantly

expanding. They are, collectively, growing at a rate of approximately 500 meters per hour. It would not make sense to reconfigure the entire sensor network each time an expansion is made. This would be both costly and inefficient. To adequately monitor new areas of mines, termed mine expansion regions, an additional sensor should be added to specifically monitor the new region. This allows the current network to remain undisturbed as well as simplifying the algorithm used to place an additional sensor.

1.2 Problem Statement

To develop a simple effective algorithm to place sensors in mine expansion regions to improve the precision of the location estimator.

1.3 Aims and Objectives

The objective of this research is to provide a simple, practical and effective method of placing additional sensors in mine expansion regions. The placement of these sensors will improve the precision of the location estimator and also be superior to other possible sensor positions. There are many possible criteria that could be used to decide which sensor position is superior to another, and two of these are presented.

The statistic of a location estimator is dependent upon the specific estimator chosen. It is clear, then, that an estimator should be chosen with a known formula for calculating one of its statistics. A single–step estimator is chosen due to its quality of robustness to short–duration events. The estimator's localization variance in a given direction has been derived, previously, and is used as the statistic.

The localization variance expression depends in part upon the seismic event's energy spectral density. To accurately estimate the variance it is necessary to have an accurate characterization of the event's energy spectral density. A characterization method using existing mine data captured by the seismic monitoring system is designed.

Though this placement algorithm can be used in any mine, the specific application of interest here is potash mines.

1.4 Organization of Thesis

The organization of this thesis is as follows. Chapter 2 provides an overview and introduction to location estimators. Estimators directly applicable to microseismic event location are highlighted. This chapter aims to provide an overview of estimators as well as justifying the use of this new type of estimator in microseismic event location. Though location estimation techniques are not directly related to sensor placement, it is necessary to have an understanding of them, since the error of a location technique is used as a measure of performance for a sensor configuration. This chapter should also clarify why chapter 3 is included. Neither Chapter 2 or Chapter 3 is necessary to understand the main focus of this thesis, presented in chapters 4 and 5. However, they both provide background information relevant to the choice and evaluation of the measure used in the sensor placement algorithms.

Chapter 3, which is fairly self–contained, describes the procedure used to characterize the energy spectral density for a given mine. It also presents the results of two mine characterizations using two different spectral estimation techniques. These spectral estimation techniques are described at the beginning of the chapter. At the end of the chapter an example is presented which uses the results to evaluate the variance expression for the chosen location estimator. This provides a practical comparison between the two mines and the two spectral estimation techniques.

Chapter 4 presents the the two sensor placement algorithms. Initally, at the beginning of the chapter, an overview of previous work in seismic sensor positioning is presented. A justification of the presented algorithms is also included to differentiate between those in the literature and those presented in the chapter. At the end of the chapter some additional modifications to the algorithms are suggested.

Chapter 5 presents three examples. The first two are synthetic while the third is taken from Lanigan Potash Mine. In each example one and two sensors are placed using the algorithms with

and without the modifications.

Finally, Chapter 6 summarizes the thesis, presents the conclusions and outlines some possible avenues for future work.

Chapter 2

Location Estimation

Estimating the location of a passive acoustic source is an important topic in a wide range of fields. Recently, much attention has been paid to acoustic source localization using microphone arrays for camera pointing in video conferencing [5]. Underwater acoustics has also been a very active research area in the past [6] and of specific relevance to this thesis, seismic and microseismic earthquake location [7].

Specifically, the problem here can be termed as passively locating a wideband vibrational source in the near–field region; wideband means a source whose ratio of highest to lowest frequencies is much greater than unity. The near–field region is the region about a source where the wavefronts appear curved unlike the far–field region where the wavefronts appear planar and parallel (figures 2.1 and 2.2). As the names indicate the far field region is much further away than the near field region. It should be clearly stated that figures 2.1 and 2.2 are not to scale. Even though the distance between source and sensors appear the same, this is not the case. For a wavefront to appear planar to a group of sensors the source must be a long way away or the inter–sensor distance must be small compared to the distance to the source. If the sensors are in the far field region then it is not possible to obtain both the distance and direction to the source, such as in the near field region, but only the direction to the source.

Although this location problem is present in many other areas of study, this chapter will focus on the application of interest: seismic and microseismic event analysis. At a risk of generalizing, estimation algorithms can be separated into two broad groups: two–step estimation and one–step estimation. Strictly speaking these two groups may not span the complete set of location algorithms, although for the purpose of explanation in this thesis they will be sufficient.

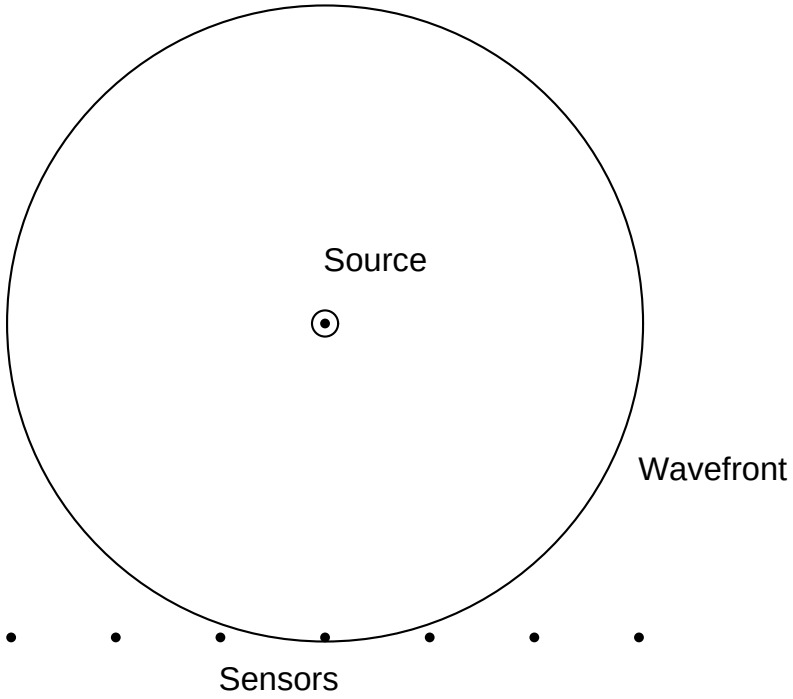

Figure 2.1: Sources in the near field produce curved wavefronts

The aim of this chapter is to introduce the location estimation problem and two types of location estimators. It may, at first, not be clear how this relates to the placement of sensors. However, the choice of the location estimator determines the performance measure of a sensor configuration. As will be seen in this chapter a natural choice for the two–step estimator is the determinant of the covariance matrix while for the one–step estimator the variance of error in a particular direction is more appropriate. This chapter also attempts to explain the differences between the two techniques so that the use of the one–step estimator is justified. Finally, the choice of the one–step estimator necessitates characterizing the Energy Spectral Density (ESD) of microseismic events, which is performed in Chapter 3. It is necessary because the variance expression, which will be used as a performance measure of a sensor configuration, depends upon the ESD of microseismic events.

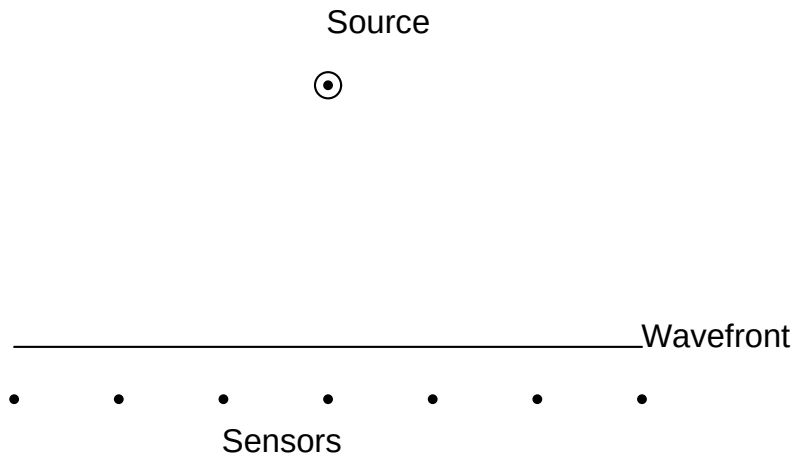

Figure 2.2: Sources in the far field produce planar wavefronts

2.1 Two–Step Estimators

The two–step estimator, as the name implies, uses two steps to estimate the location for an acoustic source. The first step estimates the time of arrivals (TOAs) at each sensor. The second step locates a source position and origin time, which would produce the TOAs estimated in step one.

2.1.1 Time of Arrival Estimation

In the early days of seismology arrival times were picked by eye. This practice has continued to the present day where many systems let a person make initial arrival time 'picks' and then refine them using cross–correlation techniques. It has only been in the last 20 years that there has been a concerted effort to develop automatic arrival time 'pickers,' mainly for convenience rather than to improve the accuracy. When these new techniques are presented they are often compared to human 'pickers' to assess their accuracy. An example of one of these newer techniques is presented in [8]. The technique used in the potash mines in Saskatchewan use a technique similar to that presented in [9]. The details of arrival time refinement, or picking, will not be discussed here and it will be assumed for the next section that N arrival times, $\hat{\mathcal{T}}_i$, are available.

2.1.2 Localization

Once the arrival times are estimated the problem of locating an event can be stated as follows: Given a set of arrival times $\hat{\mathcal{T}}_i$ and an assumed propagation model, find a location which minimizes the error between the measured arrival times and those calculated using the assumed location.

The location estimation procedure outlined in this section follows that of seismic event location; for some examples of algorithms not related, specifically, to seismology see [10–13].

The arrival time values can be calculated as

$$\mathcal{T}_i = T_i(\theta, \mathbf{P}) + t_o, \tag{2.1}$$

where θ is the vector of unknown parameters (x_s, y_s, z_s, t_o), $T_i(\cdot)$ is the propagation time from the source to sensor i and t_o is the origin time. The set of sensors is $\mathbf{P} = [\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_N]$. The origin time is the absolute time at which the event takes place, *ie* the time of day. It is usually only included in the estimation in seismic location procedures. This is because, as mentioned in Chapter 1, the time an event takes place is an important piece of information for seismologists. It can be estimated prior to source localization by centering techniques [14]. For the following analysis, however, we assume the origin time still needs to be estimated.

The residuals can then be defined as

$$r_i = \hat{\mathcal{T}}_i - \mathcal{T}_i, \tag{2.2}$$

which is the difference between the calculated and observed arrival times. A common objective function is a sum of squares

$$\Theta = \sum_{i=1}^{N} r_i^2, \tag{2.3}$$

and the solution which minimizes this is known as the Least Squares solution. Since $\mathcal{T}_i$ is a non–linear function of the source location parameters, the optimization problem is non–linear.

One of the first methods used for earthquake location to minimize this objective function was that by Geiger [15], which uses an iterated Gauss–Newton method. It attempts to find an adjustment vector $\delta\theta$ which when added to a trial source position θ^* produces a new estimate of the source position, that decreases the objective function.

An excellent description of this algorithm is presented in [4] and it will be followed closely here. First, write the objective function in matrix notation

$$\Theta(\theta^*) = \mathbf{r}^T\mathbf{r}, \tag{2.4}$$

where $\mathbf{r}$ is a column vector containing the N residuals and T is the transpose operator. The adjustment vector can be found to be

$$\delta\theta = -(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{r}, \tag{2.5}$$

where $\mathbf{A}$ is the Jacobian evaluated at the trial source location θ^*.

$$\mathbf{A} = -\begin{bmatrix} 1 & \partial T_1/\partial x_s & \partial T_1/\partial y_s & \partial T_1/\partial z_s \\ 1 & \partial T_2/\partial x_s & \partial T_2/\partial y_s & \partial T_2/\partial z_s \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \partial T_N/\partial x_s & \partial T_N/\partial y_s & \partial T_N/\partial z_s \end{bmatrix} \tag{2.6}$$

Solving for $\delta\theta$ in equation (2.5) gives an adjustment vector to the initial trial location θ^* which leads to a new trial solution. The new trial location θ' can be found by setting it equal to $\delta\theta + \theta^*$. This procedure can be repeated with $\theta^* = \theta'$ in an iterative fashion until some termination criteria have been met. The iterative procedure is necessary due to the nonlinearity; specifically, the propagation time equations are nonlinear with respect to hypocenter[1] coordinates. A simple modification to equation (2.5), which arises by dismissing the assumption that the reliability of all observed arrival times are the same, can be made. One possible estimate of the reliability is the standard deviation of the time residuals σ_{r_i}. Using these, and the assumptions that the time residuals are statistically independent at each station the covariance matrix is

$$\mathbf{C_r} = \mathrm{diag}(\sigma_{r_1}^2, \cdots, \sigma_{r_N}^2). \tag{2.7}$$

The inverse of this can then be used as the weighting matrix in weighted–least–squares form of equation (2.5),

$$\delta\theta = -(\mathbf{A}^T{\mathbf{C_r}}^{-1}\mathbf{A})^{-1}\mathbf{A}^T{\mathbf{C_r}}^{-1}\mathbf{r}. \tag{2.8}$$

[1] Hypocenter is a seismological term which refers the earthquake location specified by the three spatial coordinates

One of the drawbacks of the algorithms presented above is that the initial 'guess' of the source location has to be quite good. If it is a poor guess then the algorithm may converge to a local minimum in the objective function rather than the absolute minimum. To address this issue a nonlinear search algorithm, the Nelder–Mead Simplex procedure, has been applied to this problem [16, 17]. The method described in [16] also modifies equation 2.3 to be the L1 norm as opposed to the L2 norm,

$$\Theta = \sum_{i=1}^{N} |r_i|, \tag{2.9}$$

where $|\cdot|$ is the absolute value. The L1 norm tends to deemphasize effects of single large errors and performs better when the outlier values are large [18].

2.1.3 Estimate of the Error

The estimate of the error in the case of least–squares is given by the covariance matrix of the sought parameters θ [19]

$$\mathbf{C}_\theta(\mathbf{P}) = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{C_r}\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}. \tag{2.10}$$

The above equation simplifies to

$$\mathbf{C}_\theta(\mathbf{P}) = (\mathbf{A}^T{\mathbf{C_r}}^{-1}\mathbf{A})^{-1}. \tag{2.11}$$

Both of the above equations show how the error depends in part upon the sensor positions $\mathbf{P}$, because $\mathbf{A}$ depends on sensor location.

It will be seen in later chapters that when comparing sensor configurations, it is useful to compare a single number and not a group of numbers, *ie* a matrix. To this end the determinant of the covariance matrix is often used. The determinant of the covariance matrix is proportional to the volume in the four dimensional (4D) error ellipse. A sensor configuration which minimizes this determinant for a given source location is known as D–optimal. This will be discussed in greater detail in Chapter 4.

2.2 Single–Step Estimators

Single–step, or direct process, estimators forego the initial time delay estimation step. They typically have, as inputs, the recieved sensor signals. The algorithms are search based and maximize a function over a grid of hypothesized source locations [20]. Some recent examples of these algorithms are [21–24]. In this section, the Total Signal Energy estimate (TSE) [21], will be discussed and modified to include weighting. The Weighted Total Signal Energy (WTSE) estimate, described in the following subsection, is the location estimator that is used throughout the remainder of this thesis. Because of its central role in this document the underlying signal model will be explained as well as how the weightings are determined. In subsection 2.2.2 the variance of error in the direction of the unit vector $\vec{e}$ for the WTSE estimate is presented. Subsection 2.2.3 looks at the variance of error expression evaluated with sensor positions at or near the source location.

2.2.1 Weighted Total Signal Energy Estimate

This subsection describes the system model and the microseismic event localization algorithm. The material presented here is an extension of the research published in [20] where the localization algorithm is described and a lengthy derivation of the variance expression is presented. In this thesis the original work is extended to include signal weighting to improve the performance of the algorithm.

The spatial coordinate system used to develop the localization algorithm is a three dimension x, y, z Cartesian system with arbitrary orientation. In the analysis the origin of the coordinate system is defined to be the location of the microseismic event. Shifting the origin of the coordinate system to the location of the event does not affect the variance of error since the error only depends on the relative distances between the event location and the sensor locations.

The geometry of the sensor system is given in Figure 2.3. There are M sensors in the system with the location of sensor m denoted as $\vec{p}_m = (x_m, y_m, z_m)$. The location of the source of the microseismic event is $\vec{p}_s = (0, 0, 0)$ and the distance from the event to sensor m is denoted d_m. The

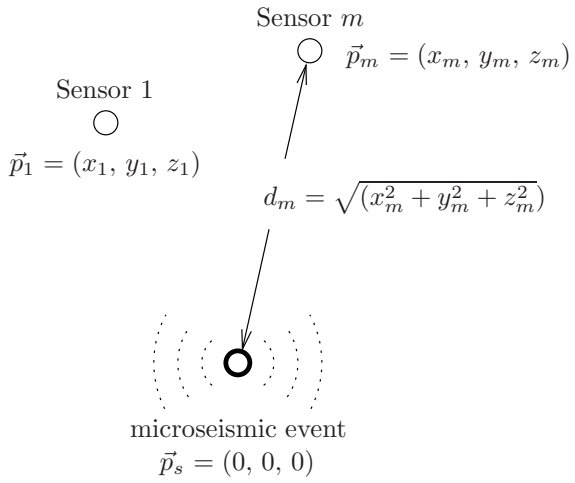

Figure 2.3: Coordinate system for the sensors

time origin is defined as the start time of the event.

The output of sensor m is modelled as a stochastic process,

$$\mathbf{r}_m(t) = \frac{s(t - \tau_m)}{d_m} + \mathbf{n}_m(t - \tau_m) \tag{2.12}$$

where

$$\tau_m = \frac{d_m}{c} \tag{2.13}$$

is the propagation time from the event location to sensor m and c is the speed of propagation of the mechanical wave. This model is based on the assumption of spherical spreading in a lossless, homogeneous medium.

