CrossMark

Assessment of tsunami hazards in ports and their impact on marine vessels derived from tsunami models and the observed damage data

Abdul Muhari[1,2] · Ingrid Charvet[3] ·
Futami Tsuyoshi[4] · Anawat Suppasri[1] ·
Fumihiko Imamura[1]

Abstract This paper presents a detailed study of tsunami hazard in ports and its corre-lation with the damage suffered by marine vessels. The study aims to develop a new loss function to estimate the potential damage of marine vessels due to tsunami attack based on a novel multivariate statistical modeling method, which used several explanatory variables simultaneously to estimate an outcome or the probability of such outcome. In the first part of the paper, tsunami heights and velocities are numerically modeled by using high-resolution bathymetry and topography data for the southern part of Honshu Island. We apply statistical methods to the complete sequence of spatially distributed time series of tsunami parameters in order to obtain the best fit with the observed damage data. In the

✉ Abdul Muhari
abdul.muhari@gmail.com

Ingrid Charvet
i.charvet@ucl.ac.uk

Futami Tsuyoshi
Tsuyoshi.Futami@willis.com

Anawat Suppasri
suppasri@irides.tohoku.ac.jp

Fumihiko Imamura
imamura@irides.tohoku.ac.jp

[1] Tsunami Engineering Lab, International Research Institute of Disaster Science (IRIDeS), Tohoku University, 6-6-11-1106 Aoba, Sendai 980-8579, Japan

[2] Present Address: The Ministry of Marine Affairs and Fisheries, Mina Bahari Bld. 7th Floor, Jakarta 10110, Indonesia

[3] Earthquake and People Interaction Centre (EPICentre), University College London-UCL, London, UK

[4] Willis Research Network, Willis-Re, Tokyo, Japan

second part, we develop loss functions for marine vessels by using ordinal regression, which uses simultaneously several explanatory variables. We perform several statistical tests to determine the appropriate model variables to be used in developing three-dimensional loss estimation surfaces, which provide probability of loss for each combination of measured or simulated values of tsunami parameters. The main feature of the developed loss functions presented in this study is their capability to integrate the key factors influencing the damage probability, such as tsunami parameters, characteristics of marine vessels and the impact of collision experienced by the vessels during the tsunami. Such a robust method, therefore, is crucially important to understand the tsunami impact on ports and, particularly on marine vessels.

1 Introduction

Maritime and fishery were among the most affected sectors during the 2011 Great East Japan Earthquake–tsunami. A report published by the Ministry of Agriculture, Forestry and Fisheries (MMAF 2012) stated that 28,612 marine vessels, 319 ports and approximately 1,725 marine facilities were damaged by the tsunami, which brought economic losses of approximately US$12 billion. Among those losses, only US$2.3 billion were covered by insurance (Kajitani et al. 2013). In Japan, such a huge loss has increased the urgency to understand tsunami hazards in ports and their potential damage to the maritime sectors, particularly to marine vessels (e.g., Shuto 1987; Aketa et al. 1994; Kawata et al. 2004; Kazama et al. 2006; Tsubota et al. 2007; Kato et al. 2009).

In line with the increasing amount of data availability, research to understand the potential impacts of tsunamis on ports and marine vessels has become widespread. Lynett et al. (2014) developed a damage index of ports and their facilities based on the numerically predicted or instrumentally recorded data of tsunamis current in California. Previously, Suppasri et al. (2013a) and Muhari et al. (2013) introduced loss functions to indicate the damage probability of marine vessels subjected to tsunami hazards. They used the observed and simulated tsunami height and velocity and the observed damage data of approximately 21,000 marine vessels during the 2011 Great East Japan–tsunami. The latter can be considered to be possibly the most comprehensive damage data that have ever existed.

The study to understand tsunami hazards and their associated impacts in ports requires not only a large number of observed damage data of marine vessels but also reliable tsunami observation data. In ports, even without flooding, maritime assets are vulnerable to significant damage from strong currents and the associated drag forces (Lynett et al. 2014). Unlike the commonly used tsunami height data, the observed tsunami current, in particular, has rarely been utilized for tsunami research due to the lack of the observed tsunami current data during tsunami events. The latter has led to a poor understanding of the mechanisms of tsunami-induced current hazards (lynett et al. 2012, 2014). To address this problem, numerical simulations with specific conditions have been recommended as a potential solution (e.g., Admire et al. 2014; Lynett et al. 2014). By using a high-approximation numerical scheme and/or high-resolution bathymetry data, a reliable tsunami
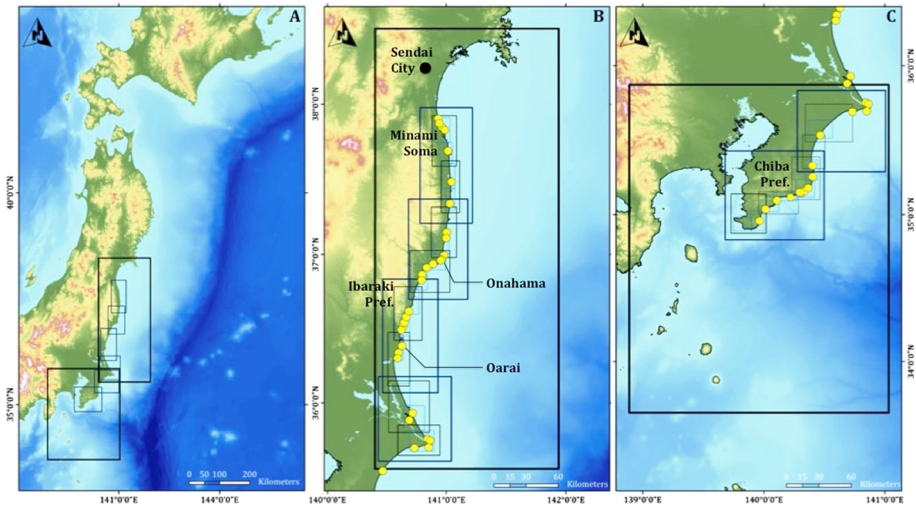
Fig. 1 Domain setting for the numerical model used in this study. **A** is the first domain with 270-m grid accuracy and the second domain with 90-m grid accuracy are indicated by *thick line boxes*. *Boxes with thin lines* in **A** are the third domain with 30-m grid accuracy. *Boxes with blue lines* in **B** and **C** indicate the smallest domains with 10-m grid accuracy. *Yellow dots* in **B** and **C** are locations of the ports analyzed in this study

simulation around the port can be obtained. With reliable tsunami simulation and complete damage data of port and marine vessels after a tsunami, our understanding of tsunami hazards and their associated impacts for maritime sectors can be improved.

In this paper, we aim to provide a reliable numerical simulation of tsunamis in ports to be used in developing a new loss function for marine vessels. We take the 2011 Great East Japan Earthquake–tsunami as the study case. We simulate the tsunami height and velocity in 38 ports located in the southern part of Honshu Island, from Minami Soma to the southern part of Ibaraki Prefecture (Fig. 1). The simulated tsunami heights and velocities are then used as the input for statistical models to develop loss estimation surfaces, as the result of empirical regression analysis to estimate the potential loss from damage to marine vessels during the tsunami event.

2 Data and methodology

2.1 Tsunami simulation

In simulating the tsunami hazards in ports, a high-approximation numerical scheme, on the one hand, results a longer computational time that might not be convenient for practical purposes. On the other hand, the commonly used numerical scheme in tsunami modeling with shallow water approximation (e.g., Titov and González 1997; Imamura 1996) has an efficient computational time with relatively good accuracy. However, they should be equipped with high-resolution bathymetry and topographic data with an accuracy of at least 10 m if one would like to use them for detailed modeling of tsunami hazard in ports (Uslu et al. 2013; Lynett et al. 2012, 2014; Admire et al. 2014).

