TSUNAMI SOURCE ESTIMATION AND WAVEFORM FORECAST USING COMPUTATIONAL INTELLIGENCE

by

IYAN EKA MULIA

DISSERTATION

Presented to the Faculty of the Graduate School of Science and Engineering

Kagoshima University

in Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF ENGINEERING

KAGOSHIMA UNIVERSITY

2016

The Dissertation Committee for Iyan Eka Mulia Certifies that this is the approved version of the following dissertation:

TSUNAMI SOURCE ESTIMATION AND WAVEFORM FORECAST USING COMPUTATIONAL INTELLIGENCE

Committee:

Prof. Toshiyuki Asano (Supervisor)

Prof. Takahiro Adachi

Asc. Prof. Taro Kakinuma

Acknowledgements

I would like to express the deepest appreciation to my academic advisor Prof. Toshiyuki Asano for his valuable guidance and continuous supports during my doctoral study and research.

Besides my advisor, I would like to thank the thesis committee: Prof. Takahiro Adachi and Asc. Prof. Taro Kakinuma for their insightful comments and suggestions to improve the quality of my research.

I would also like to thank Dr. Tatsuhiko Saito for reviewing the thesis and for taking the time to attend my thesis defense.

My sincere thanks also goes to all member and staff of Dept. of Ocean and Civil Engineering, Kagoshima University who provided me an opportunity to join and adapt to the Japanese culture and traditions, particularly in the working environments, such as laboratory and research facilities.

I greatly appreciate the financial support from the Monbukagakusho scholarship provided by the Ministry of Education, Culture, Sports, Science and Technology, Japan.

Last but not the least, I would like to thank my family for supporting me spiritually throughout the thesis writing process and my life in general.

Abstract

In tsunami waveform inversions that use a conventional Green's function technique, an optimal solution can be difficult to obtain for a number of reasons. This study proposes a new method to both optimize the determination of the unknown parameters and introduce computational intelligence for tsunami waveform inversion. The main objective is to enhance the traditional least squares inversion used for two different applications. A genetic algorithm combined with a pattern search method is used for the first application (tsunami source estimation), and an extreme learning machine is used for the second application (forecasting tsunami waveforms). Based on the results, computational intelligence methods have been found to successfully enhance the standard least squares inversion both for estimating the tsunami source and for forecasting the tsunami waveforms.

In the first part of this study, estimation of tsunami source is conducted using a genetic algorithm and pattern search. The method aims to either find a suitable distribution of unit source locations or optimal least square parameters prior to the inversion. The method has been tested using both an artificial tsunami source and the real tsunami that occurred during the 2011 Tohoku event. Unlike the conventional method, which is characterized by equidistant unit sources, the proposed method generates random and scattered unit sources inside the inverse region or the source-influenced area. This leads to a better approximation of the initial profile of a tsunami. Additionally, it significantly reduces the number of parameters and suppresses the negative effects of regularization schemes that decrease the plausibility of the model. Furthermore, the stochastic approach used to derive the time delays (in the real case) is a more flexible strategy for simulating actual phenomena that occur in nature. In general, the proposed method has improved the ability of such models to reveal the underlying physics associated with the tsunami generating processes.

The second part of this study introduces an extreme learning machine to produce rapid forecasts of tsunami waveforms in coastal areas using tsunami signals recorded at specified locations. The remarkable training speed of the algorithm means that it can run in real-time, and therefore it is suitable for early warning systems in near-field tsunami events. Additionally, as a universal function approximator, the proposed method can capture nonlinearities exhibited by the tsunami. Therefore, it provides advantages over the standard inversion analysis used in many existing studies that are typically developed under a linear assumption. The proposed method is applied to the 2011 Tohoku tsunami. The results demonstrate that the proposed method is more accurate and does not significantly increase the computation time when compared with the standard method based on a tsunami waveform inversion. Furthermore, the model uncertainty analysis proves that the method is robust and reliable, despite its dependency on random input weights and biases.

Table of Contents

Acknowledgements ..... **iii**

Abstract ..... **v**

Table of Contents ..... **vii**

List of Tables ..... **x**

List of Figures ..... **xi**

Abbreviations ..... **xiv**

Chapter 1 ..... **1**

1.1 Tsunami Source Estimation 2

1.2 Tsunami Waveform Forecasting 4

1.3 Thesis Outline 5

Chapter 2 ..... **8**

2.1 Basic Formulation of Tsunami Waveform Inversion 8

2.2 Computation of Tsunami Waveforms 11

2.3 Single Time Window Inversion 13

2.4 Multiple Time Window Inversion 14

Chapter 3 ..... **16**

3.1 Various Methodologies in Computational Intelligence 16

3.2 Optimization algorithms 17

3.2.1 Genetic Algorithm 17

3.2.2 Patterns Search Method 20

3.3 Extreme Learning Machine 21

Chapter 4 24

4.1 Problem Description 24

4.2 Numerical Experiments 26

4.2.1 Design Parameters 27

4.2.2 Fitness Function 29

4.2.3 Auxiliary Basis Function 31

4.2.4 Waveform Interpolation 32

4.3 Model Development 33

4.4 Results and Discussion 36

Chapter 5 42

5.1 Problem Description 42

5.2 Brief Reviews on the 2011 Tohoku Tsunami 43

5.3 Waveforms and Bathymetry Data 44

5.4 Determination of the Source Area 45

5.5 Green's Function 50

5.6 Regularization Scheme 50

5.7 Selection of Model Parameters 51

5.8 Time Delays Determination 53

5.9 Results and Discussion 55

5.9.1 Selected Model Parameters 56

5.9.2 Results for Different Time Delays 61

5.9.3 Effects of the Hybrid Optimization 67

Chapter 6 69

6.1 Problem Description 69

6.2 Tsunami Data and Bathymetry 70

6.3 Model Development 72

6.4 Model Evaluations 74

6.5 Uncertainty Analysis 75

6.6 Results and Discussion 76

6.6.1 General Results 77

6.6.2 Model Uncertainty 81

Chapter 7 85

7.1 Tsunami Source Estimation (Ideal case) 85

7.2 Tsunami Source Estimation (Real case) 86

7.3 Tsunami Waveform Forecasting 87

References 89

List of Tables

Table 4.1 Summary of statistical evaluations based of the proposed cost function. ... 37

Table 5.1 GA and PS default parameter values. 53

Table 6.1 Statistical evaluation results. 79

Table 6.2 Maximum amplitude and arrival time discrepancies at coastal gauges. Negative original values indicate underestimates and positive values indicate overestimates. Gauge names with brackets are the stations with incomplete data. 80

List of Figures

Figure 2.1 Tsunami waveform inversion using the Green's function technique. 9

Figure 3.1 Schematic illustration of a GA. 17

Figure 4.1 Study area and bathymetry profile. Red dots indicate unit sources located throughout the inverse region. Green triangles with numbers are artificial observation stations. 26

Figure 4.2 Example of a unique identification number of the forward model computational grid used for the second design parameters. 29

Figure 4.3 Gaussian basis function. 31

Figure 4.4 Waveform interpolation. Left-hand figure shows selected location of a unit source indicated by a black square. Blue dots represent the four nearest unit sources used in the interpolation. Right-hand figures are comparisons of waveforms between numerical model (solid black line) and interpolation (dashed red line) at the artificial gauges. 34

Figure 4.5 Flowchart of model development procedure. 35

Figure 4.6 Tsunami initial wave sources. Gray dots indicate the centroid of unit sources. (a) Artificial tsunami source as the target to be approximated, (b) inverted source using least squares method, (c) GAPSu model, (d) GAPSr model. 39

Figure 4.7 Comparison of waveforms at Gauges. Gray bar above the time axis indicates the time range for the inversion. 40

Figure 4.8 Scatter plots of each method with respect to the waveforms of the target source at all gauges. 41

Figure 5.1 The scheme of travel time computation. 46

Figure 5.2 The 16-dot template for tsunami travel time computation. 48

Figure 5.3 Model domain and source area (yellow). The black star indicates the location of the epicenter of the 2011 Tohoku earthquake. The contour lines are the backward-propagated tsunami arrival times from their corresponding observation stations. 49

Figure 5.4 The L-curve plot used to choose an appropriate spatial smoothing quantity in the inversion (γ =1.833) 56

Figure 5.5 Inverted sea surface deformation (left) and corresponding statistical measures of the waveforms fit at all gauges (right). (a) Initial Green's function based on 528 unit sources. (b) After GA-PS optimization with 200 unit sources. The gray dots in the left figures indicate the centroids of the unit sources. The broken and continuous black contour lines show subsidence (0.5 m interval) and uplift (2 m interval). The green contour lines indicate the 0 m elevation inside the source area. 59

Figure 5.6 Comparisons of observed (black) and inverted waveforms at all gauges obtained using the initial Green's function based on 528 unit sources (red) and after GA-PS optimization with 200 unit sources (blue). Solid black lines indicate the time interval of the data being used for the inversion. 60

Figure 5.7 Cumulative inverted sea surface deformation after 180 s, overlaid with estimated time delay contours (continuous black) and effective perturbation front contours (broken green) at 30-s intervals (left), and the corresponding statistical measures of the waveforms fit at all gauges (right). (a) Using a constant velocity of 2 km/s (Model 1); (b) using a direct GA-PS optimization of the time delays (Model 2); and (c) using varied velocities from the GA-PS optimization (Model 3). 62

Figure 5.8 Cumulative inverted sea surface deformation at 180 s for 30-s time intervals, obtained from the optimized model parameters and time delays (Model 3). The broken and continuous black contour lines show subsidence (0.5 m interval) and uplift (2 m interval). The green contour lines indicate the 0 m elevation inside the source area. 64

Figure 5.9 Comparisons of observed (black) and inverted waveforms at all gauges obtained using Model 1 (red), Model 2 (green), and Model 3 (blue). Solid black lines indicate the time interval of the data being used for the inversion. 66

Figure 5.10 Surface errors of the GA (blue) and PS (red) according to the specified design parameters. (a). Least squares model parameters selection. (b). Time delays selection in Model 2. (c). Time delays selection in Model 3. 67

Figure 6.1 Bathymetry profile of the study area and gauge locations. The red dots are the OBP gauges, the green squares are the GPS buoys, and the blue triangles are the coastal gauges (T: tide gauges and W: wave gauges). The black star is the epicenter of the 2011 Tohoku earthquake. 71

Figure 6.2 Equidistant unit sources at 20-km intervals, which are used to develop the synthetic waveforms database. The dots are the centroids of the unit sources. Orange dots indicate the unit sources selected for the 35-min training period. The black star is the epicenter of the 2011 Tohoku earthquake. 73

Figure 6.3 Comparison of waveforms from the observations, TWI, and ELM. Vertical dashed lines separate the 35-min training period from the subsequent forecasting period. Light red and light blue shadings indicate the 95% confidence interval for the TWI and ELM, respectively. The left figures are the offshore gauges and the right figures are the coastal gauges.... 77

Figure 6.4 Boxplots that indicate the differences between the observed and forecasted maximum amplitudes in the first leading waves and their arrival times, for 20 model runs at nine coastal gauges. The red bars represent the TWI results. Gauge names with brackets are the stations with incomplete data..... 82

Abbreviations

BC	Balance Criterion
CI	Computational Intelligence
COMCOT	Cornell Multi-grid Coupled Tsunami Model
DART	Deep-ocean Assessment and Reporting of Tsunamis
ELM	Extreme Learning Machine
GA	Genetic Algorithm
GEBCO	General Bathymetric Chart of the Oceans
GPS	Global Positioning System
MAE	Mean Absolute Error
MSDE	Mean Square Derivative Error
OBP	Ocean Bottom Pressure
PEP	Percent Error in Peak
PS	Pattern Search
RMSE	Root Mean Square Error
TEWS	Tsunami Early Warning System
TUNAMI	Tohoku University's Numerical Analysis Model for Investigation of Near-field tsunamis
TWI	Tsunami Waveform Inversion

Chapter 1

Introduction

Tsunamis are catastrophic natural events that can cause tremendous destruction of property and loss of lives. They can be categorized as one of the most destructive forces in nature, threatening vulnerable areas with little warning. There are several causes of tsunamis, including submarine earthquakes, landslides, volcanic eruptions, meteorite impacts, and other disturbances to the equilibrium state of oceanic hydrodynamics. However, a distinct proportion (> 80%) of tsunamis are generated by earthquakes (http://www.ngdc.noaa.gov/hazard/tsu.shtml, last accessed, July 19, 2015).

Over the last few decades, there have been several major earthquake-generated tsunamis that have greatly affected coastal communities and subsequently stimulated new tsunami research. Various methods and measures have been proposed that have led towards a better understanding of tsunami behaviors (e.g., those associated with generation, propagation, run-up, and inundation stages). Computer-based modeling and simulations are one of the most popular approaches used in tsunami research (e.g., Imamura et al., 2006; Wang, 2009). Such approaches can be effective in addressing the challenges associated with the tsunami disaster countermeasures. However, many of

these have been developed based on conventional computation methods, and thus are unable to reach the full potential of modern computer technologies.

Recent developments in computing have resulted in a new paradigm of modeling and data-handling techniques that help to reveal the underlying dynamics of natural phenomena, which are sometimes not accessible by conventional approaches. Such methods are known as "computational intelligence", and they have been extensively developed and used within scientific and engineering communities. Unfortunately, the applications of such methods in tsunami research are limited. Motivated by that fact, this study proposes novel approaches for estimating the initial tsunami source and forecasting tsunami waveforms using computational intelligence. Although both of these goals use different methods, the cores of the algorithms are actually similar; both are based on a least-squares tsunami waveform inversion (TWI) that is further enhanced by computational intelligence approaches. Genetic algorithm (GA) and pattern search (PS) methods are used for the first goal (tsunami source estimation), and an extreme learning machine (ELM) is used for the second goal (tsunami waveform forecasting).

1.1 Tsunami Source Estimation

Direct observations of an initial tsunami source or sea surface deformation following the occurrence of an earthquake are rarely possible; therefore, estimations of such movements are often performed by considering relevant seismic information or the hydrodynamic response of the sea determined from recorded tsunami waveforms. More specifically, such estimations can be made by characterizing co-seismic deformations of the ocean floor, using either an analytical solution of the stationary problem in elastic theory, or the development of detailed inverse models of geodetic, seismic, or tsunami waveforms.

Determination of the initial tsunami source generated by earthquakes is crucial to the success of tsunami modeling. One of the most frequently used methods is to presume it from a fault model based on the elastic theory (Okada, 1985). However, this model assumes an average uniform slip, which is probably adequate only for far-field simulations. A more realistic approach was proposed by Satake (1987), who analyzed recorded waveforms to infer earthquake source parameters, or particularly coseismic slip, using the Green's function technique. Even though the fault model is still required, the division of a fault into smaller sub-faults allows the slip to be estimated in a heterogeneous manner, which leads to a better approximation of seafloor displacement.

The main limitation of using fault models to estimate an initial tsunami source is that they typically require a general assumption that is related to energy conversion processes. That is, they assume that the initial sea surface profile is similar to the vertical component of the seafloor displacement caused by the fault rupture. Such an assumption is relevant and appropriate for large interplate events occurring on gently dipping faults, but may not be adequate for tsunamigenic earthquakes characterized by a steep dip angle (Saito and Furumura, 2009). A study by Nosov and Kolesov (2007) contradicted this assumption by considering the effect of water compressibility. Furthermore, tsunami excitation is sometimes triggered by other parameters that are independent of seismic moment (Geist, 2002).

We should, therefore, develop a method that directly captures the signature of the tsunami generating processes using the recorded tsunami waveforms to accurately estimate the tsunami source. A simpler method was actually introduced earlier by Aida (1972) for which no prior assumption of a fault model was needed. The basic premise is to replace the fault model by an auxiliary basis function with unit sources, which is equivalent to the sub-fault approach by Satake (1987). More recently, several studies have been widely developed using tsunami waveform inversion to estimate the tsunami source without fault model assumptions (Baba et al., 2005; Satake et al., 2005; Saito et

al., 2010). This study is in line with those previously mentioned because the main interest here is to estimate sea surface deformation, rather than a slip on the fault plane.

1.2 Tsunami Waveform Forecasting

Reliable forecasting of a tsunami generated by a submarine earthquake is a difficult process to undertake. This is particularly true for near-field events, where there is limited time to execute a warning system and disseminate useful information. Even though the current global network of sensors can be used to immediately retrieve information on earthquake parameters, they do not always explain the exact characteristics of the resultant tsunami (Geist, 2002; Tsushima et al., 2009). Therefore, a tsunami early warning system (TEWS) based solely on seismic data will produce a relatively large number of false-positive warnings (Behrens et al., 2010).

The best way to identify a tsunami is from its incident waves as detected by water level gauges. Subsequently, we can assimilate the truncated data to forecast the tsunami waveforms at other locations. However, we must consider the trade-off between forecasting accuracy and speed. Additionally, the spatial coverage of the deployed tsunami gauges also has an important impact on the effectiveness of the forecast.

Standard tsunami forecasting methods based on recorded waveforms use the TWI. In general, this method can be independent of seismic information, and very accurately estimate the possibility of a tsunami. It can also produce a complete series of waveforms rather than only extracting certain wave features such as maximum amplitudes and arrival times. Many others have proposed different TWI methods using various designs and settings to develop both far-field and near-field TEWSs (e.g., Koike et al., 2003; Tsushima et al., 2009; Yasuda and Mase, 2013). A more advanced TEWS can even provide site-specific inundation forecasts within a reasonable time (Gusman

et al., 2014). However, this requires intensive computational efforts to construct a high-resolution inundation database, which is beyond the scope of this study. Aside from their effectiveness in detecting tsunamis, most TWI methods are not applicable to nonlinear processes because they impose a linear assumption. Therefore, a method that captures the nonlinearities in tsunamis should be introduced, particularly in coastal areas where nonlinear effects prevail.

The second part of the thesis discusses the application of a recent neural network algorithm (ELM) to produce rapid forecasts of tsunami waveforms in coastal areas using tsunami signals recorded at specified locations. The remarkable training speed of the algorithm means that it can run in real-time, and therefore it is suitable for a TEWS in near-field tsunami events. Additionally, as a universal function approximator, the proposed method can capture nonlinearities exhibited by the tsunami. Therefore, it provides advantages over the standard linear inversion analysis used in many existing studies.

1.3 Thesis Outline

Chapter 1 explains the general backgrounds, motivations, and expected contributions of the study. The chapter also establishes the main framework of the thesis that is divided into two parts: the estimation of tsunami sources and forecasting tsunami waveforms.

Chapter 2 explains the fundamental theory of the TWI method, which is the core of the algorithm used in the both parts of this study (estimation of tsunami source and waveform forecasting). The description also includes a general formulation of a regularization scheme that is typically incorporated into the TWI to maintain stable solutions. Additionally, mathematical expressions for single and multiple time window inversions are presented.

Chapter 3 describes the general theories of the CI comprised of a GA and an ELM. A brief introduction to the PS routine as a supplement for the GA is also described. For the GA and PS, the descriptions are mainly in generalized forms, and thus may not be directly related to the TWI. Details on the connections of the GA-PS algorithms to particular TWI applications are discussed separately in chapters **Chapter 4** and **Chapter 5**.

Chapter 4 discusses the first attempt to apply the proposed method using GA and PS to estimate the tsunami source. In this chapter, all data (i.e., tsunami source, waveforms) are artificial except for bathymetry. The GA-PS algorithm is first tested using the same design parameters as in the TWI. The purpose of this test is simply to compare the performance of the proposed GA-PS method with the traditional least squares inversion in the same model design and environment. The second design parameters are actually the essence of the study. The design aims to find the optimal locations of unit sources that, at the initial state, are distributed randomly around the tsunami source. The use of the artificial data in an ideal case allows assessment of the advantage of the proposed method in a more detailed manner.

Chapter 5 is an extension of **Chapter 4**. Here, the GA-PS algorithm is applied to the real case of the 2011 Tohoku tsunami. However, there are several additional steps which were previously not included in the ideal case. For example, the determination of the source area uses backward-propagated tsunami arrival times and a multiple time window inversion that includes the computation of time delays. Other than those, in the real case, instead of using randomly distributed unit sources, the model uses a very dense set of unit sources and selects the optimal ones. Therefore, in the real case, waveform interpolation is not required.

Chapter 6 is the second part of the study, in which the use of the ELM is described in relation to tsunami waveform forecasting. The descriptions are started by a brief review of existing studies of tsunami forecasting using neural networks, and

followed by explaining the potential application of the ELM, which has never been used before. Various statistical analyses are conducted to assess the performance of the proposed model. Furthermore, an uncertainty analysis is also performed to ensure that the ELM has a consistent predictive power in spite of the random characteristics of the network parameters.

Chapter 7 presents an overall summary of the conclusions of the study, then goes on to provide detail for each of the chapters (**Chapter 4**, **Chapter 5** and **Chapter 6**).

Chapter 2

Tsunami Waveform Inversion

2.1 Basic Formulation of Tsunami Waveform Inversion

A TWI is based on the principle of linear superposition of predefined segments discretizing the tsunami source-influenced area. These segments can be represented by either a fault model or an auxiliary basis function. The main goal of the TWI is to determine the magnitude of the slip on a fault plane or to directly infer the initial displacements that excite the tsunami. The advantage of using tsunami waveforms over seismic waves is that the bathymetry data is relatively easy to obtain and more accurate compared to the seismic velocity structure (Satake, 1987). Therefore, effects on the wave propagation path can be accurately evaluated using numerical computation (forward modeling).

When an earthquake occurs, seismic waves originating from the source radiate through the solid body of the earth. These waves recorded at seismometer contain information about the earthquake source. Similarly, imprints of the tsunami generating process exist in the recorded tsunami waveforms. Therefore, this information can be used to infer the tsunami source, as a proxy for a direct observation, which is difficult to obtain. It should be noted, however, that the recorded tsunami waveforms may

exhibit other patterns besides the source information. For instance, interferences from harbors or local bathymetric effects usually exist in the waveform readings. Therefore, it is necessary to conduct a pre-analysis and thorough investigation prior to the application of the method.

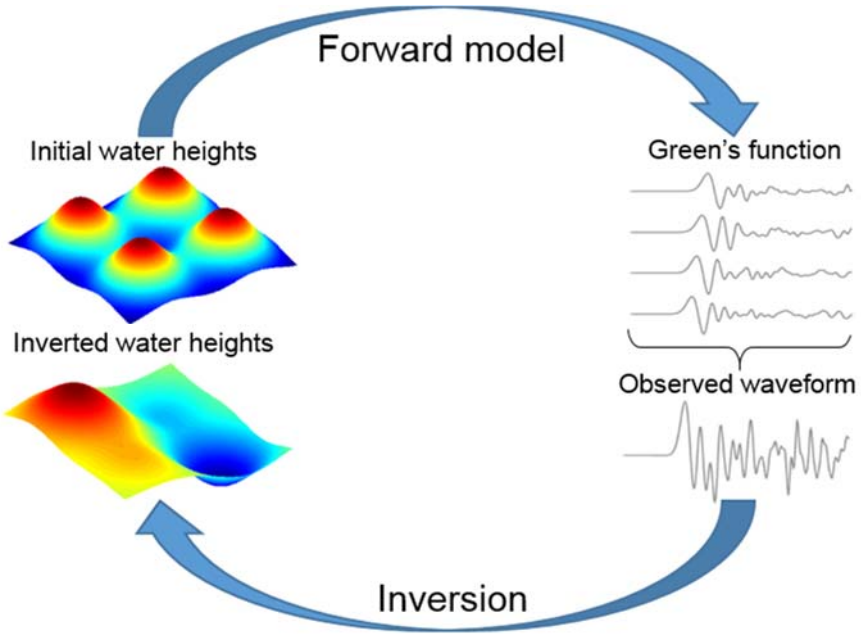

Figure 2.1 Tsunami waveform inversion using the Green's function technique.

The first step of the TWI development is to divide the source area into segments. This segments are typically termed as subfaults if one uses fault model, or in this study, the term unit sources is more relevant because the method requires no fault model. These unit sources represent the displacement of water surface in the source area after the earthquake. Then, compute the Green's function that is the synthetic waveforms recorded at measurement stations originating from each unit source with a certain initial amplitude or water height (for simplicity, it is usually set to 1 m). The actual water height on each unit source is obtained by an inversion; assuming that the observed tsunami waveforms are a linear superposition of the Green's function. A schematic

diagram on the development of a tsunami waveform inversion using the Green's function technique is presented in Figure 2.1

According to superposition theory, the measured tsunami waveforms at observation stations can be approximated using a superposition of the Green's function. That is,

$$\zeta_j(t) = \sum_{i=1}^{M} G_{ij}(t) w_i \quad (0 \le t \le T), \tag{2.1}$$

where $\zeta_j(t)$ is the observed waveform at the *j*-th location; $G_{ij}(t)$ is Green's function (which is a synthetic waveform originating from the *i*-th unit source, recorded at the *j*-th location); w_i is the displacement of the *i*-th unit source (the model parameter to be determined by the inversion); *M* is the number of unit sources; *t* is the time; and *T* is the maximum inversion time range, which defines the length of data used for the inversion at each observation station.

