2014

Direct And Evolutionary Approaches For Optimal Receiver Function Inversion

Mulugeta Tuji Dugda
North Carolina Agricultural and Technical State University

Direct and Evolutionary Approaches for Optimal Receiver Function Inversion

Mulugeta Tuji Dugda

North Carolina A&T State University

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department: Electrical and Computer Engineering

Major: Electrical Engineering

Major Professor: Dr. Jung H. Kim

Co-advisor Professor: Dr. Abdollah Homaifar

Greensboro, North Carolina

2014

The Graduate School
North Carolina Agricultural and Technical State University
This is to certify that the Doctoral Dissertation of

Mulugeta Tuji Dugda

has met the dissertation requirements of
North Carolina Agricultural and Technical State University

Greensboro, North Carolina
2014

Approved by:

Dr. Jung Hyoun Kim
Major Professor

Dr. Abdollah Homaifar
Academic Co-Advisor

Dr. John Kelly
Committee Member

Dr. Robert Y. Li
Committee Member

Dr. John Kelly
Department Chair

Dr. Jung Hee Kim
Committee Member

Dr. Sanjiv Sarin
Dean, The Graduate School

Biographical Sketch

Mulugeta Tuji Dugda was born in Butajira, South Shoa, Ethiopia. He received his bachelor's degree with distinction and his master's degree with the highest honor in Electrical Engineering from Addis Ababa University (Ethiopia) in 1990 and 1998, respectively. He later received a master's and PhD degrees in Geosciences from Pennsylvania State University in 2003 and 2007, respectively. Mulugeta joined the PhD program of Electrical and Computer Engineering Department at North Carolina Agricultural & Technical State University in the Fall of 2009.

Acknowledgements

My first and foremost gratitude goes to the almighty God who gave me the PhD opportunity and the people to help as well as the health and strength to do my PhD.

I am very thankful to my advisor Dr. Jung Hyoun Kim and my co-advisor Dr. Abdollah Homaifar for their invaluable advice, unwavering support and encouragement. The professional and technical guidance of both of my advisors has helped shape this research work and ended up in great success. I would like to express my deepest gratitude to my doctoral committee members Dr. John Kelly, Dr. Robert Li and Dr. Jung Hee Kim for their advice, critics and important suggestions at different stages of this research work as well as for their editing of this dissertation. My thanks also go to the graduate school representative Dr. Cameron Seay for his helpful suggestions and critics.

I have no words to express my appreciations to my family, especially my wife Aster Habtewold and my daughter Rebekah Mulugeta. They have endured a lot of hardship and burden throughout my study periods at NCA&T and at Penn State. I would like to say thank you to Astuka and Beckye for understanding and paying a price for the long periods of my studies.

I am extremely grateful to my friends at NCA&T and Greensboro, especially Dr. Abebe Kebede. I am deeply indebted to Dr. Abebe Kebede who has been a great help to me and my family throughout this journey of the PhD study. Dr. Ademe Mekonnen has also been very helpful during my study and I would like to say thank you. I would like to express my heartfelt appreciation to all friends in the two labs where I have been working.

Thank God for all the brothers and sisters He has prepared for me and my family when we moved to Greensboro, NC. I am very grateful to all the brothers and sisters at the Ethiopian Christian Fellowship Church at Greensboro for their fervent prayers and tremendous support.

Table of Contents

List of Figures

List of Tables

Abstract

Receiver functions are time series obtained by deconvolving vertical component seismograms from radial component seismograms. Receiver functions represent the impulse response of the earth structure beneath a seismic station. Generally, receiver functions consist of a number of seismic phases related to discontinuities in the crust and upper mantle. The relative arrival times of these phases are correlated with the locations of discontinuities as well as the media of seismic wave propagation. The Moho (Mohorovicic discontinuity) is a major interface or discontinuity that separates the crust and the mantle. In this research, automatic techniques to determine the depth of the Moho from the earth's surface (the crustal thickness H) and the ratio of crustal seismic P-wave velocity (Vp) to S-wave velocity (Vs) (κ= Vp/Vs) were developed.

In this dissertation, an optimization problem of inverting receiver functions has been developed to determine crustal parameters and the three associated weights using evolutionary and direct optimization techniques. The first technique developed makes use of the evolutionary Genetic Algorithms (GA) optimization technique. The second technique developed combines the direct Generalized Pattern Search (GPS) and evolutionary Fitness Proportionate Niching (FPN) techniques by employing their strengths. In a previous study, Monte Carlo technique has been utilized for determining variable weights in the H-κ stacking of receiver functions. Compared to that previously introduced variable weights approach, the current GA and GPS-FPN techniques have tremendous advantages of saving time and these new techniques are suitable for automatic and simultaneous determination of crustal parameters and appropriate weights.

The GA implementation provides optimal or near optimal weights necessary in stacking receiver functions as well as optimal H and κ values simultaneously. Generally, the objective function of the H-κ stacking problem displays multimodal surfaces with multiple local and global optima.

Niching mechanism permits standard GAs to identify different subpopulations representing various peaks. In multimodal optimization, fitness sharing has been commonly used to generate stable subpopulations of individuals around multiple optimum points in the search space. In this study the newly developed FPN is implemented to identify the different local and global optima regions (niches).

“Survival of the fittest” from evolutionary concepts is the basis for GA and the approximate location of the highest fitness individual (global optima) is quickly identifiable from the FPN niche masters (cluster centers). Using the approximate global optima location from the FPN as an initial point, the GPS technique provides quicker and optimal solutions for the five variables under investigation – the crustal thickness, Vp/Vs ratio and the three associated weights. Applications of GA and GPS-FPN using seismic data from seismic stations within Ethiopia and surrounding the East Africa Rift System provided results which are consistent with previously published studies. The GPS technique is among the very few provably convergent, derivative-free search methods for linearly constrained optimization problems. GPS is shown in this study to be a powerful optimization tool that provides consistent results as if it searches the parameter space exhaustively. However, GPS searches the parameter space only in a given pattern and computes objective function values at few points. Key features of GPS technique reported in this study also include repeatability of its results, unlike heuristic search approaches, repeatability of the number of iterations as well as the number of objective function evaluations as long as initial values, the lower and upper bounds, and the processing machine stay the same. GPS even produces consistently similar results irrespective of initial values. The limitation of GPS being sometimes trapped at a local optimum is solved in this study by combining it with FPN.

CHAPTER 1

Introduction

1.1 Motivation

Direct and evolutionary techniques have advanced since their original inception and have since been implemented to solve challenging problems in different areas of studies and applications. In this dissertation, a technique that combines and uses both direct and evolutionary approaches is introduced to solve a seismic optimization problem. A direct search method of generalized pattern search (GPS) and evolutionary algorithms based on genetic algorithm (GA) and fitness proportionate niching (FPN) have been effectively implemented to solve the given seismic optimization problem. A trade-off study has been performed to illustrate the effectiveness of the two methods in solving the seismic problem.

The direct method of GPS has been developed between the 1990s and early 2000 (Kolda, Lewis and Torczon, 2003; Lewis and Torczon, 2001; Lewis, Torczon and Trosset, 2000; Lewis and Torczon, 1997, 1999a, b), but has not been applied in many disciplines yet. To the best of our knowledge, the GPS technique is applied here for the first time for solving such a seismic problem. Since the seismic optimization problem introduced in this dissertation can have both global as well as local optima, finding the global solution is a challenge at times because the search technique can be trapped in local optima. In order to address this problem, one has to be able to identify the appropriate starting/initial values that would lead to obtaining the correct solution. To accomplish the identification of the best starting values, the fitness proportionate niching (FPN) algorithm is proposed in this research

The FPN algorithm is a new technique that has been developed at North Carolina

Agricultural & Technical State University in 2012 as a dynamic clustering algorithm (Workineh and Homaifar, 2012). The FPN algorithm has been implemented for finding the correct initial values in the seismic optimization problem because no a priori knowledge about the distribution of the surface is available. The techniques introduced here are used to solve a seismic inverse problem. Inverse problems are inherently difficult problems, especially when they are compared to their counterpart forward problems.

There are many incentives for developing techniques to solve seismic problems. Some of these motivations are discussed in the following paragraphs. Seismic waves are generated by earthquakes or ground vibrations. Seismic waves or seismic signals are investigated primarily to mitigate earthquake hazards, as earthquakes could cause many deaths and enormous property losses due mainly to soil failure (liquefaction) and tsunamis. China has been hit especially hard in historic earthquakes. The Shensi earthquake in China is estimated to have killed more than 800,000 on January 23, 1531. Another devastating earthquake in Tangshan, China, killed almost 250,000 people on July 28, 1976. Among the most devastating earthquakes in the last decade include the 2004 Sumatra, Indonesia, earthquake which killed more than 283,000 people. The 1990 Manjil-Rudbar Iran earthquake, which killed more than 40,000 people, is among the most devastating earthquakes in the middle-east and the Persian Gulf. Among the most overwhelming recent earthquakes in our memories are the 2010 Haiti earthquake, which killed more than 220,000 people, and the 2011 Tohoku Japan earthquake that became a cause for the death of more than 18,000 people and a nuclear meltdown.

Besides the loss of life, earthquakes could cause huge amount of destruction, and cleaning up and repairing the damages of earthquakes are very expensive. Hence, study related to seismic waves has an incentive in this regard too. The October 1989 San Francisco, California, earthquake cost was about $5.6 billion while the January 1994 Northridge, California, earthquake cost was $15 billion. The January 1995 Hyogoken-Nanbu (Kobe) Japan, earthquake cost was $150 billion, whereas the March 2011 Japan Earthquake (Fukushima Dai-ichi) and related nuclear disaster cost was more than $235 billion.

Seismic signals provide information about the internal structure and composition of the earth and this provides another motivation for seismic studies. Based on such information, seismic studies could be used for economic purposes such as exploration of oil, gas, and minerals. Seismic techniques are arguably the most powerful geophysical techniques for prospecting oil, gas and minerals.

Seismic signals have been very useful means for studying the amount of thinning and thickening of the crust at different parts of the world, in verifying the different hypotheses in the plate tectonics theory. One reason for studying crustal structure is to try to verify the thinning of crust at the rift valleys and thickening of the crust at plate collision zones, as implied by plate tectonics theory. Understanding the composition of the earth in different parts of the earth and determining changes that have occurred there is an active area of research. This is one of the motivations for our research and some details related to crustal structure studies will be covered in section 2.6.

Another reason for investigating seismic signals is because seismic signals have been useful tools for discriminating nuclear explosions from earthquakes. Since a nuclear explosion, for that matter any explosion or implosion, produces seismic waves just like an earthquake, techniques have been developed and still being improved to correctly identify whether a seismic event is an earthquake or an explosion. Hence, monitoring and verifying the implementation of the Comprehensive Nuclear-Test-Ban Treaty has been made easier due to techniques utilizing seismic signals. Some of these seismic techniques involve a closer examination of the following parameters: seismic source depth, polarity of the first arriving (primary or P-) wave, and the strength of seismic surface waves.

1.2 Problem Statement

In this dissertation, both direct and evolutionary techniques have been introduced to solve a global optimization problem of inverting receiver functions based on H-κ stacking. First, some of the seismic data and seismic approaches used in this dissertation are discussed. Receiver functions are a time series obtained by de-convolving vertical component seismograms from radial component seismograms (Langston, 1979). Receiver functions consist of a number of phases, the arrival times of which are correlated with discontinuities in the crust and upper mantle. In fact, receiver functions represent the impulse response of the receiver structure beneath the seismic station (Rondenay, 2009). H-κ stacking of receiver functions (where H represents the crustal thickness and κ designates the ratio of crustal P-wave velocity to crustal S-wave (secondary wave) velocity) is a technique useful for estimating crustal thickness H and crustal Vp/Vs ratio κ (Zhu and Kanamori, 2000). The κ is interrelated to the crustal Poisson's ratio, which is a very commonly used elastic parameter that indicates the composition and nature of the crust beneath a seismic station.

In the implementation of the H-κ stacking, the arrival times for three important converted phases in the receiver functions, the P-to-S converted phase (Ps) and the first two reverberations of P-to-S converted phases (PpPs and PsPs+PpSs) in the crust, need to be determined. Ammon (1991) has shown that a receiver function is just a scaled version of the radial component seismogram with the exclusion of P multiples. In other words, receiver functions consist of P-to-S converted phases and reverberations of converted phases. The arrival times of these converted phases and their multiples are related to the crustal thickness (H) and κ (Zandt et al., 1995). The main goal of this research is to investigate the application of GPS and GA along with FPN techniques to determine optimal values of H, κ, and weights which are important parts to compute the H-κ stacking.

Previous studies have utilized different approaches for solving the H-κ stacking of receiver functions to determine crustal thickness (H) and κ parameters. The optimization in the current study takes into account the previous studies for the employment of the H-κ stacking algorithm and the weights corresponding to different phases in the receiver functions. Besides optimality, automatic algorithm development is the aim of the current study.

More receiver functions from different seismic stations have been utilized when the GPS technique was applied as compared to the one used with the GA technique. The GPS technique is among the very few provably convergent and derivative-free methods applicable to linearly constrained optimization (minimization) problems like the H-κ stacking problem in this dissertation. The GPS method amounts to finding the steepest descent path without computing the derivative of the objective function. In the last decade, the GPS technique has emerged as an important derivative free technique, developed mainly with the support of NASA, yet has not been widely applied in many areas, including seismic studies. In order to apply the GPS technique, first the maximization problem of H-κ stacking is converted to a minimization problem. For such a minimization problem, GPS searches for the optimal solution by determining the point that provides the minimum objective function value among a sequence of points in a chosen matrix pattern. Almost, the outputs and results from the GPS implementation can be repeated unlike other heuristic search approaches. Not only the outputs, but also the number of iterations as well as the number of objective function evaluations remain the same as long as initial values, the lower and upper bounds, as well as the processing machine stay the same. It is observed that optimal values will be almost identical irrespective of the initial values, as long as those initial values are within a plausible range.

The optimization problem emanating from the work of Zhu and Kanamori (2000) has given rise to the idea of implementing the GA, GPS and FPN techniques in this research. The original problem setup is a maximization problem known as H-κ stacking technique. Since its inception, the H-κ stacking has been applied in many crustal structure studies to determine crustal thickness and κ or the Poisson's ratio (e.g., Julia and Mejia, 2004; Dugda and Nyblade, 2006; Buffoni et al., 2012). In the previous studies published so far, values of weights, which are necessary components for the H-κ stacking, have been assigned through assumptions or using Monte Carlo Simulation technique (e.g., Zhu and Kanamori, 2000; Dugda et al, 2005) or by using Genetic Algorithms (GA) (Dugda et al., 2012). One of the objectives of this dissertation is to implement the GPS technique and evaluate its performance for H-κ stacking of receiver functions. The GPS technique will enable to determine the optimal parameters using a minimization of an objective function value. One may like to compare the optimal weights resulting from the GPS employment with previously published works (e.g., Dugda et al., 2005; Dugda et al., 2012) for the determination of the weights as well as H and κ parameters. The GPS technique realized here has the potential to be used as an automatic crustal parameters determination tool along with important weights for the receiver function phases. In this research, the GPS technique, when allowed to go through complete polling, would be able to escape local optima to provide global optima. In the event when the GPS is trapped in a local optimum, the proposed FPN technique would enable to identify the right initial values necessary to arrive at the global optimum point.

1.3 Contribution

The contributions of this study include developing tools by implementing Genetic Algorithms (GA), Generalized Pattern Search (GPS) and Fitness Proportionate Niching (FPN) as automatic techniques to solve the problem of inverting receiver functions for a number of seismic stations at the same time. The inversion in this study will solve for both crustal parameters and the weights simultaneously. So far, techniques introduced have been implemented to solve for one seismic station at a time and the weight assignment has been either fixed by assumption or variable. The variable weights approach introduced in 2005 uses a Monte Carlo technique to find the variable weights for receiver functions from one seismic station (Dugda et al., 2005). Besides, verification of the picking of the phases was required to determine the correctness of the phase picking. The search method used to determine the best parameters was an exhaustive search technique but focused on the unimodal portions of the objective function at a time. Because of all these factors, the technique introduced in 2005 was time-consuming, applicable to one station at a time, and it was not suitable for automatic implementation. In fact the 2005 study has focused on the accuracy of the then newly introduced variable weights approach, on the revelation of the crustal structure of the region and most importantly on the implication of the crustal structure discovered by that study. Enhancement of the technique introduced at that time was not within the scope of that study, as there was no enough time to do that. The research and techniques developed in this dissertation take the variable weights approach to a whole new higher level by introducing new optimization techniques.

In this dissertation, the receiver function inversion problem is shown to be a multimodal problem and two different techniques have been developed to solve this problem. The techniques developed here involve evolutionary and direct optimization approaches and they are used to invert the receiver functions to determine crustal parameters and the three associated weights simultaneously. The first technique developed here makes use of the evolutionary Genetic Algorithms (GA) optimization technique. The second technique introduced in this research combines the direct Generalized Pattern Search (GPS) and evolutionary Fitness Proportionate Niching (FPN) techniques by employing their strengths. Compared to the previous Monte Carlo variable weights approach introduced about a decade ago (Dugda et al., 2005), the current GA and GPS-FPN techniques have tremendous advantages of saving time and these new techniques are suitable for automatic and simultaneous determination of the crustal parameters along with appropriate weights.

The GPS technique is among the very few provably convergent, derivative-free search techniques for linearly constrained optimization problems. The study in this dissertation has shown clearly that GPS technique offers repeatable results, especially compared to heuristic search approaches. Moreover, the number of iterations as well as the number of objective function evaluations will remain the same as long as initial values, the lower and upper bounds, and the processing machine remain the same. Thus, GPS is shown in the current study to be a very powerful optimization tool that provides consistent results as if it searches the parameter space exhaustively. The GPS, on the other hand, searches the parameter space in a given matrix pattern and computes objective function values only at those points in the pattern. Our study also shows the limitations of the GPS technique when it is applied to optimize multimodal problems.

Most real-life problems are multimodal in nature and in such cases this research has extended the power of GPS by using FPN technique to identify an initial search point for the GPS optimization search to start with. We call this combined technique GPS-FPN.

In the GPS-FPN implementation, as a first step, after the GPS inversion is complete for a set of stations, the GPS has been utilized again to test the final model to check for any initial model dependence. If the inversion is found to be sensitive to initial values for some specific seismic stations (i.e., if the objective function has different local and global optima in the given H-κ parameter space), then we invoke the FPN to identify the appropriate initial parameters. So far this seismic inversion problem has not been considered a multimodal problem to be solved by a multimodal optimization approaches. Finding the best initial values is a new approach introduced in this study and it is also a new application area for the FPN or for any other multimodal optimization approach. Thus, implementation of FPN for identifying the right initial values is another significant contribution of this research work. The tools developed and being developed can be instrumental in geological and geophysical surveys and investigations. Moreover, the technique can be used to solve similar exploration inverse problems for the petroleum/oil industry. In fact the technique can be extended to similar multimodal optimization problems.

The different parts of this research work have been published in peer-reviewed journal and conference papers:

- **Dugda,** Mulugeta, Abrham T. Workineh, Abdollah Homaifar and Jung Hyoun Kim (2012). "*Receiver Function Inversion Using Genetic Algorithms*," *Bulletin of the Seismological Society of America (BSSA)*, Vol. 102, No. 5, pp. 2245–2251, October 2012, doi: 10.1785/0120120001.

- ***Dugda,*** *Mulugeta, Abdollah Homaifar and Jung Hyoun Kim* (2013). *"Receiver Function Inversion Using Generalized Pattern Search Technique,"* the Geological Society of America (GSA) 125th year celebration and meeting in Denver Colorado, GSA Abstracts with Programs Vol. 45, No. 7, October 27-30, **2013**.
- ***Dugda,*** *Mulugeta, Abrham T. Workineh, Jung H. Kim, Abdollah Homaifar, "Optimal Receiver Function Inversion Using Generalized Pattern Search and Fitness Proportionate Niching (FPN) of Genetic Algorithms Approach*," American Geophysical Union (AGU) Fall meeting, San Francisco, CA, *Eos Trans. AGU, 94*(52), Fall Meet. Suppl. Dec. 9-13, **2013**.
- ***Dugda,*** *Mulugeta, Jung H. Kim, Abdollah Homaifar* (**2014**). *Generalized Pattern Search (GPS) and Fitness Proportionate Niching (FPN) for Optimal Multimodal Receiver Function Inversion*, Institute of Electrical and Electronic Engineers for Geosciences and Remote Sensing Society (IEEE-GRSS) Transactions (**In Preparation**).

1.4 Objectives and Goals

A non-seismologist researcher, such as a geochemist, even a geophysicist with little or no experience in seismology, a geographer, or an engineer may be interested in knowing the estimate of some important parameters of the earth like the crustal thickness (H) and the Poisson's ratio (ρ) or the Vp-to-Vs ratio (κ) beneath a certain place on Earth. One may have access to some seismic data from that specific region (available via online, or directly from a seismic station in the vicinity). How can the person come up with the solution to his problem of estimating those parameters without asking the help of a seismologist? The person may get an estimate of those parameters if an automatic crustal parameter computing system utilizing the seismic signals available at his/her disposal. Developing a programming tool that can be integrated into an automatic system is the main objective of this work. This research has shown that automatic crustal parameter determination technique can be developed using GPS and FPN.

This tools proposed here have been developed using Matlab for determining crustal parameters automatically or semi-automatically, and at some point in time the codes may be available to the public. The approach developed here will be an important step towards developing a fully automatic technique. The previous crustal parameter determination techniques using seismic signals and optimization tool need the expertise of a seismologist to endorse that the analysis algorithm and technique has properly picked the correct receiver function phases.

One stimulus for the automatic crustal parameter determination system could come from the fact that a non-seismologist may be interested in some of the crustal parameters. For instance, geochemists are usually interested in the Poisson's ratio of the crust in order to determine how much partial melt exists inside some particular crust. The automatic crustal parameter determination system planned in this study would be a friendly tool to the non-seismologist and does not need the approval of the seismologist.

1.5 Dissertation Outline

This dissertation is organized in six chapters. Chapter 1 states the problem we are attempting to address and discusses the motivation behind the research. The contributions of this study as well as the main objectives and goal of the research work are also deliberated in chapter 1.

Chapter 2 provides background and literature survey for both techniques developed in this study, the processed seismic data and the algorithm involved for processing it. A background study has been devoted to receiver functions, H-κ stacking of receiver functions, genetic algorithms (GA), pattern search methods for linearly constrained minimization problems, and fitness proportionate niching (FPN). Moreover, a section is dedicated to discuss the difference between inverse and forward problems.

Chapter 3 describes the proposed methodology – genetic algorithms (GA), generalized pattern search (GPS), and fitness proportionate niching (FPN) techniques. Implementation of H-κ stacking of receiver functions, fixed weights vs. variable weights in H-κ stacking; also, the inverse problem of finding model parameters with best fitting data are discussed in this chapter . The H-κ stacking of receiver functions is also demonstrated to be a multimodal optimization problem.

Chapter 4 considers the experimental implementation of the techniques introduced in this dissertation and displays the results of this research. Genetic algorithm implementation and test results, the GPS test results as well as the FPN results are presented here. The location of the seismic stations for testing the working of the proposed techniques is provided in this chapter. The flow chart of the GPS implementation, the GPS results, the convergence of final values and objective function are also presented here. A graphical user interface (GUI) implementation for the GPS algorithm and the implementation of the GPS convergence test algorithm is presented in here. A complete polling with the GPS is considered for global optimization; also, a partial polling versus complete polling is demonstrated too.

Chapter 5 provides discussions on the observed results and deals with performance evaluation. A discussion on the GA and the GPS implementation and results is given here. More results for GPS convergence test and a discussion has been included in this chapter. Results and a discussion on complete polling of GPS for global optimization have been incorporated here. A comparison of crustal parameters and weights obtained in this research and some previous studies, for both crustal thickness and Vp/Vs values for the same seismic stations are conveyed in chapter 5. Finally, a comparison of the GA and the GPS techniques based on speed and computational cost has been offered. A discussion is also given on the FPN technique performance speed.

Chapter 6 offers some concluding remarks and future research directions. It provides conclusions on the GA and the GPS implementations. It gives concluding remarks regarding FPN implementations too. It also describes the contributions of this particular study. Further research on a class of optimization of weighted sum problems and further research on a multimodal optimization and dynamic clustering techniques is proposed here. Further studies on optimization and data analytics (data mining techniques, like clustering and machine learning) techniques applied to both earthquake generated seismic data and other seismic data from exploration field studies. At the end, Appendix A is added into this dissertation. The Appendix lists some of my recent publications.

CHAPTER 2

Background and Literature Review

Seismic waves are created by earthquakes or ground vibrations. Seismic signals recorded in seismic stations contain information about the seismic source as well as the media through which the seismic wave has propagated. In addition to being essential tools for reducing earthquake dangers, seismic signals have been helpful to learn the internal structure and composition of the earth. The two important internal structure and composition model parameters that are commonly determined using different seismic methods are layer thickness and layer velocity. Besides, seismic phase arrival times are the regularly observed and measured data parameter exploited in different seismic techniques. In this dissertation, the crustal layer thickness and the ratio of crustal velocities are the model parameters of interest and seismic phase arrival times and amplitudes are utilized for determining those crustal model parameters. The process of discovering such earth model parameters from observed data requires an inversion process. Seismic velocity variations with depth, as one glances deeper into crust and mantle structure, are known as crustal and mantle velocity structures.

The focus of this research is developing techniques suitable for inverting processed seismic data called receiver functions and retrieve optimal crustal model parameters with appropriate weights. Initially, concepts on seismic wave propagation, seismic signals, receiver functions, and stacking of receiver functions are presented. This chapter also contains literature survey on inverse and forward problems, models of the earth structure, and the importance of crustal structure investigation. Furthermore, different ways of stacking receiver functions and state-of-the-art techniques for solving H-κ stack of receiver functions are discussed in this chapter. The rest of the literature review will embrace the techniques implemented in this dissertation. These techniques include pattern search methods for linearly constrained minimization problems, specifically the generalized pattern search (GPS) technique, genetic algorithms (GA), and fitness proportionate niching (FPN) technique.

2.1 Seismic Wave Propagation

There are two types of body waves that arise as a result of an earthquake: P-waves and S-waves. P-waves (Primary waves) are the first seismic waves to arrive at the seismic station. S-waves (Secondary waves) are the seismic waves that arrive at the seismic station after the P-wave arrival. P- and S-waves are called seismic body waves because they propagate throughout the body of the Earth. After the Secondary waves arrive at the seismic station, there is another group of waves called surface waves (Love and Rayleigh waves) which arrive at the seismic station. As their name indicates, surface waves are seismic waves that are propagating close to the surface of the earth. An earthquake basically radiates P- and S-waves in all directions and the interaction of those P- and S-waves with the Earth's surface and shallow structures produces surface waves.

A typical seismogram in Figure 2.1 shows the basic kinds of seismic waves originating from an earthquake event. The first arriving seismic wave called the Primary (P-) wave is a longitudinal wave which is similar to a sound wave. The second arriving seismic wave is called the Secondary (S-) wave which is a shear or transverse wave. P-waves are compressional waves that cause the medium through which they propagate to deform such that the particle motion in the medium is parallel to the direction of the wave propagation. On the other hand, the S-waves are shear waves that cause particles in the medium of propagation to move perpendicular to the direction of wave propagation. In both the P-wave and S-wave propagation, the material (medium of propagation) returns to its original shape after the waves pass.

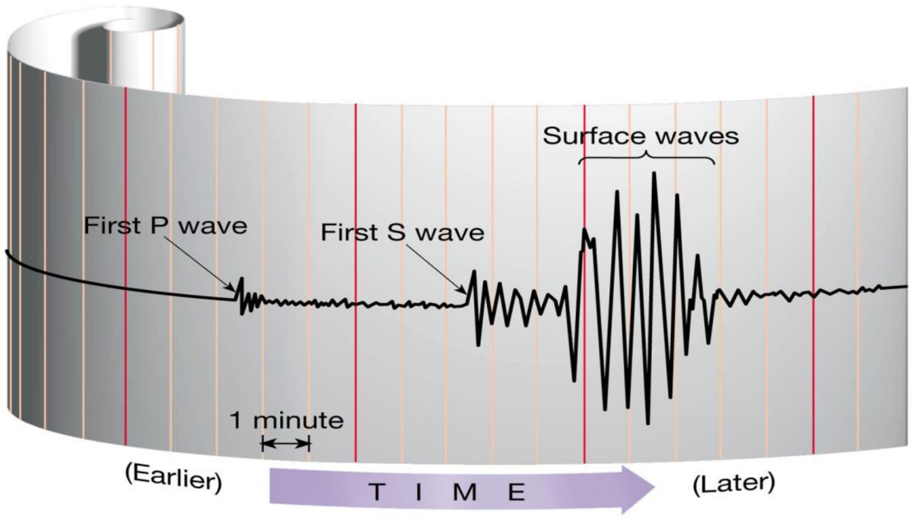


Figure 2.1 A typical seismogram showing the Primary (P) and Secondary (S) body waves and surface waves generated as a result of an earthquake (Tarbuck et al., 2013).

Generally three components of seismograms are recorded in a single seismic station (in a single seismometer). The recording of three components in three perpendicular Cartesian coordinates enables to capture the motion of the ground fully, and enables it to describe the motion of the ground completely. These three components include: two horizontal components of East-West (EW) motion and North-South (NS) motion, and a single vertical component of Z. An example of three component seismic recording by a seismic (seismographic) station called NNA (Nana, Peru) is displayed in Figure 2.2. Seismic station data contain information on magnitude of earthquake (e.g., 6.5), where the earthquake happened (e.g., near central Chile), the country or city where the seismic data is recorded (e.g., Peru), the distance between the seismic station and the earthquake source (e.g., 1993 km away), what depth the earthquake occurred from the earth surface, what time the earthquake originated, the arrival times of the P-, the S- and surface waves, and the exact locations of the seismic station (e.g., 11.9875^0 S latitude, 76.8422^0

W longitude) and the location of the earthquake (e.g., 29.2934^0 S latitude, 71.5471^0 W longitude).

