

# ガウス過程回帰による海底水圧データからの最大津波高予測

## Maximum tsunami height prediction using Ocean-Bottom Pressure Values based on Gaussian Process Regression

高橋 光太郎 \*1      柏原 健之朗 \*1      五十嵐 康彦 \*1\*2      馬場 俊孝 \*3      堀 高峰 \*4  
Kotaro Takahashi      Kenjiro Kashiwabara      Yasuhiko Igarashi      Toshitaka Baba      Takane Hori

岡田 真人 \*1  
Masato Okada

\*1 東京大学大学院新領域創成科学研究科      \*2 科学技術振興機構 さきがけ  
Graduate School of Frontier Sciences, The University of Tokyo      Japan Science and Technology Agency PRESTO

\*3 徳島大学大学院社会産業理工学研究部  
Institute of Technology and Science, The University of Tokushima

\*4 海洋研究開発機構地震津波海域観測研究開発センター  
Research and Development Center for Earthquake and Tsunami, Japan Agency for Marine-Earth Science and Technology

Tsunami early warning systems using water pressure gauges operate around the world to cope with damage caused by tsunami waves. The systems use a correlation between observed pressure gauges value around coast and tsunami height at prediction points near shore, because tsunami height basically depends on the topography (bathymetry) during its propagation. In predicting the tsunami height, it is important both avoiding underestimation and increasing accuracy in order to minimize the damage. The conventional method selects an scenario which has the largest tsunami height of near the observed pressure gauges value in the tsunami database that contains pre-computed tsunamis offshore and nearshore from 1506 earthquake scenarios. Although this conventional method can avoid under estimation, it puts the prediction accuracy second. In this study, we extended tsunami height prediction method using Gaussian Process regression and proposed a prediction method with less underestimation and higher accuracy. We investigate the prediction accuracy and the possibility of underestimation by our proposed method that uses pressure gauges data from the Dense Ocean-floor Network System for Earthquakes and Tsunamis (DONET) in the Nankai trough.

### 1. はじめに

大規模な地震に伴う津波による被害を抑えるためには津波高即時予測システムが重要となる。我が国においても津波高即時予測を目的として、東北地方太平洋沖に展開されている日本海溝海底地震津波観測網 (S-net) や和歌山県沖合で運用されている地震・津波観測監視システム (DONET) などでは海底に水圧センサーが敷設され、高密度観測網によって水圧データをリアルタイムで取得することができる。実際、2011年の東日本大震災における大津波を、沖合の水圧センサーは早期に計測しており、津波高即時予測に有効であると考えられている [Baba 14]。

津波はレイイベントであり、実観測データがごく少ないため、高精度な津波シミュレーション [Baba 16] を用いることで様々なシナリオを含むデータベースを構築する。そして、あらかじめ構築されたデータベースを用いて、沖合の水圧センサー値から沿岸の津波高を即時予測することが可能である [Baba 14, Igarashi 16, Takahashi 18, 高橋 成実 今井 健太郎 18, Masashi ss]。また、津波高の予測においては被害を最小限に抑えるために精度の良さに加えて過小評価をしないことが重要である。従来法においては水圧センサーの観測値から想定される沿岸の津波高のうち最も大きいシナリオを予測値として算出している [Takahashi 18, 高橋 成実 今井 健太郎 18]。この手法による予測は既に和歌山県で独自の津波高予測として実装されているが [Takahashi 18,

高橋 成実 今井 健太郎 18], 過小評価を避けるために実際の津波高に比べて大きく予測され、誤差の大きい手法となっている。

そこで本研究では、ガウス過程回帰を利用した津波高を予測する手法 [Igarashi 16] を拡張し、最大津波高を正確に予測することで、過小評価が少なくかつ予測精度が高い津波高予測手法を提案する。本研究では、南海トラフ上の DONET1 の観測点と紀伊半島沿岸の尾鷲市における水位をシミュレーションしたデータベースを用いる。このとき、提案手法による DONET1 センサー値から尾鷲市における津波高予測を行ない、津波高予測の精度及び、過小評価の可能性について検証する。

