

Empirical fragility assessment of buildings affected by the 2011 Great East Japan tsunami using improved statistical models

I. Charvet · I. Ioannou · T. Rossetto · A. Suppasri · F. Imamura

Received: 20 November 2013 / Accepted: 28 February 2014 / Published online: 15 March 2014
© The Author(s) 2014. This article is published with open access at Springerlink.com

Abstract Tsunamis are destructive natural phenomena which cause extensive damage to the built environment, affecting the livelihoods and economy of the impacted nations. This has been demonstrated by the tragic events of the Indian Ocean tsunami in 2004, or the Great East Japan tsunami in 2011. Following such events, a few studies have attempted to assess the fragility of the existing building inventory by constructing empirical stochastic functions, which relate the damage to a measure of tsunami intensity. However, these studies typically fit a linear statistical model to the available damage data, which are aggregated in bins of similar levels of tsunami intensity. This procedure, however, cannot deal well with aggregated data, low and high damage probabilities, nor does it result in the most realistic representation of the tsunami-induced damage. Deviating from this trend, the present study adopts the more realistic generalised linear models which address the aforementioned disadvantages. The proposed models are fitted to the damage database, containing 178,448 buildings surveyed in the aftermath of the 2011 Japanese tsunami, provided by the Ministry of Land, Infrastructure Transport and Tourism in Japan. In line with the results obtained in previous studies, the fragility curves show that wooden buildings (the dominant construction type in Japan) are the least resistant against tsunami loading. The diagnostics show that taking into account both the building's construction type and the tsunami flow depth is crucial to the quality of the damage estimation and that these two variables do not act independently. In addition, the diagnostics reveal that tsunami flow depth estimates low levels of damage reasonably well; however, it is not the most representative measure of intensity of the tsunami for high damage states (especially structural damage). Further research using disaggregated damage data and additional explanatory variables is required in order to obtain reliable model estimations of building damage probability.

I. Charvet (✉) · I. Ioannou · T. Rossetto
Earthquake and People Interaction Center, University College London, London WC1E 6BT, UK
e-mail: i.charvet@ucl.ac.uk

A. Suppasri · F. Imamura
International Research Institute of Disaster Science, Tohoku University, 6-6-11-1104,
Aramaki, Aoba, Sendai, Japan

Keywords Tsunami damage · Fragility functions · Generalised linear models · Ordinal regression · Diagnostics

1 Introduction

The fragility of buildings affected by tsunamis is commonly expressed in terms of a stochastic relationship between the level of damage caused by the tsunami and a measure of its intensity. When empirical in nature, this relationship is expressed in terms of a parametric statistical model which is fitted to available post-tsunami damage survey (e.g. Peiris 2006) or satellite data (e.g. Koshimura et al. 2009). A statistical model consists of a random and a systematic component. The random component expresses the probability distribution of the response variable (e.g. the counts of buildings reaching or exceeding a damage state) given the explanatory variable (i.e. the measured tsunami flow depth). Flow depth alone is used as the explanatory variable by most existing fragility studies (e.g. Dias et al. 2009; Leone et al. 2011; Mas et al. 2012; Suppasri et al. 2012), as it drives the main tsunami forces acting on solid bodies (hydrostatic and hydrodynamic) and can be easily determined through field surveys—for example, by measuring inundation marks on the affected buildings. The systematic component of the statistical model expresses the mean response as a function of the explanatory variable. The systematic component typically represents the probability that a damage state is reached or exceeded given the tsunami intensity, and is called the fragility curve. Fragility curves are key components in seismic and tsunami loss estimation and have been constructed following the 2004 tsunami in Indonesia (Koshimura et al. 2009), in Sri Lanka (Peiris 2006; Murao and Nakazato 2010) and Thailand (Suppasri et al. 2011), as well as after the 2011 Great East Japan tsunami (Suppasri et al. 2012, 2013a, b).

The construction of fragility curves for buildings affected by tsunami is at an early stage, and the few studies published so far follow closely the procedures widely adopted for seismic fragility assessment of buildings. Thus, most fragility curves are expressed in terms of a lognormal, and in few cases normal, cumulative distribution for which parameters are estimated by fitting a linear statistical model to the aggregated damage data. One main limitation of this model is its inability to deal with aggregated data points with probabilities of 0 or 1: such points represent bins where all the buildings suffered less than the expected damage or suffered the examined level of damage or more, respectively. This issue is addressed in some studies by re-grouping buildings into bins with approximately equal number of buildings, resulting in some bins corresponding to a rather wide range of intensity measure levels. For instance, Suppasri et al. (2011) formed even bins of 50 buildings (Phuket) and 100 buildings (Phang Nga); Suppasri et al. (2012) formed bins of 15 buildings in the study areas of Sendai and Ishinomaki. However, this re-grouping can affect the shape of the fragility curve especially at the tails. The issue of the model with extreme probabilities leads to the systematic dismissal of points (bins) yielding such probabilities of 0 or 1, which will affect not only the power of the analysis (Green 1991), but will also lead to biased estimations of the distribution's parameters (mean and standard deviation). Other shortcomings include the non-verification (and often violation) of the linear regression assumptions and the non-representation of the discrete nature of the response variable (i.e. the damage level), as shown by Charvet et al. (2013). These issues are non-trivial, as ultimately the usefulness of fragility

functions lies in their potential to be used for prediction of future damage; therefore, an adequate estimation of the population's distribution is essential. An improved model which does not require the re-grouping of the available data and takes into account the discrete nature of the response variable is the generalised linear model adopted by Reese et al. (2011). The authors applied logistic regression to the damaged buildings, so the occurrence of each damage state can be represented as a binary outcome. This also means that the fragility curve is constructed individually for each damage state, which ignores the fact that damage is an ordinal categorical variable, i.e. typically expressed by discrete damage states which increase in intensity from no damage to collapse of the structure.

Deviating from current practice, this study aims to examine the relationship of tsunami damage to buildings with flow depth using more realistic statistical models, which account for the fact that the damage is a discrete and ordered outcome. These models are then fitted to the field damage data of buildings affected by the 2011 Japanese Tsunami. The case study emphasises the need to assess and document the goodness-of-fit of the selected models, which is not adequately addressed in the literature. In particular, the assumptions of the selected statistical models are validated, and the confidence intervals around the developed fragility curves are also provided. The proposed fragility assessment aims to:

1. Produce fragility curves, which can be used for risk assessment.
2. Examine the probability of collapsed buildings to be washed away for increasing tsunami intensity levels. This could be used for future estimation of the debris generated by washed away structures, given that these structures can act as projectiles against other structures and contribute to further damage.
3. Assess the significance of buildings classification according to their structural characteristics in the assessment of fragility, given that a few studies produced fragility curves for all available buildings irrespective of their building class (Peiris 2006; Koshimura et al. 2009; Gokon et al. 2011; Reese et al. 2011).
4. Assess the ability of the flow depth to explain the damage suffered by the buildings affected by a tsunami.