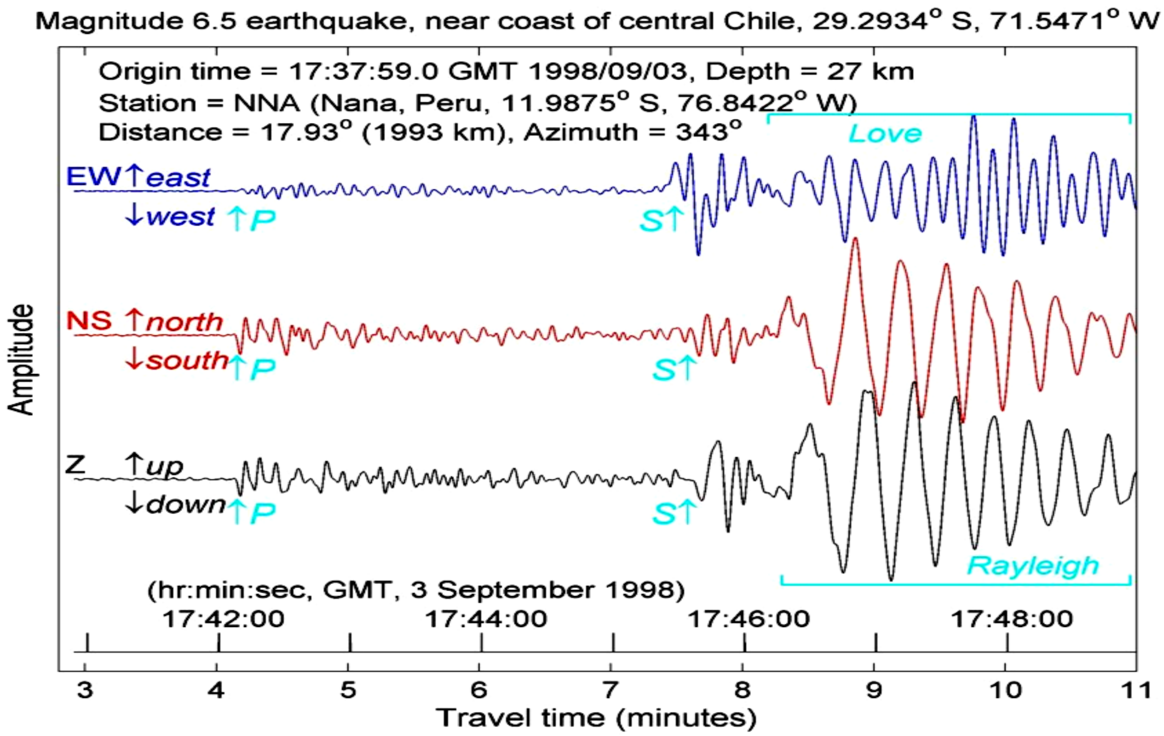


Figure 2.2 An example of three-component seismograms recording in a single seismic station (Braile, 2010).

2.2 Rotating Seismic Data

Generally, a seismometer records data along three directions: the vertical **Z**, North-South **N (NS)**, and East-West **E (EW)** directions (the **ZNE** rotation system). For the purpose of computing receiver functions, the two horizontal components **N** and **E** are projected into the radial **R** and tangential **T** components in the following way to form **ZRT** system:

$$\begin{bmatrix} R \\ T \\ Z \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\ \theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} E \\ N \\ Z \end{bmatrix} \qquad (2.1)$$

where $\theta = 3\pi/2\text{-}\xi$ and ξ is the back azimuth (BAZ).

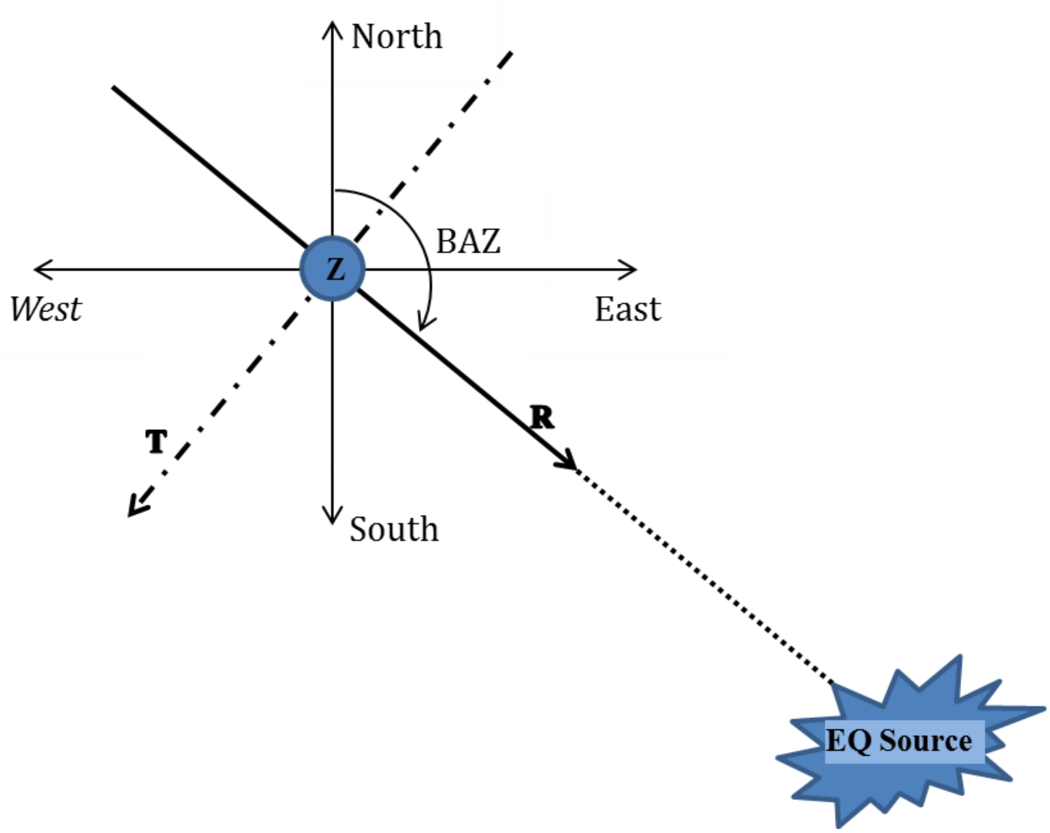


Figure 2.3 The transformation from ZNE system to ZRT system. EQ means Earthquake.

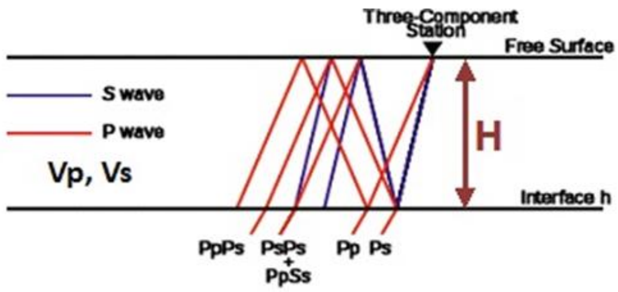
Figure 2.3 shows the transformation from ZNE system to ZRT system. The radial component R is along the direction connecting the seismometer location (Z) with the earthquake source (the seismic source) location (EQ source), while T is transverse to that direction. Equation (2.1) basically transforms the ZNE system to ZRT system.

2.3 Receiver Function

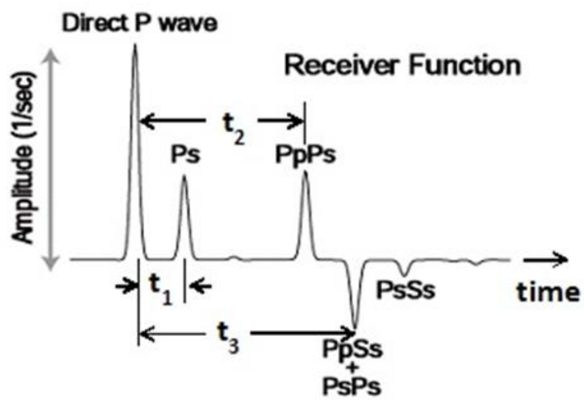
Receiver functions are time series determined by the de-convolution of vertical component seismograms from radial component seismograms (Langston, 1979). The receiver function is an approximation of the crust/mantle transfer function for an incident teleseismic or deep-event regional *P* wave (Burdick and Langston, 1977; Langston, 1977, 1979; Zhang and Langston, 1995). The receiver functions actually characterize the impulse response of the earth structure at the receiving end of the seismic signal (Ammon et al, 1990; Rondenay, 2009).

Receiver functions display a number of phases the arrival times of which are correlated with discontinuities in the crust and upper mantle. H-κ stacking of receiver functions (where H is the crustal thickness and κ is the ratio of crustal P-wave velocity Vp to S-wave velocity Vs) is a technique useful for estimating crustal thickness H and crustal Poisson's ratio which is interrelated to the crustal Vp-to-Vs ratio κ (Zhu and Kanamori, 2000). In the implementation of the H-κ stacking, the arrival times for three important converted phases in the receiver functions, the P-to-S converted phase (Ps) and the first two reverberations of P-to-S converted phases (PpPs and PsPs+PpSs), need to be determined. These arrival times are determined from the H and κ (Zandt et al., 1995).

In Figure 2.4(a), a crust with thickness H and seismic P- and S-waves are shown. Those P- and S-waves originate from an impinging plane P-wave at the crust-mantle boundary called the Moho, in honor of the seismologist Mohorovicic who discovered it. The direct P-wave and P-to-S converted wave as well as some reverberated waves are also shown using red and blue lines there. As shown in Figure 2.4(b), a receiver function time series is presented which corresponds to the given crust with the arriving direct P wave, P-to-S converted phase (Ps), PpPs phase, PpSs+PsPs phase, PsSs phase. As shown, t_1, t_2, and t_3 provide arrival times of the P-to-S converted phase Ps (converted at the crust-mantle boundary), and the two seismic wave reverberations PpPs and PpSs+PsPs in the crust relative to the direct P wave arrival Pp.



(a)



(b)

Figure 2.4 (a) A three component seismic station, a simplified one layer crustal Earth model of thickness H, and different seismic phases; (b) receiver function corresponding to the given crust and impinging P-wave (Modified from Ammon et al, 1990).

Relationships of the crustal thickness H, crustal velocities Vp and Vs with the different phase arrival times in the receiver function are given by the following equations (Zandt et al., 1995):

$$H = \frac{t_{Ps} - t_P}{\sqrt{\frac{1}{V_s^2} - p^2} - \sqrt{\frac{1}{V_p^2} - p^2}} \quad (2.2)$$

$$H = \frac{t_{PpPs} - t_P}{\sqrt{\frac{1}{V_s^2} - p^2} + \sqrt{\frac{1}{V_p^2} - p^2}} \quad (2.3)$$

$$H = \frac{t_{PpSs+PsPs} - t_P}{2\sqrt{\frac{1}{V_s^2} - p^2}} \qquad (2.4)$$

where p is the ray parameter of the seismic signal commonly specified by:

$$p = \frac{\sin\theta}{V_p} \qquad (2.5)$$

where θ is the angle of incidence between the seismic ray and a normal, and V_p is the velocity of the P-wave seismic ray.

Receiver functions are computed by the deconvolvolution of vertical component seismograms from radial component seismograms (Langston, 1979). However, in practice, an averaging function A(ω) and a Gaussian low-pass filter G(ω) are included when determining receiver functions.

As stated earlier, receiver functions are determined by de-convolving vertical component seismograms from radial component seismograms. Instead of de-convolution in the time domain, it is a standard procedure to divide the spectrum of radial component seismograms by the spectrum of vertical component seismograms. Under ideal conditions, receiver functions can be determined in the frequency domain using the ratio of radial component and vertical component motions (seismograms):

$$\mathrm{RF}(\omega) = \frac{\mathrm{R}(\omega)}{\mathrm{V}(\omega)} \qquad (2.6)$$

where RF(ω) is frequency-domain receiver function, R(ω) and V(ω) are frequency-domain radial and vertical component seismograms, respectively (Cassidy, 1992). Practically, because of noise in the data as well as the band-limited nature of the signal, the water-level technique is used to stabilize this spectral division (Clayton and Wiggins, 1976), and therefore, the receiver function estimate $\overline{\mathrm{RF}}(\omega)$ is given by:

$$\overline{\mathrm{RF}}(\omega) = \frac{\mathrm{R}(\omega)}{\mathrm{V}(\omega)} A(\omega) \qquad (2.7)$$

where A(ω) is the averaging function obtained using the relationship

$$\mathrm{A}(\omega) = \frac{\mathrm{R}(\omega)\mathrm{R}^{*}(\omega)\mathrm{G}(\omega)}{\varphi(\omega)} \qquad (2.8)$$

where

$$\mathrm{G}(\omega) = \exp\left(\frac{-\omega^{2}}{4a^{2}}\right) \qquad (2.9)$$

and

$$\varphi(\omega) = \max\{\mathrm{V}(\omega)\mathrm{V}(\omega)^{*}, \mathrm{c}.\max\{\mathrm{V}(\omega)\mathrm{V}(\omega)^{*}\}\} \qquad (2.10)$$

with **a** being the width of the Gaussian filter used to remove high-frequency noise, and **c** being the water-level parameter expressed as a fraction of the maximum vertical component power spectra (Ammon, 1991; Cassidy, 1992). Once the receiver functions are obtained in the frequency domain, the time-domain receiver function can be determined by applying inverse Fourier Transform as:

$$\overline{\mathrm{rf}}(\mathrm{t}) = \mathrm{F}^{-1}\left\{\overline{\mathrm{RF}}(\omega)\right\} \qquad (2.11)$$

This time domain receiver function is the one being employed in the GA implementation of H-κ analysis in this research.

Receiver functions encompass information about discontinuities in the crust and upper mantle, especially the crust-mantle discontinuity called the Moho. Since the arrival times of different seismic phases in receiver functions are interrelated to the different discontinuities, including the crust-mantle discontinuity, receiver function analysis has been used to study crustal structure for decades (e.g., Langston, 1979) (Figure 2.4).

Tangential receiver functions are also obtained by de-convolving vertical component seismograms from tangential component seismograms. The tangential receiver functions are usually important to check for heterogeneity of the earth structure under investigation.

2.3.1 H-κ stacking of receiver functions. The arrival time differences in the phases of the receiver functions have been utilized to determine crustal parameters (crustal thickness and crustal Poisson's ratio) since their inception in late 1970s. The arrival time difference between the direct P-phase and the first P-to-S converted phase has been used for this purpose first. Since the problem is not well constrained this approach is not a noble candidate technique. Since the mid-1990s, different efforts have been exerted to constrain the problem using more converted and reverberated phases in the receiver functions (e.g., Ammon and Zandt, 1995).

H-κ stacking of receiver functions has been applied to get an estimate of the thickness and Poisson's ratio of the crust (Zhu and Kanamori, 2000). Important information for H-κ stacking of receiver functions are arrival times for three phases: the P-to-S converted phase (Ps) and the first two reverberations of P-to-S converted phases (PpPs and PsPs+PpSs). These times are related to crustal thickness and V_p/V_s by the equations derived in Zandt et al. (1995).

H and κ have a strong trade off (Ammon et al., 1990; Zandt et al., 1995). In order to lessen the ambiguity brought about by this trade off, Zhu and Kanamori (2000) incorporated the later arriving crustal reverberations PpPs and PpSs+PsPs in a stacking procedure whereby the stacking itself transforms the time-domain receiver functions directly to objective function values in H-κ parameter space.

The relative times t_1, t_2, and t_3 as well as H and κ are used in our objective function formulation. In addition, the objective function originally given by Zhu and Kanamori (2000) is used in this study too for the implementation of the GA, GPS and FPN techniques in H-κ stacking as:

Maximize

$$S(H,\kappa) = \sum_{j=1}^{N} w_1 r_j\big(t_1(H,\kappa)\big) + w_2 r_j\big(t_2(H,\kappa)\big) + w_3 r_j\big(t_3(H,\kappa)\big) \qquad (2.12)$$

subject to

$$w_1 + w_2 + w_3 = 1 \qquad (2.13)$$

where t_1, t_2 and t_3 are obtained using the respective formulae given below (equations 2.14-2.16), and w_1, w_2, w_3 are weights (with $w_3=1-w_1-w_2$). The $r_j(t_i)$, i=1, 2, 3, are the receiver function amplitude values at the predicted arrival times of the Ps, PpPs, and PsPs+PpSs phases, respectively, for the j^{th} receiver function, and N is the total number of receiver functions used. By performing a grid search through H and κ parameter space, the H and κ values corresponding to the maximum value of the objective function can be determined. The main hypothesis behind H-κ stacking of receiver functions is that the weighted sum stack will attain its maximum value when H and κ take their proper values (Zhu and Kanamori, 2000).

In the objective function equation (2.12), t_1, t_2 and t_3 can be related to H and κ using the following formulae (Zandt et al., 1995):

$$t_1 = \left(\sqrt{\kappa^2 - V_p^2 p^2} - \sqrt{1 - V_p^2 p^2}\right) \cdot \frac{H}{V_p} \tag{2.14}$$

$$t_2 = \left(\sqrt{\kappa^2 - V_p^2 p^2} + \sqrt{1 - V_p^2 p^2}\right) \cdot \frac{H}{V_p} \tag{2.15}$$

$$t_3 = \frac{2H}{V_p} \cdot \left(\sqrt{\kappa^2 - V_p^2 p^2}\right) \tag{2.16}$$

It is worth noting that t_1, t_2 and t_3 are relative arrival times for Ps, PpPs, and PsPs+PpSs seismic phases, respectively, with respect to the direct P-wave arrival as it is shown in Figure 2.4(b).

Poisson's ratio is one of the very important elastic parameters of the Earth that seismologists and other geologists use to study the composition of the Earth. Poisson's ratios are directly related to Vp/Vs ratios. The mathematical relationships for Vp/Vs ratio and Poisson's ratio (σ), respectively, are as follows (Christensen, 1996; Zandt et al., 1995):

$$\kappa = \frac{V_p}{V_s} = \sqrt{(1 - p^2 V_p^2)(2\left(\frac{t_1}{t_2 - t_1}\right) + 1)^2 + p^2 V_p^2} \tag{2.17}$$

$$\sigma = \frac{V_p^{\,2} - 2V_s^2}{2(V_p^{\,2} - V_s^2)} = \frac{\kappa^2 - 2}{2\left(\kappa^2 - 1\right)} \tag{2.18}$$

Some studies have indicated a pattern between Moho depth and Poisson's ratio where Poisson's ratio decreases with increasing crustal thickness for Moho depths of 20 to 45 km (Egorkin, 1998; Chevrot and van der Hilst, 2000; Dugda et al., 2005). The highest Poisson's ratios are found for the thinnest crust and this is believed to reflect the mafic composition of the rifted crust. On the other hand, lower Poisson's ratios found for thicker crust may reflect an unmodified crust.

2.4 Inverse Versus Forward Problems

The seismic problem is an inverse problem. In an inverse problem, there will be an observed data and the model parameters that are responsible for creating those data are to be determined. Generating synthetic data for a given set of model parameters is the opposite process of performing inverse modeling and it is known as a forward problem. In a forward problem, the model and model parameters are known to generate data based on the model.

2.4.1 Inverse problem theory. The majority of physics and engineering problems involve solving forward problems. A common example of forward problem would be a calculation of the electromagnetic field at a given distance from a conducting rod. If every impurity was mapped and the current was known everywhere inside the rod, the electromagnetic field could be calculated uniquely for every point outside the rod. One related inverse problem to this would be to determine the structure of the rod by measuring the electromagnetic field outside the rod. The inverse problem is harder than the forward problem because several different rod structures can produce the same electromagnetic field, which is a result of the non-uniqueness behavior of inverse problems.

2.4.2 The inverse problem. Consider the relationship between an observed data d and the model parameters m as given by:

$$d = G(m) \tag{2.19}$$

where G is forward operator describing the relationship and it is known as observation operator, or observation function. Generally G denotes the governing equations that relate the model parameters to the observed data but usually it is difficult to find explicit relationships between *d* and *m*. The goal of an inverse problem is to find the best model parameters *m* from an observed data *d (*Torantola, 2004; Menke, 2012; Richardson et al., 2012).

There is an inherently more difficult family of inverse problems collectively known as non-linear inverse problems. Non-linear inverse problems have a more complex relationship between data and model (Khosrow and Sabatier, 1977; Borchers and Thurber, 2012; Press et al., 2007). Another important application is constructing computational models of oil reservoirs, and it is widely practiced in geophysics overall.

For a given condition, a theoretical forward model can be used to calculate synthetic receiver function. On the other hand, the purpose of research in this dissertation is to find the earth's crustal velocity structure that generates the observed receiver function. The inverse problem is able to find a model *m* that corresponds to some observed data d_{obs} by:

$$m = G^{-1}(d_{obs}) \tag{2.20}$$

In general, for most inverse problems it is not possible to determine G^{-1}. In particular, inversion of receiver functions is highly non-linear problem, and G^{-1} cannot be constructed (Ammon et al., 1990; Chadan and Sabatier, 1977; Borchers, Brian, and Thurber, Clifford, 2012; Press et al., 2007).

2.5 Models of the Earth Structure

Seismic waves, specially the P- and S-waves, could show the internal structure of the earth. Seismic signals are used to make seismic tomographic images of the earth. For making the tomographic images of the earth, the seismic signals are used just like x-ray signals which are used for human body tomography. The main difference between seismic tomography and x-ray tomography is the fact that there is no control over the source of the seismic signal, which is the earthquake, whereas we have the control over the sources of x-ray. The earthquake can happen anywhere on the globe. Only big earthquakes are useful for tomographic imaging of the deep structure of the earth, because those earthquakes are the ones having the capability of generating seismic waves that can crisscross the whole earth.

Crust is a very thin structure of the earth. Crustal thickness ranges between about 0 km and up to 70 km, with an average of approximately 42 km for continental crust. Two places on earth where one can find thickest crust are beneath Tibetan Plateau in Asia and under Andes mountains in Latin America. A very simplified Earth model is given on Figure 2.5. It shows the different layers of the earth and their approximate relative sizes except the crust which is exaggerated. The real size of the crust is too small (less than 1%) compared to the other layers of the earth. Out of the 6371 km radius of the earth, the crust makes up on average only the top 40 km thickness. The mantle forms more than 80% of the earth’s volume and it is more than 2800 km thick. The outer core is liquid iron and nickel and it is about 1400 km thick. The analysis of seismic waves, especially the absence of the S-waves, has led scientist to conclude that the outer core is liquid. The inner core is more than 1100 km thick and is a solid iron. Most of the weight of the earth is concentrated in the core.

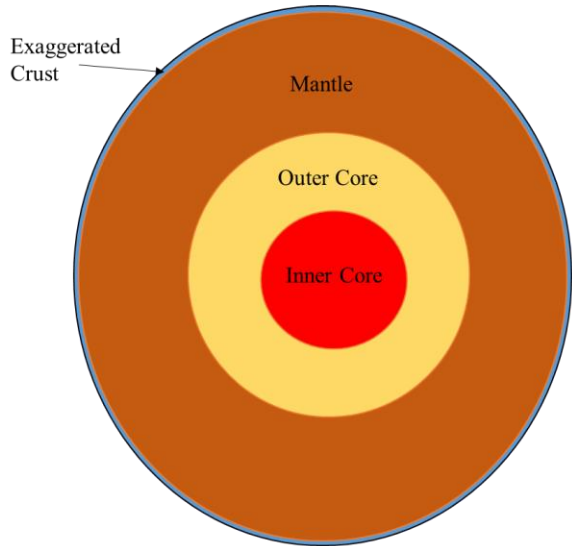
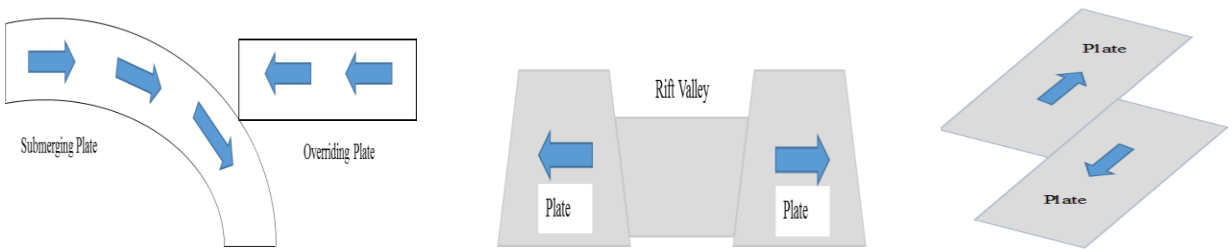


Figure 2.5 The different layers of the earth

In Figure 2.6, three kinds of plate interaction are shown. The interaction of plates is the major cause of earthquakes, and so most earthquakes occur at the plate boundaries. There are three different kinds of plate boundaries as shown in the figure. Figure 2.6(a) displays convergent plate boundary where a submerging plate going beneath another plate called overriding plate at the subduction zone, like in Japan and Philippines. Subduction zones are the origins of big earthquakes that can be classified as great (> magnitude 8) earthquakes, like that of the 2011 Japan earthquake and the 1960 largest recorded earthquake of Chile. The plates are formed from the crust and the uppermost mantle and they are rigid and brittle hard rocks. Thickness of plates vary from pace to place and can reach up to 100 km or more. Plates are sitting and ridding on top of the part of the mantle called asthenosphere which is less rigid, ductile, hotter, and more mobile. Plates are moving due to mantle convection. In Figure 2.6(b) shows a divergent plate boundary where we can observe another kind of plate interaction at the rift valley. This is the case at mid-ocean ridges and rift valley regions like the East African Rift

Valley which includes the Ethiopian Rift Valley. The seismic stations in which seismic data has been collected and used for our study here are sitting inside and around the Ethiopian Rift Valley. Figure 2.6(c) shows Transform plate boundaries with transform faults similar to the San Andreas Fault in California. Major earthquakes (with a magnitude of 7 to 7.9) can occur at transform faults. Compared to Transform Faults (Transform plate boundaries) and Subduction zones (Convergent plate boundaries), the sizes of earthquakes at Divergent plate boundaries, like the mid-ocean ridges or rift valleys, are not that big.



(a) (b) (c)

Figure 2.6 Interaction of plates at (a) subduction zones, (b) rift valleys, (c) transform plate faults

2.6 Crustal Structure Studies

Within the geoscientific community crustal structures are studied with various drives. Some of these stimuli include determining the amount of thinning and thickening of the crust in a certain region, especially related to plate tectonics theory. One incentive for studying crustal structure is to try to verify the thinning of crust at the rift valleys and thickening of the crust at plate collision zones, as implied by plate tectonics theory. This is one of the motivations for our research.

Another reason for crustal studying is to learn the composition of the earth in different parts of the earth and determining changes that has occurred in the crust. Determining how much change has occurred to the crust in a region of study is an active area of research. Ground truth development for different geophysical investigations, including Comprehensive Nuclear Test Ban Treaty monitoring & verification is one important reason for studying the crust. In fact ground truth is developed for various reasons.

There are several advantages mentioned in literature by different authors who studied crust and mantle structure in different parts of the world. Some of the specific and general advantages include: knowledge of crust and mantle structure across the Basin and Range-Colorado Plateau in the US has implications for Cenozoic extensional mechanism (Zandt et al., 1995). Another study of crustal structure of northeastern Iceland (Staples et al., 1997) placed a constraint on the flow within the Iceland mantle plume. A different study showed that continental crust composition is constrained by measurements of crustal Poisson’s ratio (Zandt and Ammon, 1995). Several studies on Ethiopia and Kenya crust have shown implications for rift development in eastern Africa (Dugda et al., 2005), and an opening of a new ocean floor in the region in few millions of years’ time (Yirgu et al., 2006). Since 2005 the Afar Triangle (Afar Depression) in Ethiopia is experiencing many earthquakes, tremors, fissure formations, and numerous volcanic activities from the dormant volcano of Erta Ale. Many geological and geophysical studies in the Afar region indicate that new ocean floor is in the process of development; and geologists state that the new crack is an indication of the birth of a significant body of water which will effectively “cut- off” the horn of Africa from the mainland in about ten million years’ time (Yang and Chen, 2010).

2.7 Receiver Functions for Crustal Structure Parameters

The crustal structure parameters investigated in this dissertation are crustal thickness (**H**) and crustal Vp/Vs ratio (**κ**). In most seismic studies, the available data parameters are arrival times of seismic phases while the model parameters to be determined are usually layer thickness and layer velocities. We are more interested in finding the ratio of Vp and Vs as the second parameter.

Receiver functions have been widely used tools for studying the crustal and upper mantle structures since its inception in 1979 by Charles Langston. Figure 2.4 shows the crustal thickness H(h) and Vp/Vs ratio or the crustal Poisson's ratio (σ) parameters and the mathematical relationship to the arrival times of the different phases in the receiver functions are given by Equations (2.14-2.18).

Receiver functions have been applied using different approaches for determining the H and Poisson's ratio (σ) or κ. The two procedures that are commonly used so far include a simple receiver function stacking and an H-κ stacking of receiver functions.

2.7.1 Simple stack of receiver functions. Simple stacking of receiver function has been among the very common approaches and still a useful technique when the available data are very scarce. The simple stacking improves the signal-to-noise ratio (SNR) in the receiver functions which enhances the estimation of the crustal parameters... Simple stacking is just averaging of receiver functions. Usually only one phase in the simple stack of receiver functions is used to determine the crustal parameters.

2.7.2 H-κ stacking of receiver functions and state-of-the-art in solving it. H-κ stacking of receiver function has been introduced in 2000 by Zhu and Kanamori (2000) in order to address the very common trade off problem between H and κ. The H-κ stacking of receiver functions reduces the tradeoff ambiguity by bringing in more (a total of three) seismic phases for the determination of the crustal parameters. There has been a need to determine weights associated with the different phases involved in the stacking. Compared to the single phase used to determine the crustal parameters in the case of simple stack of receiver functions, the H-κ stacking combines three phases of receiver functions in estimating the H and κ parameters**.**

For the three phases involved in the H-κ stacking, values for the weights could be assigned using two different ways. In one of those approaches, fixed weights are assigned to all receiver functions from one seismic stations... For example, equal weights have been assigned in many cases (e.g., Julia et al, 2005; Crotwell and Owens, 2005). In the second method, variable weights have been assigned (e.g., Dugda et al., 2005; Jeon et al., 2013), but, it has been time-consuming because the weight assignment needs to be done on a station-by-station basis and human intervention is necessary. Therefore, an optimization method that can be automated looks a natural choice. In the first part of this research, Genetic Algorithms (GA) was used to solve all five parameters simultaneously. This work was published in 2012 (Dugda et al., 2012) and is found to be suitable for automation. Moreover, the Generalized Pattern Search (GPS) and Fitness Proportionate Niching (FPN) techniques that have been developed in this study are also applied to solve all the five parameters simultaneously and they are appropriate for automation too.

