Estimate of tsunami source using optimized unit sources and including dispersion effects during tsunami propagation: The 2012 Haida Gwaii earthquake

Aditya Riadi Gusman¹, Iyan Eka Mulia², Kenji Satake¹, Shingo Watada¹, Mohammad Heidarzadeh¹,³, and Anne F. Sheehan⁴

¹Earthquake Research Institute, University of Tokyo, Tokyo, Japan, ²Ocean Alliance, University of Tokyo, Tokyo, Japan, ³Port and Airport Research Institute, Yokosuka, Japan, ⁴Department of Geological Sciences and Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado, USA

Abstract We apply a genetic algorithm to find the optimized unit sources using dispersive tsunami synthetics to estimate the tsunami source of the 2012 Haida Gwaii earthquake. The optimal number and distribution of unit sources gives the sea surface elevation similar to that from our previous slip distribution on a fault using tsunami data, but different from that using seismic data. The difference is possibly due to submarine mass failure in the source region. Dispersion effects during tsunami propagation reduce the maximum amplitudes by up to 20% of conventional linear longwave propagation model. Dispersion effects also increase tsunami travel time by approximately 1 min per 1300 km on average. The dispersion effects on amplitudes depend on the azimuth from the tsunami source reflecting the directivity of tsunami source, while the effects on travel times depend only on the distance from the source.

1. Introduction

Tsunami is a dispersive wave that may contain a wide range of wavelengths from a couple tens to several hundreds of kilometers. The longwave assumption breaks down and dispersion effect becomes significant at short wavelengths for earthquakes with a steep dipping fault plane [e.g., Gusman et al., 2009; Inazu and Saito, 2014] or submarine mass failures [e.g., Synolakis et al., 2002; Watts et al., 2003; Grilli and Watts, 2005; Løvholt et al., 2005; Tappin et al., 2014]. For such cases the Boussinesq equations are solved instead of the linear longwave equations [Tanioka, 1999; Saito et al., 2010; Kirby et al., 2013; Baba et al., 2015]. Weak dispersion at long period is due to the seawater compressibility, the elasticity of the Earth, and the gravitational potential effects [Watada et al., 2014] causes travel time delay relative to linear long wave and initial phase reversal at far field. This dispersion effect can be ignored for near-field (0–500 km) tsunami observations, and it is a common practice to generate tsunami synthetics by solving the linear shallow water equations [e.g., Titov et al., 2005; Fujii and Satake, 2006; Lorito et al., 2011; Gusman et al., 2012]. The dispersion effect must be considered when tsunami observations are in the midfield (500–2000 km) and far field (>2000 km) [Watada et al., 2014; Allgeyer and Cummins, 2014; Yoshimoto et al., 2016].

In this paper, the tsunami source of the 2012 Haida Gwaii earthquake (Mc 7.8) is first estimated using midfield (500–2000 km) tsunami observations. We apply a genetic algorithm (GA) to find the optimal number and distribution of unit sources. We describe the features in the optimum initial sea surface elevation model that is obtained by tsunami waveform inversion and compare the initial sea surface elevation with the ones computed from existing fault slip models inverted from tsunami and seismic waves. We explore the consequences of ignoring the dispersive effects in tsunami source estimation and tsunami wave prediction. To evaluate the dispersion effects on tsunami propagation, we compare the simulation results of linear wave and dispersive wave from the best tsunami source model in terms of maximum amplitude and travel time.

2. The 2012 Haida Gwaii Earthquake and Tsunami

An earthquake with moment magnitude (Mw) 7.8 occurred off Haida Gwaii, British Columbia, Canada, on 28 October 2012. The earthquake source mechanism [Lay et al., 2013], aftershock relocation [Kao et al., 2015], and its tsunami impact [Leonard et al., 2014; Fine et al., 2015] have been previously studied. The tsunami generated by this earthquake was recorded in near field at tide gauges, in midfield at Deep-ocean Assessment
and Reporting of Tsunamis (DART) buoy systems, the NEPTUNE cabled bottom pressure gauges, and bottom pressure gauges on an ocean bottom seismometer (OBS) array in the Cascadia subduction zone, and in far field at DART buoy systems [Lay et al., 2013; Fine et al., 2015; Sheehan et al., 2015; Gusman et al., 2016]. The bottom pressure gauges consist of Absolute Pressure Gauges (Lamont Doherty Earth Observatory, LDEO) and Differential Pressure Gauges (Scripps Institution of Oceanography, SIO, and Woods Hole Oceanographic Institution, WHOI) [Sheehan et al., 2015; Gusman et al., 2016]. The peak amplitudes at midfield DART and OBS stations ranged from 2 to 5 cm. In this study we use the midfield tsunami waveforms at eight DARTs and 19 LDEOs for tsunami waveform inversion (orange and blue circles in Figure 1). Then, we employ four tide gauge waveforms in the near-field, four WHOI waveforms in the midfield, and three DART waveforms in the far field for tsunami source model validation (green circles in Figure 1).