By referring to the above-mentioned conditions, this study chooses to use the world standard tsunami model (IOC-UNESCO 1997), which utilizes the shallow water approximation and nested grid system as a basis for detailed simulations including the advance friction factor to accommodate tsunami inundation in built environment (Imamura 2009). This model has been extensively used in tsunami research worldwide, and it has been benchmarked with both experimental and the observed tsunami data (i.e., Imamura 1996). Therefore, the details of the model will not be presented in this paper. The bathymetry data for the numerical model were obtained from the Disaster Management Agency, The Cabinet Office of Japan, which compiled all available sources in Japan prior to the 2011 Great East Japan Earthquake. These data were published with a 10-m grid resolution, which is suitable to simulate tsunami hazards in ports as required by the aforementioned references. We selected 38 ports in the southern part of Honshu Island to be modeled because these ports and their facilities, including breakwaters, were not severely damaged by the tsunami. Such a condition allows us to model the tsunami by integrating breakwaters and the other built structures around the ports into the digital elevation model or the so-called topographic model (Muhari et al. 2011). Such model is not applicable if one simulates the tsunami in the ports located in the northern part of Honshu Island or the so-called Sanriku coast, where most of the ports' structures were damaged by the tsunami.
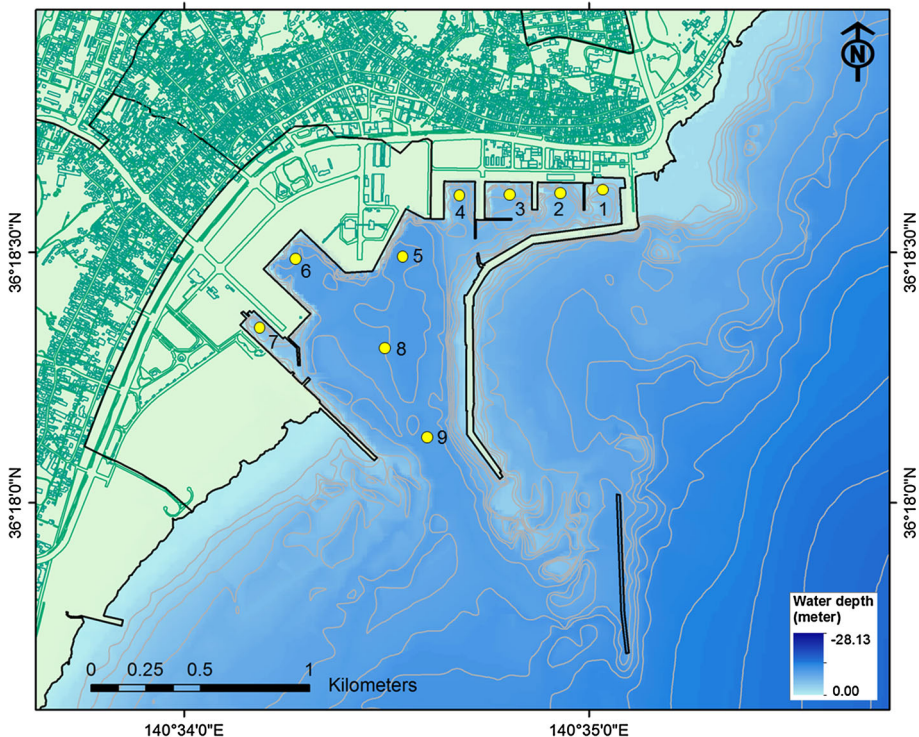

Fig. 2 Example of locations of the extracted points in Oarai Port, Ibaraki, Japan. Points 1–3 are locations of small- to medium-size fishing boats, points 4–6 are the ferry ports, point 7 is the private marina, and points 8–9 are the representation of the navigation channel

The common method to analyze tsunami hazards in ports, particularly for non-static objects such as marine vessels, is to correlate the resulting impact observed after tsunami with the maximum simulated and/or observed tsunami parameters—tsunami heights and velocities in the region. This assumption might not be generally accepted. For instance, the tsunami height and velocity that initiates movement of small marine vessels should be different than the values that influence cargo ships or large marine vessels. Moreover, in both cases, the maximum tsunami height and velocity might not correspond to the actual wave conditions that triggered the initial ship motion due to the time variation of those parameters at the point where the ships are previously moored. Depending on its size, a ship might have been swept away by the tsunami before the maximum height or velocity occurred. Secondly, the mechanism of damage to the ship by the tsunami is unclear. The tsunami may initiate the movement of the ship, but evidence that the tsunami force is responsible for the destruction of the ships is rarely found in the available literature, as would happen for a static object such as coastal structures or buildings (e.g., Koshimura et al. 2009; Suppasri et al. 2013b; Charvet et al. 2014a, b).

Our strategy to address these problems was to extract the tsunami height and velocity from both spatial and time series format in order to capture the complete figure and sequence of the tsunami attack. In each port, five to ten extraction points are placed in locations where the marine vessels were most likely anchored before tsunami arrived. Figure 2 shows nine initial locations of the vessels studied in Oarai Port, which is one of the 38 ports modeled in this study. This covers all possible locations of ships in each port, regardless of the type of ship (i.e., ferry, yacht, cargo, and fishing boat). Locations of the points in each port were determined through a thorough analysis using satellite imagery, port documentation, and aerial photographs, including oblique photographs provided by the Geospatial Survey Agency of Japan.

During the tsunami, several peaks of tsunami height and velocity might occur, and the values of tsunami parameters that initiated the movement and are eventually sweeping away of the vessels are not known. We, therefore, take into account all peak values of tsunami height and velocity for each time series and each location until it reaches a maximum value for each parameter. The statistical model used to provide an estimation of loss probability, that is the probability that a marine vessel will suffer a given loss or belong to a certain loss category, was then applied to each relevant combination of ship location. Through sensitivity analysis, we select values of tsunami height and velocity from the wave conditions that give the best fit to the loss data, which were then chosen to develop the final loss estimation functions. Loss data refer to the sample of reported financial loss ratios (or loss ratio intervals) suffered by the marine vessels as a result of the tsunami.

2.2 Loss data of the marine vessels

We used the surveyed loss data of marine vessels in Suppasri et al. (2013a) and Muhari et al. (2013) that comprise detailed damage and loss information for approximately 21,000 marine vessels from approximately 200 ports along the east coast of Japan. The data provide information about the situation of the boat before the tsunami, location of the boat after the tsunami, tonnage, materials, type of engine, and type or cause of damage, such as sinking, collision, stranding (Fig. 3). Such details of damage data bring valuable information to be used later in the statistical analysis.

In contrast to previous research (Suppasri et al. 2013a; Muhari et al. 2013), this study applied several conditions to reduce uncertainty of the selected vessel data. For instance,
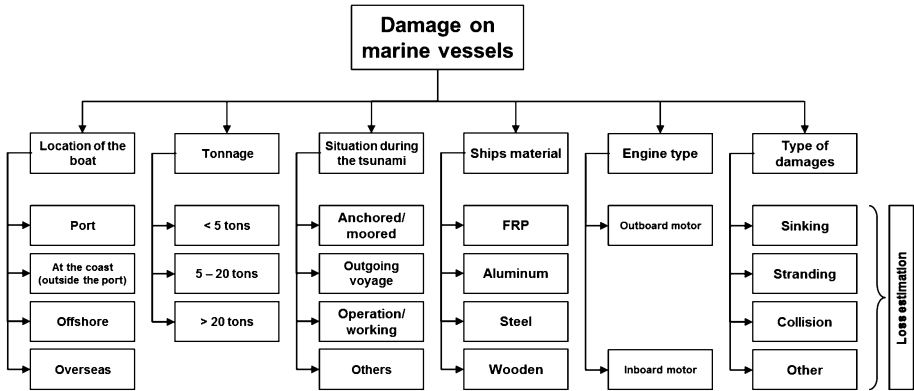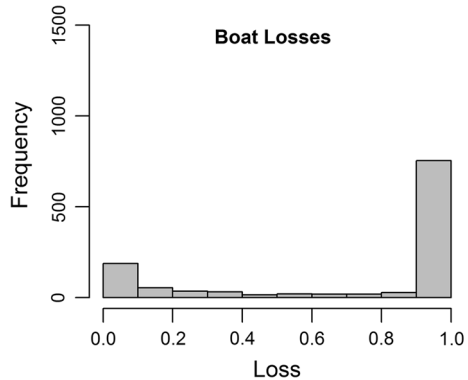
Fig. 3 Classification of the observed damage data of marine vessels in the ports along the east coast of Japan during the 2011 East Japan tsunami

Fig. 4 Frequency distribution of vessels' loss data

we only choose marine vessels that were located inside the port before the tsunami. Next, we used only ones that were previously anchored or moored in the port before the tsunami. This means the ships were in a static condition before they were hit by the tsunami. We excluded the vessels that were located outside of the port or offshore and the ones that were not moored before the tsunami arrived because the values of tsunami height and velocity used to assess their loss become highly uncertain.

