TSUNAMI SOURCE INVERSION USING GENETIC ALGORITHM

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

CANER ŞEN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ENGINEERING SCIENCES

FEBRUARY 2011

Approval of the thesis:

TSUNAMI SOURCE INVERSION USING GENETIC ALGORITHM

submitted by **CANER ŞEN** in partial fulfillment of the requirements for the degree of **Master of Science in Engineering Sciences Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences**　———————————

Prof. Dr. Turgut Tokdemir
Head of Department, **Engineering Sciences**　———————————

Assoc. Prof. Dr. Utku Kânoğlu
Supervisor, **Engineering Sciences Department**　———————————

Examining Committee Members:

Prof. Dr. Ayşen Ergin
Department of Civil Engineering, METU　———————————

Assoc. Prof. Dr. Utku Kânoğlu
Department of Engineering Sciences, METU　———————————

Prof. Dr. Mehmet Zülfü Aşık
Department of Engineering Sciences, METU　———————————

Assoc. Prof. Dr. Hakan Işık Tarman
Department of Engineering Sciences, METU　———————————

Prof. Dr. Ahmet Cevdet Yalçıner
Department of Civil Engineering, METU　———————————

Date:　———————————

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: CANER ŞEN

Signature   :

ABSTRACT

TSUNAMI SOURCE INVERSION USING GENETIC ALGORITHM

Şen, Caner

M.Sc., Department of Engineering Sciences

Supervisor   : Assoc. Prof. Dr. Utku Kânoğlu

February 2011, 77 pages

Tsunami forecasting methodology developed by the United States National Oceanic and Atmospheric Administration's Center for Tsunami Research is based on the concept of a pre-computed tsunami database which includes tsunami model results from M_w 7.5 earthquakes called tsunami source functions. Tsunami source functions are placed along the subduction zones of the oceans of the world in several rows. Linearity of tsunami propagation in an open ocean allows scaling and/or combination of the pre-computed tsunami source functions. An offshore scenario is obtained through inverting scaled and/or combined tsunami source functions against Deep-ocean Assessment and Reporting of Tsunami (DART^{TM}) buoy measurements. A graphical user interface called Genetic Algorithm for INversion (GAIN) was developed in MATLAB^{TM} using general optimization toolbox to perform an inversion. The 15 November 2006 Kuril and 27 February 2010 Chile tsunamis are chosen as case studies. One and/or several DART^{TM} buoy measurement(s) is/are used to test different error minimization functions with/without earthquake magnitude as constraint. The inversion results are discussed comparing the forecasting model results with the tide gage measurements.

Keywords: Tsunami forecasting, tsunami source functions, inversion, genetic algorithm

ÖZ

TSUNAMİ KAYNAĞININ GENETİK ALGORİTMA İLE TERS ÇÖZÜMÜ

Şen, Caner

Yüksek Lisans, Mühendislik Bilimleri Bölümü

Tez Yöneticisi   : Doç. Dr. Utku Kânoğlu

Şubat 2011, 77 sayfa

Amerikan Ulusal Okyanus ve Hava İdaresi Tsunami Araştırma Merkezi tarafından geliştirilen tsunami tahmin yöntemi, önceden hesaplanmış M_w 7.5 depremler sonucu oluşan tsunami kaynak fonksiyonları veritabanına dayanmaktır. Bu kaynak fonksiyonları dünya okyanuslarının dalma-batma alanlarında birkaç sıra halinde dizilmiştir. Açık okyanustaki tsunami yayılımının doğrusallığı, tsunami kaynak fonksiyonlarının ölçeklenmesine ve/veya birleştirilmesine olanak tanır. Açık okyanus senaryosu, ölçeklenmiş ve/veya birleştirilmiş tsunami kaynak fonksiyonlarının Derin-okyanus Değerlendirme ve Tsunami Raporlama (DART[TM]) şamandıra ölçümlerine dayalı olarak ters çözümü ile elde edilir. Bu amaçla, MATLAB[TM] genel optimizasyon araç kutusu kullanılarak Ters Çözüm için Genetik Algoritma (GAIN) adlı grafik arayüzü geliştirilmiştir. 15 Kasım 2006 Kuril adaları ve 27 Şubat 2010 Şili tsunamileri örnek olarak seçilmiştir. Farklı hata minimizasyon fonksiyonları bir veya birkaç DART[TM] ölçümü ve deprem büyüklüğü kısıtlama olarak kullanılıp kullanılmadan test edilmiştir. Ters çözüm sonuçları gelgit ölçer kayıtları ile karşılaştırılarak tartışılmıştır.

Anahtar Kelimeler: Tsunami tahmini, tsunami kaynak fonksiyonları, ters çözüm, genetik algoritma

To my precious family and friends

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Tsunamis are long waves -their wavelength much larger than the depth of the ocean on which they are propagating- resulted from underwater earthquakes, volcanic eruptions, submarine or sub-aerial landslides and less common explosions or asteroid impacts. In case of an earthquake generated tsunami, underwater subduction zones between the tectonic plates cause disturbances on the ocean floor and this deformation appears as a displacement of the sea surface. Gravity tends to restore the mean sea level; hence tsunamis are called gravity waves. Tsunamis propagate across ocean basins at speeds exceeding 700 km/hour (Bernard and Robinson, 2009) and even propagate into different ocean basins as in the case of the 26 December 2004 tsunami (Titov et al., 2005b).

Tsunami is a Japanese word; 'tsu' meaning harbor and 'name' meaning wave probably belies first and unusual observation of these waves in harbors. Systematic historical documentation of tsunamis dates back to the 9th Century AD in Japan. However, the first historical report of coastal inundation by tsunamis refers to the eruption of the Thera volcano in the eastern Mediterranean, now believed around 1620 BC (Bruins et al., 2008).

Tsunamis differ from wind-generated waves and tidal waves, i.e., with a few centimeters amplitude at deep-ocean with extremely long wavelength. Observers on the deep-ocean might not realize a passing tsunami since its amplitude is small compare to the tidal wave. However, as a tsunami approaches shallow water, its wavelength decreases and amplitude increases since they are not dissipated during propagation across the ocean and their energy is conserved until they reach shallow water. When the tsunami reaches to shoreline energy amplifies the height and might cause substantial inundation.

Tsunamis can be listed as one of the world's most hazardous natural events. For example, people along the shorelines of the Indian Ocean faced with a well-known tsunami event after the earthquake at the west coast of northern Sumatra on the 26 December 2004 at 00:58:53 UTC. This is the one of the largest earthquake recorded since 1900s with a magnitude of 9.2 (Stein and Okal, 2008) and ruptured zone is estimated more than 1000 km long. This event caused 227.898 casualties and 1.7 million people were displaced due to the large inundation (U.S. Geological Survey, Sumatra, 2004). According to the United States National Oceanic and Atmospheric Administration's National Geophysical Data Center (NOAA's NGDC), over 400,000 people have died from tsunamis since 1800 (Whitmore, 2009). In order to save lives, tsunami warning systems have been established. The first warning system known as the Honolulu Observatory was established in 1949 after the 1946 Aleutian tsunami (Whitmore, 2009). After the 1960 Chile tsunami, Pacific nations coordinated under the supervision of the United Nations and established a Pacific-wide warning system. In 1968, the Honolulu Observatory is granted a task to provide warnings to the nations throughout the Pacific and became known as the Pacific Tsunami Warning Center (PTWC) (Whitmore, 2009).
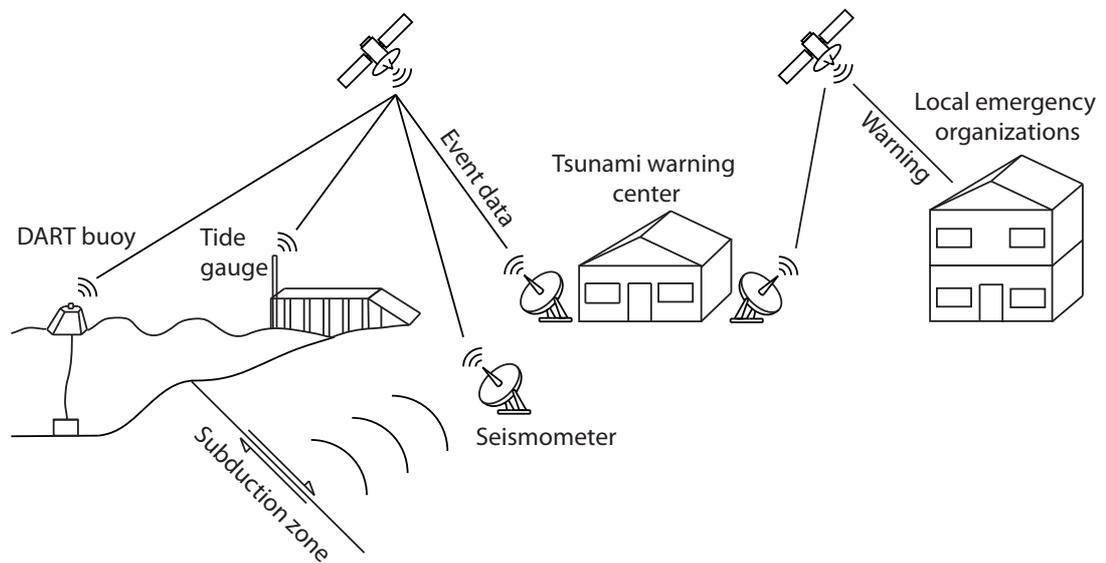
Figure 1.1: Schematic of a tsunami warning system (Whitmore, 2009).

Tsunami warning systems are composed of three basic components (Figure 1.1) (Whitmore, 2009). The first component is a tsunami warning center (TWC) which acquires tsunami data,

processes and analyzes the data and delivers event information to the authorities and public. These centers quickly resolve potential tsunami-generating events and inform coastal residents prior to wave impact. Warning communication ways are the second component of a tsunami warning system. Civil defense and local emergency organizations which respond to warnings received from a warning center are the third component of a tsunami warning system.

In this study, it is focused to develop a tool which could be used as a part of a real-time tsunami forecasting methodology in a tsunami warning center. A graphical user interface (GUI) is developed for an inversion of a tsunami source using DART measurements. First, tsunami forecasting methodology of the United States National Oceanic and Atmospheric Administration's Center for Tsunami Research (NOAA's NCTR) is described briefly. In short, their methodology based on pre-computed tsunami source functions and linearity of tsunami in an open-ocean which allows combination of these pre-computed tsunami source functions. Gica et al. (2008) explains tsunami generation, propagation, pre-computed tsunami source functions and forecast propagation database in detail. Some other tsunami warning systems are also using similar methodology (Behrens et al., 2010). In contrast, the Australian Tsunami Warning System uses pre-computed different magnitude events which do not require combination (Greenslade et al., 2011, 2009, 2007; Greenslade and Titov, 2008).

Question remains to use pre-computed scenario events is that what combination of tsunami source functions might provide satisfactory description of a tsunami source. NOAA's NCTR uses DART buoy measurements to invert the tsunami source. Percival et al. (2009) developed inversion methodology; two buoys and two tsunami source functions are used for inversion procedure and tested on four different error minimization norms using the 15 November 2006 Kuril Islands tsunami event as an example. Minimization of the sum of square of the residuals between combination of tsunami source functions and DART buoy measurements is suggested as a preferred norm. Also, it is emphasized that this norm can be used in the least square estimation since its derivative is continuous.

Here, genetic algorithm optimization is used as a solver in a graphical user interface which is developed for tsunami source inversion. Altunkaynak and Esin (2004) presented an idea about parameter estimation in complex problems. An example about genetic algorithm method for

parameter estimation in nonlinear regression is given in Altunkaynak and Esin (2004). Also, the genetic algorithm methodology is explained in detail in Haupt and Haupt (2004); Coley (1999).

A graphical user interface (GUI) is developed based on genetic algorithm for tsunami source inversion using MATLAB^{TM} global optimization toolbox (MATLAB, 2010). Even though it was not the aim for this study to develop a methodology for the tsunami source inversion, the 15 November 2006 Kuril Islands and the 27 February 2010 Chile tsunamis are used to exercise the GUI. In these exercises several features are investigated, e.g., number of DARTs used in the inversion, different norms, constraining inversion with earthquake magnitude, parameters used in genetic algorithm. Preliminary inversion results for these two events are used in forecasting models to compare the numerical modeling results with tide gage data. It is preliminary conclusion that one DART measurement provides sufficient data to estimate the tsunami source (Moore et al., 2010).

CHAPTER 2

TSUNAMI FORECASTING METHODOLOGY OF NOAA's CENTER FOR TSUNAMI RESEARCH

The United States National Oceanic and Atmospheric Administration (NOAA)'s Center for Tsunami Research (NCTR) at Pacific Marine Environmental Laboratory (PMEL) has been developing a tsunami forecasting system for the NOAA's two Tsunami Warning Centers (TWCs), located in Hawaii and Alaska (Titov et al., 2005a; Titov, 2009). The NCTR's real-time forecasting methodology is based on a constrained inversion of tsunami source using real-time tsunami measurements from Deep-ocean Assessment and Reporting of Tsunami (DART) buoys (González et al., 2005; Bernard et al., 2006; Bernard and Titov, 2006). First, tsunami source is constrained using pre-computed tsunami source functions against DART buoy measurements. Once a tsunami source is constrained high-resolution forecast inundation models are used to evaluate effects of tsunami at target coastlines in real-time(Titov et al., 1999, 2005a; Tang et al., 2009). In short, the NCTR's forecasting methodology involves three components, i.e., real-time DART measurements, pre-computed tsunami source functions and high-resolution coastal forecast inundation models.

2.1 Deep-ocean Assessment and Reporting of Tsunamis (DART)

Several real-time data sources, such as seismometers, coastal tide gages and deep-ocean wave height recordings have been used for tsunami forecast and warning (Satake et al., 2008; Whitmore, 2003; Titov, 2009). NOAA's preference for real-time forecasting is to use DART buoy measurements since it has several advantages: One, it provides a direct measure of tsunami

heights, i.e., in general it is the earliest tsunami information available. Two, DART buoy measurements are interference free from harbor and local shelf effects and allow for the application of efficient inversion methods (Tang et al., 2010). Currently, there are 52 DARTs at the oceans of the world and 40 of them belong to NOAA's National Data Buoy Center (NDBC) (Figure 2.1, NDBC (2011)).
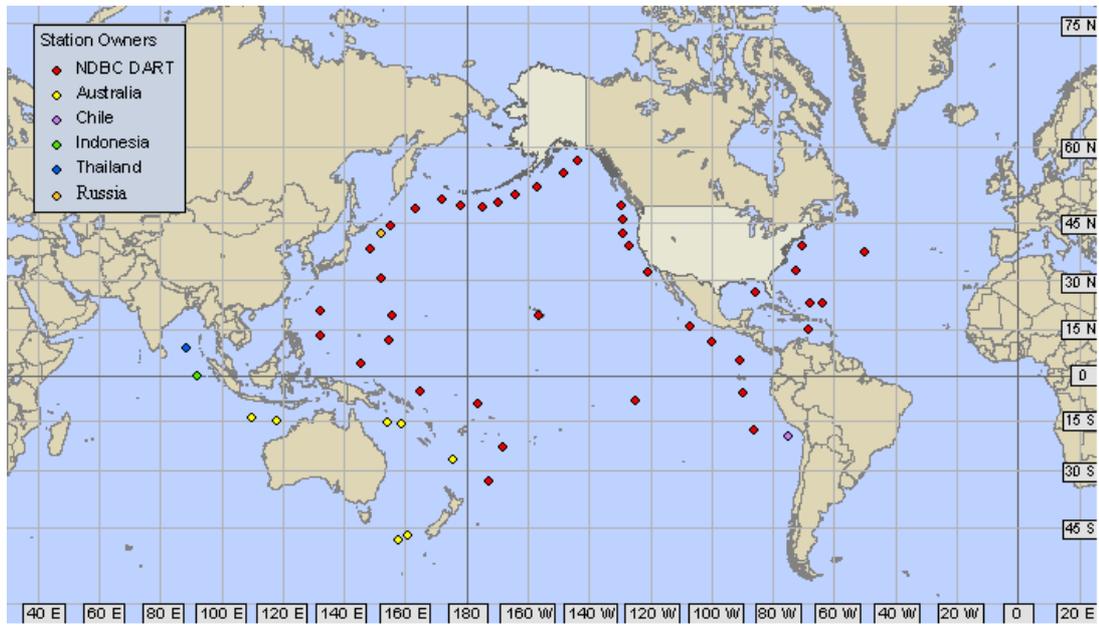
Figure 2.1: DARTs at the oceans of the world (NDBC, 2011).

DART stations are composed of two parts, i.e., tsunameter and surface buoy. Tsunameter is anchored to the ocean floor and has a central processing unit on it. Tsunameter's main task is to record pressure at the bottom of the ocean and communicate with the surface buoy bi-directionally via acoustic signals. It can be self or manually triggered in case of an earthquake. Surface buoy is also anchored to the bottom of the ocean. It carries Global Positioning System (GPS) antennas and communicates bi-directionally with Tsunami Warning Centers (TWCs) via iridium satellite in order to bridge between bottom pressure recorder (BPR) and TWCs (Figure 2.2). DART buoy measurements include both tidal wave and tsunami, thus before starting inversion process, DART buoy measurements should be de-tided (Tolkova, 2009). The raw buoy data and de-tided buoy measurements of the 27 February 2010 Chile tsunami recorded at the DARTs 43412 and 51406 are shown in Figures 2.3 and 2.4 respectively.

6

Iridium
satellite

Bi - directional
communication &
control

Tsunami warning
center

Iridium &
GPS
antennas

Lift
handle

Radar
deflector

Electronic systems &
batteries

DART buoy

1.8 m

2.6 m diameter
4000 kg displacement

25 mm chain (3.5 m)

Acoustic transducers
(2 each)

Swivel
100 m Nilspin®
25 mm nylon
22 mm nylon

1000 - 6000 m

Bottom pressure
recorder

Glass ball
floatation

75 m

19 mm nylon

13 mm chain (5 m)

Acoustic
transducer

Acoustic
release

Batteries

Anchor 3100 kg total

Anchor 340 kg

Figure 2.2: Schematic of a DART system (NDBC, 2011).
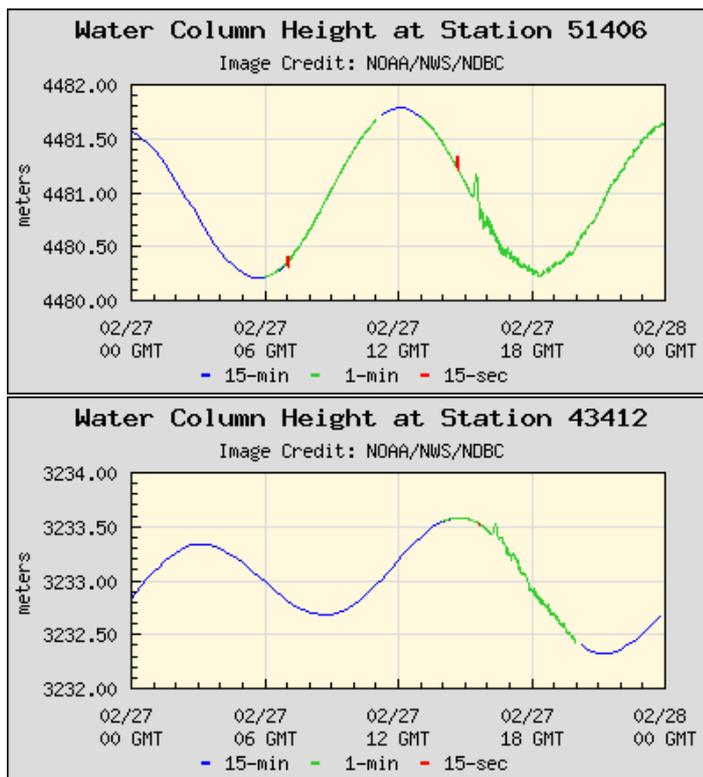
Figure 2.3: The measurements for the 27 February 2010 Chile tsunami at the DARTs 51406 and 43412 without de-tiding (NDBC, 2011). Blue line: tidal wave, green line: tsunami, red marks: station triggered.
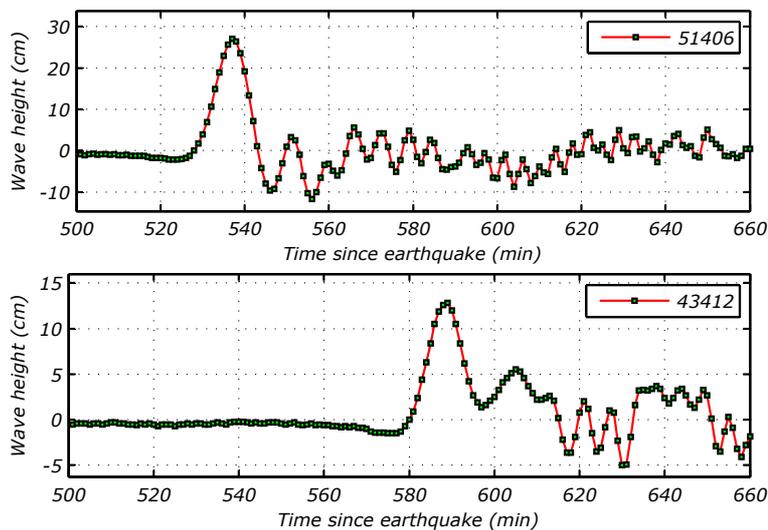
Figure 2.4: The de-tided tsunami recorded at the DARTs 51406 and 43412 for the 27 February 2010 Chile tsunami. Since the DART 51406 is closer to the earthquake epicenter, tsunami is recorded at the DART 51406 first, then it reaches to the DART 43412.

2.2 Pre-computed tsunami source functions

NCTR defined tsunami source functions along the subduction zones of the oceans of the world (Gica et al., 2008). Subduction zones and number of tsunami source functions are listed in Table 2.1 and shown in Figure 2.6. Scaling and/or combination of pre-computed tsunami source functions constrained by DART buoy measurements provides an offshore boundary and initial condition for high-resolution forecast models without additional time-consuming ocean base model run. Tsunami source function databases for the Pacific, Atlantic and Indian Oceans have been generated using tsunamigenic earthquakes with a moment magnitude M_w 7.5,

$$M_w = \frac{2}{3} \cdot log(M_0 \cdot \frac{1}{dynes \cdot cm}) - 10.7, \tag{2.1}$$

where $M_0 = \mu \cdot u_0 \cdot L \cdot W$ with predefined parameters of length $L = 100km$, width $W = 50km$, average slip $u_0 = 1m$, rake $\lambda = 90°$ (Figure 2.5) and rigidity $\mu = 4.0 \cdot 10^{11} dynes/cm^2$ resulting $M_0 = 2x10^{27} dynes·cm$. Other parameters and details of the propagation databases are explained in Gica et al. (2008). Titov et al. (1999, 2001) carried out sensitivity studies for far-field tsunamis to various parameters of the elastic deformation model explained in Gusiakov (1978) and Okada (1985). The results revealed that source magnitude and location is enough to define a far-field tsunami effect for a large group of subduction zone earthquakes.
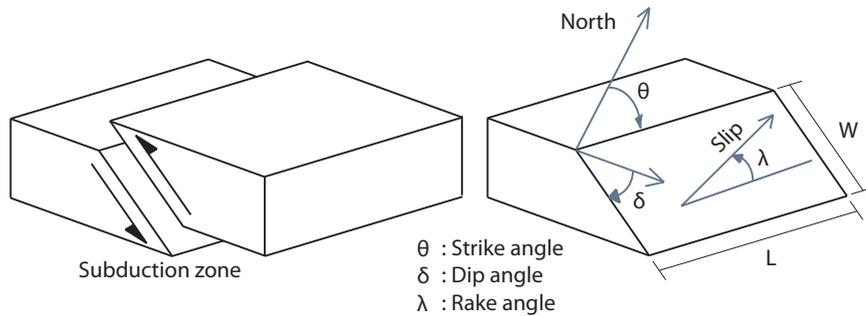
Figure 2.5: Schematic of a fault geometry.
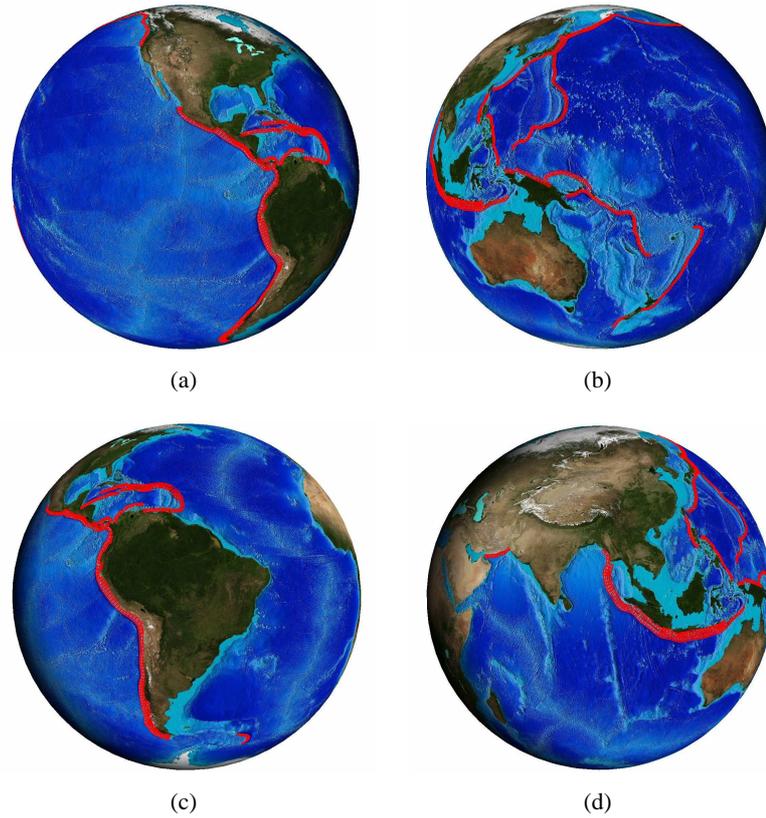
Figure 2.6: Subduction zones and tsunami source functions at (a) East Pacific, (b) West Pacific, (c) Atlantic and (d) Indian Oceans (NCTR, 2010).