In vector form, we can write Eq. (2.1) as

$$\boldsymbol{\zeta} = \mathbf{G}\boldsymbol{w}, \tag{2.2}$$

where $\mathbf{G} \in \mathbb{R}^{m \times n}$ and $m > n$. To find $\boldsymbol{w}$, then minimize the norm of residual $\boldsymbol{r} = \mathbf{G}\boldsymbol{w} - \boldsymbol{\zeta}$, that is,

$$\|\boldsymbol{r}\|^2 = \boldsymbol{w}^T \mathbf{G}^T \mathbf{G}\boldsymbol{w} - 2\boldsymbol{\zeta}^T \mathbf{G}\boldsymbol{w} + \boldsymbol{\zeta}^T \boldsymbol{\zeta}, \tag{2.3}$$

setting the gradient with respect to $\boldsymbol{w}$ to zero, we have

$$\nabla_{\boldsymbol{w}} \|\boldsymbol{r}\|^2 = 2\mathbf{G}^T \mathbf{G}\boldsymbol{w} - 2\mathbf{G}^T \boldsymbol{\zeta} = 0. \tag{2.4}$$

Equation (2.4) leads to a normal equation $\mathbf{G}^T \mathbf{G}\boldsymbol{w} = \mathbf{G}^T \boldsymbol{\zeta}$. Assuming that $\mathbf{G}^T \mathbf{G}$ is invertible, we can get the least squares solution for $\boldsymbol{w}$,

$$\boldsymbol{w} = \left(\mathbf{G}^T\mathbf{G}\right)^{-1}\mathbf{G}^T\boldsymbol{\zeta}. \tag{2.5}$$

TWIs without the fault model assumption typically require a high spatial resolution to properly capture the complexity of the heterogeneous initial water profile. It therefore needs a large number of model parameters, which are proportional to the degrees of freedom in the optimization. To obtain a stable solution, we must accordingly apply a proper treatment such as regularization schemes. An example of a well-known regularization is the Tikhonov's regularization (Honerkamp and Weese, 1990). After introducing the regularization scheme, we now minimize both the norm of the residual and the additional regularization term, that is,

$$\min_{\boldsymbol{w}}\left[\left\|\boldsymbol{r}\right\|^2 + \gamma\left\|\boldsymbol{w}\right\|^2\right], \tag{2.6}$$

where $\gamma > 0$ is the regularization parameter. The solution of Eq. (2.6) is

$$\boldsymbol{w} = \left(\mathbf{G}^T\mathbf{G} + \gamma\mathbf{S}\right)^{-1}\mathbf{G}^T\boldsymbol{\zeta}, \tag{2.7}$$

where $\mathbf{S}$ can be either the identity matrix or a matrix that represents the derivative operator that measures spatial smoothness.

2.2 Computation of Tsunami Waveforms

Generally, the characteristics of tsunami propagation in deep water are linear. According to Satake (1987), even in shallow coastal areas, the first leading waves recorded at coastal tide gauges are still well simulated by the linear long wave model. Therefore, a typical linear non-dispersive shallow water equation is used in the forward modeling to compute time histories of sea surface elevation at the specified observation points, that is,

$$\begin{cases} \dfrac{\partial \boldsymbol{V}}{\partial t} = -g\nabla\eta \\ \dfrac{\partial \eta}{\partial t} = -\nabla\cdot\left\{\left(d+\eta\right)\boldsymbol{V}\right\} \end{cases}, \tag{2.8}$$

where η is the water elevation of the tsunami, $\boldsymbol{V}\left(u,v\right)$ is the depth-averaged horizontal fluid velocity vector, d is the water depth, and g is gravitational acceleration. Equation (2.8) is completed with two types of boundary conditions, as follows

$$\begin{cases} \boldsymbol{V}\cdot\boldsymbol{n} = \dfrac{g}{c}\eta \\ \boldsymbol{V}\cdot\boldsymbol{n} = 0 \end{cases}, \tag{2.9}$$

where $c=\sqrt{g\left(d+\eta\right)}$ is the wave speed and $\boldsymbol{n}$ is the unit vector normal to the boundary. The upper and lower expression in Eq. (2.9) represent open and closed boundary respectively. These set of equations can be solved numerically using various numerical schemes.

The propagation stage of a tsunami covers the most extensive area, and is presently the better understood stage that can be explained by a relatively simple theory and analysis. Detailed numerical models of tsunami propagation are available in the literatures. The finite difference methods are probably the most common method to solve the governing equation of the tsunami propagation. There are several tsunami propagation models developed based on the finite difference methods, for instance, Tohoku University's Numerical Analysis Model for Investigation of Near-field tsunamis (TUNAMI) by Imamura et al. (2006), and Cornell Multi-grid Coupled Tsunami Model (COMCOT) by Wang (2009).

In this study, COMCOT model is used to compute the synthetic waveforms for the Green's function. COMCOT is able to solve both linear and nonlinear shallow water equations in Cartesian or Spherical coordinate systems, and it also supports the nested

grid system that allows the finer grid to be placed on a coarser grid to increase local resolution. It has been widely used to study many historical tsunami events (Liu et al., 1995; Wang and Liu, 2006).

2.3 Single Time Window Inversion

A TWI typically assumes an infinite rupture velocity for relatively small rupture areas or limited temporal resolutions of tsunami waveforms. This approach considers that the water surface deformation occurs instantaneously with the earthquake. The validity of this assumption can be briefly analyzed by comparing the tsunami celerity and rupture velocity. For example, the speed of a tsunami propagated on a depth of 3000 m is approximately 0.17 km/s (using the long wave approximation). The rupture velocity of a relatively small earthquake with 100 km fault length and the source process time of 30 s is 3 km/s. Therefore, the source process is finished before the tsunami starts to propagate.

For the above mentioned case, the temporal variation due to the source process can be neglected. The inversion for the case is straightforward by means of a single time window inversion, which is the original and standard TWI as derived from Eq. (2.1) to Eq. (2.7). To easily differentiate between a single and multiple time window inversion, it is better to rewrite Eq. (2.7) in matrix form, that is

$$\begin{bmatrix} \mathbf{G}(t) \\ \gamma \mathbf{S} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_M \end{bmatrix} = \begin{bmatrix} \zeta(t) \\ 0 \end{bmatrix}. \tag{2.10}$$

2.4 Multiple Time Window Inversion

For large tsunami events, the rupture velocity should affect the generated tsunami waveforms. In such cases the rupture start time (often called the time delay) varies throughout the fault plane. In this study, the rupture velocity is considered to be comparable to the propagation speed of the water surface perturbation. Note that the water surface perturbation term used here is different from the water wave propagation, because it is considered as a direct projection of the rupture process onto water surfaces. The time delay can easily be determined by assuming a constant rupture velocity (Tanioka et al., 2006; Satake et al., 2013). A different and more complex approach was proposed by (Piatanesi and Lorito (2007), where the rupture velocity was considered as a model parameter together with the slip distribution. This results in a nonlinear inverse problem.

If the temporal resolution of the recorded tsunami waveforms is adequate, a time-dependent model is applicable to simulate the transient deformation of large tsunamis. In such cases, we can calculate multiple time window inversion by shifting the initial Green's function for a certain time span. The start times of the Green's functions (τ) should also be taken into account. With the addition of a damping constraint, the single time window in Eq. (2.10) is modified to

$$\begin{bmatrix} \mathbf{G}(t-\tau) & \mathbf{G}(t-(\tau+\Delta t)) & \cdots \mathbf{G}(t-(\tau+(K-1)\Delta t)) \\ \gamma\mathbf{S}(t-\tau) & \gamma\mathbf{S}(t-(\tau+\Delta t)) & \cdots \gamma\mathbf{S}(t-(\tau+(K-1)\Delta t)) \\ \lambda\mathbf{I}(t-\tau) & \lambda\mathbf{I}(t-(\tau+\Delta t)) & \cdots \lambda\mathbf{I}(t-(\tau+(K-1)\Delta t)) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_{MK} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\eta}(t) \\ 0 \\ 0 \end{bmatrix}, \quad (2.11)$$

where Δt is the time span, K is the number of time windows, λ is the damping parameter, and $\mathbf{I}$ is the identity matrix. The number of Green's function used in each window may vary according to the estimated rupture/perturbation front, which also

affects the size of the matrix $\mathbf{S}$ and $\mathbf{I}$. This aids the analysis of the time-dependent variations of the water surface throughout the duration of the rupture.

Chapter 3

Computational Intelligence

3.1 Various Methodologies in Computational Intelligence

Computational intelligence (CI) is a set of nature-inspired computational methodologies and approaches that can provide a practical alternative for solving intractable and complex real-world problems. The main advantages and mathematical power of the CI are commonly attributed to their ability in avoiding local optimum solutions, and the neural-like system architecture that allows to address a considerable complexity of problems that do not have an algorithmic solution. The former advantage is possessed by evolutionary computations (such as GA), and the latter refers to neural networks.

A GA is used in the first part of the study, which is to estimate the tsunami source, particularly to find optimum locations of unit sources prior to the inversion. GAs are global optimization methods that do not require the gradient of the problem to be optimized and are part of the CI. However, in some cases, GA alone may not be sufficient, because a global optimization method can explore a broad search space, but is not appropriate for fine tuning the approximation of the expected solution. Therefore, a hybrid approach based on a combination of a global and local search algorithms is

generally preferable. To that end, PS is employed as a local search algorithm to locate other nearby solutions that could possibly be better than the result of the previous search by the GA.

In the second part, an ELM is utilized to produce rapid forecasts of tsunami waveforms in coastal areas using tsunami signals recorded at specified locations. The ELM has a remarkable training speed compared to gradient-based methods such as back propagation, and produces better generalizations. Additionally, neural network models are well known for their ability to approximate most nonlinear functions demanded by practical applications. Therefore, the ELM can be used as a better alternative to the standard inversion analysis that is strictly constrained by a linear assumption.

3.2 Optimization algorithms

3.2.1 Genetic Algorithm

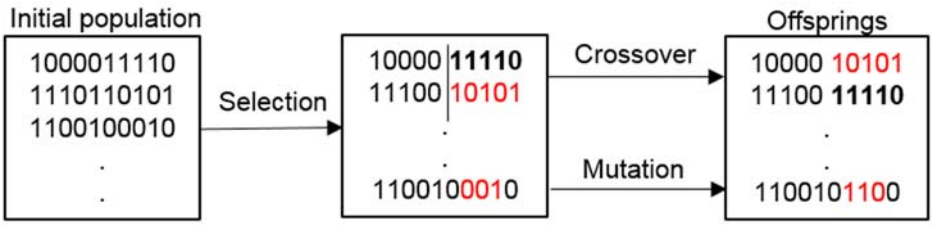

Figure 3.1 Schematic illustration of a GA.

GAs are optimization methods that search for an optimal value of a complex function by adopting the process of natural evolution (Goldberg, 1989). It can be categorized as a type of stochastic optimization method and as a part of CI. In GA, the model parameters or decision variables in the optimization are first transformed into a chromosome-like data structure that later evolves to form a better offspring. The most common representation of design parameters in GA is their encoding into a binary

string. There are three basic genetic operators in GA: selection, crossover, and mutation (Figure 3.1).

Selection process in a GA equates to survival of the fittest. Two individuals are chosen randomly based on a certain probability function corresponding to their fitness level. There are many types of selection in GA, but the most common one is probably the roulette wheels selection, also known as fitness proportionate selection. The probability of selection for individuals can be formulated as,

$$\vartheta_i = \frac{f_i}{\sum_{j=1}^{N} f_j}, \tag{3.1}$$

where ϑ is the selection function, f_i is the fitness function of individual i in the population, N is the number of individual in the population.

Crossover operator merges or swaps genetic information of two individuals to produce a new and different individual (offspring). The main purpose of the crossover is to generate a better individual, which can be achieved if it takes the best characteristics from each of the selected parents. This feature is a prime distinguished factor of GAs from other optimization techniques.

The last genetic operator is mutation. During the mutation, genes of individuals are altered from its initial state. The mutation process is able to avoid the loss of a particular piece of information, and maintain the diversity within the population. Therefore it inhibits premature convergence.

Mathematical formulations and procedures of GA is clearly described in Wetter and Wright (2003). In the GA, the finite lower and upper bounds of $\boldsymbol{X}$ are discretized to the mesh $\mathbb{M}(l,1)$, and all operators are such that the real value of any point is an element of $\mathbb{M}(l,1) \cap \boldsymbol{X}$. Given a non-empty set $\boldsymbol{X} \subset \mathbb{R}^n$, and non-zero $M \in \mathbb{N}$, $\underline{\boldsymbol{X}}_M$ is the set of all sequence $\boldsymbol{X}$ with M elements,

$$\underline{\boldsymbol{X}}_M \doteq \left\{ \{\boldsymbol{x}_i\}_{i=1}^{M} \middle| \boldsymbol{x}_i \in \boldsymbol{X},\, i \in \{1,...,M\} \right\}. \tag{3.2}$$

Let $\boldsymbol{B}$ be the set containing all elements of $\mathbb{M}(l,1) \cap \boldsymbol{X}$. Similarly, $\underline{\boldsymbol{B}}_M$ is the set of all sequences in $\boldsymbol{B}$ with M elements,

$$\underline{\boldsymbol{B}}_M \doteq \left\{ \{\boldsymbol{\chi}_i\}_{i=1}^{M} \middle| \boldsymbol{\chi}_i \in \boldsymbol{B},\, i \in \{1,...,M\} \right\}. \tag{3.3}$$

Denoting $k \in \mathbb{N}$ as the generation number, $M \in \mathbb{N}$ as the population size, $\underline{\boldsymbol{x}}_k \in \underline{\boldsymbol{X}}_M$ as the M points in the k-th generation, $\underline{\boldsymbol{\chi}}_k \in \underline{\boldsymbol{B}}_M$ as the binary representation of $\underline{\boldsymbol{x}}_k$, and $\underline{\chi}_{k,i}$ denotes the i-th element of $\underline{\boldsymbol{\chi}}_k$ for $i \in \{1,...,M\}$, which is a binary representation of a point in $\mathbb{R}^n$, the step-by-step operation of the GA is described in the following:

Step 1: Randomize the initial population $\underline{\boldsymbol{\chi}}_o \in \underline{\boldsymbol{B}}_M$ of M randomly generated points. Then, evaluate all $\underline{\boldsymbol{\chi}}_k$ according to the specified fitness function; thus, a fitness function $\Theta: \mathbb{N} \times \underline{\boldsymbol{B}}_M \to \mathbb{R}^M$ computes a fitness value of $\underline{\boldsymbol{\chi}}_k$.

Step 2: Select a pair of individuals with the selection function $\vartheta: \mathbb{N} \times \underline{\boldsymbol{B}}_M \to \underline{\boldsymbol{B}}_2$. Individuals with better fitness values have a higher probability of being selected.

Step 3: Perform crossover by swaping the bits (genes) between the selected individuals (chromosomes) to produce a new offspring, $\phi: \mathbb{N} \times \underline{\boldsymbol{B}}_2 \times [0,1] \to \underline{\boldsymbol{B}}_2$ with the crossover probability $p_r \in [0,1]$.

Step 4: The final genetic operator is the mutation that aims to maintain genetic diversity. The mutation function $\psi: \mathbb{N} \times \boldsymbol{B} \times [0,1] \to \boldsymbol{B}$ alters a bit on individuals from its initial state.

The algorithm stops when one of the stopping criteria is met. These stopping criteria can be based on the predefined maximum number of generation or the average relative change in the fitness function value over a certain generation.

3.2.2 Patterns Search Method

PS is also a derivative-free optimization method, although it may not be categorized as part of the CI. It was first introduced by Hooke and Jeeves (1961). Later, Torczon (1997) proved the convergence of PS using the theory of positive bases. The algorithm of PS used in this study is similar to that of Wetter and Write (2003), while the design parameters are identical to the GA optimization, because the output of the GA optimization result is treated as the initial condition for the PS optimization.

For the same optimization problem as solved by the GA, the algorithm searches a lower cost function value than $f(\boldsymbol{x}_k)$, where $\boldsymbol{x}_k \in \boldsymbol{X}$ denotes a solution at the current iterate and $k \in \mathbb{N}$ denotes the iteration number. The search takes place on the points in the set

$$\gamma_k \doteq \left\{ \boldsymbol{x} \in \boldsymbol{X} \middle| \boldsymbol{x} = \boldsymbol{x}_k \pm \Delta_k s^i e_i,\ i \in \{1,...,n\} \right\}, \tag{3.4}$$

where $\Delta_k > 0$ is the mesh size factor, $s \in \mathbb{R}^n$ is fixed and scales the parameters, and e is the unknown approximation error. The rule to select a finite number of points in $\boldsymbol{X}$ on a mesh can be defined by:

$$\mathbb{M}(\boldsymbol{x}_0, \Delta_k) \doteq \left\{ \boldsymbol{x} + m\Delta_k s^i e_i \middle| i \in \{1,...,n\}, m \in \mathbb{Z} \right\}, \tag{3.5}$$

where $\boldsymbol{x}_0 \in \boldsymbol{X}$ is the initial iterate, and $\mathbb{Z}$ is the set of nonnegative integers.

The overall procedures of PS optimization can be elaborated as follows:

Step 1: Initialize $\boldsymbol{x}_0 \in \boldsymbol{X}$ and $\Delta_0 > 0$, for which in our case, $\boldsymbol{x}_0$ is obtained from the GA's output.

Step 2: If $f(\boldsymbol{x}') < f(\boldsymbol{x}_k)$ for some $\boldsymbol{x}' \in \mathbb{M}(\boldsymbol{x}_0, \Delta_k)$, then set $\boldsymbol{x}_{k+1} = \boldsymbol{x}'$ and $\Delta_{k+1} = \Delta_k$.

Step 3: If $f(\boldsymbol{x}') \geq f(\boldsymbol{x}_k)$ for all $\boldsymbol{x}' \in \gamma_k$, the search continues with $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k$ and reduced mesh factor $\Delta_{k+1} = \frac{\Delta_k}{2}$.

The search continues until a stopping condition is satisfied (e.g., the mesh size or changes are less than a given tolerance). Thus, the smallest $\boldsymbol{x}''$ satisfies $f(\boldsymbol{x}'') \leq f(\boldsymbol{x}')$ for all $\boldsymbol{x}' \doteq \left\{ \boldsymbol{x} \in \boldsymbol{X} \middle| \boldsymbol{x} = \boldsymbol{x}'' \pm \Delta_{min} s^i \boldsymbol{e}_i,\ i \in \{1,...,n\} \right\}$, where Δ_{min} is the smallest mesh size factor.

3.3 Extreme Learning Machine

A completed ELM method was proposed by Huang et al. (2006). It has received a growing amount of attention from scientific and engineering communities because of its remarkably fast training speed, when compared to classical training algorithms. The ELM can be trained up to 170 times faster than gradient-based methods such as back propagation, and produces better generalizations. In this study, computing speed is very important because it enables immediate forecasts of tsunami waveforms. However, unlike conventional neural networks, ELM may require more hidden neurons because of the random determination of the input weights and biases (Zhu et al., 2005).

Using the same variables as in TWI, the training set for the ELM is

$$L = \left\{ (\boldsymbol{\eta}_i, \boldsymbol{\zeta}_i) \middle| \boldsymbol{\eta}_i \in \square^{M}, \boldsymbol{\zeta}_i \in \square^{m}, i = 1, 2, ..., N \right\}, \tag{3.6}$$

where $\boldsymbol{\eta}_i = (\eta_{i1}, \eta_{i2}, ..., \eta_{iM})^T$ is the synthetic wave originating from the M unit sources used to construct the Green's function in the TWI, and m is the number of output neurons (in this study, there is only one node). An ELM with $\tilde{N}$ hidden neurons and activation function g can be analytically modeled as

$$\sum_{j=1}^{\tilde{N}} \beta_j g\left(\tilde{\boldsymbol{w}}_j \cdot \boldsymbol{\eta}_i + b_j \right) = \zeta_i, \quad i = 1, ..., N, \tag{3.7}$$

where $\tilde{\boldsymbol{w}}_j = \left(\tilde{w}_{j1}, \tilde{w}_{j2}, ..., \tilde{w}_{jM}\right)^T \in \mathbb{R}^M$ is the weight vector connecting the input layer to the *j*-th hidden neuron (we modified the symbol to distinguish between TWI weights and ELM input weights), b_j is the bias of the *j*-th hidden neuron, and β_j is the output weight connecting the *j*-th hidden neuron to the output layer. A sigmoid transfer function

$$g(x) = 1/\left(1+\exp(-x)\right), \tag{3.8}$$

is used in the hidden layer, which introduces nonlinearities to the network.

Rewriting Eq. (3.7) in matrix form we get

$$\boldsymbol{\beta}^T \mathbf{H} = \boldsymbol{\zeta}, \tag{3.9}$$

where

$$\mathbf{H} = \begin{bmatrix} g(\tilde{\boldsymbol{w}}_1 \cdot \boldsymbol{\eta}_1 + b_1) & \cdots & g(\tilde{\boldsymbol{w}}_1 \cdot \boldsymbol{\eta}_N + b_1) \\ g(\tilde{\boldsymbol{w}}_2 \cdot \boldsymbol{\eta}_1 + b_2) & \cdots & g(\tilde{\boldsymbol{w}}_2 \cdot \boldsymbol{\eta}_N + b_2) \\ \vdots & \ddots & \vdots \\ g(\tilde{\boldsymbol{w}}_{\tilde{N}} \cdot \boldsymbol{\eta}_1 + b_{\tilde{N}}) & \cdots & g(\tilde{\boldsymbol{w}}_{\tilde{N}} \cdot \boldsymbol{\eta}_N + b_{\tilde{N}}) \end{bmatrix}_{\tilde{N} \times N}, \tag{3.10}$$

and

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1^T \\ \beta_2^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{bmatrix}, \text{ and } \boldsymbol{\zeta} = \begin{bmatrix} \zeta_1 & \zeta_2 & \cdots & \zeta_N \end{bmatrix}. \tag{3.11}$$

Because $\tilde{N} < N$, the output weights can be calculated using the least squares solution of Eq. (3.9), that is,

$$\tilde{\boldsymbol{\beta}} = \mathbf{H}^{\dagger} \boldsymbol{\zeta}, \tag{3.12}$$

where $\mathbf{H}^{\dagger}$ is the Moore-Penrose generalized inverse matrix of $\mathbf{H}$ (Rao and Mitra, 1971). If $\mathbf{H}$ is nonsingular, Eq. (3.12) can be written as

$$\tilde{\boldsymbol{\beta}} = \left(\mathbf{H}\mathbf{H}^T\right)^{-1}\mathbf{H}^T\boldsymbol{\zeta}. \tag{3.13}$$

We can see that the output weights of the ELM are very similar to the unknown parameters in the TWI (Eq. (2.5)). The difference between the two methods is that ELM incorporates input weights and biases, and a nonlinear transfer function.

The ELM requires an additional regularization to suppress erratic small-scale variations that are probably induced by the complex network structure. Therefore the graph regularization (Peng et al., 2015) is used in addition to the Tikhonov's regularization scheme. The final objective function that must be solved is

$$\min_{\boldsymbol{\beta}}\left[\left\|\boldsymbol{\beta}^T\mathbf{H}-\boldsymbol{\zeta}\right\|^2+\mu\mathrm{Tr}\left(\boldsymbol{\beta}^T\mathbf{H}\mathbf{L}\mathbf{H}\boldsymbol{\beta}\right)+\gamma\left\|\boldsymbol{\beta}\right\|^2\right], \tag{3.14}$$

where $\mathrm{Tr}\left(\boldsymbol{\beta}^T\mathbf{H}\mathbf{L}\mathbf{H}\boldsymbol{\beta}\right)$ is the graph regularization term, $\mathbf{L}$ is the graph Laplacian, and μ is the additional regularization parameter (μ =1 in this paper). The solution to Eq. (3.14) is

$$\boldsymbol{\beta} = \left(\mathbf{H}\mathbf{H}^T+\mu\mathbf{H}\mathbf{L}\mathbf{H}^T+\gamma\mathbf{I}\right)^{-1}\mathbf{H}^T\boldsymbol{\zeta}. \tag{3.15}$$

Details on the mathematical derivation on the graph regularization can be found in Smola and Kondor (2003).

