

Maximum tsunami height prediction using pressure gauge data by a Gaussian process at Owase in the Kii Peninsula, Japan

Yasuhiko Igarashi¹ · Takane Hori² · Shin Murata¹ · Kenichiro Sato² · Toshitaka Baba³ · Masato Okada¹

Received: 10 March 2016 / Accepted: 17 September 2016 / Published online: 8 December 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract We constructed a model to predict the maximum tsunami height by a Gaussian process (GP) that uses pressure gauge data from the Dense Oceanfloor Network System for Earthquakes and Tsunamis (DONET) in the Nankai trough. We found a greatly improved generalization error of the maximum tsunami height by our prediction model. The error is about one third of that by a previous method, which tends to make larger predictions, especially for large tsunami heights (>10 m). These results indicate that GP enables us to get a more accurate prediction of tsunami height by using pressure gauge data.

Keywords Tsunami prediction · Gaussian process · Cabled seafloor observatory · DONET · Nankai trough

Introduction

To mitigate tsunami damage, early warning systems are operated around the world. The Pacific Tsunami Warning System is responsible for monitoring the occurrence of earthquakes from seismological and tidal stations

throughout the Pacific Ocean estimating their tsunamigenic potential, and disseminating tsunami warning information to 26 participating countries and regions. Recently the countries located near the Indian Ocean have established three regional tsunami warning institutes, the German Indonesian Tsunami Early Warning System, Joint Australian Tsunami Warning Centre, and the Indian National Tsunami Early Warning System of the Indian Center for Ocean Information Services (Allen and Greenslade 2010; Münch et al. 2011; Kumar et al. 2012). In Japan, Japan Meteorological Agency immediately issues a tsunami early warning after the occurrence of an earthquake (Kamigaichi 2009).

The tsunami warning systems employ seismic and sea-level monitoring, in combination with a database of past tsunami events and premade numerical simulations. Another monitoring system for tsunami warning is the direct use of offshore tsunami data, such as data from the Deep-ocean Assessment and Reporting of Tsunami (DART) buoys. The National Oceanic and Atmospheric Administration (NOAA) has developed a far-field tsunami forecasting system using data assimilation by DART buoys in real time (Titov et al. 2005; Percival et al. 2011, 2014).

In Japan, the Dense Oceanfloor Network System for Earthquakes and Tsunamis (DONET) was recently developed in the Nankai trough (Kaneda et al. 2015). DONET1 (east of Kii Peninsula) is equipped with seismometers and ocean-bottom pressure gauges at 20 points on the sea floor. Fiber-optic cables connect them to a station on land so that submarine data can be acquired in real time. These data are useful for early prediction of tsunamis caused by earthquakes and submarine landslides. Such data are available not only in Japan but also in other countries such as Canada and Taiwan (Thomson et al. 2011; Fine et al. 2015; Hsiao et al. 2014).

✉ Masato Okada
okada@k.u-tokyo.ac.jp

¹ Department of Complexity Science and Engineering, The University of Tokyo, 5-1-5, Kashiwanoha, Kashiwa, Chiba 277-8561, Japan

² Research and Development Center for Earthquake and Tsunami, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 3173-25, Showa-machi, Kanazawa-ku, Yokohama, Kanagawa 236-0001, Japan

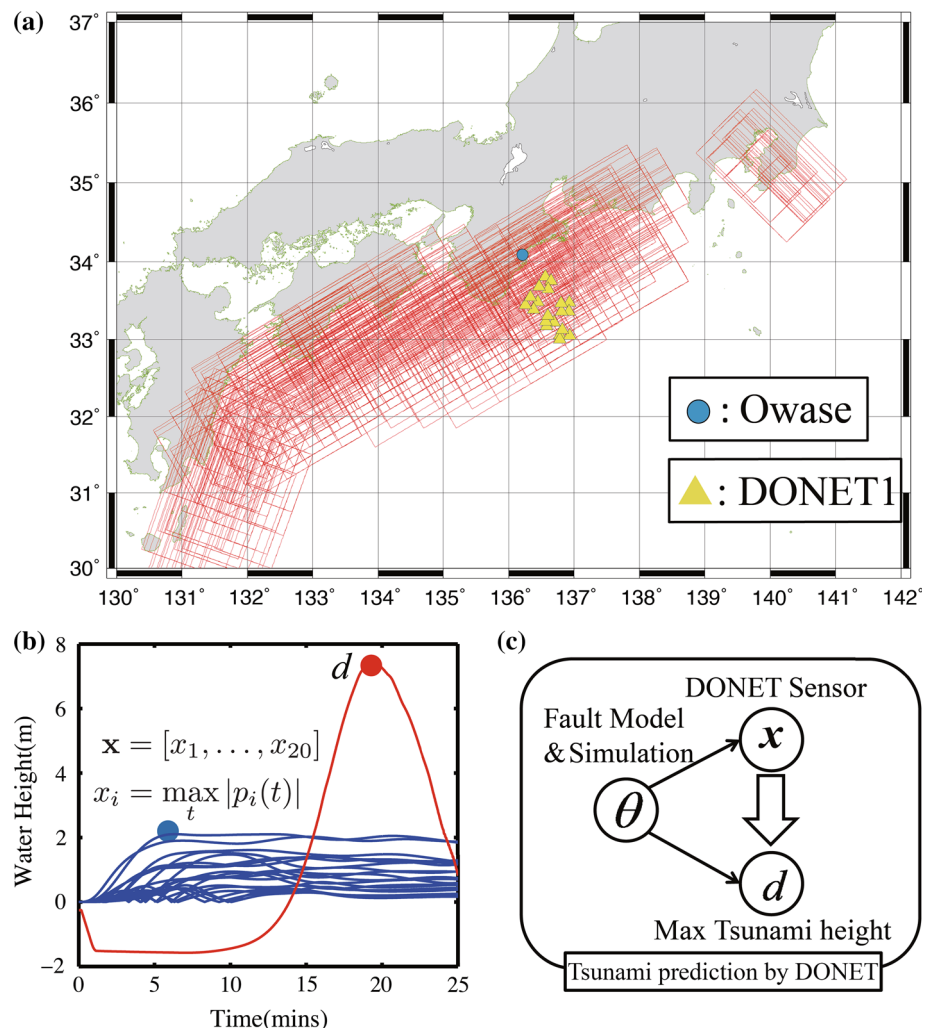
³ Institute of Technology and Science, The University of Tokushima, 2-1 Minami-jyosanjima-cho, Tokushima 770-8506, Japan

Previous studies have used hydrostatic pressure gauges on the sea floor to investigate real-time, fast-forecasting methods. Tsushima et al. (2009, 2011, 2014) combined hydrostatic pressure measurements and estimated tsunami sources to determine the spatial distribution of initial sea-surface displacements in the near-field tsunami source region. This procedure provided a highly accurate prediction within 20 min after the occurrence of the near-field Tohoku earthquake (Tsushima et al. 2011). However, faster prediction of an arriving tsunami is needed since it takes only a few minutes for a tsunami to arrive at the coastal area of the Kii Peninsula from the Nankai trough (Baba et al. 2004; Hayashi 2010; Baba et al. 2013a).