Table 2.1: Tsunami source functions by region (NCTR, 2010).

Fault name	Subduction zone	Number of sources
PACIFIC OCEAN		
acsz	Aleutian-Alaska-Canada-Cascadia	184
cssz	Central-South America	382
epsz	East Philippines	44
kisz	Kamchatka-Kuril-Japan-Izu-Bonin-Marianas-Yap	222
mosz	Manus Ocean Convergence Boundary	34
ngsz	North New Guniea	30
ntsz	New Zealand-Kermadec-Tonga	78
nvsz	New Britain-Solomons-Vanuatu	74
nzsz	South New Zealand	14
rnsz	Ryukuy-Kyushu-Nankai	44
wpsz	West Philippines	22
INDIAN OCEAN		
iosz	Andaman-Nicobar-Sumatra	307
mksz	Makran	20
ATLANTIC OCEAN		
atsz	Atlantic	214
sssz	South Sandwich	22
Total number of tsunami source functions		1691

2.3 DART constrained inversion

The 27 November 2010 Chile tsunami is resulted from the magnitude of M_w 8.8 earthquake. The DART 43412 buoy measurement for this event is shown in Figure 2.4 (see also Section 5.2). In addition, time series at the DART 43412 location for the two (cssz087a and cssz091a) of the tsunami source functions generated for a magnitude of M_w 7.5 earthquakes close to the epicenter of the Chile event are shown in Figure 2.7 (see also Section 5.2). It is clear that to recover DART buoy measurement, tsunami source functions need to be scaled with some weight(s) and/or combined (sum) (Figure 2.7). Linearity of tsunami propagation in the deep-ocean allows scaling and/or combination of tsunami source functions. In other words, time series of tsunami measurements in the deep-ocean can be decomposed into a set of tsunami source functions in the time domain. Coefficients of tsunami source functions obtained from inversion process are called tsunami source coefficients (weights or scalings). The magnitude computed from the scaling and/or combination of tsunami source functions is called the tsunami moment magnitude (T_{M_w}) (Okal and Titov, 2007). During a real-time tsunami forecast, earthquake waves move much faster than tsunami, thus initial earthquake magnitude can be obtained earlier than the DART measurements. Since time is important in tsunami warning, the initial tsunami forecast is based on the earthquake magnitude only. The tsunami moment magnitude will update the forecast when it is calculated after the DART inversion using the tsunami source functions (Tang et al., 2010).

2.4 Inversion and residual minimization norms

De-tided tsunami record at a DART can be modeled by the linear combination of source functions with some residual represented as

$$b(t) = \sum_{i=1}^{n} w_i g_i(t) + e_t, \qquad (2.2)$$

where $b(t)$ is the DART measurement, $g_i(t)$s are the pre-computed source functions at the DART location with w_i scalings (weights), n is the number of tsunami source functions and e_t

11
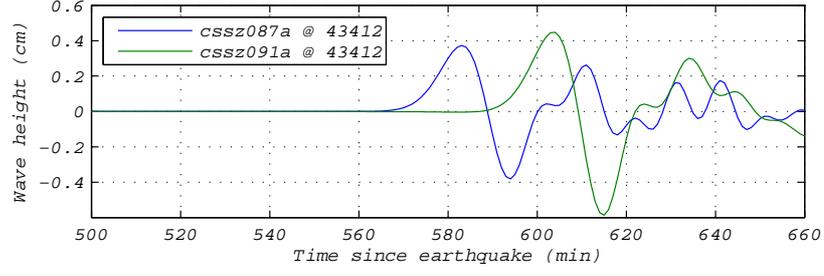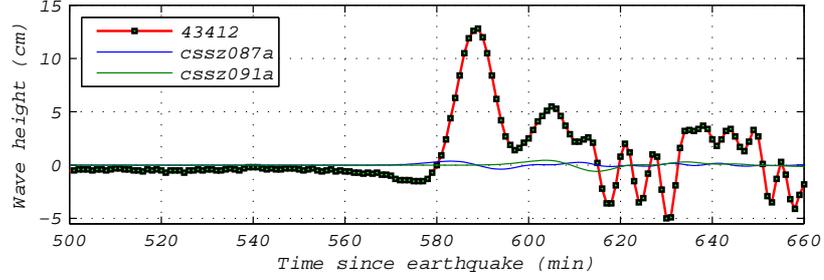
Figure 2.7: cssz087a and cssz091a tsunami source function time series (a) at the location of the DART 43412 and (b) tsunami source functions together with DART buoy record. Tsunami source functions are generated for a magnitude of M_w 7.5 earthquakes and DART buoy record is belong to the 27 November 2010 Chile tsunami with a magnitude of M_w 8.8.

is the residual. The purpose of an inversion is to minimize the residual e_t based on a measure. Following measures are used in this study;

1. make the sum of the magnitudes of the residuals as small as possible:

$$L_1(w) \equiv \sum_{t \in T} |e_t| = \sum_{t \in T} \left| b(t) - \sum_{i=1}^{n} w_i g_i(t) \right|, \qquad (2.3)$$

2. make the sum of the square of the residuals as small as possible:

$$L_2(w) \equiv \sum_{t \in T} e_t^2 = \sum_{t \in T} \left[b(t) - \sum_{i=1}^{n} w_i g_i(t) \right]^2, \qquad (2.4)$$

3. and make the largest magnitude of the residuals as small as possible:

$$L_\infty(w) \equiv max |e_t| = max \left| b(t) - \sum_{i=1}^{n} w_i g_i(t) \right|, \quad t \in T, \qquad (2.5)$$

to invert the tsunami source as given in Percival et al. (2009). In this study, inversion is performed with the genetic algorithm optimization to search for tsunami source coefficients

that will result best fit to DART buoy measurement based on the norms in Equations 2.4, 2.3 and 2.5. Genetic algorithm optimization is explained briefly with its main features in Chapter 3.

2.5 High-resolution tsunami forecast models

Once tsunami source combination is determined through an inversion (see Chapter 5) high-resolution tsunami forecast numerical models for target coastlines can be evaluated. Pre-computed time series from NCTR's database are used as the dynamic boundary conditions for the high-resolution forecast models. Although tsunami inundation is highly nonlinear, boundary/initial conditions incorporated from linear combination of source functions provide robust solutions (Tang et al., 2010). The forecast models are generated with the Method of Splitting Tsunami (MOST) model which solves nonlinear long wave equations (Titov and González, 1997). MOST model is validated through Synolakis et al. (2008). At this point, several high-resolution forecast models developed for the Pacific coastlines of the United States (Uslu et al., 2010; Arcas and Uslu, 2010; Righi and Arcas, 2010; Wei and Arcas, 2010). High-resolution forecast models involve three nested grids from low to high grid resolutions (Figure 2.8). They are designed to simulate four hours of actual propagation in approximately ten minutes and rigorously tested and validated with historical events where available.

PACIFIC OCEAN,
KURIL SUBDUCTION ZONE AND
HAWAIIAN ISLANDS

A - GRID

B - GRID

C - GRID
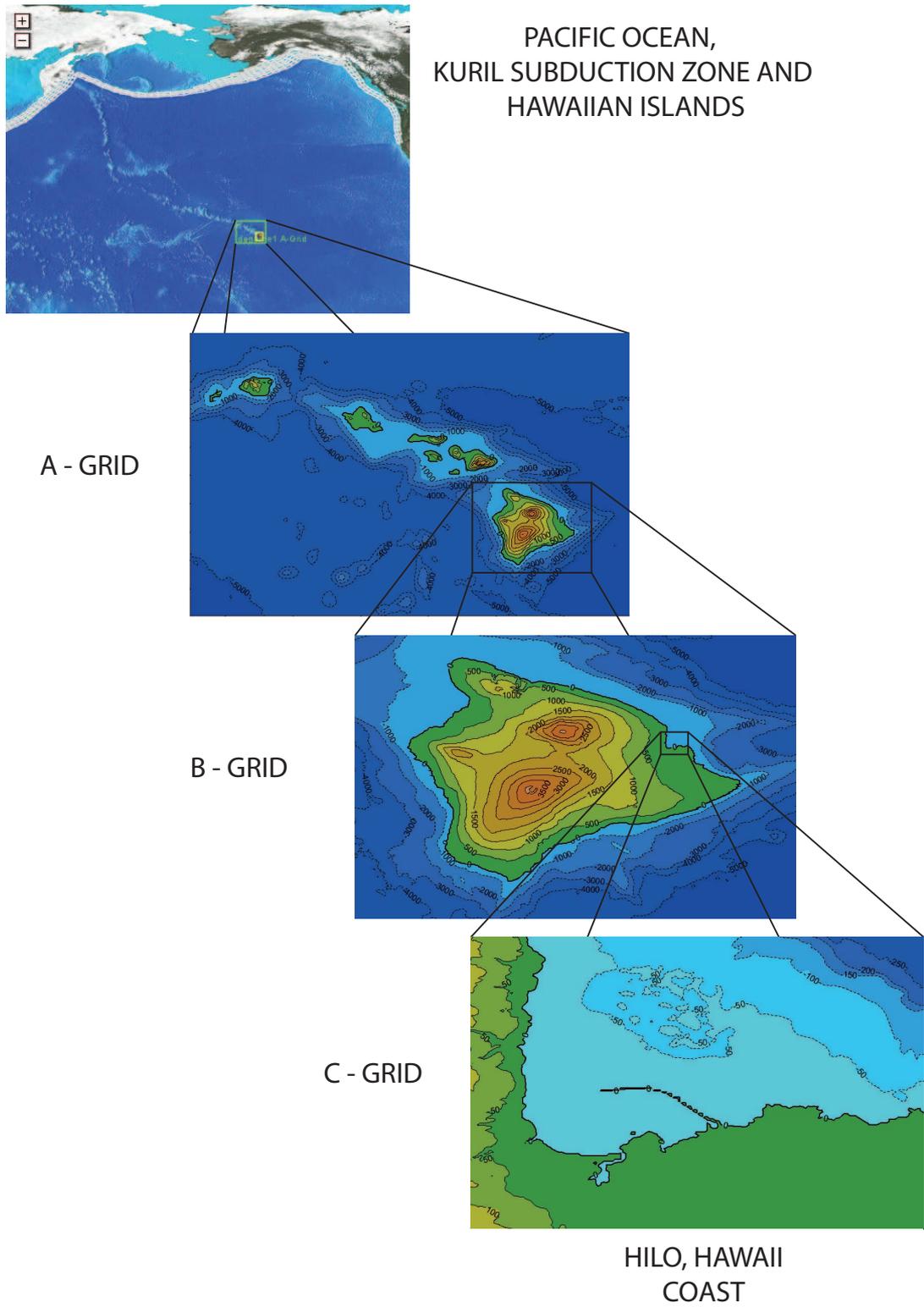
HILO, HAWAII
COAST

Figure 2.8: Three nested grids (A, B and C from low to high resolutions) are used in forecast modeling for Hilo, Hawaii coast (Tang et al., 2010).

CHAPTER 3

GENETIC ALGORITHM OPTIMIZATION

Optimization is a method of choosing best variable combination from a set of available solutions. Fitnesss or cost function is the objective of the optimization, e.g., optimum point can be minimum/maximum of a fitness (cost) function when an optimization problem is mathematically defined. Its name, variables and definition varies from one branch of science to another. Generally, real life problems might involve more than one objective and all objectives may be taken into account by giving weights depending to their importance or some of them may be neglected. For example, in design engineering, objectives can be rigidity and weight of the design. If cross-section of a structure is increased, its rigidity will increase but its weight also increases. Thus trade-off between weight and rigidity of the design is needed (Rao, 1996; Weise, 2009). A set of solutions exist in decision of this trade-off is called *pareto set* and the curve passing through the best of these sets is called *pareto frontier*. In addition to having multi objectives, variables may be bounded with lower/upper limits in order to ensure physical rules. This brings more complexity to a problem with increasing variable size. When a problem is complex and solution space is very large, genetic algorithm optimization is a very efficient method to search for variables in a large solution domain (Weise, 2009). In this chapter, main features of genetic algorithm as an optimization tool are explained over an example.

3.1 Genetic algorithm

Genetic algorithm is listed under the evolutionary algorithms family which mimics the natural selection process occurs in nature (Mitchell, 1998). Thus, genetic algorithm terminology is similar to genetics, i.e., *selection*, *crossover*, *mutation* and *elitism*. Individuals in nature are forced to evolve to survive in changing conditions and they are naturally selected since unfit ones are elected by time.

In the aspect of optimization, changing nature can be modeled as a fitness function that needs to be optimized and individuals are represented as variables that are driven on the fitness function to its optimum locus. Since similar process and operators have been applying in optimization problems similarity between optimization and genetic algorithm is revealed in 1970s (Mitchell, 1998). Genetic algorithm search in optimization starts with an initial group of variable sets called *population* which are suggested as solutions to a fitness function. The scores of the variable sets are calculated for the fitness function. After evaluating the scores, the variable sets are sorted according to their scores; designated number of suggested solutions is kept and the others are discarded. Selected variable sets called *parents* are mated amongst them to produce new variable sets in place of the discarded variable sets. This process is continued until the variable sets meet termination criteria. Sequence of genetic algorithm operators is given in the flowchart in Figure 3.1 (solid lined path). One additional step called *decoding* is added to the flowchart when the variables are binary type (dash lined path).

Functions of the genetic operators are described briefly over a single objective genetic algorithm to find the minimum of the following fitness function $z(x, y)$ with two variables, x and y,

$$z(x, y) = \frac{\left[x^2 y^2 (x + 25)(y + 25)\right]\left[\sin(x)\tanh(y) + \sin(y)\tanh(x)\right]}{10^5}, \qquad (3.1)$$

over the intervals of $-25 < x < 0$ and $-25 < y < 0$. Mesh and contour plots of the sample fitness function are shown in Figures 3.2(a) and 3.2(b) respectively. Genetic algorithm is used iteratively to find the global minimum. The initial and the final populations are tabulated in Table 3.1 and shown in Figure 3.3, including some mid-steps.

16
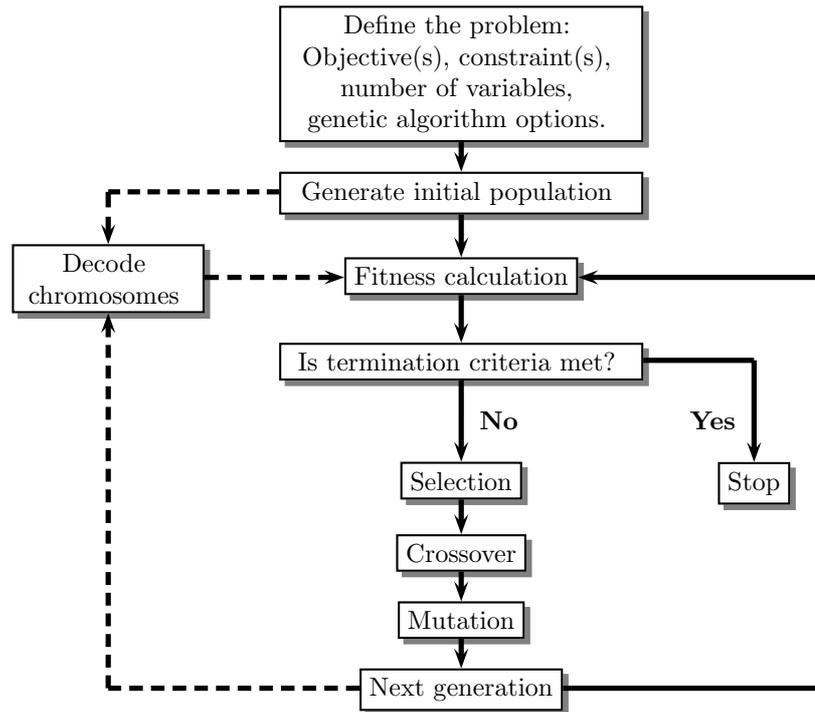
Figure 3.1: Genetic algorithm flowchart. Dash lined path is followed when the variables are binary type.

(a)   (b)
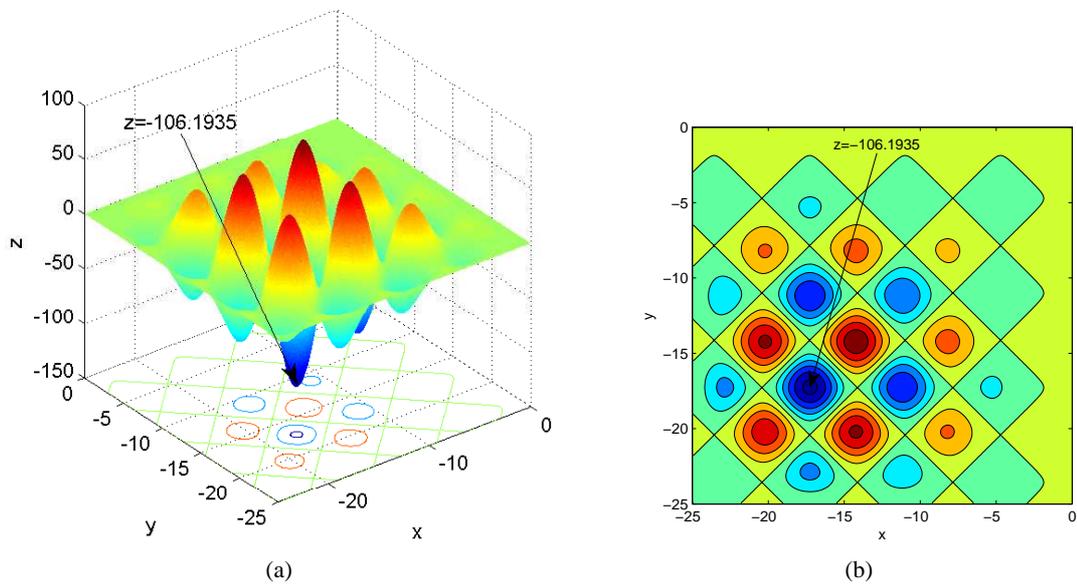
Figure 3.2: (a) Mesh and (b) contour plots of sample fitness function given in Equation 3.1. Global minimum is shown by an arrow and occurs at ($x = -17.3$, $y = -17.3$) with the value of $z = -106.1935$.

Table 3.1: Generations in the solution of the sample fitness function given in Equation 3.1. (a) The initial and (b) the sorted initial populations based on scores, (c) the potential parents, (d) the first generation and (e) the final population.

(a)

Generation=0, initial population			
Individual no	x	y	Score
1	0.0000	0.0000	0.0000
2	-2.5235	-15.8919	1.3134
3	-0.7176	-14.8062	0.3186
4	-0.3616	-1.1475	0.0006
5	-1.6274	-7.1687	0.9729
6	-14.3305	-3.4188	3.8998
7	-18.8273	-12.2248	-14.9224
8	-3.5801	-16.6320	-7.7620
9	-16.4096	-5.1976	-18.9333
10	-19.3015	-2.5441	3.0669
11	-4.5735	-12.8897	-5.7831
12	-9.1894	-16.4233	-13.0544
13	-13.3692	-0.0713	0.0003
14	-8.4356	-0.8175	0.2462
15	-2.9099	-16.8799	-2.9706
16	-7.1527	-1.4214	0.7259
17	-11.5980	-11.5084	-54.6113
18	-9.0818	-15.5383	15.1506
19	-16.3539	-17.2709	-85.4007
20	-15.2747	-14.7967	61.3401

(b)

Generation=0, sorted initial population			
Individual no	x	y	Score
19	-16.3539	-17.2709	-85.4007
17	-11.5980	-11.5084	-54.6113
9	-16.4096	-5.1976	-18.9333
7	-18.8273	-12.2248	-14.9224
12	-9.1894	-16.4233	-13.0544
8	-3.5801	-16.6320	-7.7620
11	-4.5735	-12.8897	-5.7831
15	-2.9099	-16.8799	-2.9706
1	0.0000	0.0000	0.0000
13	-13.3692	-0.0713	0.0003
4	-0.3616	-1.1475	0.0006
14	-8.4356	-0.8175	0.2462
3	-0.7176	-14.8062	0.3186
16	-7.1527	-1.4214	0.7259
5	-1.6274	-7.1687	0.9729
2	-2.5235	-15.8919	1.3134
10	-19.3015	-2.5441	3.0669
6	-14.3305	-3.4188	3.8998
18	-9.0818	-15.5383	15.1506
20	-15.2747	-14.7967	61.3401

(c)

Individual no	Binary x	Binary y	x	y	Score
19	0101100010001001	0100111100100101	-16.3539	-17.2709	-85.4007
17	1000100100111100	1000101000100111	-11.5980	-11.5084	-54.6113
9	0101011111110111	1100101011000110	-16.4096	-5.1976	-18.9333
7	0011111100110101	1000001011010001	-18.8273	-12.2248	-14.9224
12	1010000111100110	0101011111010011	-9.1894	-16.4233	-13.0544
8	1101101101010110	0101010110110000	-3.5801	-16.6320	-7.7620
11	1101000100101010	0111110000000010	-4.5735	-12.8897	-5.7831
15	1100001000110011	0101001100100110	-2.9099	-16.8799	-2.9706
1	1111111111111111	1111111111111111	0.0000	0.0000	0.0000
13	0111011100011001	1111111101000100	-13.3692	-0.0713	0.0003

(d)

	Individual no	x	y	Score
Mutated parents	19	-16.3539	-17.2709	-85.4007
	17	-11.5980	-11.5084	-54.6113
	9	-16.4096	-5.1976	-18.9333
	7	-18.8273	-12.2248	-14.9224
	12	-9.1863	-16.4233	-12.9551
	8	-3.5801	-16.6320	-7.7620
	11	-4.5735	-12.8897	-5.7831
	15	-1.3474	-16.8799	0.1692
	1	0.0000	0.0000	0.0000
	13	-13.3692	-0.0713	0.0003
Mutated offsprings	o1	-16.3413	-12.6932	-21.3349
	o2	-4.5861	-17.4674	-19.4797
	o3	-13.3692	-4.7711	-2.6713
	o4	-16.3539	-12.5711	-27.1253
	o5	-16.3394	-12.5498	-27.5150
	o6	-6.2646	-4.7211	-3.3850
	o7	-16.421	-5.1122	-18.9400
	o8	-16.3424	-17.3564	-84.6379
	o9	-7.8138	-4.8188	0.0232
	o10	-2.9725	-6.6896	0.8955

Generation=1, first generation

. . .