Chapter 4

Tsunami Source Estimation in an Artificial (Ideal) Case

4.1 Problem Description

Recently, several studies using tsunami waveform inversion to estimate the tsunami source without fault model assumptions have been widely developed. For instances, Baba et al. (2005) used a simplified fault model by disregarding actual earthquake parameters to produce the initial profile on each unit source, whereas Satake et al. (2005) proposed a more direct approximation using a pyramidal shape with a flat top. Other studies by Liu and Wang (2008) and Saito et al. (2010) demonstrated attempts to use Gaussian function, whereas Wu and Ho (2011) adopted a top-hat small unit source to represent the initial profile. The same approach was proposed by Tsushima et al. (2011) and Yasuda and Mase (2013) for the more practical purpose of a tsunami early warning system. However, this inversion method for tsunami waveforms possesses a limitation, in that the inverse matrix does not always exist because of the non-uniqueness of the solution. In addition to the large number of unknown parameters, which might produce many local optima on the misfit function

measure, the search towards optimality is confined by the uniform distance of unit sources used in the regular Green's function.

TWI sometimes falls into an ill-posed problem, in which small errors in the observed waveforms are exceptionally amplified in the solution. Therefore, both the uniqueness and the stability of solutions are sometimes difficult to attain without appropriate treatments. The most frequently employed techniques to maintain a stable solution is to use a smoothing constraint (Saito et al., 2010; Gusman et al., 2013). Other than that, Koike et al. (2003) suggested reducing the unknown parameters using the wavelet base to guarantee the uniqueness of the solution. However, they found later that the selection of the wavelet base was not straightforward. Another effort to overcome the issue was discussed by Voronina (2011). The study promoted a method to control numerical stability for the ill-posed problem in tsunami waveform inversion by means of singular value decomposition and r-solutions techniques.

This study proposes a new approach to tackle the same problem by determining the optimal position or spatial distribution of unit sources located around the source area or epicenter. A GA as a global optimization method, combined with a PS method, is employed to search the mentioned positions prior to the inversion. As the selected positions are probably located in between the initial unit sources, interpolations are performed during the optimization. Therefore, the Green's function evolves dynamically at each generation of the GA and PS iteration.

The ultimate purpose of a global optimization method is to find the extreme value of a given non-convex function in a certain feasible region. Following the growth of computer science, new types of optimization based on CI have been developed extensively and used by scientific and engineering communities. The reason for this is that the new optimization methods possess the interesting feature of being able to avoid local optimum solutions, which is something classical methods fail to do.

4.2 Numerical Experiments

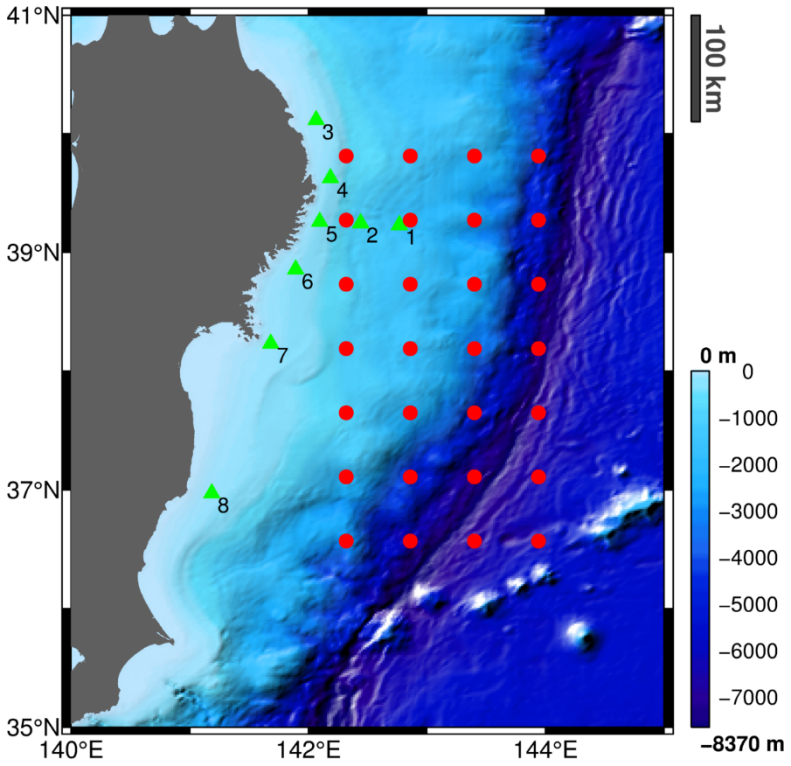

Figure 4.1 Study area and bathymetry profile. Red dots indicate unit sources located throughout the inverse region. Green triangles with numbers are artificial observation stations.

Numerical experiments using an artificial tsunami source propagated on an actual bathymetry profile have been conducted. An area extending from 140–145°E and 35–41°N is chosen as the domain of interest (Figure 4.1). The selected domain resembles that used in most studies of tsunami waveform inversion for the 2011 Tohoku tsunami. The resolution of the numerical model is 1 arc minute, which is consistent with the resolution of the bathymetry data obtained from the ETOPO1. ETOPO1 is a global relief model of the earth's surface, which includes ocean bathymetry, available from the National Geophysical Data Center of the National Oceanic and Atmospheric Administration (Amante and Eakins, 2009). There are eight artificial observation

stations associated with the actual location of gauges within the study area. The tsunami source is divided into 28 unit sources that are distributed uniformly around the actual epicenter of the 2011 Tohoku tsunami (Figure 4.1).

4.2.1 Design Parameters

The use of global optimization methods in tsunami waveform inversion is not new, relevant discussions can be found in Piatanesi and Lorito (2007) and Romano et al. (2010). They used a simulated annealing technique to solve the inverse problems. This study proposes a different algorithm based on a hybrid optimization of GA and PS. The hybrid technique is preferred because global optimization methods, such as GA, are capable of exploring broader search space, but not as good in fine tuning the approximation of the expected solution. Therefore, PS is employed as a local search algorithm to locate other nearby solutions that could possibly be better than the result of the previous search by GA (Payne and Eppstein, 2005; Costa et al., 2010).

The hybrid algorithm proposed in this study works by simply treating the output of the GA optimization result as the initial condition for the PS optimization. The technique is proven effective even though more fitness function evaluation is required; hence, it costs extra computational efforts. However, parallelization of either GA or PS can be easily implemented to expedite the computing time and gain substantial performance enhancement.

The formulation of an optimization problem can be expressed as

$$\min_{\boldsymbol{x}\in\boldsymbol{X}} f(\boldsymbol{x}), \tag{4.1}$$

where $\boldsymbol{x} \in \boldsymbol{X}$ is the vector of parameters, and $f : \boldsymbol{X} \to \mathbb{R}$ is the cost function. $\boldsymbol{X} \subset \mathbb{R}^n$ is the constraint set or bounds defined as

$$\boldsymbol{X} \doteq \left\{ \boldsymbol{x} \in \mathbb{R}^n \middle| l^i \le x^i \le u^i,\ i \in \{1,...,n\} \right\}, \tag{4.2}$$

where $-\infty \le l^i \le u^i \le \infty$, for all $i \in \{1,...,n\}$, and l and u are the lower and upper bounds.

In this study, the proposed algorithm is developed with two different design parameters. First, is simply to search the water elevation of each unit source without including a search of the optimum locations, thus similar to that of the ordinary least squares method. This experiment aims to gauge the significance of the global optimization method applied to a linear system compared to the conventional method using least squares. Second, is to search for the optimal spatial distribution of unit sources, which is the main topic of this study.

The TWI is basically an optimization problem. Therefore, by rearranging the Eq. (2.2), the first design parameter can be expressed as an optimization problem as follow:

$$\min_{\boldsymbol{x} \in \boldsymbol{X}} f(\boldsymbol{x}) = \boldsymbol{\zeta} - \boldsymbol{x}\mathbf{G}, \tag{4.3}$$

where $\boldsymbol{x}$ is the unknown parameters $\boldsymbol{w}$ in the Eq. (2.2) and $\boldsymbol{X}$ is the constraint or bounds representing plausible values of subsidence and uplift of the water surface acted as lower and upper bounds respectively.

The purpose of the second design parameters is to search the optimal location of unit sources, while the initial water heights are calculated using the least squares inversion. The formulation of the optimization problem for the second design parameters can be expressed as

$$\min_{\boldsymbol{x} \in \boldsymbol{X}} f(\boldsymbol{x}) = \boldsymbol{\zeta} - \boldsymbol{w}\mathbf{G}(\boldsymbol{x}), \tag{4.4}$$

where $\boldsymbol{w}$ is the calculated water elevations using least squares without a regularization, Eq.(2.5), and $\boldsymbol{x}$ define a unique identification for the computational grid of the forward model located inside the inverse region $\boldsymbol{X}$ (Figure 4.2). In this part, the combination of the algorithm with the least squares method is necessary to save the computational efforts.

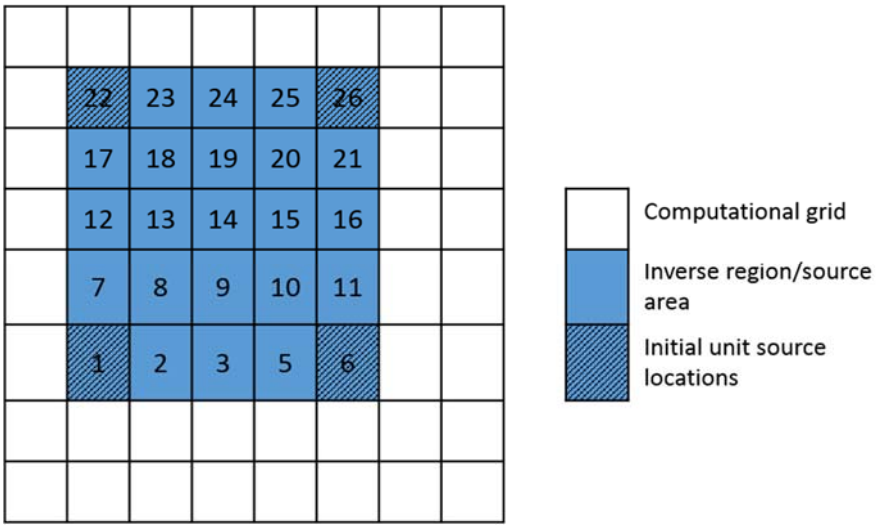

Figure 4.2 Example of a unique identification number of the forward model computational grid used for the second design parameters.

4.2.2 Fitness Function

The selection of a fitness function or cost function is essential because it directs the fate of the optimization towards the optimal solution. Here, a combination of root mean square error (RMSE) and Pearson correlation coefficient (r) is used. While the RMSE is sensitive to amplitude matching, the correlation coefficient is more sensitive to phasing between the compared series (Barnston, 1992).

The RMSE serves to aggregate the individual differences of data points into a single measure of predictive power, which is defined as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(d_i - y_i\right)^2}\ , \tag{4.5}$$

where y is the predicted value by the model, d is the measurement data for each *i*-th data point, or in our case, the waveform generated by the artificial tsunami source, and n is the total number of data.

The Pearson correlation coefficient is defined as a division of the covariance of the two variables by the product of their standard deviations, that is

$$r = \frac{\sum_{i=1}^{n}\left(d_i - \bar{d}\right)\cdot\left(y_i - \bar{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(d_i - \bar{d}\right)^2 \cdot \sum_{i=1}^{n}\left(y - \bar{y}\right)^2}}, \tag{4.6}$$

where $\bar{d}$ and $\bar{y}$ represent the means of d and y, respectively. The correlation or relevance of the data is measured using the value $R = 0.5\left(r+1\right)$ to avoid negative values in the case of a decreasing linear relationship between the compared series.

The fitness evaluation is subject to noise from various factors that might lead the optimization towards unexpected solutions. The easiest technique to overcome such problems is by means of explicit averaging over a number of samples to smooth the cost function (Jin and Branke, 2005). As the best solution or closest fit is indicated by $RMSE \to 0$ and $R \to 1$, the final cost function is a summation of the mean of the *t*-th sample over the time window T of Eqs. (4.5) and (4.6), which can be written as

$$E = \sum_{k=1}^{N}\left[\frac{1}{T}\sum_{t=1}^{T} RMSE_t + \left(1 - R_t\right)\right]_k, \tag{4.7}$$

where k denotes the respective time window and N is the total number of windows.

4.2.3 Auxiliary Basis Function

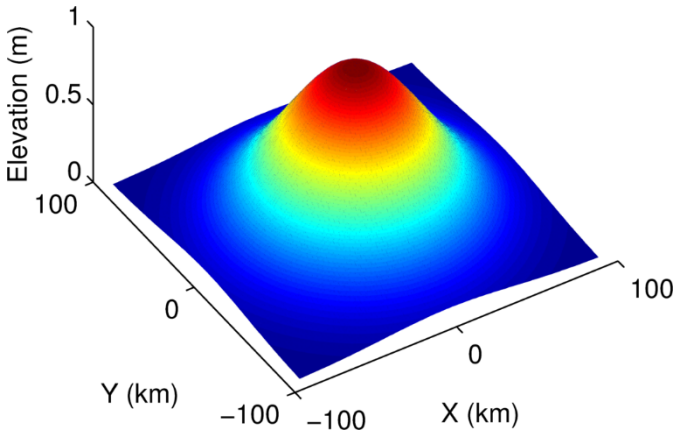

Figure 4.3 Gaussian basis function.

A Gaussian shape with 1-m amplitude is used as the basis function for each unit source (Liu and Wang, 2008). Providing A_i as the amplitude of a unit source with the centroid positions of x_i and y_i, the basis function can be written as

$$z_i(x,y) = A_i \exp\left[-\frac{(x-x_i)^2+(y-y_i)^2}{2L^2}\right], \tag{4.8}$$

where $z_i(x,y)$ is the initial water surface corresponding to the *i*-th unit source, *x* and *y* are the locations of the computational grid points, *L* is the spread of the basis function with a length of 40 km. The specified length should satisfy the long wave assumption, where the wavelength should be greater than 20 times the average water depth. An example of the Gaussian profile used as the auxiliary basis function to replace the fault model is shown by Figure 4.3.

4.2.4 Waveform Interpolation

Since the selected positions are likely located in between the initial unit sources, interpolations are performed as a substitute for the numerical forward model to produce the synthetic waveforms. This is necessary because performing numerical computation during the optimization is too time consuming. The waveform interpolation has two stages. First is interpolation of the arrival time, and then followed by waveform amplitude interpolation. In the first stage, the arrival time of the selected unit sources are interpolated based on the four nearest unit sources in the initial Green's function. A nearest neighbor weighted interpolation is used to estimate the properties of interest. The general form of the method is given by

$$P = \begin{cases} \dfrac{\sum_{i=1}^{N} \dfrac{P_i}{\sqrt{(x_i - x)^2 + (y_i - y)^2}}}{\sum_{i=1}^{N} \dfrac{1}{\sqrt{(x_i - x)^2 + (y_i - y)^2}}} & x_i \neq x \; or \; y_i \neq y \\ P_i & x_i = x \; and \; y_i = y \end{cases}, \tag{4.9}$$

where P is the property of interest (either the arrival time or wave amplitude), i is the index of the nearest data points (which has a total of N=4), and x and y represent the position of the unit source.

After computing the arrival time of the selected unit source location, the waveforms of the four nearest unit sources are shifted according to the estimated arrival time. As a result, the waveforms of the four nearest unit sources then have the same phase. Subsequently, the wave amplitude is interpolated using the same algorithm. The phase-shifting process was necessary because it is difficult to perform a direct interpolation of waveforms with different phases.

4.3 Model Development

For the first design parameters, the optimization is performed merely to search the water elevation or initial amplitudes of each unit source. For this case, the Green's function is constructed based on the initial 28 unit sources, separated by a uniform distance of 60 km, which is identical to that used in the least squares inversion. Hereafter, the first model will be termed the Genetic Algorithm Pattern Search for uniform source distribution (GAPSu). The purpose of this model is simply to compare the performance of the proposed global optimization method with the traditional least squares method in the same model design and environment.

The second design parameters aim to find the optimal locations of unit sources that, at the initial state, are distributed randomly around the tsunami source. The second model will be termed the Genetic Algorithm Pattern Search for random source distribution (GAPSr). In the GAPSr model, the amplitudes are computed using the least squares inversion and therefore, the second model is actually a combination of a deterministic and stochastic optimization. This approach increases the flexibility of the Green's function to better approximate any complex water surface profiles. However, interpolations are required to produce the synthetic waveform at the selected locations.

An example of the interpolation results, complete with statistical evaluations in terms of RMSE and r, is shown in Figure 4.4. Despite the satisfying results, based on the measure of fitness shown by the interpolation method (overall RMSE < 0.048 m and $r > 0.966$), the small errors might be amplified in the solution because of the ill-posed problem. Consequently, further improvements should be made to suppress the generated errors.

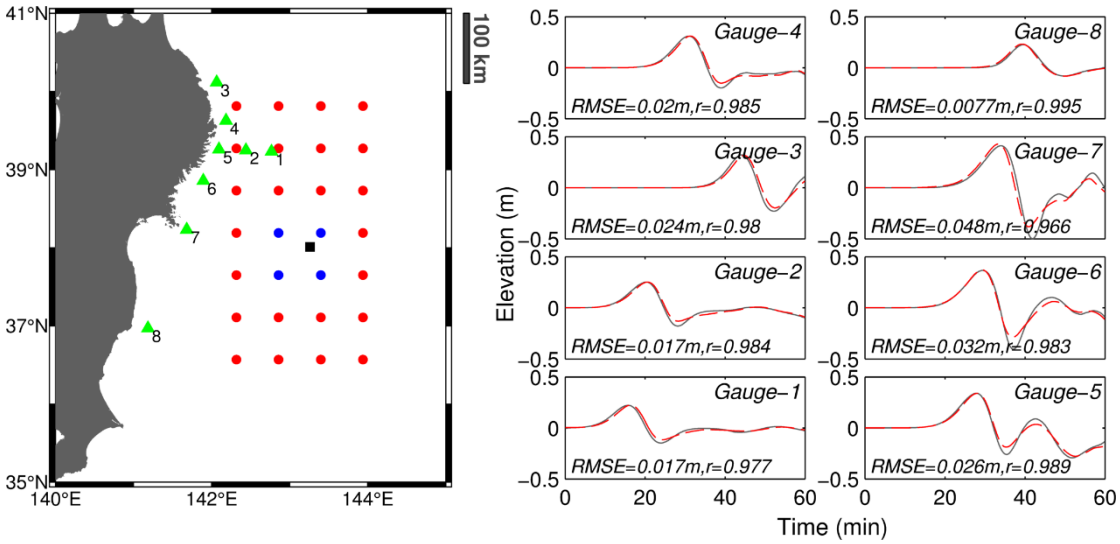

Figure 4.4 Waveform interpolation. Left-hand figure shows selected location of a unit source indicated by a black square. Blue dots represent the four nearest unit sources used in the interpolation. Right-hand figures are comparisons of waveforms between numerical model (solid black line) and interpolation (dashed red line) at the artificial gauges.

An artificial tsunami source is used to test the method (Figure 4.6a). Instead of using a simpler profile produced by the Okada's solution, a more complex shape is generated from a superposition of 10 unit sources with random amplitudes and positions located inside the inverse region. The number of unit source is limited to avoid generating shorter wavelengths than the prescribed long wave assumption. For the regular Green's function (first design parameter), using only 10 unit sources is insufficient to reconstruct the target profile. Consequently, the number of unit source should be increased. The decision of using 28 unit sources in both design parameters is for the purpose of model performance comparisons that will be further discussed in the next section. The approximation of this initial profile is performed using unconstrained, traditional least squares inversion, GAPSu, and GAPSr. However, GAPSr is the most important part in this study and therefore, the discussion will be focused on the GAPSr model.

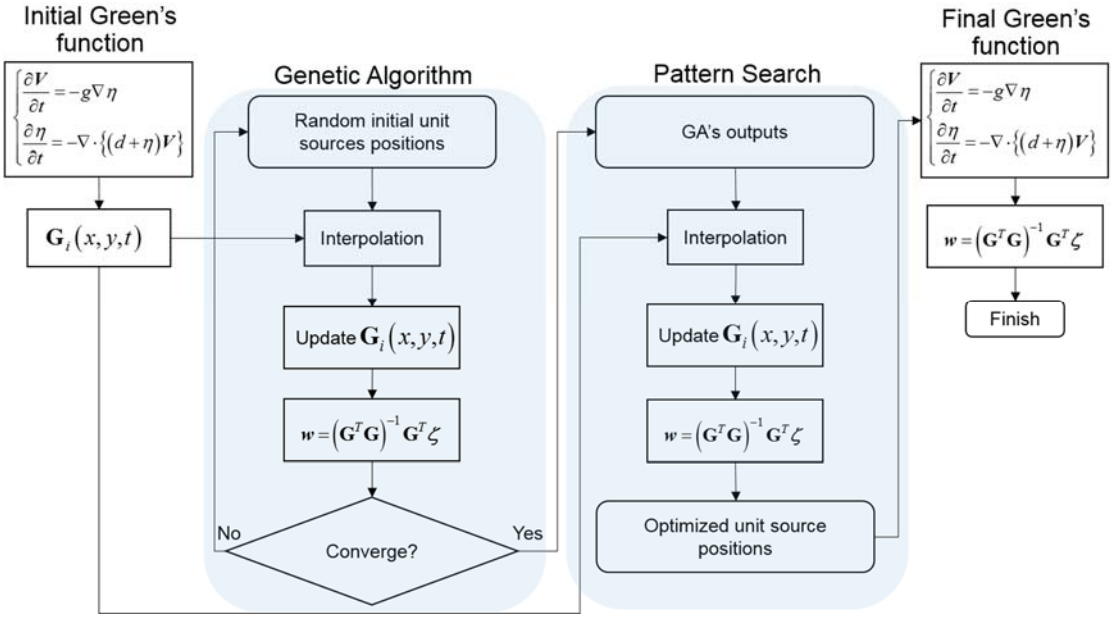

Figure 4.5 Flowchart of model development procedure.

The development of the GAPSr model depicted in the flowchart (Figure 4.5) can be summarized as follows:

Step 1: Construct the initial Green's function.

Step 2: Initialize the GAPSr model by randomly distributing the unit source locations. The search of the optimal location is bounded by the area of the inverse region.

Step 3: Perform interpolation and update the Green's function.

Step 4: Evaluate the fitness by performing the least squares inversion.

Step 5: After reaching the stopping criteria, the forward numerical modeling is run again for each of the optimized unit sources to avoid errors generated from the interpolation result. Subsequently, the inversion is performed for the final time.

4.4 Results and Discussion

Overall, all models can produce a relatively good estimation of the targeted tsunami source. This is likely because they are applied to an ideal case with artificial conditions and settings, except for bathymetry. For instance, the target source was generated from the same Gaussian shape as that used to construct the Green's function. Therefore, the task is more straightforward as less complexities are encountered. Nevertheless, in this study, the use of the ideal case has allowed us to assess the advantage of the proposed method in a more detailed manner.

The GAPSu model yields a slightly better fit of waveforms compared with the least squares method (Table 4.1). This means that the global optimization method locates a better minimum value in the cost function, which is situated beyond the reach of the least squares method. However, the slight refinement by the GAPSu model over the least squares method makes it difficult to gauge the benefits of employing the method. This should not be a surprise because the model is applied to determine the coefficients in a linear system, which is relatively easy to solve using a conventional method. Moreover, the waveforms used to invert the initial tsunami source are generated from an artificial tsunami source instead of real measurements. Therefore, the linearity is well conserved and thus, the use of more advanced methods becomes redundant and unnecessary. In the study by Piatanesi and Lorito (2007), a global optimization method was successfully promoted for the case of tsunami waveform inversion. This was because the optimization method was applied to a nonlinear inverse problem of actual measurement data. Accordingly, the appraisal of such a method can be clearly defined.

Table 4.1 Summary of statistical evaluations based of the proposed cost function.

Station	Least squares		GAPSu		GAPSr	
	RMSE (meter)	***r***	**RMSE (meter)**	***r***	**RMSE (meter)**	***r***
Gauge 1	0.0260	0.9973	0.0256	0.9974	0.0094	0.9996
Gauge 2	0.0375	0.9876	0.0375	0.9877	0.0099	0.9992
Gauge 3	0.0822	0.9744	0.0792	0.9761	0.0250	0.9976
Gauge 4	0.0646	0.9809	0.0643	0.9811	0.0203	0.9982
Gauge 5	0.0600	0.9829	0.0591	0.9835	0.0180	0.9984
Gauge 6	0.0659	0.9854	0.0653	0.9857	0.0241	0.9981
Gauge 7	0.0937	0.9954	0.0940	0.9954	0.0417	0.9992
Gauge 8	0.0354	0.9879	0.0346	0.9885	0.0114	0.9987

Spurious uplifts and subsidence of the water surface profile are generated in both the least squares and GAPSu model results (see Figure 4.6b and Figure 4.6c). One may argue that the specified spatial resolution of the unit sources is too coarse to represent the complete form of the target source. A denser distribution of unit sources should improve the results; however, it might also introduce other problems. A large number of model parameters (unknown parameters), which are proportional to the degrees of freedom in the optimization, are liable to cause the solution to become easily entrapped in a local optimum. Without a smoothing constraint, the result of the tsunami waveform inversion might be bumpy and non-physical, especially for cases with high spatial resolution (Wu and Ho, 2011). In other studies on tsunami waveform inversion by Baba et al. (2005) and Wu and Ho (2011), an equality constraint was imposed to maintain the smoothness of the inverted parameters to satisfy the long wave assumption, while by Saito et al. (2010) the constraint was used to obtain stable

solutions. However, such constraint might restrict the exploration throughout the feasible search space and render the discovery of an optimum solution more difficult. Another plausible explanation for the unsatisfactory results of the least squares and GAPSu model is simply that the equidistant unit sources in the regular Green's function confines the search for optimality. This can be proven by the result of the GAPSr model, for which the same number of model parameter (28 unit sources) with random and scattered locations yields a much better estimation.

