AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2017GL076031

Key Points:
- The Green's Function based Time Reverse Imaging method is implemented in the 2009 Samoa tsunami triggered by a doublet earthquake
- An adjoint sensitivity method is applied to find the optimal set of stations for tsunami source inversion
- Results show that the GFTRI with adjoint method is able to recover the tsunami source associated with both normal and thrust faultings

Supporting Information:
- Supporting Information S1

Correspondence to:
M. J. Hossen,
mjhossen@eri.u-tokyo.ac.jp

Citation:
Hossen, M. J., Gusman, A. R., Satake, K., & Cummins, P. R. (2018). An adjoint sensitivity method applied to time reverse imaging of tsunami source for the 2009 Samoa earthquake. *Geophysical Research Letters*, *45*, 627–636. https://doi.org/10.1002/2017GL076031

Received 12 OCT 2017
Accepted 21 DEC 2017
Accepted article online 27 DEC 2017
Published online 17 JAN 2018

An Adjoint Sensitivity Method Applied to Time Reverse Imaging of Tsunami Source for the 2009 Samoa Earthquake

M. Jakir Hossen[1,2], Aditya Gusman[1], Kenji Satake[1], and Phil R. Cummins[3]

[1]Earthquake Research Institute, University of Tokyo, Bunkyo, Japan, [2]Department of Mathematics and Natural Sciences, BRAC University, Dhaka, Bangladesh, [3]Research School of Earth Sciences, Australian National University, Acton, ACT, Australia

Abstract We have previously developed a tsunami source inversion method based on "Time Reverse Imaging" and demonstrated that it is computationally very efficient and has the ability to reproduce the tsunami source model with good accuracy using tsunami data of the 2011 Tohoku earthquake tsunami. In this paper, we implemented this approach in the 2009 Samoa earthquake tsunami triggered by a doublet earthquake consisting of both normal and thrust faulting. Our result showed that the method is quite capable of recovering the source model associated with normal and thrust faulting. We found that the inversion result is highly sensitive to some stations that must be removed from the inversion. We applied an adjoint sensitivity method to find the optimal set of stations in order to estimate a realistic source model. We found that the inversion result is improved significantly once the optimal set of stations is used. In addition, from the reconstructed source model we estimated the slip distribution of the fault from which we successfully determined the dipping orientation of the fault plane for the normal fault earthquake. Our result suggests that the fault plane dip toward the northeast.

1. Introduction

The M_W = 8.0 2009 Samoa earthquake occurred at 17:48:10 GMT on 29 September 2009, triggering a local tsunami that reached Samoa and American Samoa in about 15 to 20 min, causing severe destruction in American Samoa, Samoa, and the northern Tonga island of Niuatoputapu, with most of the 189 fatalities and US$200 million in economic loss occurring in Samoa. This earthquake was caused by the crustal stresses associated with the subduction of the Pacific Plate beneath the Australian Plate.

Preliminary results indicated that this tsunami was triggered only by rupture on a fault in the outer rise, where the extensional stress associated with flexure of the subducting Pacific plate promotes failure with normal faulting mechanism parallel to the oceanic trench axis, as indicated by the Global CMT solution. Fritz et al. (2011) have studied the tsunami waveform recorded in Pago Pago Harbor considering both outer rise normal fault and megathrust sources and concluded that normal fault source alone can satisfactorily explain the tsunami waveform data.

However, other studies show that the 2009 Samoa earthquake involved both outer rise normal and megathrust fault rupture. Beavan et al. (2010) have shown that normal faulting on the outer rise region alone cannot explain estimates of the event's coseismic displacement made using campaign Global Positioning System (GPS) measurements on Niuatoputapu. Lay et al. (2010) have also found that seismic waveforms generated by the event cannot be explained by the outer rise normal faulting alone, but are consistent with two interplate underthrusting subevents occurring within 100 s of the initial outer rise earthquake rupture. To examine further the causes of this tsunami, Tonini et al. (2011) have studied five candidate normal fault mechanisms determined by USGS and GCMT and concluded that none of them are by themselves able to provide a good fit to the observed tsunami waveforms. Finally, Fan et al. (2016) used backprojection of seismic waves to show that the 2009 Samoa earthquake involved rupture of both an outer rise normal fault and the Tonga Trench megathrust.

Knowledge of the slip on a fault of known geometry can be used to calculate tsunami generation by using elastic dislocation theory (e.g., Okada, 1985) to model the deformation of the sea floor caused by the fault

rupture and translating this to the sea surface displacement that acts as a tsunami source. Many studies have used tsunami waveform data to estimate fault slip by inverting this relationship between fault slip and sea surface displacement (Fujii et al., 2011; Gusman et al., 2015; Heidarzadeh et al., 2016; Newman et al., 2011; Satake et al., 2013; Satake, 1987, 1989; Wei et al., 2013). This procedure requires knowledge of the fault geometry and is particularly problematic if, as is the case for the 2009 Samoa earthquake, it is not clear what fault ruptured to generate the observed tsunami. It is also possible to use tsunami data to estimate the tsunami source (i.e., initial sea surface displacement) directly, without making any assumption about fault rupture at all (e.g., Baba et al., 2005; Ho et al., 2017; Hossen et al., 2015b). Recently, the Time Reverse Imaging (TRI) method, widely used in medical science and acoustics, has been applied in tsunami science to estimate the tsunami source models (An & Meng, 2017; Hossen, Cummins, Roberts, et al., 2015; Hossen et al., 2017) from tsunami data alone, and has been shown capable of reconstructing the tsunami source with few if any assumptions about the nature of the source—that is, the tsunami can be generated by any fault or set of faults, or by some nonseismic source such as a submarine landslide.