As a result of such sorting procedures, only 1,167 vessels of the total 21,000 vessels data were used in the analysis. Among them, 82 % are small marine vessels of less than 5 tons, 16 % are ones between 5 tons and 20 tons, and the other 2 % are large vessels of more than 20 tons. The large number of small marine vessels brings fact that 97 % of the vessels' materials are fabricated reinforced plastic (FRP), 1 % aluminum, 1 % steel, and 1 % wood.

The loss ratio is defined as the ratio between the amounts of insurance paid over the total insurance value. The loss ratio is categorized into 10 classes that represent a 10 % increment from no damage at all, or less than 10 % loss (class 1), to complete damage, or 100 % loss (class 10), where the total insured value is paid. Approximately 90 % of the data correspond to 100 % losses because most of the small marine vessels cannot be used after the tsunami or were completely damaged by the tsunami (Fig. 4). The ships that

suffered small damages were the large marine vessels (up to 400 tons), where the paid insurance is less than 10 % of the total insured value. By using these criteria, we attempt to simplify the understanding of the probability of exceeding certain critical loss ratio ranges, as opposed to a measure of average loss, which requires more user knowledge to make an assessment and understand the implications (Suppasri et al. 2013a).

By previously having the detailed simulated tsunami parameters at hand, we developed a new loss function to estimate the loss probability of marine vessels subjected to tsunami parameters based on the advanced statistical modeling method proposed by Charvet et al. (2014a). To be unique, treatment in predicting damage probability for static structures (e.g., Koshimura et al. 2009; Suppasri et al. 2013b; Charvet et al. 2014a, b) was enhanced in order to better predict loss probability of marine vessels. As a first step, the loss is associated with the tsunami height, the tsunami velocity, and the combination of both. In this step, we use values from the extracted points in each port that gave the best fit with the loss data. Next, we improved the loss estimation function by introducing the "internal aspects" of the vessels, such as tonnage and vessel material. Finally, we take into account the "external aspects," such as the impact of a ship's collision with either static or non-static objects during the tsunami, which is likely the most influential damage factor. Because the statistical method used in this study allows the simultaneous consideration of several predictor variables that have potential to influence the boat damage, it is necessary to develop three-dimensional loss estimation curves, or surface, to represent loss probability. The developed loss estimation surfaces here are therefore capable of achieving a more realistic prediction of loss probability of marine vessels during the tsunami event.

3 Modeling of the 2011 Great East Japan Earthquake-tsunami

We designed the numerical domains for computation with four different bathymetric and topographic data resolutions (see Fig. 1). The first and the second domains are gridded with 270- and 90-m cell size, respectively. Next, the detailed data from Minami Soma in the southern part of Sendai City until the area in the southern part of Chiba Prefecture are covered by 13 numerical domains with 30-m grid size and 24 domains with 10-m grid size and data resolutions. This numerical domain consists of 38 ports where a total of 1,167 marine vessels were damaged by the 2011 tsunami. In general, the tsunami propagation model was running for 2 h of modeled time, but in some ports, like Oarai and Onahama, the modeled time was extended up to 4 h for spatial validation of the model.

To start the numerical simulation, an earthquake source model should be determined. It is acknowledged that after the 2011 Great East Japan Earthquake, there have been many earthquake source models proposed based on different methods and data sets. Because the purpose of this study is not intended to contribute in developing the earthquake source model, we decided to use a composite source model proposed by the Japan Nuclear Energy Safety (JNES 2011; Sugino et al. 2013). This source model was developed based on the inversion of the observed tsunami waves in GPS buoys along the Tohoku coasts, DART buoys offshore Japan, and tide gauges in the north of Tohoku, Hokkaido, and in the south of Japan. The overall predicted source area is divided into 48 segments with dimensions of 50×50 km in general and 50×30 km near the Japan Trench line. The estimated amount of slip each segment varied from 0.0 m to a maximum of 77.9 m near the trench. This source model was previously developed to reproduce the observed tsunami in the nuclear power plant Fukushima Daichi in the Fukushima area, which is located inside the area of interest of this paper.
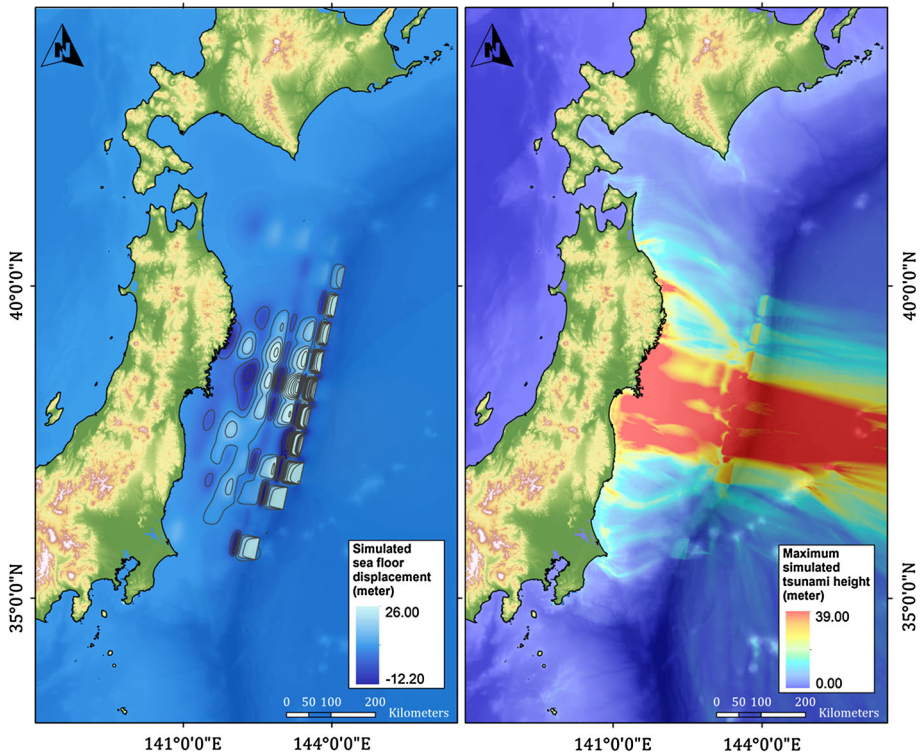
Fig. 5 The calculated sea floor displacement (*left*) and the simulated maximum tsunami height (*right*) along the east coast of Japan

Based on the predicted fault parameters, the initial sea surface condition was assumed to be the same as the sea bottom deformation, which was then calculated instantaneously by using the seismic deformation model proposed by Okada (1985). The calculated seafloor displacement yields a maximum 26 m uplift near the trench and up to 12 m subsidence along the coasts and in the offshore Tohoku region (Fig. 5). The simulated tsunami height along the Tohoku coast has a maximum value of 39 m at the southern part of the Sanriku coast (~38°–39°N).

By using an earthquake source similar to the one used in this study, previous research has provided the source model validation through the comparison between the resulting simulated tsunami heights and the observed sea-level data at the offshore GPS buoys and tsunami run-up height from all affected areas (JNES 2011; Sugino et al. 2013). Additionally, a satisfactory comparison between the simulated tsunami velocity in the northern part of Honshu Island and the observed data analyzed from amateur video in Kesennuma, and the Sendai plain has also been shown in Suppasri et al. (2013a). This study, therefore, focuses only on the validation of the simulated tsunami heights with the observation data (Mori et al. 2011) around the ports in the southern part of Honshu Island.