2.8 Pattern Search Techniques

Dennis and Torczon introduced a multidirectional search algorithm in 1989 in Torczon's PhD work (Torczon, 1989). That multidirectional search algorithm was considered a first step headed to a general purpose optimization algorithm with promising characteristics for parallel computation. Succeeding work based on the multidirectional search algorithm was then bound for a class of algorithms that allow more flexible computation (Torczon, 2001). One great unexpected accomplishment of that research was a global convergence theorem for the original multidirectional search algorithm (Lewis and Torczon, 2001). Subsequent to the multidirectional search algorithm, the generalized pattern search (GPS) has been developed between the 1990s and early 2000 (Lewis and Torczon, 2001; Lewis, Torczon and Trosset, 2000; Lewis and Torczon, 1997, 1999a, b). The GPS is a direct search optimization technique which does not require the gradient or higher derivative of the objective function to solve the optimization problem. Traditional optimization methods, on the other hand, utilize the gradient or higher derivatives information in their search for an optimal solution. Direct search methods search a set of points around the current point to find a point where the value of the objective function is lower than the value at the current point. The direct methods of pattern search are useful tools when the problem at hand has an objective function that is not differentiable and/or not continuous (Kolda, Lewis and Torczon, 2003; Lewis, Torczon and Trosset, 2000; Conn, Gould and Toint, 1991; Hooke and Jeeves, 1961).

The GPS technique has been applied in this research for solving a seismic optimization problem, especially employing receiver function inversion. The kind of problem attempted to solve here has the potential of having multiple peaks or troughs. Thus, occasionally application of GPS alone may not provide the best or global optimum point as the optimization problem can be trapped in a local optimum point rather than the global optimum point. Thus, it is necessary to have the right starting values in order to find the right solution. The fitness proportionate niching (FPN), a clustering algorithm developed at North Carolina Agricultural & Technical State University (Workineh, 2013), has been implemented in this research work for pinpointing the right initial model in the seismic optimization problem. The optimization problem setup here actually resolves a receiver function inversion problem. Inverse problems are generally difficult, especially when they are compared to their forward counterparts.

2.8.1 Pattern search methods for linearly constrained minimization problems. Lewis and Torczon (1999a, b) proposed to extend pattern search methods for linearly constrained minimization problems. As a result, a general class of feasible point pattern search algorithms was developed and global convergence to a Karush-Kuhn-Tucker point has been proven to hold for such an approach (Lewis and Torczon, 1997). For the case of minimization with general constraints and simple bounds, similar pattern search methods employ augmented Lagrangian (Conn et al., 1999; Lewis and Torczon, 2001). The pattern search methods for linearly constrained cases, just like in the case of unconstrained problems, achieve the searching objective without explicit resort to gradient or directional derivative.

In our particular study, pattern search algorithms are implemented for the following kind of optimization problem with linear constraints:

Minimize

$$f(x) \tag{2.21}$$

subject to

$$\ell \leq Ax \leq u,$$

where f(x) is an objective function and $f: R^n \rightarrow R$, $x \in R^n$, matrix $A \in Q^{m \times n}$, $\ell, u \in R^m$, and upper and lower bounds $\ell \leq u$. Here, the sets of real, rational, integer, and natural numbers are represented by R, Q, Z, and N, respectively. In general, we may allow the possibility that some of the variables can be unbounded either above or below by permitting $\ell_i, u_i = \pm\infty$, $i \in \{1, \cdot\cdot\cdot, m\}$. In this approach, we also admit equality constraints by allowing $\ell_i = u_i$. If the objective function f is continuously differentiable, then a subsequence of the iterate produced by a pattern search method for linearly constrained minimization converges to a Karush-Kuhn-Tucker point of problem (Equation 2.21). We do not attempt to estimate Lagrange multipliers for this implementation. , The pattern of points over which we must search in the worst case scenario will, when we are close to the boundary, conform to the geometry of the boundary (Lewis and Torczon, 1996). The general idea is that the pattern must contain search directions that comprise a set of generators for the cone of feasible directions. We must also take into account the constraints that are almost binding in order to be able to take sufficiently long steps. Though in the bound constrained case it turns out to be simple to ensure, in the case of general linear constraints the situation is more complicated. In practice, pattern search methods are most applicable in cases where there are relatively few linear constraints besides simple bounds on the variables.

Jerrold May (1974) expanded the then existing derivative-free algorithm for unconstrained minimization problems into linearly constrained problems. May proves global convergence and superlinear local convergence for his method. Lewis and Toczon (1998) claim that May’s method is the only other provably convergent derivative-free method for linearly constrained minimization problems introduced before their method. Both May’s approach and the methods described here use only values of the objective function at feasible points to conduct their searches.

We must place additional algebraic restrictions on the search directions since pattern search methods require their iterates to lie on a rational lattice. This requires that the matrix of constraints A in Equation (2.21) be rational. The mild restriction is a price paid for not enforcing a sufficient decrease condition. The i^{th} standard basis vector will be denoted by e_i. Unless otherwise noted, norms are assumed to be the Euclidean norm.

Ω denotes the feasible region for problem given by Equation (2.21) and defined by:

$$\Omega = \{x \in R^n \mid \ell \leq Ax \leq u\} \quad (2.22)$$

2.8.1.1 The pattern. Basically two matrices are required to define a pattern: a basis matrix and a generating matrix. The basis matrix can be any nonsingular matrix $B \in R_{n\times n}$, whereas the generating matrix is a matrix $G_k \in Z_{n\times p}$, where $p > 2n$. We partition the generating matrix into components

$$G_i = [M_i \quad -M_i \quad L_i] = [\Gamma_i \quad L_i] \quad (2.23)$$

We require that $M_i \in M \subset Z_{n\times n}$, where M is a finite set of nonsingular matrices, and that $L_i \in Z_{n\times(p-2n)}$ and contains at least one column, the column of zeros.

A pattern to search P_i is computed by the columns of the matrix $P_i = BG_i$. Because both B and G_i have rank n, the columns of P_i span R_n. For convenience, we use the partition of the generating matrix G_i given in Equation (2.23) can be partitioned P_i as follows:

$$P_i = BG_i = [BM_i \; -BM_i \; BL_i] = [B\Gamma_i \; BL_i] \quad (2.24)$$

Given $\Delta_i \in R$, $\Delta_i > 0$, we define a trial step s_{ji} to be any vector of the form

$$s_{ji} = \Delta_i B g_{ji} \quad (2.25)$$

where g_{ji} is a column vector of $G_i = [g_{1i} \cdot \cdot \cdot g_{pi}]$. Bg_{ji} determines the direction of the step, while Δ_i serves as a step length parameter. At iteration i, we define a trial point as any point of the form $x_{ji} = x_i+s_{ji}$, where x_i is the current iterate.

The pattern for linearly constrained minimization is defined in a way that is only slightly less flexible than for patterns in the unconstrained case. According to Lewis and Torczon (1997), at each iteration the pattern P_i is specified as the product $P_i = BG_i$ of two components, a fixed basis matrix B and a generating matrix G_i that can vary from iteration to iteration. For linearly constrained problems, B can be ignored to be B = I and $P_i = G_i$.

A pattern P_i is a matrix $P_i \in Z_{nx}p_i$, where $p_i > n+1$ and no upper bound for p_i. The generating matrix can be divided into components $P_i = [\Gamma_i \; L_i]$ where $\Gamma_i \in Z_{n\times ri}$ which belongs to a finite set of matrices Γ with certain geometrical properties, and that $L_i \in Z_{n\times(pi-ri)}$ contains at least one column, a column of zeroes. The inclusion of a column of zeroes is simply a formalism to allow for a zero step, i.e., $x_{i+1} = x_i$, and $n + 1 \leq r_i < p_i$.

Given $\Delta_i \in R$, $\Delta_i > 0$, a trial step s_{ji} can be any vector of the form $s_{ji} = \Delta_i c_{ji}$, where c_{ji} is a column vector of $P_i = [c_{1i} \cdot \; \cdot \; \cdot \; c_{pji}]$. The trial step s_{ji} feasible for $(x_i + s_{ji}) \in \Omega$. At iteration i, a trial point is any point of the form $x_{ji} = x_i + s_{ji}$, where x_i is the current iterate.

2.8.1.2 The linearly constrained exploratory moves. A pattern search or GPS method conducts a sequence of exploratory moves about the current iterate x_i to determine a feasible step s_i and a new iterate $x_{i+1} = x_i + s_i$. The hypotheses for linearly constrained exploratory moves shown in Algorithm 2.1, provide a broad choice of exploratory moves while ensuring the properties required to prove convergence. These hypotheses assume the role played by sufficient decrease conditions in quasi-Newton methods.

The choice of exploratory moves must ensure two things:

1. The direction of any step s_i accepted at iteration i is defined by the pattern P_i, and its length is determined by Δ_i.

2. If simple decrease on the function value at the current iterate can be found among any of the 2n trial steps defined by $\Delta_i B \Gamma_i$, then the exploratory moves must produce a step s_i that also gives simple decrease on the function value at the current iterate. In particular,

$f(x_i + s_i)$ need not be less than or equal to $\min\{f(x_i + y), y \in \Delta_i B\Gamma_i\}$. The iterates must be feasible.

1. $s_i \in \Delta_i P_i = \Delta_i [\Gamma_i \; L_i]$.
2. $(x_i + s_i) \in \Omega$.
3. If $\min \{ f(x_i + y) \mid y \in \Delta_i\Gamma_i$ and $(x_i + y) \in \Omega\} < f(x_i)$,

   then $f(x_i + s_i) < f(x_i)$.

Algorithm 2.1. Hypotheses on the result of the linearly constrained exploratory moves.

2.8.1.3 The generalized pattern search (GPS) method. The generalized pattern search (GPS) method for minimization with linear constraints is displayed in Algorithm 2.2. In order to define a particular pattern search method, we must specify the pattern P_i, the linearly constrained exploratory moves for generating a feasible step s_i, and the specific algorithms for updating P_i and Δ_i.

Let $x_0 \in \Omega$ and $\Delta_0 > 0$ be given.

For i = 0, 1, · · · ,

a) Compute $f(x_i)$.

b) Determine a step s_i using a linearly constrained exploratory moves algorithm.

c) If $f(x_i + s_i) < f(x_i)$, then $x_{i+1} = x_i + s_i$. Otherwise $x_{i+1} = x_i$.

d) Update P_i and Δ_i.

Algorithm 2.2. The generalized pattern search (GPS) method for linearly constrained problems.

2.8.1.4 The updates. The rules of updating Δ_i are specified in Algorithm 2.3. The goal of updating Δ_i is to make sure that the objective function f(x) decreases. An iteration is successful if

$f(x_i + s_i) < f(x_i)$; otherwise, the iteration is considered unsuccessful. For any pattern search method to be acceptable, a step needs only produce simple decrease, and not sufficient decrease.

Let $\tau \in Q$, $\tau > 1$, and $\{w_0, w_1, \cdot \cdot \cdot, w_L\} \subset Z$, $w_0 < 0$, and $w_j \geq 0$, $j = 1, \cdot \cdot \cdot, L$. Let $\theta = \tau^{w0}$, and $\lambda_i \in \Lambda = \{\tau^{w1}, \cdot \cdot \cdot, \tau^{wL}\}$.

a) If $f(x_i + s_i) \geq f(x_i)$, then $\Delta_{i+1} = \theta\Delta_i$.

b) If $f(x_i + s_i) < f(x_i)$, then $\Delta_{i+1} = \lambda_i\Delta_i$.

Algorithm 2.3. Updating Δ_i.

The conditions on θ and Λ ensure that $0 < \theta < 1$ and $\lambda_i \geq 1$ for all $\lambda_i \in \Lambda$. Thus, if an iteration is successful it may be possible to increase the step length parameter Δ_i, but Δ_i is not allowed to decrease. The algorithm for updating G_i depends on the pattern search method.

2.9 Genetic Algorithms

A Genetic algorithm (GA) is an adaptation procedure based on the mechanics of natural genetics and natural selection (Goldberg, 1989). The GA discovers new solutions among a population of candidate solutions using mutation and crossover operators. Unlike gradient based search techniques, the GA search is not trapped in local optima since the search is performed over the entire domain in parallel. Basically GA involves three major operations: reproduction, crossover, and mutation. Reproduction is done by mating better solutions from the population. There are different selection mechanisms in the literature (roulette wheel, tournament, elitist and ranking selection) and a comparative analysis of the selection schemes is given in Goldberg and Deb (1991). In this research, a roulette wheel selection method is used to select parents for reproduction. The roulette wheel selection is a probabilistic selection method where the chance of an individual being selected for reproduction is proportional to its fitness. Accordingly, individuals with a higher fitness (objective function value) will reproduce more often and, hence, dominate the population in the coming generations. The crossover operator allows the GA to come up with a better candidate solution by exploiting its experience through genetic combination of existing solutions. The mutation operator gives the power to discover new solutions from the search space, and, hence, provides a mechanism to get out of local optima.

When there is lack of diversity in the population, Genetic algorithms (GAs) can converge prematurely. Mutation and crossover operations play an important role in GAs to discover new solutions as well as accelerate convergence. The mutation assists the GA in the new discovery by getting out of local optima, while the crossover accelerates its convergence by enhancing already acquired solutions. In GAs there should be an equilibrium maintained between exploration and exploitation. High mutation rate means too much exploration and this would destroy good solution and delays convergence. Higher crossover (higher crossover rate) means too much exploitation and this brings about a premature convergence which overlooks perhaps better or even global optimal solutions at the unexplored space (Floudas and Pardalos, 1996; Pardalos and Romeijn, 2002).

GAs can be categorized into a bigger class of population-based search algorithms called Evolutionary Approaches (EAs). Though traditional search methods employing gradients and/or higher derivatives are susceptible to be trapped in local optima, generally Evolutionary Approaches (EAs) utilize the mutation operator to escape such local optima (Floudas and Pardalos, 1996; Pardalos and Romeijn, 2002). On the other hand, no one can warranty that EAs would always find out the global optima every time they are executed. This is because finding a global optima basically rests on various dynamics such as initial population and type of fitness landscape (like uni-modality, multimodality, ruggedness, etc.) (Jong, 1975). The observation and study on H-κ stacking surfaces in this research shows that those surfaces are generally characterized by multimodality.

2.9.1 Genetic algorithms in seismology. In the last two decades, genetic algorithms (GAs) have been successfully applied in various areas of seismology (e.g., Stoffa and Sen, 1991; Sambridge and Drijkoningen, 1992; Shibutani et al, 1996; Chang et al., 2004; Wu et al., 2008). These seismic studies include investigations of crustal velocity structure (e.g., Jin and Madariaga, 1993; Zhou et al., 1995; Yamanaka and Ishida, 1996; Bhattacharyya et al., 1999; Lawrence and Wiens, 2004; Pezeshk and Zarrabi, 2005), studies of mantle velocity structure (Lomax and Sneider, 1995; Curtis et al., 1995; Neves et al., 1996), hypocenter (the exact location of the *earthquake* origin) relocations (e.g., Sambridge and Gallagher, 1993; Billings et al., 1994), and determinations of source parameters (e.g., Yin and Cornet, 1994; Sileny, 1998; Jimenez et al., 2005; Wu et al., 2008). Genetic algorithms (GAs) have been effective techniques for searching parameter spaces to obtain optimal values. The GA uses the mechanics of natural selection and natural genetics in their search and is based on the principle of survival of the fittest (Goldberg, 1989). GAs are population based stochastic search algorithms that start with a random population of candidate solutions and evolve it to a new and possibly better population using crossover and mutation operators. The techniques which GAs employ in their attempt to mimic "natural selection" are the use of roulette-wheel and tournament selection to choose the fittest individuals from a mating pool (Homaifar et al., 1987; Goldberg, 1988).

One of the major goals of this research has been to implement genetic algorithms to determine optimal or near optimal values of weights which are essential to compute the H-κ stacking of receiver functions along with searching for optimal H and κ values. Since its introduction, the H-κ stacking has been applied in many crustal structure studies to determine crustal thickness and Vp-to-Vs ratio,κ, or the Poisson's ratio (e.g., Julia and Mejia, 2004; Dugda and Nyblade, 2006; Buffoni et al., 2012).

2.10 Fitness Proportionate Niching (FPN)

Traditional GAs are appropriate for discovering the optimum of unimodal functions as they converge to a single solution of the search space (Sareni and Krähenbühl, 1998). The H-κ stacking problem in this study, on the other hand, has a very good similarity to a clustering problem. Due to its resemblance, not as a unimodal clustering problem, but as a multimodal clustering problem, it is proposed in here to approach the problem at hand as a clustering problem. Then, each cluster center corresponds to a local or global optimum of the fitness function (objective function). Algorithms that allow the formation and the maintenance of different solutions can be used to solve such multimodal problems. By retaining useful diversity in a population, GAs can avoid early convergence so that they can explore the search space very well and locate multiple optima at the same time (Horn and Goldberg, 1996; Shir and Back, 2006).

Niching methods have been established for the purpose of maintaining population diversity in GAs so that they will not be trapped by a single optimum solution. The fundamental notion of niching methods comes from natural ecosystems which preserve population diversity (Forrest et al., 1993; Goldberg and Richardson, 1987; R. Smith et al., 1993). The diverse population in a GA enables it to investigate many optima in parallel. A typical ecosystem is consisting of various physical niches that display different attributes to be able to allow both the development and the preservation of different types of species. Individuals in a species have similar biological characteristics and are capable of reproducing among themselves, but unable to reproduce with individuals of other species (Chang et al., 2010; Mayr, 1942). By analogy, in artificial systems, a niche corresponds to a local optimum of the fitness/objective function, and the individuals in one niche show similar characteristics in terms of certain metrics. Among niching techniques proposed so far, fitness sharing (FS), or traditional fitness sharing (TFS) as some studies call it, is among the best known and widely applied methods (Workineh, 2013; Goldberg and Richardson, 1987; Deb and Goldberg, 1989; Smith et al., 1992; Forrest et al., 1993; Darwen and Yao, 1996).

Fitness sharing (FS) has been widely applied in multimodal optimization because of the common need of multiple optima in real-world optimization problems (Sareni and Kräahenbühl, 1998). In fitness sharing (FS), the shared fitness of individual i with an actual fitness value of f_i is given by

$$f_i^{'} = \frac{f_i}{m_i} \qquad (2.26)$$

where m_i is the niche count that measures the approximate number of individuals with whom the fitness is shared. The niche count is obtained by summing a sharing function over all members of the population

$$m_i = \sum_{j=1}^{N} sh(d_{i,j}) \qquad (2.27)$$

where N represents the population size and $d_{i,j}$ represents the distance between individual i and individual j. The sharing function (*sh*) measures the similarity level between two individuals in the population and the most widely used sharing function is:

$$sh(d_{i,j}) = \begin{cases} 1 - (d_{i,j}/\sigma_\sigma)^\alpha \\ 0 \end{cases} \qquad (2.28)$$

By decreasing the payoff in densely populated regions, FS (TFS) can modify the search site. Thus, TFS cuts the fitness of an individual in the population by an amount nearly equal to the number of similar individuals in the population. With such a strategy, TFS keeps population diversity and allows the exploration of many peaks (optima) in a given feasible domain.

Fitness sharing has been widely used for clustering data and for finding multiple optimal solutions (Chang et al., 2010; Sareni and Kräahenbühl, 1998). Over the last few decades, clustering has turned out to be an essential part of data mining. The main goal of clustering is to partition a given set of data or objects into clusters or subsets depending on some patterns so that objects in one cluster will have some resemblance to each other in some sense. Based on the kind of metric (genotypic or phenotypic similarity) we define, the distance between individuals is the norm by which we categorize to one group or to the other. Genotypic similarity is associated with bit-string representation and it is mostly the Hamming distance. On the other hand, phenotypic similarity is interrelated to real parameters of the search space and it can be the Euclidian distance.

In FS or TFS, the fitness is the resource for which the individuals in the same niche group are competing (Goldberg and Richardson, 1987). However, TFS should define a similarity metric on the search space and an appropriate niche radius, and usually providing an operative value for the niche radius without any a priori knowledge is very challenging. Deb and Goldberg (1989) proposed a criterion for estimating the niche radius given the heights of the peaks and their distances (Deb and Goldberg, 1989). Since in most of the real applications there is little prior knowledge about the fitness landscape, it is difficult to estimate the niche radius.

Clustering analysis has become a central problem in data mining and its applications are crossing various disciplines (Tou and Gonzalez, 1974; Jain and Dubes, 1988). A clustering algorithm based on dynamic niching with niche migration (DNNM-clustering) is proposed by Chang et al. (2010). DNNM-clustering introduces a dynamic niching with niche migration to preserve the diversity of the population. DNNM-clustering is claimed to be robust to noise and cluster volumes. The niche migration is introduced to overcome the dependence of the niche radius. In DNNM-clustering, all the niches presented in the population at each generation are automatically and explicitly identified. The application of TFS is then limited to individuals belonging to the same niche.

Another new approach that overcomes the two major challenges of prior knowledge of number of niches and dependence on the niche radius has been proposed and it is called fitness proportionate niching (FPN) (Workineh, 2013). Fitness Proportionate niching (FPN) is a new niching approach in GA when we don't have a prior knowledge on the number of niches. Like the traditional fitness sharing (TFS) based on niche counts, the FPN technique is based on the notion of limited resources where individuals in a given niche share the resource of that niche. However, resource sharing in FPN is in proportion to strength (fitness value) (Workineh, 2013; Workineh and Homaifar, 2012).

According to the FPN sharing scheme, the sharing function and the derated fitness of individual i are given by the following equations, respectively:

$$sh(d_{i,j}) = \begin{cases} F_{j,} & if \ d_{i,j} < \sigma_{sh} \\ 0, & otherwise \end{cases} \tag{2.29}$$

$$F_{sh,i} = \frac{F_i}{\sum_{j=1}^{M} sh(d_{i,j})} \tag{2.30}$$

where M is the number of individuals in a given niche, $d_{i,j}$ is the distance between individuals i and j, $sh(d_{i,j})$ is the sharing function which measures the similarity between the two individuals, and σ_{sh} is the niche radius.

CHAPTER 3

Proposed Methodology

One major goal of this study has been to harness genetic algorithm (GA) to determine optimal or near optimal values of weights that will be used to compute the H and κ crustal parameters along with the optimal values of H and κ. The GA has been shown to be a useful tool to obtain those optimal weights and crustal parameters simultaneously (Dugda et al, 2012). Another objective of this research is to implement a Generalized Pattern Search (GPS) technique to determine those optimal values of crustal parameters as well as the weights and to verify the performance of these different approaches. In addition, we consider the H-κ stacking problem as a multimodal problem for the first time. Thus, the Fitness Proportionate Niching (FPN) has been proposed and implemented to operate with GPS whenever the GPS method is trapped in local optima instead of a global optima point.

The implementation of both the evolutionary (GA and FPN) and direct (GPS) techniques for solving the receiver function inversion problem in this dissertation is to pave a way to the development of a complete automatic receiver function inversion system. These techniques don't involve any assumptions about weights involved and they can be applied for a number of seismic stations at the same time, which have not been done before.

3.1 H-κ Stacking of Receiver Functions

For the implementation of the H-κ stacking, the arrival times for three important converted phases in the receiver functions, the P-to-S converted phase (Ps) and the first two reverberations of P-to-S converted phases (PpPs and PsPs+PpSs), need to be determined (see

Figure 2.4(b)). These arrival times can be obtained from the crustal thickness (H) and V_p-to-V_s ratio (κ) (equations 2.21-2.23).

3.1.1 Fixed weights vs variable weights. Many researchers have attempted to solve this H-κ stacking of receiver function problem just one seismic station at a time. Over the last decade, different approaches have been utilized for estimating reasonable values of weights to be used in the H-κ stacking. For weights assignment or computation, the different approaches introduced so far can be categorized into two categories: fixed weights and variable weights approaches.

3.1.1.1 Fixed weights. Please take a look at Figure 2.4(b). Almost equal weights for all the three receiver function phases (0.34, 0.33, 0.33) have been apportioned by Crotwell and Owens (2005) for weights w_1, w_2, and w_3, respectively. Equal weights of 0.5 for any two of the three phases, usually for the first two phases, has been assigned (e.g., Julia et al, 2005). Normally when one or two of the receiver function phases are clear and the other(s) are not, assigning equal weights doesn't seem a reasonable solution or approach. We may illustrate this point in Figure 3.1 with a receiver function (Dugda et al., 2005).

Figure 3.1 displays an example of a receiver function from a seismic station CHEF in Ethiopia (Dugda et al., 2005). In Figure 3.1, the peaks of Ps and PsPs+PpSs phases are clear in the receiver function. However, the exact timing (location) of PpPs phase is not clear, in this case it forms a plateau. Allocating equal weights to the different phases, as in the aforementioned fixed weights approaches does not seem a good idea.

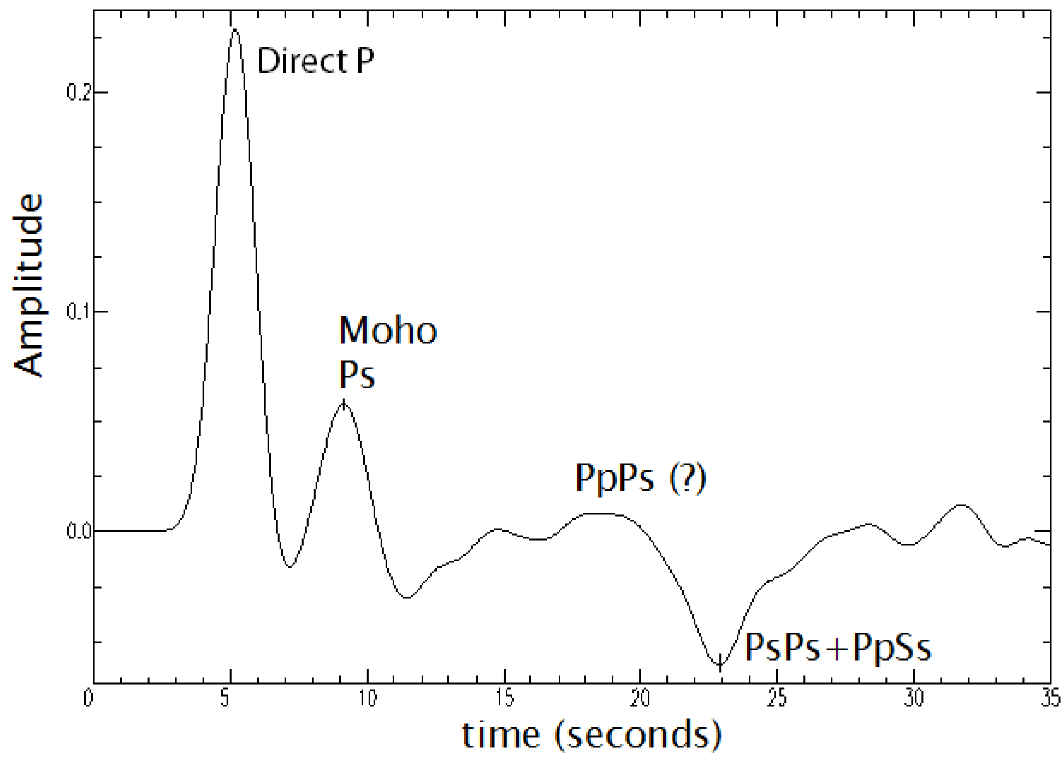


Figure 3.1 An example of a receiver function with two distinct and one indistinguishable phases

3.1.1.2 Variable weights. Different combinations of weights based on the quality of receiver functions and clarity of phases could be a better solution for such a problem. Variable weights approach using Monte Carlo technique has been introduced in 2005 (Dugda et al., 2005; Dugda and Nyblade, 2006) and some studies are still following similar variable weights approaches (e.g., Jeon, 2013; Moidaki et al., 2013). One disadvantage of applying such variable weights approach was that it has been time-consuming. This is because appropriate weights had to be searched on a station-by-station basis for every set of receiver functions and at the completion of each search the picking of the phases has to be verified for its correctness.

In the past, a variable weights approach was introduced that uses a Monte Carlo technique to find the variable weights for H-κ stacking of receiver functions (Dugda et al., 2005).

The earlier technique has been implemented to solve one seismic station at a time, and the confirmation of the picking of the phases was required to determine the correctness of the corresponding phase picking. The search method used to determine the best crustal parameters was focused on uni-modal portions of the objective function at a time. Therefore, the technique we introduced then was a new method but it was time-consuming compared to a fixed weights approach. The Monte Carlo technique was applicable to one station at a time and it was not suitable for automatic implementation. The techniques developed in this dissertation are variable weights approaches with new capabilities such as being appropriate for automatic application.

In this study, the techniques developed involve evolutionary and direct optimization approaches and they are used to invert receiver functions for determining crustal parameters and the three associated weights simultaneously. The first technique developed here makes use of the GA optimization technique. This H-κ stacking of receiver function inversion is shown to be a multimodal problem and a second technique involving both direct and evolutionary techniques has been developed to solve this problem. The second technique introduced in this research combines the direct GPS and evolutionary FPN technique by employing their strengths. Compared to the previous Monte Carlo variable weights approach introduced in the previous study, the current GA and GPS-FPN techniques have remarkable advantages of saving time and these new techniques are suitable for automatic and simultaneous determination of the crustal parameters with variable weights.

In this dissertation the first part of the study involves the GA and it was published in a leading seismology journal (Dugda et al., 2012). The GA introduced here solves for all the five parameters (the two model parameters and the three weights) simultaneously. Such a GA implementation has taken care of the variable weights approach without being protracted. Thus, an optimization method like the GA that may be automated looks a natural choice for a complete solution of this problem, without assuming equal weights or equally identifiable phases. The implementation of GPS with FPN is another alternative variable weights approach which will also take care of the weights as variables to be determined by the optimization technique as in the GA which reduces the painstaking effort of applying variable weights.

3.1.2 The inverse problem of finding model parameters with best fitting data. For an inverse problem, we determine the model parameters which are best fit the observation (the observed data). In the current receiver function inversion problem the important parameters to be utilized in the receiver functions (which are actually processed seismic signals/data) are relative arrival times of the P-to-S converted phases and their reverberations within the crust, and the receiver function amplitudes. Thus, what we are accomplishing in this research for the inversion process is finding the model parameters that would result in the best fitting of the three relative arrival times for the set of receiver functions in each seismic station.

The weighted sum objective function used in this study has been developed on the premises that when the right model parameters H and κ are picked they would lead to the picking of the best fitting arrival times. Then the weighted sum of the receiver function amplitudes at the rightly picked arrival times would produce the highest value of the objective function. Therefore, the objective function can be calculated directly from the weighted summation of the receiver function amplitudes.

3.1.3 Inversion of H-κ stacked receiver function. To discuss the receiver function inversion process, we consider an inversion implementation result using seismic data from seismic station ARBA of Ethiopia. Figure 3.2 shows two inversion results, one using the GA of this study and the other result by an exhaustive search for the limited search space with assumed values of weights (w_1=0.5, w_2=0.4, w_3=0.1). The weights used in the exhaustive search are the same as the weights determined in a 2005 study for comparison purposes (Dugda et al., 2005). The horizontal axis represents crustal thickness H and the vertical axis denotes the Vp/Vs ratio (κ). The highest of the objective function S occurs at the center of the yellow ellipse.