The different design parameters in the GAPSr model have considerably improved the inversion accuracy. For instance, at Gauge 1, where the best fit of the waveform is attained, the measurement of accuracy as RMSE is 0.0260, 0.0256, and 0.0094 m, and as r is 0.9973, 0.9974, 0.9996, for the least squares, GAPSu, and GAPSr models, respectively (Table 4.1). For further qualitative or visual assessments, comparisons of time series of the waveforms at each gauge are presented in Figure 4.7. In addition, statistical evaluation results for the waveforms at all gauges are shown by scatter plots in Figure 4.8. Overall, the statistical analyses on the waveforms suggest that the GAPSr model shows very good agreement with the target. These results conform to the inverted tsunami source resulted by the GAPSr model, which can be seen in Figure 4.6. The random and scattered location of the unit sources allows the approximation to capture the exact profile of the target source. Other than this, despite no smoothing constraint being used, the inverted sea surface deformation remains smooth and coherent.

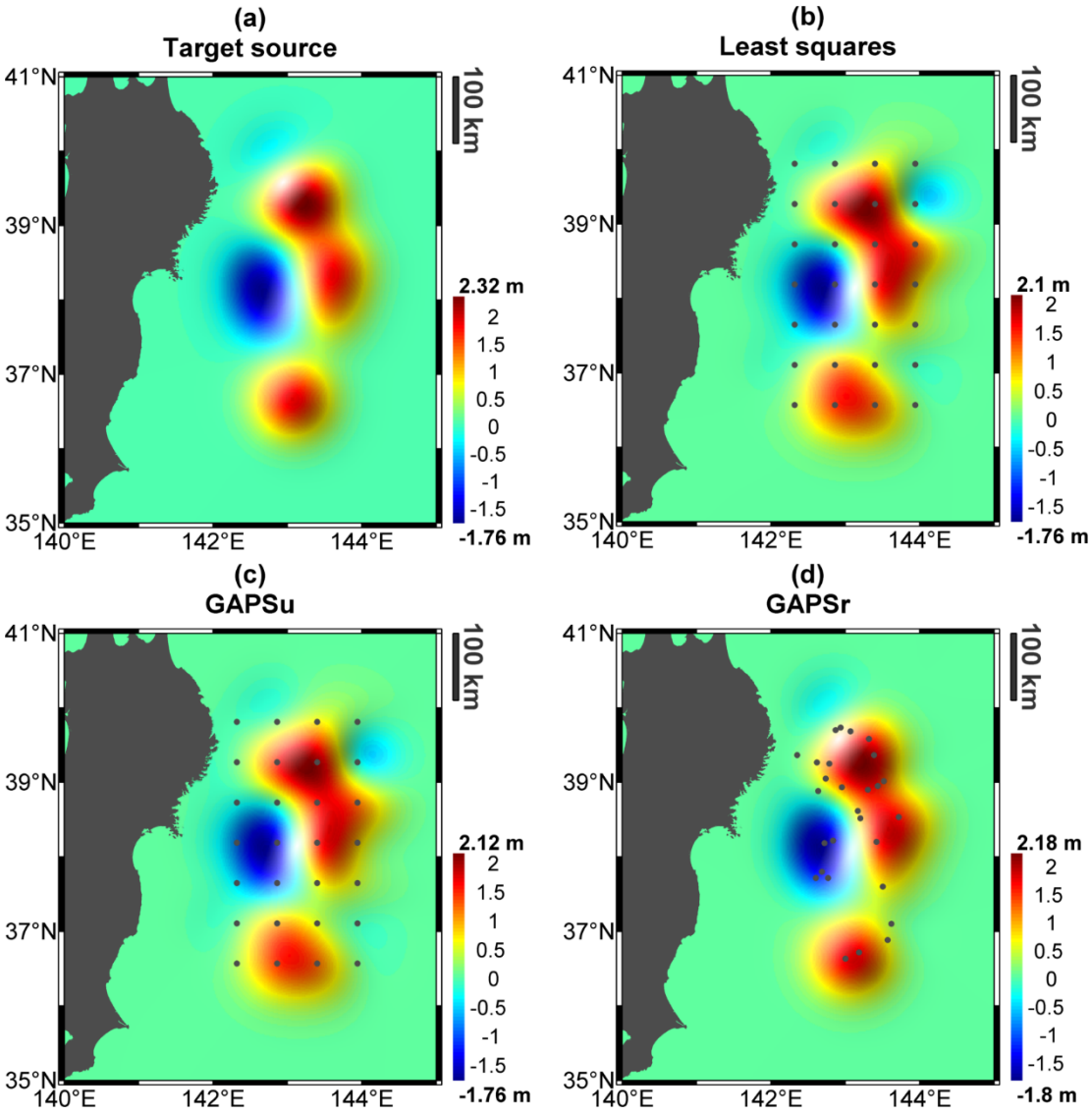

Figure 4.6 Tsunami initial wave sources. Gray dots indicate the centroid of unit sources. (a) Artificial tsunami source as the target to be approximated, (b) inverted source using least squares method, (c) GAPSu model, (d) GAPSr model.

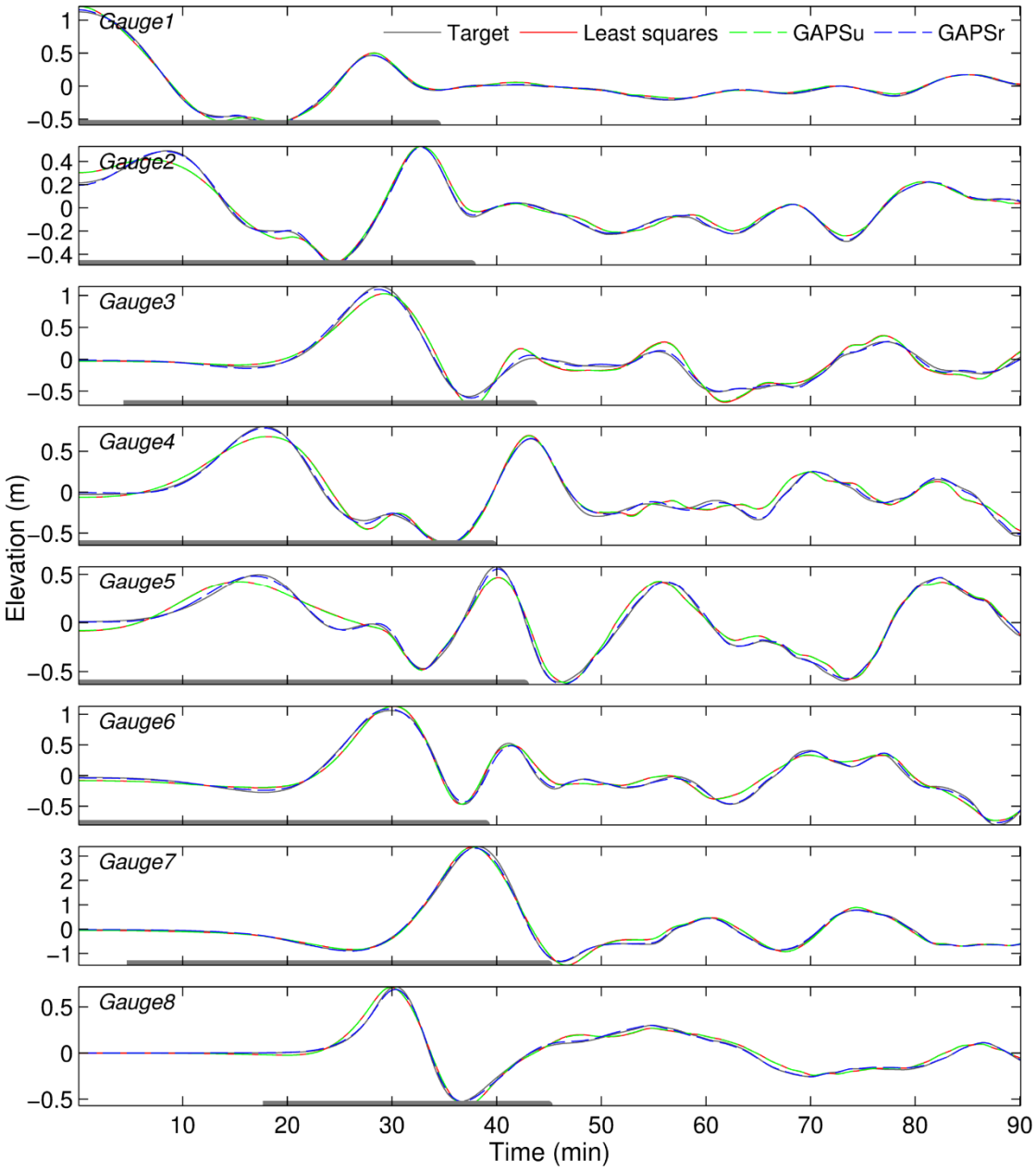

Figure 4.7 Comparison of waveforms at Gauges. Gray bar above the time axis indicates the time range for the inversion.

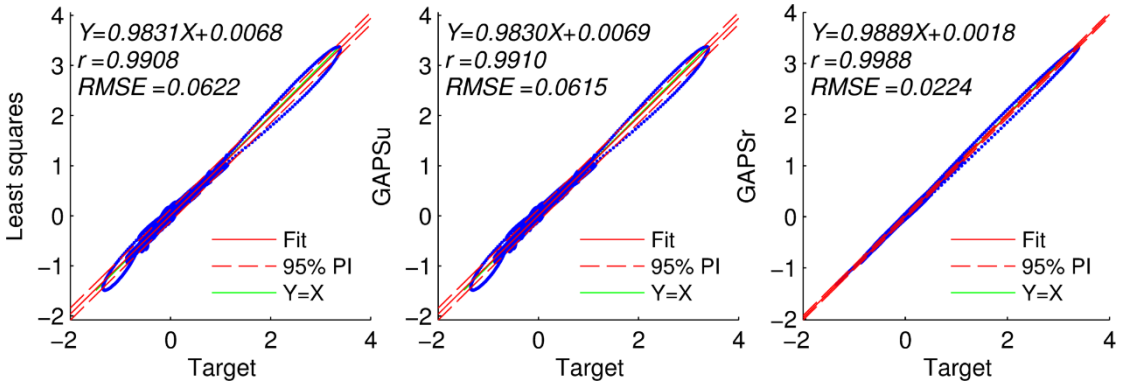

Figure 4.8 Scatter plots of each method with respect to the waveforms of the target source at all gauges.

The search for the optimal locations of the unit sources allows the least squares method to find the unique and optimal solution. Such an approach is difficult to achieve deterministically using conventional gradient methods, because there is the possibility that the constructed design parameters in the GAPSr model produce a discontinuous or non-differentiable error surface owing to the random characteristics exhibited by the artificial tsunami source (target source). The same characteristic is very likely to occur in nature. A high degree of uncertainty has been observed in tsunami sources leading to significant variations in the nearshore tsunami amplitude (Geist, 2005). Therefore, the use of the model in real case applications is encouraged to reveal the underlying dynamics in tsunami generation. However, as with typical stochastic methods, the proposed model cannot ensure a constant optimization response time. The solution and convergence are strongly dependent on the random initial state.

Chapter 5

Tsunami Source Estimation in a Real Case – the 2011 Tohoku Tsunami

5.1 Problem Description

Reconstructing the tsunami source caused by an underwater earthquake is the important first stage that determines the fate of subsequent tsunami models. Combined with other prominent factors that contribute to the tsunami's behavior, the initial wave source defines the characteristics of the tsunami. It is particularly important in near-field events, where the tsunami impact is more dependent on the source geometry (Okal and Synolakis, 2004). The initial tsunami source can be considered a hydrodynamic response to the perturbed seafloor caused by the earthquake. The response defines the source characteristic and can be revealed from the recorded tsunami waveforms.

This chapter discusses the application of the proposed method to the 2011 Tohoku tsunami, which uses near-field recorded tsunami waveforms that contain important information regarding the characteristics of the initial tsunami-generated wave profile. The waveforms are first used to identify the source area using a backward-

propagated travel time, and are later used for the inversion. Here, GA-PS algorithm is used to determine the optimal least squares model parameters and time delays. The time delays determination is performed because in the real case of the 2011 Tohoku tsunami, the time dependent variation of the tsunami generating process is not negligible (Satake et al., 2013; Hossen et al., 2015). The computed time delays are part of the components of the multiple time window inversion, which indicate the time response of the unit sources to perturbed seafloor. This approach helps the traditional least squares inversion method to find a unique and global optimum solution, and reduces problems caused by the regularization scheme.

5.2 Brief Reviews on the 2011 Tohoku Tsunami

According to the Japan Meteorological Agency, a great earthquake occurred on March 11, 2011 at 5:46:18 UTC with the epicenter at 38.104^{o} N and 142.861^{o} E off the east coast of Tohoku. Its moment magnitude (Mw) was 9.0, which was probably one of the largest earthquakes ever recorded in Japan. The earthquake located in the Japan Trench subduction zone where the Pacific plate subducts beneath the Okhotsk plate, generating an extremely large tsunami. Maximum tsunami runup heights of nearly 40 m were measured along the Pacific coast of Tohoku (Mori et al., 2011). The National Police Agency of Japan has confirmed casualties of about 16,000 deaths, 3000 people missing, and 6000 injured (http://www.npa.go.jp/archive/keibi/biki/higaijokyo_e.pdf, last accessed, October 10, 2015). Other than that, the tsunami also hit four nuclear power stations located at Fukushima.

The Pacific coasts of Tohoku have suffered from large tsunamis in the past. For example, the 869 Jogan earthquake (Minoura and Nakaya, 1991), the 1896 Sanriku earthquake (Kanamori, 1972), and the 1933 Sanriku earthquake (Kanamori, 1971). However, the 2011 earthquake possessed unique characteristics that responsible for the

unprecedented high tsunamis along the northern Sanriku coastline. Several studies have proposed various explanations on the cause for the said occurrence. Gusman et al. (2012) reported a motion of sedimentary wedge generating an additional uplift near the trench during the earthquake, while Kawamura et al. (2012) and Tappin et al. (2014) suggested a submarine mass failure. Other study by Satake et al. (2013) proposed that the occurrence can be explained by a delayed shallow slip to the north of the largest slip. This study is expected to describe the aforementioned issue in terms of water surface displacements initiating the tsunami.

5.3 Waveforms and Bathymetry Data

The 2011 Tohoku tsunami was probably one of the most well-documented megathrust events in history. The tsunami waveforms were recorded at many locations by various instruments, i.e., ocean-bottom pressure (OBP) gauges, global positioning system (GPS), coastal tide/wave gauges, and other sophisticated devices. This study mainly focuses on the measurements from deep water instruments. This is to avoid interferences from harbors or local shelf effects. There is only one coastal wave gauge included at Tomakomai Port (W602) to confine the source area in the north (Figure 5.3). Other than that, the waveforms are obtained from two OBP gauges off the coast of Tokachi (KPG1 and KPG2), two OBP gauges off the coast of Iwate (TM1 and TM2), ten GPS buoys (801, 802, 803, 804, 806, 807, 811, 812, 813, 815), and four Deep-ocean Assessment and Reporting of Tsunamis (DART) buoys (21401, 21413, 21418, 21419). All the data were pre-processed to remove signals from other sources (i.e., tides, wind waves, and noise), and resampled into 15-s intervals for numerical purposes.

The model domain covers an area between 133–158°E and 28–48°N (Figure 5.3). Three different grid resolutions are used to compute the tsunami propagation and simulate synthetic waveforms at the wave gauge (10 arc sec), OBP and GPS buoys (30

arc sec), and DART buoys (90 arc sec). For the smallest grid (10 arc sec), the bathymetry data are resampled from the J-EGG500, which is a 500-m gridded bathymetric data set provided by the Japan Oceanographic Data Center. The bathymetry data for the 30 and 90 arc sec grids are obtained and resampled from the GEBCO_08 Grid, which is available for public access from the British Oceanographic Data Center.

5.4 Determination of the Source Area

Magnitude (M) to area (A) scaling is often used to estimate earthquake sizes. For instance, Well and Coppersmith [1994] found a linear relation between M and $\log_{10} A$. Assuming a simple rectangular fault model, the fault length (L) and width (W) are calculated from the estimated rupture area using L=2W. These parameters can alternatively be calculated from the aftershock distributions. This approach is required for tsunami waveform inversions that use the fault model assumption to derive the rupture-influenced area projection onto the sea surface, called the source area. Another approach that can be used with and without the fault model assumption is to restrict the source area according to the rupture velocity and duration (Takagawa and Tomita, 2012).

In this study, only information from the recorded waveforms are used to determine the source area, so that it can avoid making assumptions that are not necessarily true. Based on the Huygen's principle, a finite region of the source area can be identified using backward-propagated tsunami arrival times from each observation station. Hayashi et al. (2011) and Haidarzadeh and Satake (2014) successfully implemented this method using commercial software called Geoware Tsunami Travel Times (TTT. v. 3). Here, a simpler ray tracing algorithm is developed using a 16-dot template for calculating tsunami travel times on a rectangular computational grid (Marchuk, 2008).

According to the Huygen's principle, the wave energy radiates omnidirectionally from its source. Therefore, the algorithm searches through all adjacent points to generate wave front and determines the minimum travel time among these points to be used in the subsequent computation. Simple rectangular computational grids with known water depth values can be used to compute the tsunami travel time throughout the area of interest. An example of the grid system and travel time computing scheme is shown Figure 5.1 (Marchuk, 2008). The black squares indicate points on the grid, where the wave has arrived and the travel time values at these points are known.

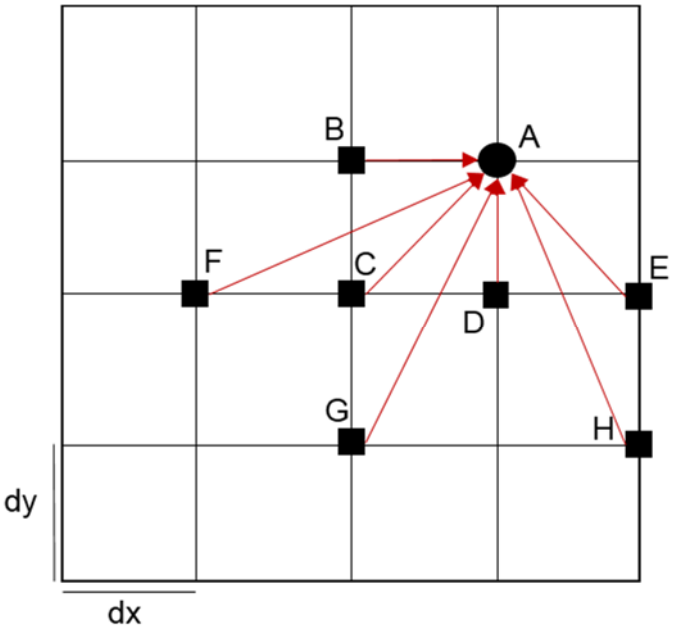

Figure 5.1 The scheme of travel time computation.

Point A is the location where the travel time will be computed, whereas points B, C, D, E, F, G, and H are the neighboring points, in which the travel time are known. Assuming that the depth between adjacent points (d_1 and d_2) is varied linearly, and

L is the distance between those points, the angle of bottom declination can be expressed as,

$$tg(\alpha) = (d_2 - d_1)/L. \tag{5.1}$$

The tsunami travel time can be computed based on the long wave approximation, that is,

$$\begin{aligned} T &= \int_0^L \frac{dl}{\sqrt{g(d_1 + tg(\alpha)l)}} \\ &= \frac{1}{\sqrt{g\,tg(\alpha)}} \int_0^L \left(l + \frac{d_1}{tg(\alpha)} \right)^{-1/2} d\left(l + \frac{d_1}{tg(\alpha)} \right) \\ &= \frac{2}{\sqrt{g\,tg(\alpha)}} \left(l + \frac{d_1}{tg(\alpha)} \right)^{1/2} \Bigg|_0^L = \frac{2}{\sqrt{g\,tg(\alpha)}} \cdot \frac{\sqrt{d_2} - \sqrt{d_1}}{\sqrt{tg(\alpha)}} \\ &= \frac{2}{\sqrt{g}\,tg(\alpha)} \cdot \frac{d_2 - d_1}{\sqrt{d_2} + \sqrt{d_1}} = \frac{2L}{\sqrt{gd_2} + \sqrt{gd_1}} \end{aligned} \tag{5.2}$$

Denoting the tsunami travel time at the neighboring points B to H as T_B, T_C, T_D, T_E, T_F, T_G, and T_H, the travel time from each of this point with respect to the point A can be computed as follow,

$$
\begin{aligned}
T_1 &= T_B + \frac{2\Delta x}{\sqrt{gd_B} + \sqrt{gd_A}}; T_2 = T_C + \frac{2\sqrt{(\Delta x)^2 + (\Delta y)^2}}{\sqrt{gd_C} + \sqrt{gd_A}} \\
T_3 &= T_D + \frac{2\Delta y}{\sqrt{gd_D} + \sqrt{gd_A}}; T_4 = T_E + \frac{2\sqrt{(\Delta x)^2 + (\Delta y)^2}}{\sqrt{gd_E} + \sqrt{gd_A}} \\
T_5 &= T_F + \frac{2\sqrt{(2\Delta x)^2 + (\Delta y)^2}}{\sqrt{gd_F} + \sqrt{gd_A}}; T_6 = T_G + \frac{2\sqrt{(\Delta x)^2 + (2\Delta y)^2}}{\sqrt{gd_G} + \sqrt{gd_A}}, \\
T_7 &= T_H + \frac{2\sqrt{(\Delta x)^2 + (2\Delta y)^2}}{\sqrt{gd_H} + \sqrt{gd_A}}
\end{aligned}
\tag{5.3}
$$

where Δx and Δy are the grid size. The travel time at point A is the minimum values obtained from the T_1 to T_7.

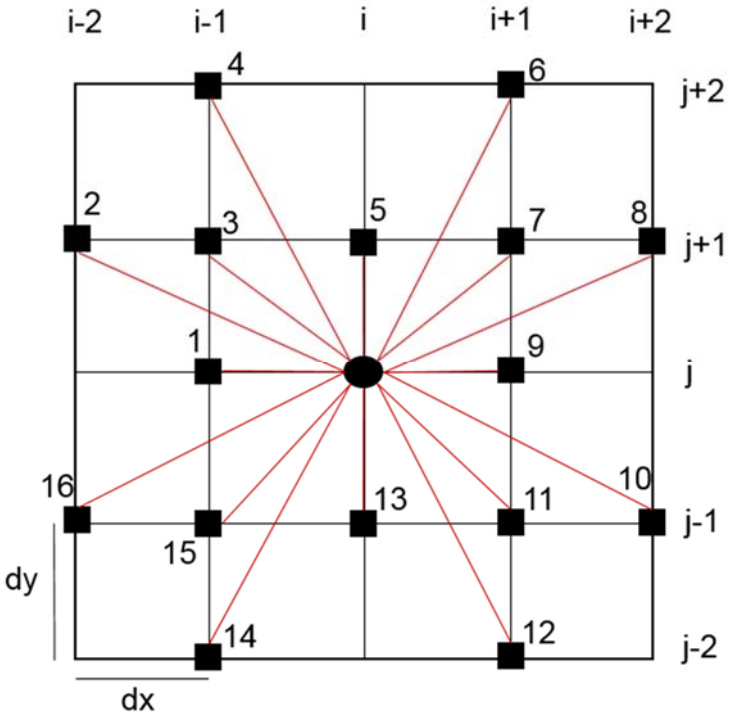

Figure 5.2 The 16-dot template for tsunami travel time computation.

The above description is just an example of the travel time computation procedure, and it does not represent the actual system. The real computation comprises of a certain template for determining the neighboring points. The 16-dot template used in this study is shown in Figure 5.2, where the travel time at the center of the template is computed based on the 16 surrounding points using Eq. (5.3).