3.1 Tsunami parameters in the southern part of Honshu Island

Tsunami inundation models in the ports are run by using the topographic data that integrate the constructed environment in ports, such as breakwaters and jetties into the digital

elevation model. This means we assumed that these structures withstand the tsunami. Through a careful investigation by using various data sources, we ensured that the majority of the main structures of the ports analyzed in this study survived the tsunami. Some damage occurred in some ports particularly in the sea side of offshore breakwater, but such a small change cannot be accommodated by the existing model and such a factor is neglected due to incomplete information.

In each port, we extracted the simulated tsunami heights and compared them with the available observation data of tsunami run-ups surrounding the ports. In total, 211 observation points located near the coastline or near the ports' breakwater were used for model verification (Fig. 6). The geo-reference data of those verification points are available in supplementary data (Table S1).

In general, the simulated tsunami run-ups are consistent with the observed data. Some points have significant discrepancy, particularly the ones near the breakwater or in the back of coastal structures, which may indicate the limitation of the shallow water approximation. We validated the model by using the geometric mean of the ratio between the observed and simulated data (Aida 1978). The equations of this validation method are given as follows:

$$\log K = \frac{1}{n} \sum_{i=1}^{n} \log K_i \tag{1}$$

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

$$K_i = \frac{R_i}{H_i} \tag{3}$$

In the above equations, R_i and H_i are the observed and simulated data, respectively; K is the geometrical mean of K_i; κ is the derivation or the variance of K. From a total of 211 observational points, we obtained the values of K and κ as 1.17 and 1.37, respectively. These values are considered close to the threshold provided by the Japanese Society of Civil Engineers (JSCE 2002), which suggests that $0.95 < K < 1.05$ and $\kappa < 1.45$ are necessary for a model to have "good agreement" with the observational data.

Fig. 6 The comparison between the observed and simulated tsunami run-ups at the ports in the study area
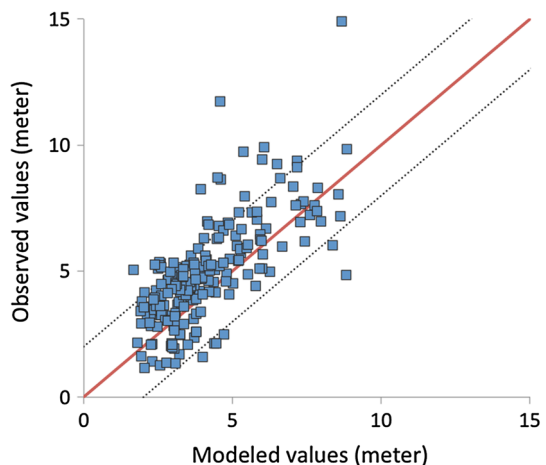
3.2 Tsunami parameters in the ports

We plotted the maximum simulated tsunami height and velocity and the snapshot of tsunami propagation inside the harbor to retrieve the spatial distribution of the simulated tsunami parameters. The snapshot of tsunami propagation inside the port (i.e., at Oarai Port in Fig. 7) demonstrates the capability of the model to draw the dynamics of tsunami current inside the port, including large vortex behavior that is observed several times during the simulation as it was recorded by many amateur videos (e.g., FNN 2012). For ports where the physical structures, such as breakwaters, withstand the tsunami, the maximum simulated tsunami velocity is mostly observed along and near the edge of the breakwater. The maximum simulated tsunami velocity was rarely located at or around the anchoring position of the ships, even though high values are sometimes observed inside a long and narrow marina. Such phenomena can lead to an overestimated prediction when one correlates the observed damage with the maximum tsunami parameters because the difference between the maximum tsunami velocities around the breakwater can be of the order of the square of the velocity values around the location of anchored ships.

To avoid such an overestimated prediction of the tsunami parameters, we analyzed the extracted tsunami time series to capture the peak values of tsunami heights and velocities along the simulation in each port (see e.g., locations of extracted points in Fig. 2). At each
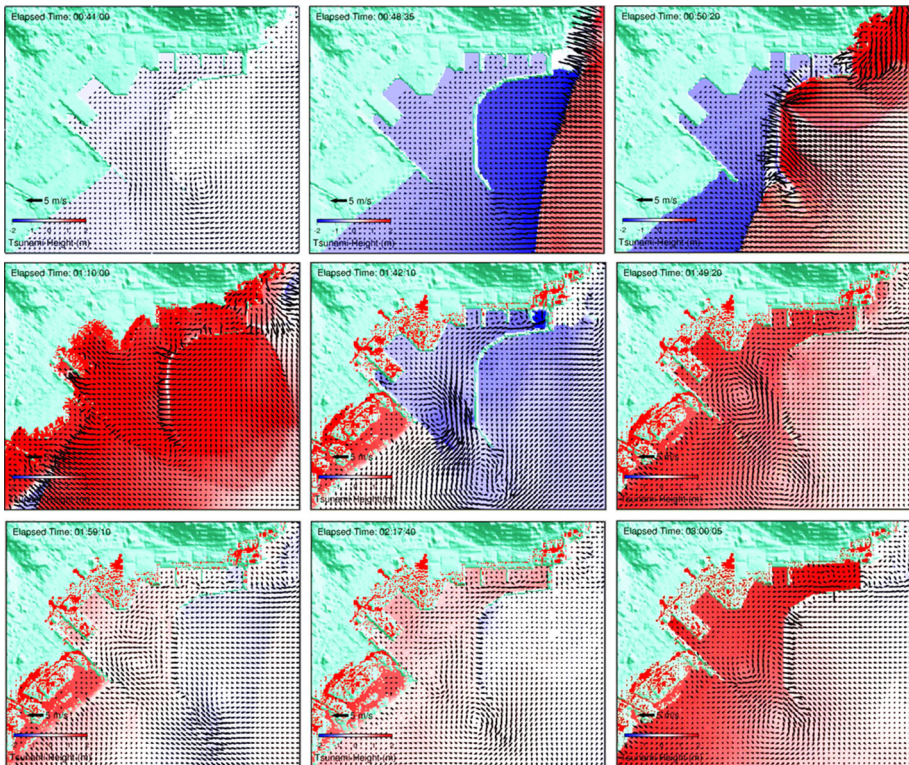

Fig. 7 Snapshot of tsunami propagation around the Oarai Port, Ibaraki, Japan. Elapsed time indicates snapshot time after the earthquake

point in each port, we took into account the peak values of tsunami height and tsunami velocity until they reached the maximum value recorded during the simulation. Once the maximum simulated tsunami height and velocity were obtained, the next peak values were no longer considered because the maximum height and velocity are the threshold for condition that initiates ships' movement. For instance, at point 7 in Oarai Port (see Fig. 2 for point location), the maximum value of tsunami height is obtained at Peak 4 (P4) and maximum simulated tsunami velocity is obtained in the third peak by considering that the captured time of tsunami velocity is similar to tsunami height. In this case, the fifth peak of the simulated tsunami height and the fourth and fifth peaks of the simulated tsunami velocity will not be used (Fig. 8).

At ports in study area, the maximum tsunami heights are mostly obtained in the third peak, while the maximum tsunami velocity is obtained in the fourth peak. The average height of the first to the fifth peaks of each point from 38 ports is 2.71, 3.75, 3.91, 3.93, and 3.94 m, respectively. For the tsunami velocity, the average values of the first to the fifth peaks are 2.29, 3.12, 3.36, 3.45, and 3.45 m/s, respectively (Table 1).

