Near-field tsunami inundation forecast method assimilating ocean bottom pressure data: A synthetic test for the 2011 Tohoku-oki tsunami

Yuichiro Tanioka\textsuperscript{a,\textasteriskcentered}, Aditya Riadi Gusman\textsuperscript{b}

\textsuperscript{a} Institute of Seismology and Volcanology, Hokkaido University, N10W8 Kita-ku, Sapporo Hokkaido 060-0810, Japan
\textsuperscript{b} GNS Science, 1 Fairway Drive, Avalon, Lower Hutt 5010, New Zealand

\section*{A B S T R A C T}

An approach for forecasting near-field tsunami inundation was developed by combining two methods. The first method computes tsunami by assimilating pressure data observed at numerous ocean bottom sensors without tsunami source information, and the second method forecasts near-field tsunami inundation by selecting a site-specific scenario from a precomputed tsunami inundation database. In order to evaluate the validity of this combined method, we performed a synthetic forecast test for the 2011 Tohoku-oki tsunami along the Pacific coast in Japan. A tsunami computation test performed using the assimilation of synthetic pressure data reveals that the method reproduced well the tsunami field for the 2011 Tohoku-oki tsunami. A synthetic near-field tsunami inundation forecast at four sites, Kamaishi, Rikuzentakata, Minamisanriku, and the Sendai Plain for the 2011 Tohoku-oki tsunami also worked. The results indicate that an accurate tsunami inundation forecast method by this combined approach using pressure data from numerous ocean bottom sensors is now available.

\section*{1. Introduction}

The 2011 Tohoku-oki earthquake (Mw9.1) generated a large tsunami that caused catastrophic destruction along the Pacific coast of the Tohoku region in Japan (Mori et al., 2012). Although the Japan Meteorological Agency (JMA) issued major tsunami warning messages and forecasts of tsunami heights along the coast of Japan (Ozaki, 2011), approximately 18,000 people were killed by the tsunami. The JMA's tsunami early warnings are based on a hypocenter and a magnitude of the source area such as the S-net. However, those systems require a long time to operate, and a tsunami forecast system using the ocean bottom pressure sensors in and around Japan Trench (S-net) in 2012 and completed the installation in 2018. This network consists of 150 ocean bottom stations in which ocean bottom pressure gauges and seismometers are installed (Uehira et al., 2012; Kanazawa, 2013). These ocean bottom stations in the S-net are connected by cables at 30 km intervals. These cables are distributed along a wide portion of the seafloor offshore of Kanto, Tohoku, and Hokkaido. The next step after the installation of this network was to develop an accurate tsunami computation method using these dense tsunami observation data.

Numerical simulations for tsunami have been used extensively in research on tsunami warning and forecasting methods (Geist et al., 2016; Rabinovich et al., 2017). In such tsunami simulations, a computation of tsunami propagation starts from an initial ocean surface displacement. This initial displacement is typically calculated from the ocean bottom deformation caused by the faulting of a large earthquake (Tanioka and Seno, 2001; Gusman et al., 2012, 2014). In this case, the source model of the earthquake must be estimated from other data, such as seismic or geodetic data, before the tsunami simulation is conducted.

The initial ocean surface displacements can be estimated directly from the observed tsunami waveforms. Tsunami forecast system using ocean bottom pressure observations has been developed and is operated by the National Oceanic Atmospheric Administration (NOAA) (Tang et al., 2009, 2012). Tsushima et al. (2009, 2012) developed a tsunami forecast system using the ocean bottom pressure sensors in and around the source area such as the S-net. However, those systems require a long...
time to produce the appropriate tsunami warning in near-field. Those methods have been improved by incorporating geodetic observations made on land with the ocean bottom pressure observations, which facilitated more rapid forecasting (Tsushima et al., 2014; Melgar and Bock, 2013, 2015; Wei et al., 2014). However, such methods continue to estimate the tsunami source, the initial surface displacement, from the observations. It is more accurate to compute tsunami directly from the observations made by ocean bottom pressure sensors without...
computing tsunami inundation is a time-intensive process. Only supercomputers can handle computations in real time for a near-field forecast on tsunami inundation, such as the forecast of the Sanriku coast for the 2011 Tohoku-oki earthquake (Musa et al., 2018). The other method for a near-field forecast on high resolution bathymetry in real-time is a precomputed tsunami inundation and tsunami waveform database (Gusman et al., 2014). In this method, after tsunami is computed using linear shallow-water equations at nearshore points in real-time, a scenario that gives the most similar tsunami waveforms at the nearshore points is selected as a site-specific best scenario. The tsunami inundation from this model becomes the tsunami inundation forecast.