3. Methodology

3.1. Genetic Algorithm to Estimate the Initial Sea Surface Elevation

Without using earthquake fault parameters, initial sea surface elevation in the source region can be estimated by inversion of tsunami waveforms [Satake et al., 2005; Soito et al., 2010; Hossen et al., 2015; Mulia and Asano, 2015]. A combination of genetic algorithm (GA) methods for tsunami source inversion [Mulia and Asano, 2015, 2016] is used in this study to determine the initial sea surface elevation in the source region of the 2012 Haida Gwaii earthquake. The method uses a two-dimensional Gaussian shape water surface displacement with a characteristic horizontal wavelength of 40 km as a unit source inside the source area. Initially, we distribute 189 unit sources at 15 km equidistant interval covering the source area (green dots in Figure 2). Unlike most of other tsunami inversion techniques that fix the distribution of unit sources (Figure 2a), our GA uses the least squares method iteratively to find the optimal number and distribution of unit sources. In the first stage, the GA selects the optimal unit sources among the initial ones. This leads...
Figure 2. Initial sea surface elevations from (a) the initial source model with fixed unit sources (green dots), (b) the first stage source model for which GA reduced the number of unit sources (black dots) of the initial distribution, (c) the second (final) stage source model for which GA optimized the distribution unit sources (black dots), (d) a fault slip model of the 2012 Haida Gwaii earthquake estimated using tsunami waveforms [Gusman et al., 2016], (e) a fault model of the 2012 Haida Gwaii earthquake estimated using teleseismic waveforms [USGS], and (f) the submarine mass failure model, green rectangle indicates the failure area. The positive contour interval is 0.2 m, the negative contour interval is 0.1 m, and the bathymetric features are shown in the background.
to a reduction of the unit sources because the GA removes any unit source that has similar information in terms of surface height from the adjacent source points (black dots in Figure 2b) [Mulia and Asano, 2016]. In the second stage, the GA adjusts the locations of the selected unit sources from the first stage in order to further improve the waveform fit [Mulia and Asano, 2015]. The GA selects the next distribution of unit sources that produces a better waveform fit than the previous distribution. This is done iteratively until the conditions for convergence are met, which is when the number of GA generations is larger than 500 and the average fitness change over 50 GA generations is less than or equal to $1 \times 10^{-6}$. As a result, the spatial distribution of the unit sources will be scattered throughout the source area nonequidistantly (black dots in Figure 2c).

### 3.2. Cost Function

The cost function for the GA measures the fit between observed and synthetic seafloor pressure waveforms. We quantify the waveform fit based on a combination of root-mean-square error (RMSE) and Pearson correlation coefficient ($r$) [Mulia and Asano, 2015] (see Supporting Information S1). The correlation of the data is normalized as $R = 0.5(r + 1)$, so that it falls in the range of $[0, 1]$. The cost function ($E$) is a summation of RMSE and $R$ for all time windows, which can be written as follows:

$$E = \sum_{k=1}^{N} [\text{RMSE}_k + 1(1 - R_k)]$$

(1)

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

### 3.3. Synthetic Tsunami Waveforms

We construct two sets of tsunami Green's functions. The first Green's function is built from linear long waves that are produced by solving the linear shallow water equations [Satake, 1995]. The tsunami source model estimated by GA inversion using this Green's function is named as the LM source model. The second Green's function is built from synthetic tsunami waveforms that include the dispersive effects of the surface gravity wave and those imposed from the Earth model (i.e., the elasticity of the Earth, compressibility of seawater, and gravitational potential change due to water and Earth mass movement) [Watada et al., 2014]. The tsunami source model estimated by GA inversion using this Green's function is named as the DM source model.

The linear long wave is simulated by a finite difference method with a staggered grid scheme [Satake, 1995]. The size of the modeling grid is 1 arc min and the time interval is 1 s. To include the dispersion effects, the simulated linear long waves are corrected by a phase correction method [Watada et al., 2014]. The phase correction method keeps the linearity of tsunami waves and its computational cost is low. These features make the method suitable for building tsunami Green's functions for tsunami waveform inversion [e.g., Gusman et al., 2015; Yoshimoto et al., 2016].