2 The MLIT tsunami damage database

The fragility curves constructed here are based on the Ministry of Land, Infrastructure, Transport and Tourism (2012) damage database. This database contains aggregated information for 178,448 buildings located in the prefectures bordering the East coast of Japan, affected by the 2011 tsunami (Fig. 1a, c). The database includes counts of buildings in each of seven discrete levels of damage, for 13 levels of tsunami intensity (flow depth, measured in metres). Each level of tsunami intensity is taken as the median of each bin (0.5 m wide), and the consequent damage is expressed in terms of the seven-state quantitative damage scale depicted in Table 1. The buildings are classified by construction material and include mostly wooden structures followed by reinforced concrete (RC) and steel as shown in Fig. 1b.

The quality of the database is assessed by examining the size of non-sampling errors and in particular the size of the non-coverage and partial non-response error according to the guidelines for the construction of empirical fragility curves proposed by Rossetto et al. (2014). The former error is considered negligible as the database includes both the

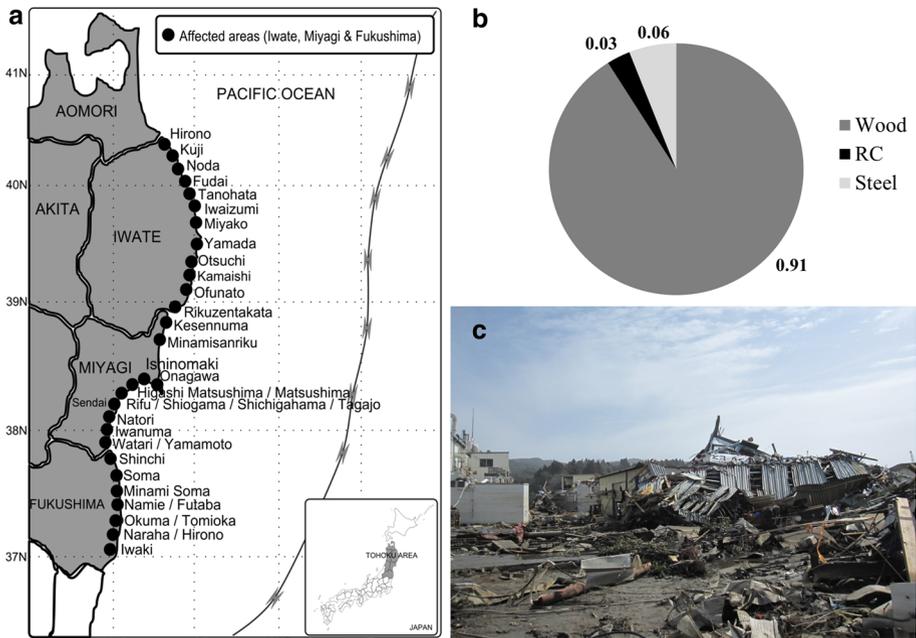


Fig. 1 a Map of the prefectures and towns affected by the 2011 Great East Japan tsunami, where the data was collected; b distribution of buildings’ construction materials (building class) in the database; c example of a damaged area in Miyagi prefecture (Kesenuma, 20th March 2011)

Table 1 Description of each damage state as provided by MLIT and used by the survey teams for building damage classification

MLIT DS	DS	Classification of damage	Description of observed damage	Condition of the structure
7	ds ₀	No damage	No damage	Promptly reusable without repair or floor cleaning
6	ds ₁	Minor	Flood only	Promptly reusable after floor cleaning
5	ds ₂	Moderate	No structural damage	Reusable after minor repair
4	ds ₃	Major	Partly damaged walls, but no damage in columns	May be reusable after great repair
3	ds ₄	Complete	Extensive damage in walls, some damage in columns.	May be reusable after great repair only in some cases
2	ds ₅	Collapse	More than half of walls were destroyed, most columns were bent or destroyed	Non-repairable
1		Washed away	Washed away, only foundation remained, overturned	Non-repairable

MLIT DS represents the original damage scale, and DS represents the damage scale used in this study

damaged and undamaged buildings in the surveyed areas. The latter error is also considered negligible as the buildings for which the damage state is unknown consist of less than 5 % of the total number of buildings in the database.

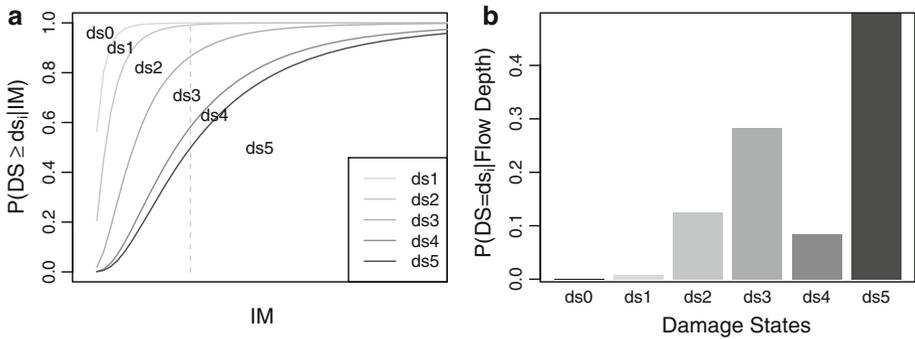


Fig. 2 **a** General example of fragility curves for the case of five damage states; **b** probability of a given damage state at a chosen IM estimated from the difference between two successive cumulative probability estimations

This extensive database can be used for the empirical fragility assessment of the three buildings classes. The objectives of fragility assessment vary from the investigation of the increase in damage with the increase in tsunami intensity as well as its use to estimate the economic loss or casualties for future events. The latter, however, requires that damage is expressed in terms of an ordinal categorical variable, where damage states represent all possible outcomes. This means that the states should be mutually exclusive and collectively exhaustive, and they should increase in intensity (McCullagh and Nelder 1989), i.e. from no damage to collapse. In this way, the damage probability matrices, necessary for the estimation of loss for a given level of intensity, can be evaluated in terms of the difference between two successive fragility curves as presented in Fig. 2.

However, Table 1 demonstrates that damage states of “collapse” and “washed away” essentially represent different collapse modes, violating the above requirements. This limitation is addressed here by aggregating the counts of buildings of these two states for each level of tsunami intensity. This allows for the transformation of the seven-state damage scale (MLIT DS in Table 1) into a six-state scale (DS), in accordance with the aforementioned recommendations.