According to previous research, the damages suffered by small ships of less than 5 tons are associated with a tsunami height greater than 2 m (Shuto 1987; Aketa et al. 1994). The mooring rope for such small ships might not withstand a flow velocity higher than 2 m/s (Tsubota et al. 2007; Kato et al. 2009). The Japan Association of Marine Safety (2003) indicates that small marine vessels weighing less than 10 tons can lose their stability if the flow velocity is higher than 2 m/s. Moreover, in similar findings summarized by Lynett et al. (2014), using damage data of ports in California during the recent tsunamis, it was
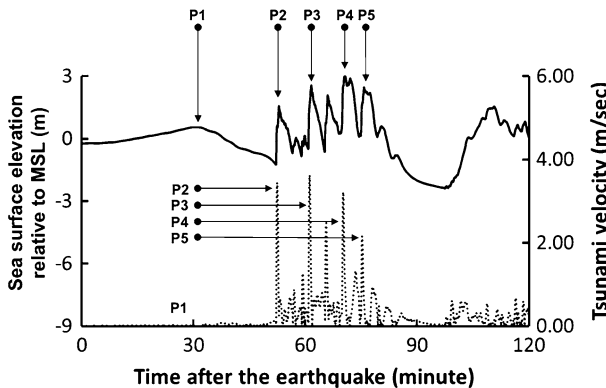

Fig. 8 Example of the extracted values of simulated tsunami heights (*solid line*) and tsunami velocities (*dashed line*) at point 7 in Oarai Port

Table 1 Minimum, mean, and maximum values of tsunami heights and velocities extracted from all points in the entirety of ports inside the study area

Value	Tsunami height					Tsunami velocity				
	Peak 1	Peak 2	Peak 3	Peak 4	Peak 5	Peak 1	Peak 2	Peak 3	Peak 4	Peak 5
Min	0.20	0.98	1.36	1.59	1.89	0.37	1.08	1.36	2.07	2.27
Mean	2.71	3.75	3.91	3.93	3.94	2.29	3.12	3.36	3.45	3.45
Max	5.80	8.20	8.20	8.20	8.20	7.55	7.55	7.55	7.55	7.55

concluded that flow velocity between 1.5 and 3 m/s can cause minor to moderate damage in ports and their facilities, including small boats. In this sense, it is highly possible for the ships in our study area to be swept away by the tsunami before it reaches its maximum height and/or velocity. These results indicate that having a complete sequence of tsunami parameters is necessary in order to use an appropriate value for statistical modeling, which will be described in detail in the next chapter.

4 Development of the loss estimation function of marine vessels

Loss estimation functions for marine vessels have been developed using ordinal regression analysis. Like other types of regression techniques, ordinal regression derives a relationship between one or several explanatory variables and a response variable, or the probability of obtaining this response. Here, the explanatory variables considered are the intensity measures (IMs) (i.e., the tsunami parameters: height and velocity, impact, ship tonnage, and ship material) in addition to the possible interactions between those IMs. The response variable is taken to be ship losses, discretized into bands of loss ratios ranging from 1 (less than 10 % loss) to 10 (more than 90 % loss). Ordinal regression consists in applying generalized linear models (Nelder and Wedderburn 1972; McCullagh and Nelder 1989) to scenarios where the response variable is categorical and ordered (Gelman and Hill 2007), as in the present case. It can been seen as a series of logistic regressions—one for each response category, the main difference being that the probability of being in a response category depends on the previous category (in simple logistic regression, a category's outcome is supposed independent of all other outcomes). Generalized linear models have been successfully applied to the estimation of tsunami damage probability in the context of building damage (Reese et al. 2011; Charvet et al. 2014a). Specifically, ordinal regression has successfully been applied to predict the probability of tsunami-induced building damage when buildings are classified according to a damage severity scale, following post-tsunami surveys (Charvet et al. 2014b; Leelawat et al. 2014).

In this study, the response consists of 10 discrete categories of loss, which make the multinomial distribution (Forbes et al. 2011) an adequate representation of the data:

$$Y_{k,i} \prod_{i=0}^{I} \frac{N_i!}{Y_{k,i}!} P(L_c = \text{loss}_k | x_i)^{Y_{k,i}} \tag{4}$$

In Eq. (4), $Y_{k,i}$ corresponds to the counts of boats being at a level of loss loss_k ($k \in \mathbb{N}$; $1 \le k \le 10$) for each value of tsunami intensity measures x_i; N_i is the total number of boats. The mean response for each category of loss, or mean probability for a boat to be at level loss_k, is μ_k and expressed as (McCullagh and Nelder 1989; Rossetto et al. 2014):

$$g(\mu_k) = \eta_k = \theta_{0,k} + \sum_{j=1}^{J} \theta_{j,k} X_j \tag{5}$$

In Eq. (5), X_j are the J explanatory variables to be used for the regression analysis, and $\{\theta_{0,k}, \ldots, \theta_{j,k}\}$ are the model parameters to be estimated through maximum-likelihood estimation (McCullagh and Nelder 1989; Myung 2003). Categorical explanatory variables such as collision or ship material are dummy coded. $g()$ is the link function linking the linear function of all explanatory variables to the mean response probability. Through

Eq. (5) complete ordering (only $\theta_{0,k}$ varies across categories, and $\theta_{j,k}$ becomes θ_j) or partial ordering of the outcomes may be considered. Although the loss categories are inherently ordered, the multinomial model (partial ordering) is more complex than the ordinal model (complete ordering), and thus it may provide a significantly better fit to the data. Therefore, both types of models should be considered and quantitatively compared using a likelihood ratio test, as shown in Rossetto et al. (2014) and Charvet et al. (2014a). The principle of the likelihood ratio test is to compare the deviances of two models to be evaluated against each other. The deviance being a measure of each model's error, where the model that has a statistically smaller deviance, should be chosen over the alternative. Further insights into the meaning and applications of deviance in generalized linear modeling are available elsewhere (i.e., Fahrmeir and Tutz 2001; Gelman and Hill 2007). The likelihood ratio test is also used to reveal which tsunami parameters have a significant influence on the response and should be kept (or not) and which interaction terms should be included in the model. The interaction term for each combination of IMs is a multiplicative function of those IMs. Finally, the best link function for the model is chosen based on the Akaike information criterion (AIC) (Akaike 1974).

The logic tree (Fig. 9) illustrates the steps taken in choosing the intensity measures (IM) that have a significant effect on the outcome to be predicted (loss category). The result of the likelihood ratio test is given in Table 2 by displaying the model's deviance (a measure of the model error) and the AIC for the relevant combinations of variables. The best-fitting model is the one that displays the significantly smallest deviance ($p < 0.05$) or smallest AIC value (for choosing the link function).
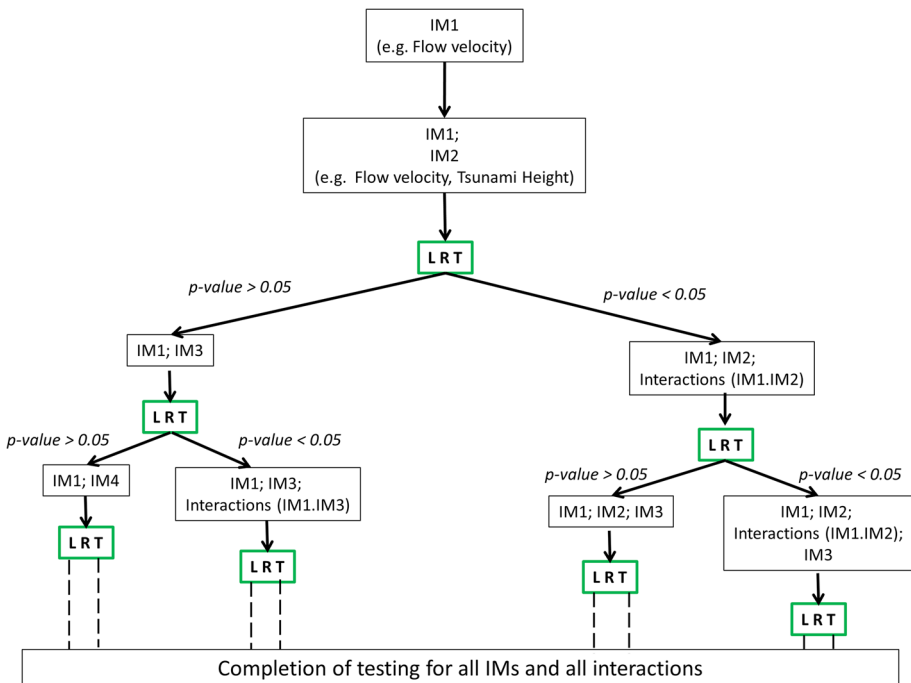

