Improving sequential Bayesian update for tsunami scenario detection by using geodetic data learning

Reika Nomura, Tohoku University
Yu Otake, Tohoku University
Shuji Moriguchi, Tohoku University
Diego Melgar, University of Oregon
Randall J. LeVeque, University of Washington
Kenjiro Terada, Tohoku University
E-mail: nomura@irides.tohoku.ac.jp

In this study, we investigate the effectiveness of geodetic data monitored by GNSS (Global Navigation Satellite System) as the prior learning to observational ocean data for the improved tsunami scenario detection. For the case study targeting Nankai-trough, 600 earthquake/tsunami scenarios are generated by GeoClaw and fakequake software. With the synthetically generated geodetic displacements by fakequake software, we examine the reasonable initial probability setting which is prior to the learning on the ocean wave data.

1. Introduction

Tsunamis are one of the most significant coastal risks in the international community. Especially in Japan, which is frequently threatened by earthquakes and its subsequent tsunami inundation, considerable attention has been paid to the expected Nankai-Trough earthquake and tsunami in the next few decades, which will be an event with extremely short evacuation time (<30 mins) for the residents. Although there have been many tsunami forecast systems proposed, there is still room to improve minimizing the uncertainties relevant to the deterministic numerical simulations or erroneous/noisy data acquired in real-time.

This study examines the utilization of geodetic data, which can be obtained immediately after the fault rupture by GNSS (Global Navigation Satellite System), as the prior learning to in-situ ocean data along the lines of the previously developed tsunami scenario detection method[1]. For that purpose, we set up 600 scenarios of fault ruptures and subsequent tsunamis, targeting the Nankai trough. In addition to the tsunami wave history data sampled in some synthetic ocean gauges, geodetic displacements supposed to be monitored at 18 GNSS stations are synthetically generated by the fakequake software[4),5)]. With the database, we examine the reasonable initial probability setting for the sequential tsunami scenario detection methods.

2. Earthquake and tsunami scenario generations

The fakequakes software[2),3),4)] is used to generate 600 synthetic kinematic slip distributions for Nankai-trough-going ruptures along with the corresponding seafloor deformations. Fig. 1(a) shows one of the slip distribution generated under Mw 8.3 situations. Fig. 1(b) and (c) shows the snapshot of the fault rupture, which are calculated from Okada model[5)]. We choose 18 locations that are identical to that of GEONET[6)], as shown in Fig. 2(a). The nearest stations to KOCHI city, where are threatened by the huge tsunami risk triggered by Nankai mega-thrust, are also provided in Fig. 2(b). In those points, geodetic motions in every 3 directions, North-South, East-West, and vertical are synthetically generated at 1.0 Hz with 1024 seconds data duration based on Green functional methods.

The subsequent tsunami propagation and inundation are realized by GeoClaw[7)], as shown in Fig. 2(c) and (d). The 4 hours wave history data are sampled at 5 seconds for POD and Bayesian update scheme. Fig. 2(d) shows the wavefoam data calculated at the black dots described in Fig. 2(c). We can understand that the first waves do not reach those offshore gauge locations when the GNSS sampling durations are indicated by the blue region in graphs at Fig. 2(d). From that, we can expect that GNSS data can provide rich information for scenario detection before the offshore wave observations.

3. Geodedic data as the prior information for offshore observations

According to the previous method[1)], we detect the most probable tsunami scenario from N_s pre-computed tsunami scenarios based on the following Bayesian theorem:

$$P(E_j \mid \varepsilon^{(t)}) = \frac{P(\varepsilon^{(t)} \mid E_j)}{\sum_{i=1}^{N_s} P(\varepsilon^{(t)} \mid E_i) P(E_i)} P(E_j \mid \varepsilon^{(t-1)}) \qquad (1)$$

Here, E_j represents the event that the tsunami equivalent to the j-th scenario occurs. Also, $\varepsilon^{(t)}$ means the events that we ob-

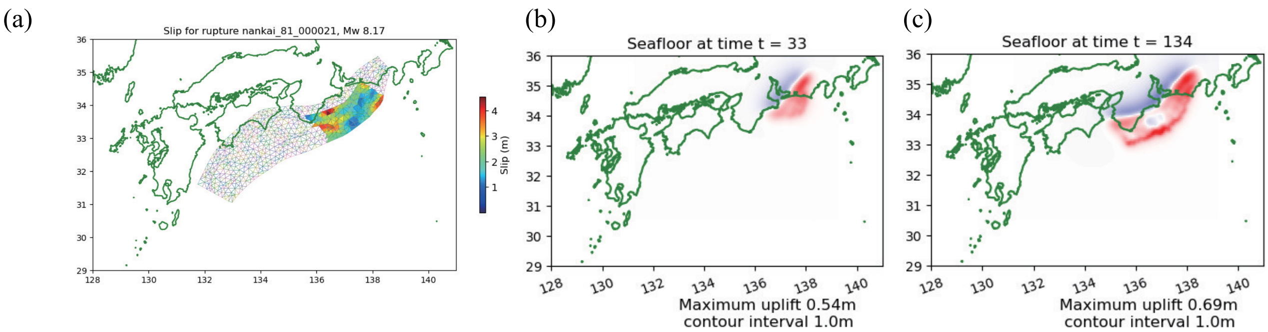

Fig. 1 Example of the synthetic fault rupture generated by fakequake and Okada model

tain some kinds of information related to the occurring tsunami. Since $P(A \mid B)$ means the conditional probability of A given B, $P(E_j \mid \varepsilon^{(t)})$ means the probability of j-th pre-computed tsunami scenario would be the occuring events. Here, the recent ocean observational network, such as DONET, would realize $\varepsilon^{(t)}$ to be the sequentially updated information per dozens of seconds. The conditional probability $P(E_j \mid \varepsilon^{(t)})$ can be updated by the following formulation:

At the first step of the Bayesian update, we can impose the uniform probability $1/N_s$ on each tsunami scenario according to the principle of insufficient reason. However, this prior probability can be determined more reasonably if we can rely on geodetic information measured immediately after the fault ruptures. In fact, the geodetic data reduct the high sensitivity of the offshore gauges located along the main tsunami propagating path in the process of machine learning done by Makinoshima et al.[8].