We also note two additional shortcomings of the damage scale as defined and used by MLIT for data collection, which unfortunately cannot be addressed in the present analysis by modifying the damage scale. Firstly, the failure mode “overturned” has been aggregated with the failure mode “washed away”. This will increase the uncertainty in the estimated probability associated with this particular failure mode. Secondly, the definition of ds_4 , even though it does not technically contradict the requirements highlighted above, is relatively subjective or even unclear and may lead to a lot of ds_4 building misclassification in the field. For example, both ds_4 and ds_5 are described by significant damage in walls and columns, in some cases both kinds of building damage will make the structure impossible to repair. A closer examination of the database reveals that out of the totality of buildings surveyed by MLIT, only 4 % correspond to a damage level of ds_4 , while the preceding and following levels contain a higher proportion of data (16 % for ds_3 and 13 % for ds_5), confirming the damage definition needs to be improved.

3 Fragility assessment methodology

The framework of the empirical fragility assessment procedure proposed here consists of three steps. Firstly, the appropriate statistical models are selected. Then they are fitted to

the available data, and finally their goodness-of-fit is assessed. The procedure is based on two, commonly found in the literature, assumptions that field response data are of high quality and the measurement error of the explanatory variable is negligible.

The procedure consists of two independent stages, developed in order to address the first two aims of this study. In the first stage, the fragility curves corresponding to the five levels of damage highlighted in Table 1 (DS), which can be used for risk assessment, are constructed in terms of the probability of a damage state being reached or exceeded by a building of a given class for different levels of tsunami intensity. The five curves are constructed by fitting a generalised linear model (GLM), originally developed as an extension to linear statistical models (McCullagh and Nelder 1989). This model is considered a more realistic representation of the relationship between a discrete ordinal (i.e. the counts of buildings being in the six damage states) and continuous (i.e. flow depth) or discrete (i.e. building class) explanatory variables. Indeed, such models recognise that the fragility curves are ordered, and therefore, they should not cross. They also recognise that the fragility curves are bounded between 0 and 1, and they take into account that some points have a larger overall number of buildings than others.

In the second stage, the probability that a collapsed building will be washed away, conditioned on tsunami intensity, is assessed. This conditional probability is also estimated by fitting the GLM to the collapsed data and their corresponding intensity measure levels. For this stage, the GLM is considered the most appropriate model to relate a discrete response variable (i.e. counts of collapsed buildings being washed away) with a continuous (i.e. flow depth) or discrete (i.e. building class) explanatory variable.

The importance of accounting for the structural characteristic of buildings in the fragility assessment is explored by the construction of generalised linear models of increased complexity in line with the recommendations of Rossetto et al. (2014). Initially, a simple parametric statistical model is introduced in order to relate the damage sustained by buildings irrespective of their class to a measure of tsunami intensity. The complexity of this model is then increased by adding a second explanatory variable expressing the structural characteristics of each building class, and the potential superiority of the most complex model is assessed using appropriate statistical measures.

Finally, the ability of the flow depth to predict the damage data, taking into account the building class, is investigated by assessing whether the constructed models adequately fit the available damage data through graphical model diagnostics.

3.1 Selection of generalised linear models

The GLMs methodology implies firstly, the selection of an appropriate statistical distribution for the response variable (i.e. damage state), which will be referred to as the *random component* of the model; secondly, the expression of the damage probability exceedance (or fragility function μ) through a chosen linear predictor η and link function g (this will be referred to as the *systematic component* of the model). When using GLM, the linear predictor is an additive function of all explanatory variables included in the model.

3.1.1 Overview of model components

In the first stage of the analysis, the damage response consists in six discrete damage levels: each building in the sample may be described by 1 out of 6 possible outcomes, which makes the *multinomial distribution* (Forbes et al. 2011) an adequate random component for this stage. Through the expression of the systematic component, complete

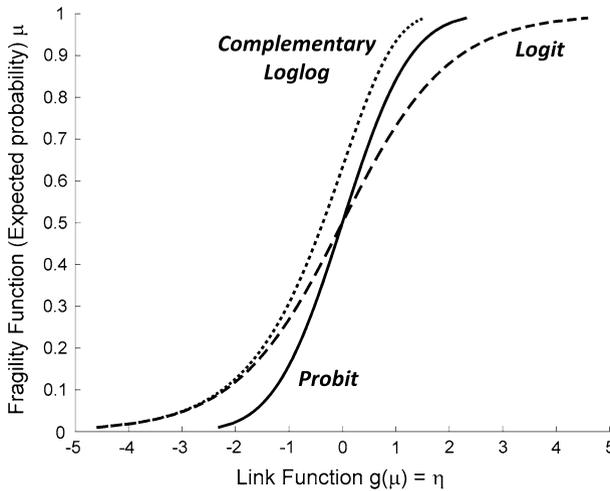


Fig. 3 Example link functions and their relationship with the damage exceedance probability (i.e. fragility function) and the linear predictor (a linear function of the explanatory variables included in the model). Probit (solid line): $\Phi^{-1}(\mu)$; logit (dashed line): $\log\left(\frac{\mu}{1-\mu}\right)$; complementary Log–log (dotted line): $\log(-\log(1 - \mu))$

ordering (ordered model) or partial ordering of these outcomes may be considered. Although the damage scale presented in Table 1 presents fully ordered outcomes (i.e. damage states of strictly increasing severity), the partially ordered model is more complex; thus, it may provide a significantly better fit to the data. Therefore, both types of models should be considered and quantitatively compared. In the second stage of the analysis, the damage response consists in two discrete damage levels: each building which has collapsed may have been washed away or not. This makes the *binomial distribution* the most appropriate random component for the second stage.

The relationship between the estimated fragility curve, the linear predictor and the link function is illustrated in Fig. 3 for the three link functions applicable to multinomial and binomial distributions, namely the probit, logit and complementary log–log functions. Note that the complementary log–log link shall not be considered in this study, unless the model appears to systematically underestimate observations for high damage probabilities. Indeed, this link provides estimations close to the logit link except for a heavier right tail and cannot be differentiated when the linear predictor takes the value of zero.

For each stage of the analysis, the chosen random components and possible expressions for the systematic components will be introduced. Candidate models will be evaluated through measures of relative goodness-of-fit in order to select the best fitting model.

3.1.2 First stage

The construction of fragility curves in the first stage of analysis is based on the selection of an ordered and partially ordered probit model of increasing complexity. For this stage, the response variable is expressed in terms of the counts of buildings suffering a given damage state ($DS = ds_i$ for $i = 0, \dots, n$). The explanatory variables are expressed in terms of a tsunami intensity measure and the building class.