Fig. 9 Logic tree decision for performing the likelihood ratio test

Table 2 Deviances of each model tested through the likelihood ratio test

Explanatory variables and model type	Velocity M1	Height + velocity M2.1	Height + velocity + HV M2.2	Height + velocity + impact M3.1	Height + velocity + impact + HI M3.2	Height + velocity + impact + HI + VI M3.3	Height + velocity + impact + HI + Tonnage M4	Height + velocity + impact + HI + material M5.1	Height + velocity + impact + HI + material + HM M5.1	Height + velocity + impact + HI + material + HM + VM M5.2	Height + velocity + impact + HI + material + HM + IM M5.3
Ordinal	2 k = 20	2 k = 22	2 k = 24	2 k = 24	2 k = 26	2 k = 28	2 k = 26	2 k = 32	2 k = 38	2 k = 44	2 k = 44
Multinomial	2 k = 36	2 k = 54	2 k = 72	2 k = 72	2 k = 90	2 k = 108	2 k = 90	2 k = 144	2 k = 198	2 k = 252	2 k = 252
Ordered probit	2940	2775**	2774	2465**	2460** 2486	2460	2459	2454	2447	2445	2442
Multinomial probit	2930	2759	2747	2432	2428	2399* 2507	2425	2415	2392	2373	2372
Ordered logit	2942	2776**	2774	2465**	2460** 2486	2459	2459	2454	2445**	2443	2442
Multinomial logit	2931	2759	2748	2437	2458	2397* 2505	2452	2441	2420	2398	2398
Ordered comp log-log	2946	2784**	2779	2448**	2442**	2442	2443	2436** 2468	**2419**** 2457	2417	2414
Multinomial comp log-log	2932	2765	2760	2448	/	2403	/	2446	2390	2372	2370

The value of AIC is also included in cases where the performance of the model under different link functions needs to be compared. For each scenario, the number of statistical parameters k included in the model (either ordinal or multinomial) is provided

* The multinomial model is significantly better than the corresponding ordinal model, according to the likelihood ratio test ($p < 0.05$)

** The extra predictor significantly improves the fit compared to the simpler model (to the left), according to the likelihood ratio test ($p < 0.05$)

Table 3 Penalized CV accuracy results for the different points (i.e., geographical location in the ports) and different peak wave conditions (tsunami heights and velocities)

	Peak 1 (%)	Peak 2 (%)	Peak 3 (%)	Peak 4 (%)	Peak 5 (%)
Point 1	84.27	85.58	85.23	85.05	85.00
Point 2	83.10	85.29	85.19	85.30	85.33
Point 3	84.20	84.61	84.56	84.55	84.53
Point 4	84.67	84.55	84.41	84.33	84.36
Point 5	84.37	84.49	84.42	84.45	84.44

Based on these results, we chose the ordinal model, with tsunami height, velocity, impact, boat material, and the interaction terms height–material and height–impact, with a complementary loglog link function.

The overall model's performance is then assessed by calculating predictive accuracy through a repeated tenfold cross-validation (CV) for each point and for each tsunami wave conditions (peak height and associated peak velocity), following the CV and penalized accuracy calculation methodology proposed by Charvet et al. (2014c). Tenfold CV consists in randomly splitting the data into ten equal subsamples, using nine of these subsamples as a "training set" (the set to which the model is fitted), and the remaining subsample as a "test set" (the set which observed categories are being compared to the model's predicted categories). This process is repeated a number of times, typically until the CV results become stable. We started with point 1/peak 5 in each port, using the derivation of the results in Table 3. We followed the method proposed by Charvet et al. (2014c) to evaluate a penalized accuracy measure, defined as a measure of the distance between predicted and observed classes for each point and averaged over the total sample size.

To determine the number of repetitions for the cross-validation algorithm to yield a stable penalized accuracy results, a number of shuffles between 2 and 100 were tested. We found that $i = 26$ repetitions are necessary to obtain stable results. We therefore applied a repeated tenfold cross-validation with 26 repetitions to each point and wave condition in each port.

The penalized accuracy measure (Table 3) is indicative of the most likely values of tsunami parameters (peak height and velocity) at the location considered and at the time the vessels were struck by the tsunami.

Table 3 shows that the most accurate results are obtained in point 1/peak 2. However, we can also observe that the model yields consistently high-accuracy results, for all wave conditions, and with little variability. This result could be slightly surprising considering the uncertainty associated with the actual location of each boat at the time of damage occurrence, potentially quite different from the points of estimated tsunami heights and velocities. However, this result can be explained by the fact that, on a large scale, the wave conditions at one point span a large enough geographical area to include the location where the boat was damaged (so they are representative of a vessel's loss predictors), and the uncertainty is not as large as it may appear at first. In addition, because most of the data are contained in the most severe loss category (i.e., loss >90 %), the model may always yield consistent results in predicting the most common category.

By completing the steps described above, we finally derive loss estimation surfaces. For clarity of the graphical output, we choose to plot only the minimum and maximum loss data, corresponding to the probability of a ship suffering less than 10 % and more than

90 % loss (Fig. 10). The loss exceedence probability surfaces here are derived for only point 1/peak 2 of each port.

The points in "cyan" represent the expected loss exceedence probability for FRP boats, in "black" for aluminum boats, in "green" for steel boats, and in "red" for wooden boats. The smooth surfaces are constructed by approximating the discrete expected probabilities on a two-dimensional grid (D'Errico 2010).

The results in Fig. 10 show that collision is a major factor in the determination of vessel loss. Indeed, it appears that a ship will suffer at least minor losses (>10 %) for any wave conditions when a collision occurs, whereas when collision does not occur, the probability of such losses only becomes high for large tsunami heights and velocities. Similarly, the increase in loss probability is much steeper for large losses (>90 %) when collision occurs, compared to when it does not.

Here, it is important to note a major difference in the interpretation of (three-dimensional) loss estimation surfaces compared to two-dimensional loss estimation curves (Suppasri et al. 2013a; Muhari et al. 2013). A 3-D loss estimation surface provides a probability of loss for each combination of measured or simulated values of tsunami parameters in opposition to a 2-D loss estimation curve as provided by Suppasri et al. (2013a), which gives such probability only for one tsunami parameter at a time. In other words, loss estimation surfaces reveal that there is more than one unique value/combination of values for the explanatory variables that can yield the same or higher probability of loss, while the 2-D loss estimation curve is a straightforward for one value of the hazard
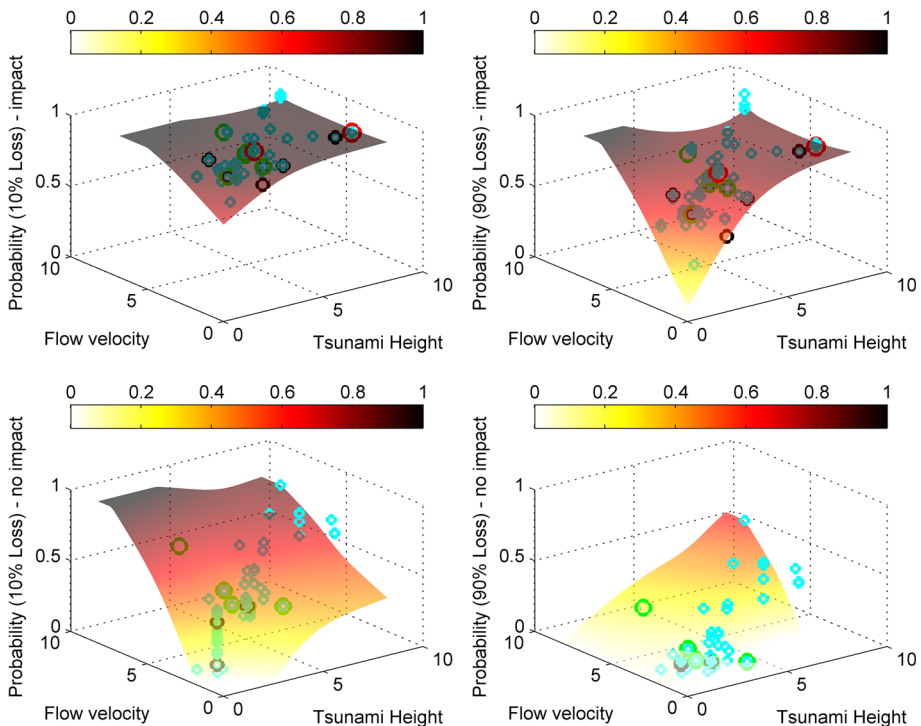