The consequences of the model assumptions are briefly explored here. These three assumptions are rarely used in detailed geophysical analysis due to their poor correspondence to actual physical processes. For instance, spherical spreading is inconsistent with many failure mechanisms which produce radiation patterns of a dipole or even quadripole nature. This could cause some sensors to not receive any signal while another sensor, equal distance from the source, receives a clear signal. This would be difficult to account for in a simple model, but with sufficient sensors in the region of the event it would be more likely to attain a reasonable estimate of the location, even with the assumption of spherical spreading. The effects of the homogeneous medium assumption can be combatted by searching for an unknown velocity, as well as the spatial coordinates, in the algorithm.

The aim of this is to take account for the varying properties of the medium. The algorithm used in this thesis does assume a known velocity, but adding an unknown velocity to the algorithm is very simple, as explained below.

The localization algorithm is based upon a global search of a finite grid. An energy function is computed as a function of hypothesized source position for each point on the grid. The point which maximizes the energy function is the estimate of the position of the source. It is assumed the grid is very fine or that interpolation is used to find the location of the true peak. The energy function, which includes signal weighting factors, w_m, is

$$W(\vec{\tilde{p}}_s) = \int_0^T \left[\sum_{m=1}^{M} w_m r_m(t + \tilde{\tau}_m) \right]^2 dt, \tag{2.14}$$

where $\vec{\tilde{p}}_s$ is the hypothesized source location, $r_m(t)$ is a sample function from $\mathbf{r}_m(t)$, T is the length of the event $s(t)$, which in practice can be estimated from the strongest sensor signal at the closest sensor, and $\tilde{\tau}_m$, a function of $\vec{\tilde{p}}_s$, is the time required for the wave to propagate from the hypothesized source position to sensor m. Note, to simplify this analysis the velocity c used in the algorithm in expression (2.14) is assumed to be constant and known; however, the algorithm is easily extended to include searching over an additional dimension, which would be velocity.

The energy function, seen in equation (2.14), is maximized when the received signals are in alignment. In the abscence of noise this shift is consistent with the actual event location. In the presence of noise the global maximum will shift away from the actual event location. This degrades the performance of the algorithm.

The performance of the algorithm can be improved by defining suitable signal weights, w_m. If the true source location is known and is $\vec{p}_s = (0, 0, 0)$ then $\tilde{\tau}_m = \tau_m$ and the energy function in (2.14) can be expanded by substituting (2.12) to give

$$\mathbf{W}_s(\vec{p}_s) = \int_0^T \left[\sum_{m=1}^{M} \frac{w_m}{d_m} s(t) + \sum_{m=1}^{M} w_m \mathbf{n}_m(t) \right]^2 dt. \tag{2.15}$$

The noise term in (2.15) adversely affects the peak of $\mathbf{W}_s$. The effect of the noise can be reduced

by selecting the signal weights, w_m, to maximize the signal to noise ratio, γ,

$$\gamma = \frac{\displaystyle\int_0^T \left(\sum_{m=1}^{M} \frac{w_m}{d_m} s(t) \right)^2 dt}{E\left[\left(\displaystyle\sum_{m=1}^{M} w_m \mathbf{n}_m(t) \right)^2 \right]} \tag{2.16}$$

where $E(\cdot)$ is the expected value. Simplifying (2.16) gives

$$\gamma = \frac{\displaystyle\int_0^T s^2(t)dt}{\sigma^2} \frac{\displaystyle\sum_{l=1}^{M}\sum_{m=1}^{M} \frac{w_l w_m}{d_l d_m}}{\displaystyle\sum_{m=1}^{M} w_m^2} \tag{2.17}$$

where the denominator is simplified by noting that the noise term $\mathbf{n}_m(t)$ is an independent, zero mean, process and $E(\mathbf{n}_m^2(t)) = \sigma^2$ is the power in the noise. The term $\int_0^T s^2(t)dt$ is the signal energy.

The weights, w_m, are determined by solving the set of equations

$$\frac{\partial \gamma}{\partial w_m} = 0 \quad m = 1, \ldots, M \quad ,$$

but a constraint is required to limit the choices for w_m. Choosing the constraint as

$$\sum_{m=1}^{M} w_m^2 = 1$$

reduces the search to

$$\text{maximizing} \quad \sum_{i=1}^{M}\sum_{j=1}^{M} \frac{w_i w_j}{d_i d_j} \quad \text{subject to} \quad \sum_{m=1}^{M} w_m^2 = 1 \tag{2.18}$$

Here the method of Lagrange multipliers is used to solve (2.18). Define the function

$$F(w_1, \ldots, w_M, \lambda) = \sum_{i=1}^{M}\sum_{j=1}^{M} \frac{w_i w_j}{d_i d_j} + \lambda \left(\sum_{m=1}^{M} w_m^2 - 1 \right). \tag{2.19}$$

Taking the derivative of (2.19) with respect to the arguments gives

$$\frac{\partial F}{\partial w_m} = \sum_{i=1}^{M} \frac{2w_i}{d_m d_i} + \lambda w_m = 0 \quad m = 1, \ldots, M \tag{2.20}$$

$$\frac{\partial F}{\partial \lambda} = \sum_{m=1}^{M} w_m^2 - 1 = 0. \tag{2.21}$$

Solving (2.20) for λ gives

$$\lambda = \frac{-\frac{1}{d_m}\sum_{i=1}^{M}\frac{w_i}{d_i}}{w_m} \quad m = 1, \ldots, M \tag{2.22}$$

Equating the equations in (2.22) gives

$$\frac{1}{w_1 d_1} = \frac{1}{w_2 d_2} = \cdots = \frac{1}{w_M d_M}. \tag{2.23}$$

Substituting $w_j = \frac{w_1 d_1}{d_j}$ for $j = 2, \ldots, M$, which was derived from (2.23), in (2.21) gives

$$w_1^2 + \sum_{j=2}^{M}\left(\frac{w_1 d_1}{d_j}\right)^2 - 1 = 0. \tag{2.24}$$

Simplifying gives

$$w_1 = \frac{\frac{1}{d_1}}{\sqrt{\sum_{j=1}^{M}\frac{1}{d_j^2}}}.$$

Similarly, the other weights are determined, to give

$$w_m = \frac{\frac{1}{d_m}}{\sqrt{\sum_{j=1}^{M}\frac{1}{d_j^2}}}. \tag{2.25}$$

Note that the signal weights in (2.25) are inversely proportional to the distance between the source and the sensors. These intuitively simple weights maximize the signal to noise ratio, which is a common approach used in electronic communications. These weights, for high SNR's, should achieve the theoretical lower bound given by the Cramer–Rao Bound.

2.2.2 Estimate of the Error

An expression for the variance of location, in the direction of unit vector $\vec{e}$, for the unweighted ($w_m = 1, \forall m$) localization algorithm is developed in [20]. This derivation is lengthy, but it can easily be modified to include signal weighting to give

$$
\begin{aligned}
E[\hat{\epsilon}^2] = \quad & \left(\frac{4}{2\pi c^2} \int_{-\infty}^{\infty} \omega^2 S_{ss}(\omega) S_{nn}(\omega) d\omega \right. \\
& \times \left[\sum_{l=1}^{M} w_l^2 \left(\sum_{m=1}^{M} \frac{w_m}{d_m} q_m \right)^2 \right. \\
& -2 \sum_{k=1}^{M} \frac{w_k}{d_k} q_k \sum_{l=1}^{M} w_l^2 q_l \sum_{m=1}^{M} \frac{w_m}{d_m} \\
& \left. \left. + \left(\sum_{l=1}^{M} \frac{w_l}{d_l} \right)^2 \sum_{m=1}^{M} (w_m q_m)^2 \right] \right. \\
& + \frac{4T}{2\pi c^2} \left(\sum_{l=1}^{M} w_l^2 \sum_{m=1}^{M} (w_m q_m)^2 \right. \\
& \left. \left. - \left(\sum_{m=1}^{M} w_m^2 q_m \right)^2 \right) \int_{-\infty}^{\infty} \omega^2 S_{nn}^2(\omega) d\omega \right) \\
& \Bigg/ \left(- \frac{1}{2\pi c^2} \int_{-\infty}^{\infty} \omega^2 S_{ss}(\omega) d\omega \right. \\
& \left. + \sum_{l=1}^{M} \sum_{m=1}^{M} \frac{w_l w_m}{d_l d_m} (q_l - q_m)^2 \right)^2
\end{aligned}
\tag{2.26}
$$

where $\hat{\epsilon}$ is a random variable representation of the error in the global maximum of the stochastic process energy function in the direction of $\vec{e}$. The energy spectrum, $S_{ss}(\omega)$, is given by $S_{ss}(\omega) = |S(\omega)|^2$, where $|S(\omega)|$ is the magnitude of the Fourier transform of $s(t)$. $S_{nn}(\omega)$ is the power spectrum of the noise. $q_m = \frac{\vec{p}_m \cdot \vec{e}}{|\vec{p}_m|}$, where $\vec{p}_m \cdot \vec{e}$ is the dot product of the vector of sensor m's coordinates with the unit vector $\vec{e}$, and $|\vec{p}_m|$ is the distance from the origin to sensor m. Note that q_m is the cosine of the angle between $\vec{p}_m$ and $\vec{e}$.

Equation (2.26) is the basis of the sensor placement algorithms presented in chapters 4 and 5 of this thesis. That is, given a set of sensor positions, the speed of wave propagation, a direction in which to calculate the variance and the signal's and noise's energy spectral densities, a value for the variance of a source position is obtained. A set of sensor positions that produce a lower value of this variance would be deemed better at monitoring events located at the origin. Clearly, to monitor a region, instead of just one event location, some tradeoffs must be made. This is the subject of chapters 4 and 5. Because of the assumptions discussed in the previous section, this estimate of the error should not be taken as the absolute error. This expression is merely used to express the relative improvement of one configuration over another. This should be noted when observing the results of the examples in Chapter 5 and will be stated again in that chapter.

2.2.3 Sensor Positions Near Source

It may be unclear from the model presented in section 2.2.1, or that presented in [20], how to evaluate the variance of error expression when sensors are within one meter of the source location. Appendix A is devoted to this topic and this subsection summarizes the main results of that appendix. The appendix, as well as this section, are for mathematical consistency and are not intended to accurately model the signal a sensor would receive if it were close to a microseismic event. In fact, since an event takes place over an area of tens of meters, modelling of the signal within this area would be very difficult.

A large portion of Appendix A is devoted to showing that if the assumption that $|\tilde{p}_s| \ll |\vec{p}_m|$ is not made, the same variance expression will result. The appendix also clarifies the evaluation of the variance expression when a sensor is located at or within one meter of the source position. A portion of this clarification is presented here. To clarify the evaluation of the variance expression when a sensor is located at or within one meter of the source position it is necessary to separate the attenuation and the time delay of the received sensor signals. Implied by the model, though not stated explicitly, is that a signal received at sensors within one meter of the source location will have the same attenuation as that received at one meter; the time delay, however, will be consistent with the distance travelled.

Recognizing the above, how then should one evaluate the final expression for the variance? What terms are related to the attenuation and what terms are related to the time delay? The answer, fortunately, is quite simple. The d_m values relate to attenuation and the $\frac{\vec{p}_m \cdot \vec{e}}{|\vec{p}_m|}$ term relates to time delay. This amounts to setting $d_m = 1$ if the mth sensor is closer than one meter to the event. It must be stated clearly that the distance between the sensor and event remains the same but the attenuation is changed. It so happens that the attenuation equals $1/d_m$.

The final point to note is to set $q_m = 1$ for $|\vec{p}_m| = 0$. Once again, this is explained in Appendix A. The clarifications helpful when evaluating the variance expression are summarized in Table 2.1

Case	For	Action
1	$\lvert\vec{p}_m\rvert > 1$	No change
2	$0 < \lvert\vec{p}_m\rvert < 1$	set $d_m = 1$
3	$\lvert\vec{p}_m\rvert = 0$	set $d_m = 1$ and set $\frac{\vec{p}_m \cdot \vec{e}}{\lvert\vec{p}_m\rvert} = 1$

Table 2.1: Summary of clarifications

2.2.4 Advantages of One–step Estimation

The one–step estimator has one major advantage over two–step estimators. The TOA estimation step introduces errors into the arrival times. These errors are then propagated through a non–linear error function, which could result in large errors in the final source estimation. The one–step estimator foregoes this step and is therefore able to control errors. It could be for this reason that the single–step estimator seems to perform well for short duration events.

2.3 Summary

Two classes of localization algorithms were discussed, the two–step and one–step procedures. In addition, a specific estimator from each class was discussed in detail. For the two–step class of algorithms it was Geiger's method that solved the set of nonlinear equations using a nonlinear least squares technique. The error in this estimator was presented in the form of a covariance matrix of the unknown hypocentre location and origin time. In the case of the one–step, the WTSE estimator was presented. It's variance of localization error was presented which was seen to partly depend upon the signal's energy spectrum. The WTSE shows promise as a robust localization algorithm that performs well with finite–duration events that may be of short time duration. The WTSE estimator and its corresponding variance expression are used throughout the remainder of this thesis as the estimator of choice.

Chapter 3

Spectral Characterization

The signal's Energy Spectral Density (ESD) is required in order to evaluate the variance of error of a single step estimator, for a given sensor configuration. The ESD of a signal, $s(t)$, is defined to be

$$S_{ss}(\omega) = |S(\omega)|^2 = \left| \int_{-\infty}^{\infty} s(t) \mathrm{e}^{-j\omega t} dt \right|^2 . \tag{3.1}$$

Because real world signals are of finite length there are a number of ways to estimate the actual ESD of a signal. Two such methods are described in this chapter which is devoted to the characterization, or estimation, of microseismic event ESDs for two potash mines.

3.1 Method of Analysis

The data used were gathered, over a 12 month period, from C block Lanigan Potash Mine and over a 9 month period from Allan Potash mine, both of which are located near Saskatoon. The sensors used to gather the data are geophones. Geophones measure the ground motion relative to that of an inertial mass [18]. Because the operation of a geophone is based on a mass–spring–damper system its output is proportional to velocity. The geophones used are critically damped with a natural frequency of 4.5 Hz and have a flat response above this frequency. A low–pass 8–pole Bessel anti–aliasing filter, with cutoff frequency at 250 Hz, is used prior to digitizing the microseismic signal. The signals are sampled at 1000 times per second implying that the sampling frequency, $F_s = 1000$ Hz.

Each microseismic event that occurs is recorded by 8, 12 or 16 sensors, of which, usually, a few signals are usable. The other signals, which are not usable are too corrupted by noise. The

signals which are not usable, the location of the event and the event's energy are determined using the CANSEIS software [25]. The event energy varies from a fraction of a Joule to thousands of Joules. The majority of the events are less than 10 Joules in energy. The energy values calculated by CANSEIS determine the effective radiated energy of the event, not the energy released. The values of energy should therefore be used more of a relative measure of event strength, rather than an absolute measure.

The remainder of the processing discussed in this chapter is performed using MATLAB [26]. The usable signals are first ranked according to their respective distance from the event. Because noise, multipath effects, and signal attenuation are more pronounced at greater distances, the signals closest to the event are the most suitable for analysis. Therefore, only signals closer than 610 meters (2000 feet) were analyzed.

In the following subsection the classification algorithm used to separate the signals into two types is presented. It is important to classify the two types of signals so that it can be seen whether the time–domain differences correspond to differences in their ESDs. In sections 3.1.2 and 3.1.3 the multitaper and combined time and lag weighting methods, respectively, are outlined. Section 3.1.4 presents the results.

3.1.1 Classification Algorithm

After visually examining a large number of signals it was decided to separate the data into two classes of signals. The two classifications were, an impulse type and a successive event type. Figure 3.1 shows an example of each type of signal, where F_s is the sampling frequency and n is the sample index. The impulse type waveform has a sharp peak and decays very quickly. The successive event type waveform, however, is much longer in duration and does not have a well defined peak; the waveform may also have multiple peaks in succession as the event dies out. Comparing the upper and lower plots it can be seen that the energy in the impulse is much greater than for the successive event type signal. This energy discrepancy is often the case and is related to the type of signal. A MATLAB script was written to automatically perform the task of classification.

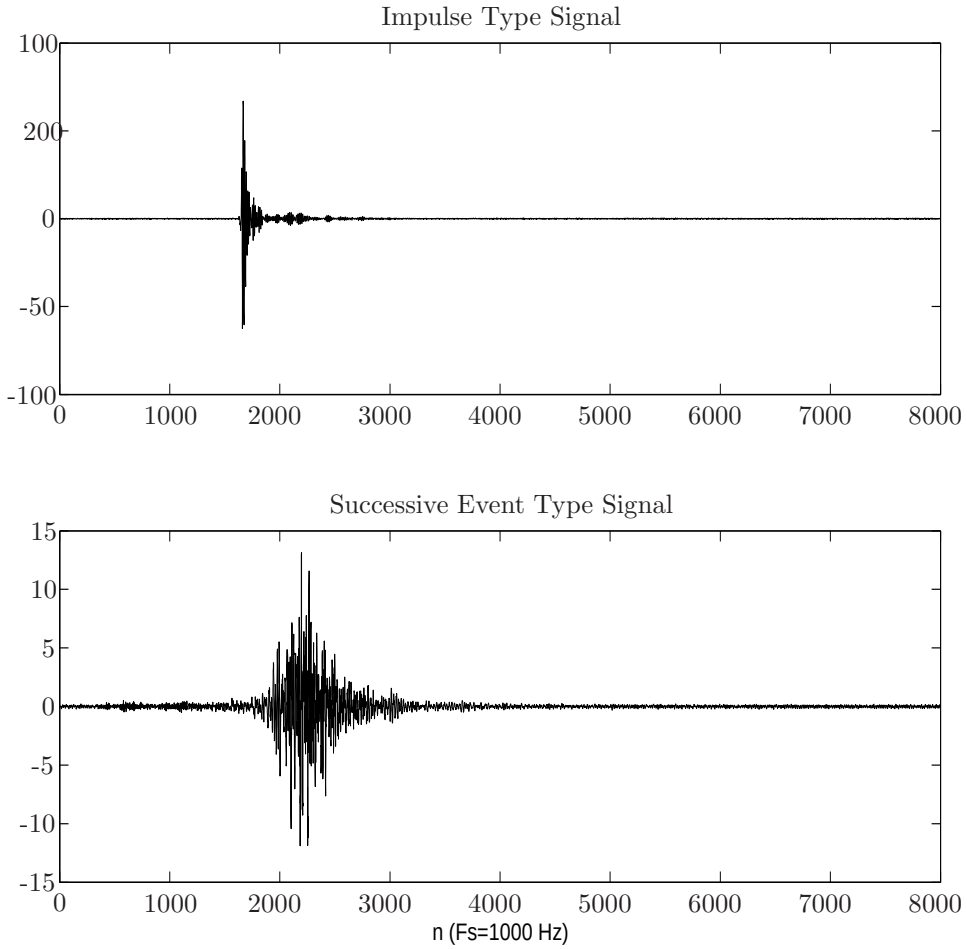

Figure 3.1: Typical signals of each classification

This classification algorithm separates each group into impulse and successive event type signals. Signals that are determined not to be one of these types are discarded because they are too corrupted by noise. In Lanigan and Allan mines there were sixty and seven signals, respectively, that were too corrupted by noise to fall into either of these types. The algorithm proceeds as follows:

1. If the energy is below 80 Joules (determined using the CANSEIS software [25]) discard the signal.

2. Subtract the DC bias from the original signal to give the zero mean signal $s(n)$.

3. If the sample variance of the first 1000 samples, Var, is greater than 1 discard the signal. The

sample variance is calculated, for the signal $s(n)$ using the following formula,

$$\text{Var} = \frac{1}{1000}\sum_{n=1}^{1000} s^2(n). \tag{3.2}$$

The first 1000 samples are used since these are known not to contain the signal due to the triggering method used to capture the event. This is a simple way to estimate the noise power of the trace and therefore ensure that noisy signals are not included in the characterization.