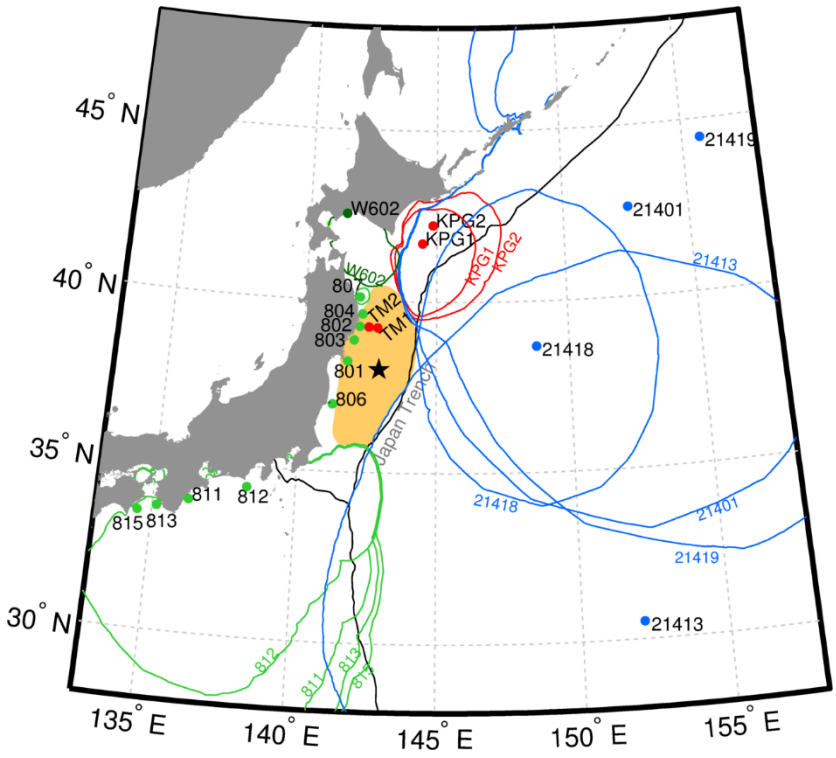

Figure 5.3 Model domain and source area (yellow). The black star indicates the location of the epicenter of the 2011 Tohoku earthquake. The contour lines are the backward-propagated tsunami arrival times from their corresponding observation stations.

To determine the source area, each observation station is considered as the source of the wave. The travel time is computed from all stations, and the duration is taken from the arrival time information on the recorded tsunami waveforms. The waveforms from GPS buoys (801, 802, 803, and 804) and OBP gauges (TM1 and TM2)

are excluded from the calculations, because the arrival times of the seismic and tsunami waves at those locations are almost identical (Hayashi et al., 2011). In other words, these stations are located inside the source area. Figure 5.3 shows the source area bounded by the back-propagated lines.

5.5 Green's Function

The surface deformation is expected to occur throughout the source area, so equidistant unit sources at 15-km intervals are distributed inside the area. Five hundred and twenty-eight unit sources are required to develop the initial Green's function. The same auxiliary basis function as in the ideal case is used, which is based on the 2D Gaussian shape to represent the initial water height of each unit source.

The COMCOT model is employed to calculate the synthetic waveforms. It uses an explicit staggered leap-frog finite difference scheme to solve the linear shallow water equation in spherical coordinates (Wang, 2009). The linear model is used for the three different nested grid system resolutions of 90, 30, and 10 arc sec. This is motivated by the general notion that the tsunami propagation characteristics in deep water are linear. The linear assumption was also validated by the fact that both the linear and nonlinear models produced similar waveforms at the OBP and GPS gauges (Satake et al., 2013).

5.6 Regularization Scheme

In the real case, a high spatial resolution of unit sources is used. Therefore, to obtain a stable solution, we must accordingly apply a proper treatment such as regularization schemes, although this can also cause other problems (Minson et al., 2013). Several studies have proposed various techniques and regularization schemes to overcome such issues. The most tractable method uses a smoothing constraint to maintain the stability of the solution (e.g., Baba et al., 2005; Gusman et al., 2013).

Koike et al. (2003) introduced a more advanced method that reduces the model parameters using wavelets. Singular value decomposition has also been used to control the numerical stability of the tsunami waveform inversion (Voronina, 2011; Takagawa and Tomita, 2012).

For simplicity, the smoothing constraint by Tsushima et al. (2009) is used, which is based on a type of Laplacian derivative operator. Traditionally, the extent of the regularization parameter γ in Eq. (2.7) is often calculated using a trial-and-error approach, which is somewhat subjective. Here, the L-curve criterion is utilized to determine the optimal regularization parameter that balances the fitting error and regularization error (Hansen, 1992).

The L-curve plot is a plot of the (semi)norm $\|\mathbf{I}\boldsymbol{w}\|_2$ of the regularized solution versus the corresponding residual norm $\|\mathbf{G}\boldsymbol{w}-\boldsymbol{\zeta}\|_2$. The subscripts correspond to what is known in the literature as the L_2 norm. The L-curve displays the compromise between minimization of these two quantities, which is the essence of regularization methods. The optimal value of the weighting quantity is shown by a distinct corner separating the vertical and horizontal part of the curve.

5.7 Selection of Model Parameters

In the ideal case (Chapter 3), the algorithm was used to find the optimal spatial distribution of unit source locations. The unit sources were randomly distributed inside the source area to broaden the search space, so that they can provide any possible complex shapes of the water surface profile. However, this approach requires an accurate interpolation to synthesize waveforms that originated from the optimized unit sources located between the initial Green's functions. The forward model was re-run for each optimized unit source to avoid interpolation errors, which was computationally expensive. In the real case application, instead of using randomly distributed unit

sources, a very dense spatial resolution is developed, which produces a large number of model parameters. However, only the optimal ones are selected.

To find the optimal least squares model parameters, an identification number is assigned for each unit source to be either included (1) or excluded (0) from the inversion. The aim of the GA-PS optimization is to determine the set of vectors containing these numbers that yields the best least squares fit. It is expected to reduce the number of model parameters by removing duplicate information from adjacent unit sources, while still maintaining the model's ability to capture the heterogeneous water surface profile. Fewer model parameters are analogous to lower degrees of freedom, which means the unique and optimal solution is more likely to obtain than when using the initial Green's function. The following algorithm selects the optimal model parameters.

Set the initial random vector to $\boldsymbol{x} \in \boldsymbol{X}\{0,1\}$. *The length of the vector is equal to the number of the initial unit sources (528).*

while stopping criteria is not met do

Apply the vector as an index for the initial Green's function $\boldsymbol{\zeta} = \mathbf{G}_x \boldsymbol{w}$.

Use only the selected component on the Green's function $(x = 1)$.

Perform the single time window least squares inversion, $\boldsymbol{w}' = \left(\mathbf{G}'^T \mathbf{G}'\right)^{-1} \mathbf{G}'^T \boldsymbol{\zeta}$,

where $\mathbf{G}' = \mathbf{G}_1$.

Evaluate the fitness function in Eq. (4.7).

Update $\boldsymbol{x}$ *using the genetic operators/pattern search optimization*.

The optimized model parameters ($\boldsymbol{w}'$) are expected to be the unique and optimal solutions with smaller vector dimensions than the initial parameter vector, $\boldsymbol{w}$.

The optimization settings and parameters for the GA and PS are presented in TABLE, which are the default values used in many applications. However, slight modifications are introduced in the stopping criteria for both GA and PS.

Table 5.1 GA and PS default parameter values.

Genetic algorithm				
Population size	**Elite count**	**Crossover function**	**Mutation function**	**Stopping criteria**
20 for model parameters selection, and 50 for time delays determination	0.05 times population size	• Scattered • Crossover rate = 0.8	• Uniform • Mutation rate = 0.01	• Maximum generation > 500 • Average fitness change over 50 generations <= 1e-6
Pattern Search				
Mesh size	**Mesh size factor**	**Stopping criteria**		
Initial = 1, and maximum allowed size = ∞	• Contraction = 0.5 • Expansion = 2	• Mesh tolerance <= 1e-6 • Maximum iteration > 1500 • Minimum fitness change between two consecutive iterations <= 1e-6		

5.8 Time Delays Determination

The time delays can be calculated using three different approaches. The first (Model 1) uses a constant perturbation velocity of 2 km/s originating from the epicenter, which is based on the rupture velocity used by Satake et al. (2013). In the second approach, the time delays are directly considered as model parameters in the GA-PS optimization (Model 2). According to the Satake et al. (2013), a delayed large slip occurred at approximately 3 min after the earthquake, so the optimization of this parameter is bounded between 0 and 180 s. This approach is independent of the source

origin, because the time delay is randomly assigned for each of the parameters optimized in the previous stage. The last approach computes the time delays from spatially varied velocity values (Model 3). Similarly to the second approach, the values are first randomly distributed throughout the selected parameters, and then the time delays are calculated using the ray tracing algorithm used in the travel time computation. Therefore, the time delays are constrained to start at the epicenter and form a systematic path through the source area. A minimum velocity of 1 km/s and a maximum of 3 km/s as bounds for the optimization are considered.

The algorithms for determining the optimal time delays using Models 2 and 3 are very similar. The only difference is that Model 3 uses a ray tracing computation.

Set the initial random vector to $\boldsymbol{x} \in \boldsymbol{X}$ *, where*

$\boldsymbol{X} \doteq \left\{ \boldsymbol{x} \in \mathbb{R}^n \middle| l^i \le x^i \le u^i,\ i \in \{1,...,n\} \right\}$. *For Model 2, l=0 and* u *=180, and for Model 3, l=1 and* u *=3. The vector length (n) is equal to the number of optimized parameters in the previous stage.*

while stopping criteria is not met do

For Model 2, the time delay is $\boldsymbol{\tau} = \boldsymbol{x}$.

For Model 3, compute the time delay $\boldsymbol{\tau}$ *from* $\boldsymbol{x}$ *using ray tracing.*

Apply the vector to the optimized Green's function, $\boldsymbol{\zeta} = \mathbf{G}'\left(\boldsymbol{t} - \boldsymbol{\tau}\right)\boldsymbol{w}'$.

Perform the multiple time window inversions, $\boldsymbol{w}'' = \left(\mathbf{G}''^T\mathbf{G}''\right)^{-1}\mathbf{G}''^T\boldsymbol{\zeta}$, *where*

$\mathbf{G}'' = \mathbf{G}'\left(\boldsymbol{t} - \left(\boldsymbol{\tau} + \Delta t\right)\right)$.

Evaluate the fitness function in Eq. (4.7).

Update $\boldsymbol{x}$ *using the genetic operators/pattern search optimization.*

The final model parameters ($\boldsymbol{w}''$) are the solutions obtained from the parameters selected in the previous stage based on their respective time delays.

5.9 Results and Discussion

The rough estimate of the source area using backward-propagated arrival times helps to avoid the laborious route of finding suitable parameterizations of the rupture process, which often have large uncertainties. The method is actually a simplified version of the reciprocity principle for harmonic waves applied to a tsunami (Korolev, 2011), because it locates the source without giving any information on the amplitude. Unfortunately, it also possesses some drawbacks. For instance, we should modify the actual arrival times to account for the time interval between the main shock and the generation of the tsunami in a large earthquake event (Hayashi et al., 2011; Heidarzadeh and Satake, 2014). Additionally, the probable nonlinearity of the phase velocity is not considered in the standard long wave velocity approximation used to calculate the tsunami travel time. Therefore, the result may be slightly inconsistent with previous studies (e.g., Maercklin et al., 2012; Fujii et al., 2011). In this study, the estimated source area cannot cover all of the immediate aftershocks of the 2011 Tohoku-Oki earthquake, especially in the north and the midway between the 806 GPS buoy and the coastline.

The L-curve plot (Figure 5.4) shows that the optimal quantity of the spatial regularization factor is γ =1.833. However, the model results show that we can still improve the accuracy using the proposed method without adjusting the resulted regularization quantity displayed by the L-curve. This is one of the benefits of the method that is related to the reduction of the spatial smoothing effect, which will be further discussed in the next subsection.

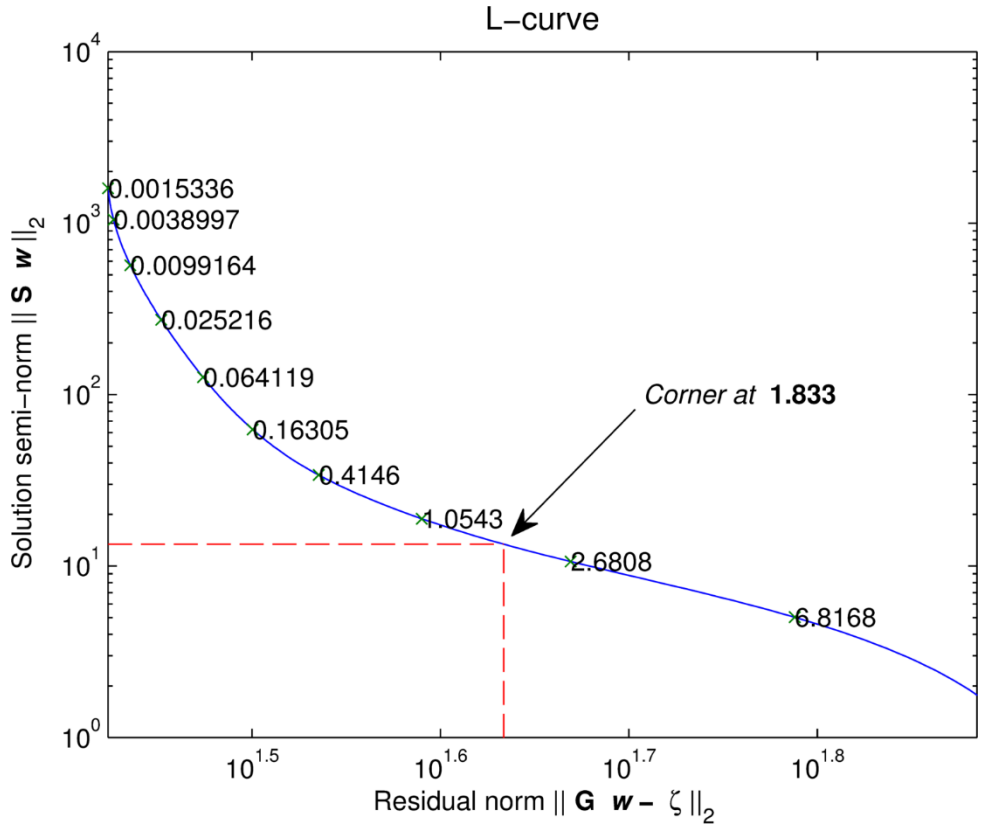

Figure 5.4 The L-curve plot used to choose an appropriate spatial smoothing quantity in the inversion (γ =1.833).

5.9.1 Selected Model Parameters

The inverted tsunami source using the initial Green's function (528 unit sources) has a very smooth profile, and a maximum uplifted area of approximately 7 m located east of the epicenter near the trench (left of Figure 5.5a). Despite the difference in amplitude, the location of this maximum water surface height corresponds to the maximum vertical seafloor displacement reported in previous studies (Fujii et al., 2011; Gusman et al., 2012; Satake et al., 2013). Adjacent to the uplifted area, the troughs of the initial waveform are distributed to the west of the epicenter along the coastline, and have a smallest value of approximately −1.5 m. The accuracy of the inverted waveforms at all gauges is calculated using the root mean square error (RMSE=0.5138 m) and the

Pearson correlation coefficient (r=0.8549), as shown in the right of Figure 5.5a. Comparisons of observed and inverted waveforms in time series plot is shown in Figure 5.6. Although it is difficult to quantitatively compare the result, it resembles other inferred surface deformations produced by tsunami waveform inversions without the fault model assumption (e.g., Saito et al., 2011; Takagawa and Tomita, 2012). However, the lack of spatial heterogeneity in the water profile (which is probably caused by the imposed regularization scheme) may mean that it is not considered a plausible model.

In contrast, GA-PS reduced the number of parameters to 200 and improved the accuracy (RMSE=0.4781 m and r=0.8798), as shown in Figure 5.5b (left). This improvement suggests that the excess model parameters were redundant and constituted a higher degrees of freedom, which limited the search for an optimal solution. It may also imply that the spatial smoothing constraint has been suppressed, because more than half of the model parameters were removed. This is a desirable gain, because using such a regularization without a proper and thorough analysis may reduce the methods ability to distill the underlying physics of the tsunami generating process (Minson et al., 2013). Traditionally, we often modify the extent of the regularization scheme using a trial-and-error approach, which is somewhat subjective. This study uses the L-curve criterion to determine the optimal regularization parameter that balances the fitting error and regularization error. However, further improvement can be achieved using the proposed method without adjusting the resulted regularization quantity displayed by the L-curve (γ =1.833).

Figure 5.5b (left) shows that the location of the maximum uplifted area remains the same, but it is more distinct and has a higher amplitude of more than 8 m. Ito et al. (2011) and Hayashi et al. (2011) suggested that the uplift at this particular site was responsible for the impulsive crest observed in TM1 and TM2. The subsidence of water surfaces are distributed near the coast, which are consistent with the geodetic data (Nishimura et al., 2011). The minimum subsidence inside the source area is now less

than −2 m. This range increment is probably because of the reduced regularization. Overall, the selected model parameters produces a rougher sea surface profile. This is more realistic according to Satake et al. (2013), because a large horizontal displacement of the seafloor that generated shorter wavelength components was responsible for 20–40% of the 2011 Tohoku-Oki tsunami amplitudes. A similar finding was also reported in Gusman et al. (2012), where an additional uplift caused by a co-seismic horizontal displacement excited a short wavelength and high peak movement in the sea surface near the trench. This cannot be observed in the initial Green's function, where the spatial smoothing influences the entire source area. Other studies have also implied that inaccuracies in inferring the tsunami source may be attributed to smoothing of water surface perturbations (Saito and Furumura, 2009; Nosov and Kolesov, 2007).

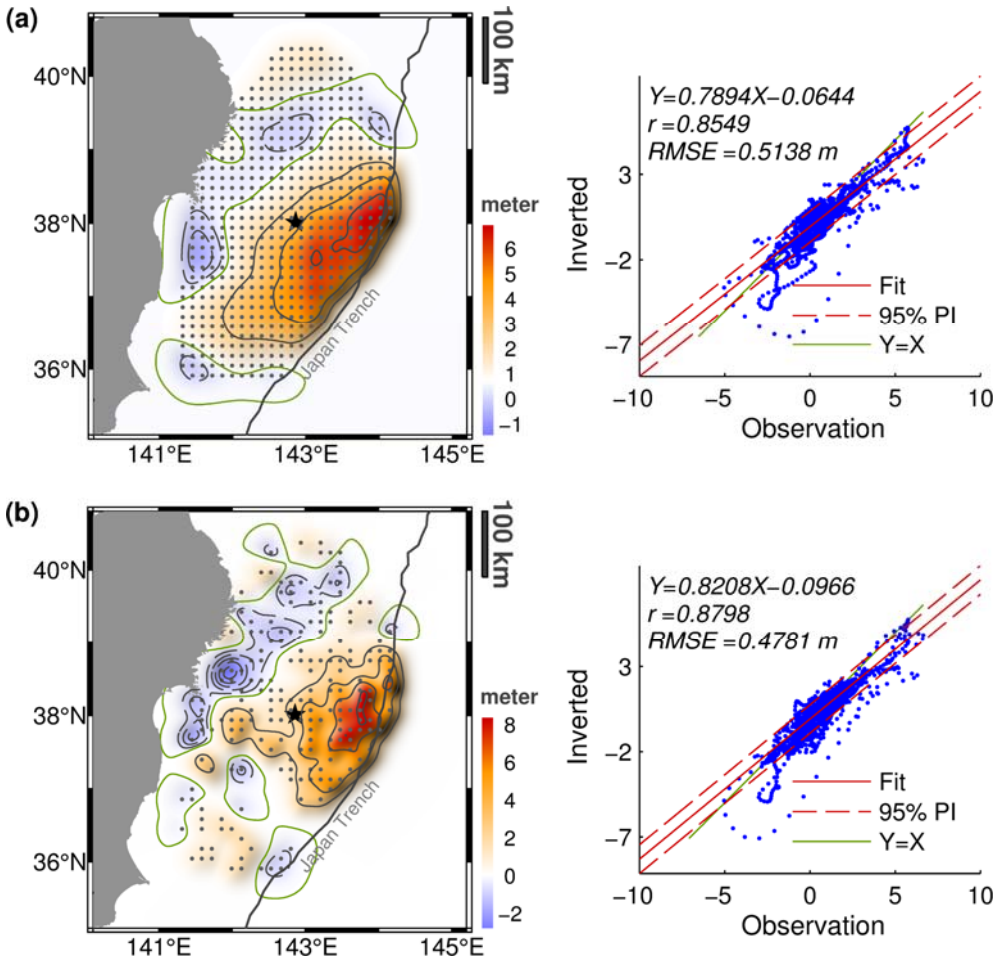

Figure 5.5 Inverted sea surface deformation (left) and corresponding statistical measures of the waveforms fit at all gauges (right). (a) Initial Green's function based on 528 unit sources. (b) After GA-PS optimization with 200 unit sources. The gray dots in the left figures indicate the centroids of the unit sources. The broken and continuous black contour lines show subsidence (0.5 m interval) and uplift (2 m interval). The green contour lines indicate the 0 m elevation inside the source area.

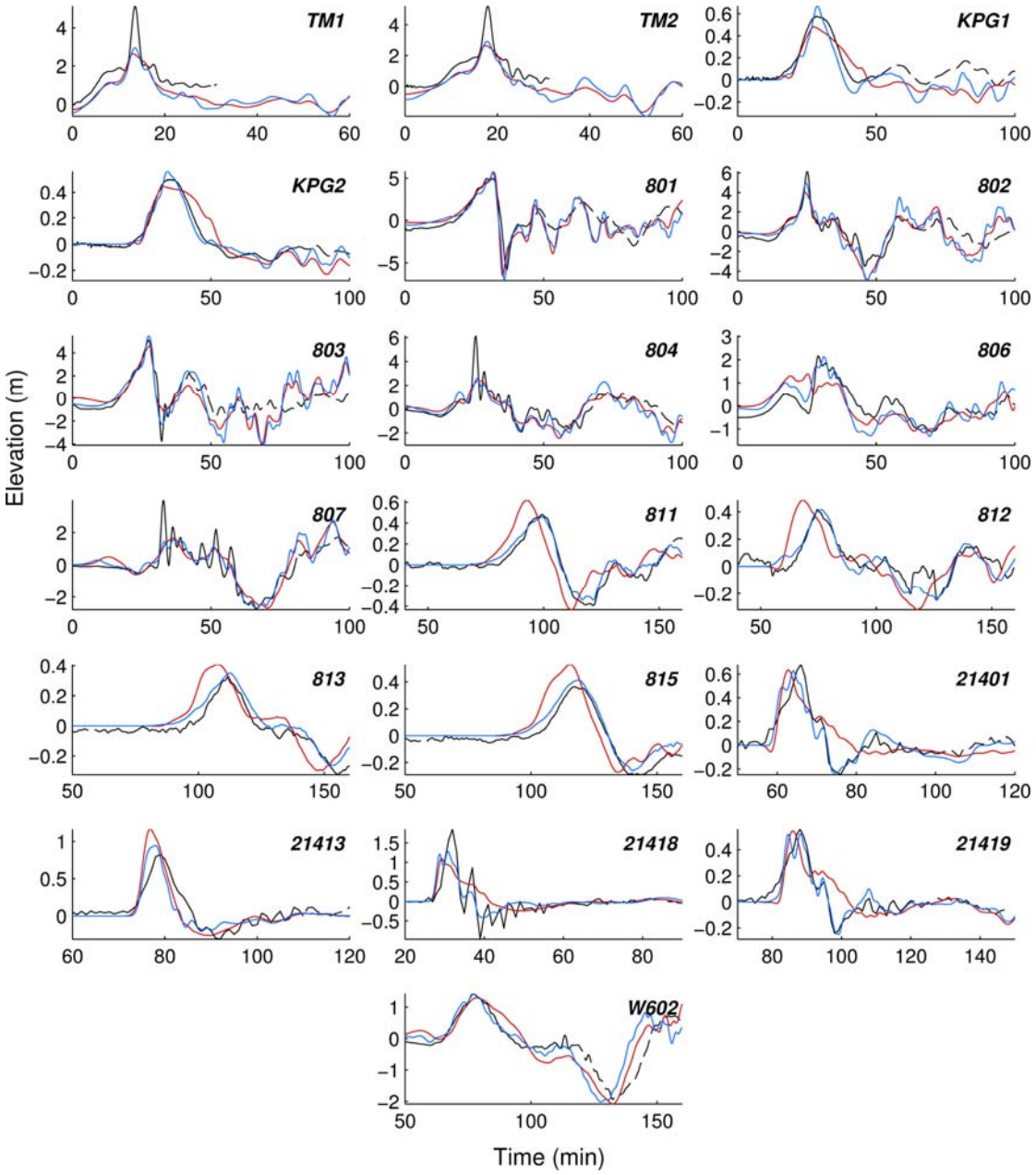

Figure 5.6 Comparisons of observed (black) and inverted waveforms at all gauges obtained using the initial Green's function based on 528 unit sources (red) and after GA-PS optimization with 200 unit sources (blue). Solid black lines indicate the time interval of the data being used for the inversion.