Fig. 10 Loss estimation surfaces for marine vessels subjected to tsunami height and velocity with collision aspects included (*upper figures*) and without collision parameters (*lower figures*)

parameter, or explanatory variable (read on the *x* axis), to retrieve a unique value of the damage or loss probability (on the *y* axis). Such factor explains discrepancies if the results of this study are compared with the previous results that use a similar data set (e.g., Suppasri et al. 2013a) in terms of the threshold of tsunami parameters that is related to the loss probability as follows; our results obtained a lower range of tsunami flow depth and velocity for loss probability larger than 50 %. For instance, a 75 % probability for 90 % loss (complete damage) in this study is obtained starting from 3.65 m tsunami height and 3.80 m/s tsunami velocity, while Suppasri et al. (2013a) obtained a 75 % probability for 90 % loss starting from 4.34 m tsunami height and 4.59 m/s tsunami velocity for ports located in the same area as in this study. Such discrepancies stressed the necessity to bring together all possible influential factors in the loss estimation function.

To make these results applicable in practice, the authors choose here to extract the minimum value of the tsunami height and velocity, for each band of loss, which would yield specific loss probabilities in the worst case scenario—when the ship suffers some form of impact (Table 4).

It can be said that, in general, marine vessels will start to have more than a 50 % probability of being damaged by the tsunami (>90 % loss) if they are struck by a tsunami with a height and velocity of more than 1.9 m and 2.1 m/s, respectively. The loss probability will significantly increase to 75 % if the ships are moored in the area where the tsunami height is more than 3.65 m and tsunami velocity is more than 3.8 m/s. From the overall data, the loss probability will reach up to 90 % if the tsunami height is more than 8 m and tsunami velocity is more than 5 m/s.

Our results in this study are consistent with the findings of damage data in the past (Shuto 1987; Aketa et al. 1994; Tsubota et al. 2007; Kato et al. 2009; Dengler and Uslu 2011). It is also consistent with the range of tsunami velocity associated with the developed damage index based on data in California (Lynett et al. 2014). They indicate that tsunami velocity ranging from 1.5 to 3 m/s can cause moderate damage on port facilities and small boats. Major damage in docks and boats are correlated with the range of 3–4.5 m/s

Table 4 Tabulation of boat losses and the associated tsunami parameters

Loss (%)	Probability of loss							
	0.25		0.50		0.75		0.90	
	Min. height	Min. velocity	Min. height	Min. velocity	Min. height (m)	Min. velocity (m/s)	Min. height (m)	Min. velocity (m/s)
>10	N/A	N/A	N/A	N/A	0.80	0.90	2.70	3.00
>20	N/A	N/A	N/A	N/A	1.60	1.80	3.80	3.60
>30	N/A	N/A	0.40 m	0.45 m/s	2.20	2.40	4.40	4.00
>40	N/A	N/A	0.75 m	0.80 m/s	2.50	2.80	4.80	4.20
>50	N/A	N/A	0.90 m	1.10 m/s	2.60	2.90	5.20	4.25
>60	N/A	N/A	1.20 m	1.20 m/s	2.80	3.10	5.60	4.40
>70	N/A	N/A	1.40 m	1.60 m/s	3.00	3.20	5.80	4.60
>80	0.40 m	0.50 m/s	1.70 m	1.95 m/s	3.40	3.50	6.80	4.75
>90	0.70 m	0.85 m/s	1.90 m	2.10 m/s	3.65	3.80	8.00	5.00

velocity, and the 4.5–6 m/s velocity range is related with extreme damage. Although their approach is different than that in our study, the range for moderate, major, and extreme impact in Lynet et al. (2014) might be proportional to 50 % loss probability (>2.1 m/s tsunami velocity), 75 % loss probability (>3.65 m/s tsunami velocity), and larger than 90 % loss probability (>5 m/s) in this study, respectively.

The limitation of the loss estimation function developed in this study is that the three-dimensional surfaces are based on the overall vessel data without any distinction based on a ship's tonnage. Although the sensitivity analysis using the likelihood ratio test indicates that there is no significant influence of the ship's tonnage, such a result might not indicate a universal condition but is instead more likely because our data are dominated by small marine vessels of less than 5 tons, and thus, the influence of higher tonnage cannot be adequately assessed. However, our results explain why the developed surfaces match well with the past historical data in Japan, which is mostly based on the observed damage of small marine vessels. On the other hand, this limitation indicates a necessity for future study to elaborate more on the different loss estimation surfaces between small, medium, and large marine vessels with an appropriate number of observed data for those three classes.

5 Final remarks

We have developed new loss estimation surfaces to estimate the damage probability of marine vessels associated with tsunami parameters, boat characteristics (material and tonnage), and the collision factor when they are subjected to tsunami conditions. Tsunami parameters in ports are simulated by using high-resolution topography data, which brought results consistent with the observed tsunami heights around the ports. We note that the simulated tsunami velocity in this study is not validated due to the lack of observational data of tsunami current along the southern part of Honshu Island. However, snapshots of tsunami propagation around the port demonstrated the capability of the model to represent the dynamics of tsunami current, including large vortex behavior during the tsunami run-up and run-down.

The statistical test for the simulated tsunami parameter associated with the observed damage data in this study reveals the necessity to have a complete sequence of the spatially distributed time series of the tsunami parameters for assessing tsunami impact on marine vessels. We show that the damage experienced by marine vessels might not be associated with the maximum values of tsunami conditions in the ports. Our simulated tsunami heights and velocities around the location of moored ships show lower values compared to the maximum tsunami height and velocity in the entire port, which is usually found along the breakwater and in the navigation channel at the opening of the harbor. Taking the maximum values of tsunami parameters, therefore, can yield an overestimated prediction to the order of square with the actual maximum simulated values at the location of moored ships.

The developed loss estimation functions indicate that marine vessels will start to have more than 50 % probability of being damaged by the tsunami (>90 % loss) if they are struck by a tsunami with a height and velocity of more than 1.9 m and 2.1 m/s, respectively. Next in the sequence, the loss probabilities of 75 and >90 % are associated with tsunami heights larger than 3.65 and 8 m, respectively, and tsunami velocities larger than 3.8 m and 5 m/s, respectively. The key capability of the developed loss estimation functions in this study is the integration of almost all factors that may influence the damage probability, starting from the initial state of the ship hit by the tsunami, during the time that

the vessels are being swayed by the tsunami, and ending with the condition when the ships are eventually found to have survived or to have been damaged by the tsunami. The results of this study, therefore, are crucially important for a better understanding of tsunami impacts on marine vessels and their future mitigation.