4. Calculate the envelope of the signal, $s_h(n)$ using the Hilbert Transform, $\mathcal{H}$, as in equation (3.3) and then smooth it out using a lowpass filter.

$$s_h(n) = \big|s(n) + j\mathcal{H}(s(n))\big|. \tag{3.3}$$

The filter is a 5th order lowpass constant phase finite impulse response filter designed using the window method and has a corner frequency of 150 Hz. Its purpose is to ensure that there are no fast changes in the envelope, that is, smooth it out. The output of the lowpass filter is $s_e(n)$.

5. Find the index, n_m, at which the peak value of $s_e(n)$ occurs.

6. If $s_e(n_m)/\text{Var} \leq 700$ discard the signal. This step ensures that the peak value is well above the noise power. The purpose of this step is similar to that of 3, in that it will catch signals with low SNR.

7. If $s_e(n_m + 40) \leq s_e(n_m)/4$ and $s_e(n_m - 30) \leq s_e(n_m)/20$ then the signal is an impulse signal. The step recognizes an impulse type event which has the properties that the peak is high, and that it is steeper on the rising edge than the falling edge.

8. The signals that are left are assumed to be of a successive event type signal.

The threshold values used in steps 6 and 7 were determined through a trial and error process to ensure that the signals were classified properly. Once the signals are classified it is simply a matter of estimating the spectrum for each signal, normalizing the spectrum, and finally averaging over all spectrums. The important decision in this procedure is the selection of an estimation algorithm; the chosen algorithms, used here, will be described in the following two sections.

3.1.2 Multitaper method

The multitaper method, first proposed by Thomson [27], has been used frequently in geophysical applications [18], [28], [29]. To assist in the simplification of notation and to be consistent with the previously cited papers on this topic the frequency variable f is assumed to be continuous in the following description. It should be noted, though, that in the actual implementation it will be a sampled version of the continuous spectrum.

The multitaper method involves forming K spectral estimates,

$$\hat{S}_k^{(mt)}(f) = T_s \left| \sum_{n=1}^{N} h_k(n) s(n) e^{j2\pi f n T_s} \right|^2, \tag{3.4}$$

each using a different taper; where $s(n)$ is the data record with DC bias removed, T_s is the sampling period, N is the length of the data record and the sequence $h_k(n)$,

$$h_k(n) = \begin{cases} v_k(n-1; N, W) & n = 1, \ldots, N \\ 0 & \text{otherwise} \end{cases}, \tag{3.5}$$

is called the kth data taper; where $v_k(n; N, W)$ is the kth order discrete prolate spheroidal sequence (dpss). The value of W is calculated from the time–bandwidth product which in this case is chosen as four. The kth dpss is the eigenvector of the equation [30],

$$\mathbf{Ag} = \lambda_k \mathbf{g}. \tag{3.6}$$

Here, $\mathbf{A}$ is an NxN matrix with elements

$$A_{nm} = \frac{\sin[2\pi W(m-n)]}{\pi(m-n)}. \tag{3.7}$$

The kth dpss is the eigenvector $v_k(n; N, W)$ corresponding to the kth eigenvalue λ_k. Some methods of determining the eigenvalues and eigenvectors are presented in Chapter 8 of [30].

By weighting the spectral estimators, the adaptive multitaper can be formed

$$\hat{S}^{(amt)}(f) = \frac{\sum_{k=0}^{K-1} b_k^2(f) \lambda_k \hat{S}_k^{(mt)}(f)}{\sum_{k=0}^{K-1} b_k^2(f) \lambda_k}, \tag{3.8}$$

where

$$b_k(f) = \frac{S(f)}{\lambda_k S(f) + (1 - \lambda_k)\sigma^2}. \tag{3.9}$$

Since the weights, $b_k(f)$ depend on the actual spectrum, $S(f)$, and variance, σ^2, they need to be estimated. The variance σ^2 is estimated as the sample variance of the signal $s(n)$. It is calculated as in equation (3.2) but with the index covering the length of the record, $N = 8000$,

$$\sigma^2 = \frac{1}{N} \sum_{n=1}^{N} s^2(n). \tag{3.10}$$

The initial estimate of $S(f)$ is the energy spectral density estimated using a basic Fast Fourier Transform. Using this estimate of $S(f)$, and the variance, the weights can be calculated. The weights can then be used in equation (3.8) to obtain the multitaper spectral estimate. This spectral estimate with the original variance σ^2 can then be used to recalculate the weights. This process can be repeated many times, though typically it is done only twice [30].

3.1.3 Combined Time and Lag Weighting

The combined time and lag weighting (CTLW) method can be seen as a general case of both the Weighted Overlapped Segment Averaging (WOSA) and the Blackman–Tukey methods [31]. As in the previous section the frequency variable f is assumed continuous and the $\mathcal{F}$ and $\mathcal{F}^{-1}$ represent the Discrete Time Fourier Transform and the Inverse Discrete Time Fourier Transform, respectively. The actual implementation, however, implements these through the use of the Fast Fourier Transform and inverse Fast Fourier Transform.

The method is a five step, or two stage, process. Like WOSA, the data $s(n)$ is first divided into P segments each of length L_1 and then multiplied by a weighting function $w_1(n)$. The pth segment after weighting is

$$y_p(n) = s(n)w_1(n - \frac{L_1}{2} - pS), \tag{3.11}$$

where S is the time shift. Each segment is then transformed into the Fourier domain and averaged to give

$$\hat{S}_1(f) = \frac{1}{P} \sum_{p=0}^{P-1} \left| \mathcal{F}[y_p(n)] \right|^2, \tag{3.12}$$

where $\mathcal{F}$ is the Fourier transform. The estimated autocorrelation $\hat{R}_1(n)$ is then obtained by taking the inverse Fourier transform,

$$\hat{R}_1(n) = \mathcal{F}^{-1}[\hat{S}_1(f)]. \tag{3.13}$$

A second–stage correlation estimate $\hat{R}_2(n)$ is obtained by multiplying the first-stage correlation estimate by a window $w_2(n)$,

$$\hat{R}_2(n) = \hat{R}_1(n)w_2(n). \tag{3.14}$$

Finally, the second–stage correlation estimate is transformed back into the Fourier domain to obtain the spectral density estimate

$$\hat{S}^{(ctlw)}(f) = \mathcal{F}[\hat{R}_2(n)], \tag{3.15}$$

where $(ctlw)$ refers to an estimate using the combined time and lag weighting method.

An important characteristic of this method is that if a rectangular weighting is chosen for $w_1(n)$, then by using the appropriate lag window $w_2(n)$ a result similar to one obtained with WOSA can be computed, but with half the multiplications. This is explained further in [31]. The two windows used in the implementation are discussed and explained in the following results section.

3.1.4 ESD Characterization Results

The two methods were implemented in MATLAB [26]. The multitaper method was performed by calling the built–in MATLAB function *pmtm* with a time–bandwidth product of 4 and a FFT size of 8192. The value of 4 was chosen as the time–bandwidth product as it is the common value used in many implementations of the multitaper method [30]. The FFT size of 8192 covers the entire record length with the addition of zeroes. This length of FFT was chosen for convenience, since picking only the portion of the record containing the signal, and zero padding, makes no difference to the final results. The characterization of successive event signals using this method for the Lanigan and Allan potash mines is given in Figures 3.2 and 3.3, respectively. The signals from Allan mine contained a strong 60 Hz component so this was filtered out using a notch filter and this is apparent in Figure 3.3. In Figures 3.2 and 3.3 the units are milli–Joules per square meter per Hz which is the energy flux density per frequency. Since the signals being analyzed are velocity

signals, they had to be multiplied by two constants, ρ and V which are equal to 2100 kg/m^3 and 2500 m/s [3]. The two constants are the rock density and the seismic velocity, respectively. In the characterizations presented in this paper these constants have no effect on the final results since each signal's ESD is normalized before being averaged with the others.

The combined time and lag weighting method was written as a custom function in MATLAB using a desired window, $w_2(n)$, known as the minimum 4–term window, described in [32]. The parameters L_1 and L_2 were set as 4096 and 2048, respectively, with an overlap of 50%. Also, the initial window, $w_1(n)$, was rectangular, so no multiplications were needed. The results of this characterization of successive event signals for Lanigan and Allan mines are plotted in Figures 3.4 and 3.5, respectively. Each mine used a total of 37 signals in the averaging process. The impulse type signal characterizations using the pmtm method for the Lanigan and Allan mines can be seen in Figures 3.6 and 3.7, respectively. Only the pmtm characterized impulse type waveform ESDs are shown since the differences between the two estimation methods can be adequately seen in the other figures. The impulse type waveforms are much less common; in Allan mine only two were used in the characterization and in Lanigan mine only nine were used.

3.2 Comparison

This section compares each signal type's ESD, the two mines, as well as the two estimation algorithms, using the results of the previous section. To further investigate the estimation algorithms the ESDs of each algorithm are used to evaluate the variance expression given in equation (2.26) in section 3.3.

3.2.1 Signal Types

Despite the apparent differences in the time domain, the ESDs of both types of signals are similar, they can be compared, for example, in Figures 3.2 and 3.6, and could be assumed to be the same type, for ESD purposes. This assumption is made throughout the remainder of this document so that the only type of signal considered is that of the successive event type. The most prevalent

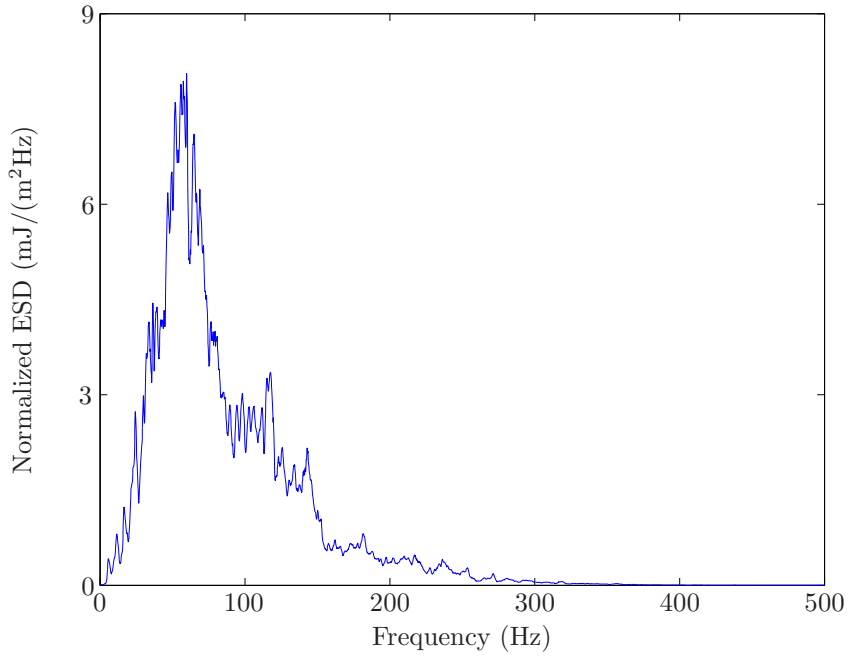

Figure 3.2: Successive Event type ESD Characterization of Lanigan Mine using the multitaper method

type of signal, by a large margin, is also the successive event type, so this assumption could also be applied for this reason.

3.2.2 Estimation Algorithms

Quantitative comparison of the two methods, so implemented, is not a trivial matter. Commonly, for a given frequency resolution, the bias and variance of spectral estimators can be compared. It is accepted that under reasonable assumptions that both estimators are unbiased [31] [30]. By taking the difference between the two spectral characterizations at each frequency index and then calculating the mean of that difference, it can be seen that it is approximately zero ($\sim 10^{-13}$). This indicates that if the estimators are biased, then they are biased by the same amount. It is reasonable to assume, then, that they are unbiased.

The two estimated ESDs are similar and therefore it is assumed that each estimator is equally suitable. The following subsection uses both mine's ESDs in an example. The remainder of this section provides references and reasons as to why it is difficult to compare the variance of each

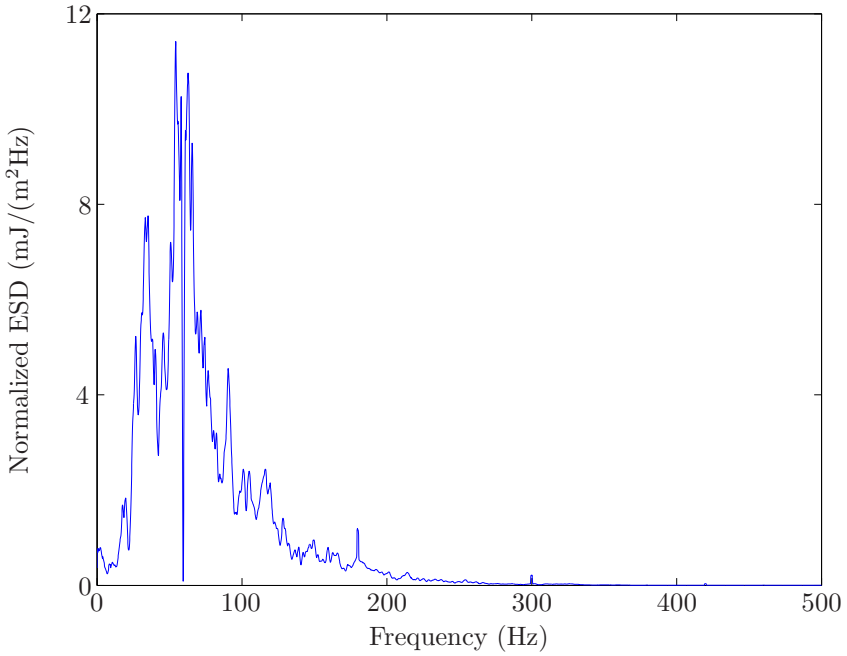

Figure 3.3: Successive Event type ESD Characterization of Allan Mine using the multitaper method

estimator. A closed form expression, assuming Gaussian noise, for the variance of the combined time and lag weighting estimator is given in [31]. There is no such expression for the multitaper method. The original work by Thomson [27] does give an expression of variance, but this is for the direct averaging of the K spectral estimates, and does not include the adaptive weighting used here. Walden [33] also gives an expression more amenable to computation, but it too, is just for the direct average of the spectral estimates. Bronez [34] compares two other estimators with direct averaging multitaper. It should be noted, once again, that in this thesis each estimator is assumed to be equally suitable.

3.2.3 Mine Comparison

There are clear differences between the characterized ESDs of Lanigan and Allan potash mines. In Allan mine, see figures 3.3, 3.5 and 3.7, the majority of the energy content is contained in two distinct frequency peaks. Lanigan mine, see figures 3.2, 3.4 and 3.6, by contrast, has a single peak whose base has similar bandwidth to that of Allan mine. The following subsection uses both mine

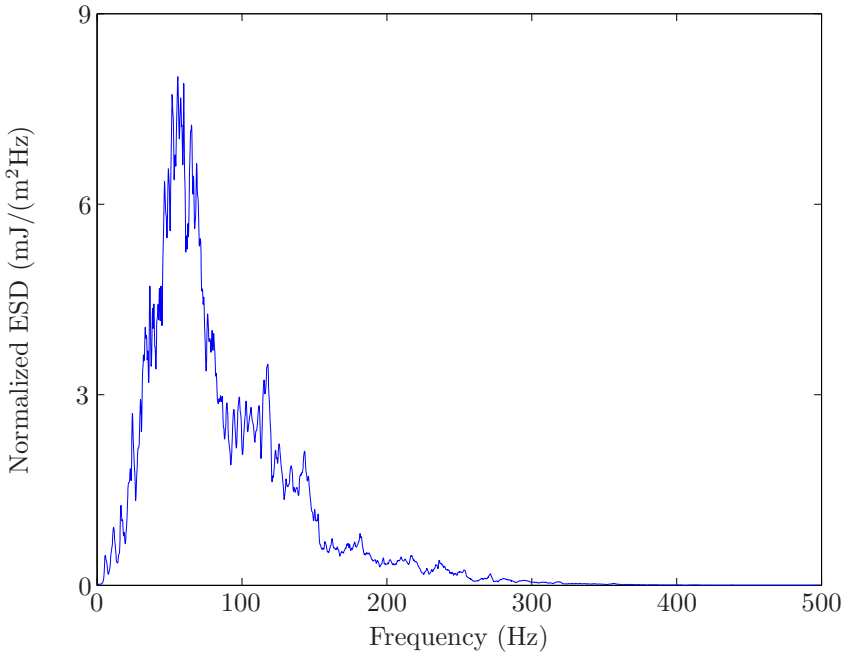

Figure 3.4: Successive Event type ESD Characterization of Lanigan Mine using the combined time and lag weighting method

characterizations to illustrate how this difference in frequency content affects their evaluation.