In this study, we introduce a method of near-field tsunami inundation forecasting that combines the computation of tsunami by assimilating tsunami observation data without the source information (Tanioka, 2018), and near-field tsunami inundation forecasting by selecting site-specific scenarios from a precomputed tsunami inundation database (Gusman et al., 2014). To evaluate the validity of this combined approach, we performed a synthetic forecast test for the 2011 Tohoku-oki tsunami along the Pacific coast from the Sendai Plain to Kamaishi City, Japan (Fig. 1).

2. Methods

2.1. Method for tsunami simulation from ocean bottom pressure data

A detailed description of the data assimilation method is presented in Tanioka (2018). Here, we explain the method briefly. As shown in Fig. 2, we defined the water-depth fluctuation, \( h_i(x, y, t) \), which is proportional to the ocean bottom pressure, as:

\[
\frac{\partial h_i(x, y, t)}{\partial t} = \frac{1}{\rho g} \Delta p(x, y, t) = \eta(x, y, t) - B(x, y, t)
\]

(1)

where \( p \) is the density of sea water, \( g \) is the gravitational acceleration, \( \eta(x, y, t) \) is the ocean surface displacement or tsunami height, and \( B(x, y, t) \) is the sea floor deformation caused by the faulting of a great earthquake. After the earthquake is completed, the sea floor deformation, \( B(x, y, t) \), is no longer time-dependent. Then, the time derivative of tsunami height, \( \frac{\partial \eta(x, y, t)}{\partial t} \), becomes the time derivative of the water-depth fluctuation, which can be obtained from the ocean bottom pressure data, \( \Delta p(x, y, t) \). A governing equation for the tsunami simulation, the wave equation of the linear shallow-water approximation, is described as:

\[
\frac{\partial^2 \eta_i(x, y, t)}{\partial t^2} = \frac{\partial \eta_i(x, y, t)}{\partial t} = gd \left( \frac{\partial^2 \eta_i(x, y, t)}{\partial x^2} + \frac{\partial^2 \eta_i(x, y, t)}{\partial y^2} \right)
\]

(2)

where \( d \) is the depth of the ocean. The equation is suitable for the time after the earthquake or the tsunami generation is completed.

From Eq. (2), we obtain the finite difference equation of equation as:

\[
\frac{h_i^{k+1} - 2h_i^k + h_i^{k-1}}{\Delta t^2} = gd \left( \frac{\eta_{i+1,j} - 2\eta_i,j + \eta_{i-1,j}}{\Delta x^2} + \frac{\eta_{i,j+1} - 2\eta_i,j + \eta_{i,j-1}}{\Delta y^2} \right)
\]

(3)

where \( i \) and \( j \) are indices of the points for the x- and y-directions in the computed domain, \( k \) is an index for time steps, \( \Delta x \) and \( \Delta y \) are space intervals of the x- and y-directions, respectively, and \( \Delta t \) is a time interval.