References

Admire AR, Dengler LA, Crawford GB, Uslu BU, Borrero JC, Greer SD, Wilson RI (2014) Observed and modeled currents from the Tohoku-Oki, Japan and other recent tsunamis in Northern California. Pure appl Geophys. doi:10.1007/s00024-014-0797-8

Aida I (1978) Reliability of a tsunami source model derived from fault parameters. J Phys Earth 26:57–73

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6):716–723

Aketa S, Yano K, Mizuno Y, Sato J, Terauchi K (1994) Reduction effect of port and fishing port facilities by tsunami damage. In: Proceedings of the coastal engineering conference (JSCE), vol 41, pp 1176–1180 (in Japanese)

Charvet I, Ioannou I, Rossetto T, Suppasri A, Imamura F (2014a) Empirical fragility assessment of buildings affected by the 2011 Great East Japan tsunami using improved statistical models. Nat Hazards. doi:10.1007/s11069-014-1118-3

Charvet I, Suppasri A, Imamura F (2014b) Empirical fragility analysis of building damage caused by the 2011 Great East Japan Tsunami in Ishinomaki City using ordinal regression, and influence of key geographical features. Stocha Environ Res Risk Assess. doi:10.1007/s00477-014-0850-2

Charvet I, Suppasri A, Kimura H, Sugawara D, Imamura F (2014c) Fragility estimations for Kesennuma City following the 2011 Great East Japan Tsunami based on maximum flow depths, velocities and debris impact, with evaluation of the ordinal model's predictive accuracy. IRIDeS internal report

D'Errico J (2010) Surface fitting using gridfit. MATLAB Central File Exchange

Dengler L, Uslu B (2011) Effects of harbor modification on Crescent City, California's tsunami vulnerability. Pure appl Geophys 168(6–7):1175–1185

Fahrmeir L, Tutz G (2001) Multivariate statistical modelling based on generalized linear models. Springer series in statistics, 2nd edn. Springer, New York

Forbes C, Evans M, Hastings N, Peacock B (2011) Statistical distributions. Wiley series in probability and statistics, 4th edn. Wiley, Hoboken, p 135

Fuji News Network-FNN (2012) Oarai Port just after the Tsunami, aerial shot (Day of the earthquake). http://www.fnn-news.com/en/311/articles/201103110047.html April, 2012

Gelman A, Hill J (2007) Data analysis using regression and multilevel/hierarchical models. Analytical methods for social research series. Cambridge University Press, Cambridge

Imamura F (1996) Review of tsunami simulation with a finite difference method. In: Yeh H, Liu P, Synolakis C (eds) Long wave run-up models. World Scientific, Singapore, pp 25–42

Imamura F (2009) Tsunami modeling: calculating inundation and hazard maps. In: Bernard EE, Robinson RA (eds) The sea, volume 15: Tsunamis. Harvard University Press, USA, pp 321–332

Intergovernmental Oceanographic Commission-IOC UNESCO (1997) Numerical method of tsunami simulation wit the leap-frog scheme. IUGG/IOC TIME project, manuals and guides, 35

Japan Association of Marine Safety (2003) Guideline for facility design of fishing port. Japan Association of Marine Safety, Tokyo (in Japanese)

Japan Nuclear Energy Safety-JNES (2011) Cross check analysis of the simulation of the 2011 Tohoku-Oki earthquake tsunami by the electric power company. http://www.nsr.go.jp/archive/nisa/shingikai/800/26/003/3-4.pdf. Retrieved April 2014

Japan Society of Civil Engineers-JSCE (2002) Tsunami assessment method for nuclear power plants in Japan, The Nuclear Evaluation Sub-Committee, 72 p

Kajitani Y, Chang SE, Tatano H (2013) Economic impact of the 2011 Tohoku-Oki earthquake and tsunami. Earthq Spectra 29(S1):S457–S478

Kato H, Miyake K, Saito M, Fujima K, Shigihara Y (2009) Proposal and application examples of a formula for calculating the tensile force acting on berthing ropes of small vessels such as fishing boats during tsunamis. Research report of the Japanese Institute of Fisheries Infrastructure and Communities, vol 21, pp 51–57 (in Japanese)

Kawata Y, Nina K, Harada K, Suzuki S (2004) Proposal for evaluation method of ship damage by tsunami. In: Proceedings of the coastal engineering conference (JSCE), vol 51, pp 316–320 (in Japanese)

Kazama T, Nakamura T, Ito T, Otsuka K, Sato K, Imazu Y (2006) A study on evacuation area for mitigating ship damage by tsunami. In: Proceedings of the coastal engineering conference (JSCE), vol 53, pp 1356–1360 (in Japanese)

Koshimura S, Oie T, Yanagisawa H, Imamura F (2009) Developing fragility function for tsunami damage estimation using numerical model and post-tsunami data from Banda Aceh, Indonesia. Coast Eng J 51:243–273

Leelawat N, Suppasri A, Charvet I, Imamura F (2014) Building damage from the 2011 Great East Japan tsunami: quantitative assessment of influential factors. Nat Hazards 73(2):449–471

Lynett P, Borrero J, Weiss R, Son S, Greer D, Renteria W (2012) Observations and modeling of tsunami-induced currents in ports and harbors. Earth Planet Sci Lett 327–328:68–74

Lynett PJ, Borrero J, Son S, Wilson R, Miller K (2014) Assessment of the tsunami-induced current hazard. Geophys Res Lett 41. doi:10.1002/2013GL058680

McCullagh P, Nelder JA (1989) Generalized linear models, 2nd. Chapman and Hall/CRC, Boca Raton

Ministry of Agriculture, Forestry and Fisheries-MMAF (2012) The Damage caused by the Great East Japan Earthquake. http://www.maff.go.jp/e/quake/press_since_120605.html. April 2014

Mori N, Takahashi T, Yasuda T, Yanagisawa H (2011) Survey of 2011 Tohoku earthquake tsunami inundation and run-up. Geophys Res Lett 38:L00G14. doi:10.1029/2011GL049210

Muhari A, Koshimura S, Imumura F (2011) Examination of three practical run-up models for assessing tsunami impact on highly populated areas. Nat Hazards Earth Syst Sci 11:3107–3123

Muhari A, Suppasri A, Murakami H, Futami T, Imamura F (2013) Measuring fragility of ships based on numerical model of the 2011 East Japan tsunami. In: Proceedings of international session. Coastal engineering conference (JSCE), Dec 2013, Kyushu, Japan

Myung IJ (2003) Tutorial on maximum likelihood estimation. J Math Psychol 47:90–100

Nelder JA, Wedderburn RW (1972) Generalized linear models. J R Stat Soc Ser A (General) 135(3):370–384

Okada Y (1985) Surface deformation due to shear and tensile faults in a half-space. Bull Seismol Soc Am 75:1135–1154

Reese S, Bradley BA, Bind J, Smart G, Power W, Sturman J (2011) Empirical building fragilities from observed damage in the 2009 South Pacific Tsunami. Earth Sci Rev 107:156–173

Rossetto T, Ioannou I, Grant DN, Maqsood T (2014) Guidelines for empirical vulnerability assessment, GEM Technical report 2014-X, GEM Foundation, Pavia, Italy

Shuto N (1987) Changing of tsunami disasters. Tsunami Engineering Technical report, Tohoku University 4:1–41 (in Japanese)

Sugino H, Wu C, Korenaga M, Nemoto M, Iwabuchi Y, Ebisawa K (2013) Analysis and verification of the 2011 Tohoku Earthquake Tsunami at nuclear power plant sites. J Jpn As Earthq Eng 13:2–21

Suppasri A, Muhari A, Futami T, Imamura F, Shuto N (2013a) Loss functions for small marine vessels based on survey data and numerical simulation of the 2011 Great East Japan Tsunami. J Waterw Port Coast Ocean Eng. doi:10.1061/(ASCE)WW.1943-5460.0000244

Suppasri A, Mas E, Charvet I, Gunasekera R, Imai K, Fukutani Y, Abe Y, Imamura F (2013b) Building damage characteristics based on surveyed data and fragility curves of the 2011 Great East Japan Tsunami. Nat Hazards 66(2):319–341

Titov VV, González FI (1997) Implementation and testing of the Method of Splitting Tsunami (MOST) model NOAA Technical Memorandum ERL PMEL-112, 11 p

Tsubota Y, Miyake K, Saito M (2007) Tension in a Mooring Rope of Fishing Boat under the Influence of Tsunami, Research report of the Japanese Institute of Fisheries Infrastructure and Communities, vol 19, pp 69–75 (in Japanese)

Uslu B, Admire AR, Dengler L (2013) Investigating effects of model resolution on computing tsunami currents in Northern California. International Tsunami Symposium (ITS), abstract book, 122 p