The random component expresses the probability of a particular combination of counts y_{ij} of buildings, irrespective of their class, suffering damage $ds_{i=0,\dots, n}$ for a specified tsunami intensity level, im_j , which is considered to follow a multinomial distribution:

$$y_{ij} \sim \prod_{i=0}^n \frac{m_j!}{y_{ij}!} P(DS = ds_i | im_j)^{y_{ij}} \tag{1}$$

Equation (1) shows that the multinomial distribution is fully defined by the determination of the conditional probability, $P(DS = ds_i | im_j)$, with m_j being the total number of buildings for each value of im_j . This probability can be transformed into the probability of reaching or exceeding ds_i given im_j , $P(DS \geq ds_i | im_j)$, essentially expressing the required fragility curve, as:

$$P(DS = ds_i | im_j) = \begin{cases} 1 - P(DS \geq ds_i | im_j) & i = 0 \\ P(DS \geq ds_i | im_j) - P(DS \geq ds_{i+1} | im_j) & 0 < i < n \\ P(DS \geq ds_i | im_j) & i = n \end{cases} \tag{2}$$

The fragility curves are expressed in terms of an ordered probit link function using the logarithm of the tsunami intensity measure (which is essentially the lognormal distribution, adopted by the majority of the published fragility assessment studies), forming a model which will be referred to as M.1.1.1, in the form:

$$\Phi^{-1} [P(DS \geq ds_i | im_j)] = \theta_{0,i} + \theta_1 \log(im_j) = \eta \tag{3}$$

where $\theta = [\theta_{0,i}, \theta_1]$ is the vector of the “true”, but unknown parameters of the model; θ_1 is the slope, and $\theta_{0,i}$ is the intercept for fragility curve corresponding to ds_i ; η is the linear predictor; $\Phi^{-1}[\cdot]$ is the inverse cumulative standard normal distribution, expressing the probit link function. In order to build an ordered model, the systematic component, expressed by Eq. (3), assumes that the fragility curves corresponding to different damage states have the same slope, but different intercepts. This ensures that the ordinal nature of the damage is taken into account leading to meaningful fragility curves, i.e. curves that do not cross.

The assumption that the fragility curves have the same slope for all damage states is validated by comparing the fit of the ordered probit model with the fit of a partially ordered probit model, which relaxes this assumption. It should be noted that the latter model considers that the damage is a nominal variable and disregards its ordinal nature by allowing the slope of the fragility curves for each damage state to vary independently. This model (M.1.1.2) is constructed by transforming the linear predictor, η , in the form:

$$\Phi^{-1} [P(DS \geq ds_i | im_j)] = \theta_{0,i} + \theta_{1,i} \log(im_j) = \eta \tag{4}$$

The complexity of the aforementioned systematic components [i.e. eqs. (3) and (4)] is then increased by adding a new explanatory variable which accounts for building class. Thus, a nested model (M.1.2) is obtained, which systematic component has the form:

$$\Phi^{-1} [P(DS \geq ds_i | im_j, \text{class})] = \theta_{0,i} + \theta_1 \log(im_j) + (\theta_2, \theta_3) \text{class} \tag{5.1}$$

$$\Phi^{-1} [P(DS \geq ds_i | im_j, \text{class})] = \theta_{0,i} + \theta_{1,i} \log(im_j) + (\theta_{2,i}, \theta_{3,i}) \text{class} \tag{5.2}$$

In Eqs. (5.1) and (5.2), (θ_2, θ_3) and $(\theta_{2,i}, \theta_{3,i})$, respectively, are parameters of the model corresponding to the $K - 1$ categories of class, which is a discrete explanatory variable representing the various building classes. Let us note that in this case (inclusion of a

categorical explanatory variable), class is *dummy coded* (i.e. the $K - 1$ values of the variable class become a binary variable, to each combination of 0–1 corresponds a different building class). The model whose systematic component is expressed by Eq. (5.1) assumes that the slope of the fragility curves is not influenced by the building class. By contrast, the intercept changes according to the building class.

The model further expands in order to account for the interaction between the two explanatory variables (the resulting model is termed M.1.3). Accounting for the interaction means that the slope changes according to the building class. This means that the rate of increase in the probability of exceedance of a given damage state, as flow depth increases, depends on the construction type of the building (i.e. the two variables do not act independently). Therefore, the systematic component is rewritten in the form:

$$\Phi^{-1} [P(DS \geq ds_i | im_j, \text{class})] = \theta_{0,i} + \theta_1 \log(im_j) + (\theta_2, \theta_3)\text{class} + \theta_{12} \log(im_j)\text{class}_1 + \theta_{13} \log(im_j)\text{class}_2 \tag{6.1}$$

$$\Phi^{-1} [P(DS \geq ds_i | im_j, \text{class})] = \theta_{0,i} + \theta_{1,i} \log(im_j) + (\theta_{2,i}, \theta_{3,i})\text{class} + \theta_{12,i} \log(im_j)\text{class}_1 + \theta_{13,i} \log(im_j)\text{class}_2 \tag{6.2}$$

3.1.3 Second stage

The second stage focuses on buildings which collapsed under tsunami actions (i.e. damage state ds_5 as presented in Table 1). Based on the information found in the database, the failure mode of the collapsed buildings is a binary variable (i.e. they can either be washed away or not). This is reflected by the selection of the random component of the generalised linear models developed below. For this stage, the response variable is expressed in terms of the counts of collapsed buildings which are washed away and the explanatory variables are the same as in the first stage.

A simple statistical model is introduced first which relates the response variable with a measure of tsunami intensity, ignoring the contribution of the building class. Thus, for a given level of intensity im_j , the counts x_j of washed away collapsed buildings are assumed to follow a discrete binomial distribution, as:

$$x_j \sim \binom{m_j}{x_j} \mu_j^{x_j} [1 - \mu_j]^{m_j - x_j} \tag{7}$$

In Eq. (7), m_j is the total number of buildings, and μ_j is the mean of the binomial distribution expressing the probability of a collapsed building being washed away given im_j . The mean, μ_j , is essentially the systematic component of the selected statistical model which is formed here in line with the linear predictors developed for the first stage as:

$$g(\mu_j) = \eta = \begin{cases} \theta_4 + \theta_5 \log(im_j). & 1 \\ \theta_4 + \theta_5 \log(im_j) + (\theta_6, \theta_7)\text{class}. & 2 \\ \theta_4 + \theta_5 \log(im_j) + (\theta_6, \theta_7)\text{class} + \theta_{56} \log(im_j)\text{class}_1 + \theta_{57} \log(im_j)\text{class}_2. & 3 \end{cases} \tag{8}$$

where $\theta = [\theta_4, \theta_5, \theta_6, \theta_7, \theta_{56}, \theta_{57}]$ is the vector of the “true”, but unknown parameters of the models; $g(\cdot)$ is the link function, which can have the form of a probit as well as logit function:

$$\begin{aligned}
 g(\mu_j) &= g(P(\text{DS} = \text{washed away} | \text{collapse}, im_j, \text{class})) \\
 &= \begin{cases} \Phi^{-1}(P(\text{DS} = \text{washed away} | \text{collapse}, im_j, \text{class})) & .1 \text{ (probit)} \\ \log\left(\frac{P(\text{DS} = \text{washed away} | \text{collapse}, im_j, \text{class})}{1 - P(\text{DS} = \text{washed away} | \text{collapse}, im_j, \text{class})}\right) & .2 \text{ (logit)} \end{cases} \quad (9)
 \end{aligned}$$