We studied the relationship between offshore and coastal tsunami heights with the aim of using DONET1 ocean-bottom pressure gauges for tsunami prediction. We assumed various tsunami models, including fault models and tsunami sources (Fig. 1a), and created a large number of simulations to reveal the relationship between DONET1 ocean bottom pressure gauge measurements and coastal tsunami heights.

Due to crustal deformation associated with faulting, hydrostatic pressure fluctuations may not accurately reflect sea surface fluctuations. The seafloor pressure observation also reflects elastic wave effects in the crust and seawater (Nosov and Kolesov 2007; Saito 2013), and seafloor vertical deformation below the observatory (Baba et al. 2006) as well. To remove the former effects, we derived these waveforms with a band-pass filter of 0.01–0.0001 Hz. Due to the latter effects, almost no hydrostatic pressure change is expected during earthquakes because the sea surface and the ocean bottom are dislocated equally, the result being no change in the total depth. The hydrostatic pressure suddenly decreases afterward as the tsunami propagates. A change in hydrostatic pressure corresponding to the vertical displacement of the seafloor remains after the tsunami has passed. Then, to catch the surface fluctuations from ocean-bottom pressure gauges, a previous study focused on the maximum absolute values of the hydrostatic pressure changes (Baba et al. 2013a), $\mathbf{x} = [x_1, x_2, \dots, x_S]$, recorded at DONET1 stations (S points) during a tsunami (Fig. 1b). They found a clear relationship between the average

Fig. 1 **a** Tsunami computational domain and the output stations used in this study. The 1506 fault models (rectangles) were prepared by making various changes to fault parameters. Triangles show the locations of DONET1 and a circle indicates the Owase tide gauge. **b** Absolute values of the hydrostatic pressure changes, $|p_i(t)|$ ($i = 1, \dots, S = 20$), at DONET1 stations (blue lines) and Owase (red line) for a near-field earthquake. The blue circle shows $x_i = \max_t |p_i(t)|$ for the i th DONET1 point and the red circle indicates the maximum wave height at Owase. **c** Schematic diagram of the relationship between DONET1 ocean-bottom pressure gauge measurements and coastal tsunami heights



waveforms of DONET1, $(1/S) \sum_i^S x_i$, and maximum tsunami heights d (Fig. 1c).

However, predictions by this method tend to be larger than simulated tsunami heights by up to about 5 m, especially for large tsunami heights (e.g., 10 m). Thus, further investigation is needed to improve the accuracy of the maximum tsunami height forecasting method based on limited DONET1 sensing data.

Here we construct an algorithm to predict maximum tsunami height based not on the average value but on individual values, $\mathbf{x} = [x_1, x_2, \dots, x_S]$. Let us consider the maximum tsunami height of the scenario n , $d(n)$, using given many sets of DONET1 sensing data and tsunami heights.

Since a previous study has reported that the maximum absolute values of the hydrostatic pressure changes have clearly positive correlations with the coastal maximum tsunami height regardless of the non-linear effects of the hydrodynamic equations (Baba et al. 2013a), we assumed that from two scenarios which are sufficiently similar in hydrostatic pressure change out of 20 observation points, the maximum tsunami height at Owase, d , will be similar and can be determined by interpolation. To be more precise, if the DONET1 sensing data of scenarios n and m are similar, $\mathbf{x}(n) \approx \mathbf{x}(m)$, we can approximate $d(n)$ by the known $d(m)$, as determined by interpolation. We then interpolate and predict tsunami height by the method of Gaussian process (GP) regression (Rasmussen and Williams 2005), using the distance between scenarios n and m , $|\mathbf{x}(n) - \mathbf{x}(m)| \equiv \sum_{i=1}^S |x_i(n) - x_i(m)|$.

To evaluate the accuracy of our method, we focus on the prediction of maximum tsunami height at Owase tide stations by DONET1 sensors (Fig. 1c). We randomly divide simulated scenarios into training and test data. Using training data, we construct an algorithm to fit d , at the Owase tide station from the maximum absolute value of the hydrostatic pressure changes, $\mathbf{x} = [x_1, x_2, \dots, x_S]$. We then apply our algorithm to test data and measure the prediction accuracy.

Simulation and dataset

Simulation

We constructed a database to estimate tsunami heights for near-field earthquakes in the Nankai trough region that would affect the Kii Peninsula region. First, we prepared 1506 fault models (Fig. 1a) where we considered the surface configuration of the Philippine Sea plate (Baba et al. 2002) and the source processes associated with the 1944 Tonankai and 1946 Nankai earthquakes (Kanamori 1972; Baba and Cummins 2005). The maximum depth of the

faults ranged from 5 to 25 km, the dip from 5 to 25°, and the magnitude from 7.2 to 8.4. We determined the fault length and width and amount of slip from the magnitude by a scaling law (Utsu 2001). Strike and rake were assumed to be constant at 240 and 90° azimuth, respectively. We assumed a uniform amount of slip on each fault to simplify the procedure, although actual earthquakes have spatially non-uniform slip distributions on fault planes.