3.3 Example

The ESD of a microseismic event was characterized so that it could be used to evaluate the variance of error expression given by equation (2.26). This calculation forms the basis of the sensor positioning algorithm, explained in Chapter 4. It is, therefore, important to see the effect that the four ESDs, calculated using different estimation algorithms and using data from two different mines, have when used in this expression. The parameters used to evaluate equation (2.26) are summarized below.

The sensor system used in the evaluation of the expression forms a rectangle with sensor coordinates: $\vec{p_1}$ = (600,300,0), $\vec{p_2}$ = (-600,300,0), $\vec{p_3}$ = (600,-300,0), and $\vec{p_4}$ = (-600,-300,0). A value of $[1/\sqrt{2},1/\sqrt{2},0]$ was used for $\vec{e}$. The results here, as in [20], use an ideal lowpass spectrum of

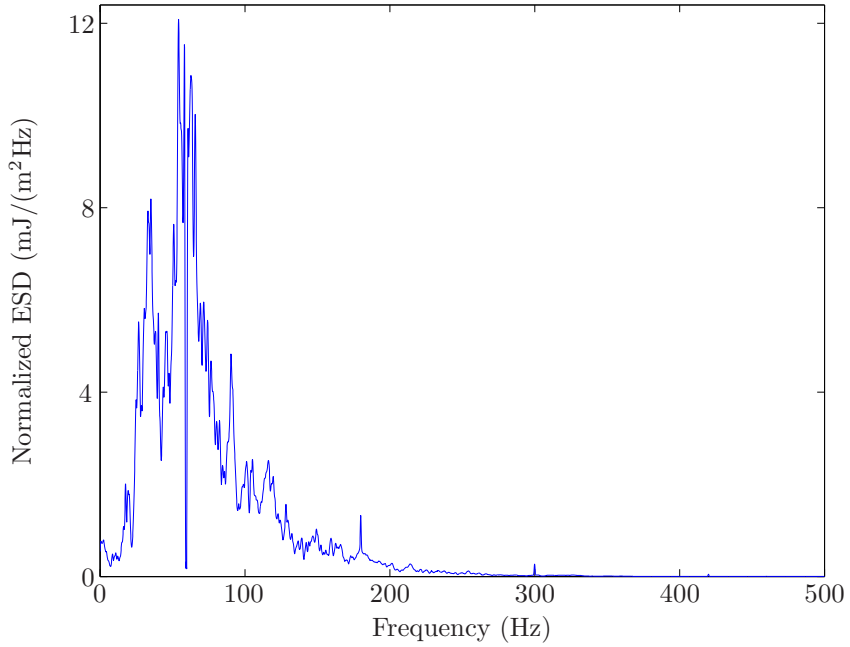

Figure 3.5: Successive Event type ESD Characterization of Allan Mine using the combined time and lag weighting method

bandwidth B Hz with the SNR defined is

$$SNR = 10 \log \left(\frac{\int_{-2\pi B}^{2\pi B} S_{ss}(\omega) d\omega / d_i^2}{T \int_{-2\pi B}^{2\pi B} S_{nn}(\omega) d\omega} \right) \tag{3.16}$$

where T is the duration of the seismic event in seconds and d_i is the distance from the source to the nearest sensor in meters. For the results presented, the SNR is kept constant at 10 dB while the signal ESD is multiplied by the appropriate constant for a given bandwidth. The noise is assumed white and therefore the spectrum is a constant. The signal duration is kept constant at 0.2 seconds. Since the signal duration remains constant the bandwidth is changed to obtain different BT products. This is done by setting the frequency content in the ESD above the chosen bandwidth to zero. The speed of the mechanical wave, c, assumed a value of 2581 m/s. The expression was evaluated and plotted in Figure 3.8 using five different signal ESDs: a white ESD, the two characterized ESDs from Lanigan mine and two characterized ESDs from Allan mine, all using the successive event type signal. The solid line is calculated using a white spectrum for both signal and noise. The dashed line and dotted line uses the CTLW characterized signal spectrum and white noise for Lanigan and Allan mines, respectively. The $*$'s and the o's represent the pmtm

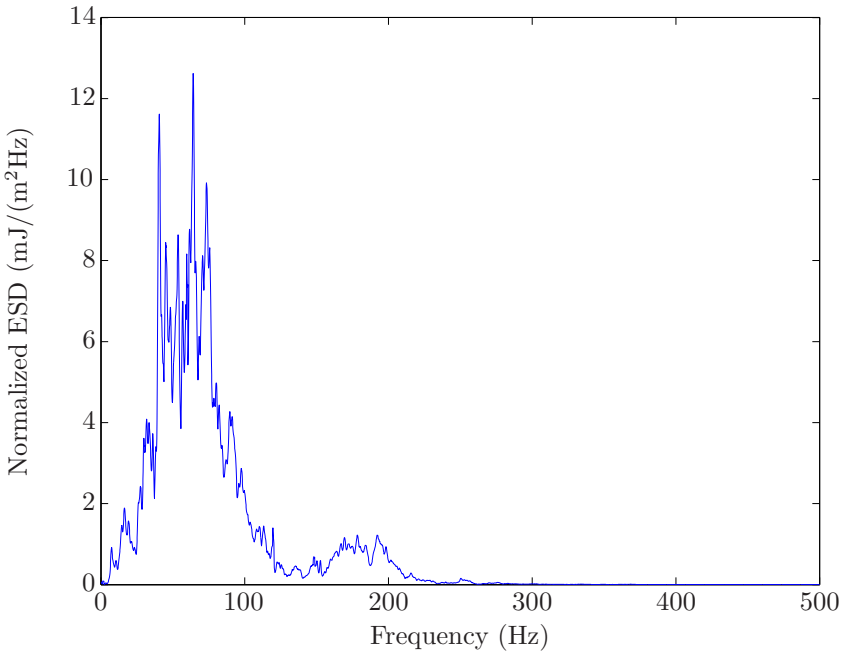

Figure 3.6: Impulse type ESD Characterization of Lanigan Mine using the multi-taper method

method characterized signal spectrum and white noise for Lanigan and Allan mines, respectively.

3.4 Summary and Discussion

As can be seen in Figure 3.8 there is little noticeable difference between the curves using the two methods of characterization. Therefore, if efficiency is a concern, the combined time and lag weighting method, should be chosen. Since the CTLW method is more efficient, the ESD characterized with this method is used exclusively in the following chapters when evaluating equation (2.26). There is, however some noticeable difference between the two mines, especially at higher bandwidth time products. Therefore, it is necessary to characterize the signal ESD in a specific mine if accurate results are desired when evaluating expression (2.26).

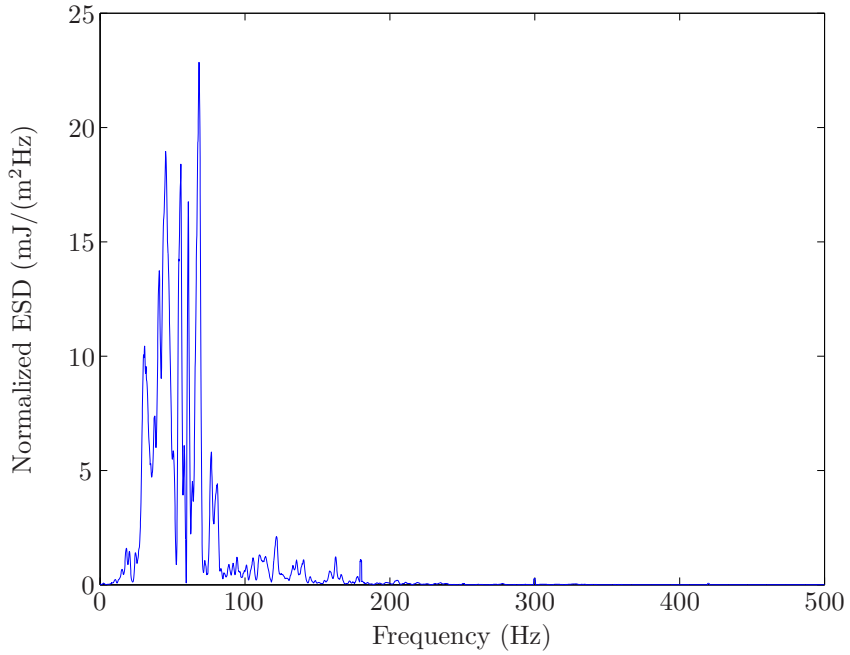

Figure 3.7: Impulse type ESD Characterization of Allan Mine using the multitaper method

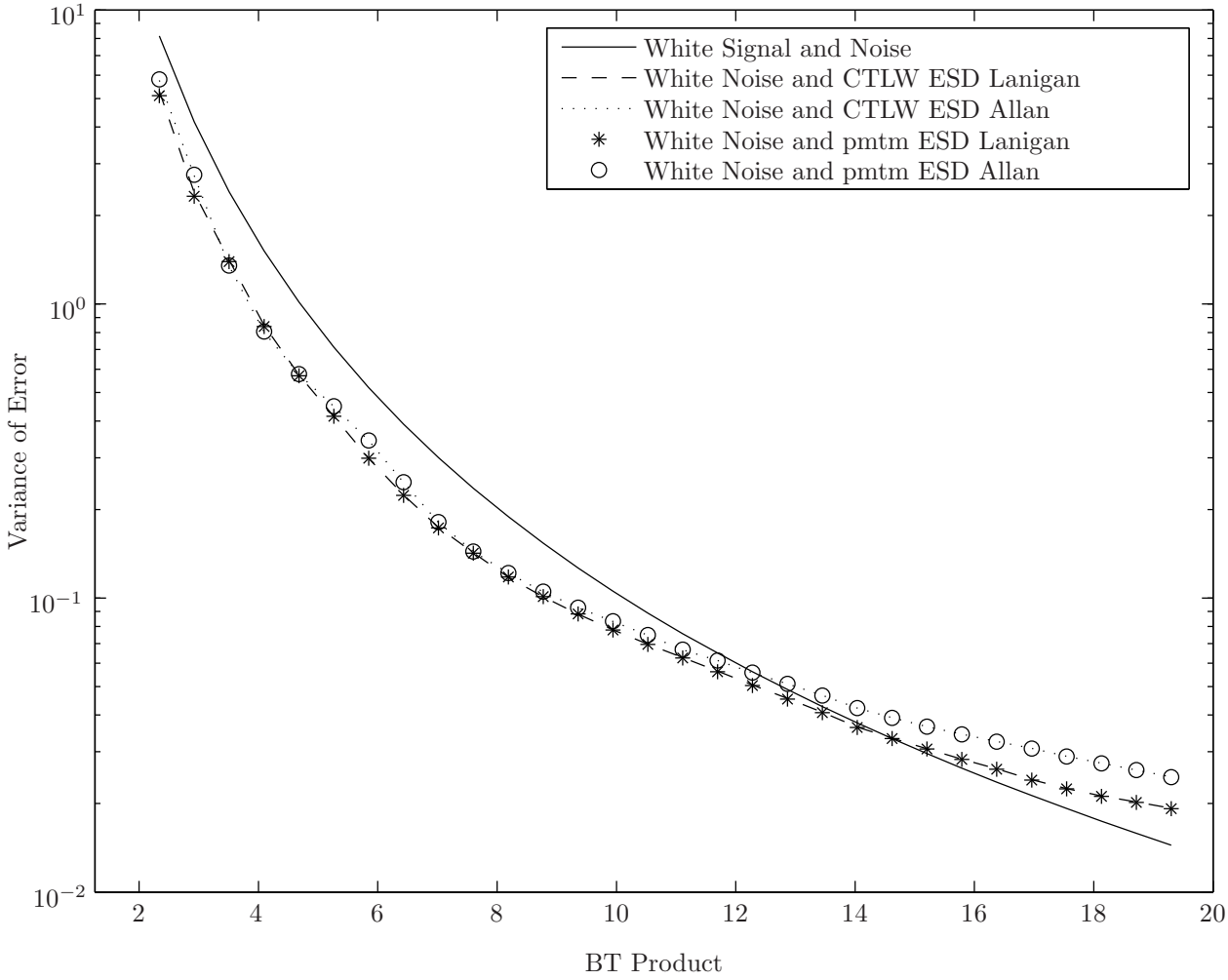

Figure 3.8: Comparison of the variance of error expression evaluated using different ESDs

Chapter 4

Sensor Placement Algorithms

The errors associated with the location of a seismic event are partially dependent upon the spatial distribution, or configuration, of geophones. The measure of performance for one configuration of sensors over another should depend upon the location error for the respective configuration. The variance expression described in equation (2.26) is used as the measure of performance in the presented algorithms. These algorithms which are proposed to select the best location for future sensors will be described in section 4.2. These algorithms are then compared with those outlined in the literature in section 4.3. Finally, some modifications to the algorithms are discussed in section 4.4. Section 4.1 reviews the current literature related to optimum sensor placement.

4.1 Background on Sensor Placement

The first effort to place sensors in optimum positions to monitor seismic events was performed in 1965 and 1966 by Sato and Skoko [35–37]. They did this by modelling the process of seismic wave recording using the Monte–Carlo method.

A general theory of optimum sensor placement was first presented by Kijko [38,39] in terms of the statistical theory of optimum experiments. In the first paper the D-criterion, that is the determinant of $\mathbf{A}^T\mathbf{A}$, where $\mathbf{A}$ is defined in equation (2.6), was introduced as the objective function to maximize. Maximizing $\det|\mathbf{A}^T\mathbf{A}|$ is equivalent to minimizing $\det|\mathbf{C}_\theta(\mathbf{P})|$ since $\mathbf{C}_\theta(\mathbf{P}) \propto (\mathbf{A}^T\mathbf{A})^{-1}$, where $\mathbf{C}_\theta(\mathbf{P})$ is the covariance matrix defined in equation (2.11). For cases where there is a region of seismicity that needs to be monitored it is first partitioned into discrete regions, each of which is assigned a weight. The weight assigned to a given partition is related to the probability of microseismic event occurance. The location of each partition, usually near the center, is used to

evaluate $\det|\mathbf{A}^T\mathbf{A}|$. This value is then multiplied by the weight assigned to that partition. Finally these values are summed. The configuration which maximizes this weighted sum is chosen. The second of his, often called landmark, papers looks at the accuracy of 7 different 'optimal' seismic networks. The networks consist of 4 to 10 stations. The accuracy sees the most improvement when a 4 station network is replaced by one of 5 or 6 stations. Burmin then looked at a nonstatistical analysis of opitmal stations [40]. He showed that a system of stations optimal for certain 'data' ceases to be optimal if the data is changed. By 'data' he implies the unknown parameters such as origin time, hypocentre, or epicentre.

In 1990 Rabinowitz and Steinberg used the DETMAX algorithm to find optimal configurations for a single source [41]. Another important contribution of this paper was the inclusion of spatially correlated error terms. As with Kijko's papers they used the D-criterion from the theory of optimal experiment design. Then in 1995 Steinberg et al. expanded their work to include multiple sources [42]. They also proposed the D-criterion for Multiple Sources (DMS), which uses a weighted sum of the logarithm of $\det|\mathbf{A}^T\mathbf{A}|$. This is opposed to Kijko's criterion which is a weighted sum of determinants. They felt that Kijko's criterion produced networks that tended to over monitor sources for which good coverage is possible but ignored sources that are difficult to monitor [42]. Also in 1995, Kijko and Sciocatti modified their criterion to take into account the energy, and therefore the detectable distance of the seismic source [43].

Bartal et al, in 2000, in the context of the Comprehensive Test Ban Treaty (CTBT), used a minmax approach to sensor placement [44]. That is, the source which produces the maximum confidence ellipse is found and then the configuration of seismic stations which minimize this maximum is chosen. He found the configuration using both a Genetic Algorithm (GA) and a Random Search (RS); both of these were then compared to an exhaustive search. The GA was found to be superior.

Rabinowitz and Steinberg, in 2000, proposed two optimzation criteria [45]. The first, which is also based on the D-criterion, is called the Average Information Criterion, and is the determinant of the weighted sum of $\mathbf{A}^T\mathbf{A}$. It is useful in problems where singular $\mathbf{A}^T\mathbf{A}$ might arise. Also, proposed was an *importance* based criteria. The importance referred to is that of a phase arrival

at a particular station. Finally, in 2003 Steinberg and Rabinowitz rigorously derived optimum (D–optimum) networks for various constraints [46].

4.2 Placement Algorithms

Using both the results of [47] and the variance expression given by equation (2.26) the performance of a sensor network can be calculated for a given microseismic event location. The functional dependence of the variance upon the M sensor positions, p_i and upon the parameter $\vec{e}$, which is the unit vector in the direction that the variance is calculated, can be explicitly expressed as

$$\mathcal{V}(\vec{e}, \vec{p}_1, \ldots, \vec{p}_M) = \sigma^2, \tag{4.1}$$

where the source position is assumed to be at the origin. To use an alternate source position the sensor positions can be shifted by the alternate source location. The following two subsections use this expression in their criteria for the placement of additional sensors in a network to adequately monitor expansion regions. The two criteria are similar in that they both seek to minimize the variance of a source position with the maximum variance. The difference lies in that the MAXSRC algorithm minimizes the maximum source position before any sensor is placed while the MINMAX algorithm minimizes the source position with maximum variance after a sensor has been placed. The two algorithms are discussed in detail in the following two subsections.