5.9.2 Results for Different Time Delays

The accuracy significantly improved after incorporating the time delays, which suggests that the transient deformation is not negligible in a large earthquake such as the 2011 Tohoku-Oki event. Referring to Eq.(2.11), the spatial smoothing quantity (γ =1.833) and damping parameter (λ =0.5) are used. Determining the time delays using a constant velocity of 2 km/s (Model 1) resulting in an RMSE=0.3874 m and r=0.9243. Figure 5.7 (left) shows the assumed constant velocity invoked from the estimated rupture front radiation formed circular patterns. However, not all unit sources produce substantial values after the inversion. Drawing the outer boundary of only the affected unit sources images the actual effective perturbation/rupture front (broken green contours). From the figure and comparison with other models, we can see that the simplification of constant velocity may not describe the actual physics of the rupture, because several factors cause variations in the rupture velocities (Zhao et al., 2011; Lay et al., 2012; Wei et al., 2012). Gusman et al. (2012) reported a motion of sedimentary wedge near the trench during the 2011 Tohoku-Oki earthquake. This region is usually characterized by a low rigidity area that caused a very slow rupture expansion (Lay et al., 2012). Another complexity of this event was reported in Maercklin et al. (2012), who noted that the earthquake originated from twin ruptures at different locations. Therefore, a more detailed parametrization to determine the time delays is required.

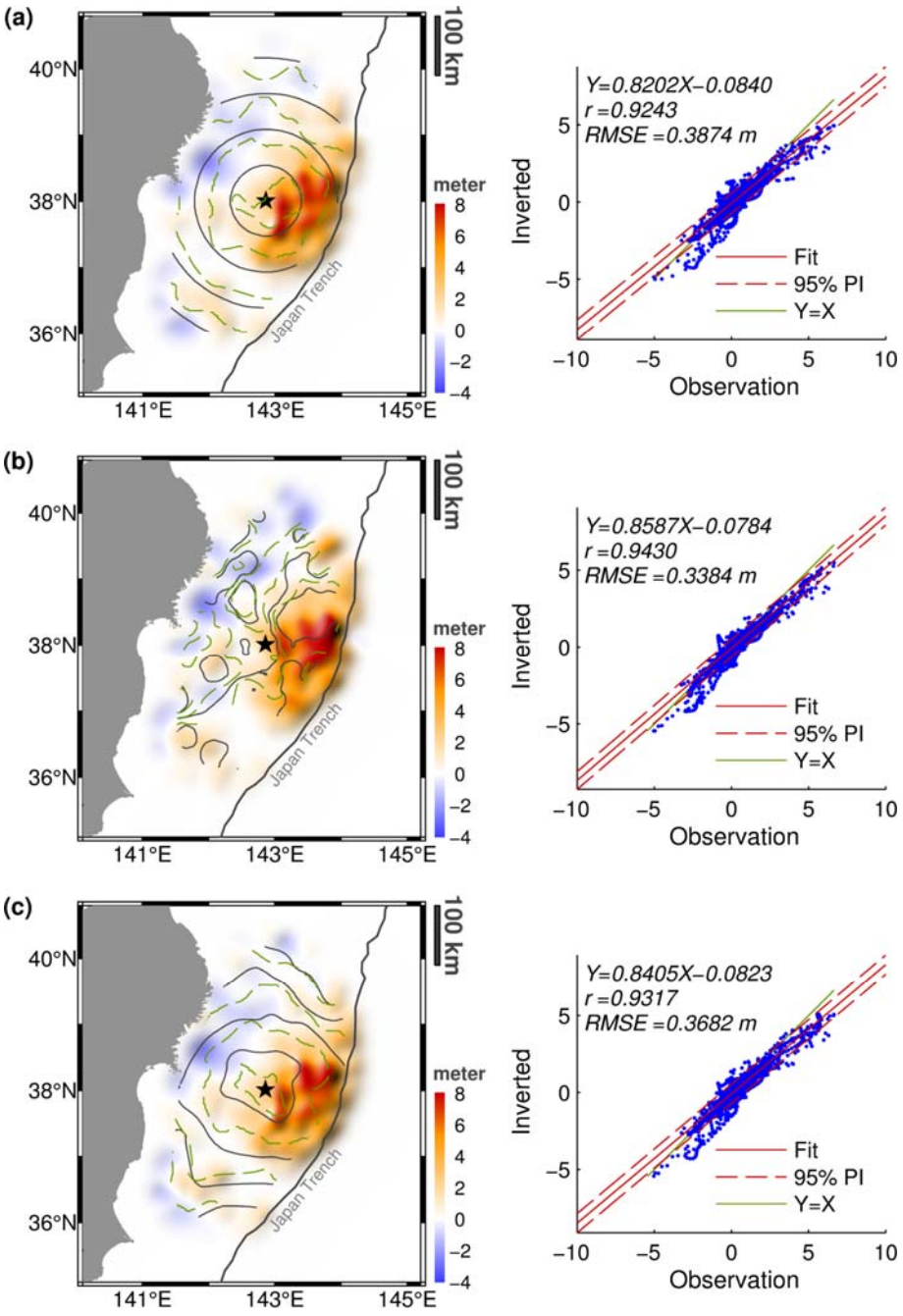

Figure 5.7 Cumulative inverted sea surface deformation after 180 s, overlaid with estimated time delay contours (continuous black) and effective perturbation front contours (broken green) at 30-s intervals (left), and the corresponding statistical measures of the waveforms fit at all gauges (right). (a) Using a constant velocity of 2 km/s (Model 1); (b) using a direct GA-PS optimization of the time delays (Model 2); and (c) using varied velocities from the GA-PS optimization (Model 3).

Model 2 is the most accurate (RMSE=0.3384 m and r=0.9430). This suggests that the tsunami generating process was too complex to assume a constant velocity. However, it is difficult to relate the distribution of time delays or effective perturbation fronts depicted in Figure 5.7b (left) with the actual physical phenomena, because the design parameter allows the water surface response to be simultaneously initiated at different locations with random directions of perturbation. The only decipherable image is of the relatively wide contour intervals in the highly uplifted areas, which indicate faster radiations than other locations. Model 2 is more appropriate than the other models for cases where there are simultaneous responses at distance locations. However, without any supporting evidence (e.g., a direct measurement of the sea surface deformation during the rupture), we cannot determine if there are any physical correlations or if the results are only the best in a statistical sense.

Model 3 yields more realistic results; the perturbation starts at the epicenter and multi-directionally propagates with different velocities (Figure 5.7c (left)). The accuracy (RMSE=0.3682 m and r=0.9317) indicates that the spatially varied velocity is better than the constant velocity used in Model 1. Statistically, Model 2 is the best model (in terms of RMSE and r), but the Model 3 result is more relevant to the physical processes. Furthermore, our visual inspection suggests that the slow perturbation velocities shown by Model 3 (narrow contour intervals for both the time delays and effective perturbation fronts) occur to the north of the source near the trench. Although the estimated time delays may not explicitly explain the actual response time of the unit sources, they can facilitate the inversion to find a better solution. This is shown by the figure, where the effective perturbation front patterns follow the estimated time delays. An example can be seen at the location very close to the small portion of the highly uplifted area to the north. Therefore, this may confirm the movement of sedimentary wedge suggested by Gusman et al. (2012) because a rupture at a shallow depth in low rigidity material enhances the uplift of the seafloor (Lay and Bilek, 2007).

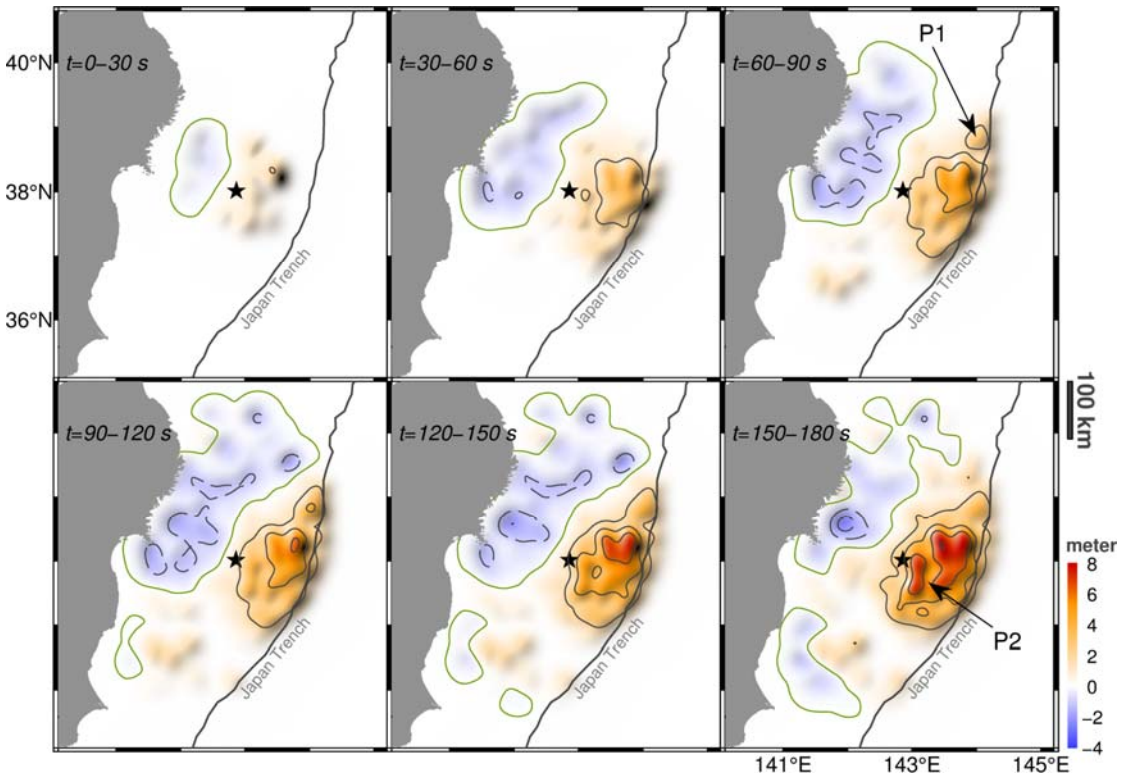

Figure 5.8 Cumulative inverted sea surface deformation at 180 s for 30-s time intervals, obtained from the optimized model parameters and time delays (Model 3). The broken and continuous black contour lines show subsidence (0.5 m interval) and uplift (2 m interval). The green contour lines indicate the 0 m elevation inside the source area.

Based on the previous discussion, Model 3 produces the most reasonable representation of the tsunami source. Therefore, the cumulative transient deformations shown in the snapshots of time-dependent variations are taken from Model 3 (Figure 5.8). The water surface uplift starts to move from the east of the epicenter, and then propagates bilaterally northwards and southwards along the trench axis, while the subsidence appears near the coast. At the period between 60-90 s, the positive water surface displacement extends westwards across the epicenter toward the coast. The model also captures a few distinctive patterns that are similar to previous studies. For example, P1 shows a northern extension of the source that may be responsible for the

high tsunami along the Sanriku coastline (Wei et al., 2013; MacInnes et al., 2013). The disturbance at this particular location was interpreted by Gusman et al. (2012) as an additional uplift caused by a motion of a sedimentary wedge, whereas Kawamura et al. (2012) and Tappin et al. (2014) suggested a submarine mass failure. The delayed rupture proposed by Satake et al. (2013) appeared at 180 s is also observed in the result (P2). However, the location of the large uplift due to the delayed rupture is closer to the epicenter compared to the Satake et al. (2013). The cause of this discrepancy was discussed in Tappin et al. (2014), which compared the rupture centroids from many publications (for the 2011 Tohoku-Oki earthquake) and showed that there were significant variations clustered within a 30-km radius.

The inverted tsunami waveforms using the optimized parameters and time delays obtained from Models 1, 2, and 3 are shown in Figure 5.9. All models significantly underpredict the maximum height of the observed waveforms at the northern GPS buoys (804 and 807). MacInnes et al. (2013) concluded that an additional source is needed to explain such high tsunamis. Tappin et al. (2014) resolved this problem by performing rigorous parameterizations of submarine mass failures located northeast of the epicenter. The secondary source in their simulation can reproduce high amplitude and frequency waves recorded at GPS buoys off Sanriku. The sea surface profile of this additional source is also captured in the proposed model, but has smaller dimensions (P1). At DART buoy 21418, all models are unable to approximate the later phases or dispersive wave train of the observed waveform, because of limitations to the non-dispersive forward model. Saito et al. (2011) and Tappin et al. (2014) confirmed that the dispersive properties of the waveform recorded at this site (situated approximately 600 km from the source) cannot be reproduced by a non-dispersive model.

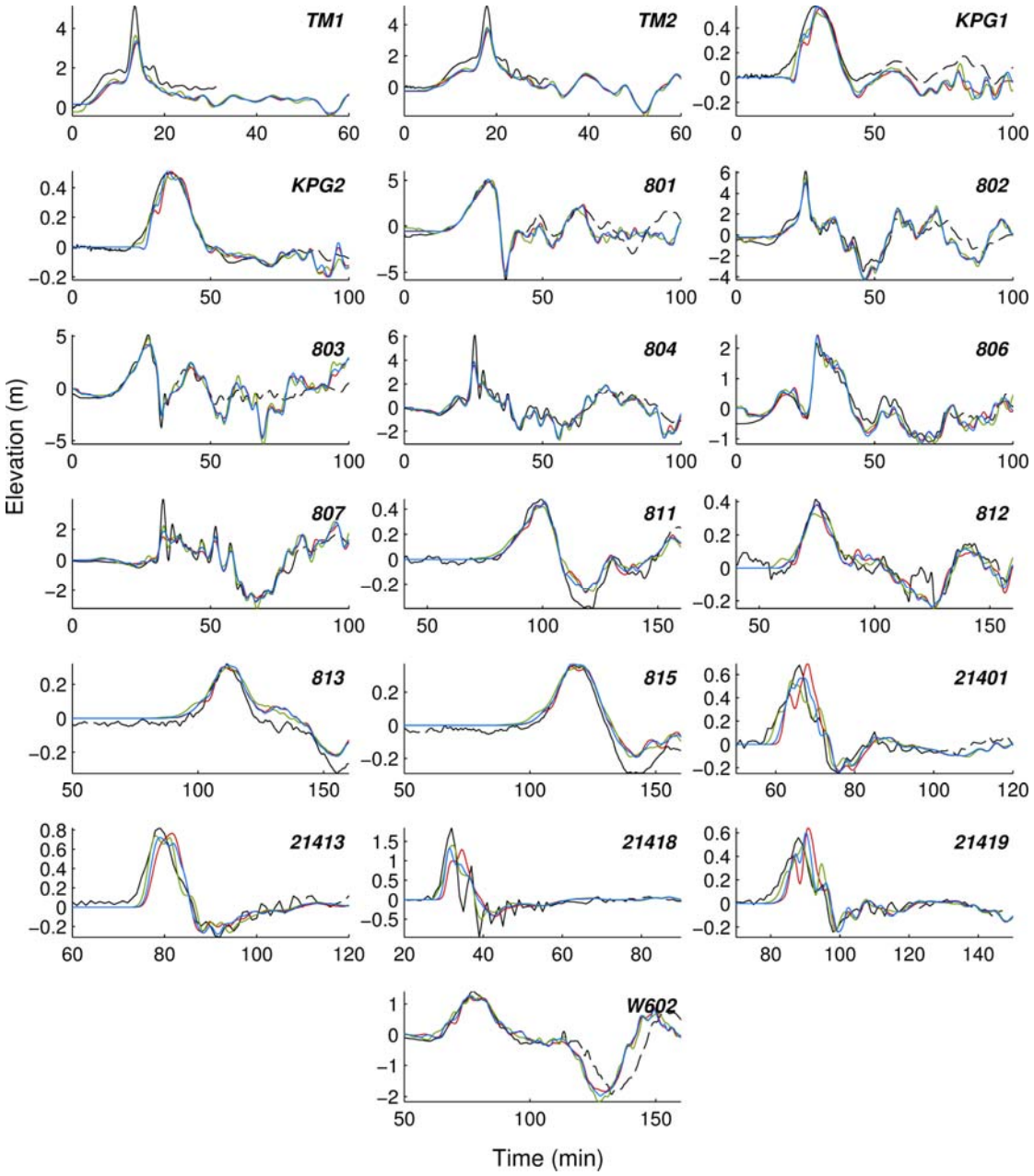

Figure 5.9 Comparisons of observed (black) and inverted waveforms at all gauges obtained using Model 1 (red), Model 2 (green), and Model 3 (blue). Solid black lines indicate the time interval of the data being used for the inversion.

Therefore, a more sophisticated forward model is required to reveal the multi-scale phenomena exhibited by the tsunami. Furthermore, at DART buoys, Model 1 shows

poor estimations of the waveform phases compared to Model 2 and Model 3. This should not be a surprise, because Model 1 assumes the constant perturbation velocity. We can adjust the velocity value to obtain better fits at the mentioned stations, but it will reduce the accuracy at the other stations. For that reason, it further confirms the previous discussion that we can enhance the model performance by incorporating the spatially varied velocity.

All model results at TM1 and TM2 underpredict the observation. This is probably because, in this study, the observed waveforms at OBP gauges are treated as surface gauges, such as GPS buoys and DART buoys. This is relevant for far-field tsunami or when the OBP gauges located outside the source area. During the 2011 Tohoku tsunami, TM1 and TM2 were located inside the source area (Hayashi et al., 2011). Therefore, the OBP sensors records not only pressure variations due to the resultant tsunami, but also those due to permanent seafloor deformation. This permanent seafloor deformation should be incorporated in the Green's function to obtain more accurate results at these particular stations (Tsushima et al., 2012).

5.9.3 Effects of the Hybrid Optimization

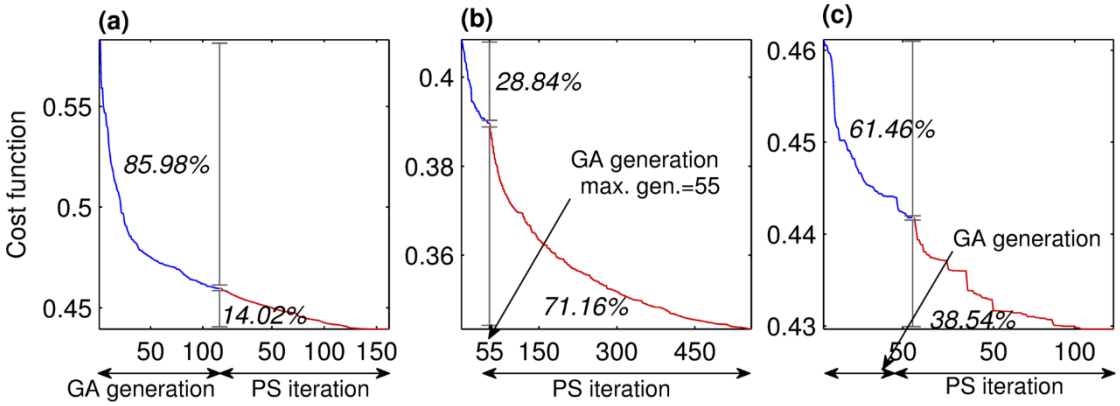

Figure 5.10 Surface errors of the GA (blue) and PS (red) according to the specified design parameters. (a). Least squares model parameters

selection. (b). Time delays selection in Model 2. (c). Time delays selection in Model 3.

To determine the benefit of the hybrid approach, the surface errors are plotted based on the proposed cost function against the generation or iteration of GA-PS optimization (Figure 5.10). Comparing the model parameters selection (Figure 5.10a) and time delays determination using Model 3 (Figure 5.10c) reveals that the GA is the predominant factor, and contributed to 85.98% and 61.46% of the overall reduction in error, respectively. However, further refinement by PS indicates that the optimal solution is still situated beyond the reach of the GA. When selecting the time delays using Model 2 (Figure 5.10b), PS has a larger effect and reduces the error by 71.16%, but take longer to converge. This is because of the broader search space of Model 2, where the time delays are not constrained to certain behaviors as in Model 3. Based on the above discussion, the hybrid algorithm is necessary, because GA explores promising areas in the nonconvex function so that it can avoid local minima, but its search is confined by their exploration through the global structure of the cost function. The PS algorithm does not have this characteristic, therefore it can get trapped in a local minimum if used from the beginning of the optimization.

Chapter 6

Real-time Forecasting of Tsunami Waveforms

6.1 Problem Description

Neural networks are part of the growing artificial intelligence branch classified as CI, and are known for their ability in approximating most practical nonlinear functions. Neural networks have been applied to tsunami forecasting in previous publications (Romano et al., 2009; Hadihardaja et al., 2010; Mase et al., 2011). However, the triggering mechanisms for those methods rely on estimates of the earthquake parameters. Namekar et al. (2009) proposed a better solution by incorporating tsunami waveforms (as in the TWI) to forecast coastal waveforms and the run-up. Their method requires no (or minimal) earthquake information. The only drawback is that the back propagation algorithm requires iterative training, which results in a considerable computation time. Consequently, it may not be appropriate for near-field tsunami events.

This study proposes using a recent neural network algorithm called ELM. The ELM can be trained significantly faster than gradient-based methods such as back propagation, and produces better generalizations. Therefore, it is a promising

alternative tool for future TEWS. The method is applied to near-field tsunami predictions of the 2011 Tohoku event, which was recorded at various measurement devices (offshore and onshore) covering the Tohoku and Hokkaido areas. To assess the performance of the proposed method, comparisons with the standard TWI results are made and rigorous statistical evaluations are also performed. Additionally, this study also conducts a model uncertainty analysis to ensure the consistency of the method, by observing the variability of several test runs. This last assessment is necessary because, as with any typical neural network methods, the resulting forecasts are subject to uncertainties caused by the random network parameters.

The limitation of this method is that we cannot interpret any physical meaning from the optimized network parameters, which is why this type of model is often called a black-box model. In TWI, the optimized model parameters represent the initial displacement of the tsunami, or the slip distribution. However, for forecasting purposes, speed and accuracy are the most important factors, therefore such an information is not necessarily required.

6.2 Tsunami Data and Bathymetry

This study uses tsunami waveforms recorded by four types of measurement devices: ocean bottom pressure (OBP) gauges, global positioning system (GPS) buoys, wave gauges, and tidal gauges. Data from four OBP gauges are used: two are deployed off the coast of Iwate (TM1 and TM2) and operated by The University of Tokyo and Tohoku University, and the other two are located off the coast of Tokachi (PG1 and PG2) and are operated by the Japan Agency for Marine-Earth Science and Technology. The other gauges used are six GPS buoys (G801, G802, G803, G804, G806, and G807), six wave gauges (W202, W203, W205, W219, W602, and W613), and three tide gauges (T618, T624, and T625), which are operated by the Ministry of Land, Infrastructure,

Transport, and Tourism. The locations of all the gauges are shown in Figure 6.1. All the data were pre-processed to remove signals from other sources (i.e., tides and wind waves), and resampled into 15-s intervals.

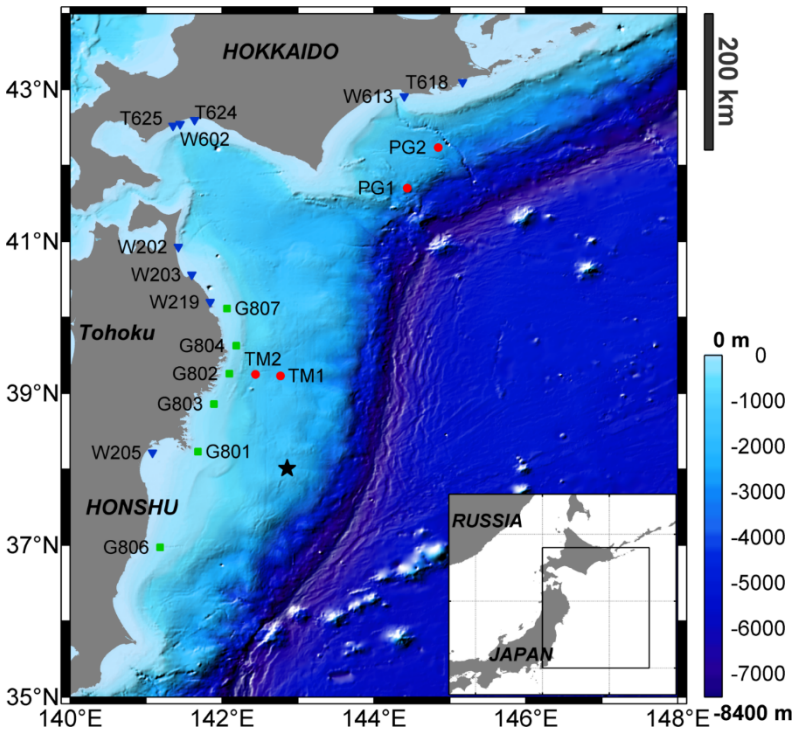

Figure 6.1 Bathymetry profile of the study area and gauge locations. The red dots are the OBP gauges, the green squares are the GPS buoys, and the blue triangles are the coastal gauges (T: tide gauges and W: wave gauges). The black star is the epicenter of the 2011 Tohoku earthquake.