Based on these fault models, we performed numerous tsunami calculations to study the relationship between coastal tsunami heights and the absolute value of the hydrostatic pressure change. We solved the nonlinear shallow water equations for the tsunami calculations expressed as,

$$\begin{aligned} \frac{\partial Q_x}{\partial t} + \frac{1}{R \sin \theta} \frac{\partial}{\partial \varphi} \left(\frac{Q_x^2}{H + \eta} \right) + \frac{1}{R} \frac{\partial}{\partial \theta} \left(\frac{Q_x Q_y}{H + \eta} \right) \\ = - \frac{g(H + \eta)}{R \sin \theta} \frac{\partial \eta}{\partial \varphi} - \frac{f^2 g}{(H + \eta)^{7/3}} Q_x \sqrt{Q_x^2 + Q_y^2}, \end{aligned} \tag{1}$$

$$\begin{aligned} \frac{\partial Q_y}{\partial t} + \frac{1}{R \sin \theta} \frac{\partial}{\partial \varphi} \left(\frac{Q_x Q_y}{H + \eta} \right) + \frac{1}{R} \frac{\partial}{\partial \theta} \left(\frac{Q_y^2}{H + \eta} \right) \\ = - \frac{g(H + \eta)}{R} \frac{\partial \eta}{\partial \theta} - \frac{f^2 g}{(H + \eta)^{7/3}} Q_y \sqrt{Q_x^2 + Q_y^2}, \end{aligned} \tag{2}$$

$$\frac{\partial \eta}{\partial t} = - \frac{1}{R \sin \theta} \left[\frac{\partial Q_x}{\partial \varphi} + \frac{\partial (Q_y \sin \theta)}{\partial \theta} \right], \tag{3}$$

where η is the water height from the sea surface at rest, t is time, φ and θ are the longitude and co-latitude, respectively. R is the earth’s radius, H is the water depth, and the variables Q_x and Q_y are depth-integrated quantities equal to $(H + \eta)u$ and $(H + \eta)v$, where u and v are flow velocities, along longitude and latitude lines, respectively. The parameter f is Manning’s roughness coefficient, and g is the gravitational constant. Equations (1) and (2) are equations of motion. Equation (3) is an equation of continuity. In this study we used a parallelized tsunami calculation code called JAGURS (Baba et al. 2013b, 2015), which solves the Eqs. (1), (2), and (3) in a staggered, leap-frog finite difference scheme with nesting algorithms. Five nesting layers were used in this analysis. Figure 1a shows the computational domain and output points. The grid spacing of the finest grid was 2 / 9 arcsec (~ 5 m) and we used it for the Owase area in Mie Prefecture in central Japan. We simulated tsunami propagation for 3 h following the earthquake and recorded waveforms at every second for all the tsunami fault models ($T = 3 \text{ h} \times 3600 \text{ s/h} = 10800 \text{ s}$).

Data set

From this simulation, we obtained time series of water height changes η and water depth H at Owase and

DONET1 stations ($S = 20$ places). We define the maximum absolute value of the tsunami height at Owase in scenario n ($n = 1, \dots, N$) as $d(n)$. The time series of total depth (depth of the sea H + tsunami height η) at DONET1 station i ($i = 1, \dots, S$) is converted to hydrostatic pressure changes $p_i(n, t)$ ($t = 1, \dots, T$) by assuming that 1hPa is equivalent to a 1-cm change in depth. In the simulations, the depth of the sea was also affected in accordance with crustal deformation due to faulting.

Figure 1b shows the tsunami wave forms at the Owase tide station, or the absolute value of hydrostatic pressure fluctuations $|p_i(n, t)|$ ($i = 1, \dots, S$) in the case of scenario $n = 1229$. In this study, we predict $d(n)$ at Owase from the maximum absolute value of the hydrostatic pressure changes, $\mathbf{x} = [x_1, x_2, \dots, x_S]$, at DONET1 stations.

To evaluate the prediction accuracy of our method, we randomly divided simulated scenarios (1506 cases) into training data (2 / 3 of all cases) and test data (1 / 3 of all cases). We set the above divide and gather statistics to both construct our prediction algorithm, using adequate training data, and to validate the prediction accuracy. Based on the training data, we optimized our algorithm to fit d from $\mathbf{x} = [x_1, x_2, \dots, x_S]$. Applying our algorithm to test data, we evaluate the prediction accuracy and compared our method to the previous one.

Methodology

Maximum mean regression

Baba et al. (2013a) took particular note of the peak absolute value in hydrostatic pressure changes, recorded at all DONET1 stations during a tsunami (Fig. 1b). They found a clear relationship between the average waveform of the values at each DONET1 point and the maximum tsunami heights at the coast. Although they used a heuristic method to choose the peak absolute values, we employed the maximum absolute value as the representative point for all the DONET1 stations, $x_i(n) = \max_t |p_i(n, t)|$ ($i = 1, \dots, S$), which produced results similar to their method. To be more precise, we can write the prediction algorithm (maximum mean (MM) algorithm), using the mean value of the maximum absolute value, $(1/S) \sum_i^S x_i$, as

$$\hat{d}_{\text{MM}}(n) = w_{\text{MM}}^1 \frac{1}{S} \sum_{i=1}^S x_i(n) + w_{\text{MM}}^0, \quad (4)$$

where $\hat{d}_{\text{MM}}(n)$ represents the predicted maximum tsunami height for scenario n and w_{MM}^1 and w_{MM}^0 are regression coefficients. MM algorithm, which is a simple linear regression, fits $d(n)$ by $\frac{1}{S} \sum_{i=1}^S x_i(n)$ and determines the

regression coefficients by using the training data set ($N_T = 1004$ cases) and the least squares approach.

Gaussian process regression

There is a prediction bias in MM because the relationship between d and $\frac{1}{S} \sum_{i=1}^S x_i(n)$ is nonlinear and the regression is strongly affected by the large number of low tsunami heights.

There is no direct causal relationship between \mathbf{x} and d , but \mathbf{x} is indirectly correlated with d since these values are the result of a common fault model and simulation. Thus, the correlation, such as a nonlinear function system or polynomial expression, remains unknown. Instead we apply GP regression as a method of interpolation (Rasmussen and Williams 2005), which is widely used for prediction or optimization in practical fields (e.g. Kocijan et al. 2004; Krause et al. 2008). GP regression estimates maximum tsunami height for a test data set using weighted sum of tsunami heights of all the training data sets. Each weight is determined by the Gaussian function of the distance between DONET1 observed values of the test data set and each training data set as denoted below. GP regression approximates a non-linear relationship and gives unbiased prediction of intermediate values, assuming not a relationship between \mathbf{x} and d but noise variance or how to calculate the distance between test and training data sets, which are determined by a few parameters.