4.2.1 MAXSRC Algorithm

The algorithm first searches for the source position with the maxium variance, expressed mathematically as

$$\max_i \mathcal{V}(\vec{e}, \vec{p}_1 - \vec{\theta}_i, \ldots, \vec{p}_M - \vec{\theta}_i), \tag{4.2}$$

where $\vec{\theta}_i \in \Omega_E$. Here Ω_E represents the mine expansion region and $\vec{\theta}_i$ are the possible hypocentre locations. The subtraction of $\vec{\theta}_i$ from the sensor positions translates the equation so that it is evaluated as if the source location was $\vec{\theta}_i$. From this expression $\vec{\theta}_m$, the source position which produces the maximum variance, is obtained. Using $\vec{\theta}_m$ as the hypocentre location additional

sensor positions are tried throughout the expansion region. The position which minimizes the error is chosen. This can be expressed mathematically as,

$$\min_j \mathcal{V}(\vec{e}, \vec{p}_1 - \vec{\theta}_m, \ldots, \vec{p}_M - \vec{\theta}_m, \vec{p}_j - \vec{\theta}_m), \tag{4.3}$$

where $\vec{p}_j \in \Omega_R$. Here Ω_R represents rooms in the mine expansion region. Note that $\Omega_R \subset \Omega_E$. The search area for the sensor is narrowed to Ω_R because placement outside of this area would require bore holes to be drilled into the surrounding ore body and would therefore be quite expensive. It was discovered, after many trials, that this algorithm would invariably place the additional sensor as close as possible to the location of the maximum source. This then leads to a more efficient criteria than that proposed in equation (4.3). After searching for the maximum position, using equation (4.2), look for the closest possible sensor position to $\vec{\theta}_m$

$$\min_j \left\| \vec{p}_j - \vec{\theta}_m \right\|, \tag{4.4}$$

where, once again, $\vec{p}_j \in \Omega_R$ and $\| \cdot \|$ denotes the Euclidian distance. It has been seen that placing a sensor using equation (4.4) instead of equation (4.3) produces little, if any, noticeable difference.

It may seem strange placing a sensor as close as possible to a source location when the model breaks down at this point. However, it must be remembered that a *region* of seismicity is desired to be monitored, not necessarily this particular source location.

If more than a single sensor is desired to be placed then one can run the algorithm again, but adding the previously placed sensor to the current configuration. Multiple sensors can, therefore, be placed in an iterative fashion.

The pseudocode for this algorithm can be seen in Algorthm 1. The function FunMaxSrc accepts the current sensor configuration $\mathbf{P}$, the direction which the variance will be calculated $\vec{e}$, the set of possible source locations Ω_E and the set of possible sensor positions Ω_R. The first loop, from lines 2 to 9 searches over the possible source locations in Ω_E for the position with the maximum variance. The second loop then searches over the set of possible sensor positions and finds the sensor position p' which is the closest to the source position with the maximum variance θ_m.

The following section will evaluate this method's performance by comparing it to the more

Algorithm 1 MAXSRC implementation

1: **function** FUNMAXSRC($\mathbf{P}, \Omega_E, \Omega_R, \vec{e}$)
2: **for all** $\theta \in \Omega_E$ **do**
3: $\mathbf{P}' \leftarrow \mathbf{P} - \theta$
4: $\mathcal{V}' \leftarrow \mathcal{V}(\vec{e}, \mathbf{P}')$
5: **if** $\mathcal{V}' \geq \mathcal{V}_t$ **then**
6: $\theta_m \leftarrow \theta$
7: $\mathcal{V}_m \leftarrow \mathcal{V}'$
8: **end if**
9: **end for**
10: **for all** $p \in \Omega_R$ **do**
11: $d \leftarrow ||p - \theta_m||$
12: **if** $d \leq d'$ **then**
13: $d' \leftarrow d$
14: $p' \leftarrow p$
15: **end if**
16: **end for**
17: **return** $(p', \mathcal{V}_m, \theta_m)$
18: **end function**

complex criteria outlined below.

4.2.2 MINMAX Algorithm

The MAXSRC algorithm, described above, attempts to minimize the variance at the source position that has the maximum variance. Unfortunately this says nothing regarding the maximum variance after the additional sensor has been placed. It could, possibly, be worse. To combat this deficiency a second criterion, termed MINMAX, is proposed. It finds the sensor position which minimizes the maximum variance *after* it has been placed. This is expressed mathematically as

$$\min_j \left[\max_k \mathcal{V}(\vec{e}, \vec{p}_1 - \vec{\theta}_k, \ldots, \vec{p}_M - \vec{\theta}_k, \vec{p}_j - \vec{\theta}_k) \right]. \tag{4.5}$$

This criterion can be expanded to two sensors quite easily,

$$\min_{j,l} \left[\max_k \mathcal{V}(\vec{e}, \vec{p}_1 - \vec{\theta}_k, \ldots, \vec{p}_M - \vec{\theta}_k, \vec{p}_j - \vec{\theta}_k, \vec{p}_l - \vec{\theta}_k) \right]. \tag{4.6}$$

The pseudocode implementation of the single sensor case is given in Algorithm 2. The inputs are the same as in Algorithm 1. The inner loop searches for the source position with the maximum variance. Then for each sensor position it checks to see if the current sensor position's corresponding maximum variance is the smallest. If it is then this sensor position's maximum variance value and source position are saved. This continues until all sensor positions in Ω_R have been checked.

Because of the increased complexity of the MINMAX algorithm it is not practical to conduct an exhaustive search for the sensor position(s). In the following chapter a Genetic Algorithm (GA) is used. Note that this criterion and the use of a GA are similar to that used by Bartal et. al. [44]. In fact they are the same with the exception that in this thesis the variance in the direction $\vec{e}$ is used as the quantity to minimize while Bartal et. al. uses the determinant of $\mathbf{A}^T\mathbf{A}$. This and other differences are discussed briefly in the following section.

4.3 Justification

There are three major differences between the algorithm presented here and those seen in the literature. As discussed in Chapter 2, a new single–step energy based algorithm shows promise as

Algorithm 2 MINMAX implementation

```
1: function FunMinMax(P, Ω_E, Ω_R, e⃗)
2:     for all p ∈ Ω_R do
3:         for all θ ∈ Ω_E do
4:             P′ ← [P : p] − θ
5:             V′ ← V(e⃗, P′)
6:             if V′ ≥ V_t then
7:                 θ_t ← θ
8:                 V_t ← V′
9:             end if
10:        end for
11:        if V_t ≥ V_m then
12:            p′ ← p
13:            θ′ ← θ_t
14:            V_m ← V_t
15:        end if
16:    end for
17:    return (p′, θ′)
18: end function
```

an alternative estimator for the location of microseismic events induced by mining activity. The variance of error in a particular direction, presented in equation (2.26), makes a natural choice for the performance measure of a given network configuration.

Potash mines in Saskatchewan are constantly expanding. They are growing at a rate of approximately 500 metres per hour. It would not make sense to reconfigure the entire sensor network each time an expansion is made. This would be both costly and inefficient. To adequately monitor new areas of mines, termed here mine expansion regions, an additional sensor should be added to specifically monitor the new region. This allows the current network to remain undisturbed as well as simplifying the algorithm used to place an additional sensor. More than one sensor can be placed in a given expansion region by running the algorithm in an iterative fashion.

In summary the three major differences are:

1. The variance of error in a particular direction is used as network configuration measure
2. Only one or two sensors are placed at a time
3. The search area for microseismic events and sensor positions is restricted to the mine expansion region

The following section discusses the possible choices that can be made regarding $\vec{e}$, the vector in the direction which the variance is calculated.

4.4 Parameter Modification

The parameter $\vec{e}$ in equation (2.26) allows extensive flexibility in implementing the algorithm described in section 4.2. This section shows some of the possibilities involving the $\vec{e}$ parameter and the following two subsections explain these modifications in more detail. Some of these possible modifications will be used in examples in the following chapter.

Changing the direction of $\vec{e}$ can be very useful in mining applications. For example, in potash mines the rooms are often parallel to each other. If one would like to know which room a microseismic event took place near, then the $\vec{e}$ vector could be chosen so that it is in the direction

perpendicular to the mine rooms. It should be stressed, however, that there is no knowledge of the other directions' variances. That is, even though the maximum variance in two possible directions is β and even though they may be perpendicular to each other there could (and likely is) another direction with a much larger variance since the orientation of the error ellipse is unknown. With that said this method is very useful in situations where the variance in certain directions, and those directions only, is wished to be controlled.

Instead of just choosing different $\vec{e}$ vectors, many can be used. By running the first part of the algorithm, finding the worst possible event location, for different $\vec{e}$ vectors a worst possible solution out of all of them can be determined. This solution can then be used, along with direction vector $\vec{e}$ which produced it, to determine the sensor position to minimize the error. An obvious choice for $\vec{e}$ vectors in this case would be the cartesian coordinate unit vectors.

Finally, an average of multiple $\vec{e}$ vectors can be used to obtain a solution which performs well in the given directions. First, the worst possible event location would be found; worst in the sense that it produced the maximum average. Then the sensor would be found which minimized the average.

In summary the three categories of modifications involving the direction vector $\vec{e}$, are:

1. The direction of $\vec{e}$ can be changed based on the error requirements of the specific area

2. A number of $\vec{e}$ directions can be chosen, and then that which is the worst can be the chosen direction in which to minimize the error

3. An average can be taken of multiple directions

Clearly the final two modifications would require more computation time than the first. To better explain how these modifications will affect the criteria the two following subsections will discuss and present the modified criteria for the MAXSRC and MINMAX alrgorithms.

4.4.1 MAXSRC

The modifications required of the MAXSRC algorithm are very simple. For the case where multiple $\vec{e}$ vectors are used the procedure is, as follows. First, the maximum variance is found for each of the $\vec{e}$ vectors. The sensor is then placed as close as possible to the position which generated the maximum variance over all $\vec{e}$ vectors.

To modify the algorithm for a weighted average of the multiple $\vec{e}$ vectors the variance at each source position is calculated for both vectors and is then averaged (or weighted averaged). The sensor is then placed as close as possible to the source position with the largest weighted average.

4.4.2 MINMAX

The MINMAX algorithm is easily extended to the multiple parameter case. Instead of minimizing the maximum variance over the possible source positions another dimension is added, the $\vec{e}$ dimension. The criterion for a single sensor placement of equation (4.5) can be modified for the additional dimension so that it becomes

$$\min_j \left[\max_{k,l} \mathcal{V}(\vec{e_l}, \mathbf{P} - \theta_k, \vec{p}_j - \vec{\theta}_k) \right]. \tag{4.7}$$

Here, the notation has been made more compact where $\mathbf{P} - \theta_k$ implies the set of current sensors $\mathbf{P}$ is being shifted to source location θ_k. In the examples presented in the following chapter only two values are used for $\vec{e}$. With only two values of $\vec{e}$ the above criterion can be written as

$$\min_j \left[\max_k [\mathcal{V}(\vec{e_1}, \mathbf{P} - \theta_k, \vec{p}_j - \vec{\theta}_k), \mathcal{V}(\vec{e_2}, \mathbf{P} - \theta_k, \vec{p}_j - \vec{\theta}_k)] \right]. \tag{4.8}$$

Algorithm 2 can be easily modified to include the special case of an additional $\vec{e}$. This modified algorithm can be seen as in Algorithm 3.

Instead of finding the maximum over the additional parameters the variances can be averaged for each $\vec{e}$ so that the maximum average can be minimized. For the general case of L different $\vec{e}$ vectors the criterion is

$$\min_j \left[\max_k \sum_{l=1}^{L} \alpha_l \mathcal{V}(\vec{e_l}, \mathbf{P} - \theta_k, \vec{p}_j - \vec{\theta}_k) \right]. \tag{4.9}$$

Algorithm 3 MINMAX implementation

1: **function** FUNMINMAX2E($\mathbf{P}, \Omega_E, \Omega_R, \vec{e}_1, \vec{e}_2$)
2: **for all** $p \in \Omega_R$ **do**
3: **for all** $\theta \in \Omega_E$ **do**
4: $\mathbf{P}' \leftarrow [\mathbf{P} : p] - \theta$
5: $\mathcal{V}'_1 \leftarrow \mathcal{V}(\vec{e}_1, \mathbf{P}')$
6: $\mathcal{V}'_2 \leftarrow \mathcal{V}(\vec{e}_2, \mathbf{P}')$
7: **if** $\mathcal{V}'_1 \geq \mathcal{V}_t$ **then**
8: $\theta_t \leftarrow \theta$
9: $\mathcal{V}_t \leftarrow \mathcal{V}'_1$
10: **end if**
11: **if** $\mathcal{V}'_2 \geq \mathcal{V}_t$ **then**
12: $\theta_t \leftarrow \theta$
13: $\mathcal{V}_t \leftarrow \mathcal{V}'_2$
14: **end if**
15: **end for**
16: **if** $\mathcal{V}_t \geq \mathcal{V}_m$ **then**
17: $p' \leftarrow p$
18: $\theta' \leftarrow \theta_t$
19: $\mathcal{V}_m \leftarrow \mathcal{V}_t$
20: **end if**
21: **end for**
22: **return** (p', θ')
23: **end function**

where α_l is the weighting factor for $\vec{e}_l$ which for the average would be $\frac{1}{L}$. If only two values of $\vec{e}$ are used in an average, then this becomes

$$\min_j \left[\max_k \frac{1}{2}\mathcal{V}(\vec{e_1}, \mathbf{P} - \theta_k, \vec{p}_j - \vec{\theta}_k) + \frac{1}{2}\mathcal{V}(\vec{e_2}, \mathbf{P} - \theta_k, \vec{p}_j - \vec{\theta}_k) \right]. \tag{4.10}$$

Algorithm 2 is similarly modified to include the average for the special case of two $\vec{e}$ vectors. This can be seen in Algorithm 4.

Algorithm 4 MINMAX implementation

function FUNMINMAX2EAVG($\mathbf{P}, \Omega_E, \Omega_R, \vec{e}_1, \vec{e}_2$)
  for all $p \in \Omega_R$ **do**
    for all $\theta \in \Omega_E$ **do**
      $\mathbf{P}' \leftarrow [\mathbf{P} : p] - \theta$
      $\mathcal{V}_1' \leftarrow \mathcal{V}(\vec{e}_1, \mathbf{P}')$
      $\mathcal{V}_2' \leftarrow \mathcal{V}(\vec{e}_2, \mathbf{P}')$
      $\mathcal{V}' \leftarrow \frac{1}{2}\mathcal{V}_1' + \frac{1}{2}\mathcal{V}_2'$
      if $\mathcal{V}' \geq \mathcal{V}_t$ **then**
        $\theta_t \leftarrow \theta$
        $\mathcal{V}_t \leftarrow \mathcal{V}'$
      end if
    end for
    if $\mathcal{V}_t \geq \mathcal{V}_m$ **then**
      $p' \leftarrow p$
      $\theta' \leftarrow \theta_t$
      $\mathcal{V}_m \leftarrow \mathcal{V}_t$
    end if
  end for
  return (p', θ')
end function

4.5 Summary

This chapter presented two classes of sensor placement algorithms for sensors in mine expansion regions. An overview of the previous literature in seismic network configuration was given and then the differences between this and previous work was highlighted. It was also shown how each of the classes of algorithms could be modified to include multiple $\vec{e}$ directions in three different ways. The following chapter will use each class of algorithms on three examples; two examples are synthetic while the third is of a potash mine located near Saskatoon, Saskatchewan.

Chapter 5

Sensor Placement Examples

Three examples are presented in this chapter. All three examples compare both MAXSRC and MINMAX algorithms. The first two examples are synthetic. Their purpose is to introduce the algorithms using simple sensor configurations. The final example applies the algorithms to the sensor network in Lanigan Potash Mine. For each example, sensors are placed using six variations of the MINMAX and MAXSRC algorithms. The first two are single $\vec{e}$ cases of MINMAX and MAXSRC. The final four use two $\vec{e}$ vectors and are referred to as: MINMAX_both, which minimizes the maximum variance calculated using each $\vec{e}$, MINMAX_avg, which minimizes the maximum variance calculated using the average of both $\vec{e}$ vectors, and similarly, MAXSRC_both and MAXSRC_avg. The final section summarizes the results of this chapter.

5.1 Simple Examples

In this section, the six algorithms described above, MINMAX, MAXSRC, MINMAX_both, MINMAX_avg, MAXSRC_both, and MAXSRC_avg, are each used to place both one and two additional sensors in two different four station configurations. The sets Ω_E and Ω_R are the same and each contain 25 elements. The search area is a 100 x 100 square area that is sampled every 25 units. In the two cases to follow the original sensor configuration is inside the 'expansion region.' The variance is calculated in the direction $\vec{e_1}$=[1 0 0], for the single $\vec{e}$ case and $\vec{e_2}$=[0 1 0] is added for the two $\vec{e}$ case. These examples assume a two dimensional system located on the $z = 0$ plane. The ESD of the signal used in the variance calculation is the CTLW successive event type characterization for Lanigan Mine. The SNR used is -11.5 dB. Both of these were chosen for convenience.

5.1.1 Configuration 1

The configuration of the initial sensor network is given in Figure 5.1. The sensors are located on a square with $p_1 = (25, 25)$, $p_2 = (25, 75)$, $p_3 = (75, 75)$ and $p_4 = (75, 25)$, where $z = 0$ is assumed. The results of placing one and two additional sensors using the six algorithms is shown in Table 5.1.

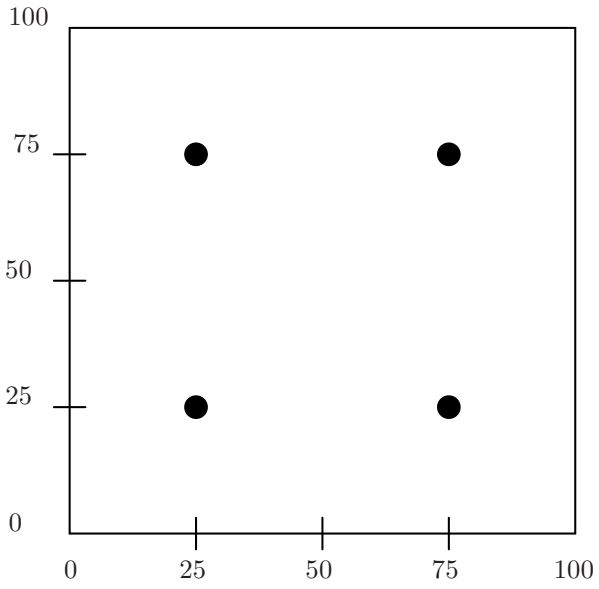

Figure 5.1: Configuration of sensors in Configuration 1

Those algorithms using a GA, MINMAX, MINMAX_avg, and MINMAX_both, were run 10 times to obtain the average and minimum values. In the first column of Table 5.1 the numbers (1) and (2) following each method indicate the number of sensors placed. The second and third columns display the average standard deviation (SD) and minimum SD, in meters, respectively. Clearly, for the MAXSRC methods these two columns will be identical, since they are solved without use of the GA. The sensor positions listed, in the fourth column, were the sensor positions that obtained the minimum variance. An illustration of the best positions to locate two additional sensors when $\vec{e}_1$ is used is given in Figure 5.2. The additional sensors are represented as circles.