The color scale variations in the figure show the value of the objective function. Thus, the red places indicate higher values of the objective function while the blue places show smaller values of the objective function. The white star and blue circle indicate the positions of optimal values for H and κ using exhaustive search and GA, respectively. The contours display 10% reductions in values of the objective function value with respect to the highest objective function value which occurs at the center of the yellow ellipse. The yellow ellipse is an error ellipse representing twice the standard deviation in the estimates of H and κ. The internal contour surrounding the white star and the blue circle is at 90% of the highest value of the objective function. The next successive contours show 80%, 70%, etc., of the highest value.

With the exhaustive search, we are considering the variations of only H and κ since the weights have assumed certain values. What the exhaustive search technique does is attempt to solve for only two parameters by computing all the objective function values in the plausible H-κ parameter space. This is usually a very time-consuming procedure. The process of receiver function inversion with the given weights is to attempt to find the best (H, κ) 2-tuples that would fit the observed receiver function data. For each possible (H, κ) point, we compute the three relative arrival times in the receiver function for each of the points in the H-κ parameter space using Equations (2.21-2.23). Then, we obtain the amplitudes of all the receiver functions at those computed arrival times in the data set. Next, the objective function values are computed for each one of the points in the H-κ parameter space. Figure 3.2 gives an example where the surface of objective function values on the H-κ parameter space looks simple. However, even for a simplified case of fixed weights, a fairly complicated objective function surface with multiple optimum points will be obtained when the H-κ parameter space is somewhat expanded. This will be discussed further in section 3.4.1.

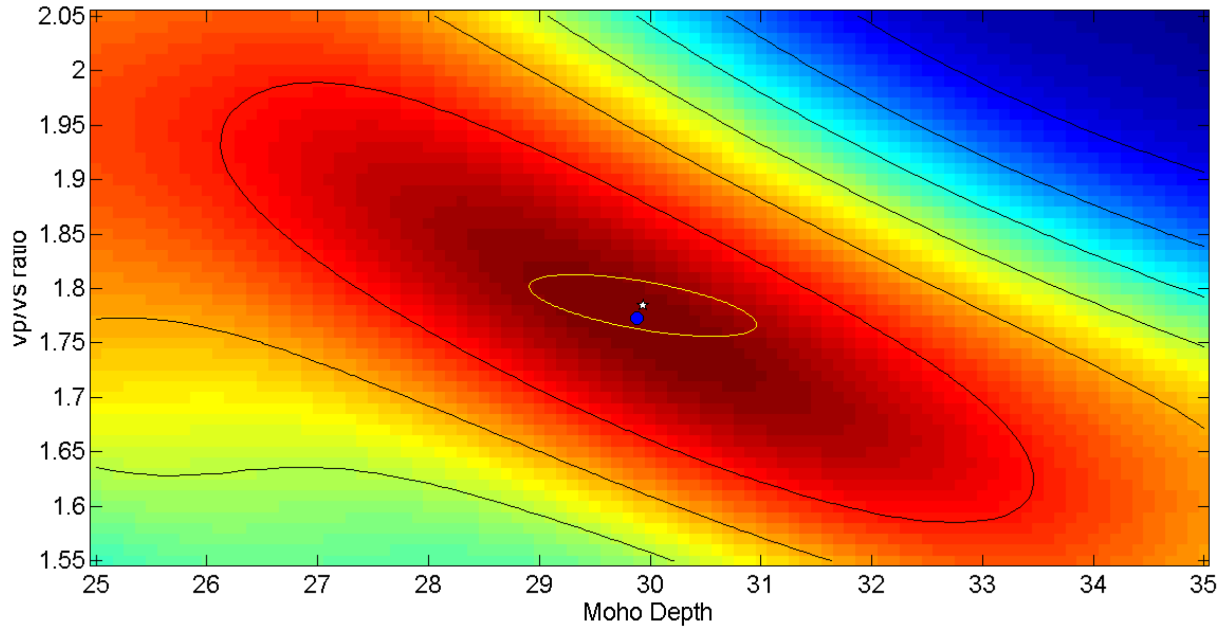


Figure 3.2 An example of H-κ receiver functions inversion process with station ARBA data

3.2 Genetic Algorithms for Optimal H-κ Inversion

Genetic Algorithms (GA) are powerful tools for finding optimal solutions from a given population of candidate solutions. Traditional GAs are fit for locating an optimum of unimodal function. GAs utilize mutation and crossover operators applied at different locations within selected string of parameters. The next subsection entails the procedure of how the lengths of the different parameters are selected in this study.

3.2.1 Implementation of GA

3.2.1.1 String lengths for the GA implementation. One of the intermediate parameters that need to be determined in the implementation of GA is the string length required to represent each of the parameters under investigation. The limits on the parameters and the precision required dictate the number of binary bits by each parameter. The number of binary bits for each parameter is directly related to the precision desired which is given by:

$$\Delta_i = \left(\frac{X_{max} - X_{min}}{2^{b_i} - 1} \right) \qquad (3.1)$$

where Δ_i is the precision sought, X_{max} and X_{min} are the maximum and minimum possible values, and b_i is the string length for the i^{th} parameter. The crustal thickness (H), Vp/Vs ratio (κ), and the weights necessary for stacking the receiver functions (w_1, w_2, and w_3) are determined through GA implementation.

Based on previous studies (e.g., Zhu and Kanamori, 2000; Julia and Mejia, 2004; Dugda et al., 2005), the precisions for the crustal thickness H, Vp-to-Vs ratio (κ), and the three weights can be taken as 0.5 km, 0.01, and 0.01, respectively. The string lengths or the number of binary bits 'b_i' which is required to obtain a certain precision can be calculated using following equation:

$$b_i = \log_2 \left[\left(\frac{X_{max} - X_{min}}{\Delta_i} \right) + 1 \right] \qquad (3.2)$$

For representing crustal thickness, the string length can be calculated as follows:

The practical limits on crustal thickness parameters, especially for the East African region are:

X_{1max} =max H (maximum crustal thickness)

=60 km, and

X_{1min} =min H (minimum crustal thickness)

=10 km,

Using the above parameters and precision Δ_1=0.5 km, b_1=6.7. Since the string length is an integer value, the next higher integer value for b_1=7. Similarly, string length for the Vp-to-Vs ratio (κ) is 6.2; string length for the all the three weights w_1, w_2, w_3 is 6.7. In all cases the next higher integer values are taken as shown in Table 3.1.

Table 3.1

String lengths used in this paper for representing crustal thickness, Vp-to-Vs ratio (κ), and weights w_1, w_2, w_3 in the GA implementation.

Parameter	Parameter Precision (Δ)	X_{max}	X_{min}	String length (b)
Crustal thickness	0.5 km	60 km	10 km	6.7 =>7
Vp-to-Vs ratio (κ)	0.01	2.2	1.5	6.2=>7
Weights (w_1, w_2, w_3)	0.01	1.0	0.0	6.7=>7

Table 3.1 provides the assumed precision, minimum and maximum values, and the corresponding string lengths of the parameters in this study. The average crustal P-wave velocity (Vp) for the Main Ethiopian Rift in which station ARBA is situated is 6.4 km/s and this value is used in our implementation (Dugda et al., 2005).

3.2.1.2 Inclusion of the constraint in the objective function. In the GA implementation of the H-κ stacking, we integrated the constraint into the objective function. From the constraint equation (Equation (2.3)) we could write one of the variables in terms of the other two variables. Thus, our new objective (fitness) function is:

$$S(H,\kappa) = \sum_{j=1}^{N} w_1 r_j\big(t_1(H,\kappa)\big) + w_2 r_j\big(t_2(H,\kappa)\big) + (1 - w_1 - w_2) r_j\big(t_3(H,\kappa)\big) \qquad (3.3)$$

The inclusion of the constraint into the objective function allows us to effectively search for four variables instead of five. Thus, the search space for the GA involves H, κ, w_1, w_2. Once the GA determines optimal values for those four variables, the value for third weight w_3 will be determined from the other two weights using the constraint equation. Any one of the weights can be written in terms of the other two.

The H-κ receiver function stacking algorithm using GA has been implemented in Matlab. First, a population is initialized, and limits on variables and search spaces are defined. Second, the receiver functions are read for a crustal structure analysis. At each generation, the whole population is updated and the process continues until there is no improvement. At each generation, the H-κ stacking or objective function values are also calculated based on the offspring or candidate solution. The candidate solutions are selected using the roulette wheel technique based on the values of the objective function. From those selected individuals or candidate solution called "parents" the offsprings are obtained after applying crossover and mutation operators. This process continues until optimal or near optimal values of weights, H and κ are obtained. Figure 3.3 displays the flow diagram of the GA implementation in this study.

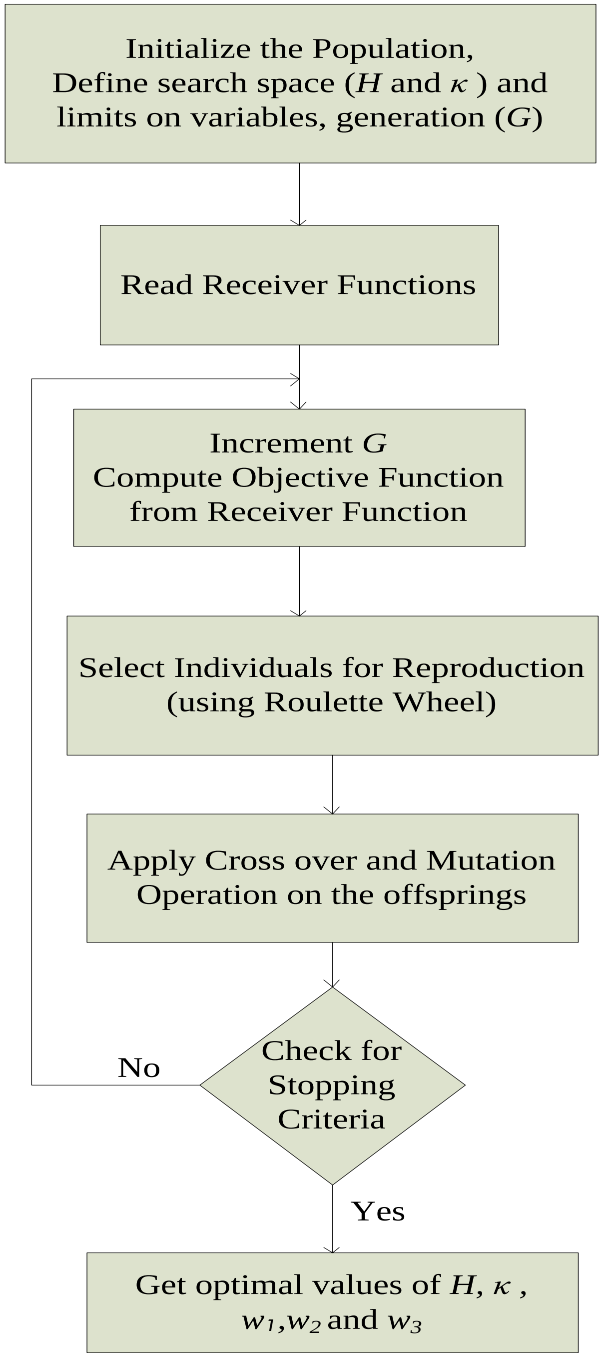


Figure 3.3 Flow diagram for the implementation of GA for H-κ stacking

3.3 Generalized Pattern Search Technique for Optimal H-κ Inversion

Generalized pattern search (GPS) is a direct search method that does not need to compute the derivative of the objective function. GPS technique has significant resemblance to steepest descent method, but unlike steepest descent scheme the GPS does not require the computation of derivatives or directional vectors. H-κ stacking optimization equation has five parameters to be determined, if we would like to solve the problem wholly, without any assumed value for those parameters. In this case, we apply the Generalized Pattern Search (GPS) technique to solve the H-κ stacking optimization problem entirely. Just like in the case of the GA implementation, we can solve for all the five parameters using GPS technique.

For objective functions which are not differentiable or not continuous, GPS is an appropriate approach for optimization. It can be used to optimize directly all the five parameters in the problem. GPS as a direct search technique is used for solving optimization problems with no information about the gradient of the objective function. Unlike the more traditional optimization methods that use information about the gradient or higher derivatives to search for an optimal point, a direct search algorithm searches for a set of points surrounding a given point, then it looks for one where the value of the objective function is lower than the value at the current iterate point (Note that we are considering a minimization problem). In general direct search techniques are applied to solve problems for which the objective function is neither differentiable nor continuous.

3.3.1 Advantages of the GPS technique for H-κ inversion. Application of GPS on H-κ stacking of receiver functions enables to almost exhaustively search for the weights w_1, w_2, w_3 as well as H and κ within the given parameter space. GPS results are repeatable. GPS provides highly repeatable outputs, especially compared to heuristic methods which may need more runs to provide similar results.

The repeatable results are important in our research because the repeatability characteristics would help us check the dependence of the inversion on initial model parameters. If the results are independent of the initial values, we should obtain almost identical results for any initial model. GPS delivers repeatable results, even if it starts at different initial conditions. Thus, the repeatable behavior can be harnessed to use GPS to test initial value dependence of final parameters.

3.3.2 A drawback of GPS and its FPN solution. One major drawback/challenge of GPS is that it may be trapped in a local optimum point. The solution may depend on our initial values. One solution for this is to apply complete polling instead of partial polling. However, complete polling may not provide the global optimum point all the time. When complete polling is not sufficient enough to obtain the global optimal solution, GPS needs support from heuristic techniques such as FPN. FPN can be used to identify the best initial values, which are close to the optimal solution. We may not know how many local minima we have in the search space and where the local optima are located.

For the specific receiver function inversion, the solution parameter space is bounded. The bounded parameter space characteristics makes our problem fit for dynamic identification of niches within Fitness Proportionate Niching (FPN) to automatically evolve the optimal number of niches. Thus, within the bounded region of parameters, we propose to apply Fitness Proportionate Niching (FPN) to identify the right starting parameters which will lead to the right optimal solution. FPN is able to automatically evolve dynamically and find niches or clusters. Those identified niche masters (cluster centers) represent approximate locations of potential local and global optimum points. Among the potential optimum points, we can identify the best or optimal solution. Only few iteration results from FPN are sufficient enough to determine the best initial parameters.

We implemented GPS as an automatic technique to solve the problem for a number of stations at the same time. After the GPS inversion is executed for a set of stations, we use the GPS again to test the final model to check any sensitivity to the initial model. If the inversion is found to be sensitive to initial models for some specific seismic stations (i.e., if the objective function has different local and global optima in the given H-κ parameter space), we invoke fitness proportionate niching (FPN) to identify the right initial model.

3.3.3 The proposed minimization problem for GPS implementation and modified bounds of weights. The H-κ stacking hypothesis has been developed based on exploiting the fact that the amplitudes of the seismic phases in the receiver functions picked will attain their maximum values when the correct H and κ values are chosen. The H-κ stacking takes advantage of the Signal to Noise ratio (SNR) improvement with employing more receiver functions. Thus, we maximize the stack of receiver functions which identify the correct H and κ values as well as the right set of weights which are appropriate to the quality of receiver functions. Originally, the pattern search algorithm has been developed for a minimization problem. There are two ways to implement the GPS algorithm for our specific problem. We have to either modify the GPS algorithm to work for a maximization problem, or else change our problem from maximization to a minimization problem. It is at ease to convert the maximization problem of H-κ stacking into a minimization problem. Since the expected values of the H-κ stacking are either zero or positive, the maximum values of the H-κ stacking are always positive. Multiplication of the H-κ stacking values by -1 will turn the data upside down. Then, the maximum value changes to a minimum value and the minimum value to a maximum. Thus, this will be equivalent to the minimization problem. Therefore, we are able to find the optimal values for H, κ, w_1, w_2, and w_3 using GPS technique.

Since each of the three phases has to contribute for the stacking, the minimum weight each phase can contribute is 0.1. The maximum weight of the three phases will be 0.8 as the total must be 1. Therefore, these new bounds are used for the upper and lower bounds of weights for the GPS implementations. When the original maximization problem with objective function S is changed to a minimization problem, we may assign this new minimization variable $\bar{S}$. Thus our new objective function for GPS implementation is:

Minimize

$$\bar{S} = -\left[\sum_{j=1}^{N} w_1 r_j\big(t_1(H,\kappa)\big) + w_2 r_j\big(t_2(H,\kappa)\big) - w_3 r_j\big(t_3(H,\kappa)\big)\right] \quad (3.4)$$

where r_j is the j^{th} receiver function amplitude at the particular expected arrival time, N is the number of receiver functions for the seismic stations, and t_1, t_2 and t_3 are arrival times related to H and κ using equations (2.21-2.23).

Or we can minimize the following:

Minimize

$$\bar{S}(x) = -\left[\sum_{j=1}^{N} x_3 f_1(x_1, x_2) + x_4 f_2(x_1, x_2) - x_5 f_3(x_1, x_2)\right] \quad (3.5)$$

where f_1, f_2, and f_3 are functions relating t_1, t_2, and t_3, respectively, to x_1 = H and x_2= κ; and x_3= w_1, $x_4 = w_2$ and $x_5 = w_3$ are weights which is subject to:

Subject to

$$x_3 + x_4 + x_5 = 1 \quad \text{(Linear Equality Constraint)} \quad (3.6)$$

And using the following bounds:

$$0.1 \le x_3 \le 0.8 \tag{3.7}$$

$$0.1 \le x_4 \le 0.8 \tag{3.8}$$

$$0.1 \le x_5 \le 0.8 \tag{3.9}$$

The average crustal P-wave velocity (Vp) for the Main Ethiopian Rift in which station ARBA is situated is 6.4 km/s (Mackenzie et al., 2005; Makris and Ginzburg, 1987) and the value is used for the implementation.

The H-κ receiver function stacking algorithm using GPS has also been implemented by Matlab. The algorithm for the implementation is given in the next subsection (section 3.3.4). The receiver functions and parameters such as times and amplitudes of the receiver functions, given seismic station are read for a crustal structure analysis. At each iteration, the H-κ stacking or objective function values are calculated.

3.3.4 Algorithm 1: The general pattern search (GPS) for linearly constrained problems

a) Compute $f(x_i) = -S$.

b) Determine a step s_i defined by the pattern P_i and its length is determined by Δ_i. ($s_i = \Delta_i P_i$ and $(x_i + s_i) \in \Omega$, where Ω denotes the feasible region for the problem.)

c) If $f(x_i + s_i) < f(x_i)$, then $x_{i+1} = x_i + s_i$. Otherwise $x_{i+1} = x_i$.

d) Update P_i and Δ_i.

Updating Δ_i:

a) If $f(x_i + s_i) < f(x_i)$, then $\Delta_{i+1} = \lambda_i \Delta_i$.

b) If $f(x_i + s_i) \ge f(x_i)$, then $\Delta_{i+1} = \theta \Delta_i$.

c) $0 < \theta < 1$ and $\lambda_i \ge 1$.

In the GPS algorithm implementation, the first step in the procedure is to convert our maximization problem into a minimization problem. Next, we select step size s. Computation of $f(x_i + s_i)$ and comparison to $f(x_i)$ follows at iteration i. If $f(x_i + s_i) < f(x_i)$, the search will be considered successful at iteration i, and then we update the iterate for next iteration:

$$x_{i+1} = x_i + s_i . \qquad (3.10)$$

If the search cannot provide an objective function value which is not less than the value at the current iterate, then the search will be considered unsuccessful at iteration i, and we continue the iteration:

$$x_{i+1} = x_{i.} \qquad (3.11)$$

On the next step, the step length control parameter Δ_i at iteration i will be determined. Δ_i can take either of two values: when the search is successful, it will attain

$$\Delta_{i+1} = \lambda_i \Delta_i, \qquad (3.12)$$

where λ_i has a value greater than 1. Otherwise, when the search is unsuccessful, and

$$\Delta_{i+1} = \theta \Delta_i \qquad (3.13)$$

where θ takes a value between 0 and 1. θ is called a contraction factor while λ is an expansion factor.

3.3.5 Generalized pattern search (GPS) algorithm using pseudo code. Let Δ_i be the step length control parameter, and Δ_{tol} be the tolerance used to test for convergence. Assume that the algorithm starts with an initial guess x_0 that has a finite function value and an initial step length $\Delta_0 > \Delta_{tol}$. Then, the GPS technique can be conveyed using a pseudo code as follows:

1: Choose generating set G, for example, $G = \{e_1, e_2, . . . , e_n, -e_1, -e_2, . . . , -e_n\}$

2: Choose Δ_0

3: **for** i = 1, 2, . . . **do**

4: **if** there exists $g_i \in G$ such that $f(x_i + \Delta_i g_i) < f(x_i)$

5: then

6: Set $x_{i+1} = x_i + \Delta_i g_i$ (update the iterate)

7: Set $\Delta_{i+1} = \Delta_i$ (no change to the step length control parameter)

8: **else**

9: Set $x_{i+1} = x_i$ if $f(x_i + \Delta_i g_i) \geq f(x_i)$ for all $g_i \in G$; (do not update the iterate)

10: Set $\Delta_{i+1} = ½ \Delta_i$ (contract the step length control parameter)

11: **if** $\Delta_{i+1} < \Delta_{tol}$

12: **then**

13: GPS algorithm has converged

12: **end if**

13: **end if**

14: **end for**

Algorithm 3. The Generalized Pattern Search (GPS) Algorithm

3.3.6 How the GPS implementation works. The Generalized Pattern Search (GPS) optimization procedure is suitable to solve different kinds of optimization problems as well as the standard optimization methods. Generally, GPS is reasonable and efficient in computation. Unlike heuristic algorithms, such as genetic algorithms (Goldberg, 1989; Michalewicz, 1996), GPS is a flexible technique to improve and adapt the global and fine tune local searches.

The Generalized Pattern Search (GPS) algorithm ensues by computing a succession of points to approach the optimal value. In GPS, the objective function value should decrease or stay the same from one point to the following point in the succession. The algorithm commences by establishing a set of points which form a mesh around an initial point. The starting point could either be provided or computed from a previous step of the algorithm. The mesh is actually constructed by adding the current point to a scalar multiple of a set of vectors known as patterns (pattern vectors). If a point in the mesh is found to improve the objective function at the current point, the new point becomes the current point at the next iteration.

3.3.7 The GPS strides for the H-κ stacking inverse problem. We can apply GPS on H-κ stacking of receiver functions, and GPS operates as if it searches the parameter space exhaustively for the three weights w_1, w_2, w_3 as well as for H and κ parameters. GPS does not do exhaustive search, but it searches along a given pattern in matrix.

First, the search starts at an initial point X_0 which will be usually be provided by us or by the user. At the first iteration, with a scalar equal to 1 (which is the mesh size), the pattern vectors are constructed to solve the problem. In the specific GPS implementation for receiver function inversion, we set B = I, and the pattern (direction) vectors are then:

P_1 = [1 0 0 0 0]; P_2 = [0 1 0 0 0 0]; P_3 = [0 0 1 0 0]; P_4 = [0 0 0 1 0]; P_5 = [0 0 0 0 1];
P_6 = [-1 0 0 0 0]; P_7 = [0 -1 0 0 0 0]; P_8 = [0 0 -1 0 0]; P_9 = [0 0 0 -1 0]; P_{10} = [0 0 0 0 -1].

One initial value set for the Main Ethiopian Rift area, or even for the whole region including the Ethiopian Plateau, is X_0 = [40 1.75 0.34 0.33 0.33]. The GPS algorithm adds these direction vectors to the initial point X_0 to compute the following new mesh points:

$X_0 + P_1$;	$X_0 + P_2$;	$X_0 + P_3$;	$X_0 + P_4$;	$X_0 + P_5$;
$X_0 + P_6$;	$X_0 + P_7$;	$X_0 + P_8$;	$X_0 + P_9$;	$X_0 + P_{10}$

The algorithm polls the mesh points by computing their objective function values until it finds the one with a value smaller than the objective function value of X_0. If there is such a point, the poll is successful and the algorithm sets this point as equal to X_1.

After a successful poll, the algorithm steps to the second iteration and multiplies the current mesh size by 2 (which is the mesh expansion factor). Multiplying the mesh size by 2 will lead to a faster convergence. The mesh at iteration 2 of the implementation contains the following points:

$2*P_1+X_1$; $2*P_2+X_1$; $2*P_3+X_1$; $2*P_4+X_1$; $2*P_5+X_1$;
$2*P_6+X_1$; $2*P_7+X_1$; $2*P_8+X_1$; $2*P_9+X_1$; $2*P_{10}+X_1$;

If the poll is successful and the algorithm sets this point as equal to X_2.

Next, if iteration 3 (mesh size = 4) ends up being an unsuccessful poll, i.e. none of the mesh points have a smaller objective function value than the value at X_2, the algorithm does not change the current point at the next iteration, that is, $X_3 = X_2$. At the next iteration, the algorithm multiplies the current mesh size by 0.5, a contraction factor (mesh reduction factor), so that the mesh size at the next iteration is smaller. The algorithm then polls with a smaller mesh size.

The GPS algorithm will repeat the above demonstrated steps until it discovers the optimal solution for the minimization of the objective function. The x stopping criteria can be one or combination of some of the conditions listed below:

- The mesh size is less than mesh tolerance.
- The number of iterations performed by the algorithm reaches a predefined value.

• The total number of objective function evaluations performed by the algorithm reaches a pre-set maximum number of function evaluations.

• The distance between the point found at one successful poll and the point found at the next successful poll is less than a set tolerance.

• The change in the objective function from one successful poll to the next successful poll is less than a function tolerance.

The occurrence of any one of the criteria conditions given above was sufficient to stop the algorithm from searching further. All the parameters can be pre-defined subject to the nature of the problem. Based on previous studies for the specific region and some global studies such as CRUST2.0 (Bassin et al., 2000; Zandt, G., & Ammon, C. J. (1995), there are some suggested ranges for the variables in the particular problem using the available seismic data for the Ethiopian Plateaus and the Main Ethiopian Rift region: x_1 = H ranges between about 20 and 45 km; x_2 = Vp/Vs ranges between about 1.65 and 1.95; $x_3 = w_1$ ranges between 0.1 and 0.8; $x_4 = w_2$ ranges between 0.1 and 0.8; $x_5 = w_3$ ranges between 0.1 and 0.8.

The updating parameters are permitted to take some commonly used values of $\theta = 0.5$ and $\lambda = 2.0$.

3.3.8 Algorithm 2: testing the inversion with GPS. Algorithm 2 is basically applying Algorithm 1 for at least two different extreme initial conditions. We take combinations of extreme values as initial values, especially, for H and V_p/V_s, in the H-κ parameter space and run the GPS algorithm. For H ranging between 20 and 45 km, and for κ = Vp/Vs ranging between 1.65 and 1.95, the combination of extreme values means (20, 1.65) and (45, 1.95) which can be used as initial values.

3.4 Fitness Proportionate Niching (FPN)

Over the last few decades, clustering has become the essential part of data mining. The main goal of clustering is to partition a given set of data or objects into clusters or subsets depending on some patterns so that objects in one cluster will have some resemblance to each other in some sense. Based on the kind of metric (genotypic or phenotypic similarity), the distance between individuals is the norm by which we categorize to one group or to the other. Fitness sharing has been used for clustering data and for finding multiple optimal solutions.

Traditional fitness sharing (TFS) has been widely applied in multimodal optimization because of the fact that real-world optimization problems often necessitate the determination of multiple optima in a search space. By decreasing the payoff in densely populated regions, TFS can modify the search landscape. Thus, TFS drops the fitness of an individual in the population by an amount nearly equal to the number of similar individuals in the population. By doing so, TFS maintains population diversity and allows the exploration of many peaks (optima) in a given feasible domain.

Fitness Proportionate Niching (FPN) has been developed as a dynamic clustering algorithm which is brought about by keeping population diversity in the GA execution. In FPN application, no a priori knowledge of the approximate number of clusters or any information about the distribution is necessary. Therefore, it is suitable for our receiver function inversion problem.

3.4.1 H-κ receiver function stacking as a multimodal optimization problem. Figure 3.4 is a contour plot of H-κ objective function values for seismic station ARBA. This plot demonstrates the variation of the H-κ receiver function stack for a wide range of H and κ parameters. This H-κ surface on the κ versus H axes confirms that there are some local optima points together with a global optimum point. Thus, the problem is a multimodal optimization problem rather than a uni-modal problem. Therefore, our solution for this problem should consider this fact and address this multimodality. In order to address this issue while applying GPS, we may resort to the use of FPN. FPN provides us the best initial parameters close to the global optimum point, which in turn will lead us to the global optimum point.

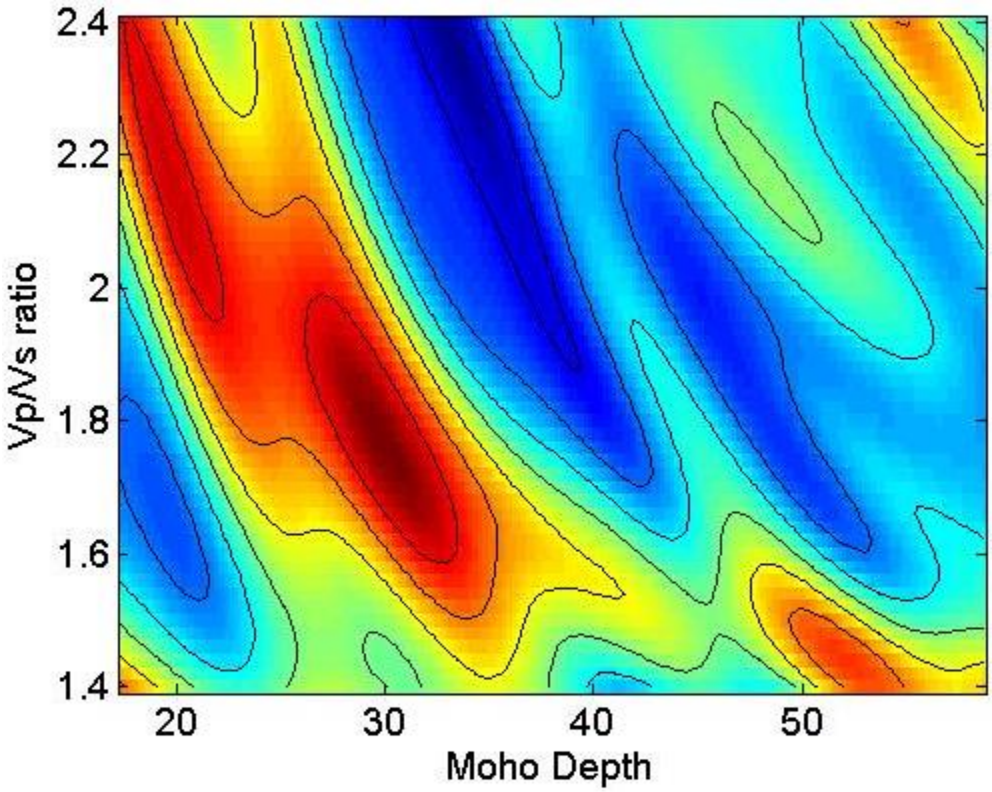


Figure 3.4 A contour plot of an H-κ Receiver Function stacks demonstrating that the problem has multiple peaks and troughs.