Let us introduce the formulation of GP regression in this paragraph. A GP is a generalization of the multivariate Gaussian probability distribution. Given the sensor values at DONET1 stations in scenario n , $\mathbf{x}(n) = \{x_1(n), \dots, x_S(n)\}$, the prediction $d_{\text{GP}}(n)$ at a point $\mathbf{x}(n)$ and its variance $v_{\text{GP}}(n)$ are described by using the Gaussian kernel function,

$$k(\mathbf{x}(n), \mathbf{x}(m)) = \exp\left(-\beta|\mathbf{x}(n) - \mathbf{x}(m)|^2\right), \quad (5)$$

where m is a scenario number and β represents the inverse of the length-scale of the Gaussian kernel. If we assume a normal distribution with noise variance, σ^2 as a prior, we obtain

$$d_{\text{GP}}(n) = (\mathbf{k}(n))^T (\mathbf{K} + \sigma^2 \mathbf{I})^{-1} \mathbf{d}, \quad (6)$$

where $\mathbf{k}(n) = [k(\mathbf{x}(1), \mathbf{x}(n)), \dots, k(\mathbf{x}(N_T), \mathbf{x}(n))]$ is a vector of the kernel between $\mathbf{x}(n)$ and training data, $(\cdot)^T$ represents the matrix transpose, and \mathbf{I} is the identity matrix of size N . $\mathbf{d} = [d(1), d(2), \dots, d(N_T)]^T$ represents the tsunami height of training data, and \mathbf{K} consists of the kernels between training data. As shown in Eq. (6), maximum tsunami height for a test data set is estimated by weighted sum of tsunami heights of all the training data sets and the

weight depends on the Gaussian kernel. The prediction variance is described by

$$v_{GP}(n) = k(\mathbf{x}(n), \mathbf{x}(n)) - (\mathbf{k}(n))^T (\mathbf{K} + \sigma^2 \mathbf{I})^{-1} \mathbf{k}(n). \quad (7)$$

Figure 2 shows examples of one-dimensional data interpolation by GP regression ($S = 1$). The black points represent given training data, and we want to predict a tsunami height when $x_1 = 4$. The solid line represents the estimated maximum tsunami height by using GP and Eq. (6) and the asterisk represents the estimated maximum tsunami height \hat{d}_{GP} , corresponding to $x_1 = 4$. The interpolation by GP regression (the solid line) runs along the means of the normal distribution with the two-sided 95% confidence interval represented by gray-filled area, which is derived by Eq. (7). As shown in Fig. 2b, while interpolation goes well near $x_1 = 4$, where training data are dense, estimation of GP regression is much less accurate near $x_1 = 0$, given the lack of training data.

As denoted above, GP approximates a non-linear relationship by using only two parameters, β and σ . Let us consider how do the GP parameters affect the tsunami prediction. First, we focus on the parameter β of GP, which determines the width of the Gaussian kernel of each training data as shown by gray dashed lines in Fig. 2. When it is wide, the estimated line is smooth and reduce the accuracy (Fig. 2a). On the other hand, when the Gaussian kernel is narrow, the estimated prediction is over-fitted to the training data (Fig. 2c).

Next, we consider noise variance, σ^2 . If we assume that σ^2 is large, the estimated line is smooth and reduce the accuracy, similar to the case of wide Gaussian kernel (Fig. 2a). On the other hand, if σ^2 is small, the confidence interval represented by gray-filled area is narrow and the estimated prediction is over-fitted to the training data.

Thus, both of the GP parameters greatly affect GP estimation. To avoid the over-fit to the training data and keep the accuracy of prediction, we optimize the parameters β and σ^2 as the generalized error of cross validation (CV) is minimized as is described in detail below.

Cross validation

Our goal with GP regression is not to fit the training data but to construct a prediction algorithm with small prediction errors. To construct a prediction algorithm to be trained on limited training data without severe over-fitting, we determine optimal GP parameters by using CV. In CV, the training dataset of N_T cases is divided into two parts, one to train the prediction algorithm and the other to evaluate its generalization error. The method we use for partitioning data is L-fold CV. First we divide the data set into L parts: C_1, \dots, C_L . For each $l = 1, \dots, L$, we train the GP regression using data that are not in C_l . Then, we use this trained GP regression algorithm to predict the tsunami height d for the data in C_l and calculate the root-mean-square error,

$$RMSE_{Ge} = \sqrt{\frac{L}{N_T} \sum_{n \in C_l} [d(n) - \hat{d}(n)]^2}. \quad (8)$$

This training and testing procedure was repeated using different data partitioning and we obtained the generalization error. In summary, CV is a method for evaluating the prediction accuracy generalized to unknown data that are not used in training (Kohavi 1995). We then determined optimal GP parameters to minimize the generalization error. We also calculated the generalization error of the MM method and compared GP and MM.

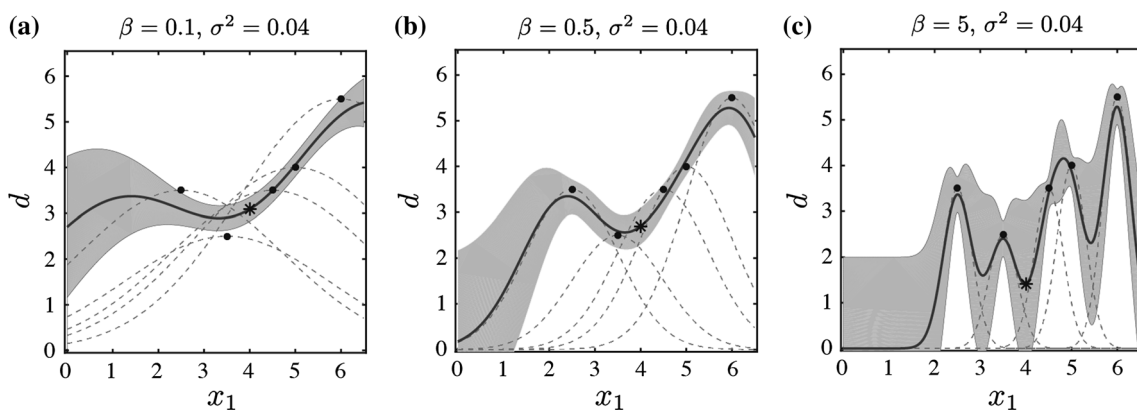


Fig. 2 Schematic diagrams of interpolation by GP regression with $\beta = 0.1, 0.5, 5$, respectively. The black points represent given training data and the gray dashed lines represent the Gaussian kernel of each training data. When β is small, the Gaussian kernel is wide. We fixed $\sigma^2 = 0.04$ and circles indicate the data locations and the asterisk

represents the estimated maximum tsunami height \hat{d}_{GP} , corresponding $x_1 = 4$. The interpolation by GP regression (the solid line) runs along the means of the normal distribution with the two-sided 95% confidence interval represented by gray-filled area