(e)

Generation=51, final generation		
x	y	Score
-17.2524	-17.2524	-106.3192
-17.2525	-17.2525	-106.3192
-17.2525	-17.2525	-106.3192
-17.2525	-17.2525	-106.3192
-17.2525	-17.2524	-106.3192
-17.2525	-17.2525	-106.3192
-17.2525	-17.2524	-106.3192
-17.2525	-17.2525	-106.3192
-17.2526	-17.2525	-106.3192
-17.2525	-17.2527	-106.3192
-17.2525	-17.2525	-106.3192
-17.2525	-17.2525	-106.3192
-17.2525	-17.2525	-106.3192
-17.2524	-17.2524	-106.3192
-17.2524	-17.2525	-106.3192
-17.2522	-17.2525	-106.3192
-17.2525	-17.2528	-106.3192
-17.2526	-17.2528	-106.3192
-17.2522	-17.2524	-106.3192
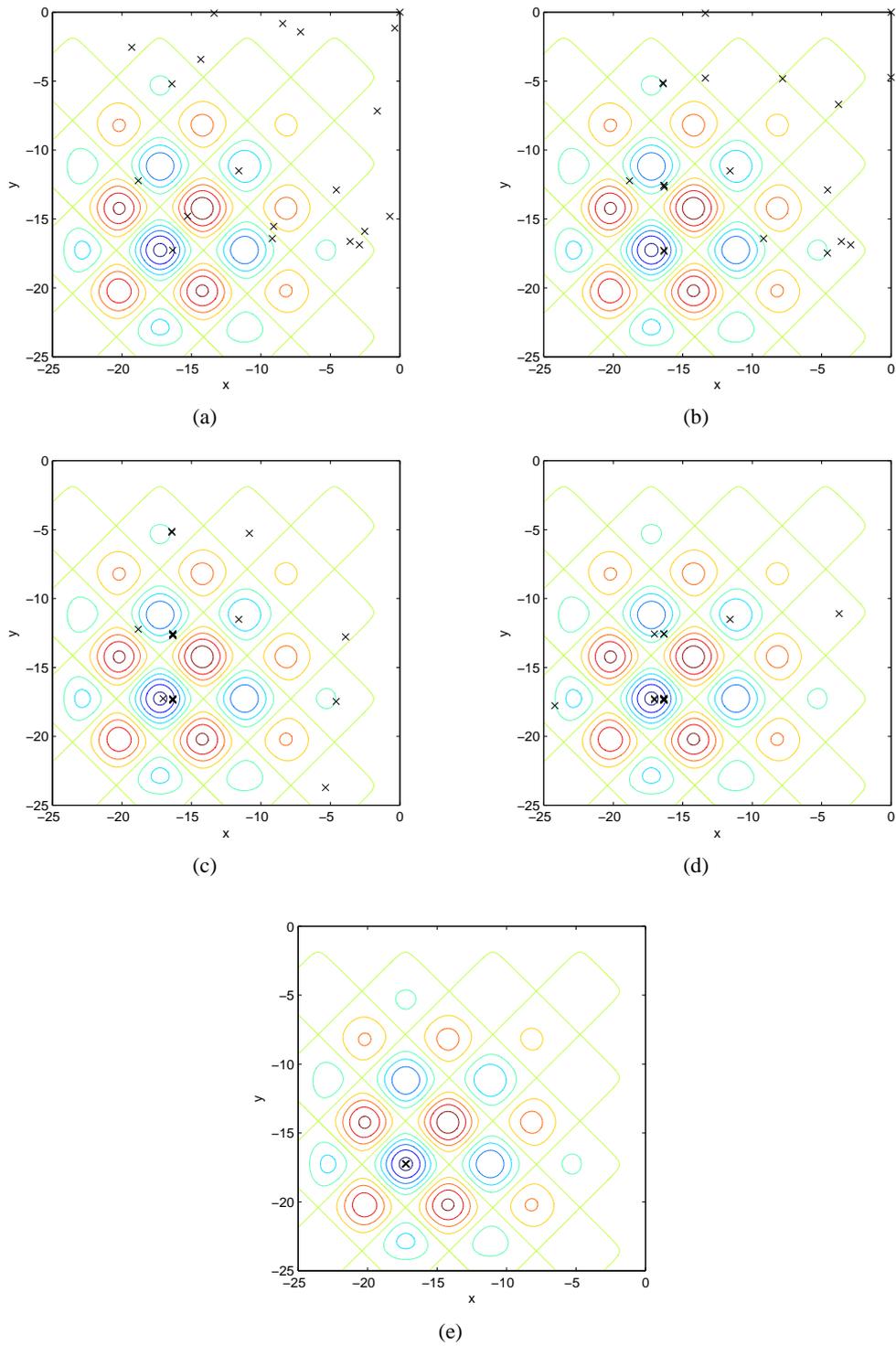
Figure 3.3: Distribution of (a) the initial population (Table 3.1(a)), (b) the first (Table 3.1(d)), (c) the second, (d) the third and (e) the final generations (Table 3.1(e)) together with the contour plot of the sample fitness function. Individuals are represented with 'x' symbols. Since some of the individuals are close to each other some of the symbols look like bold.

3.1.1 Fitness function

Fitness function is the objective of the optimization and it scores the fitness of the variable sets in the scope of the desired objective. According to the objective higher fitness means low cost so fitness function is also called cost function. Fitness function is evaluated in every iteration step and, as a fitness function becomes more complex, iteration becomes time consuming with the increasing number of variables. Therefore, fitness function for the desired problem should be wisely selected.

3.1.2 Variables

Variables in genetic algorithm are named as gene. In the example, x and y are two variables. They are written adjacently in a string form called a *chromosome* or an *individual*. Individuals are the candidate solutions and are collected in a pool of individuals. A pre-defined number of individuals called population are taken from this pool as an initial guess and their scores are calculated using the fitness function, e.g., Table 3.1(a). Scores of real type variables can be calculated directly using the fitness function. However, variables in a genetic algorithm may be represented in different forms such as binary, letter or sign. In case of variables different from real types, genetic algorithm cycle requires an additional step in which binary values are decoded to real numbers. A simple decoding function for binary variables is given as

$$v_{real} = v_{lower} + v_{decimal}\,\frac{v_{upper} - v_{lower}}{2^p - 1}, \tag{3.2}$$

(Altunkaynak and Esin, 2004). Considering the example, the real value is given as

$$v_{real} = -25 + 5\,\frac{0 - (-25)}{2^{16} - 1} = -24.9981, \tag{3.3}$$

for $v_{binary} = 0000000000000101$ provided $v_{lower} = -25$ (lower limit of the variable), $v_{upper} = 0$ (upper limit of the variable), $p = 16$ (16 bit precision) and $v_{decimal} = 5$ (decimal represen-

20

tation of the variable). Equivalents of binary variables in the defined interval of (-25, 0) are listed in Table 3.2.

Table 3.2: Variable decoding with 16 bit string.

v_{binary}	$v_{decimal}$	v_{real}
0000000000000000	0	-25.0000
0000000000000001	1	-24.9996
0000000000000010	2	-24.9992
0000000000000011	3	-24.9989
0000000000000100	4	-24.9985
0000000000000101	5	-24.9981
0000000000000110	6	-24.9977
⋮	⋮	⋮
1111111111111001	65529	-0.0023
1111111111111001	65530	-0.0019
1111111111111010	65531	-0.0015
1111111111111011	65532	-0.0011
1111111111111100	65533	-0.0008
1111111111111101	65534	-0.0004
1111111111111110	65535	0.0000

3.1.3 Constraints

Genetic algorithm optimization can produce solution in both feasible and unfeasible variable range. However, an optimum solution must be in the feasible solution domain. Thus, constraints are used to force the genetic algorithm optimization to search for a feasible region, i.e., upper and lower boundaries are defined and/or physical rules are taken into consideration. For example, in Equation 3.1, x and y are constrained in the interval of $(-25, 0)$. Generally, penalty functions are used to magnify the fitness scores of the individuals from the unfeasible region which leads to their elimination from the population.

3.1.4 Selection

Chromosomes in an initial population are scored by a fitness function (such as in Table 3.1(a)). Later, chromosomes are sorted relative to their fitness scores; individuals with minimum scores at the top of the list and maximum at the bottom (Table 3.1(b)). Then certain number of individuals with undesired scores are eliminated from the population list and the ones with the possible desired scores are kept to become potential parents according to the crossover

fraction (Table 3.1(c)). The number of individuals that are kept in the list as parents ($N_{parents}$) is determined by

$$N_{parents} = N_{pop} \cdot CF,$$ (3.4)

where N_{pop} and CF are the population size and the crossover fraction respectively. In the example, $N_{pop} = 20$ and $CF = 0.5$. Individuals with higher scores are eliminated so empty slots are opened for new offsprings. This procedure is very similar to natural selection in nature. Selecting crossover fraction is important. Small crossover fraction means fewer parents participate in mating and it limits the available genes in the offspring. On the other hand, higher crossover fraction means more parents participate in mating and this allows a chance to contribute unfeasible parents' genes to the next generation.

Next step is to select parent candidates from the population and to pair the parents. Pairing is applied to the selected parent's pool to reproduce new offsprings to fill the empty slots in the whole population. Pairing can be done in different ways, e.g., *roulette wheel selection* and *tournament selection*.

3.1.4.1   Roulette wheel selection

The probabilities are assigned to the potential parents in inversely proportional to their costs. Individual with the lowest and highest costs have the greatest and the lowest probability of being selected as a parent respectively. A random number generator is used to generate a weight in an interval of $0 < w < 1$ and this weight determines which parent is being selected among the assigned weights. Weighting can be done according to rank or cost;

1. Rank weighting: Probabilities of each individual according to their rank are calculated from

$$P_n = \frac{N_{parents} - n + 1}{\sum_{i=1}^{N_{parents}} i},$$ (3.5)

where $N_{parents}$ is the number of parents ($N_{parents}$ = 10 in the example) and listed in Table 3.3 and shown in Figure 3.4 (Haupt and Haupt, 2004). Cumulative probabilities of individuals are calculated by summing probability of the chromosome with the probabilities of the chromosomes above this chromosome. A random number between zero and one is generated. Starting from the top of the list, first chromosome with a greater cumulative probability than the random number is selected as a parent. For example, when the random number is 0.5, then it is in the range of $0.490909 < w < 0.618182$, i.e., random number is greater than the probability of third chromosome but smaller than the fourth one (Table 3.3). Therefore, chromosome in the fourth row is selected as the parent. This procedure is continued until all the parents are mated. If the chromosome is mated by itself then the selection can be repeated. Big populations have a low probability of mating by itself.

Table 3.3: The probabilities for the rank weighting for the generation=0 given in Table 3.1(c).

n	Individual no	Chromosome		Score	P_n	$\sum_{i=1}^{n} P_i$
		x	y			
1	19	-16.3539	-17.2709	-85.4007	0.181818	0.181818
2	17	-11.5980	-11.5084	-54.6113	0.163636	0.345455
3	9	-16.4096	-5.1976	-18.9333	0.145455	0.490909
4	7	-18.8273	-12.2248	-14.9224	0.127273	0.618182
5	12	-9.1894	-16.4233	-13.0544	0.109091	0.727273
6	8	-3.5801	-16.6320	-7.7620	0.090909	0.818182
7	11	-4.5735	-12.8897	-5.7831	0.072727	0.890909
8	15	-2.9099	-16.8799	-2.9706	0.054545	0.945455
9	1	0.0000	0.0000	0.0000	0.036364	0.981818
10	13	-13.3692	-0.0713	0.0003	0.018182	1.000000
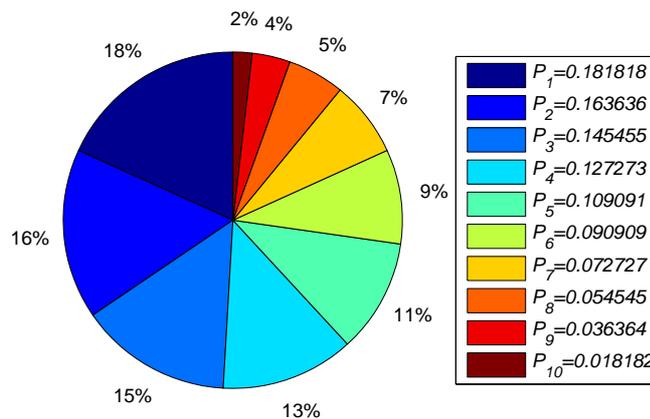
Figure 3.4: The roulette wheel for the rank weighting based on probabilities in Table 3.3.

2. Cost weighting: Probability of each individual according to their cost is calculated from

$$P_n = \left| \frac{C_n}{\sum_i^{N_{parents}} C_i} \right|, \tag{3.6}$$

where $C_n = c_n - c_{N_{parents}+1}$. Probabilities for the example are listed in Table 3.4 and shown in Figure 3.5 (Haupt and Haupt, 2004). Here, C_n is the normalized cost for each individual, c_n is the cost for each individual and $c_{N_{parents}+1} = 0.0006$ which is the cost of the lowest among the discarded individuals from sorted initial population (Table 3.1(b)). Apart from the rank weighting, probabilities must be recalculated each generation since fitness score is changing from one generation to another. Also, pies in the roulette wheel (Figure 3.5) are unfairly distributed when there are big differences in cost weighting.

Table 3.4: The probabilities for the cost weighting for the generation=0 given in Table 3.1(c).

n	Individual no	Chromosome		Score	P_n	$\sum_{i=1}^{n} P_i$
		x	y			
1	19	-16.3539	-17.2709	-85.4007	0.419779	0.419779
2	17	-11.5980	-11.5084	-54.6113	0.268437	0.688216
3	9	-16.4096	-5.1976	-18.9333	0.093067	0.781284
4	7	-18.8273	-12.2248	-14.9224	0.073352	0.854636
5	12	-9.1894	-16.4233	-13.0544	0.064170	0.918806
6	8	-3.5801	-16.6320	-7.7620	0.038156	0.956962
7	11	-4.5735	-12.8897	-5.7831	0.028429	0.985391
8	15	-2.9099	-16.8799	-2.9706	0.014605	0.999996
9	1	0.0000	0.0000	0.0000	0.000003	0.999999
10	13	-13.3692	-0.0713	0.0003	0.000001	1.000000

3.1.4.2 Tournament selection

Tournament selection scheme picks a small group of individuals from the mating pool. Then the individual having the lowest cost from this group becomes a parent. For example, three individuals -14.9224, -7.7620 and 0.0003 from Table 3.1(c) are selected as the tournament group. The individual with the cost -14.9224 is the lowest cost so it is elected as the parent. This process is continued to pair all the parents. Since sorting is a time consuming step, tournament selection is good for large populations. In this study, roulette wheel and tournament
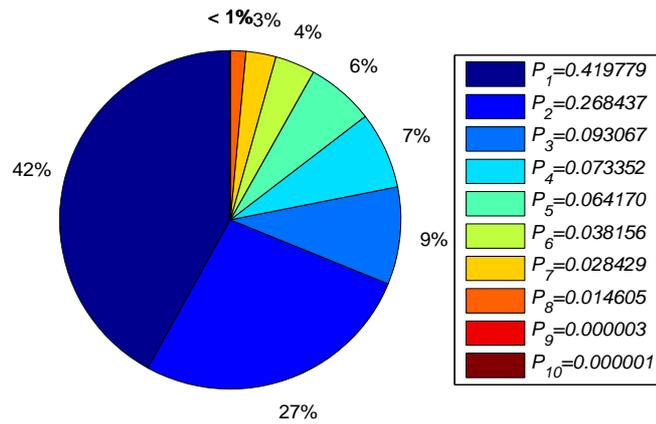
Figure 3.5: The roulette wheel for the cost weighting based on probabilities in Table 3.4.

selections are used as default operators in single-objective and multi-objective optimizations respectively.

3.1.5 Crossover

Crossover operator mates two parents to breed new individuals called offspring by exchanging bits between the random point(s) of parent chromosome strings. When a single point is selected it is called single point crossover (Figure 3.6(a)). In single point crossover, chromosome part on the left of the point are taken from first parent and chromosome part on the right of the point are taken from second parent. They are combined to form first offspring and the remaining chromosome parts are used to form the second offspring. When two points are selected it is called two points crossover where parents exchange chromosomes between two points similar to the single point crossover (Figure 3.6(b)). Again, the remaining chromosome parts are used to form the second offspring. Also, there is a scattered crossover (Figure 3.6(c)). In this type, a random mask string with a size equal to the parent chromosome is selected. Then from head to tail bits on the mask are checked, if the bit on the mask is one gene is taken from the first parent otherwise from the second parent; the remaining chromosome parts are used to form the second offspring. Crossover operator generates new individuals exist in the solution domain and drives a searching mechanism.

Parent A	0	0	1	0	▌	1	1	1	0	0	0	1	0	0	0	1	1
Parent B	0	1	1	0	▌	1	0	1	1	0	1	1	0	1	1	0	0
Offspring 1	0	0	1	0	▌	1	0	1	1	0	1	1	0	1	1	0	0
Offspring 2	0	1	1	0	▌	1	1	1	0	0	0	1	0	0	0	1	1

(a)

Parent A	0	0	1	▌	0	1	1	1	0	0	0	1	0	▌	0	0	1	1
Parent B	0	1	1	▌	0	1	0	1	1	0	1	1	0	▌	1	1	0	0
Offspring 1	0	0	1	▌	0	1	0	1	1	0	1	1	0	▌	0	0	1	1
Offspring 2	0	1	1	▌	0	1	1	1	0	0	0	1	0	▌	1	1	0	0

(b)

Parent A	0	0	1	0	1	1	1	0	0	0	1	0	0	0	1	1
Parent B	0	1	1	0	1	0	1	1	0	1	1	0	1	1	0	0
Mask	0	1	1	1	1	0	0	1	0	0	1	0	1	1	0	0
Offspring 1	0	0	1	0	1	0	1	0	0	1	1	0	0	0	0	0
Offspring 2	0	1	1	0	1	1	1	1	1	0	0	1	0	1	1	1

(c)

Figure 3.6: Different crossover types; a) single point, b) two points and c) scattered crossovers.

Table 3.5: The first population after scattered crossover of the potential parents in Table 3.1(c). The prefix 'o' shows offsprings.

	No	Binary x	Binary y	x	y	Score
Parents	19	0101100010001001	0100111100100101	-16.3539	-17.2709	-85.4007
	17	1000100100111100	1000101000100111	-11.5980	-11.5084	-54.6113
	9	0101011111110111	1100101011000110	-16.4096	-5.1976	-18.9333
	7	0011111100110101	1000000101101001	-18.8273	-12.2248	-14.9224
	12	1010000111100110	0101011111010011	-9.1894	-16.4233	-13.0544
	8	1101101101010110	0101010110110000	-3.5801	-16.6320	-7.7620
	11	1101000100101010	0111110000000010	-4.5735	-12.8897	-5.7831
	15	1110001000110011	0101001100100110	-2.9099	-16.8799	-2.9706
	1	1111111111111111	1111111111111111	0.0000	0.0000	0.0000
	13	0111011100011001	1111111101000100	-13.3692	-0.0713	0.0003
Offsprings	o1	0101100010101010	0111111000000101	-16.3413	-12.6932	-21.3349
	o2	1101000100001001	0100110100100010	-4.5861	-17.4674	-19.4797
	o3	0111011100011001	1100111100100100	-13.3692	-4.7711	-2.6713
	o4	0101100010001001	0111111101000101	-16.3539	-12.5711	-27.1253
	o5	0101100010101111	0111111101111101	-16.3394	-12.5498	-27.5150
	o6	1111111111011001	1100111110100111	-0.0145	-4.7211	0.0000
	o7	0101011111011001	1100101110100110	-16.421	-5.1122	-18.9400
	o8	0101100010100111	0100111001000101	-16.3424	-17.3564	-84.6379
	o9	1010111111111100	1100111010100111	-7.8138	-4.8188	0.0232
	o10	1101100100111111	1011101101111111	-3.7842	-6.6896	-0.5088

Table 3.6: The population after the mutation of the individuals in Table 3.5 with the probability rate of 0.006. The prefix 'o' shows offsprings, the mutated parts are highlighted and the mutated bits are bolded. The first two individuals with scores -85.4007 and -54.6113 are defined as elite and they are excluded from mutation.

	No	Binary x	Binary y	x	y	Score
Mutated parents	19	0101100010001001	0100111100100101	-16.3539	-17.2709	-85.4007
	17	1000100100111100	1000101000100111	-11.5980	-11.5084	-54.6113
	9	0101011111110111	1100101011000110	-16.4096	-5.1976	-18.9333
	7	0011111100110101	1000001011010001	-18.8273	-12.2248	-14.9224
	12	1010000111101110	0101011111010011	-9.1863	-16.4233	-12.9551
	8	1101101101010110	0101010110110000	-3.5801	-16.6320	-7.7620
	11	1101000100101010	0111110000000010	-4.5735	-12.8897	-5.7831
	15	1111001000110011	0101001100100110	-1.3474	-16.8799	0.1692
	1	1111111111111111	1111111111111111	0.0000	0.0000	0.0000
	13	0111011100011001	1111111101000100	-13.3692	-0.0713	0.0003
Mutated offsprings	o1	0101100010101010	0111111000000101	-16.3413	-12.6932	-21.3349
	o2	1101000100001001	0100110100100010	-4.5861	-17.4674	-19.4797
	o3	0111011100011001	1100111100100100	-13.3692	-4.7711	-2.6713
	o4	0101100010001001	0111111101000101	-16.3539	-12.5711	-27.1253
	o5	0101100010101111	0111111101111101	-16.3394	-12.5498	-27.5150
	o6	1011111111011001	1100111110100111	-6.2646	-4.7211	-3.3850
	o7	0101011111011001	1100101110100110	-16.421	-5.1122	-18.9400
	o8	0101100010100111	0100111001000101	-16.3424	-17.3564	-84.6379
	o9	1010111111111100	1100111010100111	-7.8138	-4.8188	0.0232
	o10	1101100110111111	1011101101111111	-2.9725	-6.6896	0.8955

3.1.6 Mutation

Crossover operator generates offsprings, which inherit the genetic map of the parents, and sustains the searching mechanism until it reaches to an optimum point. When there are more than one minimum/maximum points mechanism can be stuck in a local minimum/maximum. Mutation operator takes control at this point to jump out of a local minimum/maximum. Mutation makes small changes on the population to bring genetic diversity and generates individuals out of family which results in strengthened search. Mutation operator makes its job by inverting some random bits in the population with a predefined probability rate. This rate is used to calculate number of mutated bits (N_M) by using,

$$N_M = MR \cdot N_{pop} \cdot N_{string}, \tag{3.7}$$

where MR, N_{pop} and N_{string} are the mutation probability rate, the population size and the length of chromosome respectively. First population after mutation of individuals in Table 3.5 with the probability rate of 0.006 are presented in Table 3.6.

3.1.7   Elitism

In genetic algorithm search, a predefined number of top scored individuals in the population are guaranteed to survive to the next generation by elitism. This is done by excluding defined number of best individuals from mutation operator. Thus, their genetic behavior remains same and they keep providing offsprings with good fitness in next generation. In Table 3.6, the first two individuals with scores -85.4007 and -54.6113 are defined as elite and they are excluded from mutation.

3.1.8   Termination criteria

A termination criterion defines the conditions to end the iteration. Iteration is terminated when the iteration number reaches to a desired maximum number of generation or maximum time is elapsed or the cumulative change in the fitness function is below a defined tolerance.

CHAPTER 4

INTERFACE DESIGN FOR TSUNAMI SOURCE INVERSION

A graphical user interface (GUI) explained in this chapter is designed in order to use DART buoy measurements and pre-computed tsunami source functions to evaluate possible tsunami source. Inversion is processed through a genetic algorithm by minimizing the residual between DART buoy measurements and combination of tsunami source functions based on the norms defined in Chapter 2, so that weights for tsunami source functions are determined. In order to investigate for a better inversion methodology substantial number of scenarios need to be tested. This necessitated an interface to make the task simple.