During the second stage of GA inversion the locations of unit sources are moving within the area of the initial 189 unit sources. For every new location of unit source, synthetic tsunami waveforms at the stations are computed by applying nearest neighbor-weighted interpolation of waveforms from four nearest initial unit sources [Mulia and Asano, 2015]. The weights are determined by the distances from the new unit source location to the four nearest initial unit sources. For the final distribution of unit sources, the synthetic tsunami waveforms are computed by the method described in the previous paragraph.

### 3.4. Dispersive Tsunami Propagation Model

To simulate oceanwide dispersive tsunami propagation, the phase correction method is applied for all grids in the modeling domain. We first store the simulated linear long waves at all grids and then we apply the phase correction method to all of them. This process needs a large computer memory. For efficiency, we choose 15 s of time interval, 6 arc min for the computational grid size (the grid dimension is $701 \times 551$), and a total simulation time of 10 h. With this computation setup, we need to apply the phase correction method for 386,251 tsunami traces, and the matrix size to store the whole waveforms is $2400 \times 386,251$.

### 4. Results and Discussion

#### 4.1. Initial Sea Surface Elevation

For the DM source model, the GA produces an optimum distribution of 41 unit sources (black dots in Figures 2c and 3b). This optimum distribution is obtained after 1040 GA runs in the first stage and 2132 F3
Figure 3. Two possible instantaneous sea surface elevations for the 2012 Haida Gwaii earthquake tsunami. (a) Sea surface elevation of the source model (LM) estimated using linear longwave synthetics. (b) Sea surface elevation of the source model (DM) estimated using dispersive wave synthetics that consider the dispersive surface gravity wave and the Earth model. Green dots represent the initial unit source distribution, and the black dots represent the final unit source distribution which is estimated by the genetic algorithm. Stations in Figures 3c and 3d are used for the inversion, while stations in Figures 3e and 3f are used only for validation. Gray traces indicate observed tsunami waveforms. Blue traces indicate tsunami waveforms simulated by (c) the linear longwave propagation model and (e) the dispersive wave propagation model from the LM source model. (d, f) Red traces indicate tsunami waveforms simulated by the dispersive wave propagation model from the DM source model.
GA runs in the second stage. We show that our GA method has the ability to represent nonuniform distribution of unit sources. More unit sources are located around region B near the coast than around region A near the trench (Figures 2c and 3b). A tsunami source model with fixed number and location of unit sources (Figure 2a) failed to capture the complexity of the tsunami source.

Our preferred tsunami source model (DM) shows that secondary sea surface elevation in region B (Figure 2c) is distinctly separated from main sea surface elevation in region A near the trench. The main uplift region A has a maximum uplift of 1.1 m that is above the trench and the secondary uplift region B has a maximum uplift of 0.9 m that is located above the unique and complicated steep bathymetry near the Queen Charlotte Fault (Figure 2c). Our previous result for this event assuming a fault model [Gusman et al., 2016] shows a significant slip on the shallowest fault near the trench (which corresponds to region A) and bathymetric slope displacement effect near the coast (region B) (Figure 2d). We interpreted that the sea surface elevation near the coast was almost entirely from the horizontal motion of the steep slope, rather than vertical deformation from faulting [Gusman et al., 2016].

We compare our initial sea surface elevation with that from a fault slip distribution obtained by the USGS (United States Geological Survey) (http://earthquake.usgs.gov/earthquakes/eventpage/usp000juhz#) vertical deformation from faulting which was inverted from teleseismic body and surface waves. The initial sea surface elevation pattern near the trench between our model (Figure 2c) and the USGS model (Figure 2e) are similar, but around region B, the USGS model does not produce a sea surface elevation unlike our tsunami source model. One possible explanation is that the sea surface elevation is produced by a source mechanism that does not generate teleseismic waveforms, such as submarine mass failure (SMF) which may occur on a steep bathymetric slope [Grilli and Watts, 2005; Ma et al., 2013; Tappin et al., 2014]. Wide area in region B has bathymetric slope angle larger than 20° (Figure S1), which is one of the factors that make the region susceptible to slope failure [Varnes, 1984; Highland and Bobrowsky, 2008].

The USGS source model underestimates the observed amplitude of the first tsunami peak in the midfield by a factor of almost a half (Figure S2a). We attempt to add a SMF source with parameters of width = 40 km, length = 5 km, thickness = 250 m, slope = 15°, and slide direction = 225° by the method described in previous studies [Watts et al., 2005; Heidarzadeh and Satake, 2015]. This SMF model produces sea surface uplift near region B, similar to the DM source, but also produces subsidence, which is not modeled in the DM source (Figure 2f). The maximum and minimum sea surface deformations are 4 and −5 m, respectively. Simulation results show that the combined USGS fault slip model and the SMF model produce the larger tsunami peak amplitude at midfield stations (Figure S2b) than the USGS model, but the computed waveforms are not as close to the observations as the DM source. Therefore, this SMF model should not be considered as a realistic SMF model.