3.2 Statistical model fitting technique

Having formed the aforementioned generalised linear models, they are then fitted to the field data. This involves the estimation of their unknown parameters by maximising the log-likelihood functions, as:

$$\begin{aligned}
 \theta^{\text{opt}} &= \arg \max[\log(L(\theta))] \\
 &= \begin{cases} \arg \max \left[\log \left(\prod_{j=1}^M \prod_{i=0}^n \frac{m_j!}{y_{ij}!} P(\text{DS} = ds_i | im_j, \text{class})^{y_{ij}} \right) \right] & \text{1st stage} \\ \arg \max \left[\log \left(\prod_{j=1}^M \binom{m_j}{x_j} \mu_j^{n_j} [1 - \mu_j]^{m_j - x_j} \right) \right] & \text{2nd stage} \end{cases} \quad (10)
 \end{aligned}$$

Apart from the mean estimates of the unknown parameters, θ , the uncertainty around these values is also estimated in this study. Given that the data are grouped and that the water depth is the only explanatory variable, the uncertainty in the data is likely to be larger than the uncertainty estimated by the adopted generalised linear model. This over-dispersion is taken into account in the construction of the confidence intervals by the use of bootstrap analysis (Rossetto et al. 2014). According to this numerical analysis, 1,000 random samples with replacement are generated from the available data points. For each realisation a generalised linear model is fitted. For specified levels of tsunami intensity, the 1,000 values of the fragility curves corresponding to a given damage state are ordered and the values with 90 and 5 % probability of exceedance are identified.

3.3 Goodness-of-fit assessment

The proposed procedure is based on developing a number of realistic statistical models, which are then fitted to the available data. Which one provides the best fit? To answer this question, the relative as well as the absolute goodness-of-fit of the proposed models is assessed. The relative goodness-of-fit assessment aims to identify the model which provides the best fit compared with the available alternatives. The absolute goodness-of-fit aims to explore whether the assumptions on which the model is based upon are satisfied, thus providing a reliable model for the given data.

3.3.1 Relative goodness-of-fit

The relative goodness-of-fit procedure is based on whether the two candidate models are nested (i.e. the more complex model includes at least all the parameters of its simpler counterpart) or not.

The goodness-of-fit of two nested models is assessed using the likelihood ratio test. Generally the more complex model fits the data better given that it has more parameters. This raises the question as to whether the difference between the two models is statistically significant. The likelihood ratio test is used in order to assess whether there is sufficient

evidence to support the hypothesis that the most complex model fits the data better than its simpler alternative. According to this test, the difference, D , in the deviances $[-2\log(L)]$ is assumed to follow a Chi square distribution with degrees of freedom $df = df_{\text{simple model}} - df_{\text{complex model}}$:

$$p = P\left(D \geq d = -2 \frac{\log(L_{\text{simple model}}) - \log(L_{\text{complex model}})}{df_{\text{simple model}} - df_{\text{complex model}}}\right) \tag{11}$$

In Eq. (11), L denotes the likelihood function, and p the p value of the statistical test. The difference can be noted by random chance and therefore is not significant if we obtain a p value >0.05 . This means that there is not sufficient evidence to reject the hypothesis; hence, the simpler model can be considered a better fit than its nested alternative. By contrast, the difference is significant, and therefore, the hypothesis can be rejected if p value <0.05 .

The comparison of the fit of non-nested models (here, two models with identical linear predictors but different link functions) is assessed by the use of the Akaike information criterion (AIC) following the recommendations of Rossetto et al. (2014). This criterion is estimated as:

$$\text{AIC} = 2k - 2\ln(L) \tag{12}$$

where k is the number of parameters in the statistical model. The model with the smallest AIC value is considered to provide a relatively better fit to the available data.

3.3.2 Absolute goodness-of-fit

The absolute goodness-of-fit of the best model is assessed by informal graphical tools. In particular, the goodness-of-fit of the best candidate models from each stage is assessed by plotting the observed counts of buildings for which $DS \geq ds_i$ with their expected counterparts. The closer the points are to the 45 degree line, the better the model.

4 Results and discussion

The proposed procedure is applied to the MLIT tsunami database in order to estimate the optimum parameters of the proposed generalised linear models and thus empirically assess the fragility of wooden, steel and reinforced concrete buildings (first stage) as well as the probability that the building’s mode of failure is to be washed away (second stage), for various tsunami intensity levels. The selected models which are then fitted to the data are summarized in Table 2. The case study emphasises the diagnostics which assess the goodness-of-fit of the proposed models, and strategies to improve the model following the recommendations presented in Rossetto et al. (2014).

4.1 Fragility assessment—first stage

For the first stage, we use the available damage data for the three buildings classes, classified in six damage states ranging from no damage (ds_0) to collapse (ds_5) according to the scale presented in Table 1 and aggregated in 13 ranges of flow depth. The number of buildings in each range (across all damage states) varies from 1,166 to 25,929.

Table 2 Nomenclature used for the models, their systematic component (equation number), description and link function used

Model	Variables					Components		
	Response			Explanatory		Random	Systematic	
	Ordinal	Nominal	Binary	Flow depth	Class		Linear predictor	Link function
<i>1st stage</i>								
M.1.1.1	✓	–	–	✓	–	Eq. (1)	Eq. (3)	Eq. (3)
M.1.1.2	–	✓	–	✓	–	Eq. (1)	Eq. (4)	Eq. (4)
M.1.2.1	✓	–	–	✓	✓	Eq. (1)	Eq. (5.1)	Eq. (5.1)
M.1.2.2	–	✓	–	✓	✓	Eq. (1)	Eq. (5.2)	Eq. (5.2)
M.1.3.1	✓	–	–	✓	✓	Eq. (1)	Eq. (6.1)	Eq. (6.1)
M.1.3.2	–	✓	–	✓	✓	Eq. (1)	Eq. (6.2)	Eq. (6.2)
<i>2nd stage</i>								
M.2.1	–	–	✓	✓	–	Eq. (7)	Eq. (8.1)	Eq. (9.1)
M.2.2	–	–	✓	✓	✓	Eq. (7)	Eq. (8.2)	Eq. (9.1)
M.2.3.1	–	–	✓	✓	✓	Eq. (7)	Eq. (8.3)	Eq. (9.1)
M.2.3.2						Eq. (7)	Eq. (8.3)	Eq. (9.2)

Table 3 Log-likelihood, *dfs* and *p* values for the ordered and probit model

Model	$-2\log(L)$	<i>df</i>	<i>p</i> value
M.1.1.1	358,206	178,442	~0.00
M.1.1.2	348,998	178,438	

The fragility of buildings irrespective of their class is assessed first by fitting the ordered probit model, M.1.1.1, to the aggregated damage data. One important and often ignored question is whether it is an acceptable model for the available damage data. This question is addressed by comparing its fit with the fit of the partially ordered model M.1.1.2 which relaxes the assumption of common slope for all fragility curves; thus, the latter model has four additional parameters. The likelihood ratio test is adopted in order to assess whether the improved fit provided by M.1.1.2 is significantly better than the M.1.1.1 fit. The test results in *p* values ~0.0, indicating that the partially ordered probit model M.1.1.2 can be considered a significant improvement over its simpler alternative (Table 3).