In this study, we assumed that there were ocean bottom pressure sensors at each point, as indicated in Fig. 3. A space interval for both x- and y-directions is 15 arc-min, or approximately 30 km. The total number of sensors are assumed to be 425. The left-hand side of Eq. (3) was obtained from the observed pressure data at each ocean bottom pressure sensor. The unknown parameters are the ocean surface displacement field or tsunami height distribution, \( \eta_i,j \), above each ocean bottom pressure gauge at a particular time. The tsunami heights outside the pressure gauge network are assumed to be 0. Subsequently, we
could compute the tsunami height distribution, $\eta_{ij}^{\text{r}}$, above the ocean bottom pressure sensor network by solving Eq. (4) which includes Eq. (3) at all ocean bottom pressure sensors:

$$\mathbf{A} \mathbf{m} = \mathbf{d}$$

where $\mathbf{A}$ is a matrix consisting of coefficients of the right-hand side of Eq. (3); $\mathbf{m}$ consists of the unknown parameters or tsunami height distribution, $\eta_{ij}^{\text{r}}$, above the ocean bottom pressure gauge network at a particular time; and $\mathbf{d}$ is the observed data or the left-hand side of Eq. (3). Detailed description of the matrix, $\mathbf{A}$, is available in Tanioka (2018). The inverse of the matrix, $\mathbf{A}$, is calculated using the singular value decomposition technique (Press et al., 2007).

By solving Eq. (4), the tsunami height field can be reproduced at any time as long as the tsunami remain within the ocean bottom pressure gauge network. The estimated tsunami height field at each time with an interval, $\Delta t$, is used to replace the tsunami height field in the tsunami numerical simulation with a finite difference scheme. The tsunami numerical simulation continues as long as the necessary tsunami wave field is obtained. The result of this tsunami computation is used for the near-field tsunami inundation forecast method described in the next section.

2.2. Method for near-field tsunami inundation forecast (NearTIF)

Tsunami inundation on land is closely related to the characteristics of tsunami, such as height and period, in the nearshore environment. To forecast near-field tsunami inundation, we built a database containing pairs of precomputed inundations and precomputed tsunami waveforms at specific sites from hypothetical fault models. The detailed description of this method, NearTIF, is shown in Gusman et al. (2014).

The site-specific best scenario from the models is selected by minimizing the root-mean square error (RMSE) between the computed tsunami waveforms from the tsunami simulation method described above and those in the database. Then the corresponding precomputed tsunami inundation of the best scenario is selected as the tsunami inundation forecast.

The precomputed tsunami database is constructed from earthquake scenarios of rectangular fault models along the plate interface with moment magnitudes in the range of $M_w$ 8.0 to $M_w$ 9.0 and an interval of 0.1. Fifty-six reference points are located along the plate interface off the east coast of Honshu, Japan. A total of 532 scenarios defined by Gusman et al. (2014) were used in this study. To select the best scenario at each site, the virtual observation points at each site are defined as shown in Fig. 4. At those virtual observation points, the computed tsunami waveforms are compared with the waveforms in a database.
The tsunami waveforms at each site in a database are numerically computed by solving the linear shallow-water equations using a finite difference scheme. The numerical tsunami computation scheme is described in Johnson (1998). The size of the bathymetry grid in the simulation is 30 arc-sec. Tsunami inundations at each site from the earthquake scenarios are numerically computed by solving the non-linear shallow-water equations with a moving boundary condition at the finest-scale grid system (Imamura, 1996; Goto et al., 1997).

Fig. 5. Tsunami height distribution at 300 s, and 600 s after the origin time. (a) Tsunami numerical simulation from the original slip distribution, and (b) tsunami simulation from the assimilation of synthetic pressure data at sensors shown in Fig. 3.
Homogeneous Manning’s roughness coefficient of 0.025 m$^{-1/3}$ s was assumed at the bottom surface. Nested grid systems of 30, 10, 3.33, 1.11 arc-sec resolution were used.

To select the best scenario at each site, we selected a time shift of the tsunami waveforms at virtual observation points (Fig. 4) for each scenario. The site-specific best scenario and the optimal time shift was determined by minimizing the RMSE between the tsunami waveforms computed by the above tsunami simulation method using ocean bottom pressure data in real-time, and those in the database.