Method	Avg. SD (m)	Min. SD (m)	Sensor Positions
MINMAX(1)	5.94	5.94	(50,50)
MINMAX(2)	2.05	1.89	(0,50) (100,50)
MAXSRC(1)	6.19	6.19	(100,50)
MAXSRC(2)	1.89	1.89	(100,50) (0,50)
MINMAX_both(1)	6.07	5.94	(50,50)
MINMAX_both(2)	3.39	2.33	(0,0) (100,100)
MAXSRC_both(1)	8.27	8.27	(50,0)
MAXSRC_both(2)	5.84	5.84	(50,0) (0,50)
MINMAX_avg(1)	4.59	4.59	(50,50)
MINMAX_avg(2)	3.09	2.33	(0,100) (100,0)
MAXSRC_avg(1)	5.90	5.90	(50,0)
MAXSRC_avg(2)	4.18	4.18	(50,0) (0,50)

Table 5.1: Results for configuration 1

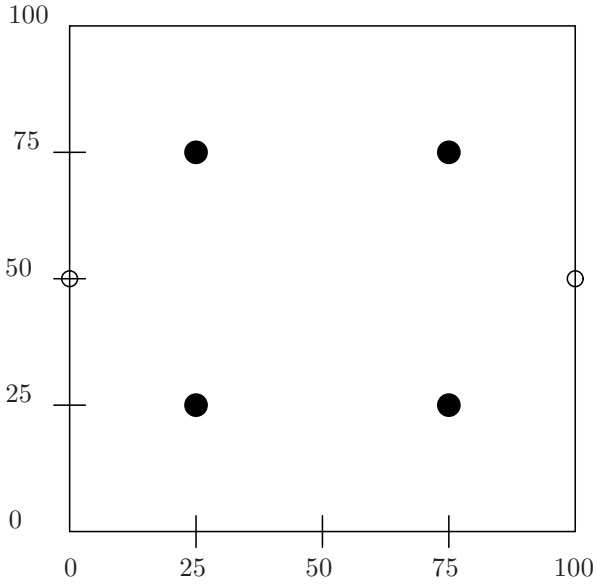

Figure 5.2: Configuration of sensors in Configuration 1 with two additional sensors, represented by open circles, placed at the best positions using MINMAX(2)

5.1.2 Configuration 2

The second configuration of sensors is illustrated in Figure 5.3. As can be seen in the figure, the sensor positions are $p_1 = (25, 25)$, $p_2 = (50, 75)$, $p_3 = (50, 50)$ and $p_4 = (75, 25)$, where it is again assumed that $z = 0$. Using the configuration in Figure 5.3, one or two additional sensors are placed for the various algorithms and the results are presented in Table 5.2. For the single $\vec{e}$ case the best positions for two additional sensors is given in Figure 5.4. The additional sensors are represented as circles. The results of both configurations are discussed in the following subsection.

5.1.3 Discussion

These examples quite clearly indicate the greater robustness of the MINMAX criterion over that of MAXSRC. It interesting to note the discrepancy between configurations for the results of the MINMAX methods and MAXSRC methods. In the first configuration the MINMAX and MAXSRC methods produce similar results. If only a single $\vec{e}$ is considered then when two sensors are placed the MAXSRC method does better, on average. When two $\vec{e}$ vectors are considered the MINMAX

Method	Avg. SD (m)	Min. SD (m)	Sensor Positions
MINMAX(1)	8.32	8.30	(50,100)
MINMAX(2)	9.09	2.39	(100,25) (0,25)
MAXSRC(1)	16.96	16.96	(0,50)
MAXSRC(2)	12.61	12.61	(0,50) (100,50)
MINMAX_both(1)	13.83	13.83	(25,0)
MINMAX_both(2)	3.84	2.92	(0,100) (100,0)
MAXSRC_both(1)	16.96	16.96	(0,50)
MAXSRC_both(2)	12.61	12.61	(0,50) (100,50)
MINMAX_avg(1)	9.98	9.98	(25,0)
MINMAX_avg(2)	3.93	2.86	(100,100) (0,0)
MAXSRC_avg(1)	12.32	12.32	(0,50)
MAXSRC_avg(2)	8.93	8.93	(0,50) (100,50)

Table 5.2: Results for configuration 2

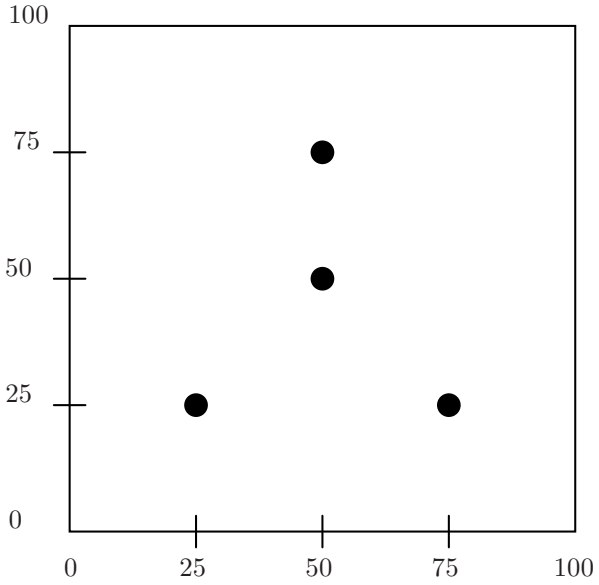

Figure 5.3: Configuration of sensors in Configuration 2

method performs consistently better; for example when the averaging methods are used to place two sensors MINMAX_avg produces an average SD of 3.09 m while MAXSRC_avg results in an average SD of 4.18 m. When considering the second configuration the results between the two methods are much wider with MINMAX consistently performing much better than the MAXSRC methods. Clearly the performance of MAXSRC methods depend upon the initial sensor configuration.

5.2 Lanigan Mine

This example calculates the best sensor positions for an actual mine sensor network in Lanigan Potash Mine, located near Saskatoon, Canada. To obtain meaningful results the average signal–to–noise ratio needs to be estimated. The following subsection outlines the SNR derivation for C block of Lanigan Potash Mine.

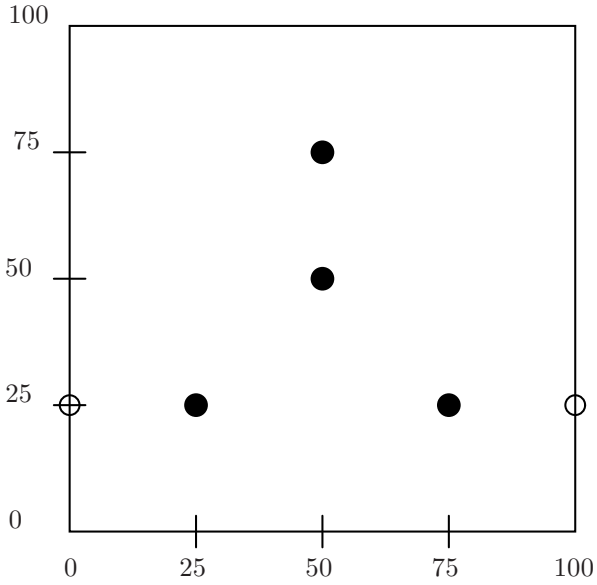

Figure 5.4: Configuration of sensors in Configuration 2 with two additional sensors, represented by open circles, placed at the best positions using MINMAX(2)

5.2.1 SNR Derivation

The derivation of the SNR in terms of the radiated energy of a seismic event is presented in this subsection. The SNR is defined as

$$\text{SNR} = 10 \log \left(\frac{\int_{-2\pi B}^{2\pi B} S_{ss}(\omega) d\omega / d_i^2}{T \int_{-2\pi B}^{2\pi B} S_{nn}(\omega) d\omega} \right), \tag{5.1}$$

where B is the bandwidth, $S_{ss}(\omega)$ is the energy spectral density of a microseismic event, $S_{nn}(\omega)$ is the noise power spectrum, T is the event duration in seconds, and d_i is the average distance between the source and sensors. It should be noted that the SNR defined here is slightly different than that defined in Chapter 3 in equation (3.16). In the previous definition, d_i was defined as the distance from the source to the closest sensor, not the average distance, as it is defined here. This change was made so that not as much emphasis would be placed on the closest sensor. It is especially useful when a source location is very close to a sensor.

The radiated energy of a seismic event $\hat{E}$ as calculated by CANSEIS software [1, 25], is

$$\hat{E} = \frac{4\pi \rho_o c_o C}{M} \sum_{n}^{M} \hat{\epsilon}_n R_n^2. \tag{5.2}$$

That is, it is the weighted average of the sensor signal's energy estimate $\hat{\epsilon}_n$. The weights are the squared distance between the sensor and the estimated source position R_n^2. The other parameters are: $\rho_o = 2100$ kg/m^3, $c_o = 2550$ m/s , $C = 1\text{x}10^{-12}$ m$^2/\mu$m^2 and M is the number of sensors used in the average. The constant C appears since the original data are in units of μm/s. The energy estimate of the signal is computed using 2000 samples about the high point of the signal. It can then be assumed that this weighted sum is an estimate of the total energy of the signal and

$$\int_{-2\pi B}^{2\pi B} S_{ss}(\omega)d\omega \approx \frac{4\pi\rho_o c_o C}{M}\sum_{n}^{M}\hat{\epsilon}_n R_n^2. \tag{5.3}$$

The SNR equation can then be rewritten as

$$SNR \approx 10\log\left(\frac{\hat{E}}{Td_i^2\int_{-2\pi B}^{2\pi B} S_{nn}(\omega)d\omega}\right). \tag{5.4}$$

The noise was estimated as follows. Since the first 800 samples of each record contain only the noise, this portion of the record was used to estimate the average noise power. Initially the noise power was assumed to be approximately white and therefore its PSD would be completely characterized by its sample variance. However, when the *cumulative periodogram test* was applied, as in Section 6.6 of [30], it was found to be an incorrect assumption. Instead, the noise PSD was estimated directly using the first 800 samples of 1901 records. Each PSD was then integrated over the bandwidth $B = 152.5$ Hz, and then these values were averaged to give the average noise estimate. This estimate was then multiplied by constants 4π, c_o, C and ρ_o to obtain the appropriate units.

Since the interest here is the average SNR which occurs in the mine, the energy estimate used was the average value taken over 1901 signals giving an average of 32 J. The parameter d_i^2 was also the average distance between source and sensor and was set equal to $3\text{x}10^6$ m^2. The value of T was set to 2 seconds. Using these values an estimate of the average SNR occuring in Lanigan Potash Mine was -10.4 dB.

5.2.2 Overview

The simulations were run in MATLAB [26] and the implementation of the GA used the Genetic Algorithm Toolbox [48, 49]. An expansion region in Block C of Lanigan Potash Mine is the area considered for sensor placement. A portion of Block C, highlighting the expansion region with diamonds, can be seen in Figure 5.5. The current sensor configuration is taken directly from the mine. The sensor positions are presented in Table 5.3, and are shown in Figure 5.5 with numerical labels. To simplify the simulations the mine and events were assumed to be on the $z = 0$ plane. It is also assumed that velocity is constant at $c = 2581$ m/s. The $\vec{e}$ is set equal to $[1/\sqrt{2},\ 1/\sqrt{2},\ 0]$ for all single $\vec{e}$ simulations and the second $\vec{e}$ was set to $[-1/\sqrt{2},\ 1/\sqrt{2},\ 0]$ for simulations using two $\vec{e}$ vectors. The signal to noise ratio is set to -10.4 dB and the average time duration of an event is set to $T = 0.8$. The derivation for the use of this value for the SNR was presented in the previous section. Using the previously characterized energy spectral density for that portion of the mine [47], the current sensor positions, and the SNR of -10.4 dB, additional sensors can be placed. In these examples Ω_E contains 208 elements and Ω_R contains 841 elements. To deal with the complicated constraints, seen in Figure 5.5, a mapping matrix was used for the MINMAX procedures that used the GA, MINMAX, MINMAX_avg and MINMAX_both. The mapping was a 29x29 matrix and the indexes were the numbers generated by the GA; that is, the GA had an upper and lower bound of 29 and 1, respectively. Each element in the matrix contains two values which correspond to x and y values in Ω_R. This mapping is illustrated in Figure 5.6. The N in the figure is 29 and each element a_{ij} is a coordinate of a point in the actual expansion region (x_{ij}, y_{ij}). Matrix elements near one another must correspond to positions in the mine map that are near each other. This is essential for proper functioning of the GA. The mapping matrix can be thought of as a sampling and rotation of the expansion region. The GA was initially run without this mapping with similar results but taking approximately 4 times longer to converge than when the mapping was used. The mapping when using a GA was adapted from the paper by Bartal et al. [44].

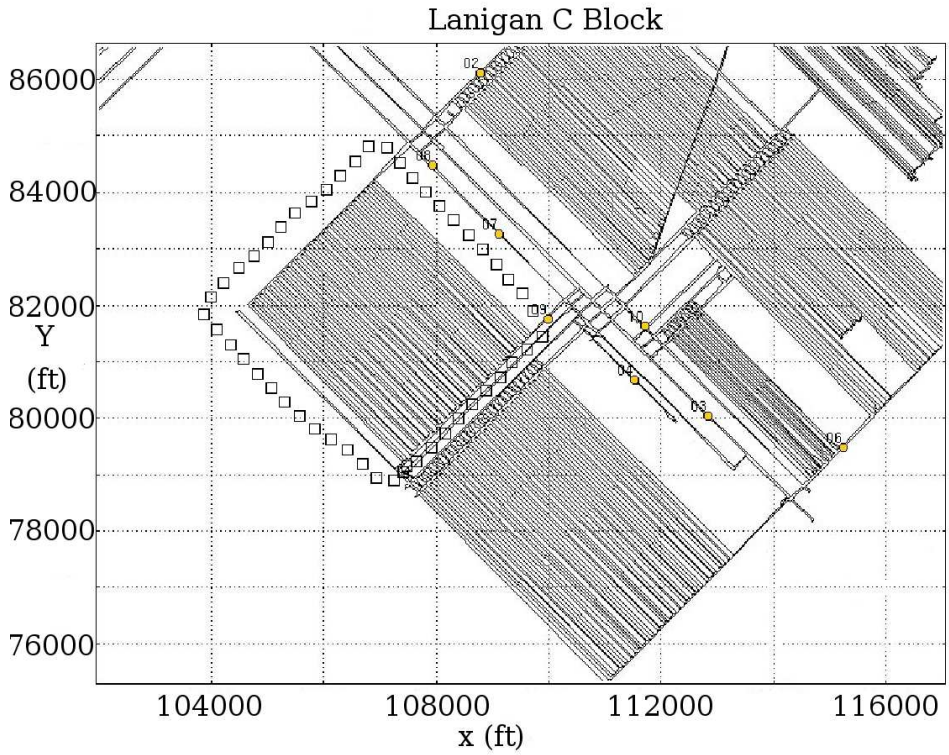

Figure 5.5: A plan view of C Block in Lanigan Potash Mine. The expansion region is outlined by diamonds.

5.2.3 Results

The results list the standard deviations (SD), in meters, after sensors have been placed using the listed algorithm. Before any additional sensors were placed the maximum SD was 6.33 meters. The results are shown in Table 5.4 for the six algorithms placing both one and two additional sensors. The GA was run 30 times to generate these results.

It should be stated, once again, that the values shown in the results are not meant to describe the actual variance observed for the given sensor configuration. Rather, what is important in this example is the relative standard deviation comparing before and after a sensor is placed.

The variance surfaces can be seen in Figures 5.7, 5.8 and 5.9 when zero, one and two additional sensors have been placed, respectively. The sensors were placed using the MAXSRC algorithm. It should be clearly noted in each of the three figures the change in the variance shading scale from one figure to another.