The fragility curves corresponding to the five damage states obtained by fitting M.1.1.2 are plotted in Fig. 4, along with the corresponding 90 % confidence intervals. It should be noted, however, that despite the very large number of buildings used in the construction of these curves, the confidence intervals for damage states 4 and 5 appear to be wide and overlap indicating that, when the building class is not taken into account, over-dispersion is potentially an issue.

The likelihood ratio test, however, does not examine whether model M.1.1.2 fits the available data satisfactorily. Thus, the absolute goodness-of-fit is assessed here by the use of graphical diagnostic plots as described in Sect. 3.3.2, where the expected counts of buildings, which reached or exceeded a damage state ($DS \geq ds_i$ for $i = \{1; 2; 3; 4; 5\}$) are plotted against their observed counterparts as presented in Fig. 5. The plots show that the points appear to be significantly away from the 45 degree line for higher damage states (i.e.

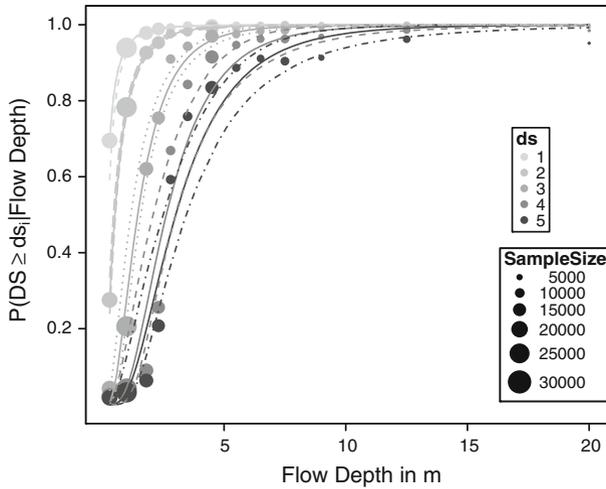


Fig. 4 Fragility curves for buildings irrespective of their building class and their 90 % confidence intervals—model M1.1.2

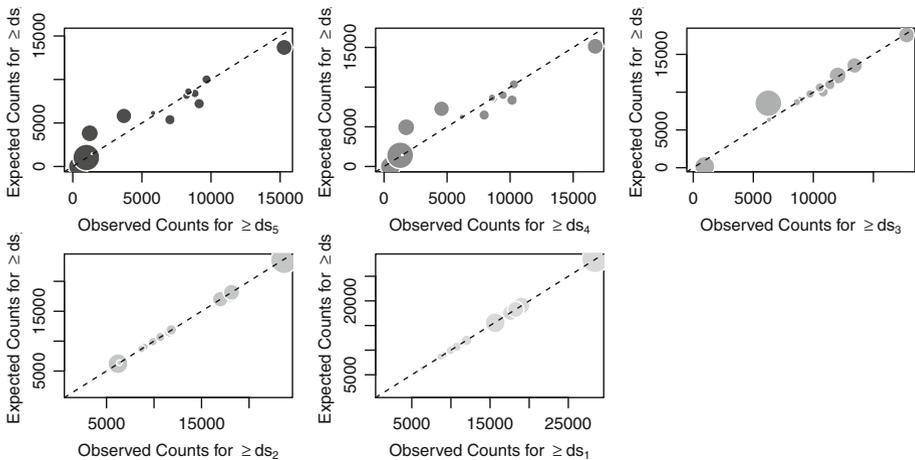


Fig. 5 Graphical diagnostics assessing the goodness-of-fit of the five fragility curves derived using M1.1.2 (probit)

DS4 to DS5), indicating the poor fit of the corresponding fragility curves. More specifically, it can be observed that the model tends to overestimate the number of damaged buildings. For lower damage states (i.e. DS1 and DS2), the model estimates the data satisfactorily.

Two alternative improved models (M.1.2.2: partially ordered model, and M.1.3.2: partially ordered model with interactions) are then adopted by adding the explanatory variable class, which accounts for the three building classes (i.e. wood, steel and reinforced concrete). Let us note here that because the ordered model M.1.1.1 has just been shown to provide a worse fit to the data at hand in comparison with the partially ordered model M.1.1.2, only the partially ordered versions of the model (Table 2) will now be considered.

Table 4 Log-likelihood, degrees of freedom and p values for the three probit models

Model	$-2 \times \log(L)$	df	p value
M.1.1.2	348,998	178,438	~ 0.00
M.1.2.2	336,368	178,428	~ 0.00
M.1.3.2	335,586	178,418	

The model which fits the damage data best is identified by the use of two likelihood ratio tests which compare the three models. The p values ~ 0.0 , shown in Table 4, indicate that M.1.3.2 provides a better fit to the data than the increasingly simpler models, M.1.2.2 and M.1.1.2.

Figure 6 depicts the fragility curves corresponding to the five damage states obtained by fitting model M.1.3.2 to the damage data. The curves for the wooden buildings appear to be shifted to the left followed by the curves for the steel buildings. This suggests that the wooden buildings are affected the most by the flow depth followed by the steel buildings, and the reinforced concrete buildings perform the best out the three classes. These results are consistent with findings from previous studies evaluating tsunami damage (e.g. Rungtassamee et al. 2006; Reese et al. 2007; Suppasri et al. 2013a, b). In previous tsunami damage assessment studies, it was found that a tsunami flow depth of 2 m would be a significant threshold for the onset of collapse of a wooden building (Suppasri et al. 2012), as well as for collapse of masonry/brick type constructions (Yamamoto et al. 2006). In line with such results, and considering that the majority of the building stock in Japan is composed of wooden buildings, we can see here that for tsunami flow depth as low as 2 m the probability of reaching or exceeding ds_4 is already high, expressing that a building will suffer at least some structural damage with a probability approaching 50 %. However, the engineering standards for the construction of wooden houses in Japan are typically higher than in other countries previously affected by tsunamis where wooden houses would often collapse at such a water depth (for example, in Thailand or Indonesia, as highlighted by Suppasri et al. 2012).