3.4.2 How FPN works for multimodal H-κ stacking surfaces. Generally, the objective function of the H-κ stacking algorithm has usually multimodal surfaces with multiple local maxima. We would like to find all local and global optimal points that could be potential solutions or best initial parameters for further searching using GPS to the H-κ stacking of receiver functions. FPN, unlike traditional Genetic Algorithms, helps to identify the different peaks in multimodal functions and multimodal H-κ stacking surfaces. FPN keeps diversity in the population by uniformly distributing the population fitness along the various peaks. Thus, we can find all the peaks of the multimodal H-κ surfaces using Fitness Proportionate Niching (FPN) of the Genetic Algorithms.

Basically the novel proposal of using FPN for identifying the best initial values comes from the idea that the FPN would always include a cluster around the global optimum point among the different clusters it is supposed to create. The FPN, as a means of clustering, will make different clusters around the different peaks in the data. By the niching mechanism it employs, FPN will increase the chances of finding alternative solutions instead of finding a single optimum point unlike the traditional GA. However, the FPN like any GA would enable to find the global optimum point in addition to other alternative local optima. Therefore, without waiting too long for the FPN to converge on all its niches and niche masters, just few runs of FPN should be able to allow locating the approximate position of the global optimum point. Thus, in every niche master set resulting from every FPN run, we would expect to find a niche master corresponding to the global optimum point. Our implementation and test on the H-κ stacking surfaces showed that sometimes we may even find two niche masters which are very close to the global optimum. The reason is because the FPN has not yet settled and when settled, the close niche masters would belong to a single niche.

3.4.3 Implementation of FPN. For the implementation of FPN, we need some weights for the receiver function set of a seismic station. An acceptable set of weights can be obtained using the initial run of the GPS technique, or the GA implementation. Once we have weights for the given set of receiver functions, the variables for the H-κ stacking objective function will be H and κ.

It is not difficult to show that the H-κ parameter space with the objective function value variation forms a multimodal surface. Clustering the objective function values together around cluster or niche centers will be done using the FPN algorithm. Once the clusters around the peak points on the H-κ parameter surface are determined, the niche masters (or the cluster centers) will be the potential optimal points. Among these multiple optima the global maximum can be an initial point for the GPS run.

A great advantage about the FPN compared to other clustering techniques such as K-mean clustering, Traditional Fitness Sharing clustering algorithms is that we don’t need to have a priori knowledge about the number of clusters upfront. FPN evolves and determines the number of clusters and makes the clustering by itself.

3.4.4 Algorithm 3: determining best initial models with FPN for rerunning GPS.

Algorithm 3 basically about applying the Fitness Proportionate Niching (FPN) and determining the best initial parameters.

If Algorithm 2 gives two different H and κ values for different extreme initial values, we resort to FPN to determine the best initial model parameters close to the global optimum point. It is observed that FPN identifies the approximate global minimum, which is very close to the actual value. Every time it runs, it may pick different local minima once or the same local minima twice.

For the case of seismic station ARBA, as an example, the FPN picks 3 niche centers. One of these centers is found to be the approximate global minimum, while the other two niche masters indicate either two different local optima or sometimes the same local minima with two different but very close niche center values.

Some FPN implementation results are given in Table 3.2. For station ARBA, the Niche Masters (Cluster Centers) identified are tabulated in the Table:

Table 3.2

Table of Some FPN Implementation Results for Seismic Station ARBA

FPN Run No.	**Niche Masters Identified**		
1	38.5451 2.1773	**30.3465 1.7673**	39.2321 2.2116
2	18.5000 1.1750	18.5000 1.1750	**30.1672 1.7584**
3	**30.1123 1.7556**	18.5000 1.1750	18.5000 1.1750
4	**30.1000 1.7550**	**30.3000 1.7650**	38.3740 2.1687
5	18.5000 1.1750	**30.3000 1.7650**	41.5000 2.3250
6	18.5000 1.1750	**30.1578 1.7579**	41.5000 2.3250
7	18.5000 1.1750	**30.3478 1.7674**	18.5000 1.1750
8	**30.1000 1.7550**	**30.3000 1.7650**	38.3740 2.1687
9	18.5000 1.1750	**30.3000 1.7650**	41.5000 2.3250
10	18.5000 1.1750	**30.1578 1.7579**	41.50002.3250

Table 3.2 gives FPN results for seismic station ARBA when it is run 10 times. Each time the FPN is run, it identifies 3 Niche centers. We can clearly observe that one of those niche masters identified is the global optimal point (**30.3465 1.7673**) or **(30.1123 1.7556).** It is very important to note that for the initial value of H we can consider only the value "**30,**" dropping fractions after the dot, and for the value of κ we may take "**1.76**" (rounding the 3rd number after the dot to the second number if that number is odd, or else keep the second number

after the dot as it is). Thus, our best initial value for station ARBA receiver functions set is **(30, 1.76).**

An algorithm that can sum up the above observation for FPN implementation is as follows:

$$H_{init} = (H_{i1}| H_{i2}| H_{i3})\&(H_{i+1,1}| H_{i+1,2}| H_{i+1,3}) \&(H_{i+2,1}| H_{i+2,2}| H_{i+2,3}) \& \ldots$$

$$\kappa_{init} = (\kappa_{i1}| \kappa_{i1|} \kappa_{i1})\& (\kappa_{i+1,1}| \kappa_{i+1,2|} \kappa_{i+1,3})\&(\kappa_{i1}| \kappa_{i1|} \kappa_{i1})\&(\kappa_{i1}| \kappa_{i1|} \kappa_{i1})\&(\kappa_{i1}| \kappa_{i1|} \kappa_{i1})\& \ldots$$

where H_{i1} represents the first of the 2-tuple (H, κ) parameter outputs representing the niche master, "i" denotes the run number (epoch) and "1" signifies the first of the niche masters identified by FPN during that single run. Thus, H_{i2} and H_{i3} represent the first of the 2-tuples that provide the 2nd and 3rd crustal thickness values which could be either local or global optima point(s). "i+1," "i+2," "i+3,"… show the $(i+1)^{th}$, $(i+2)^{nd}$, $(i+3)^{rd}$, etc. run. Similarly, κ_{i1}, $\kappa_{i+1,}$ κ_{i+2} denote the second of the 2-tuple (H, κ) parameter outputs representing the niche masters for "i," "i+1," "i+2," FPN runs. Finally, H_{init} and κ_{init} represent the best initial model to be used when rerunning the GPS technique as in Algorithm 1. This algorithm has been implemented by Matlab.

CHAPTER 4

Experimental Implementation and Results

4.1 Genetic Algorithm Implementation and Test Results

We implemented the genetic algorithm (GA) for estimating appropriate weights for the three seismic phases of receiver functions as well as the optimal values of the crustal parameters as shown in the flow diagram in Figure 3.3. To test its performance, the GA has been run for receiver functions whose H-κ stacking result was published before (Dugda et al., 2005). Seismic data from more than 25 seismic stations have been used to test the GPS and FPN implementation. Most of these data were collected by the Ethiopian Broadband Seismic Experiment (Nyblade and Langston, 2002). For testing the application of the three techniques, the GA, GPS and FPN, we employed seismic station ARBA.

Seismic data collected by station ARBA is particularly used in this research for testing the GA performance. The station ARBA lies in the southern region of Ethiopia, within the Main Ethiopian Rift. The following subsection gives the location of seismic station ARBA and other seismic stations from which seismic data are employed for testing the different tools developed in this study. Receiver functions have been computed for the seismic data obtained from those seismic stations.

4.1.1 Location of seismic stations for testing the hypotheses. The location of seismic stations from which seismic data has been obtained is shown in Figure 4.1. Seismic data collected from Ethiopia has been used for testing the hypotheses and different techniques developed in this dissertation. Black triangles represent seismic stations which run for one year between 2001 and 2002 (Nyblade and Langston, 2002). Squares with brown inside represent seismic stations which run for one year between 2001 and 2002. Completely white squares represent GSN/IRIS permanent seismic stations. IRIS stands for Incorporated Research Institutions for Seismology and GSN stands for Global Seismographic Network of the United States Geological Survey (USGS). The rift fault lines are shown by heavy lines.

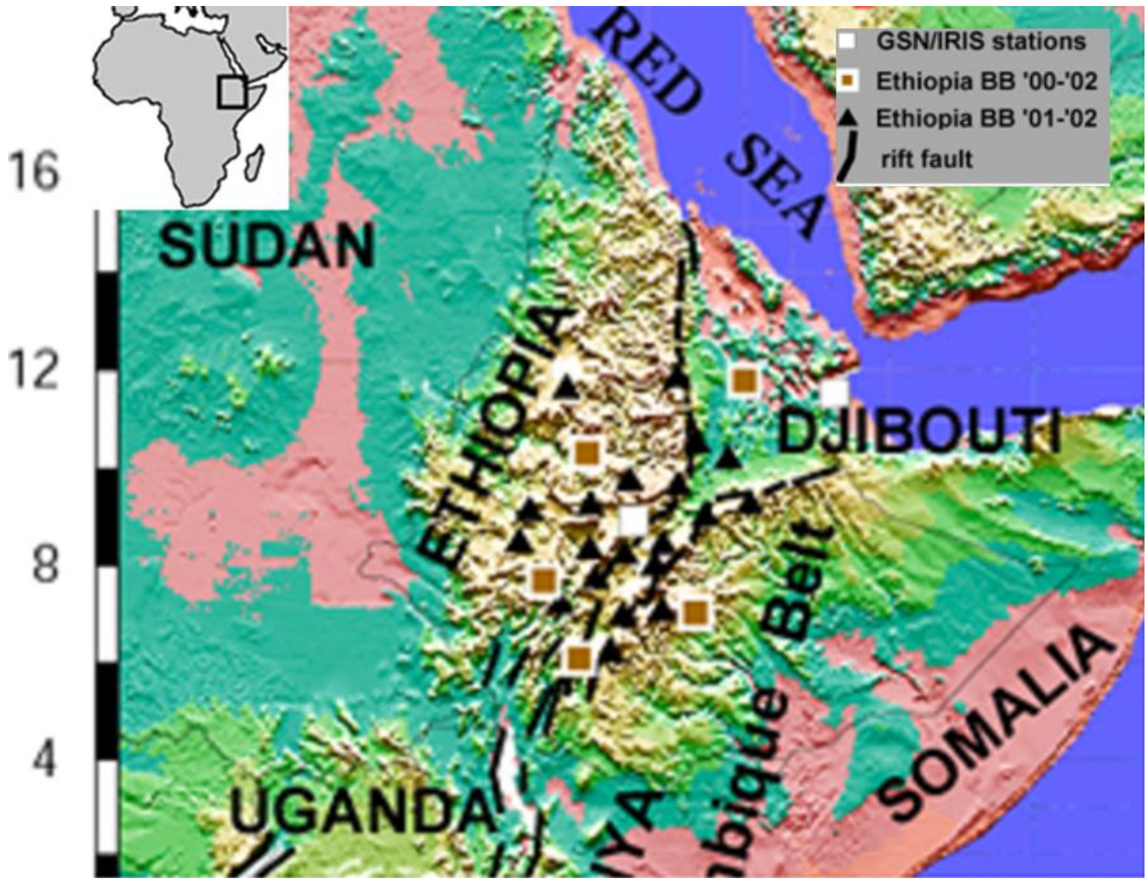


Figure 4.1 Figure showing location of 27 Seismic stations in Ethiopia (East Africa) for testing our Hypotheses.

There are different numbers of receiver functions for each station, ranging from 5 to 17. Seismic station ARBA has 13 receiver functions. Seismic station ARBA (the bottommost of the five brown squares) is a station in the Main Ethiopian Rift segment of the East African Rift System. This data has been used before in other researches, as well as used in our GA implementation before. Also we will use the same data for the GPS implementation. Then, we can compare current results using GA, GPS, and FPN with previously determined results.

4.1.2 Receiver functions from ARBA and GA results. Figure 4.2 presents the receiver functions computed using the seismic data gathered from ARBA. This data set is used for testing the GA algorithm technique that we developed. The receiver functions set contains 13 high quality receiver functions obtained from seismic data collected for 2 years and these receiver functions were used to determine crustal parameters. The three times t_1, t_2, and t_3 picked by our algorithms for the three seismic converted phases from the crust-mantle boundary the Moho are shown with short vertical lines on each of the receiver function time series on the figure.

Since weights have been obtained in Dugda et al. (2005) for the same data set using Monte Carlo technique, comparison can be easily made to observe if the weights found by the GA algorithm closely match the weights obtained in the previous study. After running the GA for a number of generations (ranging between 10 and 15) and repeating the experiment about 60 times, the average of these runs provides the optimal or near optimal weights. As discussed below, after about 60 GA runs, the average values of the three weights stabilize to give the optimal or near optimal weights.

A sample plot of the results from the GA implementation is shown in Figure 4.3. Figures 4.3 gives contour maps of the H-κ stacking objective function on a κ versus H axes just for one instance of running the GA. The different colors in Figure 4.3 show the value of the objective function. Thus, red color delineates regions where the maximum value of the objective function occurs and blue denotes minimum objective function values, and other colors represent values in between the maximum and minimum objective function values. Each successive contour represents a 10% reduction in values of the objective function with respect to the maximum at the center of an interior contour with orange or brighter colors. In Figure 4.3, the values of H and κ are 30.5 km and 1.76, respectively, for a single GA run and the average values for H and κ are 29.7 km and 1.78, after the weights stabilize and take constant average values as shown in Figure 4.4.

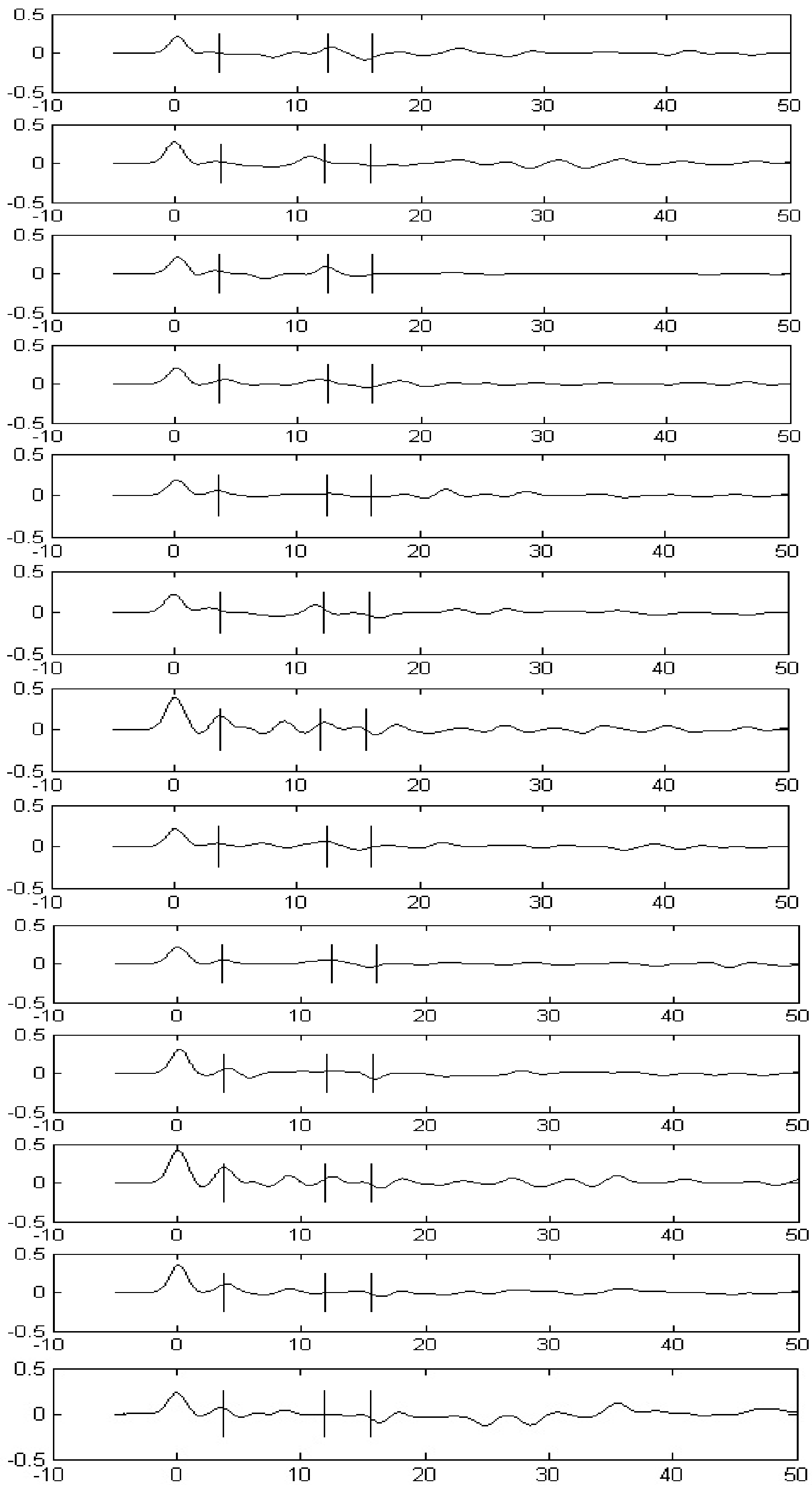


Figure 4.2 The receiver functions used for testing the GA, GPS and FPN techniques.

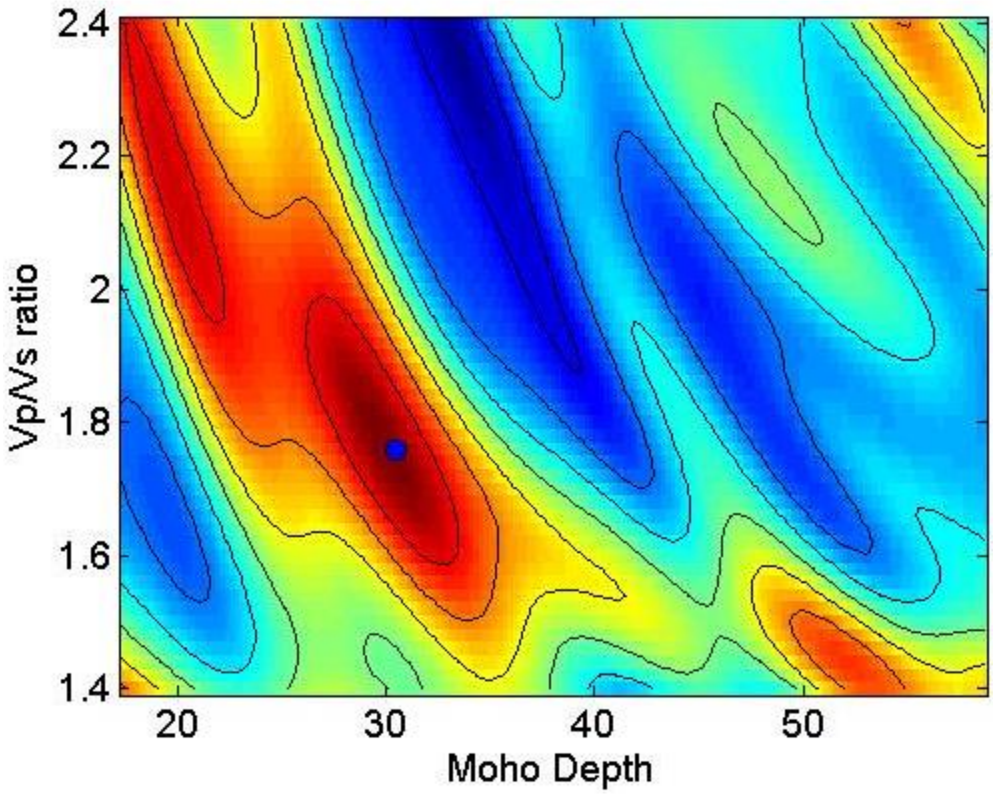


Figure 4.3 One instance of running the GA with very wide H-κ parameter space. Note that the parameter space used here is wider than that of Dugda et al. (2012).

Figure 4.4 exhibits the three GA weights obtained in this study. This figure clearly shows that average weights stabilize after about 60 GA runs of the GA implementation on the H-κ receiver function stacks. Thus, the average weights for the given set of receiver functions from the seismic station ARBA in Ethiopia are stabilized to:

$$w_1 = 0.454,\ w_2 = 0.404, \text{ and } w_3 = 0.142$$

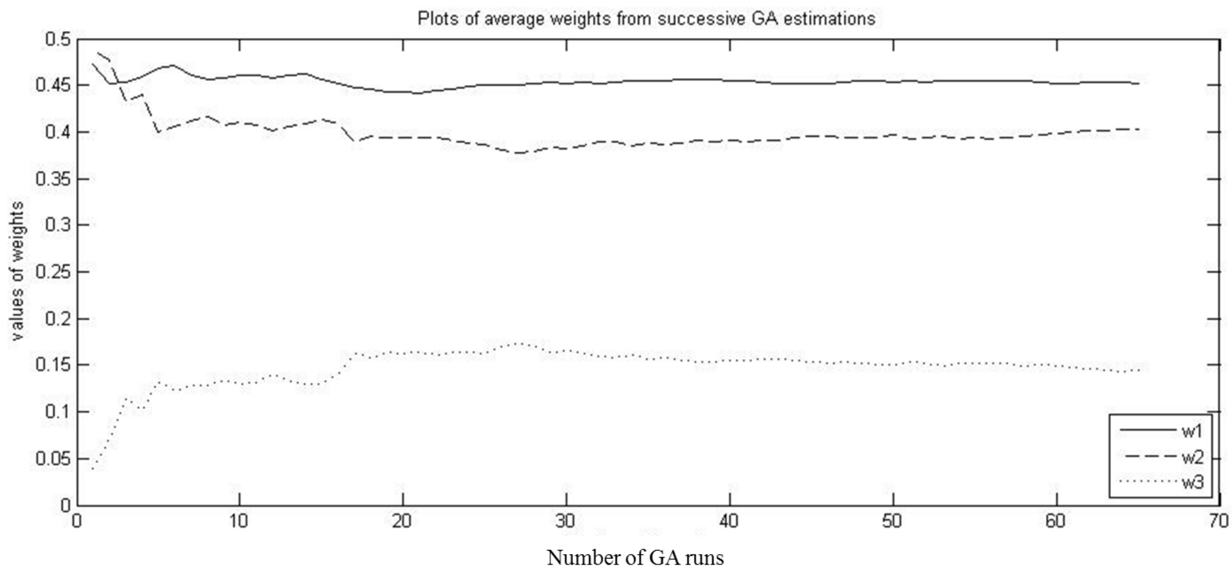


Figure 4.4 A graph showing the values of the three weights stabilized after about 60 epochs.

4.2 GPS Test Results

4.2.1 Flow chart of the GPS implementation. For the GPS implementation, first, the maximization optimization problem needs to be converted to a minimization optimization problem. The pattern search technique was originally developed for a minimization problem and so we can take advantage of previous studies if we convert the problem from maximization to minimization optimization problem. The steps of this conversion have already been discussed in chapter 3.

Since there are five different variables to be determined in the H-κ receiver function problem, a mesh size of 1 can offer 10 pattern vectors each of which equal to a unit-size direction (pattern) vector. If we add the pattern vectors to an initial iterate point, we obtain 10 new iterate (mesh) points. If we apply partial polling of GPS, we calculate objective function values for the new iterate (mesh) points until we obtain an objective function value less than the objective function at the current iterate. On the other hand, if complete polling is desired, we calculate objective function values at all the new mesh points before seeking to find the mesh point with the smallest objective function value in every iteration. Figure 4.5 gives the flow chart of the implementation of the GPS algorithm for our research.

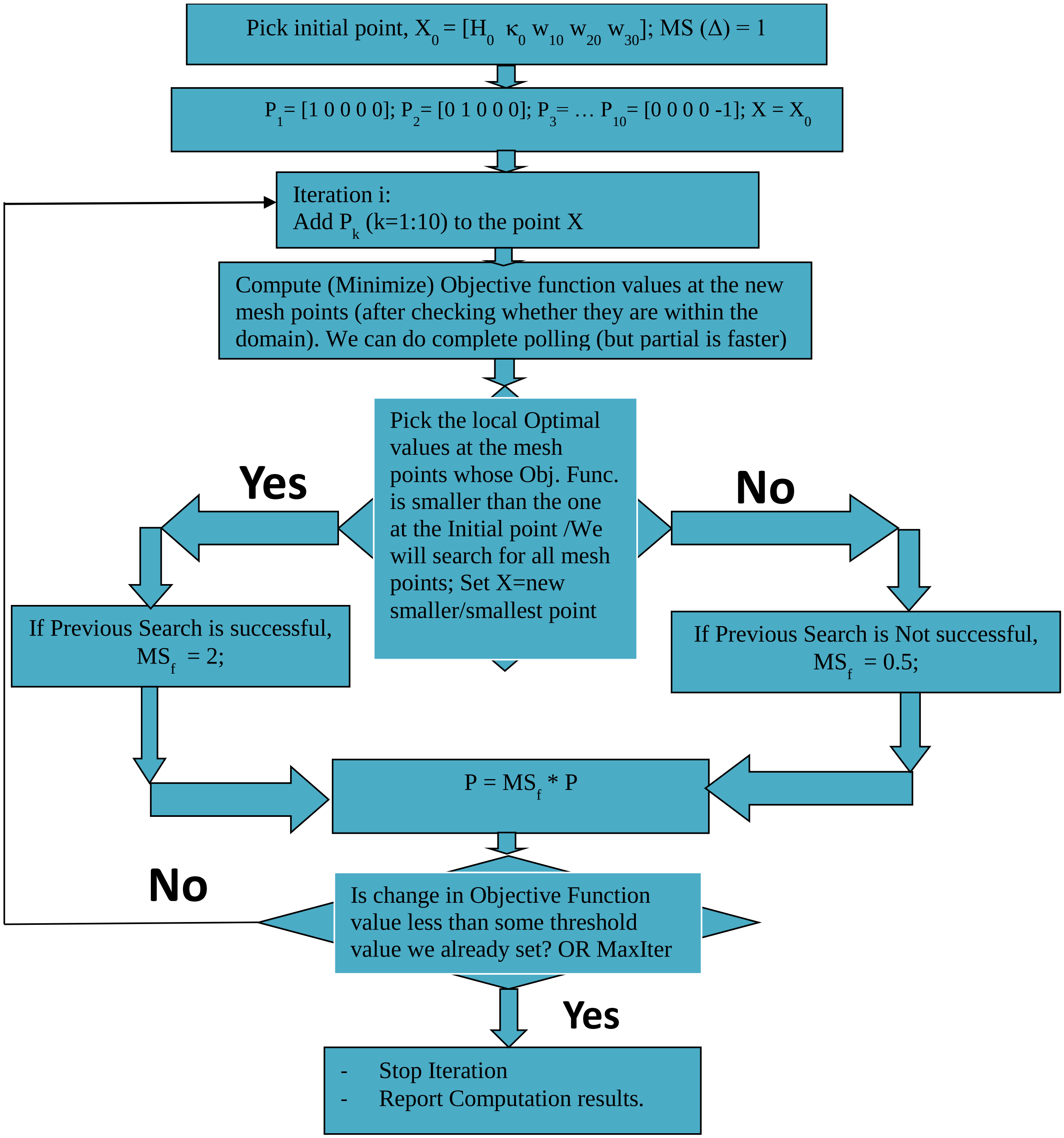


Figure 4.5 This figure shows the flow chart of the GPS Implementation. MS denotes Mesh Size.

4.2.2 GPS results - convergence of final values and objective function

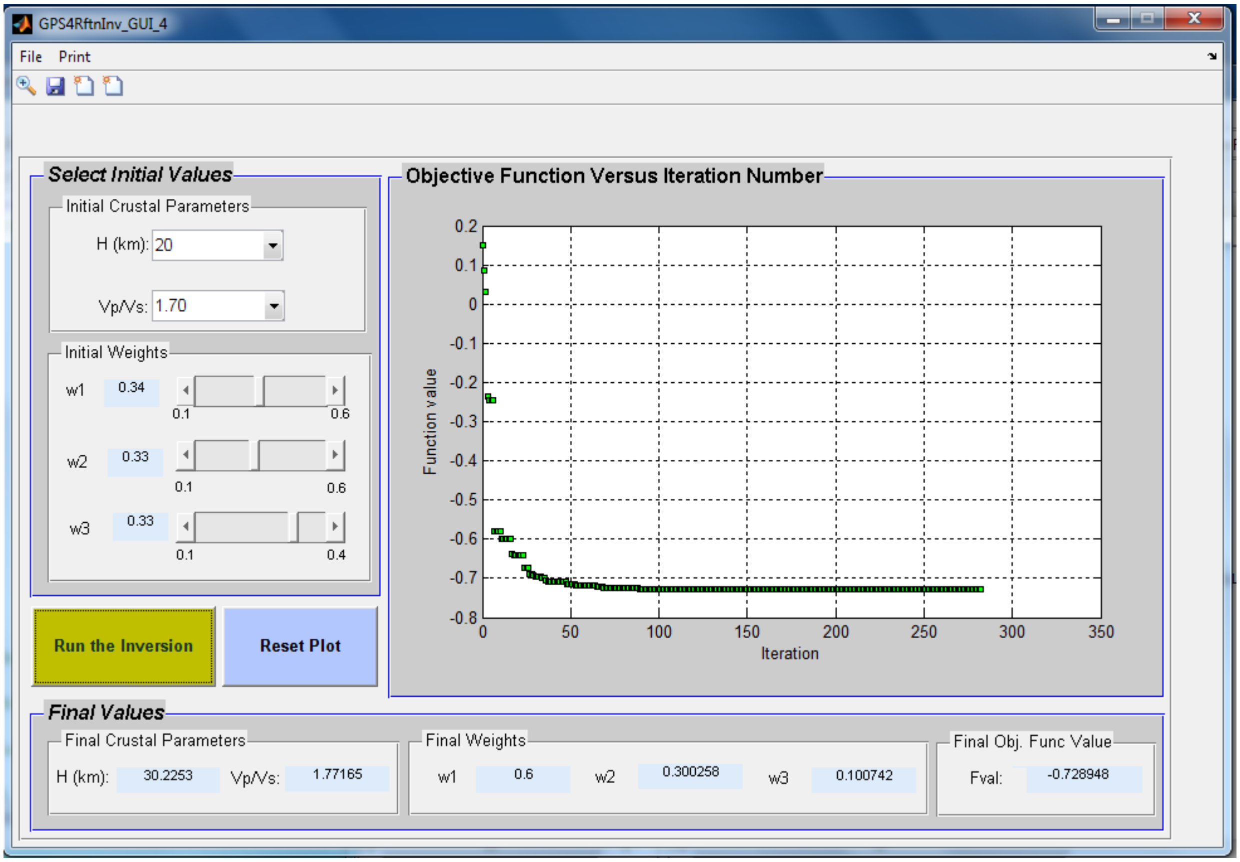


Figure 4.6 This figure displays a GUI developed for GPS implementation.