However, the performance of the TRI method depends on the choice of stations located surrounding the source. Hossen et al. (2017) carried out a sensitivity analysis of this method and found that the inversion result depends on the stations located on the area through which most tsunami energy is propagated. An and Meng (2017) applied the method to the 2015 Illapel tsunami event and found that only four stations out of 24 provides an acceptable image of the source; but the source becomes unrealistic if all the 24 stations are used. In addition, some studies suggested that increasing the number of observations used in the inversion may deteriorate the inversion result as the ocean bottom inhomogeneities and data noise may have significant impact on the inversion (e.g., Mulia et al., 2017; Voronina, 2016). Therefore, instead of using all available data, it is important to choose the station judiciously in order to obtain a realistic source model.

In this paper, we introduced an adjoint sensitivity (AS) method developed by utilizing the concept of TRI to identify the optimal set of stations for tsunami source inversion so as to obtain an accurate source model. We used a TRI-derived method known as "Green's function-based time reverse imaging (GFTRI)," developed by Hossen et al. (2015a) to estimate the source model for the 2009 Samoa earthquake tsunami triggered by a doublet earthquake having both normal and thrust fault character. They applied this GFTRI method with near-field waveforms of the 2011 Japan tsunami to estimate an initial sea surface displacement source model that is reliable enough to predict far-field tsunami waveforms with excellent agreement to those observed. Although Hossen et al. (2017) improved this method by considering cross correlation among GFs and obtained a better result using near-field data, we did not apply this improved method to the 2009 Samoa event because the method is not practical for far-field stations; but the distant stations are essential for the Samoa source inversion due to the lack of a sufficient number of near-field recordings.

2. Methodology

In this study, we applied the GFTRI method, described in Hossen et al. (2015a), to tsunami waveforms of the 2009 Samoa event. As a requirement, we divided the source region into a number of source "patches." For each source patch we computed GFs by approximating a point source using a box-type basis function with cosine tapering as defined in Hossen et al. (2015a), where sea surface displacement is 1 m near the center and smoothly tapers to zero to the outside of the source patch. This type of source has the advantage of being sharp enough to approximate a point source, but smooth enough for the numerical tsunami propagation model to be stable. The dimension of each source patch is 25×25 km, which makes in total about 144 source grid points for the source area of the 2009 Samoa tsunami, whose dimension is 300×300 km.

Green's functions (GF) associated with each unit source were computed by running a parallel numerical tsunami propagation model, named JAGURS (Baba et al., 2015), which solves the nonlinear Boussinesq dispersive equations in spherical coordinates with the respective initial source model. For the source estimation of the 2009 Samoa tsunami, we considered stations located in the central, west and east Pacific Ocean, which lie within the domain 140°E–120°W and 35°S–25°N, the computational domain of our simulation.

Two nested grids were utilized in order to obtain more accurate simulated waveforms at near-shore stations. The coordinates bounding the outer domain are (140°E–120°W, 35°S–25°N) and its resolution is 90 arcsec; those bounding inner domain are (167°E–148°W, 25°S–7°S) and its resolution is 30 arcsec. JAGURS has both linear and nonlinear options with dispersion. In GF computation, dispersion was neglected because it is computationally very expensive. We considered the linear model for the outer domain, but for the inner one,

a nonlinear model was used as all the tide gauges are located in this domain. We set the total simulation time to be 43,200 s (12 h) and used time steps $\Delta t = 1$ s, so that the stability condition is satisfied.

The GFs were then used to estimate the amplitude of each unit source by the GFTRI method. In this method, the waveforms are convolved with the corresponding time-reversed observed waveforms recorded at the same observation location and then scaled using the squared norm of the GFs (see equation (3)). This scaling is an important feature of GFTRI that overcomes one of the limitations of the original TRI method as described by Hossen, Cummins, Roberts, et al. (2015). Also, the formulation used for amplitude estimation requires an adjustment in time τ associated with tsunami propagation from the source patch (for details, see Hossen et al., 2015a). We determined the appropriate adjustment by conducting some experiments with a range of source radii. The details are given in section 4. Finally, the amplitude of the basis function at zero lag is estimated by averaging all the scaled reversed wavefields propagated back to the source area and picking up the value at the origin time, the time when the tsunami is originated, after adjusting for the time shifting. The advantage of this method is that it estimates the source directly from the information recorded by tsunami observation without additional constraints—damping or smoothing usually required by the conventional source inversion method.

Note that the quality of the reconstructed image depends on the number of stations and their location. Thus, it is necessary to consider the optimal set of observations that provides better source image. In order to determine the optimal set, we developed the AS method, which is similar to the GFTRI method. The only difference is that instead of using reversed observed wavefield, the AS method requires the reversed residual wavefield, the difference between observed and predicted waveforms. The theory of the AS method is given below.

2.1. Theory of Adjoint Sensitivity Method

The adjoint sensitivity method has been widely used in atmospheric science (numerical weather prediction, air quality forecasting, etc.) to find optimal locations for adaptive observation for improving the forecasting result (Daescu & Carmichael, 2003; Fourri'e et al., 2002; Sandu et al., 2005). The method is based on the derivative of a misfit function. Let the frequency-domain misfit function be defined as in Hossen et al. (2015a) to minimize the difference between predicted and observed waveforms at observation position $\mathbf{x}_j$ ($j = 1, 2, \ldots, p$):

$$\mathcal{J}_j(\omega) = \left| \sum_{i=1}^{n} G_i(\mathbf{x}_j, \omega) a_i(\omega) - \mathrm{d}_j(\omega) \right|^2 \tag{1}$$

where $G_i(\mathbf{x}_j, \omega)$ is the Green's function at $\mathbf{x}_j$ computed using point source at the ith of a set of n source points s_i, $d_j(\omega)$ represents the waveform recorded at $\mathbf{x}_j$, and $a_i(\omega)$ is the amplitude of the ith unit source. Assuming that rupture is instantaneous at time τ, that is, $a_i(t) = a_i\delta(t - \tau)$, where $\delta(t)$ is a Dirac delta function, Hossen et al. (2015a) showed that the derivative of the misfit function equation (1) with respect to the amplitude a_i is given by:

$$\frac{\partial \mathcal{J}_j}{\partial a_i} = G_i^*(\mathbf{x}_j, \omega) \left[\sum_{i'=1}^{n} G_{i'}(\mathbf{x}_j, \omega) a_{i'}(\omega) - \mathrm{d}_j(\omega) \right] \tag{2}$$