The 90 % asymptotic confidence intervals for M.1.3.2 are also plotted. It should be noted that despite the adoption of a model which does not account for the ordinal nature of the damage data (partially ordered), the fragility curves for each building class do not cross.

The graphical diagnostics plotted in Fig. 7 still indicate a poor fit of M.1.3.2 for the higher damage states (ds_4 and ds_5) particularly apparent for the wooden buildings, which are the largest building class in the database. By contrast, Fig. 7 shows that the model provides reasonable fit to ds_1 – ds_3 for all buildings. This result can be expected as ds_1 is a damage state representing only the flooding of the building (therefore, flow depth is intrinsically a good predictor), and ds_2 reflects only the need for minor repairs and floor cleaning—so mainly the flooding effects with added presence of small, relatively harmless debris carried inside by the flow. However, for higher damage levels, it can be noted that the increase in the statistical model's complexity fails to capture some variation in the data, in other words the points are visibly away from the 45 degree line. Most importantly, a pattern along this line is present, which gets more pronounced as the damage state increases. This indicates that the systematic component does not capture well the relationship between the explanatory and response variable. Several factors may be causing this lack of fit. Firstly, the use of the flow depth might not be the most representative measure of tsunami intensity. Indeed, this particular parameter mainly drives the

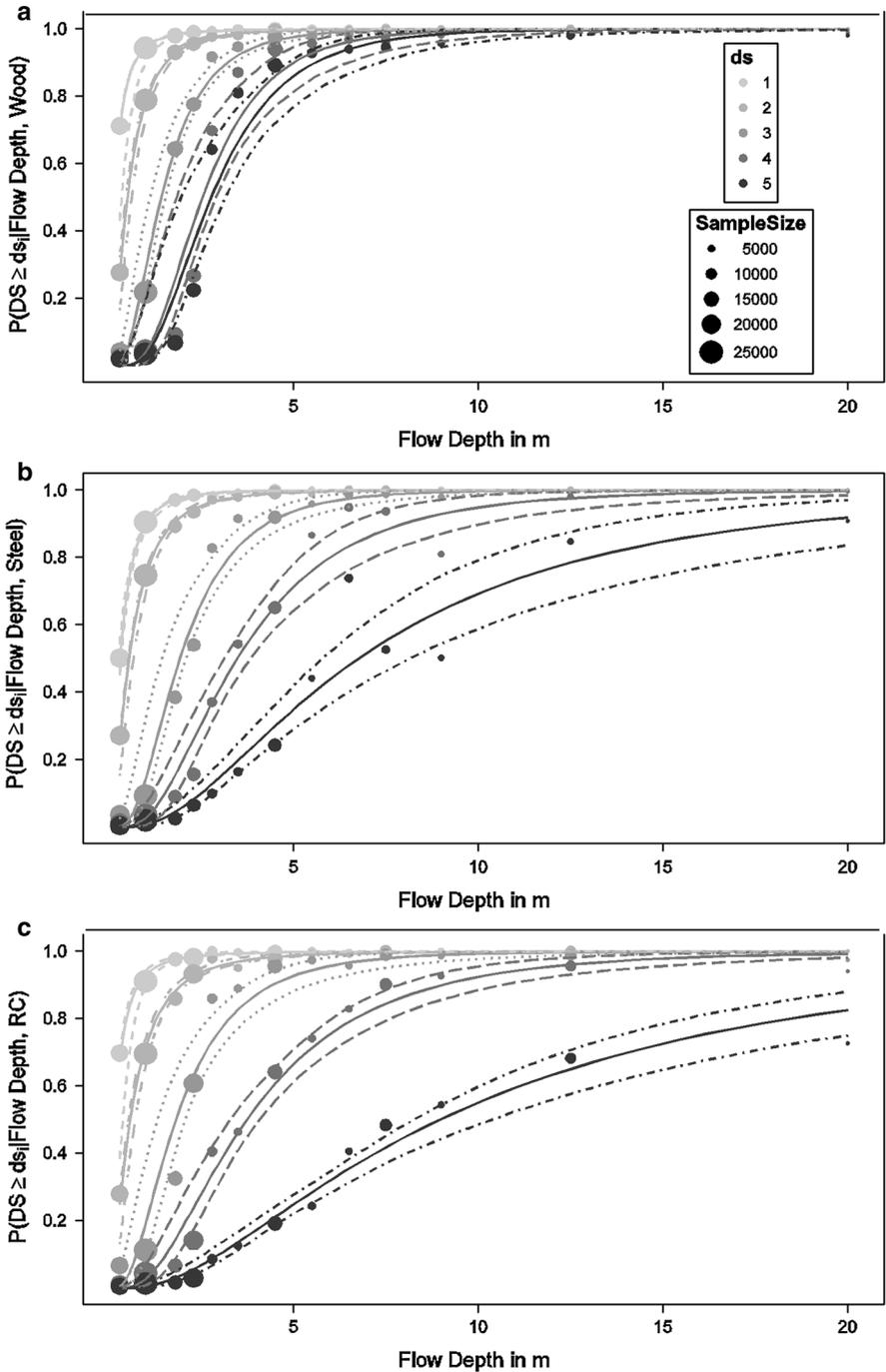


Fig. 6 Fragility curves for **a** wooden **b** steel and **c** RC buildings with their corresponding 90 % confidence intervals (*dotted lines*) constructed by fitting model M.1.3.2 (*solid lines*) to the damage data