Sensor No.	x Position (ft)	y Position (ft)
1	108776.5	86117.8
2	112819.0	80048.0
3	111507.0	80690.0
4	115231.0	79495.0
5	109111.0	83283.0
6	107915.4	84491.5
7	109984.0	81756.0
8	111700.0	81649.9

Table 5.3: Sensor positions in Lanigan potash mine

Method	Avg. SD (m)	Min. SD (m)	Sensor Position(s) (ft)
MINMAX(1)	0.871	0.832	(106159,80516)
MINMAX(2)	0.674	0.645	(106687,79988) (105420,81254)
MAXSRC(1)	1.195	1.195	(107320,79355)
MAXSRC(2)	0.776	0.776	(104651,81989)
MINMAX_both(1)	1.191	1.137	(106370,80304)
MINMAX_both(2)	0.794	0.692	(105696,83034) (106475,80199)
MAXSRC_both(1)	1.239	1.239	(107320,79355)
MAXSRC_both(2)	0.959	0.959	(104812,82150)
MINMAX_avg(1)	0.936	0.899	(105948,80727)
MINMAX_avg(2)	0.664	0.612	(106476,80199) (105214,82552)
MAXSRC_avg(1)	1.223	1.223	(107320,79355)
MAXSRC_avg(2)	0.669	0.669	(104651,81989)

Table 5.4: Results for Lanigan Mine example

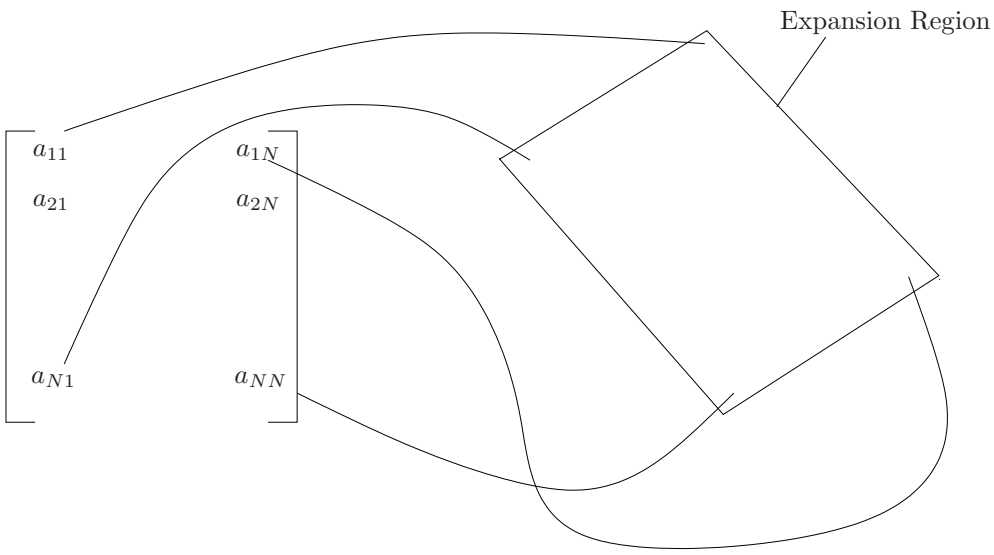

Figure 5.6: Mapping illustration

The improvement by adding a single sensor is significant; the SD goes from 6.33 meters to 120 cm for MAXSRC or 87 cm for MINMAX. Clearly the MINMAX method provides better results. However, the difference between 120 cm or 87 cm is not great. The discrepancy is even smaller when two sensors are placed with the MINMAX producing an average SD of 67 cm compared to MAXSRC's 78 cm. Similar results are seen when two $\vec{e}$ vectors are used with the notable exception of the MINMAX_avg and MAXSRC_avg algorithms which, when two sensors are placed produce results which are only approximately half a centimetre different on average between the MINMAX_avg and MAXSRC_avg algorithms. In all cases, however, the MINMAX algorithm outperforms the MAXSRC algorithm.

To iteratively place two sensors using the MAXSRC algorithm takes approximately 3.5 seconds, using MATLAB [26]. By contrast, to place a single sensor using the MINMAX criterion with the GA takes significantly more time. It takes, on average, 13 generations for the GA to converge and approximately 27 seconds per generation, giving approximately 6 minutes on average for the algorithm to find a solution. To place two sensors at a time this corresponds to 8 minutes, on average. The average times for the GA based MINMAX algorithms to converge are summarized in Table 5.5.

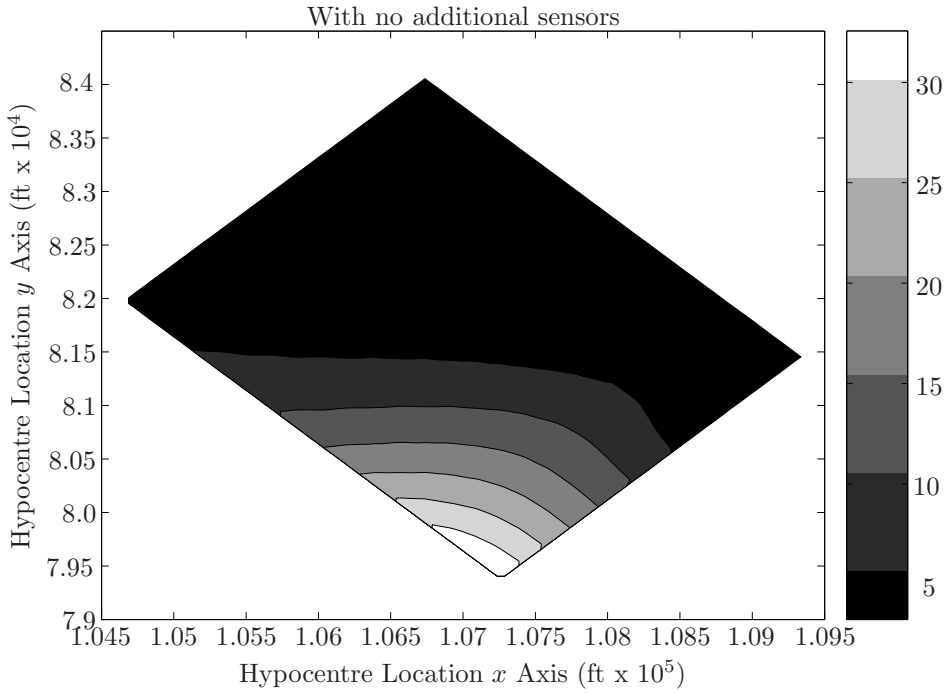

Figure 5.7: The filled contour plot of the variance for source postions in the expansion region when no additional sensors have been placed.

All of these running times are well within the constraints, which may be present at a mine. Additional sensors are usually added only a few times a year, depending on the rate of expansion.

5.3 Summary and Discussion

The first section of the chapter was devoted to two simple synthetic examples which assumed that the current sensors were inside the expansion region. Both configurations illustrated how consistently the MINMAX algorithm outperformed the MAXSRC algorithm, in terms of maximum SD. The second configuration showed this fact in exemplary fashion.

The second section presented an example of sensor placement in a Lanigan Mine expansion region. The average SNR was derived using existing mine data. Because of the complicated constraints imposed by the mine rooms a mapping matrix was created, and used when the GA was applied, which improved the speed of convergence of the GA algorithm. The results showed, once again, the superiority, in terms of maximum SD, which the MINMAX class of algorithms have over

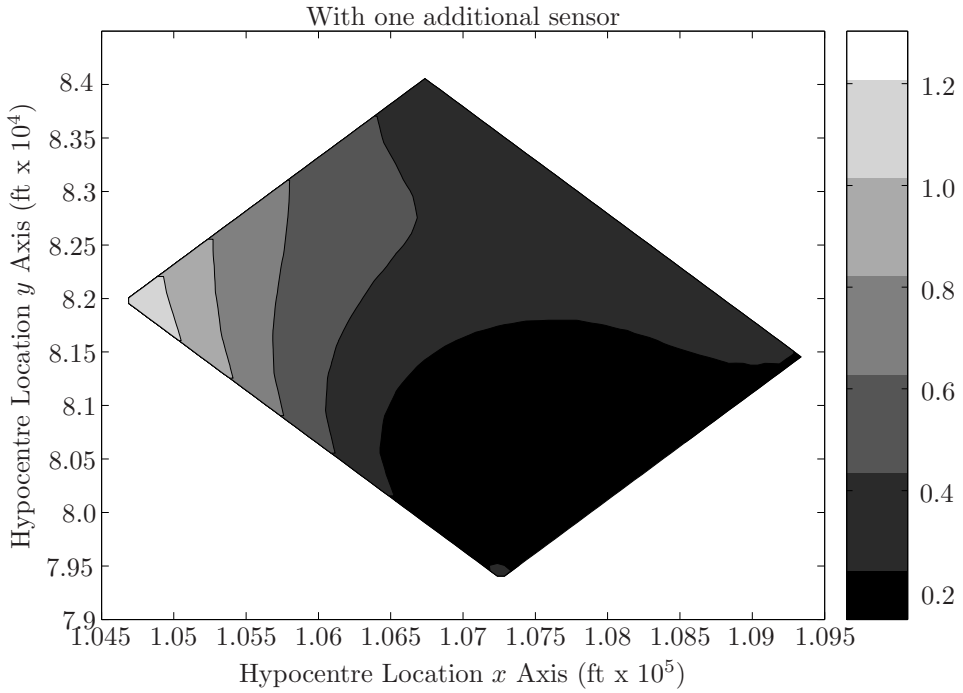

Figure 5.8: The filled contour plot of the variance for source postions in the expansion region when one additional sensor has been placed.

the MAXSRC class.

Despite the better results enjoyed by the MINMAX class of algorithms, the MAXSRC class is intuitively very simple and over 100 times faster at finding a solution. It is clear from the above three examples that the MINMAX class of algorithms, utilizing the GA, are more robust than the MAXSRC class. If, however, there are time or processing constraints then the MAXSRC class of algorithms provide acceptable solutions in approximately 1/100th the time.

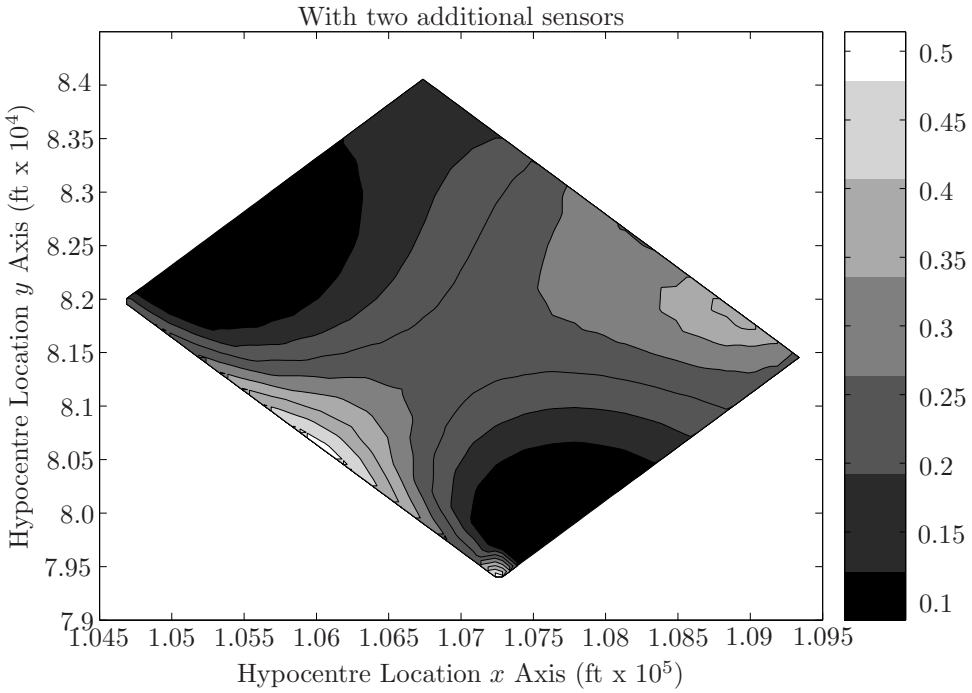

Figure 5.9: The filled contour plot of the variance for source postions in the expansion region when two additional sensors have been placed.

Method	Avg. Time (min)
MINMAX(1)	5.85
MINMAX(2)	8.10
MINAMX_both(1)	9.43
MINMAX_both(2)	11.47
MINMAX_avg(1)	8.73
MINMAX_avg(2)	11.43

Table 5.5: Average time for algorithm to converge

Chapter 6

Conclusions and Future Work

6.1 Summary

In Chapter 2, two classes of localization algorithms were discussed, the two–step and single–step procedures. In addition, a specific estimator from each class was discussed in detail. For the two–step class of algorithms it was Geiger's method which solved the set of nonlinear equations using a nonlinear least squares technique. The error in this estimator was presented in the form of a covariance matrix of the unknown hypocentre location and origin time. In the case of the single–step, the Weighted Total Signal Energy estimator was presented. Its variance of localization error was presented which was seen to partly depend upon the signal's energy spectrum. The WTSE shows promise as a robust localization algorithm that performs well with finite–duration events that may have a small time–bandwidth product. It was used as the location estimator throughout this thesis.

Chapter 3 presented a classification scheme of signals received in Potash Mines in Saskatchewan. The scheme separated the signals into a successive event type and an impulse type signal. Energy Spectral Densities (ESD) of each type of signal were estimated using the multitaper and Combined Time and Lag Weighting (CTLW) methods for two specific potash mines, Lanigan and Allen mines. The characterized ESDs of each mine of the successive event type signal were used to evaluate the variance of error expression for various bandwidth–time products showing little defference between estimation methods and a small, but possibly significant, difference between Allen and Lanigan mine.

Chapter 4 presented two classes of sensor placement algorithms for sensors in mine expansion

regions. An overview of the previous literature in seismic network configuration was given and then the differences between this and previous work were highlighted. It was also shown how each of the classes of algorithms could be modified to include multiple $\vec{e}$ directions in two different ways.

The first section of Chapter 5 was devoted to two simple synthetic examples that assumed the current sensors were inside the expansion region. Both configurations illustrated how consistently the MINMAX algorithm outperformed the MAXSRC algorithm, in terms of minimizing the maximum standard deviation (SD). The second configuration showed this fact in exemplary fashion.

The second section, of Chapter 5, presented an example of sensor placement in a Lanigan Mine expansion region. The average SNR was derived using existing mine data. Because of the complicated constraints imposed by the mine rooms, a mapping matrix was created and used when the Genetic Algorithm (GA) was applied, which improved the speed of convergence of the GA. The results showed, once again, the superiority, in terms of maximum SD, that the MINMAX class of algorithms has over the MAXSRC class.

Despite the better results enjoyed by the MINMAX class of algorithms the MAXSRC class is intuitively very simple and over 100 times faster at finding a solution. It is clear from the above three examples that the MINMAX class of algorithms, utilizing the GA, are more robust than the MAXSRC class. If, however, there are time or processing constraints then the MAXSRC class of algorithms provide acceptable solutions in approximately 1/100th the time.

6.2 Conclusions

There are two main areas which conclusions will be drawn from: the ESD characterization and the sensor placement algorithms. First, conclusions from the ESD characterization.

- The Combined Time and Lag Weighting method of spectrum estimation should be used if there are time constraints. If there are no time constraints then it was seen that both methods produced very similar results and so the choice is left to the user.

- There was noticeable difference between the ESDs of the two mines, though this difference

was not great. It is therefore advised that, in the future, an analysis and characterization of each particular mine is performed to obtain a specific ESD for that mine.

- The differences between the two mine ESDs was seen to also affect the variance calculation. This difference was seen to be larger at higher time bandwidth products.

Conclusions relating to the sensor placement algorithms are summarized below.

- The MINMAX algorithm using a Genetic Algorithm has been shown to consistently produce superior results to those of the MAXSRC algorithm. Though, it takes approximately 100 times as long to run as the MAXSRC algorithm, the execution time on a modern PC is very acceptable, being in the neighbourhood of 5.85 or 8.10 minutes, depending on whether one or two sensors are placed, respectively.

- The MINMAX algorithm is a robust sensor placement procedure which should be used instead of the MAXSRC algorithm in all situations unless relevant time constraints are present.

- The MAXSRC algorithm provides a fast and intuitively simple procedure to place additional sensors in mine expansion regions. It consistently performs the same or worse than the MINMAX algorithm and therefore should only be used when time or processing constraints are present.

6.3 Future Work

There are three main areas dealt with in this thesis that suggest future avenues of investigation: the energy spectral density, the placement algorithms and the variance expression and underlying location estimate. The following paragraphs will discuss each of these individually.

There are two areas which need work in terms of the ESD of potash mines and both are due to the need for more data. To adequately characterize the impulse type waveform requires a considerable amount of data, probably over two years from a particular mine. Using this amount of data should allow an adequate characterization of the impulse type waveforms at both mines. The second area

is the need for data from another potash mine that is not in the same geographical area and most likely a different configuration. This will allow a better ESD comparison between mines.

There are a number of areas of future work related to the sensor placement algorithms. A very useful, in terms of practicality in mining situations, is for a set of general 'rules' for placing sensors in expansion regions to be developed. This development would likely involve using both algorithms to place sensors in a large number of different examples. Hopefully from examining these examples a set of rules could be determined.

When the variance expression was derived it was assumed that the noise at each sensor was independent. In previous work, in sensor placement by Rabinowtiz and Steinberg [41] it was pointed out that this assumption has caused two sensors to be placed very near to each other. This violates the independent noise assumption. The examples presented in this thesis did not show any predeliction of the criterions to place two sensors very close to one another. This, however, is not conclusive and there needs to be further examples and analysis undertaken to determine whether this problem will arise with the algorithms presented in this thesis.

References

[1] Slawomir J. Gibowicz and Andrzej Kijko, *An Introduction to Mining Seismology*, Academic Press, 1994.

[2] Henry S. Hasegawa, Robert J. Wetmiller, and Don J. Gendzwill, "Induced seismicity in mines in canada-an overview," *Pure Appl. Geophys.*, vol. 129, no. 3–4, pp. 423 – 453, September 1989.

[3] Arnfinn Prugger and D. Gendzwill, "Microseismic monitoring in a mine, what is it and how is it useful?," Presented at CIM Meeting in Saskatoon, Canada, 1997.

[4] W. H. K. Lee and S. W. Stewart, *Principles and Applications of Microearthquake Networks*, Academic Press, 1981.

[5] Y. Huang, J. Benesty, and G. W. Elko, "Passive acoustic source localization for video camera steering," in *Proc. ICASSP 00*, 2000, vol. 2, pp. 902–912.

[6] J. C. Hassab, *underwater signal and data proceesing*, CRC Press, 1989.

[7] David P. Schaff, Gotz H. R. Bokelmann, William L. Ellsworth, Eva Zanzerkia, Felix Waldhauser, and Gregory C. Beroza, "Optimizing correlation techniques for improved earthquake location," *Bull. Seism. Soc. Am.*, vol. 94, no. 2, pp. 705–721, April 2004.

[8] Haijiang Zhang, Clifford Thurber, and Charlotte Rowe, "Automatic p–wave arrival detection and picking with multiscale wavelet analysis for single–component recordings," *Bull. Seism. Soc. Am.*, vol. 93, no. 5, pp. 1904–1912, October 2003.

[9] Brian L. F. Daku, Peter Kosteniuk, and Arnfinn F. Prugger, "Refining visually estimated arrival times of short duration signals," *Measurement*, vol. 30, pp. 297–305, 2001.

[10] Julius O. Smith and Jonathan S. Abel, "The spherical interpolation method of source localization," *IEEE Journal of Oceanic Engineering*, vol. 12, no. 1, pp. 246–252, January 1987.