Hossen et al. (2015a) also showed how a single-station estimate of the tsunami source could be obtained by "backprojecting" the waveform $\mathrm{d}_j(t)$ observed at $\mathbf{x}_j$, which in the frequency domain is expressed as follows:

$$\hat{a}_{ij}(\omega) = \frac{1}{|G_{ij}|^2} G_i^*(\mathbf{x}_j, \omega) \mathrm{d}_j(\omega) \tag{3}$$

where the superscript asterisk denotes the complex conjugate; $\hat{a}_{ij}(\omega)$ denotes the estimate of $a_i(\omega)$ that minimizes the misfit with respect to jth observation; $|G_{ij}|^2 = \int |G_i(\mathbf{x}_j, t)\|^2 \mathrm{d}t$. Here we only consider autocorrelations as we assume that the unit sources at s_i and $s_{i'}$ are well separated.

We use this same approach to estimate the contribution of the data residual at $\mathbf{x}_j$ to the model error. The data residual at $\mathbf{x}_j$ is given in the frequency domain by

$$R_j(\omega) = \sum_{i=1}^{n} G_i(\mathbf{x}_j, \omega) \hat{a}_i(\omega) - \mathrm{d}_j(\omega) \tag{4}$$

where $\hat{a}_i(\omega)$ denotes the estimate of $a_i(\omega)$ with respect to all observations. Comparison of the above equation with the derivative of the misfit function in equation (2) shows that the derivative is equivalent to the back-projection of the residual onto the model space. We use this to define the contribution of this residual to the model error by backprojecting it to the model space as

$$\delta a_{ij}(\omega) = \frac{1}{|G_{ij}|^2} G_i^*(\mathbf{x}_j, \omega) R_j(\omega) \tag{5}$$

When $\delta a_{ij}(\omega)$ is transformed to the time domain and evaluated at the time of the instantaneous rupture τ, equation (5) becomes:

$$\delta a_{ij} = \frac{1}{|G_{ij}|^2} G_i(\mathbf{x}_j, t) * R_j(-t)|_\tau \tag{6}$$

where the asterisk denotes convolution and $G_i(\mathbf{x}_j, -t) = G_i(\mathbf{x}_j, t)$ because of reciprocity (see Hossen et al., 2015a for details). That is, this equation can only be used to compute the model error if the principle of reciprocity holds.

Equation (6) allows us to compute the effect that each station residual in the observation space, $R_j(t)$, has on the model space. That is, the misfit error is propagated back to the model space to compute the station's model error. This allows us to select an optimal subset of observations to invert for a source model, by successively removing stations that make the largest contribution to the model error. To remove the stations, we calculated the error in the model space contributed by the jth station as

$$\delta E_j = \sqrt{\frac{1}{n} \sum_{i=1}^{n} |\delta a_{ij}|^2} \,. \tag{7}$$

To assess the error associated with each model, we calculated spatial distribution of the error over the model space for each inversion as

$$\delta a_i = \frac{1}{p} \sum_{j=1}^{p} \delta a_{ij} \,. \tag{8}$$

The total error contributed by all stations for each inversion is

$$\delta E = \sqrt{\frac{1}{n} \sum_{i=1}^{n} |\delta a_i|^2} \tag{9}$$

and stored them in order to determine the optimal set of stations.

This AS method was used to compute the error in energy norm in the model space, referred as "model error" hereafter, by using equation (9), which is actually the average of errors calculated by equation (7) for each individual station. The station with the largest contribution to the model error is considered as a removal for the next iteration. In the next iteration, the removal is discarded from the set, the source inversion is repeated, and a new model error is calculated. Once the process was completed, we chose the subset as the optimal set of stations that provided the smallest error in the model space.

3. Observations

The 2009 Samoa earthquake tsunami was recorded by many DART (Deep Ocean Assessment and Reporting of Tsunamis) buoys and tide gauges located in the Pacific Ocean. We considered six DART buoys (51406, 51407, 51425, 51426, 54401, and 52401) and eight tide gauges (122, 125, 127, 139, 140, 144, 152, and 348) to estimate the tsunami source model. The DART system is operated by the U.S. National Oceanic and Atmospheric Administration (NOAA), and data are available from the National Geophysical Data Center, NOAA (http://www.ngdc.noaa.gov/hazard/dart/2009samoa_dart.html). Tide gauge data were collected from the website of National Tsunami Warning Center, NOAA (http://wcatwc.arh.noaa.gov/previous.events/09-29-09-Samoa/09-29-09.htm). Among these data, we considered eight tide gauges and six DART buoys, which clearly recorded tsunami signals, for our source inversion. The locations of stations are displayed in Figure 1. In addition, we used some DART buoys located at far-field stations to validate the estimated source model reconstructed by the GFTRI method (Figure S8 of the supporting information).