3. Numerical test for the 2011 Tohoku-oki tsunami

Several source models for the 2011 Tohoku-oki earthquake have been suggested to explain the observed tsunami (e.g. Satake et al., 2013; Tappin et al., 2014; Yamazaki et al., 2018). In this study, the slip distribution estimated from tsunami waveforms and crustal deformation data (Gusman et al., 2012) (Fig. 1) was chosen to be a source model for a synthetic forecast test. The tsunami numerically computed from the slip distribution was used as an original tsunami for the synthetic test. We also assumed that the rise time of 80 s for the synthetic test. The ocean bottom pressure sensors, the observation points, are assumed to be positioned along 15 arc-min intervals, approximately 30 km intervals, as shown in Fig. 3. The original tsunami at the observation points are transferred to the water-depth fluctuation, $h_j(x, y, t)$, via Eq. (1). The water depth fluctuation was observed at 40 s intervals, $\Delta t$ in Eq. (3), because this time interval should be long enough to enable the water-depth fluctuation to be observed. The tsunami height distribution at the observation points (Fig. 3) at a particular time was computed by solving Eq. (3) at all observation points simultaneously using Eq. (4). Those tsunami height distribution was interpolated to 30 arc-sec intervals. Then, it was used to replace the tsunami height distribution at that particular time in the tsunami numerical simulation. We continued this replacement of the tsunami height distribution in the numerical tsunami computation until 120 s after the origin time in this study. The estimation of tsunami height distribution is unstable until the earthquake is completed because Eq. (2) becomes applicable after the tsunami generation or the co-seismic crustal deformation is completed. Because we assumed a rise time of 80 s for the synthetic test, it is needed to be continued to replace the tsunami height distribution in the numerical tsunami computation at least 80 s after the origin time. We computed one interval more, 120 s, to make sure the estimation of the tsunami height distribution is stable. In an actual instance, the duration of the earthquake would be unknown, therefore, we could continue the replacement as long as the indicative shaking of the earthquake continued. After those replacements of the tsunami height distribution, the tsunami numerical simulation continued and completed at 3 h of propagation time.

The tsunami waveforms for the virtual observation points at four sites shown in Fig. 4, Kamaishi, Rikuzentakata, Minamisanriku, and the Sendai Plain, were obtained as the result of the tsunami numerical simulation described the above. The computed tsunami waveforms for those virtual observation points were then compared with these waveforms in the database to select the site-specific best scenario. Finally, the corresponding precomputed tsunami inundation of the best scenario at each site was selected as the tsunami inundation forecast.

To evaluate the accuracy of this method, we computed the tsunami inundation at four sites from the original slip distribution directly. We also computed the tsunami inundation directly from the result of the above tsunami numerical simulation by assimilating pressure data at the ocean bottom sensors shown in Fig. 3 without using a precomputed...
Fig. 7. For Kamaishi, (a) the tsunami inundation computed from the original slip distribution, (b) from the assimilation of synthetic pressure data, and (c) from the NearTIF based on the data assimilation. Bars are observed tsunami heights. Purple dots are the computed maximum tsunami heights. The geometric mean ratio of Aida (1978) number $K$, and its corresponding standard deviation, $\kappa$, are also shown.

Fig. 8. For Rikuzentakata, (a) the tsunami inundation computed from the original slip distribution, (b) from the assimilation of synthetic pressure data, and (c) from the NearTIF based on the data assimilation. Bars are observed tsunami heights. Purple dots are the computed maximum tsunami heights. The geometric mean ratio of Aida (1978) number $K$, and its corresponding standard deviation, $\kappa$, are also shown.
tsunami inundation database. Tsunami inundation forecasts from these three computations are compared below.

4. Results

The results of the tsunami simulation from the synthetic ocean bottom pressure data for the 2011 Tohoku-oki tsunami are shown in Figs. 5 and 6. The comparison of the tsunami height distribution computed from the original slip distribution with that of another computed tsunami from the assimilation of synthetic pressure data at 5 and 10 min after the origin time are shown in Fig. 5. Although a short-wavelength part of the tsunami from the original slip distribution is not reproduced well by the computed tsunami from the assimilation of synthetic pressure data, the overall pattern of the tsunami height distribution from the original slip distribution is reproduced well by that from the data assimilation. The comparison of the tsunami waveforms at 7 sites, TG1 ~ TG7, in Fig. 5 are shown in Fig. 6. Short-period parts of the original tsunami waveforms are not reproduced well by these of the computed ones from the assimilation of synthetic pressure data, either. However, the overall amplitudes and shapes of the tsunami waveforms from the original slip distribution are well reproduced by the computed ones from the data assimilation. These results indicate that the tsunami simulation method using the assimilation of pressure data at numerous ocean bottom sensors without any information on the tsunami source works and can provide the tsunami wave field for near-field tsunami inundation forecasting.