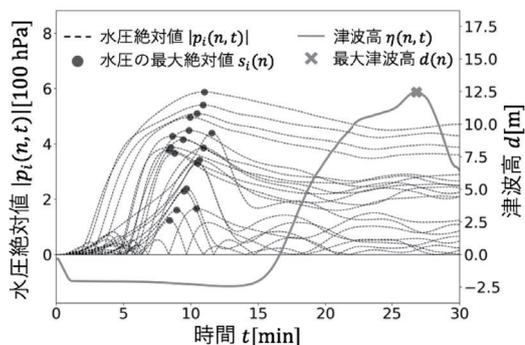


図 1: 水圧センサー値。黒色の破線は水圧センサーで観測された値の絶対値で丸は水圧センサー値の最大値を表す。実線は尾鷲市で観測された津波高でばつ印は最大津波高を示す。

連絡先: 岡田真人, 東京大学大学院新領域創成科学研究科,  
okada@edu.k.u-tokyo.ac.jp

## 2. データセット

本研究では、高精度津波シミュレーションによって南海トラフ上の DONET1 の観測点と紀伊半島沿岸の尾鷲市における水位をシミュレーションし、データベースを構築した [Baba 14]. 津波シミュレーションは津波計算コード JAGURS [Baba 16] を用いて、断層の位置やマグニチュードなど 10 個のパラメータを持つ地震シナリオを様々に設定して 1506 通りのシミュレーションを行った [Baba 14, Igarashi 16]. 図 1 のように、津波シミュレーションにより DONET1 の  $i$  番目 ( $i = 1, 2, \dots, S = 20$ ) の水圧センサーで観測された水圧値と尾鷲検潮所に到達する津波高が 1 秒ごと  $T = 180$  分間の時系列データとして記録される。

全時間予測の検証では、 $T = 180$  分間の水圧データが全て手元にあると仮定し、説明変数は水圧センサー値  $p_i(t)$  [hPa],  $0 < t \leq T$  の最大絶対値  $s_i = \max_t |p_i(t)|$  とする。また、リアルタイム予測の検証では、時刻  $\tau \leq T = 180$  分までの水圧センサー値が観測された状況を想定し、 $p_i(t)$  の時刻  $\tau$  までの最大絶対値  $s_{i,\tau} = \max_{t \leq \tau} |p_i(t)|$  を説明変数とする。予測値として、図 1 に示す尾鷲検潮所での津波高時系列データの最大値  $d$  を目的変数に設定し、各手法の予測値  $\hat{d}$  との平方根平均二乗誤差 (Root Mean Square Error, RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{n=1}^N [d(n) - \hat{d}(n)]^2} \quad (1)$$

によって予測精度を検証する。ここで  $n$  はシナリオ番号を表し、 $n = 1, \dots, N = 1506$  である。

## 3. 手法

### 3.1 従来法

まず、Takahashi らが提案し、和歌山県における津波予測システムに組み込まれている従来法について説明する [Takahashi 18, 高橋 成実 今井 健太郎 18]. 津波高  $d$  が、 $S = 20$  点の水圧センサーの最大絶対値平均  $s_{\text{mm}} = \frac{1}{S} \sum_{i=1}^S s_i$  と図 2 のように強く相関することを用いて [Baba 14] 沿岸の最大津波高予測を行う。

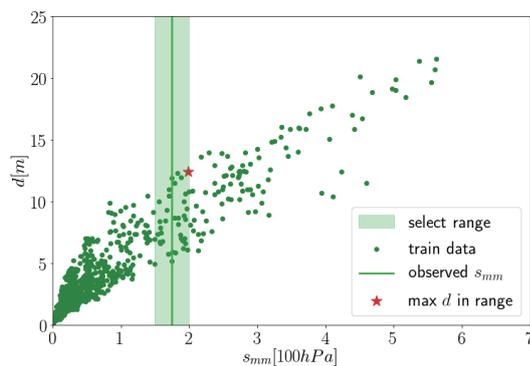


図 2: 波源シナリオの選択。緑の点は訓練データを表し、縦の緑線は観測された水圧値を示す。緑の網掛けは水圧値の閾値を表しており、本研究では 25 [hPa] と設定。赤点は範囲内の最大津波高をもたらすシナリオを表し、この手法での予測値である。

図 2 の縦緑線で表される  $s_{\text{mm}}$  が水圧センサーから得られたとする。このとき、観測によって得られた最大絶対値平均  $s_{\text{mm}}$

に近いシナリオの中で最大の沿岸津波高を選択する。本研究では観測値に近いシナリオを図 2 の網掛け部分のように、観測された最大絶対値平均  $s_{\text{mm}}$  の  $\pm 25$  [hPa] に設定した。