4.1   Graphical user interface

The GUI called Genetic Algorithm for INversion (GAIN) is developed in MATLAB$^{\text{TM}}$. GAIN is designed to use DART buoy measurements and pre-computed tsunami source functions results as an input and inverts possible tsunami source with a minimum user effort, providing weights for tsunami source functions as explained in Chapter 2. It also allows to plot the results. It can be run from command line in MATLAB$^{\text{TM}}$ and also through the GUI developed here (Figure 4.1). GAIN uses the general optimization toolboxes (MATLAB, 2010), i.e., genetic algorithm or multi-objective genetic algorithm solvers. If a single DART station is selected simple genetic algorithm is used otherwise multi-objective genetic algorithm solver is used. Interface allows choosing DART buoy measurement(s) and pre-computed tsunami source function(s) data related to the event. It also allows choosing different norms as fitness functions and earthquake moment magnitude as a constraint. Basically, it uploads DART
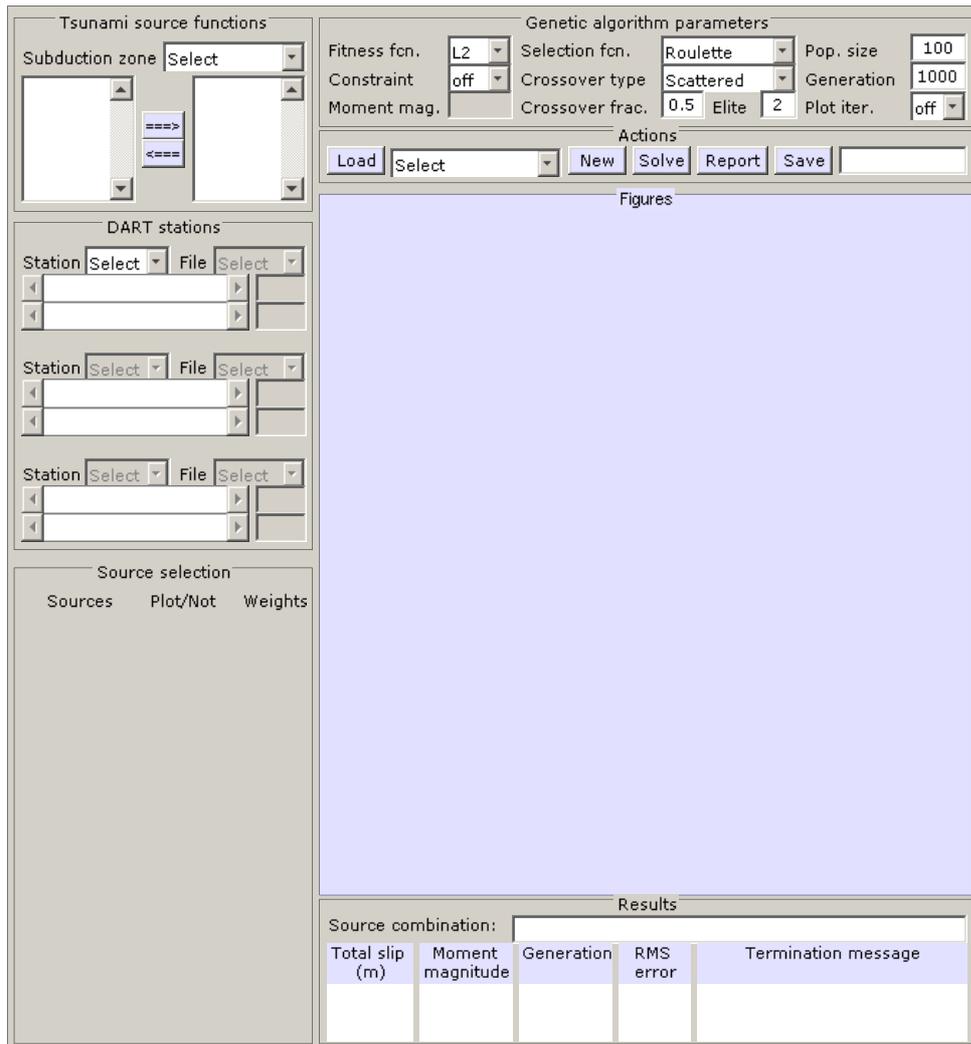
Figure 4.1: Genetic Algorithm for INversion (GAIN) interface.

buoy measurements and pre-computed tsunami source functions and considers fitness function of choice as an input. Then submits this information to the genetic algorithm solver and calculates weights for tsunami source functions.

Objects are grouped into seven panels in GAIN (Figure 4.1), i.e., tsunami source functions, DART stations, source selection, genetic algorithm parameters, actions, figures and results panels. Panels and their functions are explained in the following subsections in detail.

4.1.1 Tsunami source functions panel

Tsunami source functions panel is composed of one popupmenu, two listboxes and two push-buttons (Figure 4.2(a)). For a new event, subduction zone popupmenu is used to select the zone at which earthquakes occurred. The subduction zones and their abbreviations are listed in Table 2.1. Tsunami source function names for the selected subduction zone appear on the left listbox (Figure 4.2(b)). Right and left arrow pushbuttons are used to select/deselect the source function names, i.e., transfers from the left to the right listbox (select) or the right to the left listbox (deselect) (Figure 4.2(c)). After selecting the sources, the source function names in each listbox are sorted alphabetically. The number of selected source functions is limited to twenty at this point. At the end, the tsunami source functions that are listed in right listbox are the sources which will be uploaded from the propagation database and could be used for inversion; see Subsection 4.1.3.
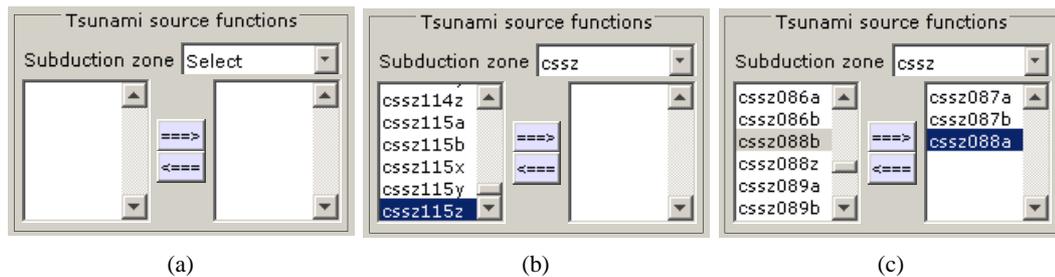
Figure 4.2: Different stages of fault panel; (a) initial view, (b) showing all the sources for the selected subduction zone (right listbox is empty) and (c) after selecting the possible tsunami source functions close to an earthquake epicenter.

4.1.2 DART stations panel

After selecting the tsunami source functions, it is required to upload pre-computed source functions time series at DART buoy locations and DART buoy measurements. DART stations panel allow this and carries two popupmenus, two edit boxes and two slider bars under each DART station (Figure 4.3(a)) buoy measurements. Popupmenus are designed to choose a DART station and DART buoy measurement for an event. Choosing a DART station allows determining the locations where tsunami source function time series are uploaded. Once the

DART station and the related DART buoy measurement file are selected from popupmenus, vertically aligned slider bars could be used to define DART buoy measurement time interval which will be used for an inversion (Figure 4.3(c)). In addition, upper and lower edit boxes could also be used to define a time interval. Selected time interval is shown with black thick solid lines in the figure panel; see Section 4.1.6. DART stations panel is designed to use up to three DART buoy measurements for inversion simultaneously. Initially, all DART stations panel objects are inactive except the first station panel (Figure 4.3(a)). Selecting the first station from the panel activates the second DART station panel and selecting the second station activates the third DART station panel. When a DART station and a file is selected from popupmenus GAIN uploads selected DART buoy measurements and time series of selected tsunami source functions at the DART station(s) (Figure 4.3(b)). Also, selected DART buoy measurements and tsunami source function time series are trimmed in the interval defined in the DART station panels. The trimmed parts are saved as variables to the workspace since these trimmed time series are required in inversion process. File and folder structure of the DART station files and tsunami source functions data are given in Appendix A.
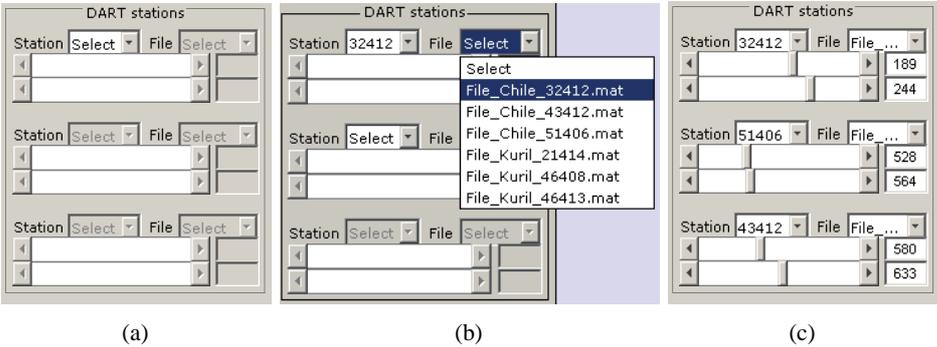
Figure 4.3: Different stages of the station panel; (a) initial view, (b) selecting DART station and DART buoy measurement file and (c) after the selection.

4.1.3   Source selection panel

Source panel is composed of checkboxes which are linked to the selected tsunami source functions. Initially, source panel is empty (Figure 4.4(a)). However, source panel is filled with source function names (Figure 4.4(b)) by selecting DART(s). Checkboxes next to tsunami

source function names are used to choose whether to include or not to include the listed tsunami source functions in the inversion process. Checkboxes under the 'Plot/Not' header is used to switch between whether time series of source function(s) is/are visible on the figures panel (see Section 4.1.6) during selection or not (Figure 4.5). After an inversion, slip values (weights for tsunami source functions) are listed under the 'Weights' header (Figure 4.4(c)).
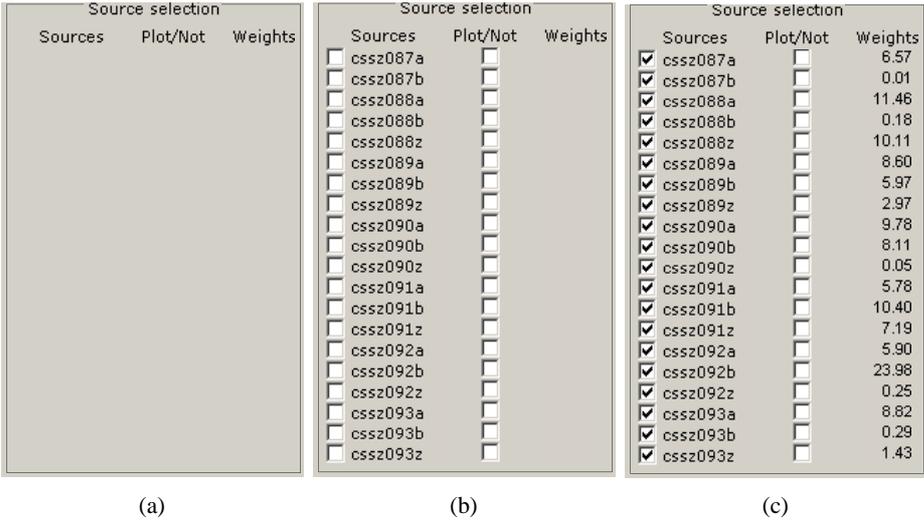
Figure 4.4: Different stages of the source panel; (a) initial view, (b) after the source selection and (c) after solution. Here, tsunami source functions with check marks are used for the inversion. The weights are listed after the inversion.
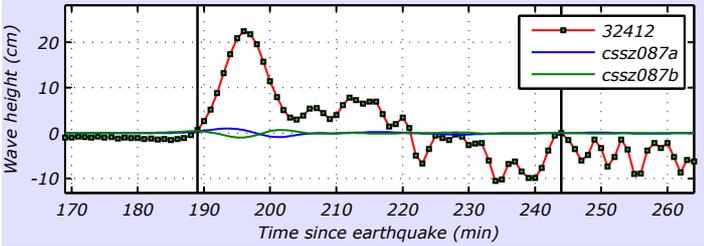
Figure 4.5: The DART 32412 buoy measurement for the 27 February 2010 Chile tsunami together with two of the tsunami source functions. 'PlotNot' header in the source selection panel could be used to enable/disable the visibility of the tsunami source function time series. The vertical solid lines define the interval for which inversion will be performed.

4.1.4 Genetic algorithm parameters panel

Genetic algorithm parameters panel allows defining the genetic algorithm parameters that will be used in an inversion (Figures 4.6(a) and 4.6(b)). These parameters are defined in detail in Chapter 3. However, in short, 'Pop. size' defines the number of sample solutions that will be compared simultaneously. 'Generation' defines the number of iteration at which program will terminate. 'Plot iter.' is used to consider whether solution is plotted during iteration or not. 'Fitness fcn.' is used to select the residual minimization norm (Equation 2.4 or 2.3 or 2.5). 'Constraint' is used to enable/disable constraint option. When it is switched to 'on' the 'Moment mag.' edit box is activated and it allows specifying for a moment magnitude of the resultant earthquake. It takes the moment magnitude and uses Equation 2.1 to calculate total slip amount in order to use as a constraint. Other inputs related to genetic algorithm parameters are explained in Chapter 3 and their default values are shown in Figures 4.6(a) and 4.6(b) for single-objective and multi-objective genetic algorithms respectively.

Figure 4.6: Genetic algorithm parameters panel. Default values for (a) single-objective and (b) multi-objective genetic algorithms.

4.1.5 Actions panel

Solve button callback in the actions panel (Figure 4.7) runs appropriate sub-function, i.e., single-objective or multi-objective inversion, depending on single or multiple DART measurement(s) usage respectively. If single station is selected, 'SingleObj' sub-function is called

otherwise 'MultiObj' sub-function is called and inversion is performed. As an output, weights of tsunami source functions are saved to the workspace as a variable and printed in the source selection panel under the 'Weights' title. At the end, GAIN generates plot of inversion result(s) vs. DART buoy measurement(s) (Figures 4.8). 'Report' button callback generates a text file in the 'REPORT' subfolder such as the one given in Appendix A including tabulated tsunami source function names used in the inversion with their weights, total slip amount and equivalent of moment magnitude. Weights are also printed as a combination into source combination box (Figure 4.9) which enables to use source combination in other programs, e.g., ComMIT^{TM} (Titov et al., 2011). 'Save' button saves current session into a file entered to the editbox in order to use it later. Load popupmenu lists saved sessions in 'SAVED' subfolder and a session can be loaded by using 'Load' button. 'New' button resets variables in the workspace and the panels for a new session.
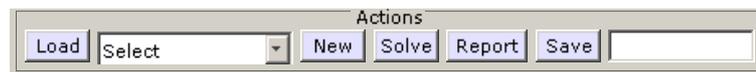
Figure 4.7: Actions panel.
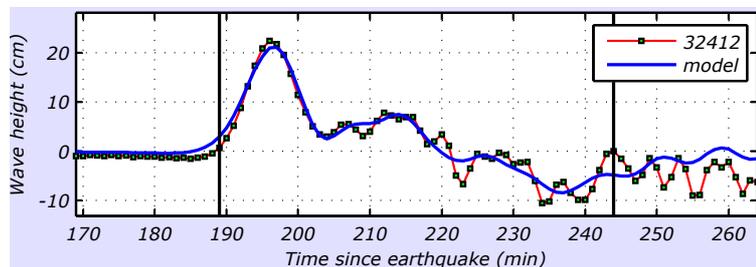
Figure 4.8: Plot of inversion results together with the DART buoy measurements for the 27 February 2010 Chile tsunami.

4.1.6 Figures panel

Figures panel is used to display the plots of solution steps. Initially it is empty. Once a DART buoy measurement file is selected only DART buoy record is shown in the figures panel. If requested, it might display tsunami source functions (Figure 4.5) and once inversion is obtained it displays inversion results together with DART buoy measurement(s) (Figure

4.8). DART stations and source selection panels are linked to the figures panel and changes affect screen display immediately.

4.1.7 Results panel

Results panel displays some additional results of an inversion, i.e., tsunami source combination, sum of tsunami source function weights, equivalent moment magnitude of this sum, generation count at which iteration terminated, RMS error(s) and termination message, (Figure 4.9).

	Results			
Source combination:	3.23*cssza87+0.01*csszb87+10.85*cssza88+1.35*css			
Total slip (m)	Moment magnitude	Generation	RMS error	Termination message
146.32	8.94	213	0.969	Optimization terminated: average change in the fitness value less than options.TolFun.

Figure 4.9: Results panel displaying inverted tsunami source function combination, sum of the tsunami source weights (total slip amount), equivalent moment magnitude of this sum, generation at which iteration terminated, RMS error in the selected interval and termination message.

CHAPTER 5

CASE STUDIES

The NCTR's tsunami forecasting methodology, genetic algorithm methodology and the GUI Genetic Algorithm for INversion (GAIN) which is developed in this study are explained in the previous chapters. Now, the 15 November 2006 Kuril and the 27 February 2010 Chile tsunamis are chosen as case studies to investigate several questions, e.g., number of tsunami source functions needs to be used, which norm could be used, number of DARTs need to be used, whether or not to constrain the inversion with earthquake magnitude, the time interval of DART measurement which inversion is performed.

Before starting to perform an inversion, one should decide how many tsunami source functions should be initially considered. Synolakis et al. (1997) provides reasonable estimates of tsunamigenic earthquake parameters (Table 5.1). Therefore, approximate fault dimensions given by Synolakis et al. (1997) will be used to estimate the initial choice of number of tsunami source functions.

Table 5.1: Estimates of fault dimensions for tsunamigenic earthquakes (Synolakis et al., 1997).

Moment magnitude M_w	Rupture width W (km)	Rupture length L (km)	Slip amount D (m)
7	30	70	0.6
8	80	200	2.7
9	240	600	9.0
9.5	250	1000	27.0
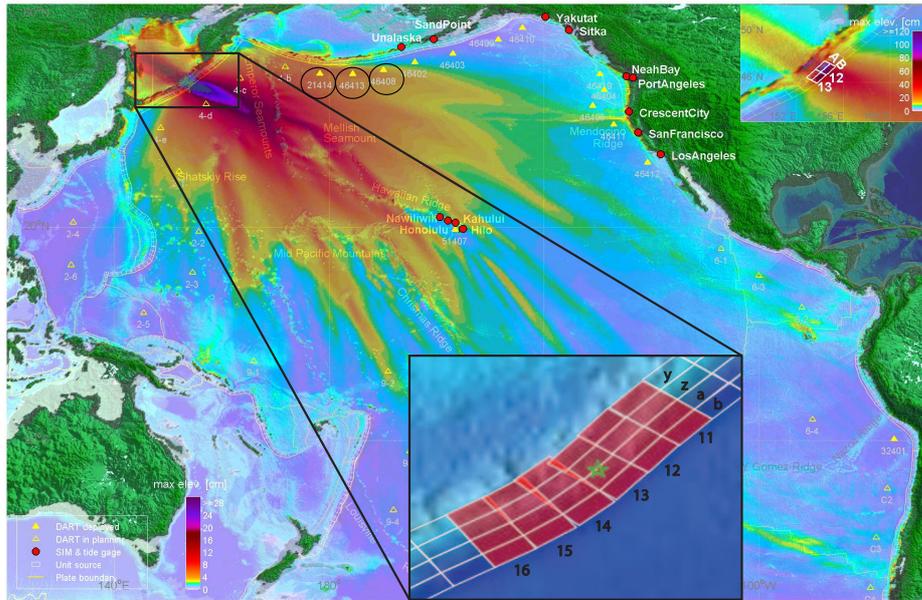
Figure 5.1: Real time energy plot of the 15 November 2006 Kuril tsunami with a source configuration 4.00·12a + 2.00·13a + 0.50·12b + 1.50·13b showing the DARTs in the region (NCTR, 2010). Green star shows the earthquake epicenter and the DARTs 21414, 46413 and 46408 are circled.

5.1 The 15 November 2006 Kuril Islands tsunami

The Kuril Islands earthquake of the 15 November 2006 (11:14:16 UTC, Longitude=46.683° N, Latitude=153.226° E) with magnitude 8.3 occurred as thrust-faulting on the boundary between the Pacific plate and the Okhotsk plate (U.S. Geological Survey, Kuril, 2006). The 15 November 2006 Kuril Islands earthquake generated Pacific basin tsunami. Energy distribution -maximum tsunami height at each computational grid point- is given in Figure 5.1 (NCTR, 2010). One person injured at Waikiki, Hawaii by a tsunami with a measured wave height of 34 cm at Honolulu, Hawaii. One parking lot was flooded at Nawiliwili, Hawaii by a tsunami with a measured wave height of 88 cm. Two docks destroyed and at least one damaged at Crescent City, California by a tsunami with a measured wave height of 176 cm (Dengler et al., 2009).

The NCTR's Kamchatka-Kuril-Japan-Izu-Mariana-Yap subduction zone (kisz) tsunami source functions are provided in Figure 5.2 (Tang et al., 2010). Now, tsunami source inversion will be performed using GAIN to test different possibilities for norms, constraint, etc.
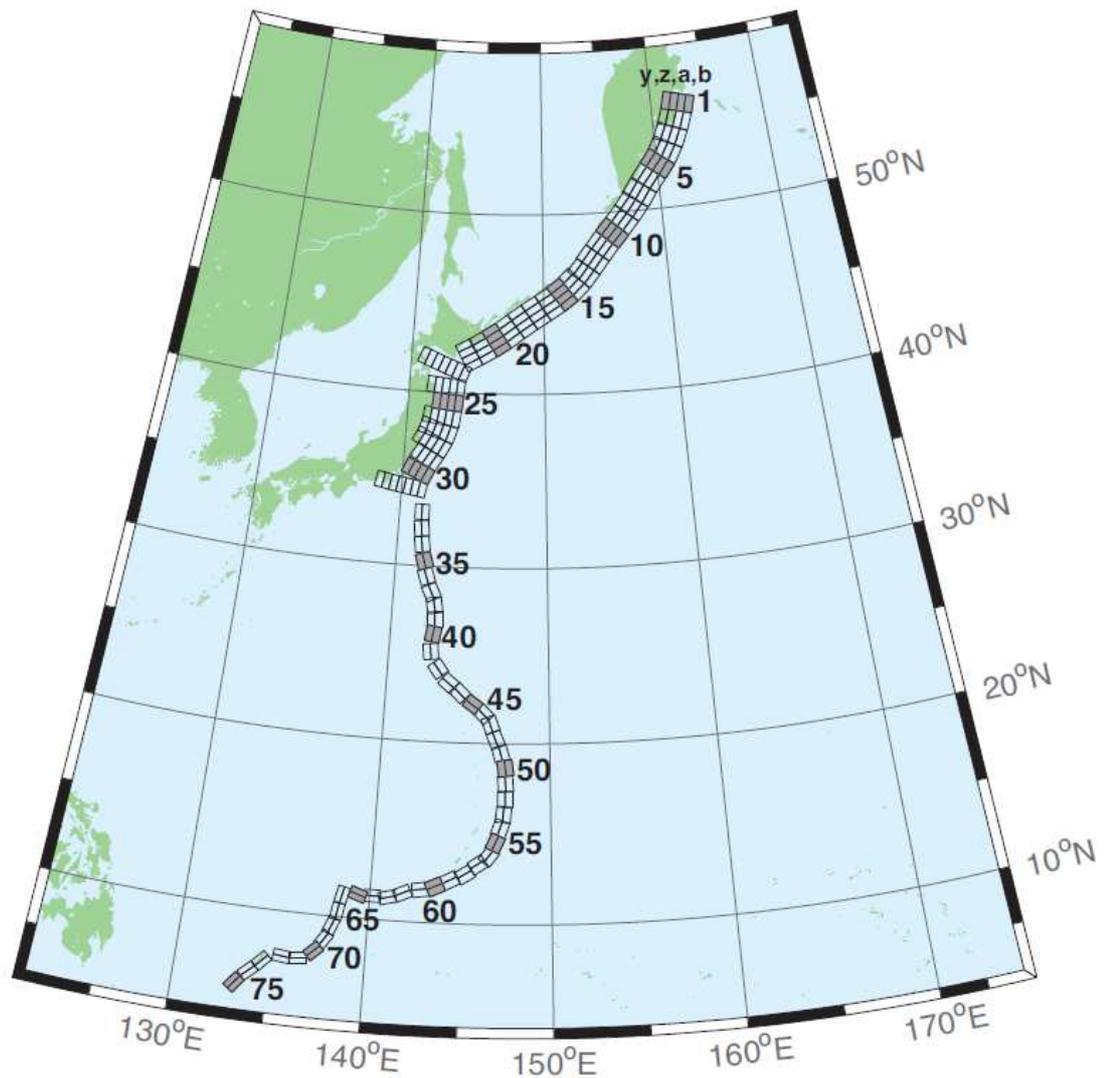
Figure 5.2: Kamchatka-Kuril-Japan-Izu-Mariana-Yap subduction zone (kisz) tsunami source functions (Tang et al., 2010).