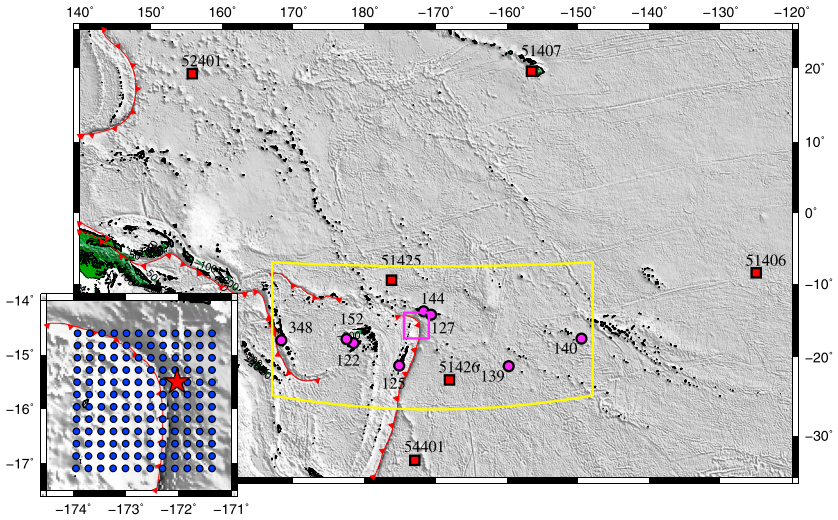

Figure 1. Map showing the computational domain (outer domain for coarse resolution) that covers the location of tide gauge (circle) and DART buoys data (square). The yellow box shows the domain for the inner nested grid (high resolution). In the inset we show the source grid points at which unit sources are created to compute Green's functions. The red star shows the epicenter and red-toothed line shows the trench axis.

To process the tsunami recordings in both DART buoys and tide gauges, we first carefully removed the gaps and repeating values and interpolated them in 30 s sample intervals. We then applied detiding by estimating tidal signals using a polynomial fit and removing them from the original records. For polynomial fitting, a polynomial of degree 12 was used for which we used a least squares fit. Apart from detiding, we filtered data using Butterworth filtering with low and high cutoff frequencies of 0.0002777 Hz (1 h) and 0.016666667 (1 min), respectively.

4. Results

Several experiments were carried out using synthetic and actual observations in order to estimate the source model using the GFTRI method, which is much simpler and more efficient than the traditional least squares method. The idea of conducting synthetic experiment was to understand the capability of this method in order to resolve the source complexity of a complex outer-rise event, such as the 2009 Samoa earthquake. The details of this experiment are given in Text S2 of supporting information. The estimated source model looks very similar to the original one and the waveform agreement between the synthetic observation and computed waveforms is excellent (Figure S1). The method is capable of recovering the source that reproduces the waveforms having good agreement with the synthetic observed waveforms.

For analysis of the actual data recorded following the 2009 Samoa earthquake, we considered 14 stations to estimate the tsunami source model. For both actual and synthetic experiments, we used full-waveforms whose length is the same for all stations (up to 500 min), though some are located near to the epicenter and some are far away from it. The tsunami reaches the near stations in a short time (less than an hour), but takes a longer time (about 8 h) to reach the distant stations. Correction of tsunami arrival times, to account for elastic loading and density stratification (explained by Allgeyer & Cummins, 2014; Watada et al., 2014), is necessary for modeling the far-field observations. Gusman et al. (2016) found that the travel time delay is about 1 min per 1,300 km which is about 0.92% of the time taken by the tsunami to reach to the observation locations. GFs associated with distant locations were corrected by adjusting the time delay of the modeled tsunami waveforms accordingly.

In addition to the adjustment of travel times to all distant stations to account for elastic loading and density stratification, we also adjusted the times of all GFs to account for the time of tsunami propagation across the spatial extent of our unit sources, as described in Hossen et al. (2015a). To find the appropriate adjustment τ, we conducted several experiments using a range of effective radii $r = 25, 20, 12, 8,$ and 4 km, which represent the size of the source area where initial water displacement is mostly concentrated and corresponding time corrections τ calculated according to $\tau = r/\sqrt{gh_i}$, where g is the acceleration of gravity and h_i is water depth