The tsunami waveforms of the virtual observation points at four sites, Kamaishi, Rikuzentakata, Minamisanriku, and the Sendai Plain (Fig. 4) were computed from the above simulation. The site-specific best scenario for each site was selected by comparing the computed tsunami waveforms at those points with those from the database, which includes 532 scenarios. A precomputed tsunami inundation for the best scenario for Kamaishi was compared with the tsunami inundation computed from the original slip distribution, and with the one computed directly from the tsunami simulation using the assimilation of synthetic pressure data (Fig. 7). In Fig. 7, the computed maximum tsunami heights from three tsunami inundation computations are compared with the actual survey tsunami heights of the 2011 Tohoku-oki tsunami (Mori et al., 2012). To qualitatively evaluate the data fit, we used the geometric mean ratio of Aida (1978) number $K$, and its corresponding standard deviation, $\kappa$, which can be written as:

$$\log K = \frac{1}{N} \sum_{i=1}^{N} \log K_i; \quad K_i = \frac{O_i}{S_i}$$

$$\log \kappa = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} ((\log K_i)^2 - (\log K)^2)$$

where $K_i$ is the ratio between the observed tsunami height, $O_i$, and the simulated tsunami height, $S_i$.

Fig. 7a reveals that the tsunami inundation at Kamaishi computed from the original slip distribution slightly overestimates the observed inundation area due to the 2011 Tohoku-oki tsunami. On an average, the observed tsunami heights on land were reproduced well by the computed ones because $K = 0.99$. Fig. 7b reveals that the tsunami inundation area computed from the assimilation of synthetic pressure data also slightly overestimates the observed inundation area of the 2011 event. The observed tsunami heights on land were also generally reproduced well by the computed ones from the assimilation of synthetic pressure data because of $K = 1.12$. Fig. 7c reveals that the tsunami inundation area of the site-specific best scenario using the NearTIF based on the data assimilation also slightly overestimates the observed inundation area of the 2011 event. The observed tsunami heights on land were also typically reproduced well by the computed ones from the NearTIF based on the data assimilation because $K = 1.19$. The standard deviations, $\kappa$, for the three computations do not change much.
The geometric mean ratio of Aida (1978) number $\kappa$, and its corresponding standard deviation, $\kappa$, to compare the observed tsunami heights with computed ones at four inundation computation sites in Figs. 7–10.

Table 1

<table>
<thead>
<tr>
<th>Sites</th>
<th>$K$</th>
<th>$\kappa$</th>
<th>$K$</th>
<th>$\kappa$</th>
<th>$K$</th>
<th>$\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kamaishi</td>
<td>0.99139</td>
<td>1.5023</td>
<td>1.1218</td>
<td>1.5381</td>
<td>1.1859</td>
<td>1.5711</td>
</tr>
<tr>
<td>Rikuzentakada</td>
<td>0.90834</td>
<td>1.2403</td>
<td>0.86298</td>
<td>1.2354</td>
<td>0.98152</td>
<td>1.1809</td>
</tr>
<tr>
<td>Minamisanriku</td>
<td>1.01</td>
<td>1.2708</td>
<td>1.0203</td>
<td>1.2939</td>
<td>1.2945</td>
<td>1.3056</td>
</tr>
<tr>
<td>Sendai Plain</td>
<td>0.66054</td>
<td>2.0282</td>
<td>0.56676</td>
<td>2.0341</td>
<td>0.59823</td>
<td>2.0221</td>
</tr>
</tbody>
</table>