### 3.2 提案法

水圧センサー値と沿岸津波高の間の相関を利用した手法として、非線形関数を近似できるノンパラメトリック手法であるガウス過程 (Gaussian Process, GP) による予測手法 [Igarashi 16] を改善した手法を提案する。ガウス過程回帰では、観測された水圧センサー値  $\mathbf{s}_*$  が与えられた時、津波高  $d_*$  の予測分布は式 (2) のようにガウス分布で得られる。

$$p(d_* | \mathbf{d}, \mathbf{s}_*) = \mathcal{N}(\bar{d}_*, \text{var}[d_*]) \quad (2)$$

$$\bar{d}_* = \mathbf{k}_*^T [K + \sigma_n^2 I_{N_T}]^{-1} \mathbf{d} \quad (3)$$

$$\text{var}[d_*] = k(\mathbf{s}_*, \mathbf{s}_*) + \sigma_n^2 - \mathbf{k}_*^T [K + \sigma_n^2 I_{N_T}]^{-1} \mathbf{k}_* \quad (4)$$

$$k(\mathbf{s}(i), \mathbf{s}(j)) = \sigma_f^2 \exp(-\beta |s(i) - s(j)|^2) \quad (5)$$

ただし、目的変数  $d$  は潜在関数  $f$  に平均 0、分散  $\sigma_n^2$  のガウス分布に従うノイズが付加されたモデルを仮定する。また、 $\mathbf{d} = [d_1, \dots, d_{N_T}]^T$ ,  $\mathbf{f} = [f_1, \dots, f_{N_T}]^T$  である。さらに、カーネル  $k$  は式 (5) のように定義する。ハイパーパラメータ  $\beta$  はガウス関数の幅の逆を表す。 $\sigma_f^2$  は信号分散と呼ばれ、値が小さいと平均値に近い関数、値が大きいほど関数を取りうる形のバリエーションが増える。式 (3), (4) の  $\mathbf{k}_* = [k(\mathbf{s}(1), \mathbf{s}_*), \dots, k(\mathbf{s}(N_T), \mathbf{s}_*)]^T$  はテストデータ  $\mathbf{s}_*$  と訓練データ  $\mathbf{s}(m)$  ( $m = 1, \dots, N_T$ ) の共分散ベクトル、 $K$  は  $i, j$  成分が  $k(\mathbf{s}(i), \mathbf{s}(j))$  ( $i, j = 1, \dots, N_T$ ) である訓練データの共分散行列、 $I_{N_T}$  は  $N_T \times N_T$  の単位行列を表す。

本研究では、過小評価を避けるため式 (6) のように 95% 信頼区間の最大値を予測値とする。その為、式 (6) において  $\alpha = 1.96$  とする。

$$\hat{d}_{\text{proposed}} = \bar{d}_* + \sqrt{\text{var}[d_*]} \times \alpha \quad (6)$$

また、ハイパーパラメータの最適化基準にはベイズ自由エネルギー最小化を用いた。ガウス過程回帰において、ハイパーパラメータの事後分布は式 (7) で表せられる。ただし、ハイパーパラメータの集合を  $\boldsymbol{\theta} = \{\beta, \sigma_f^2, \sigma_n^2\}$  と表す。ハイパーパラメータの事前分布  $p(\boldsymbol{\theta})$  を一様分布とする。さらに  $p(\mathbf{d})$  は  $\boldsymbol{\theta}$  に関して定数なので式 (8) と書くことができる。このことから式 (10) に示すベイズ自由エネルギーを最小化することでハイパーパラメータの事後確率を最大化することができ、最適なハイパーパラメータを推定することができる。

$$p(\boldsymbol{\theta} | \mathbf{d}) = \frac{p(\mathbf{d} | \boldsymbol{\theta}) p(\boldsymbol{\theta})}{p(\mathbf{d})} \quad (7)$$

$$\propto p(\mathbf{d} | \boldsymbol{\theta}) \quad (8)$$

$$= \int p(\mathbf{d} | \mathbf{f}, \boldsymbol{\theta}) p(\mathbf{f} | \boldsymbol{\theta}) d\mathbf{f} \quad (9)$$

$$-\log p(\mathbf{d} | \boldsymbol{\theta}) = \frac{1}{2} \mathbf{d}^T (K + \sigma_n^2 I_{N_T})^{-1} \mathbf{d} + \frac{1}{2} \log |K + \sigma_n^2 I_{N_T}| + \frac{1}{2} \log 2\pi \quad (10)$$