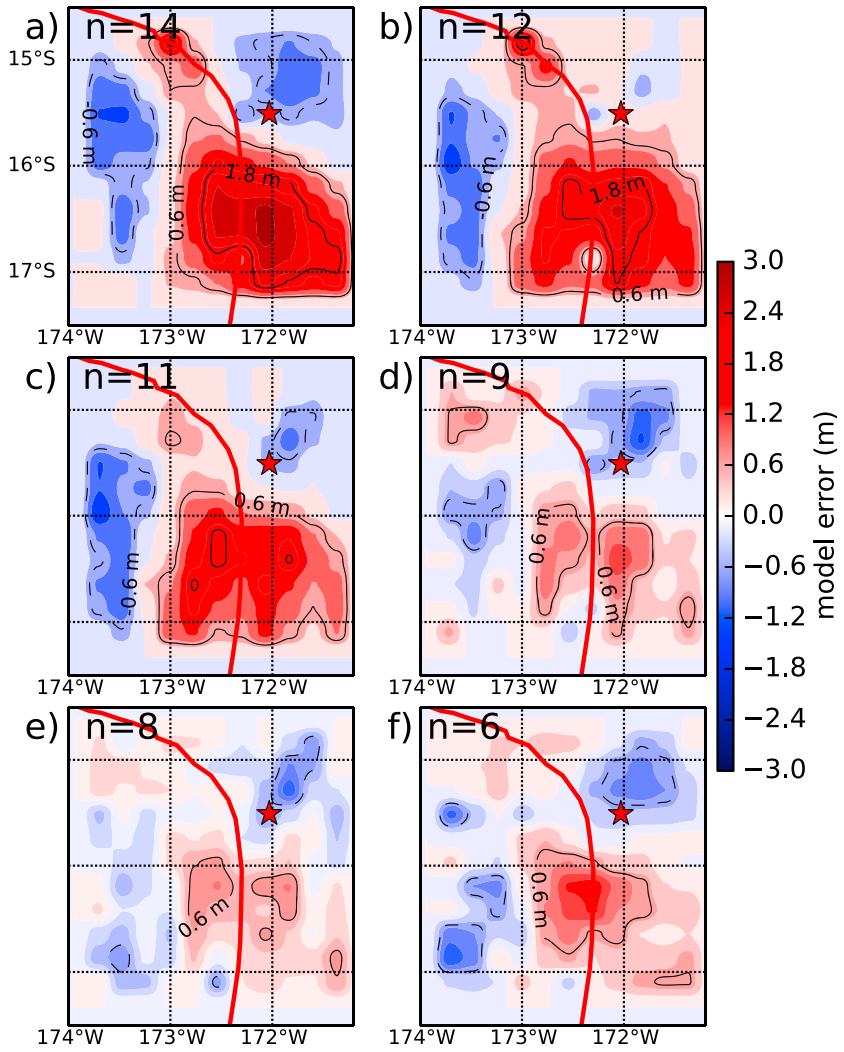

Figure 2. Spatial distribution of model error computed using different set of stations. Result shows that the smallest error occurred when the optimal eight stations are used.

at source grid point $\mathbf{x}_i$. We calculated the misfit error for each source inversion with these radius values and present the corresponding misfit values in Figure S2 of supporting information. We found that the appropriate time adjustment to account for source finiteness corresponds to 4 km radius, as this provides the smallest misfit value. This time adjustment is used for both synthetic and actual experiments.

For actual experiments, we first carried out source inversions with all 14 stations. The source model and waveform fits are presented in Figure S3 of the supporting information. Although the maximum uplift occurs to the west of the trench line, the source model (Figure S3a) seems unacceptable as water elevation extends to the east beyond the trench line. Moreover, the model error distribution (Figure 2) shows that the uncertainty in the source model is very high. Though waveform fits are good at some tide gauge stations, modeled waveforms overestimate the data at stations located to the south of the source (Figure S3b). This suggests that data at some stations are influenced by factors our modeling does not account for, such as complex shallow bathymetry near coastal tide gauge stations. It is necessary to determine the stations that degrade the inversion result and remove them from the source inversion. We applied the AS method (described above) that enables us to determine which stations have a negative impact on the inversion result.

In the AS method, the iterative source inversion was started with all the 14 stations and continued up to 5 stations remained after successively discarding the removal that yields the largest error to the inversion result at

each iteration. We present the model errors calculated for each set of stations used in that particular source inversion in Table B1 of supporting information. The table also shows the stations that make the largest contribution to model error and are, therefore, removed from the set of stations used in the subsequent inversion. We also display the model error for each set in Figure S4 which shows that the inversion with eight stations produces the smallest error. Further, we present the spatial distribution of model errors estimated by using different numbers of stations (Figure 2). The results show that the use of all the 14 stations provides the largest error over an extended area. The error is reduced significantly if the optimal set of eight stations (122, 127, 139, 140, 51425, 51426, 54401, and 51407) determined by the AS method is used in the source inversion.

Using this optimal set of observations, we estimated the source model and obtained the corresponding waveform agreement shown in Figure 3. We found that the result improves significantly; the source model looks more realistic and the agreement between observed and predicted waveforms is much better, especially at the DART buoys. However, the fit is a bit poor at tide gauge stations in the sense that there is a mismatch in arrival time and wave height, which might be due to the local bathymetry and lower-resolution grid used in simulation. In order to reproduce the waveform with good agreement, it may require to use more accurate bathymetry with high-resolution grid. The source model shows that the water height distribution is more concentrated than the one using nonoptimal stations; the maximum uplift is located in the subduction zone to the west of Tonga trench. We also present the comparison of waveforms computed using optimal (8 stations) and nonoptimal (14 stations) set in Figure S5. The waveform comparison shows that the usage of optimal set significantly improves the fit, especially at the DART stations.

5. Slip Distribution

We estimated the slip distributions on two fault geometries from the reconstructed source model with the optimal set of stations for the 2009 Samoa earthquake doublets consisting of a model with normal fault mechanism and a model with thrust fault mechanism. For the normal fault model (N model), we evaluated both nodal planes and choose the one that can best explain the estimated initial sea surface displacement. We distributed 6×2 subfaults with fault parameters obtained by a previous study (Lay et al., 2010) of strikes of 144°/315°, dips of 65°/25°, and rakes of −86°/−99°. For the thrust fault, denoted by T, we assumed a strike of 180°, dip of 29°, and rake of 90° (T model with 7×5 subfaults); the dip angle is from Lay et al. (2010); the strike angle is based on the orientation of the trench, and the rake angle is for a pure thrust faulting. The subfault size for all fault models is 25×25 km (Figure 3b). The grid interval of sea surface displacement is 30 arcsec within 185° – 195°E and 14° – 18°S. The total number of grids used for estimating the fault slip distribution for these models was 289,081.

Each subfault with a unit slip was used to generate synthetic sea surface displacements by the equations in Okada (1985) over the same domain as the source model. We assumed that the sea surface displacement is identical to the seafloor displacement. We stored the generated displacement from each subfault in a column of a matrix A and the water elevation from the reconstructed source model as a vector b. This led to a linear system: $Ax = b$, where x is the slip amounts on the subfaults, determined by using a nonnegative least squares algorithm (Lawson & Hanson, 1995). A spatial smoothness constraint for the slip distribution is applied through a Laplacian operator and a smoothness factor of 0.4 is used. The slip distribution for all the faults is displayed in Figure 3b.