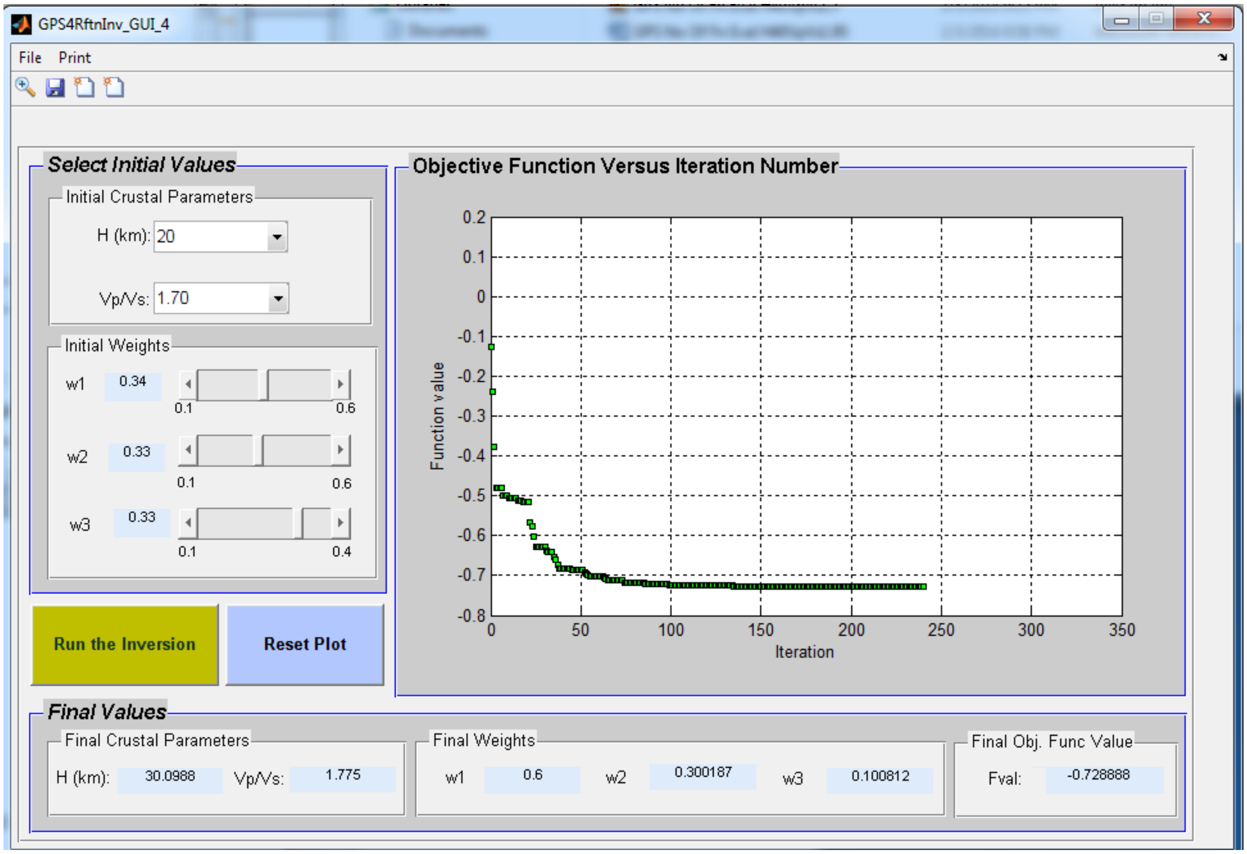
The GUI implementation of the GPS algorithm is shown in Figure 4.6. The GUI implementation provides opportunities to display the GPS results. It shows the initial values of the weights and crustal parameters on the top left hand panel, while it provides the final values of the weights and the crustal parameters at the bottom panel. The graph on the GUI displays the variation of the objective function of the minimization optimization problem versus the number of iterations. The figure particularly shows the GPS inversion results for seismic station ARBA.

The initial values are (20, 1.70, 0.34, 0.33, 0.33) and the final values are calculated by GPS algorithm. The final values of crustal parameters and weights are (30.2, 1.77, 0.6, 0.3, 0.1).

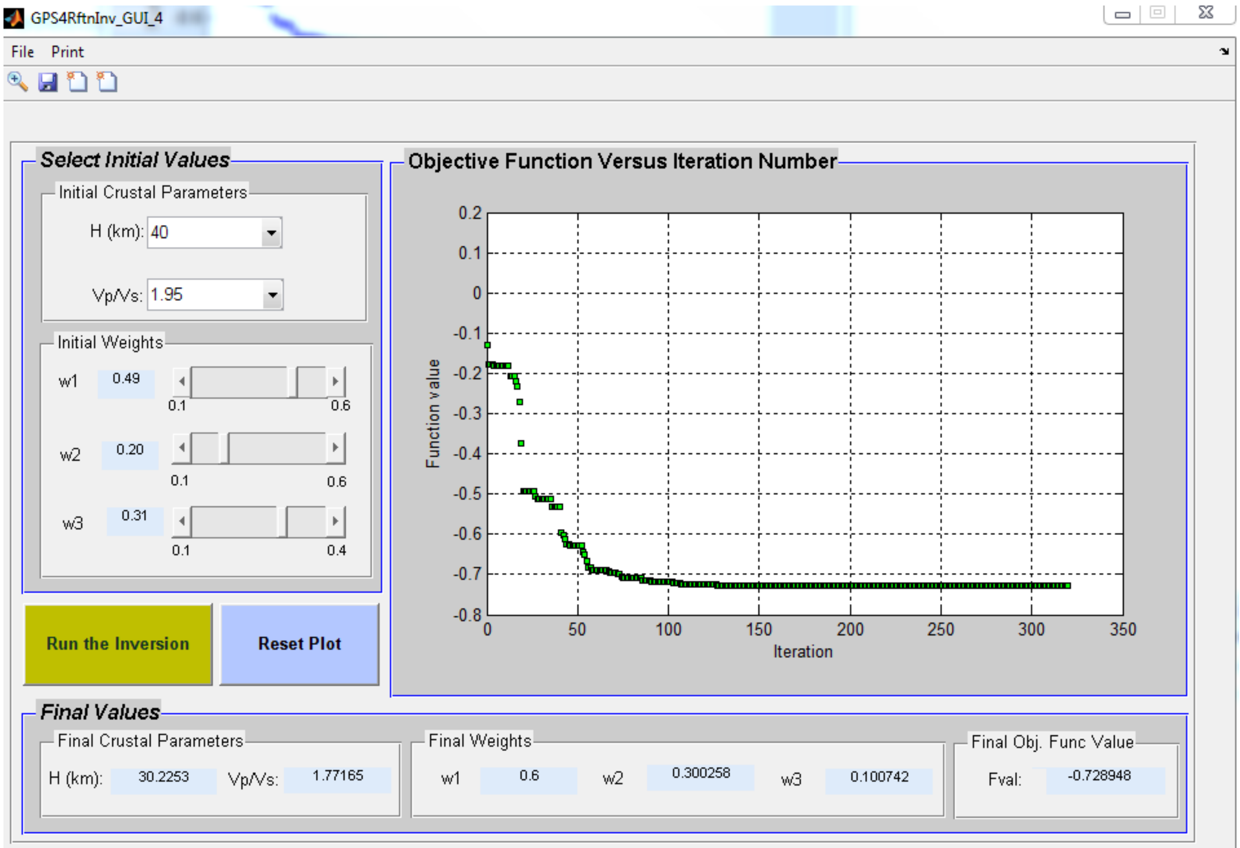
4.2.3 Graphical user interface (GUI) implementation for the GPS algorithm. A graphical user interface (GUI) implementation is important not only for examining and displaying initial parameters and final parameters, but also for demonstrating the convergence of the algorithm. A GUI has been developed using Matlab for the GPS algorithm. The GUI is very significant especially for the implementation of the GPS initial value dependence testing algorithm in this dissertation. Since the GUI clearly shows both the initial as well as final values of parameters and weights, observation of initial value dependence will be evident.

4.2.4 GPS convergence test algorithm. The algorithm of GPS Convergence Test is performed to examine final parameters and final objective function values at least for two extreme initial values. Using the GUI that is developed in this research work, the convergence for seismic station ARBA can be shown.

4.2.5 GPS convergence test for station ARBA. It is observed that, for two extreme initial values and combinations of extreme values as well as for some intermediate initial values, GPS gives similar results. Though their initial parameter values as well as their initial objective function values are different, the final parameter and objective function values are similar; i.e. the GPS results converge to the same final values. Figure 4.7 (a) and (b) show that the final values from the GPS converge to the same final or optimal values irrespective of the initial values. Therefore, in such a case we don’t need to resort to applying the FPN to find the best initial values because the inversion is independent of initial values or model parameters.



(a) Lower extreme initial H and κ (Vp/Vs) ratio values



(b) Higher extreme initial H and κ(Vp/Vs) ratio values

Figure 4.7 Figure displaying two extreme initial values for the region of investigation. (a) Smallest and (b) highest initial parameter values.

4.2.6 Complete polling of GPS for global optimization. In a given iteration, the generalized pattern search polls the points in the current mesh by computing the objective function at those mesh points. The GPS continues this polling process at the mesh points in the iteration until it finds the mesh point whose function value is smaller than the function value at the current iterate point. Generally a generalized pattern search stops its polling when it discovers a mesh point that improves the value of the objective function, and then it appoints that point as the current point for the next iteration. If the intent is to get faster computation, usually we may not allow all the available mesh points to be polled because we may want the algorithm to search only fewer points in every iteration. On the other hand, some of these unpolled mesh points might produce objective function values which are even smaller than the one the pattern search already declared to be smaller.

H-k stacking is shown in this research to be a multimodal problem with many local optima. For such multimodal problems, it is usually superior to construct the generalized pattern search technique so that it polls all the available mesh points in all the iterations and pick the one with the best objective function value. This kind of polling is known as complete polling and such kind of polling facilitates the generalized pattern search to explore more points at every iteration and in this way the GPS potentially avoids a local minimum which is not the global minimum.

The ARBA station receiver functions will be used as an example. The initial and final values for complete polling versus partial polling for station ARBA are given in the next two subsections.

4.2.6.1 Partial (non-complete) polling

Table 4.1

Initial and Final values for partial polling

variables	Initial Values	Final Values
H	45	52.0139
κ	1.6500	1.4360
w_1	0.3400	0.6000
w_2	0.3300	0.3010
w_3	0.3300	0.1000

Table 4.2

parameters of partial polling of GPS on ARBA receiver function inversion

Polling Type	Partial polling
Number of iterations	254
Number of function evaluations	1161
Time elapsed (minutes)	0.7625
The best function value	-0.550777

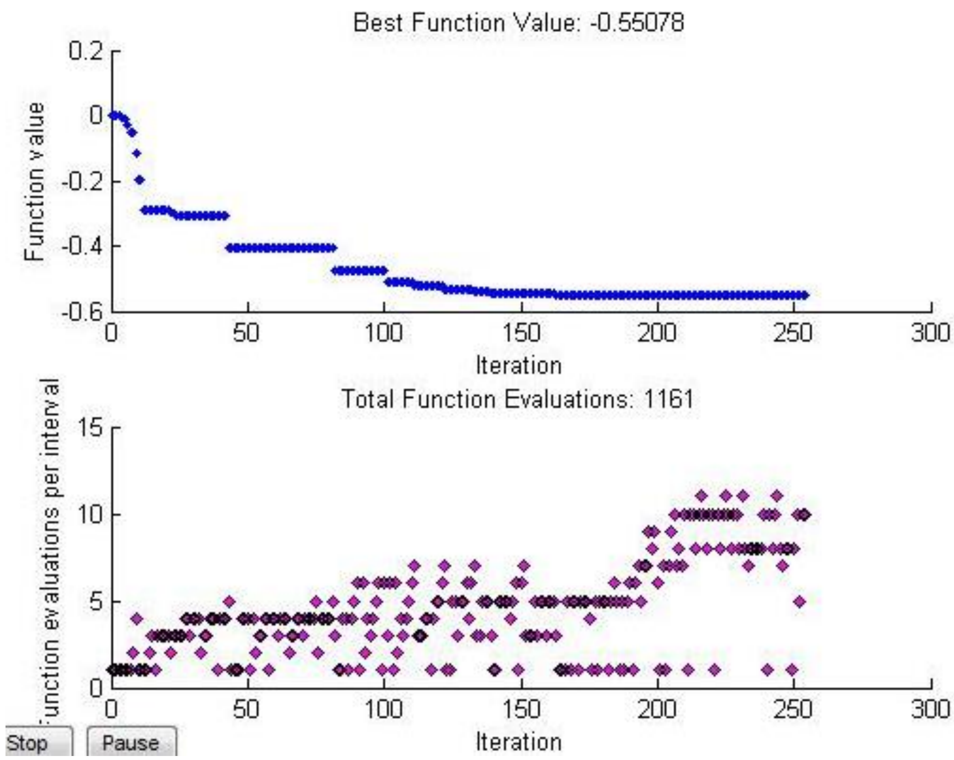


Figure 4.8 Objective function and total number of objective function evaluations versus number of iterations for a non-complete polling.

4.2.6.2 Complete polling

Table 4.3

Initial and Final values for partial polling.

variables	Initial Values	Final Values
H	45	30.1837
κ	1.6500	1.7728
w_1	0.3400	0.6000
w_2	0.3300	0.3010
w_3	0.3300	0.1000

More results from the GPS implementation will be discussed.

Table 4.4

complete polling of GPS on ARBA receiver function inversion

Polling Type	Complete Polling
Number of iterations	216
Number of function evaluations	1156
Time elapsed (minutes)	0.5847
The best function value	-0.729039

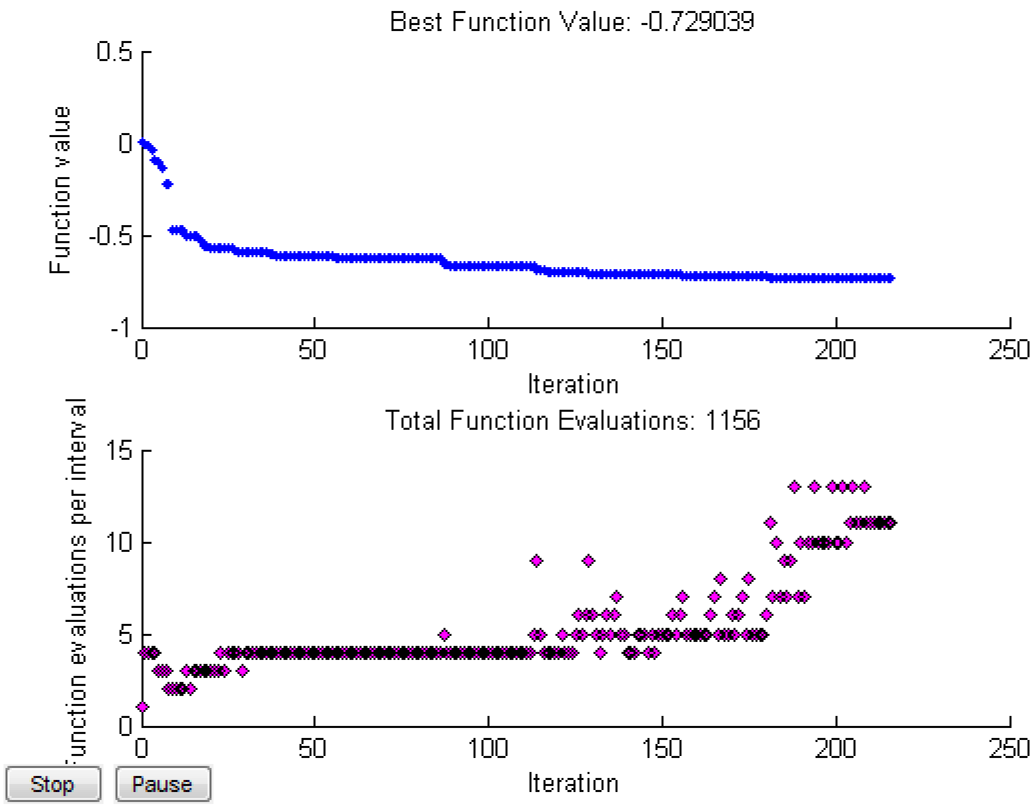


Figure 4.9 Objective function and total number of objective function evaluations versus number of iterations for a non-complete polling.

This is a 5-dimensional search problem. However, we can utilize only the 2-D cross-section to show the difference between the complete and partial polling of the GPS when they are applied to the ARBA receiver functions inversion. Figure 4.10 demonstrates a 2-D view of the location of the two optimal values when the complete polling of the GPS is applied (small solid blue circle) and when the partial polling (brown star with white circle inside) is applied. The figure clearly shows that the complete polling provides the global optimum while the partial polling is trapped in a local optimum point.

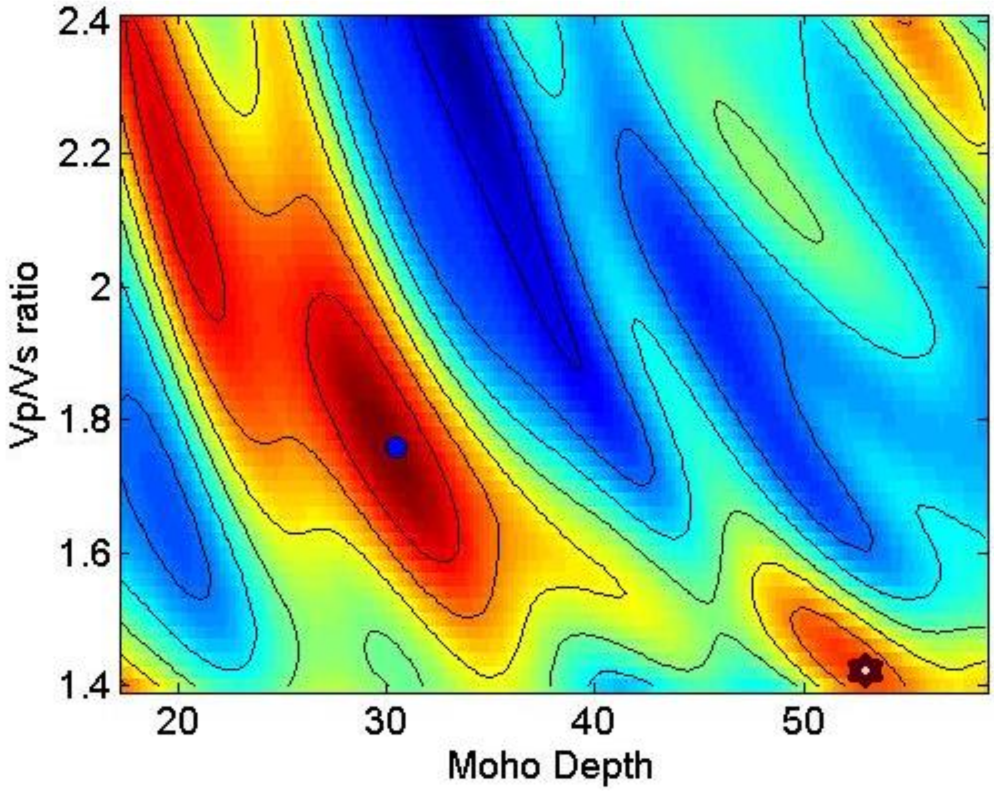


Figure 4.10 A 2-D view of the locations of the two optimal values when the GPS is applied with complete polling (blue circle) and partial polling (brown star with white inside).

Table 4.5

Solutions of crustal parameters (H and κ) and weights (w_1, w_2, w_3) using the current study applying the GPS-FPN technique for seismic stations shown on the map of Figure 4.5.

Station	H	κ	$\mathbf{w_1}$	$\mathbf{w_2}$	$\mathbf{w_3}$
ARBA	30.1837	1.7728	0.6000	0.3010	0.1000
AAUS	39.0309	1.8774	0.6000	0.3010	0.1000
BAHI	42.9995	1.8231	0.3003	0.6000	0.1007
BDAR	46.4352	1.7175	0.6000	0.3010	0.1000
BIRH	40.9531	1.8400	0.6000	0.3010	0.1000
BUTA	30.3464	2.0	0.6000	0.3010	0.1000
CHEF	34.0872	1.7258	0.3006	0.3004	0.4000
DELE	35.8398	1.7573	0.3600	0.5410	0.1000
DIYA	37.3854	1.7232	0.3005	0.6000	0.1005
DMRK	40.7179	1.8157	0.6000	0.3010	0.1000
FICH	42.9034	1.7320	0.3973	0.3014	0.3014
FURI	42.6572	1.7947	0.6000	0.3010	0.1000
GOBA	41.9530	1.7509	0.6000	0.3010	0.1000
GUDE	40.8058	1.6500	0.5510	0.3500	0.1000
HERO	42.4117	1.8312	0.6000	0.3010	0.1000
HIRN	40.5610	1.7682	0.6000	0.3010	0.1000
HOSA	36.1074	1.9676	0.6000	0.3000	0.1010
JIMA	35.2090	1.8753	0.6000	0.3010	0.1000
KARA	44.7513	1.8386	0.6000	0.3010	0.1000
NAZA	32.2413	1.9763	0.6000	0.3000	0.1000
NEKE	33.6187	1.8102	0.5510	0.3500	0.1000
TERC	34.4358	1.7474	0.6000	0.3000	0.1010
WANE	28.3318	2.0	0.6000	0.3010	0.1000
WASH	35.4584	1.9200	0.6000	0.3010	0.1000
WELK	32.9070	1.8406	0.6000	0.3010	0.1000

4.3 FPN Results

Multimodal optimization problems can be mapped into a clustering problem. FPN, as a dynamic clustering algorithm, can be used to solve our H-κ receiver function stacking problem. Results from two different seismic stations shown on the map in Figure 4-5 are tabulated here to demonstrate that the proposed methodology of FPN can be implemented to solve the multimodal optimization problem of receiver function inversion.

Table 4.6
Table of Some FPN-Hκ stacking Implementation Results. Niche Masters (Cluster Centers) identified for station ARBA are tabulated.

FPN Run No.	**Niche Masters Identified**		
1	38.5451 2.1773	**30.3465 1.7673**	39.2321 2.2116
2	18.5000 1.1750	18.5000 1.1750	**30.1672 1.7584**
3	**30.1123 1.7556**	18.5000 1.1750	18.5000 1.1750
4	**30.1000 1.7550**	**30.3000 1.7650**	38.3740 2.1687
5	18.5000 1.1750	**30.3000 1.7650**	41.5000 2.3250
6	18.5000 1.1750	**30.1578 1.7579**	41.5000 2.3250
7	18.5000 1.1750	**30.3478 1.7674**	18.5000 1.1750
8	**30.1000 1.7550**	**30.3000 1.7650**	38.3740 2.1687
9	18.5000 1.1750	**30.3000 1.7650**	41.5000 2.3250
10	18.5000 1.1750	**30.1578 1.7579**	41.5001 2.3250

Table 4.7
Table of Some FPN-Hκ stacking Implementation Results. Niche Masters (Cluster Centers) identified for station BUTA are tabulated.

FPN Run No.	**Niche Masters Identified**		
1	**28.6400 1.8320**	19.5000 1.3750	36.7036 2.2352
2	**28.7764 1.8388**	19.5000 1.3750	27.7958 1.7898
3	**28.8407 1.8420**	36.5024 2.2251	36.9901 2.2495
4	**28.7598 1.8380**	19.5000 1.3750	19.5000 1.3750
5	**28.7000 1.8350**	19.5000 1.3750	37.0653 2.2533
6	20.8374 1.4419	**28.8033 1.8402**	40.5000 2.4250
7	**28.7979 1.8399**	29.2904 1.8645	37.1663 2.2583
8	**28.7000 1.8350**	**28.8947 1.8447**	37.0000 2.2500
9	**28.7000 1.8350**	19.5000 1.3750	36.7000 2.2350
10	**28.7631 1.8382**	35.9786 2.1989	36.7226 2.2361
11	**28.6704 1.8335**	37.5351 2.2768	37.5886 2.2794
12	**28.7764 1.8388**	38.0382 2.3019	36.9378 2.2469
13	**28.7638 1.8382**	19.5000 1.3750	37.1117 2.2556
14	38.6899 2.3345	**28.8000 1.8400**	19.5000 1.3750
15	36.2703 2.2135	**28.7615 1.8381**	19.5000 1.3750

CHAPTER 5

Discussion and Performance Evaluation

5.1 Discussion on GA Implementation and Results

In this research, GA has been applied to determine optimal or near optimal weights necessary in stacking receiver functions for crustal parameters H and κ. We used receiver functions and the resulting H and κ values similar to our previous study (Dugda et al., 2005) for the purpose of comparison. The receiver functions used for the GA implementation are computed from seismic data collected at ARBA seismic station.

The GA conducts a search for the weights necessary for inverting the receiver functions and simultaneously also determines the H and κ parameters from the receiver functions. After about 60 GA runs, the three weights converge to a steady-state value, in which case w_1 = 0. 454, w_2 = 0.404, and w_3 = 0.142 for station ARBA in Ethiopia. After the weights converge, the GA provides optimal or near optimal values for the weights and the crustal parameters.

Table 5.1 provides some sample weights computed by GA, number of generations in each GA run, and the elapsed times for the different GA runs. The time it takes for the epoch or GA run depends on the number of generations. It is observable from the table that one can obtain a good estimate of the weights when the number of generations is about 15. When the number of generations is in the range of 10 and 15, the weight estimates from each GA run are approximately the same as the average weight estimates. The times are given in minutes.

Table 5.1

Sample computed GA weights, number of generations and elapsed times for different GA runs.

Attempt No.	w_1	w_2	w_3	No. of Generations	Elapsed Time (min)
1	0.465	0.378	0.157	10	2.2
2	0.492	0.374	0.134	15	3.4
3	0.425	0.496	0.079	20	4.5

For the same set of receiver functions, we found that the respective best weights for the receiver functions stacking were 0.5, 0.4, and 0.1 in the previous work (Dugda et al., 2005), which are consistent with that of the GA implementation in this study. The comparison is given in Table 5.2. Although the two sets of weights match very closely, there seems to be minor differences (ranging between 0.004 and 0.046) in the values of the different corresponding weights too.

Table 5.2
Comparison of weights obtained in this study with the study by Dugda et al. (2005) for the three phases in the receiver functions

Study	w_1	w_2	w_3
Dugda et al. (2005)	0.5	0.4	0.1
Current Study (GA)	0.454	0.404	0.142

A contour plot in Chapter 4, Figures 4.3, presents the variation of H-κ receiver function stack on an H and κ parameter space just for one run of the GA algorithm. The same figure also indicates the position of the optimal values for H and κ parameters obtained using the current GA algorithm. In order to obtain statistically sound results, we have to run the GA several times because of the unpredictability involved in the mutation and crossover, as well as the initial parameters. After 60 GA runs the weights stabilize, as shown in Figure 4.4, and then the values for H and κ are 29.8 km and 1.77, respectively. The H and κ values from the previous study

(Dugda et al., 2005) were 29.7 km and 1.79, which are strikingly similar to the results in the current study. The values of the H and κ obtained through the use of the two different approaches are very similar and as may be expected, the two studies produce consistent results.

The average values for H and κ are 29.7 km and 1.78, after the weights stabilize and take almost constant average values as shown in Figure 4.4. The average weights for the given set of receiver functions from seismic station ARBA stabilized are:

$$w_1 = 0.454,\ w_2 = 0.404,\ \text{and}\ w_3 = 0.142$$

One approach to evaluate the performance of the GA implementation was to determine the time for the GA implementation which gives good results. This could be done by computing and comparing H and κ values. The GA outputs after running the GA for 10 to 15 generations and repeating the experiment about 60 times, the average of these runs is found to offer the optimal or near optimal weights, as well as optimal or near optimal crustal parameters.

5.2 Discussion on GPS Results and Implementation

5.2.1 GPS convergence test. GPS technique requires considering initial values for the variables to be explored. When we consider initial values for the different variables, especially the crustal parameters, we will be very conservative. We would be conservative compared to what the expected value for the crustal parameters are, according to global earth models such as CRUST2.0. CRUST2.0 gives global crustal thickness estimates for a very large region of $2^0X2^0 \approx 222kmX222km$ grid over many parts of the earth (Bassin et al., 2000). The CRUST2.0 crustal model takes advantage of compilation of global sediment thickness and ice thickness on a 1x1 degree scale. The compilation of CRUST2.0 covers most of Eurasia, North America, Australia and some areas of Africa and South America and in the oceans.