一方で、ガウス過程回帰による津波高予測を行った先行研究においては、式 (3) で表される  $\bar{d}_*$  を予測値としている。また、ハイパーパラメータ  $\sigma_f^2$  は 1 と固定し、 $\beta, \sigma_n^2$  の最適化基準には 10 分割交差検定 (10-fold Cross Validation, CV) を用いている [Igarashi 16]. この先行研究と本手法による最大津波高予測の結果と比較を行う。

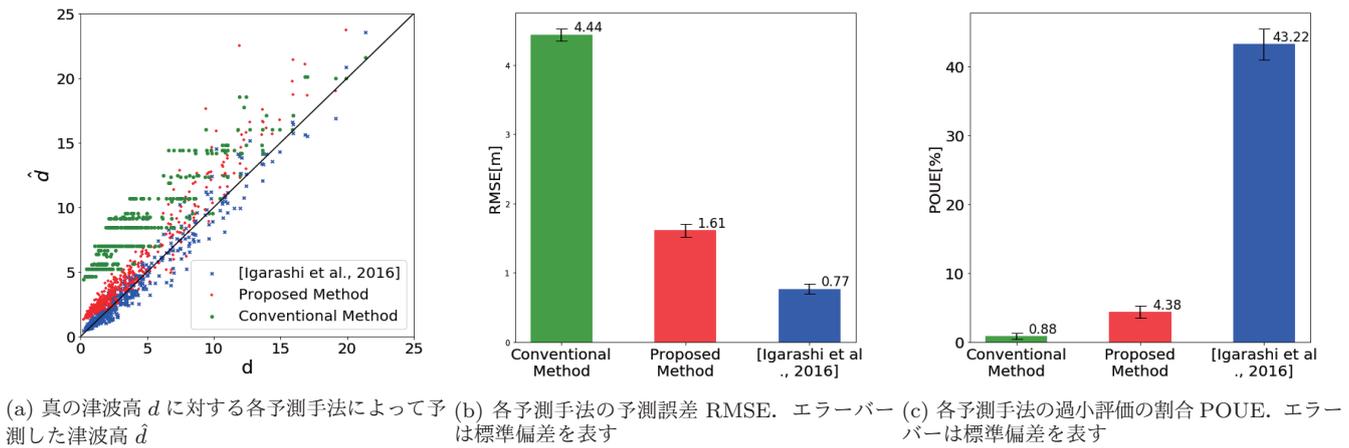


図 3: 全時間予測での結果比較図

## 4. 結果

### 4.1 全時間予測

シミュレーションにより生成された全 180 分間の水圧センサー値の最大絶対値  $\mathbf{s} = \{s_i\}_{i=1}^{20}$  を説明変数として、津波高  $d$  を予測する検証を行なった。予測誤差は、式 (1) で表される平方根平均二乗誤差 (Root Mean Square Error, RMSE) を用いる。また、過小評価の指標として以下の過小評価率 (Percentage of Underprediction Errors, POUE)[Toth 15]

$$\text{POUE} = \frac{|\{i = 1, \dots, N | d(n) > \hat{d}(n)\}|}{N} \quad (11)$$

を用いる。

全 1506 個のデータのうちランダムに選出した 1004 個のデータを訓練データとし津波高予測モデルの学習に用い、残りの 502 個のデータをテストデータとして津波高予測性能の検証に用いた。実際の津波高  $d$  に対するテストデータの的水圧から予測した津波高  $\hat{d}$  のプロットを図 3a に示す。直線  $\hat{d} = d$  に近いほど誤差が小さいことを表す。従来法に比べてガウス過程回帰を使用した提案法および五十嵐らの手法 [Igarashi 16] は、予測値が直線  $\hat{d} = d$  に近いため、より正確に予測できることが見て取れる。また、図 3a において直線  $\hat{d} = d$  の下にプロットされているものは実際の津波高よりも予測値の方が小さい過小評価となっている。五十嵐らの手法 [Igarashi 16] では平均値を予測値としているため半分程度は過小評価しているが、最大津波高を予測する提案法および従来法では過小評価が少なくなっている。

次に、データの分け方をランダムに変えながら 100 回試行した平均予測誤差と平均過小評価率を評価した結果を、図 3b, 3c にそれぞれ示す。従来法、五十嵐らの手法 [Igarashi 16] および提案手法の予測誤差は、4.44[m], 0.77[m], 1.61[m] となり、提案手法の予測誤差は、五十嵐らの手法 [Igarashi 16] の予測誤差と比べて 0.84[m] の増加したが、従来法に比べて 63.7% 減少する結果を得た。さらに、従来法、五十嵐らの手法 [Igarashi 16] 及び提案法の過小評価率は 0.88[%], 43.22[%], 4.38[%] となることから、五十嵐らの手法 [Igarashi 16] に比べて過小評価率を 38.84[m] 下げるとともに、従来法と比べても 3.50% の増加に抑えられる結果を得た。