To generate the synthetic waves used in the forecast algorithm, this study uses a numerical simulation with nested grid systems that covers an area between 140–148°E and 35–44°N. The grid size of the model for the offshore stations is 30 arc sec, and 10 arc second for the coastal gauges. For the larger grid size, the bathymetry data are obtained from the General Bathymetric Chart of the Oceans (GEBCO_08 Grid) provided by the British Oceanographic Data Center. The bathymetry for the smaller

grid (10 arc sec) is resampled from the J-EGG500, which is a 500-m gridded bathymetric data set provided by the Japan Oceanographic Data Center.

The availability of high-quality data with sufficient temporal resolution allows us to investigate dynamic rupture information of the 2011 Tohoku earthquake characterized by a relatively slow rupture propagation (Satake et al., 2013). However, this study assumes an instantaneous deformation to maintain the efficiency of the forecast algorithm when applied to smaller tsunamis or to events where high-temporal resolution data is unavailable. Moreover, the goal of this study in part is to compare the performance of the proposed method with the standard TWI using the same inputs, settings, and computational resource.

6.3 Model Development

The main goal of conventional TWI is to determine the magnitude of the slip on a fault plane or to infer the initial displacements that excite the tsunami. However, it can also be extended to estimate waveforms outside the inversion time range. The same two-dimensional Gaussian shape as in the Subsection 4.2.3 is used instead of a fault model to represent the initial water surface deformation. This is more practical for TEWS, because we do not need information about the fault geometry of the earthquake. Equidistant unit sources at 20-km intervals are distributed throughout the study area (Figure 6.2). Then, for each unit source, computation of the tsunami propagations is performed using the COMCOT model based on the linear shallow water equation (Wang, 2009). The resulting synthetic waveforms are stored for all the gauge locations.

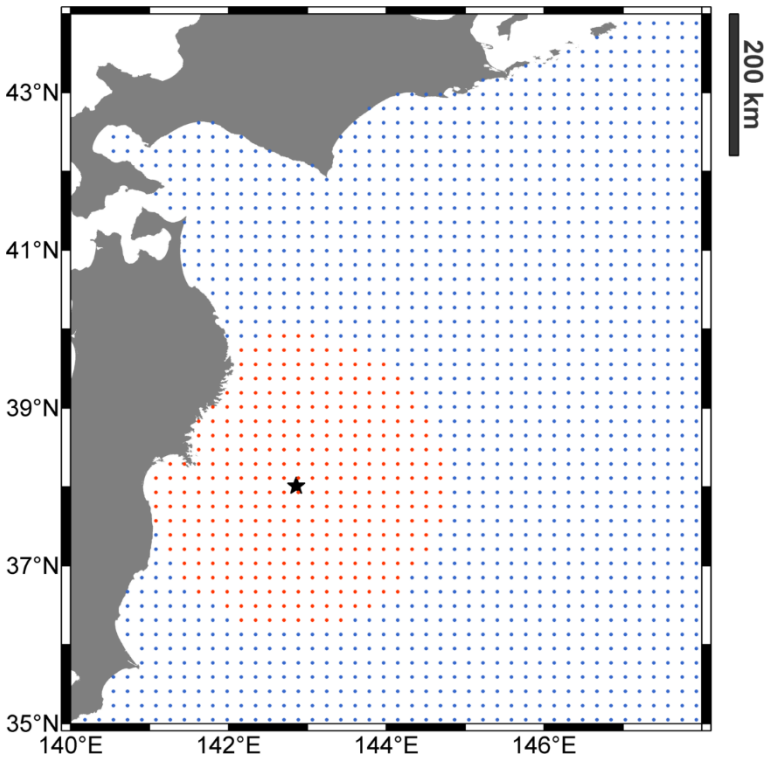

Figure 6.2 Equidistant unit sources at 20-km intervals, which are used to develop the synthetic waveforms database. The dots are the centroids of the unit sources. Orange dots indicate the unit sources selected for the 35-min training period. The black star is the epicenter of the 2011 Tohoku earthquake.

The number of unit sources in the inversion is limited by confining the source-influenced area using a very simple approach. Only $w(\tilde{t})$ unit sources is included, where $(\tilde{t} = r/\upsilon) \le T$, r (in km) is the radius from the unit sources to the epicenter, and υ is a constant that adjusts the size of the source-influenced area relative to the inversion time range. After running some trial and error tests, υ =0.07 km/s is selected. Note that υ is not similar to the rupture velocity. Using this technique, the only information we need about the earthquake is the location of the epicenter. This approach helps to reduce the number of unknown parameters in the inversion, so that we can decrease the degrees of freedom and redundancy.

As mentioned in the Section 3.3 that unlike conventional neural networks, ELM may require more hidden neurons because of the random determination of the input weights and biases (Zhu et al., 2005). Various methods have been proposed for finding the optimum number of hidden neurons (Hunter et al., 2012; Mulia et al., 2013), but they are not suitable for online training. For simplification purposes, the optimal number of hidden neurons for this case is assumed to be approximately two thirds of the input vector dimension.

Although in theory, it is not necessary to normalize inputs or outputs of the neural networks, some applications have shown that a normalization can yield more efficient training processes (Kim, 1999). In this study, the output values are normalized in a range of -1 to 1, while the inputs remain unchanged, since they already have the same range [-1,1]. The input weight and bias values are also limited between -0.5 to 0.5. Both the output normalization and the value limitations for the network parameters are expected to facilitate the ELM training process, so that it can produce a consistent predictive power.

6.4 Model Evaluations

To objectively assess the quality of the proposed method, this study quantitatively compares the observed and modeled waveforms using standard statistical measures: the mean absolute error (*MAE,* [0 to +∞], optimal value=0), which indicates the average magnitude of the model error (Willmott et al., 1985); the mean square derivative error (*MSDE*, [0 to +∞], optimal value=0), which measures the first-order derivatives of the modeled and observed values and is the difference in residuals between successive time steps (de Vos and Rientjes, 2008); the percent error in peak (*PEP*, [0 to +∞], optimal value=0), which evaluates the discrepancy between the highest modeled and observed values (Hauduc et al., 2011); the Pearson correlation coefficient

(r, [-1 to 1], optimal value=1), which measures the strength and direction of a linear relationship between two variables; and the balance criterion (BC, [-∞ to 1], optimal value=1), which measures the ability of the model in reproducing the same cumulative values as observed (Perin et al. 2001).

In addition to these statistical measures, this study also calculates a single point difference between the observed and modeled waveforms at the coastal gauges in terms of the maximum wave amplitude of the first leading waves and its arrival time (the ELM results represented by a mean value from several runs). These two criteria are crucial information for TEWS, although they do not necessarily reflect the maximum impact of a tsunami, because in some cases, the succeeding tsunamis were significantly higher than the first ones.

6.5 Uncertainty Analysis

Both TWI and ELM are subject to uncertainty due to various factors. However, in an earthquake-induced tsunami, the uncertainty is mainly attributed to the incomplete knowledge of the source process. Detailed analyses on such uncertainty using stochastic source models can be found in Fukutani et al. (2015). This study only quantifies the uncertainty in terms of lower and upper bounds of a 95% confidence interval of the predicted waveforms using a "delete-1" jackknife resampling technique (Sanchez and Cheung, 2007). Each jackknife sample produces different waveforms at the observation points that determines the correlation between observed and computed waveforms.

Another source of uncertainty exists in the ELM modeling. In ELM, the input weights and biases are not adjusted in the same way as gradient-based algorithms. They are defined using random values with a certain distribution. Therefore, it is more probable that each run produces different results (from different random seeds). This study analyzes the results from 20 consecutive runs and visualizes their distribution

using boxplots. This is a standard technique for summarizing uncertainties, and contains the median (mid-point of the data, represented by lines inside the boxes), the inter-quartile range (the middle 50% of the data, represented by the boxes), the bottom 25% and the top 25% of the data values (represented by the whiskers that extend from either side of the box), and the outliers (which indicate abnormalities, represented by the black dots). For comparison purposes, the TWI results are included using red bars. The plot uses the absolute relative error of the same variables as in the second statistical evaluation shown in Table 6.1.

6.6 Results and Discussion

In typical data assimilation processes, the forecast accuracy improves when we increase the amount of data used to calibrate the model parameters. Melgar and Bock (2013) demonstrated this by gradually increasing the inversion time range from 10 to 50 min. Tsushima et al. (2009) also conducted the same test and obtained similar results. For the sake of compactness, this study only show one forecasting result using 35 min data for the inversion, because the method behaves the same as the aforementioned studies. The 35 min truncated data were sufficient for the Tohoku 2011 tsunami, and produced accurate and timely forecasts in a well-balanced manner (Tsushima et al., 2011; Gusman et al., 2014). Hereafter, this inversion time period will be called the training period, which is the general term used in neural network modeling. The unit sources selected for the 35-min training period are shown in Figure 6.2.

6.6.1 General Results

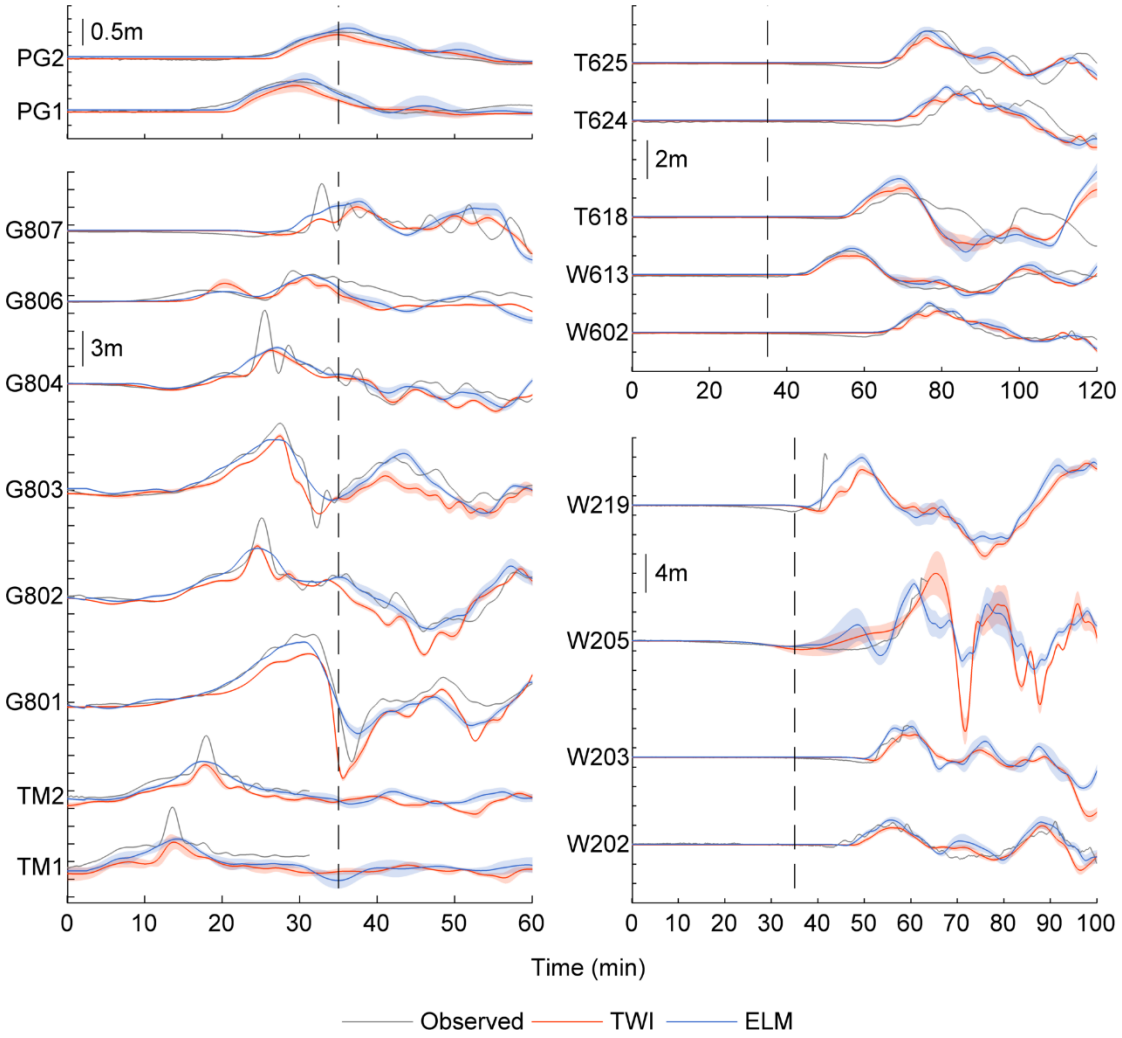

Figure 6.3 Comparison of waveforms from the observations, TWI, and ELM. Vertical dashed lines separate the 35-min training period from the subsequent forecasting period. Light red and light blue shadings indicate the 95% confidence interval for the TWI and ELM, respectively. The left figures are the offshore gauges and the right figures are the coastal gauges.

Comparisons of the observed waveforms with the TWI and ELM results are shown in Figure 6.3. Overall, the ELM produces a more reasonable fit to the observations. In this study, the fitting at the training period is purposely limited to

maintain the generalization ability of the models. Using the Tikhonov's regularization quantity (γ=3), the TWI forecasts at W203 and W205 produce very low troughs in the third and first wave respectively. The troughs will drop unrealistically if we further reduced the regularization quantity. In this case, where the observed waveforms are limited to 35 min, forcing the TWI to fit the observation by reducing the regularization amount would cause an overfitting issue with severe consequences. On the other hand, the ELM forecasts at the two stations are more plausible. Although, unfortunately there are no observed waveforms to confirm these occurrences. Furthermore, the synthetic waveforms are generated using the linear shallow water equation, therefore the TWI can only produce linear approximations. On the contrary, the ELM provides a nonlinear transformation of the input networks for the forecasted waveforms through the nonlinear transfer function (Namekar et al. 2009). The sigmoid transfer function in Eq. (3.8) transforms the linear inputs into nonlinear patterns if the nonlinearity exists in the observed waveforms. Therefore, the ELM can capture nonlinear patterns in the training dataset. However, the ELM requires an additional regularization (the graph regularization) to suppress erratic small-scale variations that are probably induced by the complex network structure.

Tsunami nonlinearities mostly occur at coastal areas and are caused by several factors, for example local bathymetric effects, coastal geometry, etc. An advanced numerical model derived from nonlinear equations must be used to spatiotemporally simulate such an occurrence, but this is computationally expensive (Saito and Furumura, 2009). In term of inversion analysis, the law of superposition is no longer valid for the nonlinear equation. Therefore, the TWI is not applicable for such cases. However, these nonlinearities can be represented by the ELM, because they are accessible through the knowledge acquired in the training process, which is not possible when using TWI. This may also imply that the offshore records contain somewhat nonlinear patterns, because the training datasets mainly consist of water fluctuations at

the offshore gauges. The superiority of the ELM over the TWI is not significantly reflected in the computational effort, because more complex and accurate models typically take more time. In this case, the computing time for the ELM is only 0.35 s, which is slightly longer that the TWI (0.2 s).

Table 6.1 Statistical evaluation results.

	MAE		***MSDE***		***PEP* (%)**		***r***		***BC***	
	TWI	**ELM**	**TWI**	**ELM**	**TWI**	**ELM**	**TWI**	**ELM**	**TWI**	**ELM**
Training	0.471	0.323	0.017	0.020	28.79	23.29	0.920	0.915	-0.594	-1.98
Forecasting	0.630	0.587	0.038	0.036	27.01	18.44	0.694	0.736	-14.14	-9.46

The model performances are quantified using statistical indexes (Table 6.1). Based on these statistical measures, the ELM performs better in the training period than TWI, except for in terms of the *MSDE* (which was 0.020 for ELM and 0.017 for TWI). This is most likely because the *MSDE* specifically penalizes timing errors and noise (de Vos and Reintjes, 2008). As discussed previously, we know that ELM results contain more small-scale variations, therefore they are prone to increased *MSDE* values. The *BC* and *r* values for the TWI in the training period are also better than the ELM values, but not in the more important forecasting period. The accuracy of the ELM extends to the forecasting period as shown by all statistical measures. This proves that the ELM has a better generalization compared to the TWI, as it can produce more accurate forecasts for the unseen patterns in the training period.

To further assess the advantages of the ELM method, the differences between the observed and modeled maximum amplitude in the first leading waves and their arrival times are compared. These are two of the most important factors in TEWSs (Table 6.2). The ELM results are based on mean values from several consecutive test runs used in Figure 6.4. This study only focus on analyzing the coastal gauges distributed along the Tohoku and Hokkaido coastlines. In terms of the maximum

amplitude, the TWI results at all gauges have considerably larger errors (with respect to the observations) than the ELM, and most of them underestimates the observations. In terms of the arrival time, both methods produce relatively comparable forecasts. The ELM estimates are only obviously poor at gauge W219, which is also observed in the TWI result. Satake et al. (2013) produced a better result at this station, because they used more than 35 min data in their inversion time period. Therefore, the water level change was properly captured by their model.

Table 6.2 Maximum amplitude and arrival time discrepancies at coastal gauges. Negative original values indicate underestimates and positive values indicate overestimates. Gauge names with brackets are the stations with incomplete data.

	Original values				**Absolute relative error (%)**			
Gauges	**Max. amplitude (m)**		**Arrival time (min)**		**Max. amplitude**		**Arrival time**	
	TWI	**ELM**	**TWI**	**ELM**	**TWI**	**ELM**	**TWI**	**ELM**
W202	-0.41	0.18	-0.25	0.61	18.06	8.05	0.45	1.09
[W203]	-0.77	0.38	1.50	1.03	23.29	11.39	2.53	1.74
[W205]	0.92	-0.49	3.25	-1.71	14.05	7.49	5.22	2.74
[W219]	-1.41	-0.61	8.00	7.67	26.24	11.28	19.28	18.47
W602	-0.19	0.09	1.50	0.92	13.24	6.05	1.94	1.19
W613	-0.19	0.15	2.00	1.12	15.17	12.26	3.60	2.02
T618	0.39	0.20	1.25	0.51	31.16	16.34	1.81	0.74
T624	-0.28	0.09	-2.25	-4.06	15.64	5.07	2.61	4.71
T625	-0.25	0.11	0.00	0.69	15.09	6.46	0.00	0.92

6.6.2 Model Uncertainty

Even though the ELM has more parameters and higher degrees of freedom, its confidence intervals are comparable to the TWI (Figure 6.3). The confidence interval spreads for both method can be considered narrow, except for the W205. According to Kawai et al. (2013), characteristics of the tsunami waveform recorded at W205 was influenced by some unknown crustal movement effects. Therefore, to address this problem, we must include additional information on the source generation mechanism. For example, Tsushima et al. (2012) incorporated permanent seafloor deformation to correct the initial Green's function and gained substantial improvements in prediction accuracy. Other study by Muhari et al. (2012) utilized arrival time information based on stopped clocks found in the tsunami affected areas to improve their source model.

The absolute relative errors of maximum tsunami amplitudes and arrival times at coastal gauges from 20 model runs are shown in Figure 6.4. The outputs normalization helps the ELM to recognize the patterns in data more efficiently, and thus leads to more consistent forecasting results. However, limiting the input weights and biases between [-0.5, 0.5] has a more significant effect in reducing the variability of the forecasts. Increasing the above range results in a higher probability of obtaining better solutions, but concurrently generates higher variabilities. On the other hand, smaller ranges would produce less variabilities, but increase the median of the absolute relative error.

For the maximum amplitude, the median values at all gauges are closer to zero than the TWI results. The ELM forecasts show various distributions for the gauges, indicated by the length of the boxes and whiskers. Nevertheless, in general, the variations are still below or close to the TWI scores (indicated by the red bars). The largest error of the ELM relative to the TWI is observed at W205, with the discrepancy of only approximately 8%. Furthermore, the number of ELM instances with larger

errors than the TWI is less than 25%, because the TWI score is located inside the upper whisker range. This demonstrates the consistency of the ELM, which results in better predictions than TWI.

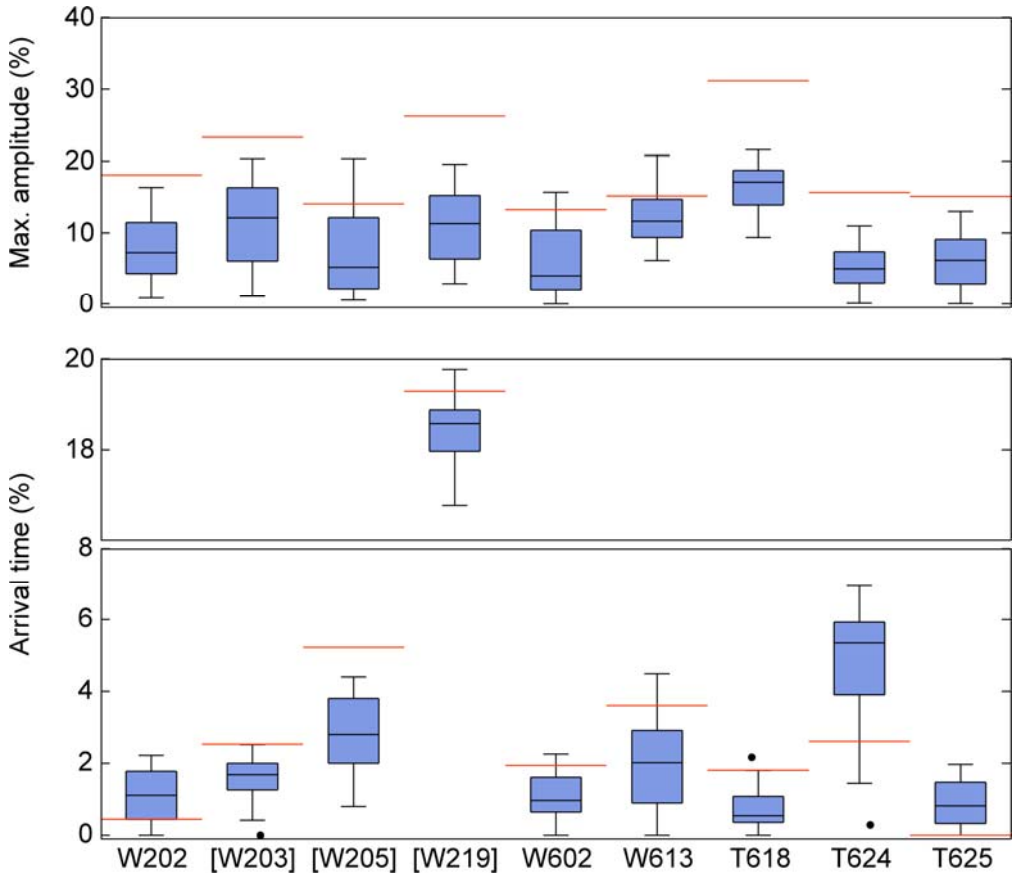

Figure 6.4 Boxplots that indicate the differences between the observed and forecasted maximum amplitudes in the first leading waves and their arrival times, for 20 model runs at nine coastal gauges. The red bars represent the TWI results. Gauge names with brackets are the stations with incomplete data.

The forecasted arrival times are less variable than the maximum amplitude, and most stations have relatively narrow interquartile ranges. This suggests that the arrival time pattern is easier to capture than the maximum amplitude, which agrees with results

from deterministic models. Therefore, the linear model using TWI can produce comparable results to the ELM model. At T264, the TWI result is even better than majority of the ELM results. The TWI and ELM results both have the largest errors at W219. We can see the distinct lag time in the ELM forecast that results in an approximately 18% error (median), and the TWI error of approximately 19%. This latency was discussed earlier as the result of the unprecedented recorded waveform at this particular gauge. Additionally, the observation at this station is incomplete, therefore the value recognized by the detection algorithm may not represent the actual arrival time of the maximum amplitude.

Tsunami source characteristics have a larger impact on the tsunami behavior in near-fields (Okal and Synolakis, 2004). Therefore, we need an appropriate method for revealing the underlying physics of the tsunami generation processes, to fully understand its propagation towards the coastline. The 2011 Tohoku tsunami was a complex megathrust event. Several authors have proposed various approaches for accurately representing its source or initial wave profile, by either incorporating overlooked physical aspects or introducing better models. Gusman et al. (2012) included an additional uplift that excited a short wavelength and high initial sea surface amplitude near the trench, leading to waveforms that better fit the observations. Tappin et al. (2014) suggested a submarine mass failure that may be responsible for the high tsunami along the Sanriku coastline. We should not expect to achieve the same performance or accuracy as these studies, because this particular case is strictly confined by the operational time. Nonetheless, the qualitative and quantitative assessments presented in this study have confirmed the reliability of the proposed method, and that it is applicable to actual TEWS.