Results

Regression

GP regression can approximate a non-linear relationship between observed values at DONET1 and maximum tsunami height better than Maximum mean (MM) method. To compare the ability to approximate the non-linear relationship of GP and MM methods, we here predict the maximum tsunami height of Owase, using 1 dimensional data for GP, referred as “1dGP”. To be more precise, we used the mean value of the maximum absolute value of scenario n , $\bar{x}(n) \equiv (1/S) \sum_i^S x_i$ from $N_T = 1004$ training data. $\hat{d}_{1dGP}(n)$ at a point $\bar{x}(n)$ are described by Eq. (6) where we used the Gaussian kernel function,

$$k(\bar{x}(n), \bar{x}(m)) = \exp\left(-\beta|\bar{x}(n) - \bar{x}(m)|^2\right). \tag{9}$$

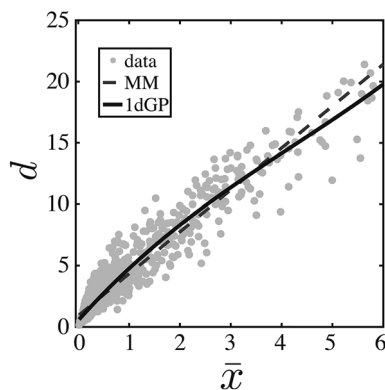


Fig. 3 Correlation between \bar{x} and d . Results of MM and 1 dimensional GP for a training data set with N_T cases. Dashed line and solid line represent the results of MM and GP method, respectively

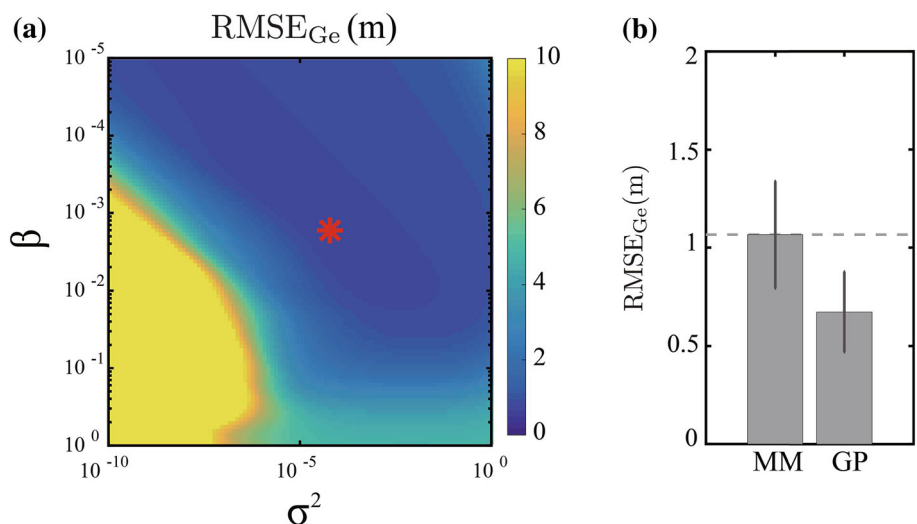
The prior parameters are β and σ^2 , optimized as the generalized error of L-fold CV ($L = 10$) is minimized. As shown in Fig. 3, \hat{d}_{MM} tends to be bigger in the case of a large tsunami height, such as 10m, since MM method is simple linear regression. On the other hand, GP regression approximates a non-linear relationship and gives unbiased prediction of intermediate values.

Next, let us consider GP regression using not the mean value of all observed points, $\bar{x}(n)$, but the individual values, $\mathbf{x} = [x_1, x_2, \dots, x_S]$. Using L-fold CV ($L = 10$), we determine optimal GP parameters, β and σ^2 , by minimizing the generalization error and constructing a prediction algorithm to be trained on limited data without severe over-fitting.

Figure 4a shows a map of the generalization error. The area of large generalization error has low noise variance ($\sigma^2 \approx 0$) and short Gaussian kernel length-scale ($\beta > 10^{-3}$) because of severe over-fitting. A quickly varying signal with low noise from this area would give rise to a white-noise process model for the signal, which is not a convincing data model. On the other hand, GP regression with moderate noise variance and long kernel length scale can support the data model and the GP regression can be generalized to unknown data not used in the training.

We thus obtain optimal GP parameters, $\beta = 1.7 \times 10^{-3}$ and $\sigma^2 = 5.7 \times 10^{-5}$, which yield the minimum generalization error (Fig. 4a). We compare the GP and MM generalization errors in Fig. 4b, and observe that both the mean and standard deviation of the GP generalization error are less than those of MM by 36 and 25 %, respectively. Based on optimal GP parameters and all the training data (1004 cases) and fitting d from $\mathbf{x} = [x_1, x_2, \dots, x_S]$ by GP regression, we construct a prediction algorithm for the maximum

Fig. 4 a Color represents the GP generalization error. The horizontal and vertical axes represent β and σ^2 respectively, on a log scale. The red star indicates the minimum point of the generalization error, $\beta = 1.7 \times 10^{-3}$ and $\sigma^2 = 5.7 \times 10^{-5}$. **b** Comparison of the GP and MM generalization errors with two-sided 95 % confidence intervals