4.2. Dispersion Effects on Estimation of Tsunami Source

The initial sea surface elevation pattern that is estimated from the longwave tsunami Green’s function (LM source model) (Figure 3a) is different from the one estimated from Green’s function that includes the wave dispersion effects (DM source model) (Figure 3b). Both LM and DM source models have main and secondary uplift regions (A and B regions), but their size and locations are different. The LM source model has a maximum uplift near the trench (0.6 m) that is almost half of that in the DM source model. The locations of uplift regions appear to be pushed away from the stations distinctively at around the trench and also around region B (Figures 3a and 3b). This is mainly because the tsunami wave computed by the linear longwave approximation arrives earlier than the one that considers the dispersive effects.

The matches between the synthetic and observed tsunami waveforms, which are used in the inversions for the LM and DM source models, are equally good (Figures 3c and 3d). The tsunami waveform match for the DM source model (dispersive propagation model was used) is slightly better with a smaller root-mean-square error of 0.0103 m compared to 0.0106 m for the LM source model (linear propagation model was used). Although the waveform matches from the LM and DM source models are equally good, their sea surface elevation patterns are different (Figures 3a and 3b) as described above.

To validate the initial sea surface elevations, we compare the observed and simulated tsunami waveforms from the LM (Figure 3a) and DM (Figure 3b) source models at near-, middle-, and far-field stations that are not used in the inversion (Figures 3e and 3f). We first simulate the tsunami from the LM source model using...
the linear longwave propagation model to maintain the consistency (Figure S3). The simulated arrival time at the far-field DART stations are earlier than the observations. Then we use the dispersive propagation model for tsunamis from the LM and DM source models. We computed waveforms at (1) near-field tide gauge stations in British Columbia, Canada, which are located 0–500 km from the source; (2) midfield WHOI stations located 500–2000 km to the south of the source; and (3) far-field DART stations located more than 2000 km from the source and near the Hawaiian and Aleutian Islands (green circles in Figure 1).

The tsunami waveforms both from the LM and DM source models fit well the observations at the near- and middle-field stations (Figures 3e and 3f). The good fits at the midfield stations (WHOIs) are expected because the tsunami waveforms used in the inversions (DARTs and LDEOs) are located around these WHOI stations, and the effects of different propagation models are not very significant in the near field. The underestimation of the first peak amplitude by the LM source model becomes more significant as the travel time increases from middle- to far-field stations (Figure 3e). In the far field, the first peak of the simulated tsunami waveforms at D46408 and D46413 are underestimated (71% and 66% of the observation, respectively), although the first peak at D51407 fits fairly well the observation (Figure 3e). The tsunami waveforms at far-field stations from the DM source model match well the observations both in terms of timing and amplitude (Figure 3f). This result suggests that the DM source model is more reliable than the LM source model.

4.3. Dispersion Effects on Maximum Amplitude and Travel Time

We further explore the dispersion effects on tsunami propagation by using the DM source model (Figure 3b) for the linear longwave and dispersive wave simulations. In this experiment we simulate the tsunami for the wider region of the Pacific Ocean to measure the maximum tsunami amplitude and travel time distributions. The maximum tsunami amplitude distributions are compared by calculating the amplitude ratio between the one simulated by the linear longwave propagation model (Figure 4a) and the one simulated by the dispersive wave propagation model (Figure 4b). The travel time difference of the two tsunami propagation models is obtained by comparing the timing of the peak amplitude of the first wave cycles from the same source model.

The distribution of maximum amplitude ratios shows that the dispersive effects are more significant in the southwest direction (Figure 4c) perpendicular to the elongated shape of the tsunami source (Figure 3b), indicating the tsunami source directivity. Compared to the dispersive wave simulation, the linear longwave simulation produces up to approximately 20% higher amplitude in the southwest direction. This indicates that the tsunami propagating to the southwest azimuth has a range of wavelengths with various phase speeds and its shorter wavelength component has a slower propagation speed compared to the longer ones. As a result, these shorter wavelengths propagate behind the longer wavelengths (Figure S4 and Movie S2), thus reducing the overall maximum amplitude. The tsunami propagating to the south has a predominant long wavelength; therefore, the computation of wave amplitude by using the linear longwave approximation is valid even for a long distance as far as 5000 km (Figures 4c and S4, and Movies S1 and S2). The area of high-amplitude ratio becomes smoothly wider from the source in Haida Gwaii in the southwest direction to the shallow bathymetry around the Hawaiian and Aleutian Islands. Because of the complex and shallow bathymetry surrounding the Hawaiian and Aleutian Islands, the amplitude ratio patterns behind these island chains are rather complicated (Figure 4c).