Our results suggest that the normal fault plane is the one that dips toward the northeast at an angle of 25°. This is consistent with the findings of Beavan et al. (2010) based on displacements at continuous GPS stations in Samoa and central Tonga. The fault slip model with northeast dipping orientation (strike/dip/rake = 315°/25°/−99°) can reproduce the initial sea surface displacement much better than the southwest dipping plane because the slip model with southwest dipping plane can only reproduce a small fraction of the estimated sea surface subsidence (Figure S7). Figure 3b also shows that the maximum slip for normal faulting is about 2 m and for thrust faulting is 4 m. It seems that the slip is smaller than the result suggested by other studies, for instance, Beavan et al. (2010). In addition, seismic moment calculated from the slip model for the normal fault earthquake with the preferred northeast dipping orientation is 1.18×10^{20} N m (M_w 7.4), by assuming a rigidity of 4×10^{10} N m^{-2}. The seismic moments calculated from the fault slip model with thrust fault mechanism is 7.33×10^{20} N m (M_w 7.9). Therefore, the total calculated seismic moment for the 2009 Samoa earthquake is 8.51×10^{20} N m which is equivalent to M_w 7.9.

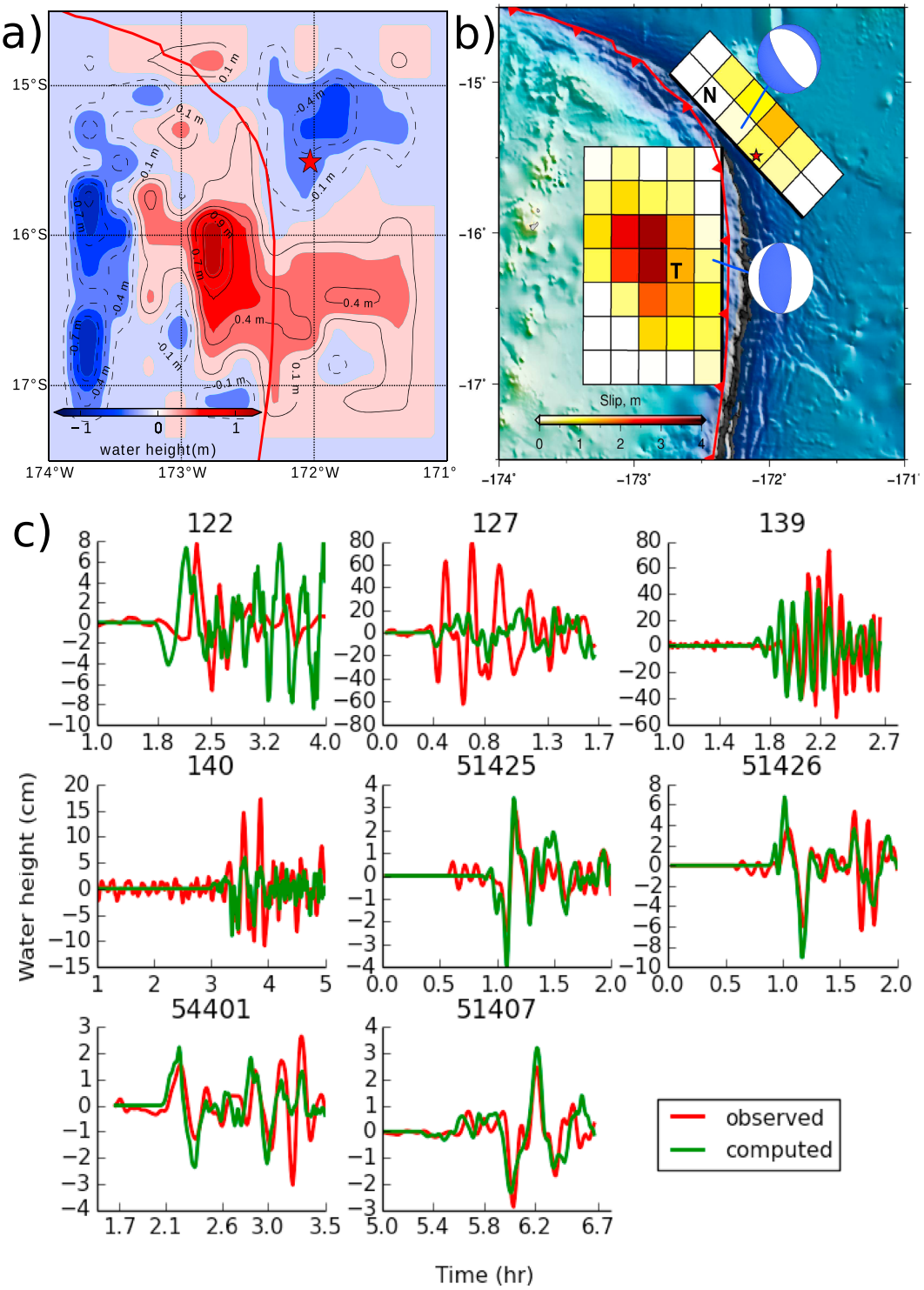

Figure 3. (a) The source model reconstructed by using the GFTRI method with the optimal set of eight stations and from this reconstructed source we estimated slip distributions on (b) two fault plane. Slip distribution results suggest that the normal fault plane dip toward the northeast (Figure S7). (c) The agreement between observed (red) and computed waveforms (green) obtained by using source model (Figure 3a).

6. Discussion

All the inversions were carried out using full waveforms, that is, the first wave with later arrivals, since only a few observations near to the source region within a reasonable range are available. However, we conducted some experiments using only the first wave recorded by the optimal set of stations. The length of data is from the tsunami origin time to the time when the first wave is recorded, and the length varies significantly because of sparse distribution of stations. The source model and waveform fits are presented in Figure S6 of supporting information; the model with the first wave shows that water elevation is more extended beyond the trench line. Further, the waveform fits are poorer at DART stations though the fit is good at some tide gauges. Overall, the performance of the method with first wave is not satisfactory; only the first wave might be sufficient for inversion if more observations were available nearby the source (Hossen et al., 2017).