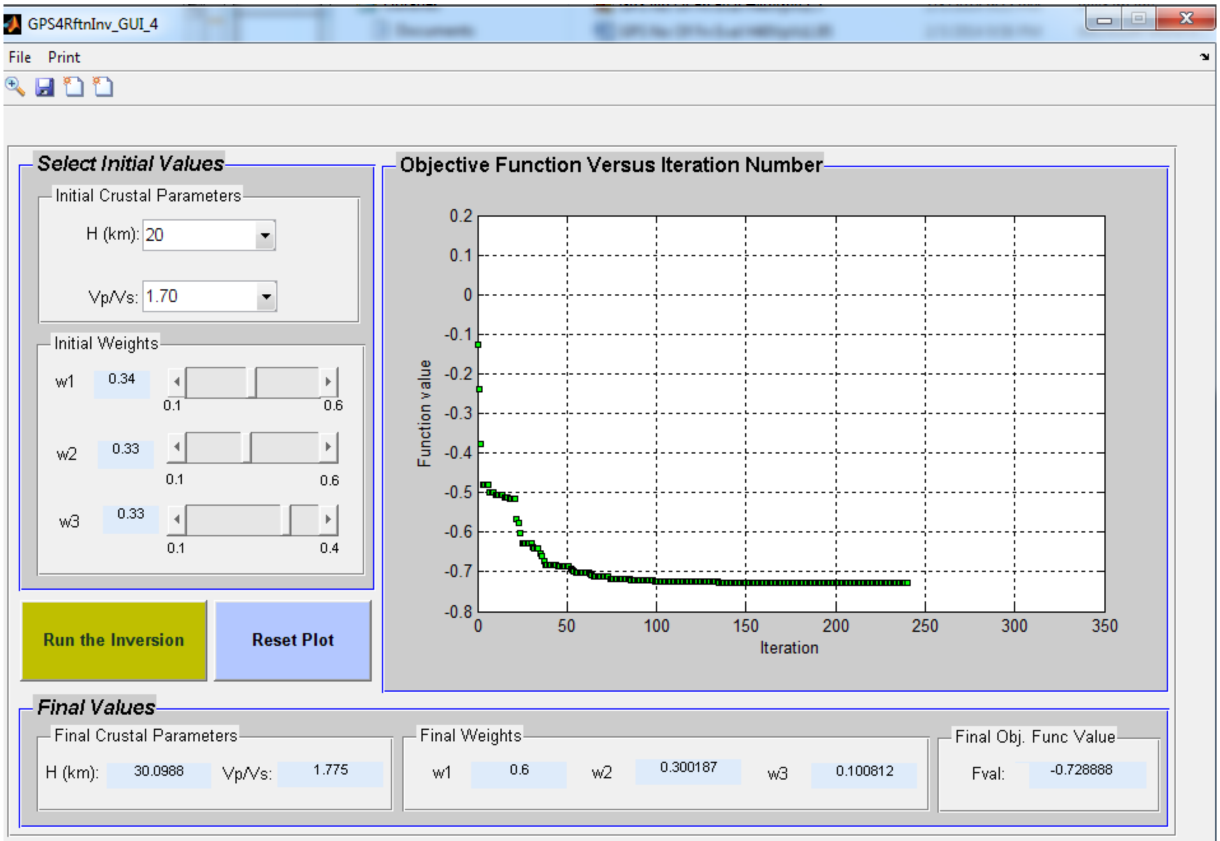
In taking the following initial value ranges, we would also be conservative from the stand point of other previous crustal studies in and around the Main Ethiopian Rift, where seismic station ARBA is situated. The crustal thickness in the Main Ethiopian Rift normally varies between about 25 and 35 km (e.g., Makris and Ginzburg, 1987; Mackenzie et al., 2005; Dugda et al., 2005). The κ values in the region vary typically between about 1.70 to about 1.90 (Zandt and Ammon, 1995; Dugda et al., 2005). Thus, the following parameter values can be used as initial values for the GPS application:

$H_{init} = 20 – 40$,

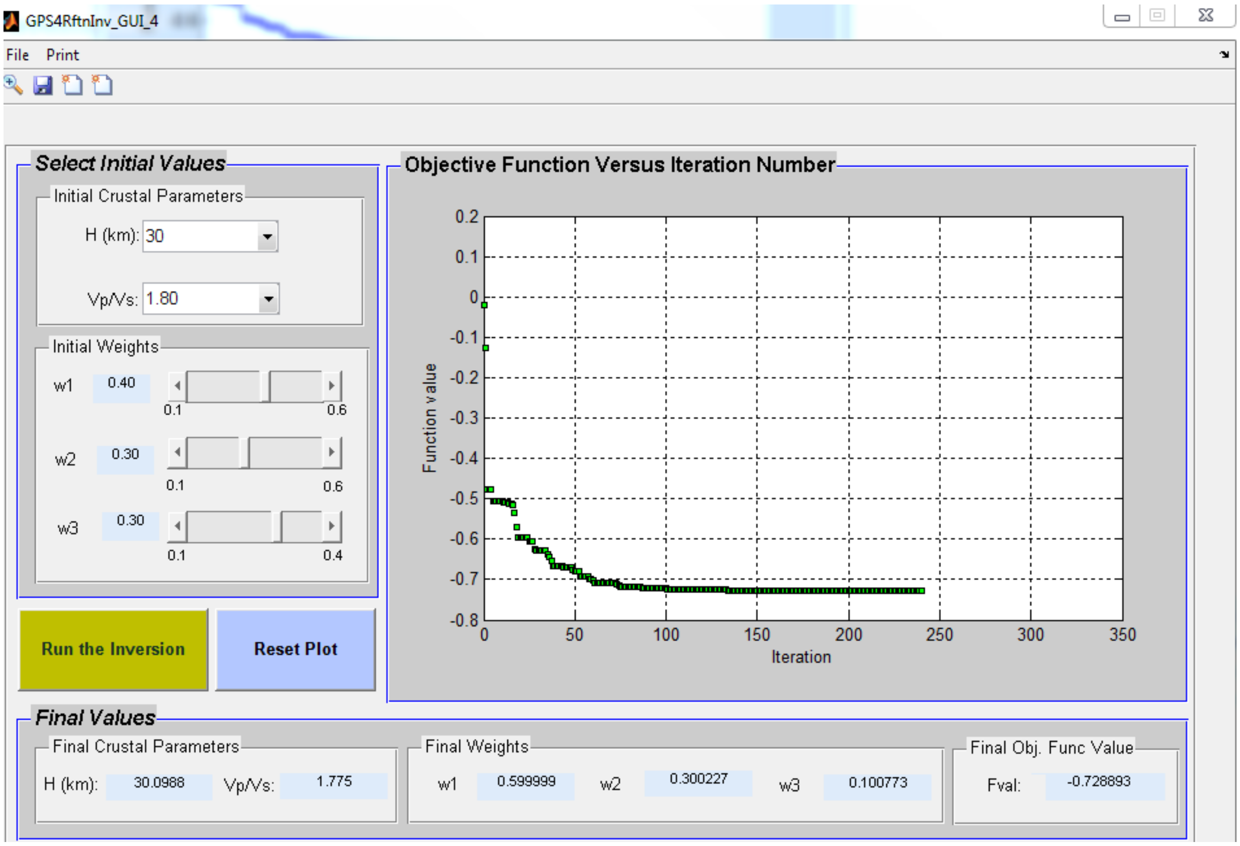
$\kappa_{init} = 1.65 - 1.95$,

w_{1init}, w_{2init}, w_{3init} can be taken from the range indicated in chapter 3.

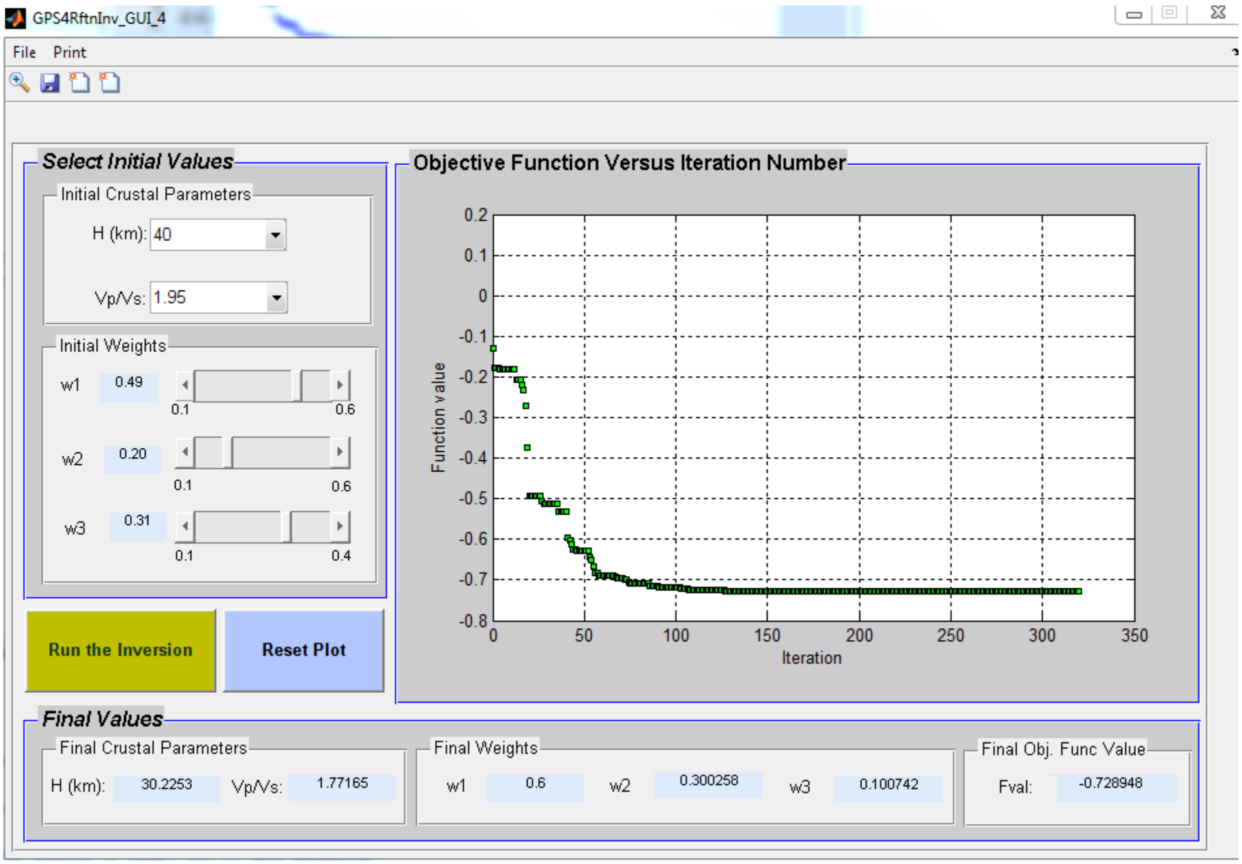
GPS Convergence has been tested by observing final parameters and also the final objective function values. We applied the test using not only two extreme initial values but also mixed extreme and some intermediate initial values. This has been performed for seismic station ARBA.



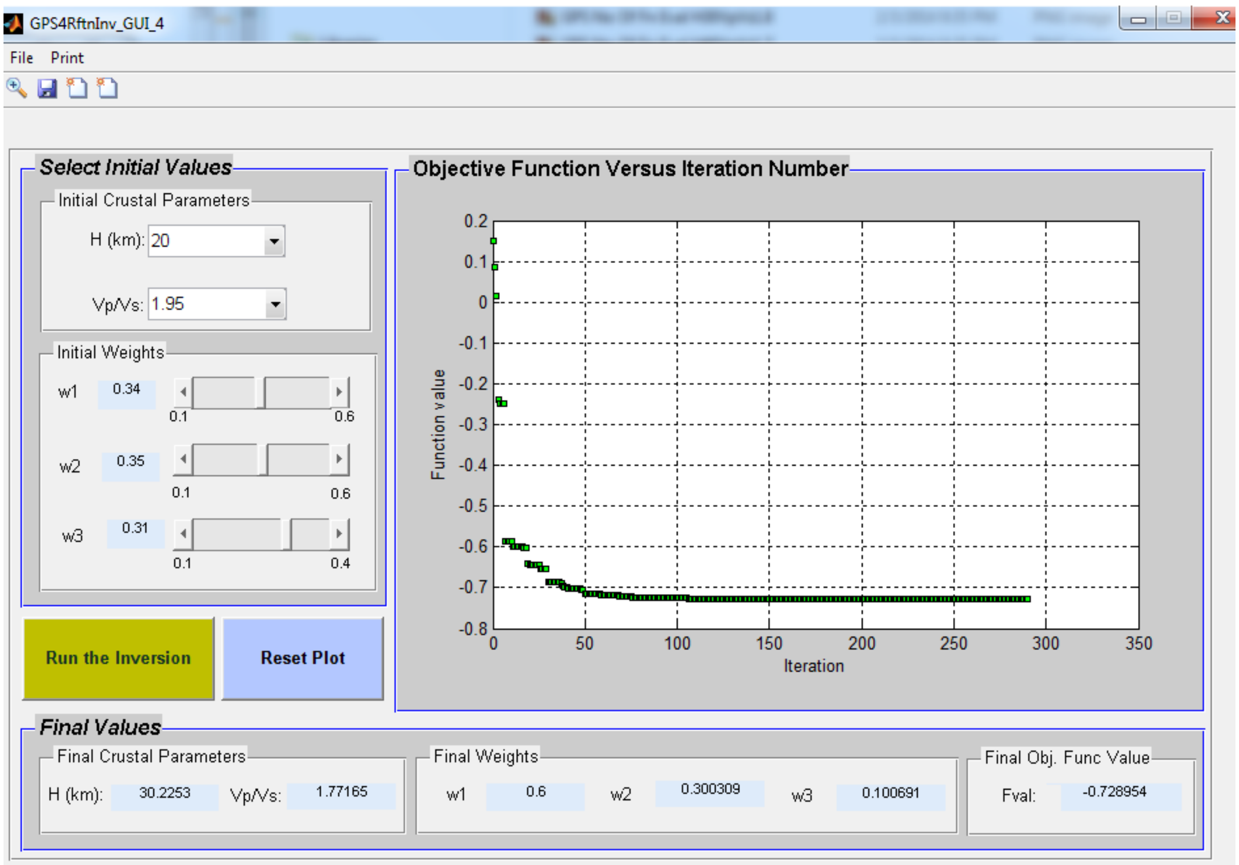
a) Lower extreme initial H and Vp/Vs ratio values



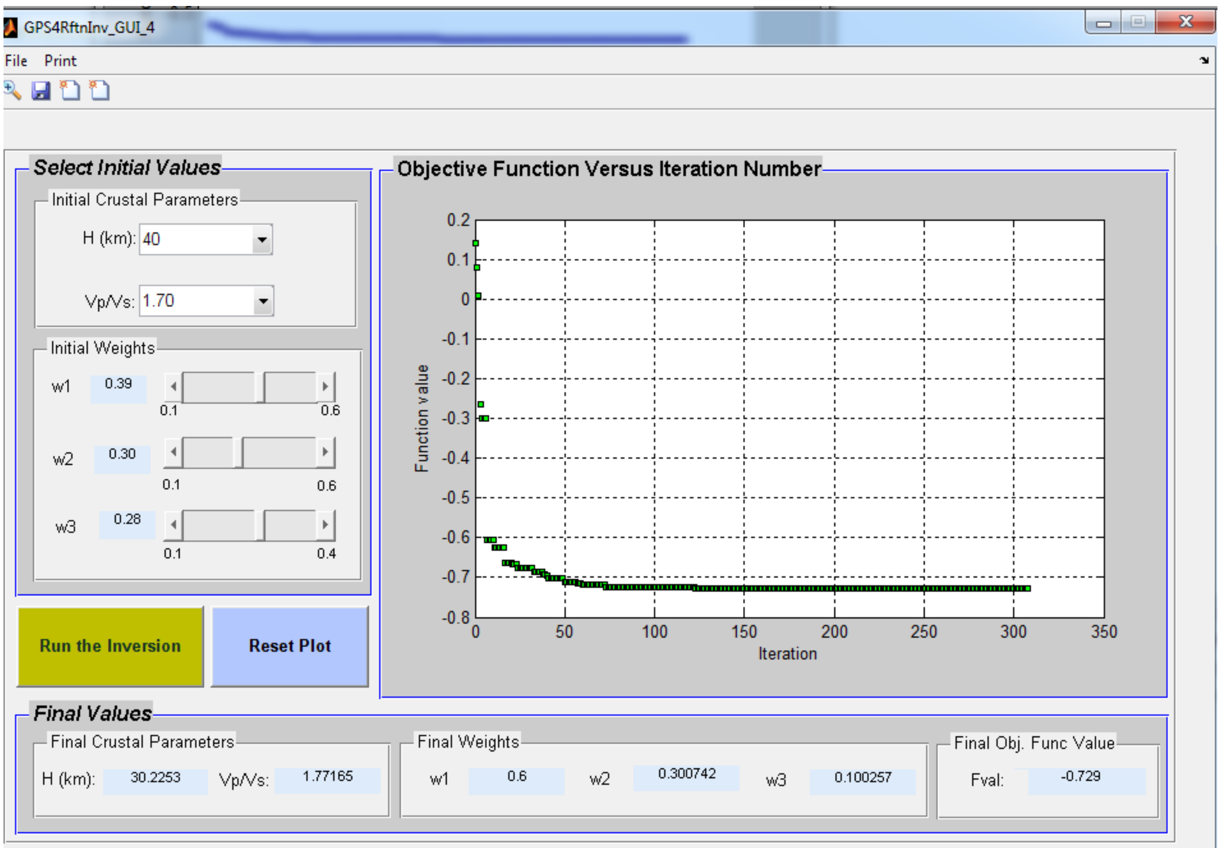
b) Medium initial H and Vp/Vs ratio values



c) Higher extreme initial H and Vp/Vs ratio values



d) Mixed of extreme initial H and Vp/Vs ratio values



e) Mixed of extreme initial H and Vp/Vs ratio values

Figure 5.1 GPS convergence test using five different combinations of initial parameters and weights.

As shown in Figure 5.1 (a)-(e) for five different initial value combinations, it is found that the GPS has converged and offers the same final values irrespective of the initial values. Table 5.3 summarizes the above results and the final values for all initial value combinations are the same.

Table 5.3
The final values of parameters, weights and objective function. FVAL is the Final Objective Function Value.

H_{opt}	30.0988
κ_{opt}	1.7750
w_{1opt}	0.6000
w_{2opt}	0.3010
w_{3opt}	0.1000
FVAL	-0.7290

Therefore, the final values are the optimal values from the GPS technique. As long as complete polling is applied, initial values can be pushed even further, even though expected parameter values for the region would not be higher than the values given in the above given conservative range.

5.2.2 On complete polling of GPS for global optimization. Since there may be many local optima in multimodal problems such as H-k stacking, we applied the GPS technique to poll all the available mesh points in every iteration. Among these we pick the one with the best objective function value. Such complete polling is observed to assist the GPS discover the global optimum instead of being trapped in a local minimum (optimum) point. As it is clearly observed in section 4.2.7 and displayed on Figure 4.10, the GPS technique provides the global optimum using a complete polling implementation. On the figure, we are able to view a 2-D H-κ surface with only H and κ parameter variation, though our search is conducted in 5-D.

Partial polling by GPS technique is considered to provide a possible optimal solution. It may be preferred when the main intent is to discover solutions faster. When partial polling is implemented, the GPS halts its polling process in a given iteration when it discovers a mesh point that improves the value of the current objective function. The point with smaller objective function value in the current iteration is appointed by the GPS to be the starting point for the next iteration. This process of GPS partial polling continues, and by so doing the GPS will involve fewer number of points in every iteration that would offer a faster computation. Even though the idea was good to find a faster and efficient GPS using partial polling, not only that the GPS ends up in a local optima in our case, but also it is not faster either as shown in Table 5.3.

Table 5.4

Comparison of complete polling and partial polling of GPS on ARBA receiver function inversion

Polling Type	Complete Polling	Partial polling
Number of iterations	216	254
Number of function evaluations	1156	1161
Time elapsed (minutes)	0.5847	0.7625

Table 5.4 clearly shows that partial polling of GPS is not faster than its complete polling counterpart. Moreover, in this particular case, the partial polling is found to be computationally more expensive, because it requires more iterations as well as more objective function evaluations. Two main reasons for many more iterations, more function evaluations and more time of computation in the partial polling compared to complete polling could be the fact that the searching speed depends on the data structure and the search starting point as well.

5.3 Comparison of Optimal Crustal Parameters and Weights with Previous Studies

Table 5.5 shows a comparison of optimal crustal parameters for the seismic station ARBA using 4 different approaches: Monte Carlo approach (Dugda et al., 2005), GA, GPS and FPN.

Table 5.5

Comparing Crustal Parameters for Station ARBA from four different approaches

Different Studies	Optimal H (km)	Optimal Vp/Vs
Genetic Algorithm implementation (Dugda et al., 2012, this study)	29.7	1.77
Monte Carlo (Dugda et al., 2005)	29.8	1.79
GPS technique	30.1	1.78
FPN technique	30.3	1.77

Table 5-5 provides an appraisal of weights obtained for seismic station ARBA using the Monte Carlo technique, GA, and GPS.

Table 5.6

Comparing Optimal Weights for seismic ARBA from three different approaches

Different Studies	Optimal w_1	Optimal w_2	Optimal w_3
Genetic Algorithm implementation (Dugda et al., 2012, BSSA):	0.5	0.4	0.1
Monte Carlo (Dugda et al., 2005)	0.5	0.4	0.1
GPS technique	0.6	0.3	0.1

5.3.1 Crustal thickness (H) values comparison for 27 seismic stations. Figure 5.2 shows a comparison of crustal thickness obtained for 27 seismic stations in Ethiopia using GPS and a Monte Carlo technique (Dugda et al., 2005). The differences for the computed crustal parameters between the two approaches are insignificant or consistent.

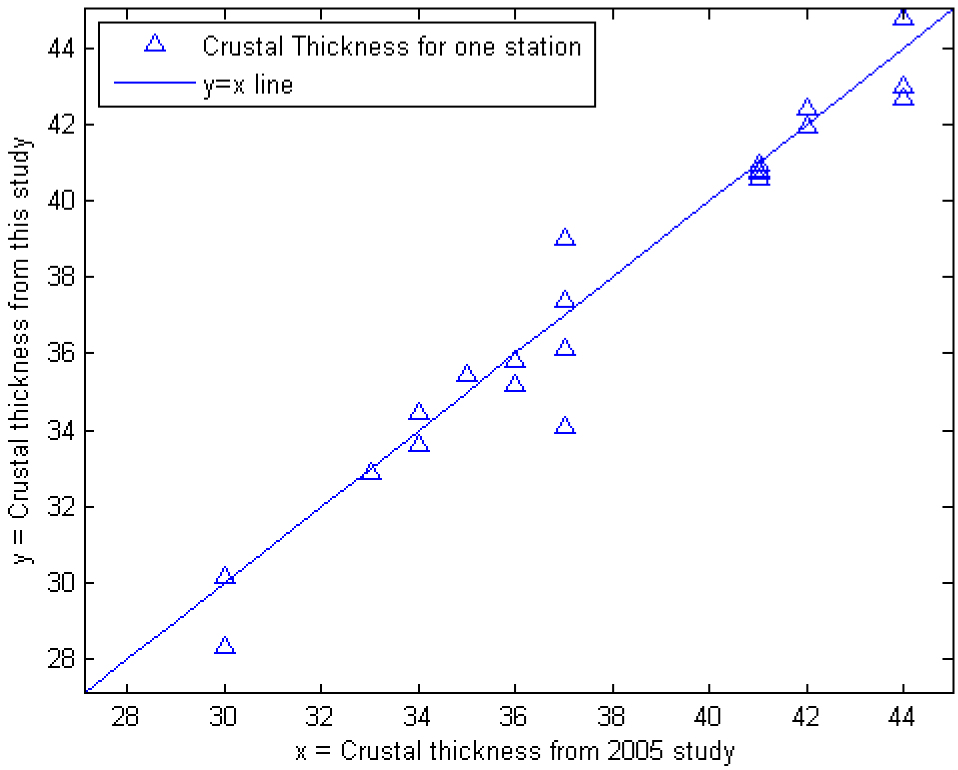


Figure 5.2 Comparison of Crustal Thickness Values from this study (GPS technique) versus results from the Monte Carlo technique (Dugda et al., 2005).

5.3.2 κ (Vp/Vs) values comparison for 27 seismic stations in Ethiopia. Figure 5.3 shows a comparison of crustal Vp/Vs ratio obtained for 27 seismic stations in Ethiopia between GPS and the Monte Carlo (Dugda et al., 2005) methods. Both crustal Vp/Vs ratio and crustal thickness results look consistent.

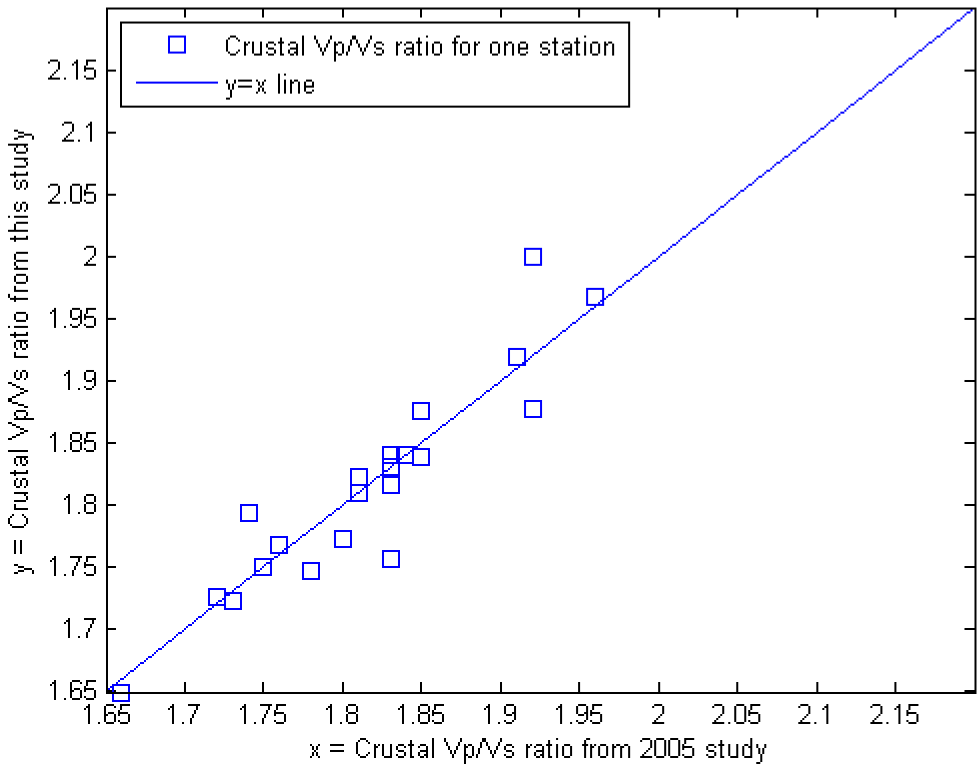


Figure 5.3 Comparison of Crustal V_p/V_s Values from this study (GPS technique) versus results from Monte Carlo (Dugda et al., 2005).

5.4 Comparison of the GA and GPS Techniques

A GA has to be run several times (up to 60) and each GA run requires 10 to 15 generations before providing a solution. As shown in Table 5.1, each GA run takes 2-3 minutes and the average of all the runs can be taken to obtain optimal solutions to the inversion problem. From several experiments, we found that GPS is faster than the GA.

However, the GA is able to escape easily from the local optima while GPS is trapped into local optima. Complete polling of the GPS technique alleviates this problem.

5.5 On FPN Technique

Since the H-κ receiver function is a multimodal surface, we can apply FPN, a niching or a dynamic clustering algorithm, to overcome the shortcomings of GPS algorithm which is sometimes caught in local optima. In just about 10 runs, each run taking less than 30 seconds, the FPN can provide the best initial parameters that should be used with GPS. In all the runs of the FPN the global optimum value will be among the few identified niche masters. We have found that the FPN obtained approximate global optimum point is very close to the right global optimum. The optimal solution by FPN is included in Table 5.5.

CHAPTER 6

Conclusions and Future Research Directions

6.1 Conclusions

In this dissertation, we have developed several techniques to solve an optimization problem of inverting receiver functions to find optimal crustal parameters and optimal weights using genetic algorithms (GA), generalized pattern search (GPS), and fitness proportionate niching (FPN). Previous study has utilized Monte Carlo technique for solving for the weights required to determine crustal parameters using the H-κ stacking of receiver functions (Dugda et al., 2005). One major objective of our work has been developing a system that is suitable for automation besides providing optimal solutions. Our algorithms have been tested using seismic data from more than twenty five seismic stations and we showed that our results are consistent with previous studies.

In the first part of this research, GA has been applied to determine optimal or near optimal weights and optimal H and κ values simultaneously. The seismic data for testing the GA implementation comes specifically from seismic station ARBA, which lies in the southern region of Ethiopia within the main Ethiopian Rift. When applied to receiver functions computed from ARBA, the three weights for the H-κ stacking converge to steady-state values within about 60 epochs. The values of the H and κ as well as the weights obtained in this dissertation match the values obtained previously with a different approach. Thus, the two studies are found to be consistent.

This study shows that GA can be used as an effective tool to determine optimal weights together with crustal H and κ parameters, as it iteratively searches many combinations of weights without being trapped in local optima due to its crossover and mutation operators. The search space also covers the entire domain with multiple candidate solutions initially generated at random. Determining the optimal parameters along with the weights simultaneously paves a way for a complete automatic crustal structure determination system. Besides, since no human intervention is required while the GA is searching to obtain the different optimal parameters, the GA can be used as an automatic crustal parameters determination tool.

In the second part of this research, an attempt has been made to implement and test the performance of the GPS technique for determining appropriate weights for the three receiver function seismic phases besides finding optimal H and κ values. GPS is a direct search method which is derivative-free, convergent and applicable to linearly constrained minimization problems (Lewis and Torczon, 2000). The GPS algorithm has been run for the same set of receiver functions from station ARBA as well as many more seismic stations. Seismic data collected in Ethiopia have been used in this study for testing the proposed GPS and FPN techniques.

Application of GPS on H-κ stacking of receiver functions enables to explore for optimal values just like it is making an exhaustive search for the weights w_1, w_2, w_3 as well as H and κ. The GPS method does the search very fast because its exploration amounts to finding the steepest descent path without computing the derivative of the objective function. Since it is not computing the derivatives, this makes the GPS to be faster in its convergence. Moreover, the GPS technique is suitable for objective functions in which finding the derivative is not easy and/or when the objective functions are not continuous. The objective function in this research satisfies both of these last conditions. Finding the derivative for our objective function is not easy and the objective function is not continuous either. Whenever the GPS is successful in its search in the current iteration, its step length increases (in our implementation the step size Δ doubles) on the next iteration. This is an important factor contributing to the faster convergence of the GPS technique.

Some of the advantages of GPS technique observed in this research include: outputs and results can be repeated seamlessly; number of iterations and number of functions evaluated are the same as long as the machine and initial conditions remain the same; and optimal values don't usually depend on the initial values. The tool developed utilizing GPS optimizes the given problem and has the following features: it is suitable for automatic processing of seismic data from all stations at the same time; it uses a user-friendly approach based on Matlab; the approach may not need much knowledge of seismology; it takes into account the quality of receiver functions through variable weights.

This study confirms that the GPS technique implemented on the H-κ stacked receiver function optimization can furnish optimal weights as well as optimal crustal parameters H and κ. The optimal crustal parameters and weights produced by GPS are consistent with the GA results. Results for station ARBA and others could make it clear that the weights and parameters found by GPS closely match those obtained previously using a different method and also the results from the GA implementation.

In conclusion, this study shows that the GPS can provide optimal weights as well as optimal crustal thickness H and κ values. Just like mutation operator helps the GA to avoid being trapped in a local optimum, the GPS uses a complete polling search to avoid being trapped in local optima. Besides, just like in the case of the GA implementation, since no human intervention is needed while the GPS is searching to obtain the different optimal parameters of H and κ together with the optimal weights. The GPS technique is a plausible technique to be used for complete automatic determination of crustal parameters.

Employment of FPN to provide best initial values for an optimization problem is a new application area introduced in this research. There are a lot of research in which initial values are needed to find a solution to their problem. When there is a potential for finding different solutions depending on the initial values, we need to find the best initial values that will lead to the right solution. Whenever an ambiguity arises in the solution we have to have a way of differentiating the best or the right solution. Among the different possible search regions in the parameter space, we have to have a way of selecting the region corresponding to the right solution. The region of the right solution is usually the region enclosing the global optimum point. This research has shown that FPN helps to cluster the different regions in the parameter space and it can help identify the region encompassing the global optimum point.