### 4.2 リアルタイム予測

実際の予測システムでは、地震発生後 180 分間の水圧データを収集するのではなく、より短時間での正確な予測を行う

ことが求められる。そこで、時刻  $\tau \leq 180$  分までの水圧センサー値の最大絶対値  $\mathbf{s}_\tau = \{s_{i,\tau}\}_{i=1}^{20}$  を説明変数として、津波高  $d$  を予測する検証を行なった。時刻  $\tau = 1, 2, \dots, 180$  分において訓練データとテストデータの分け方をランダムに変えながら予測を行い、時刻  $\tau$  における平均予測誤差を算出した。

地震発生後の時刻に対する各予測手法での予測誤差の変化を図 5 に示す。いずれの手法も地震発生後 10 分程度で予測誤差は急激に小さくなり、50 分以降ではほぼ一定となった。また、時刻  $\tau$  における平均予測過小評価率を算出し図 6 に示した。過小評価率 POUE は全ての手法において時間に関わらずほぼ一定を記録していることがわかる。 $\tau = 10, 20, 40$  分において、テストデータから予測した津波高  $\hat{d}$  を実際の津波高  $d$  に対してプロットを図 4 に示す。地震発生から短時間の予測でもガウス過程回帰と提案法は従来法よりも真値に近い予測値を取っていることが分かる。また、提案法と従来法は地震発生から短時間の予測でも POUE が 10% を下回っており、過小評価が少ないことがわかる。