We also computed fault slip distributions from the estimated sea surface displacement in Figure 3a and present the result in Figure 3b. The result shows that the 2009 earthquake was associated with rupture on both faults—normal and thrust, which is consistent with the back projection result presented in Lay et al. (2010). Though the choice of which conjugate plane is the actual rupture plane for large, well-recorded outer rise normal faulting earthquakes is often poorly determined by seismic data (see, e.g., Lay et al., 2009), we were able to distinguish which of the conjugate planes associated with such an earthquake using tsunami data. Our result shows that only rupture on the fault corresponding to the conjugate plane that dips toward the northeast can produce the same sea surface displacement (Figure S7a) as the reconstructed source model. Using this source model (Figure S7a) we computed waveforms at near- and far-field stations. In Figure S8 of supporting information, we displayed the waveform fit between observed and predicted waveforms from both models: the reconstructed sea surface model and the model from the slip distribution. The data fit looks comparable to the case of the reconstructed source, which suggests that such a two-step inversion is feasible.

7. Conclusion

In this paper, we applied the adjoint sensitivity method to identify the optimal set of stations for the GFTRI method. We utilized the method to estimate the source model of the complex event, the 2009 Samoa earthquake and tsunami. Initially, we carried out a synthetic experiment to understand whether the method is capable of recovering the source model of a complex event, such as the 2009 Samoa earthquake. In the synthetic experiment, we found that the method has the ability to recover the source associated with both normal and thrust faulting. The recovered source model looks very similar to the original model (Figure S1).

After gaining confidence from the synthetic experiment, we implemented this method with actual observations. The result shows that the method recovers the details of source model associated with the different faults, both an outer rise normal fault and the megathrust, which both appear to have ruptured in this event. The usage of optimal set of stations improves the result significantly, especially in the source model (Figure 3). The source model looks more realistic as the maximum uplift over the thrusting region is concentrated near to but west of the Tonga trench. The model error distribution (Figure 2) shows that the error in the source model is reduced significantly due to the use of optimal stations.

The waveform fits are also improved, with the wave heights and periods better fit, especially in the DART buoy data. However, the fit at some stations, for instance, station 127 remains the same though the misfit error looks very high. The method is unable to discard it. In spite of this drawback, we found that the method is able to reconstruct the source associated with a complex event that produces waveforms at far-field stations having an excellent agreement with observed waveforms (Figure S8). The peak amplitude and wave length of the computed waveforms, especially in the first wave, is almost the same as the observed one. In the same manner, the source model from the slip distribution of both faults can reproduce the waveform that explains the observations well (Figure S8).

Nevertheless, the fault slip of the events seem to be smaller than the result suggested by other studies, for instance, Beavan et al. (2010). The reason for getting smaller slip could be that the source model is smoothed out due to the use of long time series data, which contained later arrivals including reflected or reflected waves and these later arrivals may involve energy loss in shallow coastal waters. In future, we would like to carry out a further study on the effect of such attenuated arrivals on the source model using this algorithm in order to improve the performance so that it can be used for with a wider range of tide gauges.

Acknowledgments
In this study, we used tsunami waveform data (DART buoys and tide gauges) operated by the National Oceanic and Atmospheric Administration (NOAA). This research has been financially supported by Japan Society for the Promotion of Science (JSPS) KAKENHI grant 16H01838 and Australian Research Council Discovery Project DP120103207.

References

Allgeyer, S., & Cummins, P. (2014). Numerical tsunami simulation including elastic loading and seawater density stratification. *Geophysical Research Letters*, *41*, 2368–2375. https://doi.org/10.1002/2014GL059348

An, C., & Meng, L. (2017). Time reversal imaging of the 2015 Illapel tsunami source. *Geophysical Research Letters*, *44*, 1732–1739. https://doi.org/10.1002/2016GL071304

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.

Baba, T., Takahashi, N., Kaneda, Y., Ando, K., Matsuoka, D., & Kato, T. (2015). Parallel implementation of dispersive tsunami wave modeling with a nesting algorithm for the 2011 Tohoku tsunami. *Pure and Applied Geophysics*, *172*(12), 3455–3472. https://doi.org/10.1007/s00024-015-1049-2

Beavan, J., Wang, X., Holden, C., Wilson, K., Power, W., Prasetya, G., … Kautoke, R. (2010). Near-simultaneous great earthquakes at Tongan megathrust and outer rise in September 2009. *Nature*, *466*(7309), 959–963. https://doi.org/10.1038/nature09292

Daescu, D. N., & Carmichael, G. R. (2003). An adjoint sensitivity method for the adaptive location of the observations in air quality modeling. *Journal of the Atmospheric Sciences*, *60*(2), 434–450.

Fan, W., Shearer, P. M., Ji, C., & Bassett, D. (2016). Multiple branching rupture of the 2009 Tonga-Samoa earthquake. *Journal of Geophysical Research: Solid Earth*, *121*, 5809–5827. https://doi.org/10.1002/2016JB012945

Fourri'e, N., Doerenbecher, A., Bergot, T., & Joly, A. (2002). Adjoint sensitivity of the forecast to TOVS observations. *Quarterly Journal of the Royal Meteorological Society*, *128*(586), 2759–2777.

Fritz, H. M., Borrero, J. C., Synolakis, C. E., Okal, E. A., Weiss, R., Titov, V. V., … Liu, P. L.-F. (2011). Insights on the 2009 South Pacific tsunami in Samoa and Tonga from field surveys and numerical simulations. *Earth-Science Reviews*, *107*(1–2), 66–75. https://doi.org/10.1016/j.earscirev.2011.03.004

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 and Space*, *63*(7), 815–820.