Whenever the GPS stuck in a local optimum point, FPN can step in to provide the approximate location of the global optimum point which can be used as an initial value for the GPS implementation. Applying FPN on the given receiver functions set for a limited number of iterations could provide very good initial parameters within few minutes. In fact, the best initial parameters (niche masters) identified using the FPN algorithm are very close to the location of the global optimum that the GPS or the GA provide. Thus, we conclude that GPS and FPN can be integrated for complete automatic receiver function inversion.

In this dissertation, the receiver function inversion problem is shown to be a multimodal problem. This study has also shown that GPS is a very powerful optimization tool that provides consistent results as if it searches the parameter space exhaustively. However, GPS searches the parameter space in a given pattern and computes objective function values only at the points in the given pattern. GPS even produces consistently similar results irrespective of initial values. This study also clearly shows that GPS technique offers repeatable results, especially compared to heuristic search approaches. Moreover, the number of iterations as well as the number of objective function evaluations will remain the same as long as initial values, the lower and upper bounds, and the processing machine remain the same. However, the GPS limitation of being trapped at local optima at times is solved in this study by combining GPS with FPN. This new combined technique can be called GPS-FPN.

In summary, the contributions of this work include developing techniques by implementing Genetic Algorithms (GA), Generalized Pattern Search (GPS) and Fitness Proportionate Niching (FPN) as automatic techniques to solve the problem of inverting receiver functions for a number of seismic stations at the same time. The first technique developed here makes use of the GA optimization technique. This dissertation shows that receiver function inversion with H-κ stacking is a multimodal problem and a second technique involving both direct and evolutionary techniques has been developed to solve this multimodal nature of the problem. The second technique introduced combines the direct GPS and evolutionary FPN techniques by employing their strengths. The inversion here will solve for both crustal parameters and the weights simultaneously. So far, techniques introduced have been implemented to solve one seismic station at a time and the weight assignment has been either fixed by assumption or variable. The variable weights approach introduced previously (Dugda et al., 2005) uses a Monte Carlo technique to find the variable weights for receiver functions from one seismic station. Verification of the picking of the phases was required to determine the correctness of the phase picking. The search method used to determine the best parameters then was an exhaustive search technique but it was focused on unimodal portions of the objective function at a time. The main focus of our 2005 study was on the accuracy of the then newly introduced variable weights approach, on the revelation of the crustal structure of the region and most importantly on the implication of the crustal structure discovered by that study. The variable weights approach introduced in the previous study was time-consuming, applicable to one station at a time, and it was not suitable for automatic implementation. On the other hand, there was not enough time and resource for development and enhancement of techniques in our previous study similar to the ones developed in the current study. The research and techniques developed in this dissertation take the variable weights approach to a whole new higher level by introducing new optimization techniques. Compared to the previous Monte Carlo variable weights approach, the current GA and GPS-FPN techniques have significant advantages of saving time and the new techniques are suitable for automatic and simultaneous determination of crustal parameters and appropriate weights.

6.2 Future Directions

The tools developed in this dissertation can be instrumental in geological and geophysical surveys and investigations. The technique can be used to solve similar exploration inverse problems for the petroleum or oil industry. We plan to apply the techniques developed here to geological and geophysical investigations.

It appears prudent to keep on integrating FPN and GPS further for automatic optimal receiver function inversion. At the same time, further investigation on the FPN properties is important (Workineh, 2013), though we found that FPN provides both global and local optima very quickly. Studying FPN characteristics further and comparing it to other dynamic niching techniques is one plan to extend the current study.

It is a novel idea to develop an FPN based tool as a full-fledged instrument for seismologists to observe and investigate different parts of the optimization surfaces along with the picked times of seismic phases. There are many seismic research applications which are time-consuming and sometimes frustrating endeavors in which observation of some criteria such as the objective functions of H-κ stacking need to be observed along with the picked phases. The tool planned to be developed can show all the local and global optima along with the picked arrival times corresponding to those optima locations. In this way one can for sure tell that even if one has noisy receiver functions, the results are dependable.

It is also worthwhile to devise means of expanding the application areas of GA, GPS and FPN for other inverse problems. Seismic inversion is an important approach widely used in geophysical problems to infer the subsurface properties through seismic signal measurements. We plan to expand the application area of GA, GPS and FPN.

Seismic inversion could improve exploration and management success in the oil and gas industry. Specifically we plan to implement those new techniques for inverting similar geophysical exploration problems. Thus, we may implement the GA, FPN and GPS on such geophysical exploration problems.

Even though the problem we are trying to solve arises in seismology, the existence of such a weighted sum objective function optimization problem indicates that there is a potential for such weighted sum optimization problems arising in other areas and applications too. Thus, the solution sought here can be a good start to tackle those kinds (or classes) of optimization problems.

Another area of further investigation related to this research is somewhat related to multimodal optimization and dynamic clustering techniques. Clustering has become a core mechanism of data mining and multimodal optimization problems can be mapped into clustering problems. Thus, further exploration in these important interrelated areas seems a valuable endeavor of research. For the last several decades, data analytics and optimization have been very active areas of research applied almost in every discipline and industry. Data mining, clustering being its essential component and optimization are still active areas of research and they are applied almost in every field. Further investigation of the GA, GPS and FPN techniques for optimization and clustering purposes would contribute to both the wider community research effort and to various industry applications.

References

Ammon, C. J. (1991). The isolation of receiver effects from teleseismic P waveforms, *Bull. Seism. Soc. Am.* **81** 2504-2510.

Ammon, C.J., Randall, G. E., & Zandt, G. (1990). On the nonuniqueness of receiver function inversions, *J. Geophys. Res.* **95** 15,303-15,318.

Antonio, C.D., Claudio, S.D., & Angelo, M. (2007). Where are the niches? Dynamic fitness sharing, *IEEE Trans. Evol. Comput*.11 (4), 453–465.

Aster, R., Borchers, B., and Thurber, C. (2012). *Parameter Estimation and Inverse Problems*, Second Edition, Elsevier.

Bassin, C., Laske, G., & Masters, G. (2000). The Current Limits of Resolution for Surface Wave Tomography in North America, *EOS Trans AGU*, 81, F897.

Beasley, D., Bull, D.R., & Martin, R.R. (1993). A sequential niche technique for multimodal function optimization, *Evol. Comput.* 1 (2), 101–125.

Bhattacharyya, J., Sheeham, A. F., Tiampo, K., & Rundle, J. (1999). Using a genetic algorithm to model broadband regional waveforms for crustal structure in the western United States, *Bull. Seismol. Soc. Am.* **89** 202–214.

Billings, S. D., Kennet, B. L. N., & Sambridge, M. (1994). Hypocentre location: Genetic algorithms incorporating problem-specific information, *Geophys. J. Int*. **118** 693–706.

Bostok, M. G., and Sacchi, M. D. (1997). Deconvolution of teleseismic recording for mantle structure, *Geophys. J. Int.* **129,** 143–152.

Braile, L. W. (2010). Seismic Waves and the Slinky, web.ics.purdue.edu/~braile.

Buffoni, C., Sabbione, N. C., Schimmel, M., & Rosa, M. L. (2012). Teleseismic receiver function analysis in Tierra del Fuego Island: An estimation of crustal thickness and

Vp/Vs velocity ratio, *Geophysical Research Abstracts* **14**, EGU2012-837, EGU General Assembly 2012.

Burdick, L. J., and Langston, C. A. (1977). Modeling crustal structure through the use of converted phases in teleseismic body waveforms, *Bull. Seism. Soc. Am.* **67,** 677–692.

Cassidy, J. F. (1992). Numerical experiments in broadband receiver function analysis, *Bull. Seismol. Soc. Am.* **82** 1453-1474.

Chadan, K., & Sabatier, P. C. (1977). *Inverse Problems in Quantum Scattering Theory*. Springer-Verlag.

Chang, S. J., Baag, C. E., & Langston, C. A. (2004). Joint analysis of teleseismic receiver function and surface wave dispersion using the genetic algorithm, *Bull. Seismol. Soc. Am.* **94** 691–704.

Chang, D., Xian, D., Chang, W., & Dao, M. (2010). A robust dynamic niching genetic algorithm with niche migration for automatic clustering problem, *Science Direct Journal* homepage: www.elsevier.com/locate/pr, Pattern Recognition 43, pp. 1346–1360.

Chang, D., Zhao, Y., & Xiao, Y. (2011). *A Robust Dynamic Niching Genetic Clustering Approach for Image Segmentation,* GECCO'11, July 12–16, 2011, Dublin, Ireland, 2011 ACM 978-1-4503-0557-0/11/07.

Clayton, R. W., & Wiggins, R. A. (1976). Source shape estimation and deconvolution of teleseismic body waves, *Geophys. J. R. Astr. Soc.* **47** 151-177.

Conn, A. R., Gould, N. I. M., & Toint, P. L. (1991) "A globally convergent augmented Lagrangian algorithm for optimization with general constraints and simple bounds," *SIAM Journal on Numerical Analysis*, vol. 28, 545 - 572.

Curtis, A., Dost, B., Trampert, J., & Snieder, R. (1995). Shear wave velocity structure beneath Eurasia from surface wave group and phase velocities in an inverse problem reconditioned using the genetic algorithm *EOS Trans. AGU* **76** 386.

Darwen, P., & Yao, X. (1996). Every niche method has its niche: fitness sharing and implicit sharing compared, in: Voigt, H.-M., Ebeling, W., Rechenberg, I., & Schwefel, H.-P. (Eds.) (1996). *Parallel Problem Solving from Nature* - PPSNIV, Lecture Notes in Computer Science, vol.1141,Springer, Berlin, Germany,1996, pp.398–407.

Deb, K., & Goldberg, D.E. (1989). An investigation of niche and species formation in genetic function optimization in: J.D. Schaffer (Ed.),*Proceedings of the 3rd International Conference on Genetic Algorithms*, San Mateo, CA, pp. 42–50.

Der, Z. A., R. Shumway, and A. Les (1987). Multi-channel deconvolution of *P* waves at seismic arrays, *Bull. Seism. Soc. Am.* **77,** 195–211.

Dueker, K. G., and A. F. Sheehan (1997). Mantle discontinuity structure from midpoint stacks of converted *P* and *S* waves across the Yellowstone hotspot track, *J. Geophys. Res.* **102,** 8313–8326.

Dugda, M. T., Workineh, A. T., Homaifar, A., & Kim, J. H. (2012). Receiver Function Inversion Using Genetic Algorithms, *Bulletin of the Seismological Society of America (BSSA),* Vol. 102, No. 5, pp. 2245–2251, October 2012, doi: 10.1785/0120120001.

Dugda, M. T., & A. A. Nyblade (2006). New constraints on crustal structure in eastern Afar from the analysis of receiver functions and surface wave dispersion in Djibouti, in *Geological Society Special Publications,* 259, , G. Yirgu, C. J. Ebinger and P. K. H. Maguire (Editors), 239-251.

Dugda, M. T., Nyblade, A. A., J. Julia, Langston, C. A., Ammon, C. J., & Simiyu, S. (2005). Crustal structure in Ethiopia and Kenya from receiver function analysis: Implications for Rift development in eastern Africa, *J. Geophys. Res*. **110**, B01303, doi: 10.1029/2004JB003065.

Floudas, C., & Pardalos, P. (1996). *State of the art in global optimization: Computational methods and applications. Nonconvex optimization and its applications*: Kluwer Academic, Dordrecht.

Forrest, S., Smith, R.E., Javornik, B., & Perelson, A.S. (1993). Using genetic algorithms to explore pattern recognition in the immune system, *Evol. Comput*.1, 191–211.

Frankel, A., & Vidale, J. E. (1992). A three-dimensional simulation of seismic waves in the Santa Clara Valley, California, from Loma Prieta aftershock. *Bull. Seism. Soc. Am.* **82,** 2045–2074.

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, MA.

Goldberg, D. E., & Deb, K. (1991). A comparative analysis of selection schemes used in genetic algorithms, in *Foundations of Genetic Algorithms* G. Rawlins (Editor), Morgan Kaufmann, San Mateo, 69-93.

Goldberg, D.E, & Richardson, J. (1987). *Genetic algorithms with sharing for multimodal function optimization*, in Grefenstette J.J. (Ed.), Genetic Algorithms and their Applications, Lawrence Erlbaum, Hillsdale, NJ, pp.41–49.

Goldberg, D., & Richardson, J. (1987). *Genetic algorithms with sharing for multimodal function optimization.* Paper presented at the Proceedings of the Second International Conference on Genetic Algorithms on Genetic algorithms and their application.

Goodall, C., M-estimator of location: an outline of the theory, in: Hoaglin, D.C., Mosteller, F., & Tukey, J.W. (Eds.) (1983). *Understanding Robust and Exploratory Data Analysis*, NewYork, pp.339–403.

Hammer, J. K., & Langston, C. A. (1996). Modeling the effect of San Andreas Fault structure on receiver functions using elastic 3-D finite difference, *Bull. Seism. Soc. Am.* **86,** no. 5, 1608–1622.

Hampel, F.R., Ponchotti, E.M., Rousseeuw, P.J., & Stahel, W.A. (1986). *Robust Statistics: The Approach based on Influence Functions*, Wiley, NewYork.

Haskell, N. A. (1962). Crustal reflection of plane *P* and *SV* waves, *J. Geophys. Res.* **67,** 4751–4767.

Helmberger, D. V., & Wiggins, R. (1971). Upper mantle structure of Midwestern United States, *J. Geophys. Res.* **76,** 3229–3245.

Hogans IV, J. E., & Homaifar, A. (1993). Analysis of brain scan images using genetic algorithms, in *Proc. 25th Southeast. Symp. Syst. Theory* (SSST 1993), pp. 218-222.

Homaifar, A., Goldberg, D. E., & Carroll, C. C. (1988). Boolean function learning with a classifier system**,** *Proceedings of the Applications of Artificial Intelligence VI* at the *International Society of Optical Engineering and the Computer Society of the IEEE*, Orlando, FL, April 1988, pp. 264-272.

Hooke, R. & Jeeves, T. A. (1961). Direct search solution of numerical and statistical problems, *Journal of the Association for Computing Machinery (ACM)*, 8, pp. 212–229.

Horn, J., & Goldberg, D. (1996). *Natural niching for evolving cooperative classifiers.* Paper presented at the Proceedings of the First Annual Conference on Genetic Programming, Stanford, California.

Houard, S., & Nataf, H. C. (1992). Further evidences for the ‘Lay discontinuity’ beneath northern Siberia and the North Atlantic from shortperiod *P* waves recorded in France, *Phys. Earth Planet. Interiors* **72,** 264–275.

Jain, A.K., & Dubes, R.C. (1998). *Algorithms for Clustering Data*, Prentice-Hall, Englewood Cliffs, NJ.

Jeon, T. H., Park, Y., Kang, I.-B., & Kim, K. Y. (2013). Crustal Velocity Structure beneath Wonju, Korea Using the H-κ Stacking Method and Joint Inversion of Receiver Functions and Surface-wave Dispersion Curves, *AGU Fall Meeting 2013*, San Francisco, 9-13.

Jimenez, A., Garcia, J. M., & M. D. Romacho (2005). Simultaneous inversion of the source parameters and attenuation factor using genetic algorithm, *Bull. Seismol. Soc. Am.* **95** 1401–1411, doi 10.1785/0120040116.

Jin, S., & Madariaga, R. (1993). Background velocity inversion with a genetic algorithm, *Geophys. Res. Lett.* **20** 93–96.

Jong, K. (1975). Analysis of the behavior of a class of genetic adaptive systems.

Julia, J., & Mejia, J. (2004). Thickness and *Vp*/*Vs* ratio variation in the Iberian crust, *GJI* **156** 1 59–72. doi: 10.1111/j.1365-246X.2004.02127.x.

Kolda, T. G., Lewis, R. M.., & Torczon, V. (2003). *Optimization by Direct Search: New Perspectives on Some Classical and Modern Methods*, SIAM REVIEW, 2003 Society for Industrial and Applied Mathematics (SIAM), Vol. 45, No. 3, pp. 385–482.

Langston, C. A. (1977). Corvallis, Oregon, crustal and upper mantle structure from teleseismic *P* and *S* waves, *Bull. Seism. Soc. Am.* **67,** 713–724.

Langston, C. A. (1979). Structure under Mount Rainier, Washington, inferred from teleseismic body waves, *J. Geophys. Res.* **84,** 4749–4762.

Langston, C. A. (1989). Scattering of teleseismic body waves under Pasadena, California, *J. Geophys. Res.* **94,** 1935–1951.

Langston, C. A., & Ammon, C. J. (1991). Scattering of teleseismic body waves along the Hayward-Calaveras fault system, *Bull. Seism. Soc. Am.* **81,** 576–591.

Langston, C. A., & Hammer, J. K. (2001). The Vertical Component P-Wave Receiver Function, *Bulletin of the Seismological Society of America*, 91, 6, pp. 1805–1819.

Lawrence, J. F., & Wiens, D. A. (2004). Combined receiver-function and surface wave phase-velocity inversion using a niching genetic algorithm: Application to Patagonia, *Bull. Seismol. Soc. Am.* **94** 977–988, doi 10.1785/0120030172.

Lee, J.-J., & Langston, C. A. (1983). Three-dimensional ray tracing and the method of principal curvature for geometric spreading, *Bull. Seism. Soc. Am.* **73,** 765–780.

Lewis, R. M., & Torczon, V. J. (1996). Pattern search algorithms for bound constrained minimization, Tech. Rep. 96-20, ICASE, Mail Stop 403, NASA Langley Research Center, Hampton, Virginia 23681-0001.

Lewis, R. M., & Torczon, V. J. (1997). On the convergence of pattern search algorithms, *SIAM Journal on Optimization*, 7, pp. 1–25.

Lewis, R. M., & Torczon, V. J. (1999a). "Pattern search algorithms for linearly constrained minimization," *SIAM Journal on Optimization*, vol. 10, 1999, 917-941.

Lewis, R. M., & Torczon, V. J. (1999b). "Pattern search algorithms for bound constrained minimization," *SIAM Journal on Optimization*, vol. 9, 1082-99.

Lewis, R. M., & Torczon, V. J. (2001). "A globally convergent augmented Lagrangian pattern search algorithm for optimization with general constraints and simple bounds ", *SIAM Journal on Optimization*, vol. 12, 2001, 1075-89.

Lewis, R. M., & Torczon, V. J., & Trosset, M. W. (2000). "Direct search methods: then and now," *Journal of Computational and Applied Mathematics*, vol. 124, 191-207.

Li, J. P., Balazs, M.E., Parks, G.T., & Clarkson, P.J. (2002). A species conserving genetic algorithm for multimodal function optimization, *Evol. Comput.* 10 (3), 207–234.

Li, X. Q., & Nabelek, J. L. (1999). Deconvolution of teleseismic body waves for enhancing structure beneath a seismometer array, *Bull. Seism. Soc. Am.* **89,** 190–201.

Ligorria, J. P., & Ammon, C. J. (1999). Iterative deconvolution and receiver-function estimation, *Bull. Seism. Soc. Am.* **89,** 1395–1400.

Lomax, A., & Snieder, R. (1995). The contrast in upper mantle shear-wave velocity between the East European Platform and tectonic Europe obtained with genetic algorithm inversion of Rayleigh-wave group dispersion, *Geophys. J. Int.* **123** 169–182.

Mackenzie, G. D., Thybo, H., & Maguire, P. K. H. (2005), *Crustal velocity structure across the Main Ethiopian Rift: results from two-dimensional wide-angle seismic modelling*, Geophysical Journal International, Volume 162, Issue 3, pp. 994-1006.

Makris, J. & Ginzburg, A. (1987). The Afar Depression: transition between continental rifting and sea floor spreading, *Tectonophysics, 141*, 199-214.

Mahfoud, S.W. (1994). Genetic drift in sharing methods, in: *Proceedings of the 1st IEEE Conference on Evolutionary Computation*, pp. 67–72.

Mahfoud, S.W. (1995). Population size and genetic drift in fitness sharing, in: L.D. Whitley, M.D. Vose (Eds.), *Proceedings of the Foundations Genetic Algorithms*, pp. 185–223. [53] W.M. Spears, Simple subpopulation schemes

Maronna, R.A., Martin, R.D., & Yohai, V.J. (2006). *Robust Statistics: Theory and Methods*, Wiley, NewYork.

May, J. H. (1974). *Linearly Constrained Nonlinear Programming: A Solution Method That Does Not Require Analytic Derivatives*, PhD thesis, Yale University.

Mayr, E. (1942). *Systematics and the Origin of Species from the View point of a Zoologist*, Columbia University Press, New York.

Menke, W. (2012). *Geophysical data analysis*: discrete inverse theory, Matlab ed., 3rd ed., Elsevier Inc., pp. 330, 2012.

Michalewicz, Z. (1996). *Genetic algorithms, data structures, and evolution programs*, 3rd rev. and extended ed. Berlin; New York: Springer-Verlag.

Miller, B.L., & Shaw, M.J. (1996). Genetic algorithms with dynamic niche sharing for multimodal function optimization, in: Pro*ceedings of the 1996 IEEE Transactions on Evolutionary Computation,* pp. 786–791.

Moidaki, M., Gao, S. S., Liu, K. H., & Atekwana, E. (2013). Crustal thickness and Moho sharpness beneath the Midcontinent rift from receiver functions, *Research in Geophysics*, Vol 3, No 1, doi: http://dx.doi.org/10.4081/rg.2013.e1

Mooney, W. D., & Weaver, C. S. (1989). Regional crustal structure and tectonics of the Pacific Coastal States; California, Oregon, and Washington, *Geol. Soc. Am. Memoir*, Vol. 172, 129–157.

Neves, F. A., Singh, S. C., & Priestley, K. F. (1996). Velocity structure of upper-mantle transition zones beneath central Eurasia from seismic inversion using genetic algorithms, *Geophysics* **22** 523–552.

Nyblade, A. A., & Langston, C. A. (2002). Broadband seismic experiments probe the East African rift, *EOS Trans. AGU* **83** 405-408.

Oldenburg, D. W. (1981). A comprehensive solution of the linear deconvolution problem, *Geophys. J. R. Astr. Soc.* **65,** 331–358.

Owens, T. J., & Crosson, R. S. (1988). Shallow structure effects on broadband teleseismic *P* waveforms, *Bull. Seism. Soc. Am.* **78,** 96–108.

Owens, T. J., G. Zandt, & Taylor, S. R. (1984). Seismic evidence for an ancient rift beneath the Cumberland Plateau, Tennessee; a detailed analysis of broadband teleseismic *P* waveforms, *J. Geophys. Res. B* **89,** no. 9, 7783–7795.

Pardalos, P., & Romeijn, H. (2002). *Handbook of global optimization* (Vol. 2): Springer.

Pezeshk, S., & Zarrabi, M. (2005). A new inversion procedure for spectral analysis of surface waves using a genetic algorithm, *Bull. Seismol. Soc. Am.* **95** 1801–1808.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., & Flannery, B.P. (2007). "Section 19.4. Inverse Problems and the use of a Priori Information". *Numerical Recipes: The Art of Scientific Computing* (3rd ed.). New York: Cambridge University Press.

Sambridge, M., & Drijkoningen, G. (1992). Genetic algorithm in seismic waveform inversion, *Geophys. J. Int.* **109** 323–342.

Sambridge, M., & Gallagher, K. (1993). Earthquake hypocenter location using genetic algorithm, *Bull. Seismol. Soc. Am.* **83** 1467–1491.

Sareni, B., & Krähenbühl, L. (1998). *Fitness Sharing and Niching Methods Revisited*, IEEE Transactions on Evolutionary Computation, Vol. 2, No. 3, pp. 97-106, September.

Shibutani, T., Sambridge, M., & Kennet, B. L. N. (1996). Genetic algorithm inversion for receiver functions with application to crust and uppermost mantle structure beneath Eastern Australia, *Geophys. Res. Lett.* **23** 14 1829-1832.

Shir, O., & Back, T. (2006). Niche radius adaptation in the CMA-ES niching algorithm. *Parallel Problem Solving from Nature-PPSN IX*, 142-151.

Sileny, J. (1998). Earthquake source parameters and their confidence regions by a genetic algorithm with a “memory,” *Geophys. J. Int.* **134** 228–242.

Smith, R., Forrest, S., & Perelson, A. (1993). Searching for diverse, cooperative populations with genetic algorithms. *Evolutionary computation*, *1*(2), 127-149.

Stoffa, P. L., & Sen, M. K. (1991). Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane-wave seismograms, *Geophysics* **56** 1794–1810.

Tarbuck, E. J., Lutgens, F. K., & Tasa, D. G. (2013). *Earth Science*, Pearson New International Edition, Edition 13, Pearson Publishing.

Tarantola, A. (2004). *Inverse Problem Theory and Methods for Model Parameter Estimation*, the Society for Industrial and Applied Mathematics, Philadelphia, Published by Siam.

Tou, J.T., & Gonzalez, R.C. (1974). *Pattern Recognition Principles*, Addison-Wesley, Reading, MA.

Torczon, V. (1989). *Multi-Directional Search: A Direct Search Algorithm for Parallel Machines*, Ph.D. thesis, Department of Mathematical Sciences, Rice University, Houston, TX.

Torczon, V (1991). On the convergence of the multidirectional search algorithm, SIAM Journal on Optimization, 1, pp. 123–145.

Torczon, V. (2001). "On the convergence of pattern search algorithms," *SIAM Journal on Optimization*, vol. 7, 1-25x.

VanDecar, J. C., & Crosson, R. S. (1990). Determination of teleseismic *The Vertical Component P-Wave Receiver Function* 1819 relative phase arrival times using multi-channel cross-correlation and least squares, *Bull. Seism. Soc. Am.* **80,** 150–169.

Wagner, G. S., and Langston, C. A. (1992). A numerical investigation of scattering effects for teleseismic plane wave propagation in a heterogeneous layer over a homogeneous half-space, *Geophys. J. Int.* **110,** 486–500.

Workineh, A. T. (2013). *Fitness Proportionate Niching: Harnessing the Power of Evolutionary Algorithms for Evolving Cooperative Populations and Dynamic Clustering*, North Carolina A&T State University, PhD dissertation, 2013.

Workineh A., & Homaifar, A. "Fitness Proportionate Niching: Maintaining Diversity in a Rugged Fitness Landscape" The 2012 International Conference on Genetic and Evolutionary Methods, GEM'12, Las Vegas July 16-19, 2012.

Workineh A., & Homaifar, A. "Evolving Hierarchical Cooperation in Classifiers using Fitness Proportionate Niching" journal of Complex Systems *(*submitted, Jan 02, 2013*).*

Wu, Y., Zhao, L., Chang, C., & Hsu, Y. (2008). Focal-mechanism determination in Taiwan by genetic algorithm, *Bull. Seismol. Soc. Am.* **98** 2 651–661. doi:10.1785/0120070115.

Yamanaka, H., & Ishida, H. (1996). Application of the genetic algorithms to an inversion of surface-wave dispersion data, *Bull. Seismol. Soc. Am.* **86** 436–444.

Yang, Z., & Chen, W. P. (2010). Earthquakes along the East African Rift System: A multiscale, system-wide perspective, Journal of Geophysical Research, Vol. 115, B12309, doi: 10.1029/2009JB006779.

Yang, M.S., & Wu, K.L. (2004). A similarity-based robust clustering method, *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (4), 434–448.

Yin, J. M., & Cornet, F. H. (1994). Integrated stress determination by joint inversion of hydraulic tests and focal mechanisms, *Geophys. Res. Lett.* **21** 2645–2648.

Zandt, G., & Ammon, C. J. (1995). Continental Crustal composition constrained by measurements of crustal Poisson's ratio, Nature, 374, 152-154.

Zandt, G., Myers, S. C., & Wallace, T. C. (1995). Crust and mantle structure across the Basin and Range-Colorado Plateau boundary at 37 degrees N latitude and implications for Cenozoic extensional mechanism, *J. Geophys. Res.* **100** 6 10,529-10,548.

Zhou, R., Tajima, F., & Stoffa, P. L. (1995). Application of genetic algorithm to constrain near-source velocity structure for 1989 Sichuan earthquake, *Bull. Seismol. Soc. Am.* **85** 590–605.

Zhu, L., & Kanamori, H. (2000). Moho depth variation in southern California from teleseismic receiver functions, *J. Geophys. Res.* **105** 2969-2980.

.

Appendix A

1. **Dugda,** M., Workineh, A. T., Homaifar, A., & Kim, J. H. (2012). Receiver Function Inversion Using Genetic Algorithms, *Bulletin of the Seismological Society of America (BSSA),* Vol. 102, No. 5, pp. 2245–2251, October 2012, doi: 10.1785/0120120001.
2. **Dugda,** M., Homaifar, A., & Kim, J. H. *(*2014). *Generalized Pattern Search (GPS) and Fitness Proportionate Niching (FPN) Techniques for Multimodal Receiver Function Inversion,* Institute of Electrical and Electronic Engineers - Geosciences and Remote Sensing Society (IEEE-GRSS) Transactions (in preparation).
3. **Dugda,** M., & Kebede, A. (2010). *Ratios in Higher Order Statistics (RHOS) values of Seismograms for Improved Automatic P-Phase Arrival Detection*, http://arxiv.org/abs/1007.0201, July 2010.
4. **Dugda**, M., Workineh, A. T., Kim, J. H., & Homaifar, A. (2013). *"Fast and Optimal Receiver Function Inversion Using Generalized Pattern Search and Fitness Proportionate Niching (FPN) of Genetic Algorithms Approach***,"** American Geophysical Union (AGU) Fall meeting, San Francisco, CA, *Eos Trans. AGU, 94*(52), Fall Meet. Suppl. Dec. 9-13.
5. **Dugda,** M., Homaifar, A., & Kim, J. H. (**2013**). *"Receiver Function Inversion Using Generalized Pattern Search Technique,"* the Geological Society of America (GSA) 125th year celebration and meeting in Denver Colorado, GSA Abstracts with Programs Vol. 45, No. 7, October 27-30.
6. Workineh, A., **Dugda**, M., Homaifar, A., & Lebby, G. (2012). "*GMDH and RBFGRNN Networks for Multi-Class Data Classification*" The 2012 International Conference on Artificial Intelligence, ICAI'12, Las Vegas July 16-19.

7. Workineh, A., **Dugda**, M., Homaifar, A., & Lebby, G. (2011). GMDH and RBFGRNN Networks for Multi-Class Data Classification, *CEWIT2011, the 8th International Conference & Expo on Emerging Technologies for a smarter World*, Hyatt Regency Long Island, Hauppauge, New York, Nov. 3.