5.1.1 Single DART unconstrained inversion for the 15 November 2006 Kuril Islands tsunami

First, single DART 21414 buoy measurement and the tsunami source functions between 12-15 on the Kamchatka-Kuril-Japan-Izu-Mariana-Yap subduction zone (Figure 5.2) are used to invert the tsunami source. Optimization is left unconstrained and inversion time interval for the DART time series is chosen from 112 to 147 minutes after the earthquake. This time interval covers first elevation and depression part of the incoming wave. Population size is

defined as 100 and generation is limited to 1000. L_1, L_2 and L_∞ error minimization norms are used as fitness functions and results are presented in Figures 5.3(a) - 5.3(c) and weights for tsunami source functions and their root mean square (RMS)[1] errors are listed in Table 5.2.

Single DART 21414 unconstrained inversion solution results are used to evaluate estimates at the DARTs 46413 and 46408 and compared with the DART measurements for each norm (Figures 5.4(a) - 5.4(c) and Table 5.2). This is important to evaluate whether single DART inversion is satisfactory or multi DART measurements are needed. One noticeable result in Figures 5.4(a) - 5.4(c) is the approximately five-minutes time discrepancies of the arrival of the first wave at the DART 46413. This time discrepancies can be attributed to multiple causes, i.e., the errors induced by the ocean bathymetry, model approximation. Similar time difference with forecasting results and DART buoy measurements are observed during different events (Wei et al., 2008).
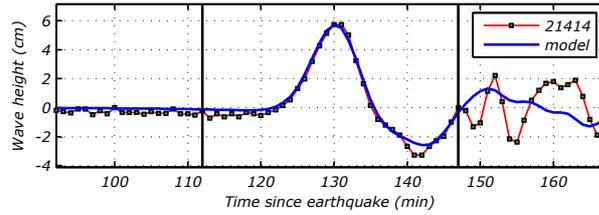
(a)
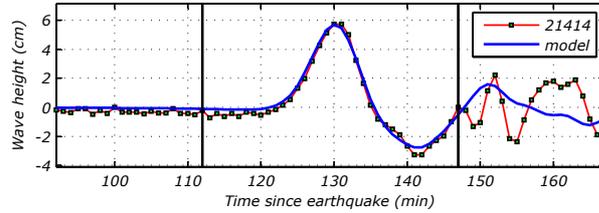
(b)

Figure 5.3: Single DART unconstrained inversion for the Kuril Islands tsunami for (a) L_1, (b) L_2 and (c) L_∞ norms. RMS errors are 0.328, 0.326 and 0.406 for L_1, L_2 and L_∞ norms respectively. Solid vertical lines show the time interval used for the inversion.

---

[1] RMS error is defined as $\sqrt{\left\{\sum_{j=1}^{m}\left(\sum_{i=1}^{n}[b(t_j) - w_i g_i(t_j)]\right)^2\right\}/m}$, where $b(t)$ is the DART measurement, $g_i(t)$s are the pre-computed source functions at the DART location with the w_i scalings (weights), n is the number of tsunami source functions and m is the number of data points.
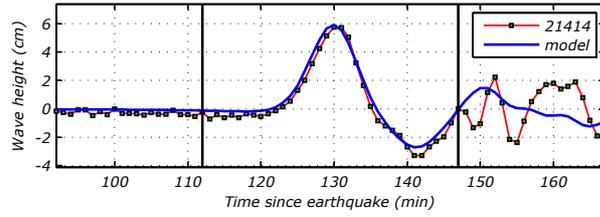
(c)

Figure 5.3: Continued
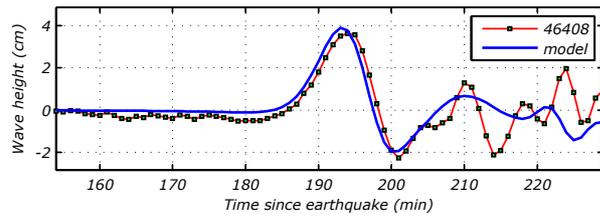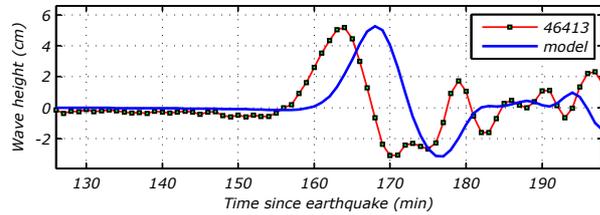
(a)
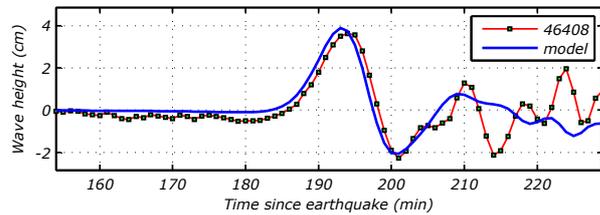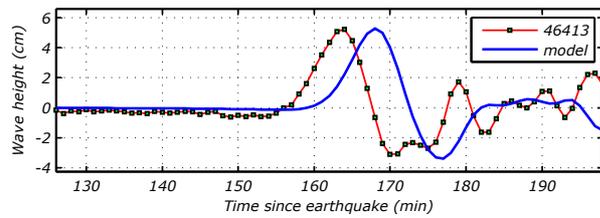
(b)

Figure 5.4: Weights calculated from single DART 21414 unconstrained inversions are used to evaluate the Kuril Islands tsunami offshore amplitudes at the DARTs 46413 and 46408 by using (a) L_1 b) L_2 and (c) L_∞ norms. RMS errors are 2.922 and 0.459 for L_1, 2.919 and 0.486 for L_2, 0.915 and 0.567 for L_∞ for the DARTs 46413 and 46408 respectively.
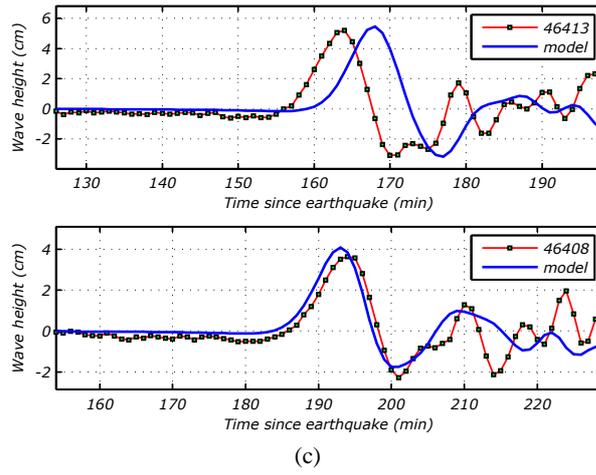
Figure 5.4: Continued.

(a)

(b)

(c)
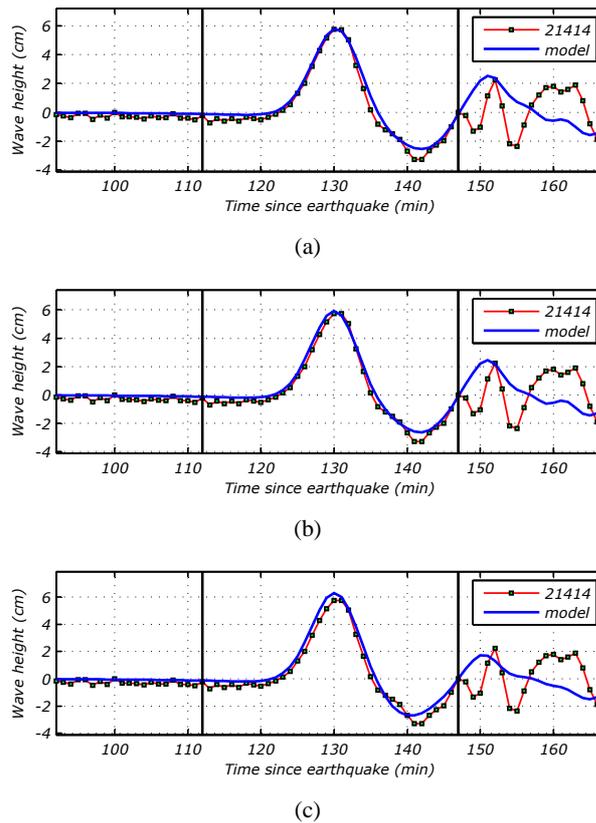
Figure 5.5: Single DART 21414 inversion for the Kuril Islands tsunami constrained with M_w 8.3 earthquake magnitude for (a) L_1, (b) L_2 and (c) L_∞ norms. RMS errors are 0.399, 0357 and 0.481 for L_1, L_2 and L_∞ norms respectively. Solid vertical lines show the time interval used for the inversion.

5.1.2 Single DART constrained inversion for the 15 November 2006 Kuril Islands tsunami

In this case, again only the DART 21414 is used to invert the tsunami source functions 12-15 on the Kamchatka-Kuril-Japan-Izu-Mariana-Yap subduction zone. However, optimization is constrained with M_w 8.3 earthquake magnitude and time the interval for the inversion is chosen from 112 to 147 minutes after the earthquake. Population size is defined as 100 and generation is limited to 1000. L_1, L_2 and L_∞ error minimization norms are used as a fitness functions and results are presented in Figure 5.5(a) - 5.5(c) and weights and RMS errors are listed in Table 5.2.
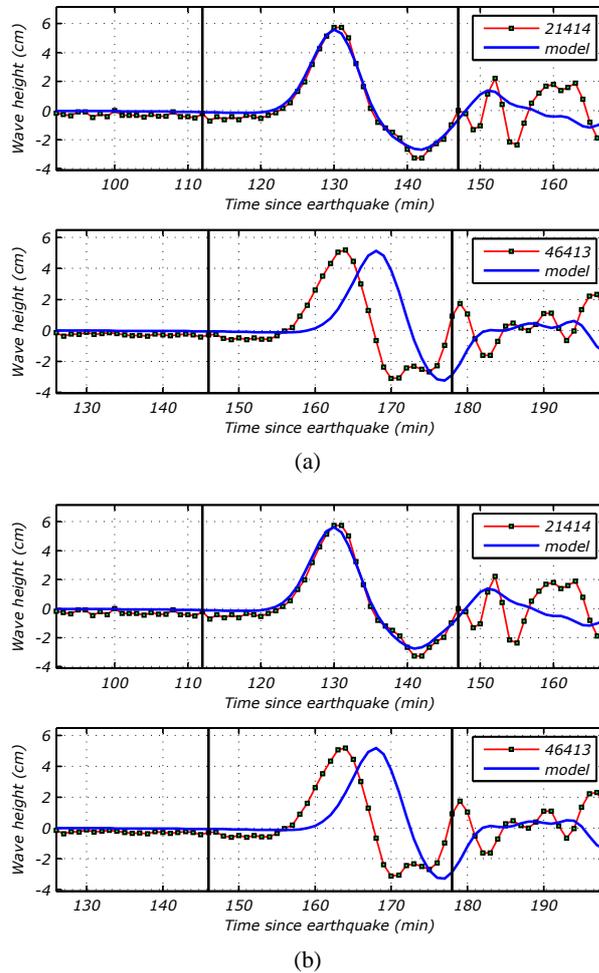
Figure 5.6: Two DARTs 21414 and 46413 unconstrained inversion for the Kuril Islands tsunami for (a) L_1, (b) L_2 and (c) L_∞ norms. RMS errors are 0.336 and 2.874 for L_1, 0.363 and 2.841 for L_2, 0.506 and 2.831 for L_∞ for the DARTs 21414 and 46413 respectively. Solid vertical lines show the time interval used for the inversion.
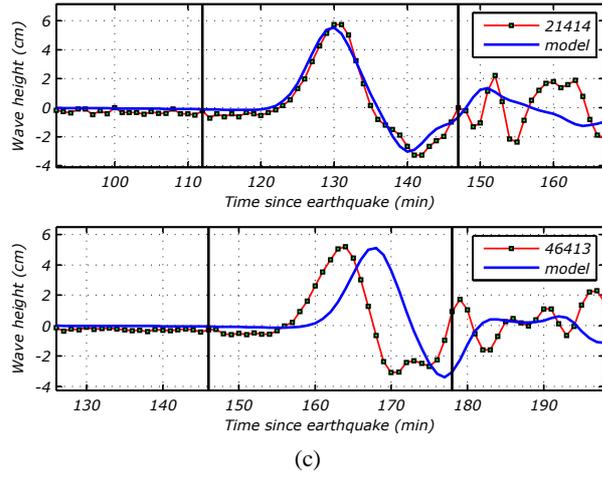
Figure 5.6: Continued.

5.1.3 Two DARTs unconstrained inversion for the 15 November 2006 Kuril Islands tsunami

Now, both DARTs 21414 and 46413 are used in inversion with the same source functions. Inversion is left unconstrained and the time interval for the inversion is defined from 112 to 147 minutes as in the previous case for the DART 21414 and from 146 to 178 minutes after the earthquake for the DART 46413. Population size is defined as 300 since this is a multi-objective inversion (see Section 4.1.4) and generation is limited to 1000. L_1, L_2 and L_∞ error minimization norms are used as fitness functions and results are presented in Figures 5.6(a) - 5.6(c) and weights and RMS errors are listed in Table 5.2.

Again, as in the one DART case, two DARTs 21414 and 46413 unconstrained inversion results for L_1, L_2 and L_∞ norms are used to evaluate tsunami form at the DART 46408 and compared for each norm (Figures 5.7(a) - 5.7(c) and Table 5.2).

44
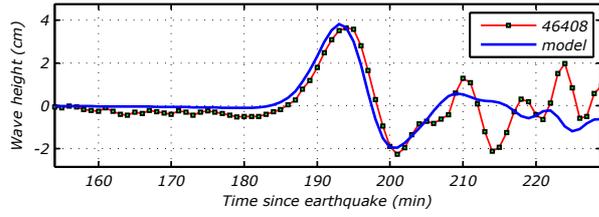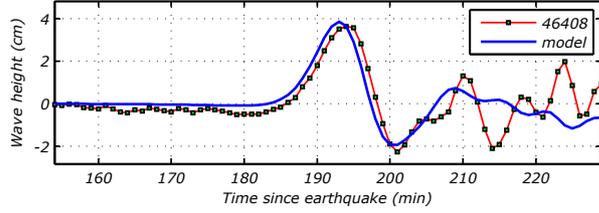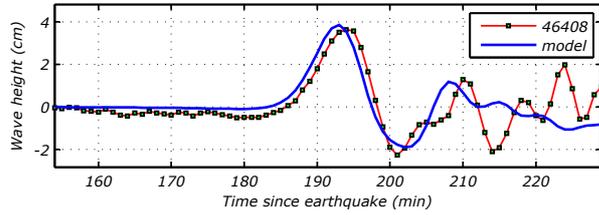
Figure 5.7: Weights for the tsunami source functions calculated from the DARTs 21414 and 46413 unconstrained inversions are used to evaluate the Kuril Islands tsunami offshore amplitudes at the DART 46408 by using (a) L_1, (b) L_2 and (c) L_∞ norms with RMS errors 0.567, 0.535 and 0.628 respectively.

5.1.4 Three DARTs unconstrained inversion for the 15 November 2006 Kuril Islands tsunami

Finally, three DARTs 21414, 46413 and 46408 are used in the inversion with the tsunami source functions 12-15 on the Kamchatka-Kuril-Japan-Izu-Mariana-Yap subduction zone. Inversion is left unconstrained and the time interval for the inversion is defined from 112 to 147 minutes and from 146 to 178 minutes after the earthquake for the DARTs 21414 and 46413 respectively as in the previous case and from 174 to 209 minutes after the earthquake for the DART 46408. Population size is defined as 300 and generation is limited to 1000. L_1, L_2 and L_∞ error minimization norms are used as fitness functions and results are presented in Figures 5.8(a) - 5.8(c) and weights and RMS errors are listed in Table 5.2.

45
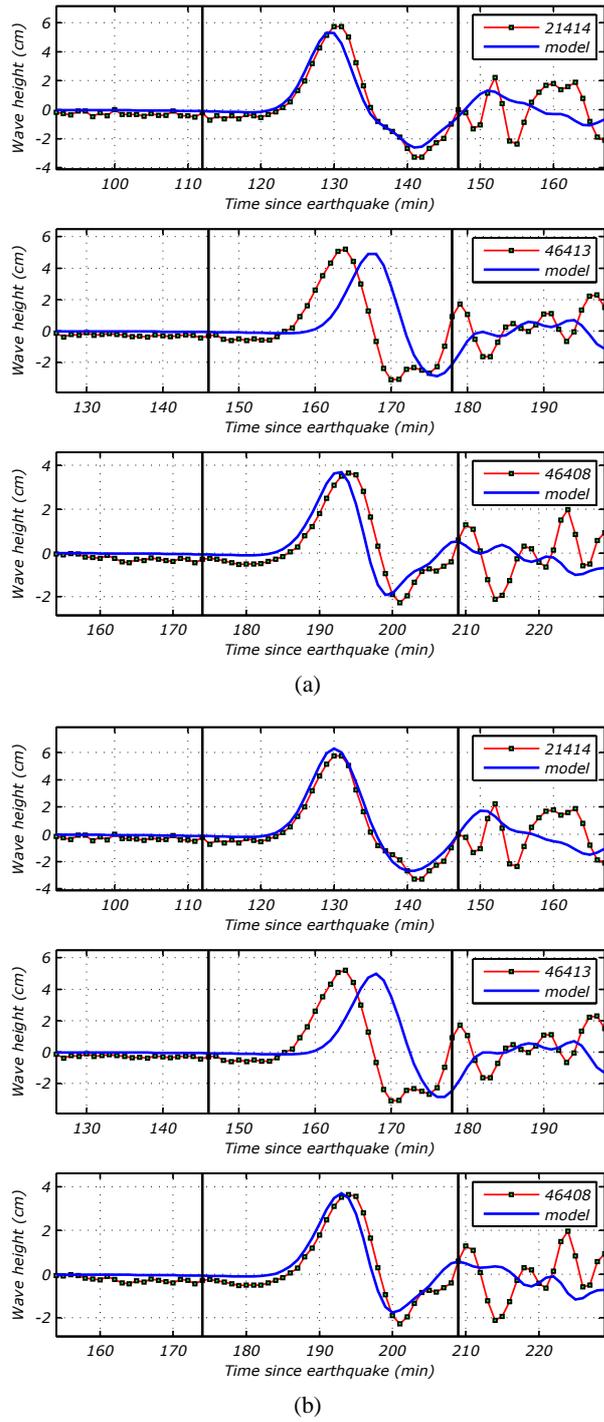
Figure 5.8: Three DARTs 21414, 46413 and 46408 unconstrained inversion for the Kuril Islands tsunami for (a) L_1, (b) L_2 and (c) L_∞ norms. RMS errors are 0.582, 2.554 and 0.765 for L_1, 0.410, 2.761 and 0.535 for L_2, 0.433, 2.787 and 0.603 for L_∞ for the DARTs 21414, 46413 and 46408 respectively. Solid vertical lines show the time interval used for the inversion.
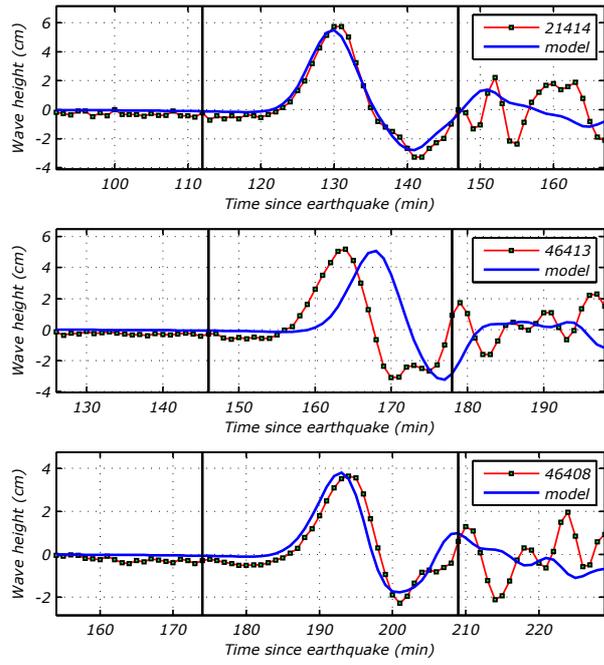
Figure 5.8: Continued.

Table 5.2: Tsunami source function weights based on different considerations in Subsections 5.1.1 - 5.1.4. Highlighted RMS errors are for the DARTs not used for the inversion, i.e., inversion results are used to evaluate tsunami height estimates at the other DARTs.