[11] Yiteng Huang, Jacob Benesty, Gary W. Elko, and Russell M. Mersereau, "Real–time passive source localization: a practical linear–correction least–squares approach," *IEEE Trans. on Speech and Audio Processing*, vol. 9, no. 8, pp. 943–956, November 2001.

[12] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. on Signal Processing*, vol. 42, no. 8, pp. 1905–1915, August 1994.

[13] H. C. Schau and A. Z. Robinson, "Passive source localization employing intersecting spherical surfaces from time–of–arrival differences," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP–35, no. 8, pp. 1223–1225, August 1987.

[14] Barry R. Lienert, E Berg, and L. Neil Frazer, "Hypocenter: An earthquake location method using centered, scaled and adaptively damped least squares," *Bull. Seism. Soc. Am.*, vol. 76, no. 3, pp. 771–783, June 1986.

[15] L. Geiger, "Probability method for the determination of earthquake epicenters from the arrival time only," *Bull. St. Louis Univ.*, vol. 8, pp. 60–71, 1912.

[16] Arnfinn Prugger and D. Gendzwill, "Microearthquake location: a nonlinear approach that makes use of a simplex stepping procedure," *Bull. Seism. Soc. Am.*, vol. 78, pp. 799–815, April 1988.

[17] N. Rabinowitz, "Microearthquake location by means of nonlinear simplex procedure," *Bull. Seism. Soc. Am.*, vol. 78, pp. 380–384, 1988.

[18] A.J. Mendecki, Ed., *Seismic Monitoring in Mines*, Chapman and Hall, 1997.

[19] Harold W. Sorenson, *Parameter Estimation: Principles and Problems*, Marcel Dekker, Inc., New York, NY, 1980.

[20] Brian L.F Daku and J. Eric Salt, "Error analysis of a localization algorithm for finite–duration events," *IEEE Trans. on Signal Processing*, To Appear.

[21] Brian L.F. Daku, J. Eric Salt, Li Sha, and Arnfinn Prugger, "An algorithm for locating microseismic events," in *Proc. IEEE CCECE*, 2005, pp. 2311–2314.

[22] Stephen W. Davies and Mariner A. Price, "Source localization by summing multiple correlator outputs," in *Proc. ICASSP 90*, 1990, vol. 5, pp. 2787–2790.

[23] N. Strobel and R. Rabenstein, "Classification of time delay estimates for robust speaker localization," in *Proc. ICASSP 99*, 1999, vol. 6, pp. 3081–3084.

[24] Joe C. Chen, Ralph E. Hudson, and Kung Yao, "Maximum–likelihood source localization and unkown sensor location estimation for wideband signals in the near–field," *IEEE Trans. on Signal Processing*, vol. 50, no. 8, pp. 1843–1854, August 2002.

[25] Peter Kosteniuk and Arnfinn Prugger, "Automated microseismic event analysis," Presented at CIM Meeting in Calgary, Canada, 1999.

[26] *MATLAB*, The Mathworks Inc., Natick, Massachusetts, version 7.0 (r14) edition, 2004.

[27] David J. Thomson, "Spectrum estimation and harmonic analysis," *Proc. IEEE*, vol. 70, no. 9, Sept. 1982.

[28] Craig R. Lindberg and Jeffrey P. Park, "Multiple–taper spectral analysis of terrestrial free oscillations: part ii," *Geophys. J. R. astr. Soc.*, vol. 91, pp. 795–836, 1987.

[29] Jeffrey P. Park, Craig R. Lindberg, and David J. Thomson, "Multiple–taper spectral analysis of terrestrial free oscillations: part i," *Geophys. J. R. astr. Soc.*, vol. 91, pp. 755–794, 1987.

[30] Donald B. Percival and Andrew T. Walden, *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques*, Cambridge University Press, 1993.

[31] Albert H. Nuttall and G. Clifford Carter, "Spectral estimation using combined time and lag weighting," *Proc. IEEE*, vol. 70, no. 9, pp. 1115–1125, Sept. 1982.

[32] Albert H. Nuttall, "Some windows with very good sidelobe behavior," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP–29, no. 1, pp. 84–91, Feb. 1981.

[33] Andrew T. Walden, E. McCoy, and Donald B. Percival, "The variance of multitaper spectrum estimates for real gaussian processes," *IEEE Trans. on Signal Processing*, vol. 42, no. 2, Feb. 1994.

[34] Thomos P. Bronez, "On the performance advantage of multitaper spectral analysis," *IEEE Trans. on Signal Processing*, vol. 40, no. 12, Dec. 1992.

[35] Y. Sato, "Optimum distribution of seismic observation points," *Seismological Society of Japan*, vol. 18, pp. 9–14, 1965.

[36] Y. Sato and D. Skoko, "Optimum distribution of seismic observation points, ii," *Bull. Earthq. Res. Inst. Tokyo Univ.*, vol. 43, pp. 451–457, 1965.

[37] D. Skoko and Y. Sato, "Optimum distribution of seismic observation points, iii," *Bull. Earthq. Res. Inst. Tokyo Univ.*, vol. 44, pp. 13–22, 1966.

[38] Andrzej Kijko, "An algorithm for the optimum distribution of a regional seismic network–i," *Pure Appl. Geophys.*, vol. 115, pp. 999–1009, 1977.

[39] Andrzej Kijko, "An algorithm for the optimum distribution of a regional seismic network–ii. an analysis of local earthquakes depending on the number of seismic stations," *Pure Appl. Geophys.*, vol. 115, pp. 1011–1021, 1977.

[40] V. Y. Burmin, "Optimal placement of seismic stations for registration of near earthquakes," *Izv. Akad. Nauk SSSR, Earth Phys.*, vol. 22, no. 5, pp. 366–372, 1986.

[41] Nitzan Rabinowitz and David M. Steinberg, "Optimal configuration of a seismographic network: A statistical approach," *Bull. Seism. Soc. Am.*, vol. 80, no. 1, pp. 187–196, Feb. 1990.

[42] David M. Steinberg, Nitzan Rabinowitz, Yair Shimshoni, and Daphna Mizrachi, "Configuring a seismographic network for optimal monitoring of fault lines and multiple sources," *Bull. Seism. Soc. Am.*, vol. 85, no. 6, pp. 1847–1857, Dec. 1995.

[43] A. Kijko and M. Sciocatti, "Optimal spatial distribution of seismic stations in mines," *Int. J. Rock Mech. and Min. Sci.*, vol. 32, no. 6, pp. 607–615, 1995.

[44] Yair Bartal, Zeev Somer, Gideon Leonard, David M. Steinberg, and Yochai B. Horin, "Optimal seismic networks in israel in the context of the comprehensive test ban treaty," *Bull. Seism. Soc. Am.*, vol. 90, no. 1, pp. 151–165, Feb. 2000.

[45] Nitzan Rabinowitz and David M. Steinberg, *Advances in Seismic Event Location*, chapter 3, Kluwer Academic Publishers, 2000.

[46] David M. Steinberg and Nitzan Rabinowitz, "Optimal seismic monitoring for event location with application to on site inspection of the comprehensive nuclear test ban treaty," *Metrika*, vol. 58, pp. 31–57, 2003.

[47] Angus Errington, Brian L.F. Daku, David Dodds, and Arnfinn Prugger, "Characterization of the energy spectral density for a potash mine," in *Proc. IEEE CCECE*, 2005, pp. 0247–0250.

[48] A. J. Chipperfield, P. J. Flemming, and C. M. Fonesca, "Genetic algorithm tools for control systems engineering," presented at Adaptive Computing in Engineering Design and Control, Plymouth, UK, September 1994.

[49] A. J. Chipperfield and P. J. Flemming, "The matlab genetic algorithm toolbox," in *from IEE Colloqium on Applied Control Techniques Using MATLAB, Digest No. 1995/014*, January 1995.

Appendix A

Evaluation of the Variance of Error: Sensor Positions at or Near Source Location

When evaluating the expression given in Daku and Salt's paper [20] for sensor postions close to or at the origin (the source location) it may be unclear how it is to be evaluated. This is also the case for the weighted variance expression used in this thesis. For this appendix, however, the analysis will be performed for the unweighted expression, but is equally applicable to the weighted expression. The purpose of this appendix is to clarify the model in such a way that when the variance expression is evaluated for sensor positions at or near zero it produces meaningful results. The first section discusses clarifications to the model which need to be made and the second section explores an assumption in the derivation which, it will be shown, does not need to be made. The meaning of the previous statement will become clear in the respective section.

A.1 Model Clarification

The received signal model given in the paper [20] at the mth sensor is

$$r_m(t) = \frac{s(t-\tau_m)}{d_m} + n_m(t-\tau_m), \tag{A.1}$$

where $n_m(t)$ is a sample function from an independent, stationary stochastic process, τ_m is the time it takes for the signal to propagate from the source to sensor m and is given by

$$\tau_m = \frac{d_m}{c}, \tag{A.2}$$

where c is the speed of propagation and d_m is the distance between the source and sensor m. The reference signal, $s(t-\tau_o)$, for the purposes of the theoretical model is defined to be a signal received at a noiseless sensor one meter from the source in the direction $\vec{e}$, where τ_o is the time it takes for the signal to propagate one meter. The question that one may ask is, what about sensors which are closer than one meter to the source position?

Though it may be unclear from the paper [20], and as described above, the model does, in fact, make allowances for sensors closer than one meter. The remainder of this section attempts to explain and clarify the model for sensors which are closer than one meter to the source location. To do this it is necessary to separate the attenuation and the time delay of the received sensor signals, for sensors closer that one meter. Implied by the model, though not stated explicitly, is that a signal received at sensors within one meter of the source location will have the same attentuation as that received at one meter; the time delay, however, will be consistent with the distance travelled.

Once the above is clear, how then should one evaluate the final expression for the variance? What terms are related to the attenuation and what terms are related to the time delay? The answer, fortunately, is quite simple. The d_m values relate to attenuation and the $\frac{\vec{p}_m \cdot \vec{e}}{|\vec{p}_m|}$ term relates to time delay. This amounts to setting $d_m = 1$ if the mth sensor is closer than one meter to the event. It must be stated clearly that the distance between the sensor and event remains the same but the attenuation is changed. It so happens that the attenuation equals $1/d_m$.

A further clarification is needed when the mth sensor is located at the same position as the event $\vec{p}_m = (0,0,0)$

$$\tilde{d}_m = |\vec{p}_m - \tilde{p}_s| = |\tilde{p}_s| = |\tilde{\epsilon}\vec{e}| = \tilde{\epsilon}. \tag{A.3}$$

Then,

$$\Delta\tilde{\tau}_m = \tilde{\tau}_m - \tau_m = \frac{\tilde{\epsilon}}{c}, \tag{A.4}$$

where $\tau_m = 0$. When this result is carried through the derivation it is seen that values of q_m which are normally equal to $\frac{\vec{p}_m \cdot \vec{e}}{|\vec{p}_m|}$ are instead equal to unity, for the case when $\vec{p}_m = (0,0,0)$.

A.2 Derivation Assumption

The assumption which causes problems is $|\tilde{p}_s| \ll |\vec{p}_m|$ which leads to the approximation that

$$\tilde{d}_m = |\vec{p}_m - \tilde{p}_s| \approx |\vec{p}_m| - \frac{\vec{p}_m \cdot \tilde{p}_s}{|\vec{p}_m|}. \tag{A.5}$$

Clearly when a sensor is within a metre of, or at, the source location this assumption is not valid. This section re–derives the variance expression without making the above assumption. It is assumed throughout this section that the paper by Daku and Salt [20] is being followed along as this section progresses. The notation used here is also the same as that in [20].

The non–trivial case where $\Delta\tilde{\tau}_m$ becomes

$$\Delta\tilde{\tau}_m = \frac{1}{c}(|\vec{p}_m - \tilde{p}_s| - |\vec{p}_m|) \tag{A.6}$$

where

$$|\vec{p}_m - \tilde{p}_s| = \sqrt{(x_m - \tilde{p}_s \cdot \hat{\mathbf{i}})^2 + (y_m - \tilde{p}_s \cdot \hat{\mathbf{j}})^2 + (z_m - \tilde{p}_s \cdot \hat{\mathbf{k}})^2}, \tag{A.7}$$

where x_m, y_m, z_m are the x,y and z coordinates of sensor m and $\hat{\mathbf{i}}$,$\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ are the three unit vectors which 'pick out' the x, y and z coordinates of $\tilde{p}_s$.

Using (A.6) in the derivation results in the same final expression as derived in [20]. For illustration purposes Appendix I of [20] is rederived using (A.6). Begin with

$$W_s(\tilde{\epsilon}\vec{e}) = \sum_{m=1}^{M}\sum_{k=1}^{M}\int_0^T \frac{1}{d_m d_k} s(t + \Delta\tilde{\tau}_m)s(t + \Delta\tilde{\tau}_k)dt. \tag{A.8}$$

As shown in Appendix I the above equation can be rewritten as

$$W_s(\tilde{\epsilon}\vec{e}) = \sum_{m=1}^{M}\sum_{k=1}^{M}\frac{1}{d_m d_k 2\pi}\int_{-\infty}^{\infty}\exp(-j\frac{1}{c}(|\vec{p}_m - \tilde{p}_s| - |\vec{p}_m| - |p_k - \tilde{p}_s| + |p_k|)\lambda)S(\lambda)S(-\lambda)d\lambda. \tag{A.9}$$

It is required to evaluate the second derivative w.r.t $\tilde{\epsilon}$ and evaluate it at $\tilde{\epsilon} = 0$. Since we are taking the derivative of an exponential the exponent will stay the same. Also note that when the exponent is evaluated at $\tilde{\epsilon} = 0$ it goes to zero. To see what happens when the derivative is taken it is useful to compare the results when the two $\Delta\tilde{\tau}_m$'s are compared. With assumption, we have

$$\frac{d\Delta\tilde{\tau}_m}{d\tilde{\epsilon}} = -\frac{\vec{p}_m \cdot \vec{e}}{|\vec{p}_m|c} \tag{A.10}$$

and then evaluate when $\tilde{\epsilon} = 0$

$$\left.\frac{d\Delta\tilde{\tau}_m}{d\tilde{\epsilon}}\right|_{\tilde{\epsilon}=0} = -\frac{\vec{p}_m \cdot \vec{e}}{|\vec{p}_m|c}, \tag{A.11}$$

which is the same. The same can be done when the assumption is not made, in that we are taking the derivative of

$$|\vec{p}_m - \tilde{p}_s| = \sqrt{(x_m - \tilde{p}_s \cdot \hat{\mathbf{i}})^2 + (y_m - \tilde{p}_s \cdot \hat{\mathbf{j}})^2 + (z_m - \tilde{p}_s \cdot \hat{\mathbf{k}})^2}, \tag{A.12}$$

which is

$$\frac{d\Delta\tilde{\tau}_m}{d\tilde{\epsilon}} = \frac{1}{2c|\vec{p}_m|}(2(x_m - \tilde{p}_s \cdot \hat{\mathbf{i}})(-\vec{e} \cdot \hat{\mathbf{i}}) + 2(y_m - \tilde{p}_s \cdot \hat{\mathbf{j}})(-\vec{e} \cdot \hat{\mathbf{j}}) + 2(z_m - \tilde{p}_s \cdot \hat{\mathbf{k}})(-\vec{e} \cdot \hat{\mathbf{k}})) \tag{A.13}$$

and then when the above is evaluated at $\tilde{\epsilon} = 0$

$$\left.\frac{d\Delta\tilde{\tau}_m}{d\tilde{\epsilon}}\right|_{\tilde{\epsilon}=0} = -2\frac{x_m\vec{e} \cdot \hat{\mathbf{i}} + y_m\vec{e} \cdot \hat{\mathbf{j}} + z_m\vec{e} \cdot \hat{\mathbf{k}}}{2c|\vec{p}_m|}. \tag{A.14}$$

Noting that the numerator is equivalent to $\vec{p}_m \cdot \vec{e}$ and cancelling the twos yields

$$\left.\frac{d\Delta\tilde{\tau}_m}{d\tilde{\epsilon}}\right|_{\tilde{\epsilon}=0} = -\frac{\vec{p}_m \cdot \vec{e}}{|\vec{p}_m|c} \tag{A.15}$$

Case	For	Action
1	$\lvert\vec{p}_m\rvert > 1$	No change
2	$0 < \lvert\vec{p}_m\rvert < 1$	set $d_m = 1$
3	$\lvert\vec{p}_m\rvert = 0$	set $d_m = 1$ and set $\frac{\vec{p}_m \cdot \vec{e}}{\lvert p_m\rvert} = 1$

Table A.1: Summary of expression evaluation

which is the same as obtained for the simplifying assumption. These results prove useful when evaluating the second derivative of $W_s(\tilde{\epsilon}\vec{e})$ w.r.t $\tilde{\epsilon}$ and evaluated at zero, to give

$$\left.\frac{d^2 W_s(\tilde{\epsilon}\vec{e})}{d\tilde{\epsilon}^2}\right|_{\tilde{\epsilon}=0} = -\frac{1}{2\pi c^2}\int_{-\infty}^{\infty}\omega^2 S_{ss}(\omega)d\omega \sum_{m=0}^{M}\sum_{k=0}^{M}\frac{1}{d_m d_k}\left[\frac{\vec{p}_m\cdot\vec{e}}{|\vec{p}_m|} - \frac{p_k\cdot\vec{e}}{|p_k|}\right]^2 \tag{A.16}$$

which is the same as when the simplifying assumption is used. It can be shown by using the same procedure on the remaining Appendices that each is equivalent whether the simplifying assumption is made or not. It can, therefore, be said that even though the variance expression derived in [20] was done with the assumption that $|\tilde{p}_s| \ll |\vec{p}_m|$ this assumption is not needed to arrive at the same expression, as given in [20].

The following section summarizes the results of this Appendix.

A.3 Summary

The clarifications helpful when evaluating the variance expression are summarized in Table A.3. Using this table the variance expression can now safely be evaluated when sensors are at or near the source location. Once again, it should be noted that the attenuation is being set to one in steps 2 and 3, not the distance between the source and sensor m. This distinction is very important as $|p_m| = d_m$ and clearly we do not want to set $|p_m| = 1$.