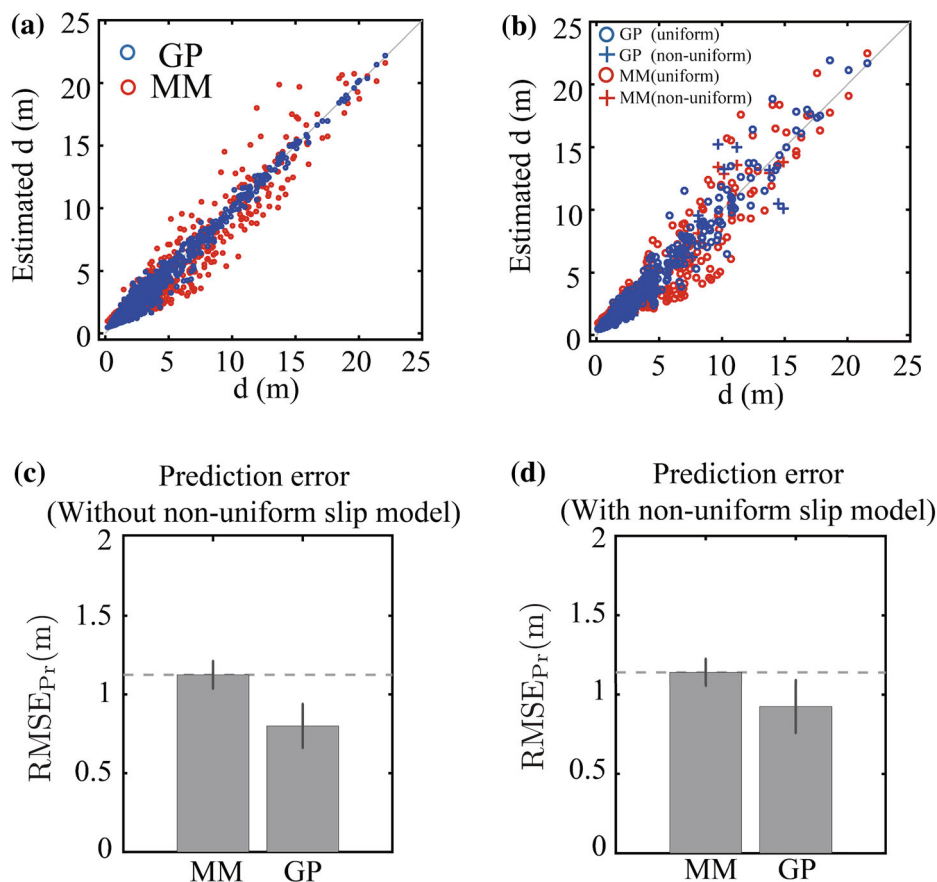


Fig. 5 **a** Relationship of estimated and simulated maximum tsunami heights at Owase using a training data set with N_T cases by GP (blue circles) and MM (red circles). The horizontal and vertical axes represent the simulated d , and estimated \hat{d} maximum tsunami heights, respectively. This figure represents that the estimation is good since many of the points are close to $d = \hat{d}$ (gray line). **b** Relationship of estimated and simulated maximum tsunami heights at Owase by MM and GP for prediction. Red circles and plus signs represent the results of MM using a test data set with N_P cases and non-uniform slip

models, respectively (Kanamori 1972; Baba and Cummins 2005; Cabinet office in Japan 2012). Blue circles and plus signs show the prediction results by GP for a test data set with N_P cases and the non-uniform slip models, respectively. The horizontal and vertical axes represent d , and \hat{d} , respectively. **c** Comparison of GP and MM prediction errors for a test data set with N_P cases with two-sided 95% confidence intervals. **d** Comparison of GP and MM prediction errors for a test data set with N_P cases and the non-uniform slip models with two-sided 95% confidence intervals

tsunami height. Figure 5a shows the relationship between estimated and simulated maximum tsunami heights at Owase. RMSE of training data by GP and MM are 0.52 and 1.07, respectively and the GP training error also dramatically decreases by 52% compared with MM.

We thus constructed GP regression model from $N_T = 1004$ training data that predicts the maximum tsunami height at Owase, d , from the sensor values of DONET1 stations, \mathbf{x} .

Prediction

Finally, by applying the algorithm to test data ($N_P = 502$ cases) of uniform slip models, we evaluate the prediction accuracy of our method, using the root-mean-square error (RMSE_{Pr}) between d and \hat{d}_{GP} , and to compare GP and MM

(Fig. 5b). We observe that the GP prediction error for uniform slip models is 0.78 and decreases by 39% compared with MM, and GP regression greatly improves the prediction accuracy. The variance of the prediction error decreases without bias, including large tsunami heights (>10 m) where the MM estimation tends to be larger.

Moreover, we show the prediction results by GP and MM about the 1944 Tonankai, 1946 Nankai earthquakes and 11 kinds of anticipated M9-class non-uniform slip models in the Nankai trough released by the Cabinet Office of the government of Japan as blue and red plus signs for GP and MM in Fig. 5b, respectively (Kanamori 1972; Baba and Cummins 2005; Cabinet office in Japan 2012). The prediction results for non-uniform slip models are generally consistent with those for uniform slip models, although there are a few cases that the prediction errors of GP are

larger than those of MM. We argue about the cause of the largest prediction error in discussion.

To indicate that the improvement of prediction accuracy is not affected by the division into training dataset and prediction dataset, we performed the same protocol 10 times to keep statistics on the mean and variance of prediction accuracy. Figure 5c, d show the average value of prediction RMSE ($RMSE_{Pr}$) of GP and MM only for uniform slip models and both for uniform and non-uniform slip models, respectively. Figure 5c shows that although the standard deviation of $RMSE_{Pr}$ of GP increases by 59% in comparison with that of MM, the average $RMSE_{Pr}$ of GP for uniform slip models decreases by 29% and prediction accuracy greatly improves independent of the divide of dataset, which is well accorded with the results of generalization error in Fig. 4b. Similar to the results for uniform slip models, the average $RMSE_{Pr}$ of GP for uniform and non-uniform slip models is 19% less than the one by MM method, as shown in Fig. 5d. Thus, the improvement of prediction accuracy is not affected by the division into training dataset and prediction dataset.