	Conditions	Unconstrained									Constrained		
	DART(s)	21414			21414 46413			21414 46413 46408			21414		
	Norms	L_1	L_2	L_∞	L_1	L_2	L_∞	L_1	L_2	L_∞	L_1	L_2	L_∞
Tsunami source functions	kisz012a	5.70	5.06	4.92	4.98	4.76	4.05	5.03	5.22	4.59	5.30	5.74	5.53
	kisz012b	0.07	0.20	0.69	0.23	0.48	1.11	1.10	0.44	0.87	0.10	0.40	0.46
	kisz012y	0.03	0.01	0.00	0.02	0.06	0.46	0.03	0.14	0.07	0.01	0.07	0.25
	kisz012z	0.15	1.54	1.82	1.38	1.76	2.09	0.08	0.10	1.39	0.97	0.14	1.24
	kisz013a	3.48	2.80	2.82	2.82	2.43	1.77	2.72	3.30	2.50	3.64	3.88	2.91
	kisz013b	0.01	0.01	0.38	0.02	0.17	1.21	0.04	0.02	0.62	0.16	0.22	0.72
	kisz013y	0.00	0.00	0.04	0.01	0.06	0.05	0.11	0.25	0.00	0.01	0.01	0.11
	kisz013z	0.41	0.16	0.05	0.08	0.11	0.29	0.05	0.06	0.01	0.06	0.03	0.44
	kisz014a	0.04	0.01	0.03	0.01	0.01	0.11	0.04	0.12	0.10	0.05	0.02	0.06
	kisz014b	1.63	0.99	1.72	1.21	0.90	0.51	1.17	1.33	1.12	1.93	2.11	1.34
	kisz014y	0.00	0.00	0.00	0.02	0.10	0.01	0.02	0.12	0.05	0.04	0.04	0.01
	kisz014z	0.00	0.00	0.01	0.00	0.01	0.03	0.04	0.03	0.01	0.01	0.03	0.05
	kisz015a	0.07	0.12	1.10	0.28	0.10	0.01	0.15	0.43	0.24	0.17	0.15	0.56
	kisz015b	1.15	0.13	0.30	0.18	0.11	0.05	0.10	0.03	0.47	0.27	0.22	0.68
	kisz015y	0.36	0.12	0.52	0.07	0.12	0.36	0.24	0.21	0.47	2.39	0.98	0.91
	kisz015z	0.28	0.48	0.30	0.04	0.02	0.09	0.13	0.31	0.24	0.70	1.75	0.54
	Total slip (m)	13.39	11.63	14.70	11.36	11.19	12.19	11.06	12.10	12.74	15.81	15.81	15.81
	M_w	8.25	8.21	8.28	8.20	8.20	8.22	8.20	8.22	8.24	8.30	8.30	8.30
RMS error	for 21414	0.328	0.326	0.406	0.336	0.363	0.506	0.582	0.410	0.433	0.399	0.357	0.481
	for 46413	2.922	2.919	0.915	2.874	2.841	2.831	2.554	2.761	2.787	2.833	2.954	3.080
	for 46408	0.459	0.486	0.567	0.567	0.535	0.628	0.765	0.535	0.603	0.628	0.508	0.619

5.1.5 Discussion of the 15 November 2006 Kuril Islands tsunami source inversion results

Time interval used in an inversion might play important role over an inversion. Time series at the DART 21414 location for the tsunami source functions close to the Kuril Islands earthquake epicenter are given in Figure 5.9 including the DART 21414 buoy measurements. In order to understand the effect of time interval, two different time intervals with tsunami source configurations are chosen to perform the inversion. Inversion results are listed in Table

5.3 and presented in Figures 5.10 - 5.11. It is clear from the results that if the initial wave of the tsunami source functions are out of the range of the inversion time interval this might produce unreasonable weights, e.g., cases 1 and 3 for the inversion time interval from 112 to 135 minutes and case 1 for the inversion time interval from 112 to 147 minutes.
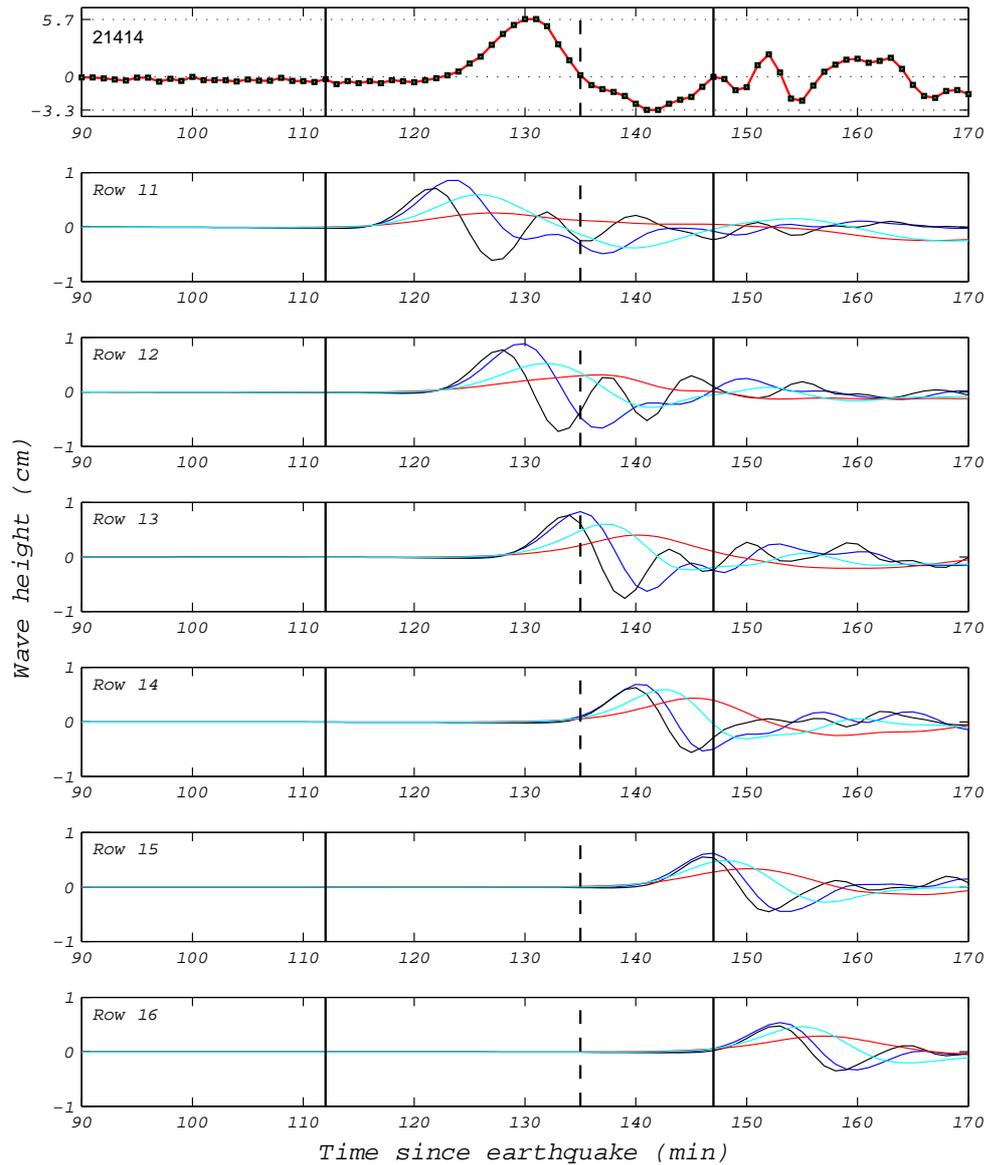
Figure 5.9: The DART 21414 buoy measurements (top inset) and time series of kisz fault tsunami source functions at the DART 21414 location. Black, blue, cyan and red show time series for tsunami source functions b, a, z and y in seaward direction, respectively. Vertical solid lines show inversion time interval.

Table 5.3: Weights of various source function combinations of the single DART 21414 unconstrained solution of the Kuril tsunami using the inversion time interval from 112 to 147 minutes and from 112 to 135 minutes after the earthquake. Population size is 100.

		Interval from 112 to 135 minutes				Interval from 112 to 147 minutes				
		Case 1	Case 2	Case 3	Case 4	Case 1	Case 2	Case 3	Case 4	Case 5
Tsunami source functions	kisz011a	0.00	0.00				0.00			
	kisz011b	0.00	0.00				0.01			
	kisz011y	0.01	0.00				0.00			
	kisz011z	0.00	0.00				0.01			
	kisz012a	5.67	4.98	5.96	5.57	4.97	5.21	4.98	5.66	5.66
	kisz012b	0.02	0.00	0.01	0.00	0.23	0.26	0.21	0.06	0.13
	kisz012y	0.02	0.00	0.00	0.00	0.10	0.01	0.00	0.00	0.00
	kisz012z	0.39	1.55	0.01	0.00	1.84	1.27	1.92	0.14	0.05
	kisz013a	0.21	0.35	0.03	0.00	2.72	3.01	2.66	2.94	3.57
	kisz013b	2.19	1.77	2.69	2.25	0.02	0.03	0.06	0.00	0.00
	kisz013y	0.32	1.65	0.01	2.86	0.01	0.00	0.03	0.00	0.00
	kisz013z	0.43	0.31	0.01	1.56	0.17	0.22	0.05	1.43	0.34
	kisz014a	4.39		1.41		0.03	0.01	0.02		0.00
	kisz014b	2.84		6.36		1.27	1.20	1.09		0.92
	kisz014y	3.66		1.60		0.03	0.02	0.00		0.61
	kisz014z	2.23		9.96		0.02	0.00	0.01		0.00
	kisz015a					0.01	0.04	0.00		
	kisz015b					0.08	0.48	0.19		
	kisz015y					0.49	0.16	0.04		
	kisz015z					0.03	0.49	0.76		
	kisz016a					7.94				
	kisz016b					7.82				
	kisz016y					0.23				
	kisz016z					7.43				
	Total slip (m)	22.38	10.61	28.05	12.14	35.44	12.43	12.02	10.23	11.28
	M_w	8.40	8.18	8.47	8.23	8.53	8.23	8.22	8.17	8.20
	RMS error	0.259	0.292	0.239	0.275	0.267	0.326	0.337	0.323	0.327
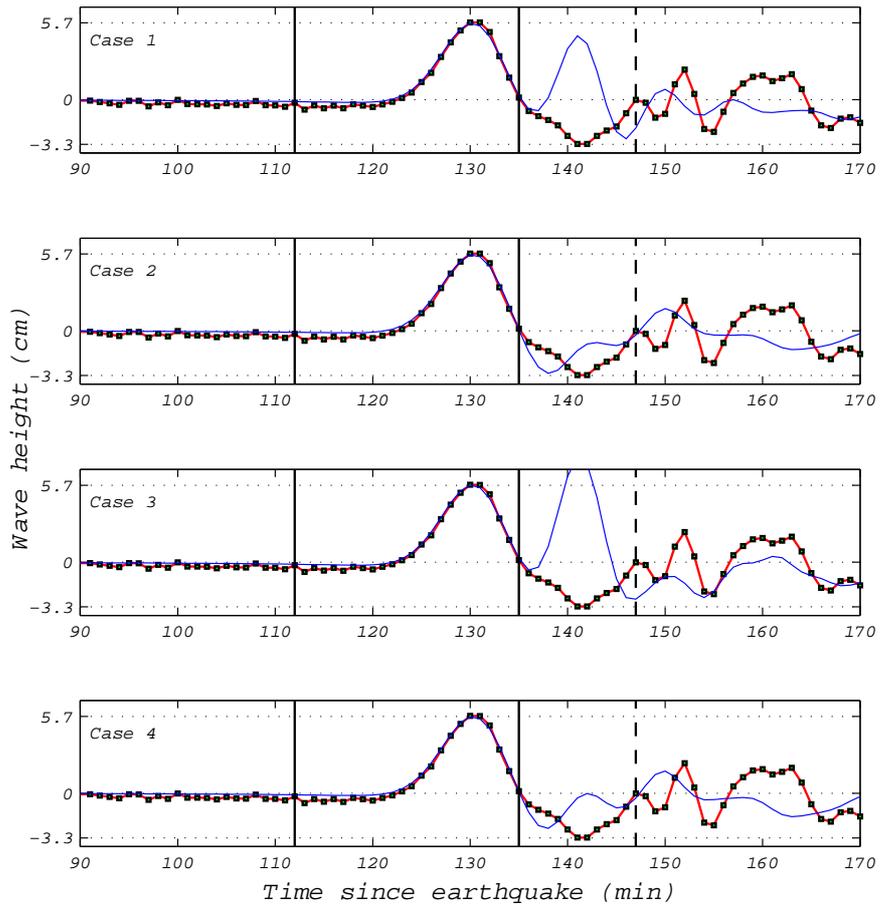
Figure 5.10: Comparisons of the DART 21414 buoy measurement and inversion results for the cases listed in Table 5.3 for the inversion time interval from 112 to 135 minutes. Solid line represents inversion result, dotted solid line represents the DART 21414 buoy measurements and solid vertical lines show the time interval used for the inversion.

Tsunami source function weights are recalculated for larger population sizes. In Table 5.4, population sizes are selected from 100 to 1000 for single DART and inversion is performed only for L_2 norm. When the results in Tables 5.2 and 5.4 are compared they are close to each other, i.e., there are slight differences resulted from randomness of genetic algorithm. However, increasing population size results in more stable weights. Solution with larger population takes more time thus obtaining solution with an optimum population might save time in real-time forecasting.

51
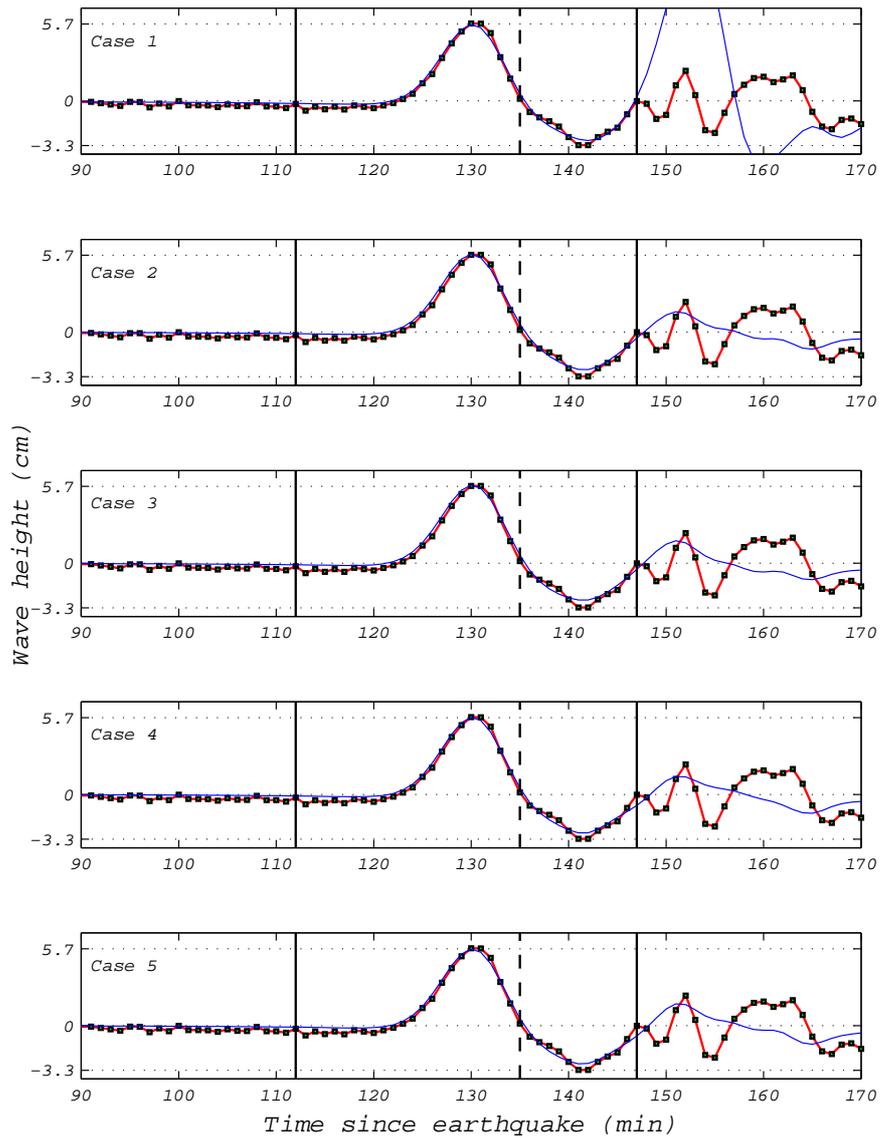
Figure 5.11: Comparisons of the DART 21414 buoy measurement and inversion results for the cases listed in Table 5.3 for the inversion time interval from 112 to 147 minutes. Solid line represents inversion result, dotted solid line represents the DART 21414 buoy measurements and solid vertical lines show the time interval used for the inversion.

Tsunami source function weights for the population size of 1000 given in Table 5.4 are averaged and used to evaluate forecast models at several coastal locations (Tang, personal communication)[2]. Forecast model results are compared with tide gage measurements. Comparisons show reasonable agreements on both tsunami arrival times and wave amplitudes (Figure 5.12).

---

[2] Model runs were performed by Liujuan Tang of NCTR using their forecasting models.

Table 5.4: Tsunami source function weights of the 15 November 2006 Kuril Islands tsunami based on L_2 norm for the population size changing from 100 to 1000. Significant ones are highligted.

		Population size											
		100	200	300	400	500	600	700	800	900	1000	1000	1000
Tsunami source functions	kisz012a	5.06	5.52	5.76	5.38	5.71	5.76	5.75	5.77	5.74	5.43	5.77	5.71
	kisz012b	0.20	0.16	0.12	0.16	0.15	0.10	0.09	0.09	0.12	0.17	0.12	0.13
	kisz012y	0.01	0.02	0.00	0.00	0.03	0.01	0.01	0.01	0.00	0.00	0.00	0.00
	kisz012z	1.54	0.56	0.01	0.84	0.02	0.02	0.04	0.00	0.03	0.72	0.01	0.04
	kisz013a	2.80	3.44	3.87	3.34	3.77	3.78	3.69	3.82	3.82	3.40	3.86	3.81
	kisz013b	0.01	0.00	0.01	0.02	0.03	0.00	0.01	0.01	0.00	0.01	0.00	0.01
	kisz013y	0.00	0.01	0.01	0.01	0.02	0.01	0.02	0.01	0.01	0.01	0.00	0.01
	kisz013z	0.16	0.11	0.01	0.01	0.07	0.07	0.28	0.05	0.05	0.03	0.01	0.01
	kisz014a	0.01	0.00	0.01	0.01	0.01	0.00	0.02	0.02	0.01	0.01	0.01	0.01
	kisz014b	0.99	1.39	1.62	1.40	1.47	1.56	1.32	1.58	1.57	1.42	1.66	1.64
	kisz014y	0.00	0.01	0.00	0.02	0.03	0.01	0.01	0.01	0.01	0.02	0.00	0.00
	kisz014z	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.02	0.00	0.00	0.02	0.01
	kisz015a	0.12	0.47	0.22	0.08	0.01	0.09	0.04	0.08	0.04	0.07	0.32	0.14
	kisz015b	0.13	0.54	0.74	0.73	0.86	1.06	1.01	1.19	1.28	0.92	1.05	1.26
	kisz015y	0.12	0.02	0.03	0.24	0.04	0.22	0.05	0.17	0.03	0.06	0.03	0.01
	kisz015z	0.48	0.15	0.63	0.25	0.37	0.23	0.30	0.15	0.13	0.21	0.16	0.07
	Total slip (m)	11.63	12.42	13.05	12.51	12.70	12.93	12.65	12.99	12.85	12.47	13.02	12.87
	M_w	8.21	8.23	8.24	8.23	8.24	8.24	8.24	8.24	8.24	8.23	8.24	8.24
	RMS error	0.326	0.310	0.301	0.314	0.306	0.301	0.304	0.302	0.300	0.310	0.300	0.300
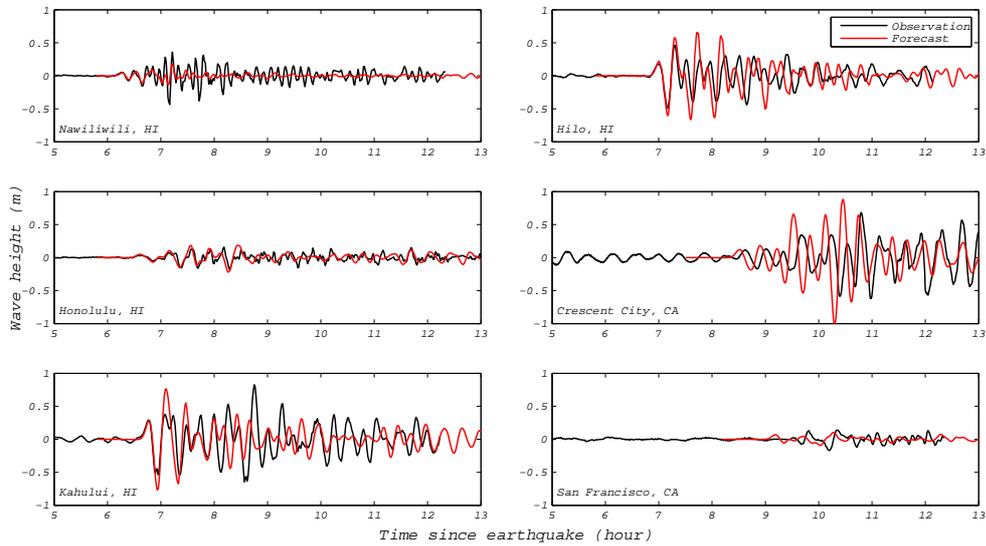
Figure 5.12: Comparison of tide gage measurements with the tsunami inversion results based on the source combination 5.64·12a + 3.69·13a + 1.57·14b + 1.08·15b at several coastal locations. Tsunami source function weights are the average weights for the population size of 1000 given in Table 5.4. Forecast model runs are provided by Liujuan Tang (personal communication).

5.2 The 27 February 2010 Chile tsunami

The Chile earthquake of the 27 February 2010 (06:34:14 UTC; Longitude = 35.826° S, Latitude = 72.669° W) with magnitude M_w 8.8 was generated at the gently sloping fault that conveys the Nazca plate eastward and downward beneath the South American plate. The Chile earthquake was felt in much of Chile and Argentina, also in parts of Bolivia, southern Brazil, Paraguay, Peru and Uruguay (U.S. Geological Survey, Chile, 2010). At least 799 people killed, many injured and around 1.5 million houses damaged by the earthquake and tsunami in the Concepcion-Valparaiso area (Fritz et al., 2011). A Pacific-wide tsunami was generated (Figure 5.13) and caused minor damage to boats and a dock in the San Diego, California area.

NCTR's Central-South America subduction zone (cssz) tsunami source functions are provided in Figure 5.14. Now, tsunami source inversions will be performed using GAIN to exercise different possibilities as in the 15 November 2006 Kuril Islands tsunami case.
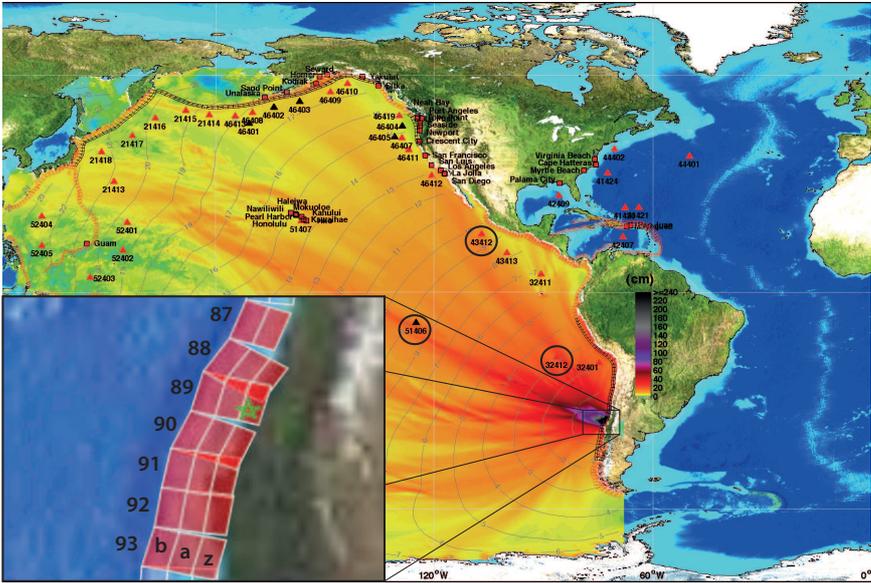
Figure 5.13: Real time energy plot of the 27 February 2010 Chile tsunami with source configuration 17.24·88a + 8.82·90a + 11.86·88b + 18.39·89b + 16.75·90b + 20.78·88z + 7.05·90z (M_w 8.83) showing the DARTs in the region (NCTR, 2010). Green star shows the earthquake epicenter and the DARTs 32412, 51406 and 43412 are circled.
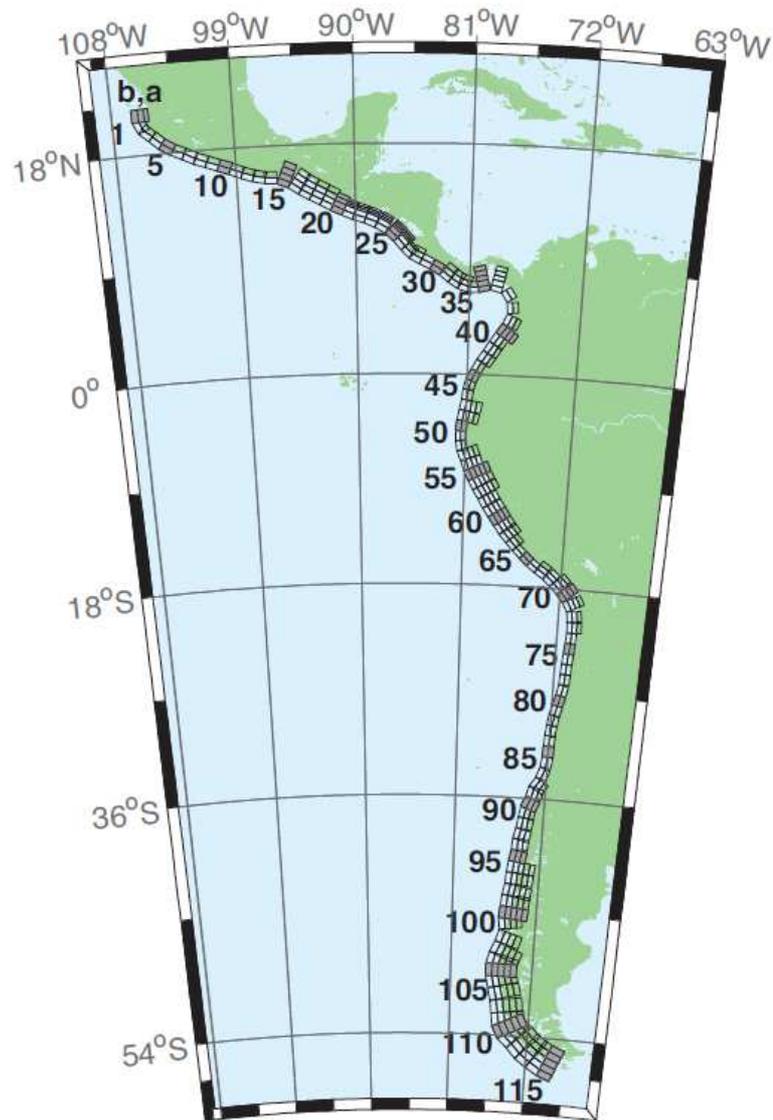
Figure 5.14: Central-South America subduction zone (cssz) tsunami source functions (Tang et al., 2010).