The phase velocity of linear long wave is generally faster than the dispersive wave and the difference is the minimum at wave period of around 1000 s [see Watada et al., 2014, Figure 5a]. The differences of phase velocities become larger for both longer and shorter periods. As a result, the travel time difference between the linear longwave simulation and dispersive wave propagation become larger at location farther from the source region (Figure 4d). For the case of the 2012 Haida Gwaii earthquake the tsunami travel time difference is approximately 1 min for every 1300 km on average (Figures 4d and S5). This value can also be obtained from the phase velocities of the linear long wave of 198 m/s and the dispersive wave of 196 m/s when assuming an average ocean depth of 4 km and a wave period of 1000 s. Tsunami travel time delay relative to the linear long wave has been observed in previous studies [Rabinovich et al., 2011; Watada et al., 2014]. Our estimated travel time delay rate is consistent with the observed travel time delay at DART stations during the 2010 Chile and 2011 Tohoku tsunamis [Watada et al., 2014] (Figure S6). This is because waves with periods of 900–2000 s will arrive at almost the same time (as shown in the dispersive curve in Watada et al. [2014]).
their phase velocity is the fastest, and most tsunamis including the 2010, 2011, and 2012 events produced waves within the period range.

5. Conclusions

Optimizing unit sources using the genetic algorithm yielded two possible initial sea surface elevation models as a tsunami source of the 2012 Haida Gwaii earthquake. The first one is computed by the linear longwave propagation model (for the LM source model), and the second one contains the dispersive effects of surface gravity wave and the Earth model (for the DM source model). The initial sea surface elevations of the DM source model is more reliable because it can satisfy the observed tsunami waveforms at tide gauges and offshore pressure gauges in the near, mid, and far fields. The linear longwave synthetics from the LM source model does not predict the arrival times and amplitudes at the far-field stations well. Our preferred sea surface elevation model has two peaks similar to our fault slip inversion result using tsunami waveforms [Gusman et al., 2016]. Because the fault slip distribution from the seismic wave analysis only produced significant uplift near the trench, our preferred sea surface model (DM source model) may hint a submarine mass failure at the steep bathymetric slope near the Queen Charlotte Fault.

Figure 4. (a) Maximum amplitude distribution from the DM source model computed by the linear longwave propagation model. (b) Maximum amplitude distribution from the DM source model computed by the dispersive wave propagation model. (c) Ratio distribution between the linear and dispersive maximum amplitudes. (d) Travel time difference map between the first tsunami cycles of the linear long wave and dispersive wave. Contours represent great circle distances in kilometers from the earthquake’s epicenter (blue star).
Compared to the dispersive wave simulation, the linear longwave simulation produces up to approximately 20% higher amplitude to the southwest azimuth perpendicular to the elongated shape of the tsunami source. This shows the directivity effect on amplitude estimate which is dependent on the shape of tsunami source.

The dispersive effects of the surface gravity wave and the Earth model can reduce the maximum tsunami amplitude. The degree of amplitude reduction at a point of interest depends on the wavelength of predominant tsunami that pass through that point. The dispersion effects on amplitude reduction is more significant for shorter tsunami wavelengths. The dispersion effects also reduce the tsunami propagation speed. For the case of the 2012 Haida Gwaii earthquake the tsunami travel time delay relative to linear long wave due to the dispersion effects is approximately 1 min per 1300 km on average. This tsunami propagation speed reduction value is likely applicable to tsunamis propagating in the deep open ocean.

Acknowledgments

6The sea bottom pressure data at the Ocean Bottom Seismograph stations used in this research were provided by instruments from the Ocean Bottom Seismograph Instrument Pool (http://www.obsip.org) which is funded by the U.S. National Science Foundation. The facilities of the IRIS Data Management System were used to access the data used in this study. The sea level observations at the DART stations used in this research were provided by the National Oceanic and Atmospheric Administration and publicly available from http://nctr.pmel.noaa.gov/Dart/.

The sea level observations at the tide gauge station in Canada were provided by the Canadian Hydrographic Service http://www.tides.gc.ca/eng. We thank Alexander Rabinovich (Institute of Ocean Sciences, Canada) for providing processed tide gauge data at the Canadian stations. We thank Lisa Beal (Editor), an anonymous reviewer, and David Tappin (British Geological Survey, UK) for the constructive comments.

References