The 2011 Tohoku had a unique characteristic, where the observed tsunami data exhibited dispersive waves, even in the near-field (Tappin et al., 2014). This is likely due to the northern extension of the source generating shorter wavelengths that are more

dispersive. A dispersive wave was also observed at Deep-ocean Assessment and Reporting of Tsunamis buoys number 21418 (not used in this study), which can only be captured with dispersive model (Saito et al., 2011). Hossen et al. (2015) confirmed the importance of dispersive effects for the 2011 Tohoku event in their inversion analysis. Therefore, one of possible improvements for the proposed method that should be considered in the future works is to develop a dispersive Green's function.

Chapter 7

Conclusions

This study investigates applications of CI to tsunami source estimation and waveform forecasting. The first method uses a GA-PS algorithm to estimate an idealized tsunami source and that of the 2011 Tohoku tsunami. The second method develops an ELM model for real-time tsunami forecasting in coastal areas, and can be considered as a potential tool for future TEWS. Overall, these methods are found to successfully enhance the standard TWI in both estimating tsunami sources and forecasting tsunami waveforms.

7.1 Tsunami Source Estimation (Ideal case)

Estimations of tsunami sources using CI with a stochastic nature have been conducted. The unique design parameters in the proposed method, applied to determine the optimum location of the unit sources prior to the inversion, demonstrate considerable improvements in accuracy. The random and scattered locations of unit sources have improved the ability of the inversion to produce a more precise approximation of the initial tsunami source without violating the general assumption of long wave theory.

The involvement of stochastic processes in the optimization increases the ability to reveal uncertainties in the tsunami source; such uncertainty is difficult to discern using deterministic approaches. However, as a signature of typical stochastic optimizations, the optimization response time is erratic because of its strong dependency on the initial state. Nevertheless, the results have demonstrated the efficacy of the method for post-event studies of tsunamis, because it can provide better estimations of the coseismic sea surface deformation compared with traditional tsunami waveform inversion methods.

7.2 Tsunami Source Estimation (Real case)

In the real case, the use of an auxiliary basis function based on a two-dimensional Gaussian shape enables the capture of tsunami generation processes that emanate from sources that are independent of the corresponding seismic characteristics. Furthermore, this approach frees us from defining earthquake parameters that are often associated with large uncertainties, and avoids the assumption of identical bottom and surface deformations. In addition to such improvements on the standard method, another objective of this study is to directly infer the initial tsunami source corresponding to an underwater earthquake. The main reason for being able to characterize the triggering earthquake is to provide an accurate initial condition for posterior tsunami models (i.e., propagation, run-up, and inundation models).

The first step is to select an appropriate number of least-squares model parameters, which leads to a reduction in the number of parameters that is proportional to the degrees of freedom. This improves the accuracy of the inversion, and decreases regularization effects. The second step is to determine the initial water movement time for each unit source in response to the perturbation. The stochastic approach provides more options for simulating various scenarios in regard to tsunami generating

mechanisms. For example, consider a case in which the propagation velocity of the surface perturbation spatially varies; similarly, consider a case in which multiple simultaneous responses occur at distant locations. In brief, both the model parameters and time delays are successfully determined, and provide a better Green's function for the inversion.

This study concisely analyzed and compared the results of current analyses with previously existing results to ensure that the method statistically improves the accuracy and that it is physically relevant. The method can describe several tangible processes associated with the 2011 Tohoku tsunami, e.g., the low rupture velocity in a low rigidity area causing significant uplift near the trench. It may also explain the additional northern source that was probably responsible for the high tsunami along the Sanriku coastline, as well as the possibilities of twin ruptures and delayed rupture.

One possible improvement that could be incorporated in the proposed method would be consideration of a more advanced numerical model that includes dispersive effects, because some observed tsunami waveforms exhibit dispersive characteristics. Other than that, additional information regarding the physical evidence of tsunami source processes should be included in future studies. For example, by incorporating permanent seafloor deformation to correct the initial Green's function. This will enable the model to reveal more meaningful physical aspects of the model and to gain substantial improvements in accuracy.

7.3 Tsunami Waveform Forecasting

This study proposes a new approach for forecasting near-field tsunami waveforms using ELM. Unlike the standard method (TWI), which is based on a linear assumption, ELM can capture nonlinearities in a tsunami using nonlinear transfer functions in the network. This feature is important, particularly when a tsunami travels

through an area of shallow water or complex coastal morphologies. Visual and statistical evaluations demonstrate that ELM outperforms TWI. This study shows the effectiveness of ELM at coastal gauges, by comparing the forecasted and observed maximum amplitudes in the first leading waves and their arrival times. The forecast accuracy of the ELM is not proportionally compensated by its computational requirement; the speed of the algorithm is only ~0.15 s less than the standard TWI.

An uncertainty analysis is also conducted, which is necessary to effectively interpret the results and usability of the model. In general, TWI and ELM show comparable confidence intervals, even though ELM involves many more parameters. Furthermore, ELM maintains a consistent predictive power over several consecutive test runs, despite its dependency on random network parameters (i.e., input weights and biases). A few distinct variances are observed at certain locations, but they are still close to (mostly better than) the TWI results with reasonable ranges.

In future studies, we must further improve the ELM performance by using a better numerical model (dispersive model) and incorporating time dependent variations of the tsunami generation, assuming that data with sufficient temporal resolution are available. This will produce a more realistic synthetic waveform database and facilitate the ELM training process. In summary, the speed and accuracy of ELM are important and desired features for a reliable and robust TEWS. With a more detailed and enhanced database, this is a promising tool that could be applied to actual TEWS. Additionally, validation of the method should be undertaken against historical tsunami events that occurred in the areas of interest to ensure the effectiveness of the method when applied to other locations, especially those with limited number of tsunami gauges.

References

Aida, I. (1972). Numerical estimation of a tsunami source. *Zisin 2 (J. Seism. Soc. Japn.), 25*, 343-352.

Amante, C., & Eakins, W. (2009). ETOPO1 1 arc minute global relief model: procedures, data sources and analysis. *NOAA technical memorandum NESDIM NGDC-24*, 19.

Baba, T., Cummins, P. R., & Hori, T. (2005). Compound fault rupture during the 2004 off the Kii Peninsula earthquake (M 7.4) inferred from highly resolved coseismic sea-surface deformation. *Earth Planets Space, 57*, 167-172.

Barnston, A. G. (1992). Correspondence among the Correlation, RMSE, and Heidke Forecast Verification Measures; Refinement of the Heidke Score. *Wea. Forecasting, 7*, 699–709.

Behrens, J., Androsov, A., Babeyko, A. Y., Harig, S., F. K., & Mentrup, L. (2010). A new multi-sensor approach to simulation assisted tsunami early warning. *Nat. Hazards Earth Syst. Sci., 10*, 1085-1100.

Costa, L., Santo, I., Denysiuk, R., & Fernandes, E. (2010). Hybridization of a Genetic Algorithm with a Pattern Search Augmented Lagrangian Method. *International Conference on Engineering Optimization.* Lisbon, Portugal.

de Vos, N. J., & Rientjes, T. H. (2008). Multiobjective training of artificial neural networks for rainfall-runoff modeling. *Water Resour. Res., 44*(W08434).

Fujii, Y., Satake, K., Sakai, S., Shinohara, M., & Kanazawa, T. (2011). Tsunami source of the 2011 off the Pacific coast of Tohoku earthquake. *Earth Planets Space, 63*, 815–820.

Fukutani, Y., Suppasri, A., & Imamura, F. (2015). Stochastic analysis and uncertainty assessment of tsunami wave height using a random source parameter model that targets a Tohoku-type earthquake fault. *Stoch. Environ. Res. Risk. Assess., 29*, 1763-1779.

Geist, E. L. (2002). Complex earthquake rupture and local tsunamis. *J. Geophys. Res., 107*(B5). doi:10.1029/2000JB000139

Geist, E. L. (2005). Local Tsunami Hazards in the Pacific Northwest from Cascadia Subduction Zone Earthquakes. *U.S. Geological Survey Professional Paper, 1661-B*(17).

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization and machine learning.* New York: Longman Publishing Co.

Gusman, A. R., Fukuoka, M., Tanioka, Y., & Sakai, S. (2013). Effect of the largest foreshock (Mw 7.3) on triggering the 2011 Tohoku earthquake (Mw 9.0). *Geophys. Res. Lett., 40*, 497–500.

Gusman, A. R., Tanioka, Y., MacInnes, B. T., & Tsushima, H. (2014). A methodology for near-field tsunami inundation forecasting: Application to the 2011Tohoku tsunami. *J. Geophys. Res. Solid Earth, 119*. doi:10.1002/2014JB010958

Gusman, A. R., Tanioka, Y., Sakai, S., & Tsushima, H. (2012). Source model of the great 2011 Tohoku earthquake estimated from tsunami waveforms and crustal deformation data. *Earth Planet. Sci. Lett., 341–344*, 234–242 .

Hadihardaja, I., Latief, H., & Mulia, I. E. (2010). Decision support system for predicting tsunami characteristics along coastline areas based on database modelling development. *J. Hydroinform. , 13*(1), 96-109.

Hansen, P. (1992). Analysis of discrete ill-posed problems by means of the l-curve. *SIAM Review, 34*(4), 561–580.

Hauduc, H., Neumann, M., Muschalla, D., Gamerith, V., Gillot, S., & Vanrolleghem, P. (2011). Towards quantitative quality criteria to evaluate simulation results in wastewater treatment - a critical review. *Proceedings 8th International IWA Symposium on Systems Analysis and Integrated Assessment in Water Management (WATERMATEX2011)*, (pp. 36-46). San. Sebastian, Spain.

Hayashi, Y., Tsushima, H., Hirata, K., Kimura, K., & Maeda, K. (2011). Tsunami source area of the 2011 off the Pacific Coast of Tohoku Earthquake determined from tsunami arrival times at offshore observation stations. *Earth Planets Space, 63*, 809-813.

Heidarzadeh, M., & Satake, K. (2014). Possible sources of the tsunami observed in the northwestern Indian Ocean following the 2013 September 24 M w 7.7 Pakistan inland earthquake. *Geophys. J. Int., 199* (2), 752-766.

Hirose, F., Miyaoka, K., Hayashimoto, N., Yamazaki, T., & Nakamura, M. (2011). Outline of the 2011 off the Pacific coast Tohoku earthquake (Mw 9.0)- Seismicity: Foreshocks, mainshock, aftershocks, and induced activity. *Earth Planets Space, 63*(7), 513–518.

Honerkamp, J., & Weese, J. (1990). Tikhonovs regularization method for ill-posed problems . *Continuum Mech. Thermodyn., 2*, 17–30.

Hooke, R., & Jeeves, T. (1961). Direct search solution of numerical and statistical problems. *J. Assoc. Comp. Mach., 8*(2), 212-229.

Hossen, M. J., Cummins, P. R., Dettmer, J., & Baba, T. (2015). Tsunami waveform inversion for sea surface displacement following the 2011 Tohoku Earthquake: Importance of dispersion and source kinematics. *J. Geophys. Res., 120*. doi:doi:10.1002/2015JB011942

Huang, G., Zhu, Q., & Siew, C. (2006). Extreme learning machine: theory and applications. *Neurocomputing, 70*, 489–501.

Hunter, D., Yu, H., Pukish, M., Kolbusz, J., & Wilamowski, B. (2012). Selection of proper neural network sizes and architectures-a comparative study. *IEEE Transactions on Industrial Informatics, 8*(2), 228–240.

Imamura, F., Yalciner, A. C., & Ozyurt, G. (2006). *Tsunami modelling manual*. Retrieved from http://www.tsunami.civil.tohoku.ac.jp/hokusai3/J/projects/manual-ver-3.1.pdf

Ito, Y., Tsuji, T., Osada, Y., Kido, M., Inazu, D., Hayashi, Y., . . . Fujimoto, H. (2011). Frontal wedge deformation near the source region of the 2011 Tohoku-Oki earthquake. *Geophys. Res. Lett., 38*(L00G05).

Jin, Y., & Branke, J. (2005). Evolutionary optimization in uncertain environments – a survey. *IEEE Trans. Evol. Comput., 9*(3), 303-317.

Kanamori, H. (1971). Seismological evidence for a lithospheric normal faulting-The Sanriku earthquake of 1933. *Phys. Earth Planet. In., 4*(289-300).

Kanamori, H. (1972). Mechanism of tsunami earthquake. *Phys. Earth Planet. In., 6*, 346-359.

Kawai, H., Satoh, M., Kawaguchi, K., & Seki, K. (2013). Characteristics of the 2011 Tohoku tsunami waveform acquired around Japan by NOWPHAS equipment. *Coastal Eng. J. , 55*(3), 1350008.

Kawamura, K., Sasaki, T., Kanamatsu, T., & Sakaguchi, A. (2012). Large submarine landslides in the Japan Trench: A new scenario for additional tsunami generation. *Geophys. Res. Lett. , 39*(L05308).

Koike, N., Kawata, Y., & Imamura, F. (2003). Far-field tsunami potential and a real-time forecast system for the Pacific using the inversion method. *Nat. Hazards, 29*, 423-436.

Korolev, Y. P. (2011). An approximate method of short-term tsunami forecast and the hindcasting of some recent events. *Nat. Hazards Earth Syst. Sci., 11*, 3081–3091.

Lay, T., & Bilek, S. (2007). Anomalous earthquake ruptures at shallow depths on subduction zone megathrusts, in The Seismogenic Zone of Subduction Thrust Faults. In T. H. Dixon, & M. J. C. (Eds.). New York: Columbia Univ. Press.

Lay, T., Kanamori, H., Ammon, C. J., Koper, K. D., Hutko, A. R., Ye, L., . . . Rushing, T. M. (2012). Depth-varying rupture properties of subduction zone megathrust faults. *J. Geophys. Res., 117*(B04311).

Liu, P. L., Cho, Y. S., Briggs, M. J., Kanoglu, U., & Synolakis, C. (1995). Runup of solitary waves on a circular island. *J. Fluid Mech., 302* , 259–285 .

Liu, P. L.-F., & Wang, X. (2008). Tsunami source region parameter identification and tsunami forecasting. *J. Earthq. Tsunami, 2*(2), 87-106.

MacInnes, B. T., Gusman, A. R., LeVeque, R. J., & Tanioka, Y. (2013). Comparison of earthquake source models for the 2011 Tohoku-oki event using tsunami simulations and near field observations. *Bull. Seismol. Soc. Am., 103*(2B), 1256–1274.

Maercklin, N., Festa, G., Colombelli, S., & Zollo, A. (2012). Twin ruptures grew to build up the giant 2011 Tohoku, Japan, earthquake. *Sci. Rep., 2*(709).

Marchuk, A. G. (2008). Minimizing computational error of tsunami wav-ray and travel time. *Sci. Tsunami Haz., 27*(4), 12-24 .

Mase, H., Yasuda, T., & Mori, N. (2011). Real-time prediction of tsunami magnitudes in Osaka Bay, Japan, using an artificial neural network. *J. Waterway, Port, Coastal Ocean Eng., 137*(5), 263–268.

Melgar, D., & Bock, Y. (2013). Near-field tsunami models with rapid earthquake source inversions from land- and ocean-based observations: The potential for forecast and warning. *J. Geophys. Res. Solid Earth, 118*, 1-17.

Minoura, K., & Nakaya, S. (1991). Traces of tsunami preserved in inter-tidal lacustrine and marsh deposits: Some examples from northeast Japan. *J. Geol., 265-287*, 99.

Minson, S., Simmons, M., & L., B. J. (2013). Bayesian inversion for finite fault earthquake source models I-Theory and algorithm. *Geophys. J. Int., 194* , 1701–1726.

Mori, N., Takahashi, T., Yasuda, T., & Yanagisawa, H. (2011). Survey of 2011 Tohoku earthquake tsunami inundation and run-up. *Geophys. Res. Lett., 38*, L00G14.

Muhari, A., Imamura, F., A., S., & Mas, E. (2012). Tsunami arrival time characteristics of the 2011 east Japan Tsunami obtained from eyewitness accounts, evidence and numerical simulation. *J. Nat. Disaster Sci., 34*, 91–104.

Mulia, I. E., Tay, H., Roopsekhar, K., & Tkalich, P. (2013). Hybrid ANNeGA model for predicting turbidity and chlorophyll-a. *J. Hydro. Environ. Res., 7*, 279-299.

Namekar, S., Yamazaki, Y., & Cheung, K. (2009). Neural network for tsunami and runup forecast. *Geophys. Res. Lett, 36*(L08604).

Nishimura, T., Munekane, H., & Yarai, H. (2011). The 2011 off the Pacific coast of Tohoku Earthquake and its aftershocks observed by GEONET. *Earth Planets Space, 63*, 631-636.

Nosov, M. A., & Kolesov, S. V. (2007). Elastic oscillations of water column in the 2003 Tokachi-oki tsunami source: in-situ measurements and 3-D numerical modelling. *Nat. Hazards Earth Syst. Sci., 7*, 243-249.

Okada, Y. (1985). Surface deformation due to shear and tensile faults in a half-space. *Bull. Seism. Soc. Am., 75*, 1435-1154.

Okal, E. A., & Synolakis, C. E. (2004). Source discriminants for near-field tsunamis . *Geophys. J. Int., 158*, 899-912.

Payne, J. L., & Eppstein, M. J. (2005). A hybrid genetic algorithm with pattern search for finding heavy atoms in protein crystals. *Proceedings of conference on genetic and evolutionary computation*, (pp. 377–384). Washington, DC, USA.

Peng, Y., Wang, S., Long, X., & Lu, B. L. (2015). Discriminative graph regularized extreme learning machine and its application to face recognition. *Neurocomputing, 149* , 340–353.

Perrin, C., Michel, C., & Andréassian, V. (2001). Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. *J. Hydrol., 242*((3-4)), 275-301.

Piatanesi, A., & Lorito, S. (2007). Rupture process of the 2004 Sumatra–Andaman Earthquake from tsunami waveform inversion. *Bull. Seismol. Soc. Am., 97*(1A), S223-S231.

Rao, C. R., & Mitra, S. (1971). *Generalized Inverse of Matrices and its Applications.* New York: Wiley.

Romano, F., Piatanesi, A., S., L., & Hirata, K. (2010). Slip distribution of the 2003 Tokachi-oki Mw 8.1 earthquake from joint inversion of tsunami waveforms and geodetic data. *J. Geophys. Res. , 115*(B11313).

Romano, M., Liong, S., Vu, M., Zemskyy, P., Doan, C., Dao, M., & Tkalich, P. (2009). Artificial neural network for tsunami forecasting. *J. Asian Earth Sci., 36*(1), 29–37.

Saito, T., & Furumura, T. (2009). Three-dimensional tsunami generation simulation due to sea-bottom deformation and its interpretation based on the linear theory. *Geophys. J. Int., 178*, 877-888.

Saito, T., Ito, Y., Inazu, D., & Hino, R. (2011). Tsunami source of the 2011 Tohoku-oki earthquake, Japan: Inversion analysis based on dispersive tsunami simulations. *Geophys. Res. Lett., 38*(L00G19).

Saito, T., Satake, K., & Furumura, T. (2010). Tsunami waveform inversion including dispersive waves: the 2004 earthquake off Kii Peninsula, Japan. *J. Geophys. Res., 115*(B06303).

Sanchez, A., & Cheung, K. F. (2007). Tsunami forecast using an adaptive inverse algorithm for the Peru-Chile source region. *Geophys. Res. Lett., 34*, L13605.

Satake, K. (1987). Inversion of tsunami waveforms for the estimation of a fault heterogenity: method and numerical experiments. *J. Phys. Earth, 35*, 241-254.

Satake, K., Baba, T., Hirata, K., Iwasaki, S.-I., Kato, T., Koshimura, S., . . . Terada, Y. (2005). Tsunami source of the 2004 off the Kii Peninsula earthquakes inferred

from offshore tsunami and coastal tide gauges. *Earth Planets Space, 57*, 173-178.

Satake, K., Fujii, Y., Harada, T., & Namegaya, Y. (2013). Time and space distribution of coseismic slip of the 2011 Tohoku earthquake as inferred from tsunami waveform data. *Bull. Seismol. Soc. Am., 103*(2B), 1473-1492.

Smola, A., & Kondor, R. (2003). Kernels and regularization on graphs. *In The Sixteenth Annual Conference on Learning Theory/The Seventh Workshop on Kernel Machines.* Berlin - Heidelberg.

Takagawa, T., & Tomita, T. (2012). Effects of rupture processes in an inverse analysis on the tsunami source of the 2011 off the Pacific Coast of Tohoku earthquake. *Proceedings of 22nd International Offshore and Polar Engineering Conference (ISOPE22)*, (pp. 14-19). Rhodes, Greece.

Tanioka, Y., Yudhicara, Kususose, T., Kathiroli, S., Nishimura, Y., Iwasaki, S.-I., & Satake, K. (2006). Rupture process of the 2004 great Sumatra-Andaman earthquake estimated from tsunami waveforms. *Earth Planets Space , 58*, 203–209.

Tappin, D. R., Grilli, S. T., Harris, J. C., Geller, R. J., Masterlark, T., Kirby, J. T., . . . MartinMai, P. (2014). Did a submarine landslide contribute to the 2011 Tohoku tsunami? . *Marine Geol., 357*, 344-361.

Torczon, V. J. (1997). On the convergence of pattern search algorithms. *SIAM Journal on Optimization, 7*(1), 1-25.

Tsushima, H., Hino, R., Fujimoto, H., Tanioka, Y., & Imamura, F. (2009). Near-field tsunami forecasting from cabled ocean bottom pressure data. *J. Geophys. Res., 114*(B06309).

Tsushima, H., Hino, R., Tanioka, Y., Imamura, F., & Fujimoto, H. (2012). Tsunami waveform inversion incorporating permanent seafloor deformation and its application to tsunami forecasting. *J. Geophys. Res., 117*, B03311.

Tsushima, H., Hirata, K., Hayashi, Y., Tanioka, Y., Kimura, K., Sakai, S., . . . Maeda, K. (2011). Near-field tsunami forecasting using offshore tsunami data from the 2011 off the Pacific coast of Tohoku earthquake. *Earth Planets Space, 63*, 821–826.

Voronina, T. (2011). Reconstruction of initial tsunami waveforms by a truncated SVD method . *J. Inv. Ill-Posed Problems, 19*, 615–629.

Wang, X. (2009). User Manual for COMCOT Version 1.7 (First Draft). Cornell University, Ithaca, NY, U.S.A.

Wang, X., & Liu, P. L. (2006). An analysis of 2004 Sumatra earthquake fault plane mechanisms and Indian Ocean tsunami. *J. Hydraul. Res., 44*, 147–154.

Wei, S., Graves, R., Helmberger, D., J.-P. A., & J. J. (2012). Sources of shaking and flooding during the Tohoku-Oki earthquake: A mixture of rupture styles. *Earth Planet. Sci. Lett., 333*, 91–100.

Wei, Y., Chamberlin, C., Titov, V. V., Tang, L., & Bernard, E. N. (2013). Modeling of 2011 Japan Tsunami: Lessons for near-field forecast. *Pure Appl. Geophys., 170*, 1309–1331.

Wells, D. L. and K. J. Coppersmith (1994), New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. Seismol. Soc. Am.*, 84, 974–1002.

Wetter, M., & Wright, J. (2003). Comparison of a generalized pattern search and a genetic algorithm optimization method. *Proceedings of the Eighth IBPSA Conference*, *III*, pp. 1401–8. NL: Eindhoven.

Willmott, C., Ackleson, S. G., Davis, R. E., Feddema, J. J., Klink, K. M., Legates, D. R., . . . Rowe, C. M. (1985). Statistics for the evaluation and comparison of models. *J. Geophys. Res. , 90*(C5), 8995-9005.

Wu, T. R., & Ho, T. C. (2011). High resolution tsunami inversion for 2010 Chile earthquake. *Nat. Hazards Earth Syst. Sci., 11*(12), 3251-3261.

Yasuda, T., & Mase, H. (2013). Real-Time tsunami prediction by inversion method using offshore observed GPS buoy data: Nankaido. *J. Waterway, Port, Coastal Ocean Eng., 139*(3), 221-231.

Zhao, D., Huang, Z., Umino, N., Hasegawa, A., & Kanamori, H. (2011). Structural heterogeneity in the megathrust zone and mechanism of the 2011 Tohoku-oki earthquake (Mw 9.0). *Geophys. Res. Lett., 38*(L17308).

Zhu, Q. Y., Qin, A., Suganthan, P., & Huang, G. B. (2005). Evolutionary extreme learning machine. *Pattern Recognition, 38*(10), 1759–63.