For that purpose, we attempt to improve our previous tsunami scenario detection method by setting the prior probability at the first step of (1), $P(E_j \mid \varepsilon^{(0)})$, as some function of the geodetic data. Our POD type learning method is also applicable to the full-time series of geodesic motion. But also the representative information extracted by those histories, such as the maximum displacement or the total displacement at the final step, will be sufficient for the learning.

4. Conclusion

In this study, we attempt to set the prior probabilty at the inital step of Bayesian update as some function related to the synthetic geodedic data generated by fakequake (MudPy) software. The predicition accuracies will be discussed with the quantative indecese such like the maximum wave heights in the presentations.

References

1) (Under review, Submitted in Dec. 2021) Nomura et al., Sequential Bayesian update to detect the most probable tsunami scenario using observational wave sequence.

2) Melgar, D., MudPy software, `https://github.com/dmelgarm/MudPy` (accessed 2021).

3) LeVeque et al., Generating random earthquake events for probabilistic tsunami hazard assessment. In Global Tsunami Science: Past and Future, Volume I, pp. 3671-3692, Birkhäuser, Cham, 2016.

4) Melgar et al., Kinematic rupture scenarios and synthetic displacement data: An example application to the Cascadia subduction zone. J. Geophys. Res.: Solid Earth, 121(9), 6658–6674, 2016.

5) Okada, Y., Surface deformation due to shear and tensile faults in a half-space. *Bulletin of the seismological society of America*, 75.4 1135-1154, 1985.

6) Sagiya, A decade of GEONET: 1994–2003 —The continuous GPS observation in Japan and its impact on earthquake studies—., *Earth Planet*, Sp 56, xxix–xli, 2004.

7) Mandli et al., Clawpack: building an open source ecosystem for solving hyperbolic PDEs, PeerJ Comput. Sci., doi:10.7717/peerj-cs.68, 2016.

8) Makinoshima et al., Early forecasting of tsunami inundation from tsunami and geodetic observation data with convolutional neural networks, Nat. Commun., 12 (1), pp. 1-10, 2021.

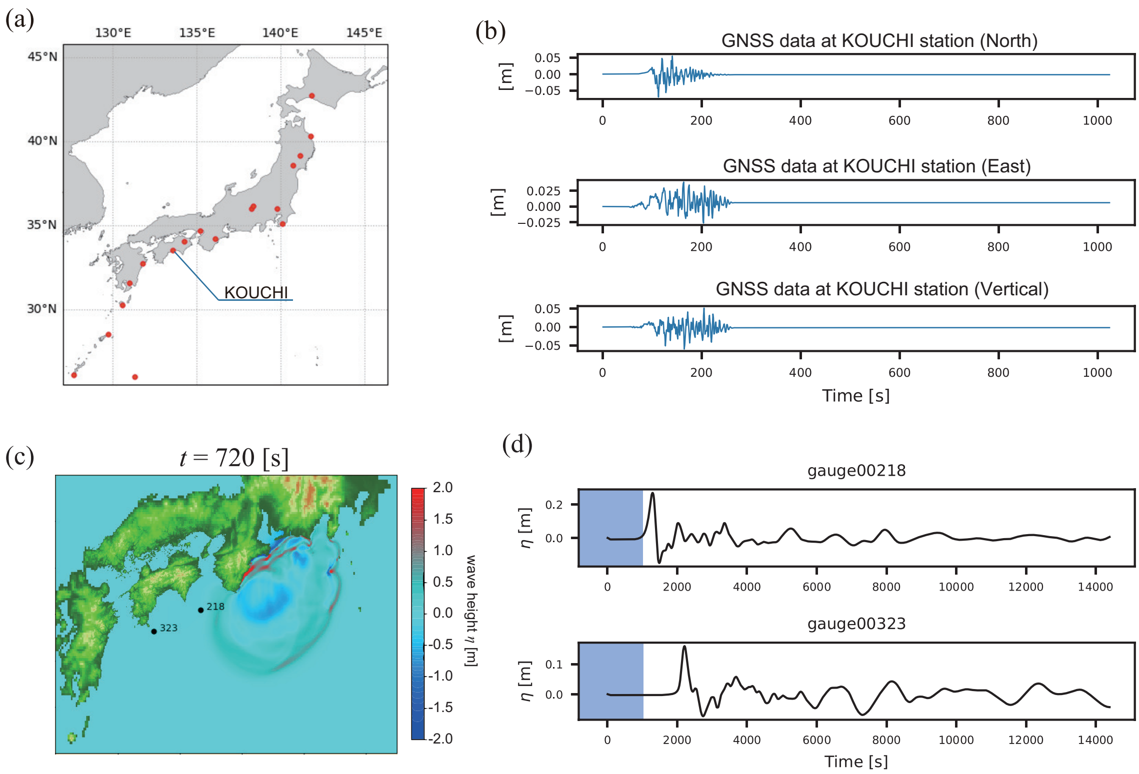

Fig. 2 Synthetic GNSS data : (a) Locations of geodetic motion monitorings, (b)(c)(d) Synthetic geodetic motion data generated by fakequake(MudPy) software