Gusman, A. R., Murotani, S., Satake, K., Heidarzadeh, M., Gunawan, E., Watada, S., & Schurr, B. (2015). Fault slip distribution of the 2014 Iquique, Chile, earthquake estimated from ocean-wide tsunami waveforms and GPS data. *Geophysical Research Letters*, *42*, 1053–1060. https://doi.org/10.1002/2014GL062604

Gusman, A. R., Mulia, I. E., Satake, K., Watada, S., Heidarzadeh, M., & Sheehan, A. F. (2016). Estimate of tsunami source using optimized unit sources and including dispersion effects during tsunami propagation: The 2012 Haida Gwaii earthquake. *Geophysical Research Letters*, *43*, 9819–9828. https://doi.org/10.1002/2016GL070140

Heidarzadeh, M., Murotani, S., Satake, K., Ishibe, T., & Gusman, A. R. (2016). Source model of the 16 September 2015 Illapel, Chile, M_w 8.4 earthquake based on teleseismic and tsunami data. *Geophysical Research Letters*, *43*, 643–650. https://doi.org/10.1002/2015GL067297

Ho, T.-C., Satake, K., & Watada, S. (2017). Improved phase corrections for transoceanic tsunami data in spatial and temporal source estimation: Application to the 2011 Tohoku earthquake. *Journal of Geophysical Research: Solid Earth*, *122*. https://doi.org/10.1002/2017JB015070

Hossen, M. J., Cummins, P. R., Dettmer, J., & Baba, T. (2015a). Time reverse imaging for far-field tsunami forecasting: 2011 Tohoku earthquake case study. *Geophysical Research Letters*, *42*, 9906–9915. https://doi.org/10.1002/2015GL065868

Hossen, M. J., Cummins, P. R., Dettmer, J., & Baba, T. (2015b). Tsunami waveform inversion for sea surface displacement following the 2011 Tohoku earthquake: Importance of dispersion and source kinematics. *Journal of Geophysical Research: Solid Earth*, *120*, 6452–6473. https://doi.org/10.1002/2015JB011942

Hossen, M. J., Cummins, P. R., Roberts, S. G., & Allgeyer, S. (2015). Time reversal imaging of the tsunami source. *Pure and Applied Geophysics*, *172*, 1–16. https://doi.org/10.1007/s00024-014-1014-5

Hossen, M. J., Cummins, P. R., & Satake, K. (2017). Complete implementation of the Green's function based time reverse imaging and sensitivity analysis of reversed time tsunami source inversion. *Geophysical Research Letters*, *44*, 9844–9855. https://doi.org/10.1002/2017GL074528

Lawson, C. L., & Hanson, R. J. (1995). *Solving least squares problems*. Philadelphia: SIAM.

Lay, T., Kanamori, H., Ammon, C. J., Hutko, A. R., Furlong, K., & Rivera, L. (2009). The 2006–2007 Kuril Islands great earthquake sequence. *Journal of Geophysical Research*, *114*, B11308. https://doi.org/10.1029/2008JB006280

Lay, T., Ammon, C. J., Kanamori, H., Rivera, L., Koper, K. D., & Hutko, A. R. (2010). The 2009 Samoa-Tonga great earthquake triggered doublet. *Nature*, *466*(7309), 964–968. https://doi.org/10.1038/nature09214

Mulia, I. E., Gusman, A. R., & Satake, K. (2017). Optimal design for placements of tsunami observing systems to accurately characterize the inducing earthquake. *Geophysical Research Letters*, *44*. https://doi.org/10.1002/2017GL075791

Newman, A. V., Hayes, G., Wei, Y., & Convers, J. (2011). The 25 October 2010 Mentawai tsunami earthquake, from real-time discriminants, finite-fault rupture, and tsunami excitation. *Geophysical Research Letters*, *38*, L05302. https://doi.org/10.1029/2010GL046498

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

Sandu, A., Daescu, D. N., Carmichael, G. R., & Chai, T. (2005). Adjoint sensitivity analysis of regional air quality models. *Journal of Computational Physics*, *204*(1), 222–252.

Satake, K. (1987). Inversion of tsunami waveforms for the estimation of a fault heterogeneity: Method and numerical experiments. *Journal of Physics of the Earth*, *35*, 241–254. https://doi.org/10.4294/jpe1952.35.241

Satake, K. (1989). Inversion of tsunami waveforms for the estimation of heterogeneous fault motion of large submarine earthquake: The 1968 Tokachi-oki and 1983 Japan Sea earthquakes. *Journal of Geophysical Research*, *94*, 5627–5636.

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. *Bulletin of the Seismological Society of America*, *103*(2B), 1473–1492. https://doi.org/10.1785/0120120122

Tonini, R., Armigliato, A., & Tinti, S. (2011). The 29 September 2009 Samoa Islands tsunami: Simulations based on the first focal mechanism solutions and implications on tsunami early warning strategies. *Pure and Applied Geophysics*, *168*(6), 1113–1123. https://doi.org/10.1007/s00024-010-0221-y

Voronina, T. A. (2016). Recovering a tsunami source and designing an observational system based on an R-solution method. *Numerical Analysis and Applications*, *9*(4), 267–276. https://doi.org/10.1134/S1995423916040017

Watada, S., Kusumoto, S., & Satake, K. (2014). Traveltime delay and initial phase reversal of distant tsunamis coupled with the self-gravitating elastic Earth. *Journal of Geophysical Research: Solid Earth*, *119*, 4287–4310. https://doi.org/10.1002/2013JB010841

Wei, Y., Chamberlin, C., Titov, V., Tang, L., & Bernard, E. (2013). Modeling of the 2011 Japan tsunami: Lessons for near-field forecast. *Pure and Applied Geophysics*, *170*(6–8), 1309–1331. https://doi.org/10.1007/s00024-012-0519-z