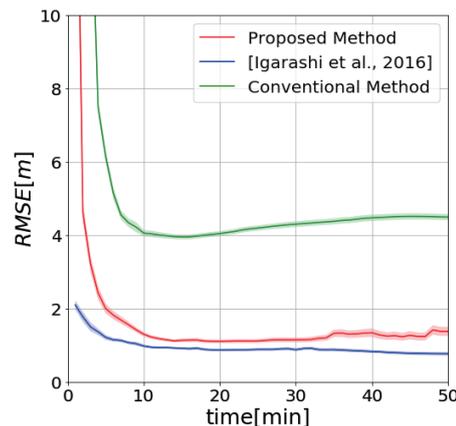


図 5: 地震発生後の時刻に対する各予測手法の予測誤差 RMSE. 青色は [Igarashi 16], 赤色は提案法, 緑色は従来法. 区間は標準偏差を表す。

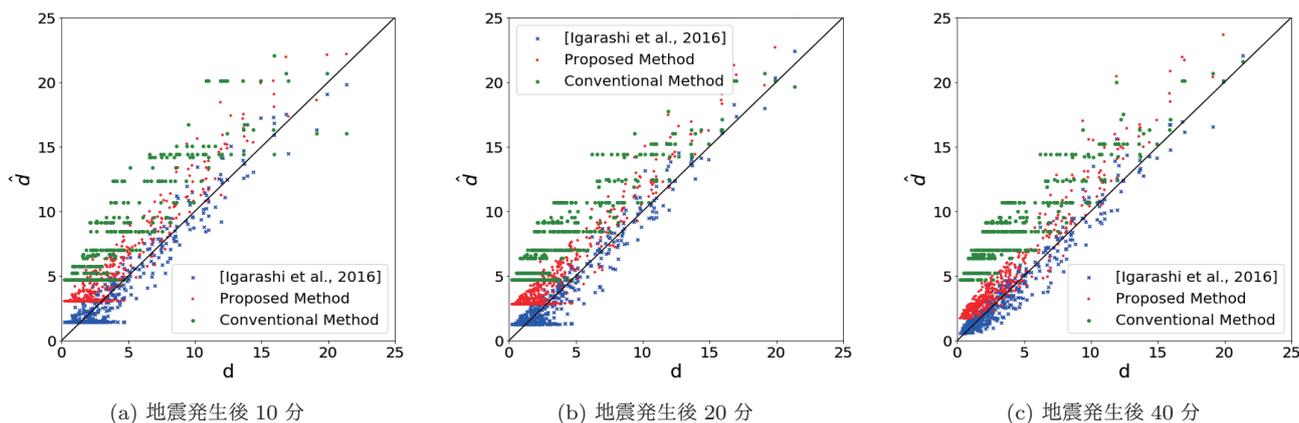


図 4: 地震発生からの時間別の真の津波高  $d$  に対する各予測手法によって予測した津波高  $\hat{d}$

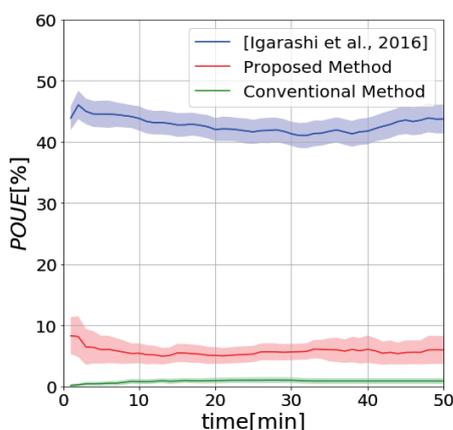


図 6: 地震発生後の時刻に対する各予測手法の過小評価率 POUE. 青色は [Igarashi 16], 赤色は提案法, 緑色は従来法. 区間は標準偏差を表す.

## 5. 結論

沿岸津波高を予測する手法には、観測された水圧センサーの値を中心とした任意の範囲内の最大津波高を予測津波高とする従来法 [Takahashi 18, 高橋 成実 今井 健太郎 18] と、水圧センサー値からガウス過程を用いて津波高の予測分布を求め、その平均値を予測値とする手法 [Igarashi 16] がある。本研究ではガウス過程回帰 [Igarashi 16] を改良し、過小評価が少なく精度も高い予測手法を提案した。提案法では、ガウス過程回帰におけるカーネルやハイパーパラメータ最適化手法の変更に加え、過小評価を防ぐために予測分布の 95% 信頼区間の最大値を予測値とした。

これらの手法を用いて DONET1 の水圧センサー値から津波高を予測する検証を行なった結果、提案法は、従来法と比べて、予測誤差は 2.83[m] も低くなり、過小評価率については、提案法は従来法に比べて、3.5% の増加に留まった。また、提案法は先行研究の手法 [Igarashi 16] に比べて、予測誤差は 0.84[m] の増加に抑え、過小評価率については、38.84% 減少した。以上から、水圧センサー値から沿岸津波高を予測する際は、過小評価が少なく、精度も高い提案手法が有効であることが分かった。

## 参考文献

- [Baba 14] Baba, T., Takahashi, N., and Kaneda, Y.: Near-field tsunami amplification factors in the Kii Peninsula, Japan for Dense Oceanfloor Network for Earthquakes and Tsunamis (DONET), *Marine Geophysical Research*, Vol. 35, No. 3, pp. 319–325 (2014)
- [Baba 16] Baba, T., Ando, K., Matsuoka, D., Hyodo, M., Hori, T., Takahashi, N., Obayashi, R., Imato, Y., Kitamura, D., Uehara, H., et al.: Large-scale, high-speed tsunami prediction for the Great Nankai Trough Earthquake on the K computer, *The International Journal of High Performance Computing Applications*, Vol. 30, No. 1, pp. 71–84 (2016)
- [Igarashi 16] Igarashi, Y., Hori, T., Murata, S., Sato, K., Baba, T., and Okada, M.: Maximum tsunami height prediction using pressure gauge data by a Gaussian process at Owase in the Kii Peninsula, Japan, *Marine Geophysical Research*, Vol. 37, No. 4, pp. 361–370 (2016)
- [Masashi ss] Masashi, Y., Igarashi, Y., Shin, M., Hori, T., Sato, K., Baba, T., and Okada, M.: A nonlinear parametric model based on power law for tsunami height prediction, *Marine Geophysical Research* (in press)
- [Takahashi 18] Takahashi, N., Ohbayashi, R., Baba, T., Imai, K., Ishibashi, M., Kaneda, Y., Sueki, K., and Tanabe, T.: Realtime Tsunami Prediction System Using Oceanfloor Network Data and its Social Implementation, in *2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO)*, pp. 1–4IEEE (2018)
- [Toth 15] Toth, E.: Asymmetric error functions for reducing the underestimation of local scour around bridge piers: application to neural networks models, *Journal of Hydraulic Engineering*, Vol. 141, No. 7, p. 04015011 (2015)
- [高橋 成実 今井 健太郎 18] 高橋 成実 今井 健太郎 石橋 正信: DONET 観測情報を活用した津波予測システムの社会実装—和歌山県の事例—, *自然災害科学*, Vol. 37, No. 1, pp. 125–142 (2018)