Although we predict the tsunami heights using 3-h time series of DONET1 data from the earthquake occurring time in this paper, we can apply this method to short-time tsunami prediction. Optimizing the GP method for only the 10-min time series of the ones used in Fig. 5a, b, we found that the prediction error of GP method is 1.09 m, including non-uniform slip models, which is 18% larger than the one given full time series and is 19% less than the one by MM method as shown in Fig. 6. These are almost the same as the results using 3-h time series data as shown in Fig. 5b, except that, if the tsunami does not arrive at DONET1 within 10 min, predicted tsunami height is constant in these cases. Thus, the GP method can also work for short-time tsunami height prediction and further investigations are needed to evaluate the short-time prediction of GP method.

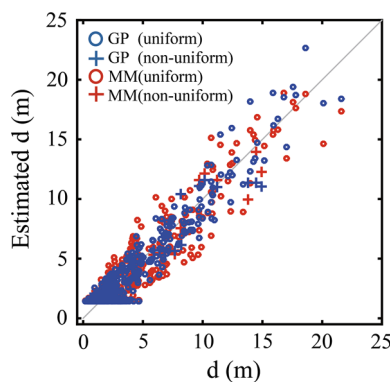


Fig. 6 Results of short-time tsunami height prediction by MM and GP using only 10-min data following the earthquake in the same way as Fig. 5b

Discussion

We developed an accurate prediction algorithm for maximum tsunami height in the Kii Peninsula of Japan by a Gaussian process (GP) that uses pressure gauge data.

It is obvious that Maximum Mean (MM) regression, which is a previously used network-averaged prediction method, will not be able to deal with future expanded observational networks such as DONET2 (west of the Kii Peninsula) (Kaneda et al. 2015) and S-net (along the Japan Trench) (Uehira et al. 2012; Saito 2013), since the large variety of scenarios measured by these systems cannot be uniformly approximated. Thus, it makes sense to construct an algorithm to predict maximum tsunami height based not on averaged but on individual measurements at DONET1 stations.

In this study we interpolate not a direct causal but a non-linear relationship between maximum tsunami height and DONET1 sensory data by using GP regression, which is an interpolation method governed by a few prior parameters (Rasmussen and Williams 2005). The prediction error by GP regression greatly decreases by about one third in comparison with MM regression, which makes predictions assuming a linear relationship with average gauge data from DONET1. Moreover, for large tsunami heights (>10 m), although MM estimates tend to be larger, GP estimates greatly improve without this bias (Fig. 5b). These results show that GP regression enables us to abstract the complex and nonlinear relationship between tsunami height and pressure gauge data and to predict tsunami height more accurately.

Although the prediction results by GP for non-uniform slip models are generally smaller than those by MM, there are a few cases that the prediction errors by GP are larger. We then discuss about the cause of the largest prediction error ($\max|d - \hat{d}| = 5.56$ m) in a non-uniform slip model as shown in Fig. 5b. We calculate the smallest exponential part of the Gaussian kernel (Eq. (5)), which is the distance between test data n and training data m ($m = 1, \dots, N_T$), ($\min_m |\mathbf{x}(n) - \mathbf{x}(m)|$), because, if it is small, test data are close to training data and GP can interpolate well. We found that the minimum distance in the case of the maximum GP prediction error is 11.6 times larger than that of the mean minimum distance. This indicates that the GP prediction error tends to be larger because test data cannot be interpolated by using training data. Furthermore, this suggests that in unexpected circumstances the tsunami height prediction of GP could be less accurate.

Therefore, to utilize GP regression for tsunami height prediction, we have to minimize unanticipated scenarios by preparing a huge variety of scenarios in advance that include actual diverse phenomena. In concrete terms, although we assume a uniform amount of slip for each fault to simplify the simulation procedure, actual earthquakes

have spatially non-uniform slip distributions on fault planes and in the simulations we deal with the non-uniformity to enhance the prediction capability. Furthermore, to expand the predictable tsunami scenario, we can add the non-uniform slip models, which are represented by blue plus signs in Fig. 5b, for training data set of GP.

However, no matter how well we prepare, in reality, unintended scenarios could happen where extrapolation is needed. If so, although test data are far from training data and we can expect the GP prediction error is apt to be large, we can prepare tsunami predictions using a conventional method such as MM, which is probably the lesser of the two methods for extrapolating to unexpected scenarios.

In this paper, using GP regression, we predict the maximum tsunami height of only one station (Owase). For practical use, different estimations are necessary to be estimated for several stations along the coast of the Kii Peninsula. Different optimal parameters are necessary to be estimated since β and σ^2 of GP regression change every point along the coasts of Kii Peninsula. Due to the smoothness of coastal topography, the optimal parameters of neighboring points could be preferred to change gradually along the coasts. The proposed method in this paper can be extended to take into account of the smoothness of coastal topography in the optimal parameters of our method. We will study how the smoothness of the coastal topography improves reliable prediction rather than one station in the future work.

Conclusion

We constructed a methodology to predict coastal maximum tsunami height by a Gaussian process (GP) that uses offshore pressure gauge data from DONET1. We found that our methodology can greatly reduce the prediction error for uniform slip models; about one third of that by a previous method which tends to make larger predictions, especially for large tsunami heights (>10 m). We can extend our method using GP to tsunami height prediction for several stations and add the non-uniform slip models for training data set. Further investigations for these extensions are needed for high-accuracy tsunami height prediction.