5.2.1 Single DART unconstrained inversion for the 27 February 2010 Chile tsunami

First, single DART 32412 buoy measurement and the tsunami source functions 87-93 on the Central-South America subduction zone (Figure 5.14) are used to invert the tsunami source. Inversion is left unconstrained and the inversion time interval for the DART buoy measurement is chosen from 189 to 244 minutes after the earthquake. Population size is defined as 100 and generation is limited to 1000. L_1, L_2 and L_∞ error minimization norms are used as

fitness functions and results are presented in Figures 5.15(a) - 5.15(c) and weights for tsunami source functions and their RMS errors are listed in Table 5.5.

Single DART 32412 unconstrained inversion results are used to evaluate estimates at the other DARTs 51406 and 43412 and compared with real DART measurements for each norm (Figures 5.16(a) - 5.16(c) and Table 5.5). This is important to evaluate whether single DART solution is enough or multi DART solution is needed. Again, as observed in the 15 November 2006 Kuril Islands tsunami case inversion based on one DART results time shift at the other DARTs, such as the one in DART 51406.

(a)

(b)

(c)
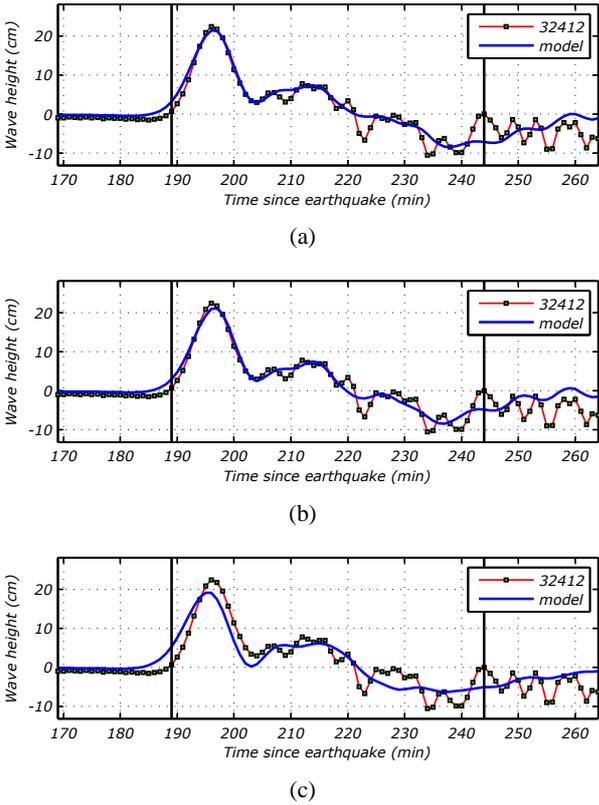
Figure 5.15: Single DART unconstrained inversion for the Chile tsunami for (a) L_1, (b) L_2 and (c) L_∞ norms. RMS errors are 2.286, 1.942 and 3.238 for L_1, L_2 and L_∞ norms respectively. Solid vertical lines show the time interval used for the inversion.
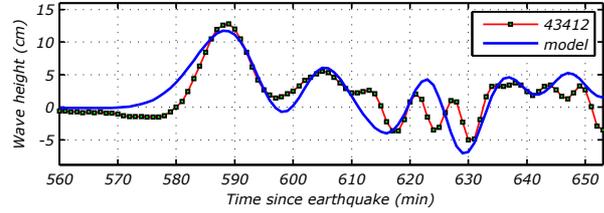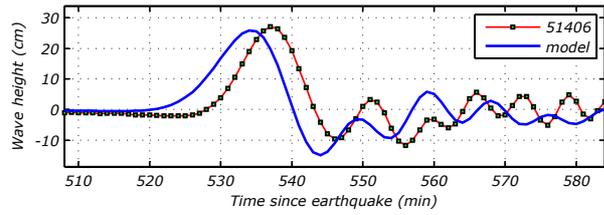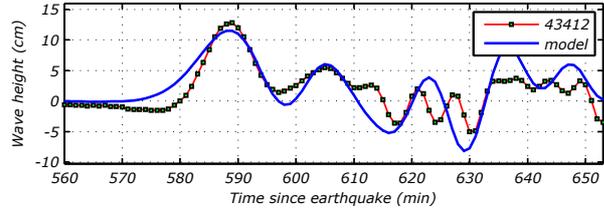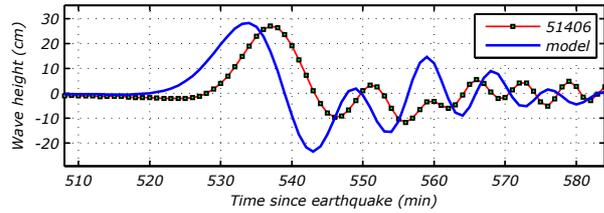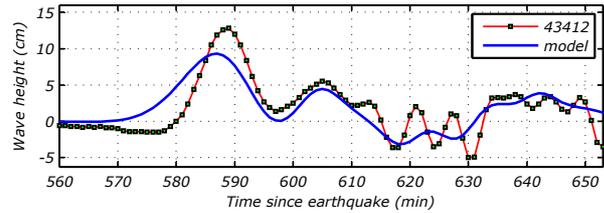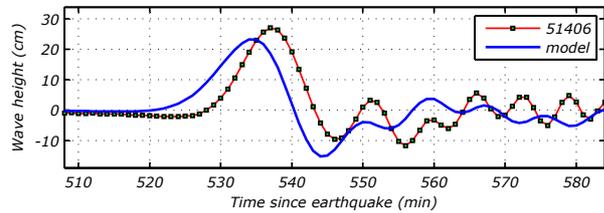
Figure 5.16: Tsunami source function weights calculated from single DART 32412 uncon-
strained inversions are used to evaluate the Chile tsunami offshore amplitudes at the DARTs
51406 and 43412 by using (a) L_1, (b) L_2 and (c) L_∞ norms. RMS errors are 9.697 and 2.671
for L_1, 14.659 and 2.876 for L_2, 8.972, 2.457 for L_∞ for the DARTs 51406 and 43412 respec-
tively.
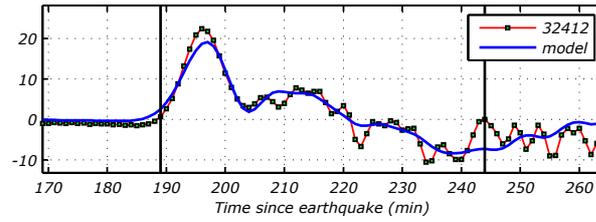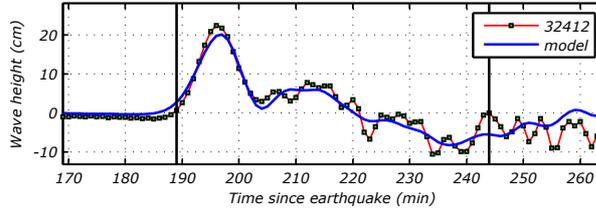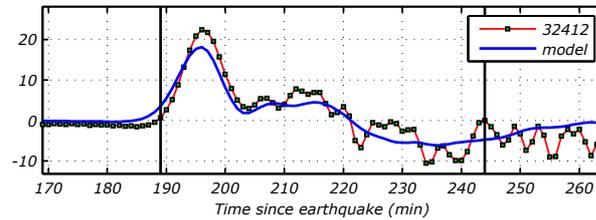
Figure 5.17: Single DART 32412 inversion for the Chile tsunami constrained with M_w 8.8 earthquake magnitude for the Chile tsunami for (a) L_1, (b) L_2 and (c) L_∞ norms. RMS errors are 2.391, 2.152 and 2.971 for L_1, L_2 and L_∞ norms respectively. Solid vertical lines show the time interval used for the inversion.

5.2.2 Single DART constrained inversion for the 27 February 2010 Chile tsunami

In this case, again only the DART 32412 is used to invert the tsunami source functions 87-93 on the Central-South America subduction zone. However, inversion is constrained with an earthquake magnitude M_w 8.8 and the time interval for the inversion is defined from 189 to 244 minutes after the earthquake. Population size is defined as 100 and generation is limited to 1000. L_1, L_2 and L_∞ error minimization norms are used as a fitness functions and results are presented in Figures 5.17(a) - 5.17(c) and weights and RMS errors are listed in Table 5.5.

5.2.3 Two DARTs unconstrained inversion for the 27 February 2010 Chile tsunami

Now, both DARTs 32412 and 51406 are used in inversion with the same tsunami source functions. Inversion is left unconstrained and the time interval for the inversion is defined from 189 to 244 minutes as in the previous case for the DART 32412 and from 528 to 564 minutes after the earthquake for the DART 51406. Population size is defined as 300 since this is a multi-objective inversion (see Section 4.1.4) and generation is limited to 1000. L_1, L_2 and L_∞ error minimization norms are used as fitness functions and results are presented in Figures 5.18(a) - 5.18(c) and weights and RMS errors are listed in Table 5.5.

Again as in the one DART case, two DARTs 32412 and 51406 unconstrained inversion results for L_1, L_2 and L_∞ norms are used to evaluate tsunami form at the DART 43412 and compared with the DART measurements for each norm (Figures 5.19(a) - 5.19(c) and Table 5.5).

(a)
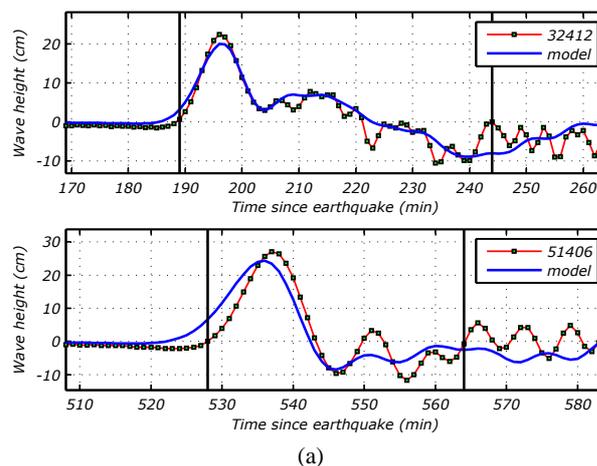
Figure 5.18: Two DARTs 32412 and 51406 unconstrained inversion results for the Chile tsunami for (a) L_1, (b) L_2 and (c) L_∞ norms. RMS errors are 2.617 and 4.674 for L_1, 2.161 and 8.057 for L_2, 3.772 and 3.952 for L_∞ for the DARTs 32412 and 51406 respectively. Solid vertical lines show the time interval used for the inversion.
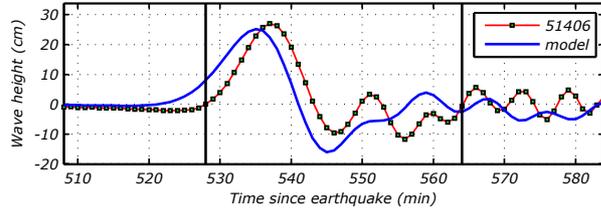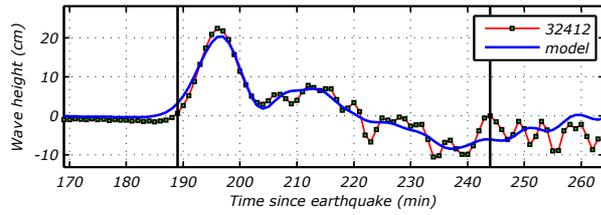
(b)
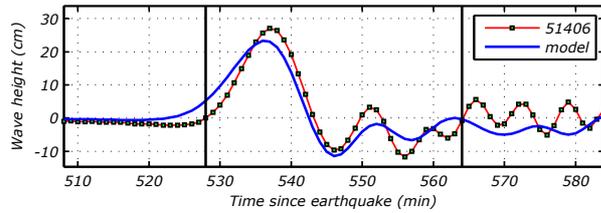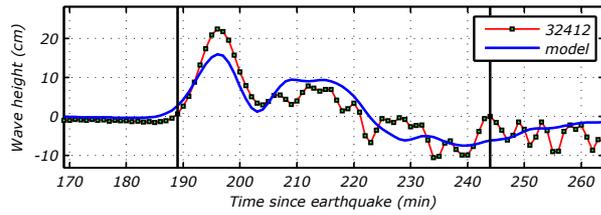
(c)

Figure 5.18: Continued.
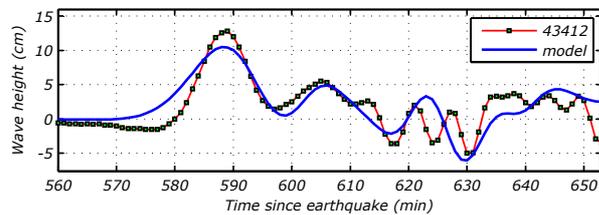
(a)

Figure 5.19: Tsunami source function weights calculated from the DARTs 32412 and 51406 unconstrained inversions are used to evaluate the Chile tsunami offshore amplitudes at the other DART 43412 by using (a) L_1, (b) L_2 and (c) L_∞ norms. RMS are errors 2.245, 2.655 and 2.548 respectively.
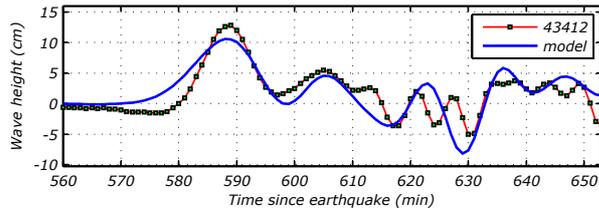
(b)
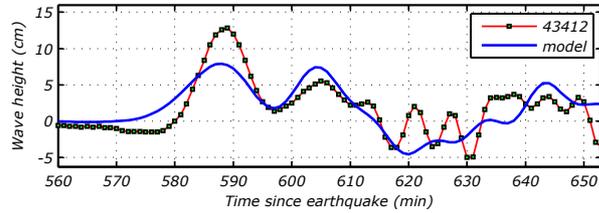
(c)

Figure 5.19: Continued.

5.2.4 Three DARTs unconstrained inversion for the 27 February 2010 Chile tsunami

Finally, three DARTs 32412, 51406 and 43412 are used in the inversion for the tsunami source functions 87-93 on the Central-South America subduction zone. Inversion is left unconstrained and the time interval for the inversion is defined from 189 to 244 minutes and from 528 to 564 minutes after the earthquake for the DARTs 32412 and 51406 respectively as in the previous case and from 580 to 633 minutes after the earthquake for the DART 43412. Population size is defined as 300 and generation is limited to 1000. L_1, L_2 and L_3 error minimization norms are used as fitness functions and results are presented in Figures 5.20(a) - 5.20(c) and weights and RMS errors are listed in Table 5.5.
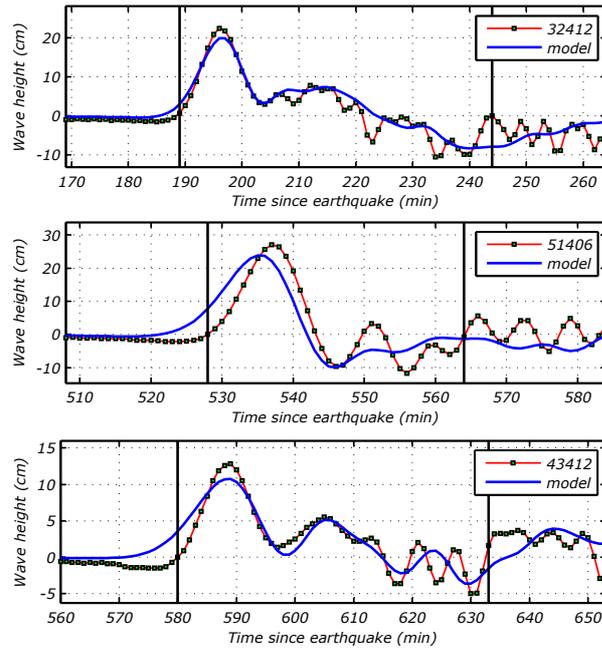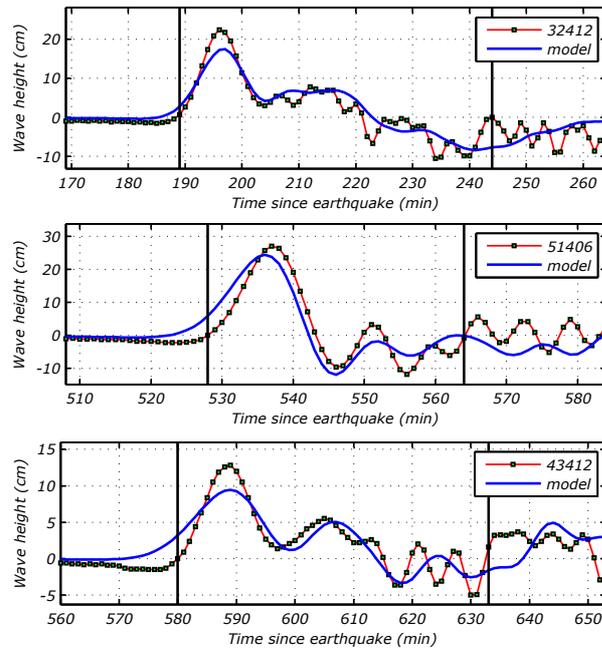
Figure 5.20: Three DARTs 32412, 51406 and 43412 unconstrained inversion results for the Chile tsunami for (a) L_1, b) L_2 and (c) L_∞ norms. RMS errors are 2.796, 5.712 and 1.788 for L_1, 3.070, 4.399 and 1.963 for L_2, 2.961, 7.554 and 2.408 for L_∞ for the DARTs 32412, 51406 and 43412 respectively. Solid vertical lines show the time interval used for the inversion.
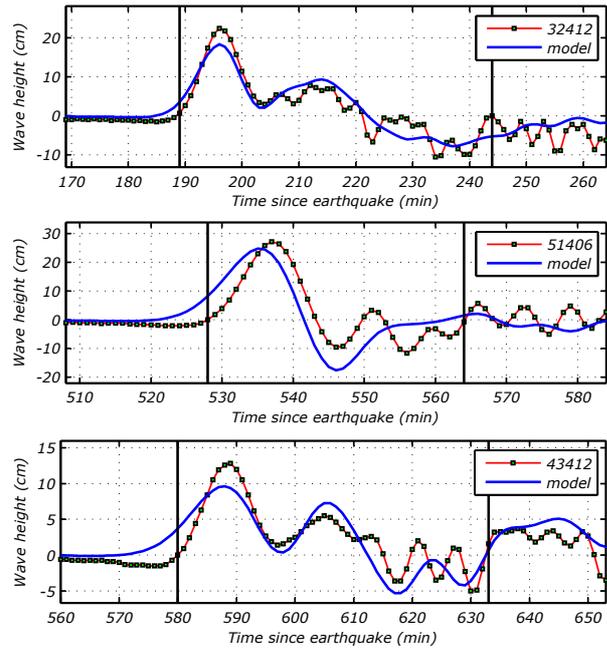
Figure 5.20: Continued.

Table 5.5: Tsunami source function weights based on different considerations in Subsections 5.2.1 - 5.2.4. Highligted RMS errors are for the DARTs which are not used for the inversion, i.e., inversion results are used to evaluate tsunami height estimates at the other DARTs.

| Conditions		Unconstrained									Constrained		
DART(s)		32412			32412 51406			32412 51406 43412			32412		
Norms		L_1	L_2	L_∞	L_1	L_2	L_∞	L_1	L_2	L_∞	L_1	L_2	L_∞
Tsunami source functions	cssz087a	5.77	5.30	10.49	7.47	7.08	6.57	6.43	6.67	6.37	5.30	5.59	5.80
	cssz087b	0.00	0.01	1.47	0.01	0.07	0.04	0.13	0.19	0.47	0.11	0.01	0.05
	cssz088a	11.08	11.53	9.54	11.69	12.22	10.38	10.90	9.36	10.32	13.09	13.51	10.45
	cssz088b	1.71	0.97	0.99	0.20	0.22	0.02	0.33	0.28	1.90	0.48	0.55	3.63
	cssz088z	11.36	13.52	7.88	13.65	11.86	9.12	10.76	9.88	10.13	9.54	13.84	10.64
	cssz089a	10.29	8.13	7.34	6.94	6.74	3.49	7.65	9.32	7.52	5.38	5.62	7.58
	cssz089b	9.80	9.47	4.28	1.44	2.09	0.68	6.25	0.18	6.01	1.42	2.14	5.59
	cssz089z	8.19	0.25	6.76	7.40	6.25	7.87	8.58	9.49	6.58	8.76	2.49	2.80
	cssz090a	6.08	8.01	6.49	5.00	7.14	5.86	8.40	4.41	5.97	9.17	7.80	5.37
	cssz090b	5.94	8.91	4.18	2.60	5.04	0.50	5.57	0.40	4.05	4.04	3.19	4.69
	cssz090z	0.40	0.11	1.13	4.12	1.51	6.78	5.21	3.90	2.53	4.27	0.88	2.01
	cssz091a	7.97	7.54	6.62	5.74	6.45	5.27	4.59	5.91	8.47	2.03	5.31	5.49
	cssz091b	4.12	10.09	5.62	0.05	0.64	0.11	0.03	1.02	0.53	0.36	0.07	7.90
	cssz091z	2.35	7.15	4.12	1.54	3.54	3.02	3.92	3.09	3.23	3.55	4.44	4.57
	cssz092a	3.03	5.36	2.86	0.55	0.20	3.25	3.19	2.35	5.77	0.89	0.62	2.34
	cssz092b	13.63	21.94	5.96	6.85	10.72	3.99	4.79	5.02	7.39	10.18	10.90	5.86
	cssz092z	4.62	0.48	2.04	5.67	1.43	2.87	3.10	5.35	0.39	1.05	1.64	0.19
	cssz093a	6.15	8.61	2.83	4.07	7.59	2.47	2.61	1.13	3.97	5.37	7.27	1.71
	cssz093b	0.04	0.08	2.59	0.20	0.03	4.65	1.12	2.34	0.52	0.04	0.03	2.18
	cssz093z	5.67	3.25	2.62	9.04	4.56	6.20	7.98	7.08	5.03	3.87	3.01	0.08
	Total slip (m)	118.23	130.70	95.79	94.20	95.36	83.12	101.55	87.38	97.16	88.91	88.91	88.91
	M_w	8.88	8.91	8.82	8.82	8.82	8.78	8.84	8.79	8.83	8.80	8.80	8.80
RMS error	for 32412	2.286	1.942	3.238	2.617	2.161	3.772	2.796	3.070	2.961	2.391	2.152	2.971
	for 51406	9.697	14.659	8.972	4.674	8.057	3.952	5.712	4.399	7.554	7.139	8.249	10.573
	for 43412	2.671	2.876	2.457	2.245	2.655	2.548	1.788	1.963	2.408	2.340	2.617	2.567

5.2.5 Discussion of the 27 February 2010 Chile tsunami source inversion results

Tsunami source function weights are recalculated for larger population sizes. In Table 5.6, population sizes are selected from 100 to 1000 for single DART and inversion is performed only for L_2 norm. As in the Kuril Islands tsunami source inversion analysis in Subsection

5.1.5, when the results in Tables 5.5 and 5.6 are compared they are close to each other, i.e., there are slight differences resulted from randomness of genetic algorithm. However, increasing population size results in more stable weights. Solution with larger population takes more time thus obtaining solution with an optimum population might save time in real-time forecasting.

Tsunami source function weights for the population size of 1000 given in Table 5.6 are averaged and used to evaluate forecast models at several coastal locations (Tang, personal communication)[3]. Forecast model results are compared with tide gage measurements. Comparisons show reasonable agreements on both tsunami arrival times and wave amplitudes (Figure 5.21).

Table 5.6: Tsunami source function weights of single DART 32412 solution of the 27 February 2010 Chile tsunami based on L_2 norm for various population sizes from 100 to 1000. Significant ones are highligted.