The three computed tsunami inundations for each region, Rikuzentakada, Minamisanriku, and Sendai Plain are compared with the observed tsunami inundation area and heights of the 2011 Tohoku-oki tsunami and shown in Figs. 8, 9 and 10, respectively. To evaluate the data fit between the observed tsunami heights and the computed ones, the geometric mean ratio, $K$, and its corresponding standard deviation, $\kappa$, are shown in Table 1. For Rikuzentakada, the observed tsunami heights are well reproduced by the three computed tsunami inundations from the original slip distribution, the assimilation of synthetic pressure data, and the NearTIF based on the data assimilation. For Minamisanriku, the observed tsunami heights are slightly underestimated by the computed tsunami inundation from the NearTIF based on the data assimilation, $K = 1.29$ in Fig. 9 and Table 1, although the observed tsunami heights are well reproduced by the other two computed inundations. For Sendai Plain, the observed tsunami heights are overestimated by all three computed tsunami inundations, $K = 0.66$, 0.57, and 0.60 from the original slip distribution, the assimilation of synthetic pressure data, and the NearTIF based on the data assimilation, respectively. However, $K$ from the three computed tsunami inundations are similar. The standard deviations, $\kappa$, for those three do not change much either. We will discuss about those overestimations in the following section.

5. Discussions and conclusions

Overall, the tsunami computation from the pressure data at numerous ocean bottom sensors without any information on tsunami source functioned well for the 2011 Tohoku-oki tsunami. By using the results from these tsunami computations, a synthetic tsunami inundation forecast using a precomputed tsunami inundation and tsunami waveform database (Gusman et al., 2014) was performed at four sites, Kamaishi, Rikuzentakada, Minamisanriku, and the Sendai Plain. The tsunami inundation forecasts for Kamaishi, Rikuzentakada, and Minamisanriku generally functioned well. Although the tsunami inundation forecasts for the Sendai Plain overestimated the observed tsunami inundation area, the reason for this overestimation does not come from the methods employed, but from the original tsunami source because the tsunami inundation computed directly from the original slip distribution also overestimated the observed data.

An average computation time of tsunami inundations for four regions, Kamaishi, Rikuzentakada, Minamisanriku, and Sendai Plain, was 103 min using a processor, 4 GHz Intel Core i7. On the other hand, it only took about 4 min for the assimilation of synthetic pressure data and about 3 min for the NearTIF operation. The total operation time of 6 min is short enough for the real time tsunami forecast. We conclude that the synthetic tsunami inundation forecast test for the 2011 Tohoku-oki tsunami using the NearTIF method (Gusman et al., 2014) based on the assimilation of synthetic pressure data was functioned well. However, it may need more time for the actual operation to remove high-frequency noise from original ocean bottom pressure data. In this synthetic test, we knew that the origin time and a rise time of 80 s for co-seismic deformation. Therefore, we decided to assimilate the synthetic pressure data until 120 s after the origin time. In an actual instance, the duration of the earthquake would be unknown, therefore, we could continue to assimilate the pressure data as long as the indicative shaking of the earthquake continued, or until the tsunami height distribution was stably estimated. Therefore, the operation time may increase more in an actual instance.

The ocean bottom pressure sensors in S-net were connected at 30 km intervals by a cable. Therefore, we distributed sensors at 15 min intervals, approximately 30 km intervals, in this study. However, in real case, a cable (S-net) was not installed in a constant interval, particularly in north-south direction. We also distributed sensors in wider area than S-net. It is necessary to improve our method to use exact locations of sensors in S-net for the real time tsunami inundation forecast. We can also improve the method by including data from existing tsunami gauges such as GPS buoys.

This study demonstrated that an accurate near-field tsunami inundation forecast method is available using data from numerous ocean bottom pressure sensors.

Acknowledgments

We thank two anonymous reviewers for helpful comments to improve the manuscript. This study was supported by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan, under its Earthquake and Volcano Hazards Observation and Research Program.
Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.pepi.2018.08.006.

References