Acknowledgments The comments from two anonymous reviewers and the editor were constructive and helped improve this manuscript. Part of this research was financially supported by a Grant-in-Aid for Scientific Research on Innovative Areas (No. 25120009) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and CREST project “Establishing the most advanced disaster reduction management system by fusion of real-time disaster simulation and big data assimilation” from the Japan Science and Technology Agency.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Allen SCR, Greenslade DJM (2010) Model-based tsunami warnings derived from observed impacts. *Nat Hazards Earth Syst Sci* 10:2631–2642
- Baba T, Cummins PR (2005) Contiguous rupture areas of two Nankai Trough earthquakes revealed by high-resolution tsunami waveform inversion. *Geophys Res Lett* 32(8):1–4
- Baba T, Tanioka Y, Cummins PR, Uehira K (2002) The slip distribution of the 1946 Nankai earthquake estimated from tsunami inversion using a new plate model. *Phys Earth Planet Inter* 132(1–3):59–73
- Baba T, Hirata K, Kaneda Y (2004) Tsunami magnitudes determined from ocean-bottom pressure gauge data around Japan. *Geophys Res Lett* 31:2–5
- Baba T, Hirata K, Hori T, Sakaguchi H (2006) Offshore geodetic data conducive to the estimation of the afterslip distribution following the 2003 Tokachi-oki earthquake. *Earth Planet Sci Lett* 241(1):281–292
- Baba T, Takahashi N, Kaneda Y (2013a) Near-field tsunami amplification factors in the Kii Peninsula, Japan for Dense Oceanfloor Network for Earthquakes and Tsunamis (DONET). *Mar Geophys Res* 35(3):319–325
- Baba T, Takahashi N, Kaneda Y, Inazawa Y, Kikkojin M (2013b) Tsunami inundation modeling of the 2011 Tohoku earthquake using three-dimensional building data for Sendai, Miyagi prefecture, Japan. In: Fandinò VS (ed) *Tsunami events and lessons learned; ecological and societal significance*. Springer, Amsterdam
- 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 Appl Geophys* 172(12):3455–3472
- Cabinet office panel (2012) Government of Japan, The 2nd report of the possible mega earthquake in the Nankai trough (in Japanese), http://www.bousai.go.jp/jishin/chubou/nankai_trough/nankai_trough_top.html
- Fine IV, Cherniawsky JY, Thomson RE, Rabinovich AB, Krassovski MV (2015) Observations and numerical modeling of the 2012 Haida Gwaii tsunami off the coast of British Columbia. *Pure Appl Geophys* 172:699–718
- Hayashi Y (2010) Empirical relationship of tsunami height between offshore and coastal stations. *Earth, Planets Space* 62(3):269–275
- Hsiao NC, Lin TW, Hsu SK, Kuo KW, Shin TC, Leu PL (2014) Improvement of earthquake locations with the Marine Cable Hosted Observatory (MACHO) offshore NE Taiwan. *Mar Geophys Res* 35(3):327–336
- Kamigaichi O (2009) Tsunami forecasting and warning. In: Meyers RA (ed) *Encyclopedia of complexity and systems science*. Springer, Heidelberg, pp 9592–9618
- Kanamori H (1972) Tectonic implications of the 1944 Tonankai and the 1946 Nankaido earthquakes. *Phys Earth Planet Inter* 5:129–139
- Kaneda Y, Kawaguchi K, Araki E, Matsumoto H, Nakamura T, Kamiya S, Ariyoshi K, Hori T, Baba T, Takahashi N (2015)

- Development and application of an advanced ocean floor network system for megathrust earthquakes and tsunamis, in *Seafloor Observatories*, Springer Praxis Books, pp 643–666
- Kocijan J, Murray-Smith R, Rasmussen CE, Girard A (2004) Gaussian process model based predictive control. In: *Proceedings of the 2004, IEEE American Control Conference, 2004*, vol 3. pp 2214–2219
- Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *International Joint Conference on Artificial Intelligence*, vol 14. pp 1137–1143
- Krause A, Singh A, Guestrin C (2008) Near-optimal sensor placements in Gaussian processes: theory, efficient algorithms and empirical studies. *J Mach Learn Res* 9:235–284
- Kumar TS, Nayak S, Kumadr P, Yadav RBS, Kumar A, Sunanda MV, Devi EU, Shenoi SSC (2012) Performance of the tsunami forecast system for the Indian Ocean. *Curr Sci* 102(1):110–114
- Münch U, Rudloff A, Lauterjung J (2011) The GITEWS project results, summary and outlook. *Nat Hazards Earth Syst Sci* 11(3):765–769
- Nosov MA, Kolesov SV (2007) Elastic oscillations of water column in the 2003 Tokachi-oki tsunami source: in-situ measurements and 3-D numerical modelling. *Nat Hazards Earth Syst Sci* 7:243–249
- Percival DB, Denbo DW, Eblé MC, Gica E, Huang PY, Mofjeld HO, Spillane MC, Tang L, Titov VV (2011) Extraction of tsunami source coefficients via inversion of DART@bouy data. *Nat Hazards* 58:567–590
- Percival DB, Denbo DW, Eblé MC, Gica E, Huang PY, Mofjeld HO, Spillane MC, Titov VV, Tolkova EI (2014) Detiding DART@-Buoy data for real-time extraction of source coefficients for operational tsunami forecasting. *Pure Appl Geophys* 172(6):1653–1678
- Rasmussen CE, Williams CKI (2005) *Gaussian processes for machine learning*. MIT Press, Cambridge
- Saito T (2013) Dynamic tsunami generation due to sea-bottom deformation: analytical representation based on linear potential theory. *Earth Planets Space* 65(12):1411–1423
- Thomson R, Fine I, Rabinovich A, Mihály S, Davis E, Heesemann M, Krassovski M (2011) Observation of the 2009 Samoa tsunami by the NEPTUNE-Canada cabled observatory: test data for an operational regional tsunami forecast model. *Geophys Res Lett*. doi:10.1029/2011GL046728
- Titov VV, Gonzalez FI, Eblé MC, Mofjeld HO, Newman JC, Venturato AJ (2005) Real-time tsunami forecasting: challenges and solutions. *Nat Hazards* 35:35–41
- Tsushima H, Hino R, Fujimoto H, Tanioka Y, Imamura F (2009) Near-field tsunami forecasting from cabled ocean bottom pressure data. *J Geophys Res Solid Earth* 114(6):1–20
- Tsushima H, Hirata K, Hayashi Y, Tanioka Y, Kimura K, Sakai S, Shinohara M, Kanazawa T, Hino R, Maeda K (2011) Near-field tsunami forecasting using offshore tsunami data from the 2011 off the Pacific coast of Tohoku earthquake. *Earth, Planets Space* 63(7):821–826
- Tsushima H, Hino R, Ohta Y, Iinuma T, Miura S (2014) tFISH/RAPiD: rapid improvement of near-field tsunami forecasting based on offshore tsunami data by incorporating onshore GNSS data. *Geophys Res Lett* 41:3390–3397
- Uehira K, Kanazawa T, Noguchi S, Aoi S, Kunugi T, Okada Y, Sekiguchi S, Shiomi K, Shinohara M, Yamada T (2012) Ocean bottom seismic and tsunami network along the Japan Trench. *AGU Fall Meeting, Abstract OS41C-1736*
- Utsu T (2001) *Seismology*, 3rd edn. Kyoritsu Pub, Tokyo (**in Japanese**)