		Population size											
		100	200	300	400	500	600	700	800	900	1000	1000	1000
Tsunami source functions	cssz087a	5.72	5.62	6.71	5.57	6.51	6.36	6.28	6.47	6.60	6.45	6.50	6.86
	cssz087b	0.00	0.05	0.11	0.06	0.00	0.03	0.08	0.00	0.03	0.09	0.02	0.07
	cssz088a	11.52	11.56	11.77	10.80	12.06	11.73	11.24	10.74	11.41	11.39	10.86	11.30
	cssz088b	1.17	0.82	0.10	0.84	0.21	0.00	0.13	0.18	0.04	0.02	0.06	0.05
	cssz088z	10.31	12.52	11.14	13.95	11.29	11.48	10.86	11.21	14.06	11.79	12.07	11.67
	cssz089a	9.93	9.17	7.92	10.07	6.31	6.50	7.44	9.72	7.53	7.30	8.24	7.68
	cssz089b	6.81	7.66	4.41	9.51	5.42	6.97	7.45	6.89	6.01	6.71	7.16	5.60
	cssz089z	2.19	0.20	0.18	0.12	0.08	0.23	0.24	0.11	0.04	0.02	0.10	0.08
	cssz090a	10.61	6.24	10.69	3.65	10.34	9.89	8.56	8.49	4.94	8.96	6.82	7.74
	cssz090b	6.44	7.42	8.31	6.22	11.63	11.84	11.88	7.81	8.57	10.56	9.74	9.98
	cssz090z	0.46	0.10	0.16	0.01	0.00	0.13	0.10	0.10	0.01	0.07	0.00	0.15
	cssz091a	7.58	8.78	6.20	9.16	4.33	4.58	6.87	7.71	7.75	6.08	6.67	6.74
	cssz091b	5.38	15.14	9.99	16.02	16.39	16.05	19.09	13.42	16.62	15.38	17.62	17.20
	cssz091z	4.13	2.25	5.77	3.87	8.11	9.34	3.28	3.18	2.55	5.47	6.09	3.85
	cssz092a	6.05	8.24	7.76	6.44	9.51	8.48	11.29	9.22	8.58	9.48	8.41	9.79
	cssz092b	20.15	26.50	26.06	26.02	29.97	28.51	30.53	28.84	27.78	29.01	29.86	30.17
	cssz092z	0.04	0.18	0.12	0.08	0.22	0.08	0.06	0.02	0.05	0.00	0.10	0.12
	cssz093a	9.63	9.31	9.58	8.78	8.28	8.08	8.55	9.82	8.89	8.76	8.57	8.96
	cssz093b	0.18	0.05	0.04	0.01	0.10	0.01	0.06	0.08	0.01	0.10	0.06	0.10
	cssz093z	1.78	3.73	2.16	4.88	2.00	2.13	2.62	2.87	5.70	2.97	3.06	3.47
	Total slip (m)	120.06	135.55	129.16	136.05	142.77	142.42	146.62	136.87	137.17	140.62	142.03	141.59
	M_w	8.89	8.92	8.91	8.92	8.94	8.94	8.94	8.92	8.93	8.93	8.94	8.93
	RMS error	1.980	1.902	1.909	1.904	1.865	1.874	1.861	1.889	1.876	1.873	1.869	1.868

---

[3] Model runs were performed by Liujuan Tang of NCTR using their forecasting models.
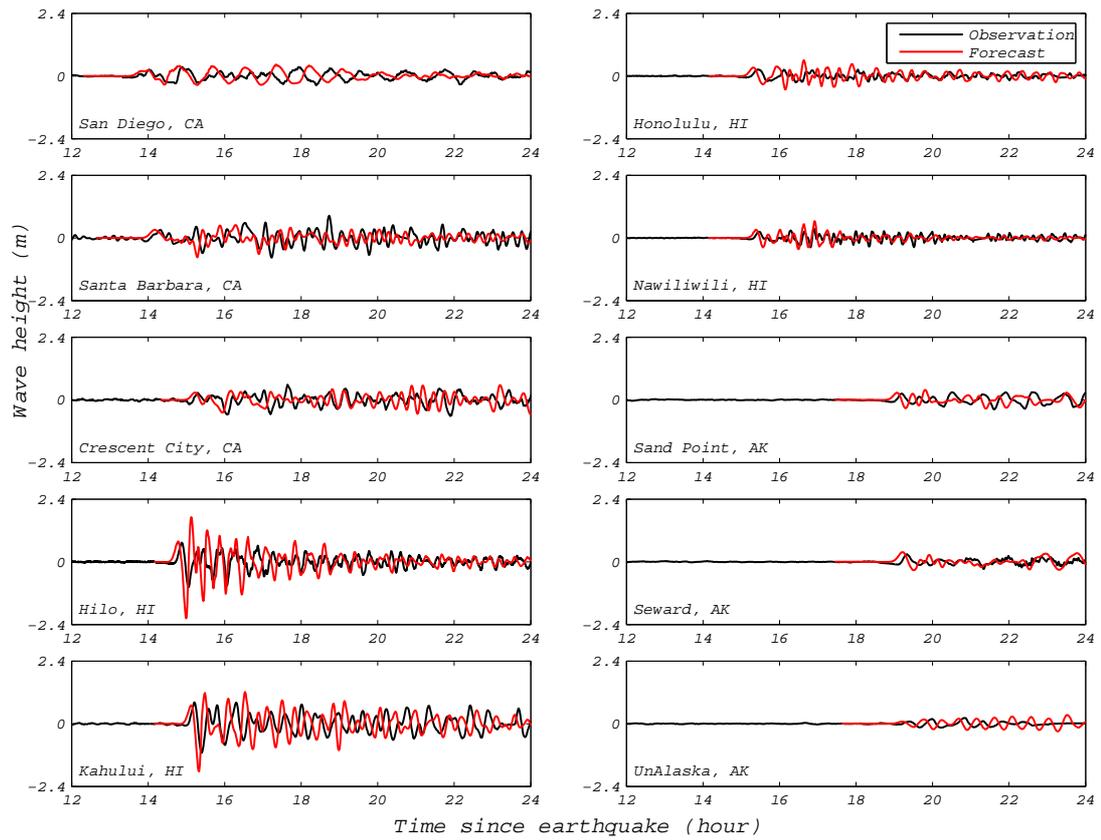
Figure 5.21: Comparison of tide gage measurements with the tsunami inversion results based on source combination 5.99·87a + 11.12·88a + 12.32·88z + 7.83·89a + 8.39·89b + 7.08·90a + 10.02·90b + 7.08·91a + 17.25·91b + 3.44·91z + 9.68·92a + 29.37·92b + 8.89·93a + 3.57·93z at several coastal locations. Tsunami source function weights are the average weights for the population size of 1000 given in Table 5.6 for the Chile tsunami. Forecast model runs are provided by Liujuan Tang (personal communication).

CHAPTER 6

CONCLUSIONS

A graphical user interface (GUI) called Genetic Algorithm for INversion (GAIN) was developed to determine tsunami source by inverting Deep-ocean Assessment and Reporting of Tsunami (DART$^{\text{TM}}$) buoy measurement(s) against pre-computed tsunami source functions. MATLAB$^{\text{TM}}$ genetic algorithm toolbox is used as a solver for GAIN. The 15 November 2006 Kuril Island and the 27 February 2010 Chile tsunamis are selected to exercise the GAIN. It was not the emphasis of the study to develop an inversion methodology. However, some preliminary additional tasks are carried out to investigate several features of the inversion such as:

- Different error minimization norms are considered for inversion, i.e., L_1 (the sum of the magnitudes of the residuals), L_2 (the sum of the square of the residuals) and L_∞ (the largest magnitude of the residuals) norms.

- Single DART, two and three DARTs inversions are obtained for each norm.

- Single DART inversion constrained with earthquake magnitudes are also obtained.

- The effect of DART measurements time interval used in the inversion is investigated.

- The effect of population size in genetic algorithm over the inversion is investigated.

In addition, obtained tsunami scenarios for the 15 November 2006 Kuril Islands and the 27 February 2010 Chile tsunamis are used to model offshore tsunami propagation and tide gage

67

data at several coastal cities (Tang, personal communication)[1]. Tide gage results are compared with the actual measurements.

At this point, the number of DARTs and tsunami source functions are limited to three and twenty respectively. In inversions with multiple DARTs, population size is kept higher since complexity is higher relative to single one. Also, tournament selection is assigned as a default operator for this case because it does not need sorting in every step which decreases elapsed time. Solution by genetic algorithm inversion has reasonable elapsed time so results can be used in coastal inundation modeling through Short-term Inundation Forecast for Tsunamis (SIFT^{TM}) and/or Community Model Interface for Tsunami (ComMIT^{TM}) for real-time forecasting.

Based on the results, the following preliminary conclusions can be drawn:

- Single DART unconstrained -without considering earthquake moment magnitude as a constrain- inversion appears to provide good estimates of tsunami amplitudes at offshore and at coastal locations, i.e., satisfactorily represents the measurements at other DARTs and tide gages in harbors.

- Earthquake moment magnitude constrained inversion does not necessarily improve tsunami source determination.

- Moment magnitudes calculated through DART inversion are close to moment magnitude calculated through seismic inversion.

- Data from multiple DARTs might be providing robust constraints for the selection of tsunami source functions which will be used in inversion. This might be needed for a more accurate forecast of tsunami arrival time and amplitude.

Currently, GAIN uses local pre-computed tsunami source function files and DART measurements. However, it is possible to upload data in real-time from NCTR's database with a slight modifications. In addition, enabling source function selection from a world map figure will enrich the interface. Also, inversion process is not automated, i.e., there is no self-selecting

---

[1] Model runs were performed by Liujuan Tang of NCTR using their forecasting models.

algorithm for tsunami source functions used, time interval, etc. If these features are added to GAIN, it may be considered as a useful additional tool to NCTR's Short-Term Inundation and Tsunami Forecasting (SIFT) system and/or Community Model Interface for Tsunamis (ComMIT).

BIBLIOGRAPHY

Altunkaynak, B. and Esin, A. (2004). The genetic algorithm method for parameter estimation in nonlinear regression. *Gazi University Journal of Science*, 17(2): 43–51.

Arcas, D. and Uslu, B. (2010). A tsunami forecast model for Crescent City, California. NOAA OAR Special Report, PMEL Tsunami Forecast Series: Vol. 2. 112 pp.

Behrens, J., Androsov, A., Babeyko, A., Harig, S., Klaschka, F., and Mentrup, L. (2010). A new multi-sensor approach to simulation assisted tsunami early warning. *Natural Hazards and Earth System Science*, 10(6): 1085–1100.

Bernard, E., Mofjeld, H., Titov, V., Synolakis, C., and González, F. (2006). Tsunami: scientific frontiers, mitigation, forecasting and policy implications. *Physical and Engineering Sciences*, 364(1845): 1989–2007.

Bernard, E. and Robinson, A. (2009). Tsunamis. In Bernard, E. and Robinson, A., editors, *The SEA*, volume 15, chapter 1, pages 1–22. Harvard University Press, Cambridge, MA and London, England.

Bernard, E. and Titov, V. (2006). Improving tsunami forecast skill using deep ocean observations. *Marine Technology Society Journal*, 40(4): 86–89.

Bruins, H., MacGillivray, J., Synolakis, C., Benjamini, C., Keller, J., Kisch, H., Klugel, A., and van der Plicht, J. (2008). Geoarchaeological tsunami deposits at Palaikastro (Crete) and the Late Minoan IA eruption of Santorini. *Journal of Archaeological Science*, 35(1): 191–212.

Coley, D. (1999). *An Introduction to Genetic Algorithms for Scientists and Engineers*. World Scientific Publishing. 185 pp.

Dengler, L., Uslu, B., Barberopoulou, A., Yim, S., and Kelly, A. (2009). The November 15, 2006 Kuril Islands-Generated Tsunami in Crescent City, California. *Pure and Applied Geophysics*, 166(1-2): 37–53.

Fritz, H., Petroff, C., Catalán, P., Cienfuegos, R., Winckler, P., Kalligeris, N., Weiss, R., Barrientos, S., Meneses, G., Valderas-Bermejo, C., Ebeling, C., Papadopoulos, A., Contreras, M., Almar, R., Dominguez, J., and Synolakis, C. (2011). Field survey of the 27 February 2010 Chile tsunami. *Pure and Applied Geophysics*, 168. (In press).

Gica, E., Spillane, M., Titov, V., Chamberlin, C., and Newman, J. (2008). Development of the forecast propagation database for NOAA's short-term inundation forecast for tsunamis (SIFT). NOAA Technical Memorandum OAR PMEL-139, NOAA / Pacific Marine Environmental Laboratory PMEL, Seattle, WA. 89 pp.

González, F., Bernard, E., Meinig, C., Eble, M., Mofjeld, H., and Stalin, S. (2005). The NTHMP tsunameter network. In Bernard, E., editor, *Developing Tsunami-Resilient Communities*, chapter 2, pages 25–39. Springer Netherlands.

Greenslade, D., Allen, S., and Simanjuntak, M. (2011). An evaluation of tsunami forecasts from the T2 scenario database. *Pure and Applied Geophysics*. (In print).

Greenslade, D., Simanjuntak, M., and Allen, S. (2009). An enhanced tsunami scenario database : T2. CAWCR technical report no. 014, The Centre for Australian Weather and Climate Research.

Greenslade, D., Simanjuntak, M., Chittleborough, J., and Burbidge, D. (2007). A first-generation realtime tsunami forecasting system for the Australian region. BMRC research report no. 126, Bureau of Meteorology, Australia.

Greenslade, D. and Titov, V. (2008). A comparison study of two numerical tsunami forecasting systems. *Pure and Applied Geophysics*, 165(11): 1991–2001.

Gusiakov, V. (1978). Static displacement on the surface of an elastic space. Ill-posed problems of mathematical physics and interpretation of geophysical data. *VC SOAN SSSR, Novosibirsk (in Russian)*, pages 23–51.

Haupt, R. and Haupt, S. (2004). *Practical Genetic Algorithms*. Wiley-Interscience, 2nd edition. 253 pp.

MATLAB (2010). *Global Optimization Toolbox User's Guide Ver. 3.1*. MathWorks Inc.

Mitchell, M. (1998). *An Introduction to Genetic Algorithms*. MIT Press, Cambridge, MA, USA, 5th edition. 158 pp.

Moore, C., Şen, C., Aydın, B., Tang, L., Titov, V., and Kânoğlu, U. (2010). Assessing the source of the 2010 Chilean tsunami using DART data. Abstract G33A-0829 presented at 2010 Fall Meeting, American Geophysical Union, San Francisco, California, 13-17 December.

NCTR (2010). NOAA Center for Tsunami Research (NCTR). Retrieved from `http://nctr.pmel.noaa.gov/propagation-database.html`. Last accessed: September 5, 2010.

NDBC (2011). DART buoy data. Retrieved from `http://www.ndbc.noaa.gov/dart.shtm`. Last accessed: January 03, 2011.

Okada, Y. (1985). Surface deformation due to shear and tensile faults in a half-space. *Bulletin of the Seismological Society of America*, 75(4): 1135–1154.

Okal, E. and Titov, V. (2007). MTSU : Recovering seismic moments from tsunameter records. *Pure and Applied Geophysics*, 164(2): 355–378.

Percival, D., Arcas, D., Denbo, D., Eble, M., Gica, E., Mofjeld, H., Spillane, M., Tang, L., and Titov, V. (2009). Extracting tsunami source parameters via inversion of DART® buoy data. NOAA Technical Memorandum OAR PMEL-144, NOAA / Pacific Marine Environmental Laboratory (PMEL), Seattle, WA. 22 pp.

Rao, S. S. (1996). *Engineering Optimization: Theory and Practice, 3rd Edition*. Wiley-Interscience, third edition. 922 pp.

Righi, D. and Arcas, D. (2010). A tsunami forecast model for Newport, Oregon. NOAA OAR special report, PMEL Tsunami Forecast Series: Vol. 5. 80 pp.

Satake, K., Hasegawa, Y., Nishimae, Y., and Igarashi, Y. (2008). Recent tsunamis that affected the Japanese coasts and evaluation of JMA's tsunami warnings. *EOS Trans. American Geophysical Union*, 89(53), Fall Meet. Suppl., Abstract OS42B-01.

Stein, S. and Okal, E. (2008). Ultralong period seismic study of the December 2004 Indian Ocean earthquake and implications for regional tectonics and the subduction process. *Bulletin of the Seismological Society of America*, 97(1A): S279–S295.

Synolakis, C., Bernard, E., Titov, V., Kânoğlu, U., and González, F. (2008). Validation and verification of tsunami numerical models. *Journal of Pure and Applied Geophysics*, 165(11-12): 2197–2228.

Synolakis, C., Liu, P., Philip, H., Carrier, G., and Yeh, H. (1997). Tsunamigenic sea-floor deformations. *Science*, 278(5338): 598–600.

Tang, L., Titov, V., and Chamberlin, C. (2009). Development, testing, and applications of site-specific tsunami inundation models for real-time forecasting. *Journal of Geophysical Research*, 114(C12):C12025+.

Tang, L., Titov, V., and Chamberlin, C. (2010). A tsunami forecast model for Hilo, Hawaii. NOAA OAR Special Report, PMEL Tsunami Forecast Series: Vol.1. 94 pp.

Titov, V. (2009). Tsunami forecasting. In Bernard, E. and Robinson, A., editors, *The SEA*, volume 15, chapter 12, pages 371–400. Harvard University Press, Cambridge, MA and London, England.

Titov, V. and González, F. (1997). Implementation and testing of the Method of Splitting Tsunami (MOST) model. NOAA Technical Memorandum ERL PMEL-112, NOAA / Pacific Marine Environmental Laboratory (PMEL), Seattle, WA. 11 pp.

Titov, V., González, F., Bernard, E., M.C. Eble, H. M., Newman, J., and Venturato, A. (2005a). Real-time tsunami forecasting: Challenges and solutions. In Bernard, E., editor, *Developing Tsunami-Resilient Communities*, chapter 3, pages 41–58. Springer Netherlands.

Titov, V., Mofjeld, H., González, F., and Newman, J. (1999). Offshore forecasting of Alaska-Aleutian subduction zone tsunamis in Hawaii. NOAA Technical Memorandum ERL PMEL-114, NOAA / Pacific Marine Environmental Laboratory PMEL, Seattle, WA.

Titov, V., Mofjeld, H., González, F., and Newman, J. (2001). Offshore forecasting of Alaska tsunamis in Hawaii. In Hebenstreit, G., editor, *Tsunami Research at the End of a Critical Decade*, pages 75–90. Kluwer Academic Publishers.

Titov, V., Moore, C., Greenslade, D., Pattiaratchi, C., Badal, R., Synolakis, C., and Kânoğlu, U. (2011). A new tool for inundation modeling: Community Modeling Interface for Tsunamis (ComMIT). *Journal of Pure and Applied Geophysics*, Accepted.

Titov, V., Rabinovich, A., Mofjeld, H., Thomson, R., and González, F. (2005b). The global reach of the 26 December 2004 Sumatra tsunami. *Science*, 309(5743): 2045–2048.

Tolkova, E. (2009). Principal component analysis of tsunami buoy record: Tide prediction and removal. *Dynamics of Atmospheres and Oceans*, 46(1-4): 62–82.

U.S. Geological Survey, Chile (2010). 2010 significant earthquake and news headline archive Chile earthquake. Retrieved from `http://earthquake.usgs.gov/earthquakes/eqinthenews/2010/us2010tfan/`. Last accessed: September 25, 2010.

U.S. Geological Survey, Kuril (2006). 2006 significant earthquake and news headline archive Kuril earthquake. Retrieved from `http://earthquake.usgs.gov/earthquakes/eqinthenews/2006/usvcam/`. Last accessed: September 25, 2010.

U.S. Geological Survey, Sumatra (2004). 2004 significant earthquake and news headline archive Sumatra earthquake. Retrieved from `http://earthquake.usgs.gov/earthquakes/eqinthenews/2004/usslav/`. Last accessed: September 23, 2010.

Uslu, B., Arcas, D., Titov, V., and Venturato, A. (2010). A tsunami forecast model for San Francisco, California. NOAA OAR Special Report, PMEL Tsunami Forecast Series: Vol. 3. 88 pp.

Wei, Y. and Arcas, D. (2010). PMEL tsunami forecast series: Vol. 4 a tsunami forecast model for Kodiak, Alaska. NOAA OAR special report, NOAA Center for Tsunami Research (NCTR), Seattle, WA. 98 pp.

Wei, Y., Bernard, E., Tang, L., Weiss, R., Titov, V., Moore, C., Spillane, M., Hopkins, M., and Kânoğlu, U. (2008). Real-time experimental forecast of the Peruvian tsunami of August 2007 for U.S. Coastlines. *Geophysical Research Letters*, 35(4): L04609.

Weise, T. (2009). Global optimization algorithms: Theory and application. Self-published.

Whitmore, P. (2003). Tsunami amplitude prediction during events: A test based on previous tsunamis. *Science of Tsunami Hazards*, 21: 135–143.

Whitmore, P. (2009). Tsunami warning systems. In Bernard, E. and Robinson, A., editors, *The SEA: Tsunamis*, volume 15, chapter 14, pages 402–443. Harvard University Press.

APPENDIX A

STEPS TO RUN GAIN

Before running the GAIN it is required to provide DART buoy measurement and tsunami source function files. DART buoy measurement files must be MATLAB^{TM} '.mat' file in the 'DART' subfolder including two variables called 't' and 'felev' which represent time in minutes and wave amplitudes in meters respectively. Tsunami source function files are required to be included in 'SOURCE' subfolder. File structure of tsunami source file is given in Table A.1. First row in tsunami source file includes DART station names, first column includes time in seconds and other columns includes tsunami source function time series at the specified DART stations in centimeters. When the DART station is selected GAIN reads tsunami source file, find the column belong to selected DART and uploads the pre-computed time series data from that column.

Steps to run GAIN is summarized as follows and shown in Figure A.1:

1. Select the subduction zone at which earthquake occurred and sources around the epicenter with an appropriate range (1a-d in Figure A.1).

2. Select DART station(s) close to the epicenter and define the inversion time interval(s) using the sliding bars or edit boxes (2a-c in Figure A.1).

3. Select the sources from source selection panel in order to use in the inversion (3 in Figure A.1).

4. Choose fitness function and genetic algorithm parameters (4 in Figure A.1).

5. All these selections are submitted to MATLAB^{TM} genetic algorithm optimization toolbox by using 'Solve' button and results are plotted (5 in Figure A.1).

6. Solution can be exported as a text file or session can be saved when it is required (6 in Figure A.1).

Table A.1: Structure of the tsunami source function files. Here, time series are given for the tsunami source function cssz085a at the location of the DARTs 21413, 21414, ..., 99902 and 99903 location (File named as 'dart_cssza85_linear.dat').

0	21413	21414	21415	$\cdots$	99901	99902	99903
0	0.0000	0.0000	0.0000	$\cdots$	0.0000	0.0000	0.0000
60	0.0000	0.0000	0.0000	$\cdots$	0.0000	0.0000	0.0000
120	0.0000	0.0000	0.0000	$\cdots$	0.0000	0.0000	0.0000
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$
86280	0.0442	-0.0133	-0.0653	$\cdots$	-0.0845	-0.1020	0.0065
86340	0.0453	-0.0192	-0.0576	$\cdots$	-0.0949	-0.1096	0.0138
86400	0.0388	-0.0247	-0.0425	$\cdots$	-0.0965	-0.1080	0.0191

Table A.2: Sample report of GAIN.

DART station(s) : 21414
Left boundary (min) : 112
Right boundary (min) : 147
Fitness function : 2
Constraint : 1
Population size : 100
Generation : 1000

Sources	Weigthts
kisz012a	5.66
kisz012b	0.13
kisz012y	0.00
kisz012z	0.05
kisz013a	3.57
kisz013b	0.00
kisz013y	0.00
kisz013z	0.34
kisz014a	0.00
kisz014b	0.92
kisz014y	0.61
kisz014z	0.00

Total slip (m) : 11.30
Moment magnitude : 8.20
FVAL : 3.85
RMS error : 0.327
Final generation : 202
Optimization terminated :
average change in the fitness
value less than options.TolFun.

Figure A.1: Steps to run GAIN.