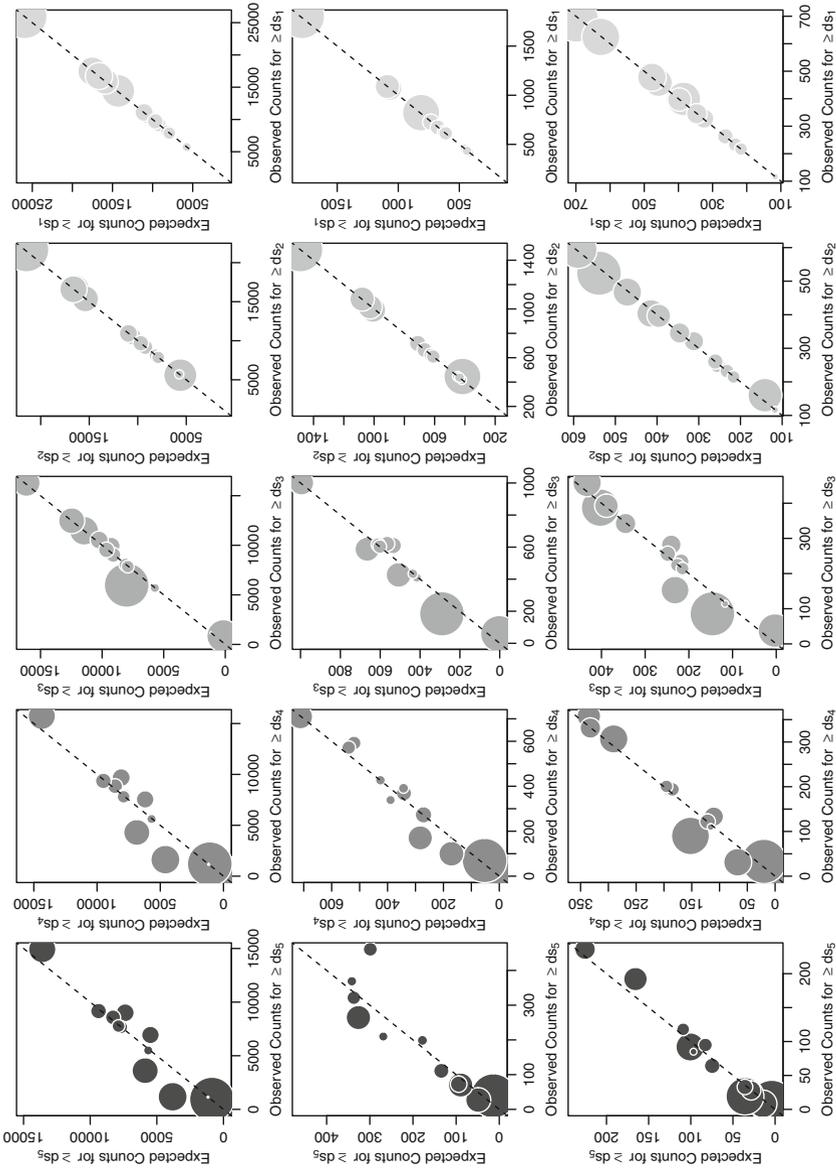


Fig. 7 Graphical diagnostics assessing the goodness-of-fit of the five fragility curves for the model having systematic component expressed by Eq. (5.2)—M1.3.2, for each building material considered (from left to right: wood, steel, RC)



Fig. 8 **a** Severe scour caused by the 2011 tsunami, exposing the buildings foundations (structural damage, at least DS4); **b** extensive damage mainly due to debris impact on the second floor (*wood*), structural damage. **a** Courtesy of Dr J. Bricker, Tohoku University

hydrostatic forces acting on a structure and to an extent the hydrodynamic forces; but it does not capture the effects of flow velocity (also determining the hydrodynamic forces) nor does it reflect the damage caused by debris impact or scour around the building's foundations (Fig. 8). The level of damage caused by debris impact will be related to an extent to flow depth (the larger the depth, the greater the amount of larger debris that can be carried) and also to the flow velocity, as well as the mass and stiffness of the projectiles (FEMA 2008). In other words, for high flow depths many buildings will suffer moderate-to-high levels of damage which can only be weakly explained by flow depth alone as an explanatory variable. The level of damage caused by scour will depend largely on soil conditions, the angle of approach of the flow and cyclic inflow and outflow (Chock et al. 2011). Therefore, additional measures of intensity such as flow velocity, debris load or local soil conditions may prove useful in capturing more uncertainty in the available data.

It is also very likely that many buildings suffered extensive damage due to the earthquake or associated soil liquefaction: such buildings may have reached a given damage state solely due to ground shaking or liquefaction significantly weakening the structure prior to the tsunami arrival (evidence of soil liquefaction in the coastal area close to Sendai airport before tsunami inundation have been reported in Kazama 2011). As it is often impossible in the field to distinguish how much of the building damage is due to the tsunami or preceding ground shaking, such cases are likely to have been classified as highly damaged even for low inundation depths (introducing bias). Finally, the level of data aggregation, both geographically (i.e. the building data has been aggregated across six prefectures) and across the range of flow depths for each bin, has possibly led to a significant loss of information. It is likely that using a dataset of individual buildings and flow depths would improve the fit. Such a dataset may also contain additional information such as geographical location, number of stories, the building's functionality, or year of construction (Suppasri et al. 2013a), some of which are likely to have a significant influence on building performance under tsunami loads.

4.2 Fragility assessment—second stage

The second stage focuses on the investigation of the probability that the failure mode of collapsed buildings is “washed away”, by fitting the models presented in Sect. 3.1.3 to the available data.

Table 5 Deviance, degrees of freedom, p value and AIC values for the examined models

Model	Probit				Logit AIC
	$-2*\log(L)$	df	p value	AIC	
M.2.1	10,179	37	~ 0.00	–	–
M.2.2	7,984	35	~ 0.00	–	–
M.2.3.1	7,540	33		7,770	–
M.2.3.2	–				7,198

The three nested models, termed M.2.1 (flow depth only, probit link), M.2.2 (flow depth and building class, probit link), M.2.3 (flow depth and building class with interactions) presented in Table 2, are fitted to the available data. Thus, we firstly identify the expression of the systematic component which fits the data best by comparing the fits of these models using the likelihood ratio test.

The link function of the aforementioned three models is expressed in terms of the probit function (M.2.3.1 for M.2.3). The suitability of the selected link function is assessed by comparing the goodness of its fit with the fit of the model with the same linear predictor but using the logit link function (M.2.3.2).

The p values in Table 5 show that similarly to the results obtained in the first stage, the more complex M.2.3.1 model fits the data better than the simpler models (M.2.1 and M.2.2), indicating that the importance of taking into account the tsunami flow depth, the construction material, and also the fact that the impact of flow depth depends on the value of class. The comparison of the goodness-of-fit of the two complex models with differing link functions is performed by the use of the Akaike information criterion (AIC)—Eq. (12), presented in Table 5. The results show that the logit link function provides the smallest AIC criterion, indicating that this function provides a relatively better fit to the available data.

Figure 9 depicts the three fragility functions, obtained from model M.2.3.2, and their 90 % confidence intervals. The general shape of the fragility curves show that the wooden buildings are the most likely to be washed away, followed by steel buildings. The reinforced concrete buildings appear to be the least likely to be washed away.

Having identified the link function which best fits the data, the absolute goodness-of-fit of the model is assessed using the graphical diagnostics presented in Fig. 10. A pattern is evident which indicates that the fit is not acceptable, for the same reasons as in the first stage. Obviously for a flow depth close to zero, no buildings should be washed away. Nonetheless, Fig. 9 shows a rather high probability of steel and RC buildings being washed away for flow depths ranging between 0 and 0.5 m. In addition, Table 6 shows the probability of wooden, steel and RC buildings to be washed away for a flow depth of 2 m, with steel buildings surprisingly displaying a slightly higher fragility for this depth, compared with wooden buildings. The curve for steel buildings, and to an extent for RC buildings, appears to be rather flat, indicating large uncertainty. This reveals that the data cannot tell us much about the behaviour (i.e. the mode of collapse) of steel and RC buildings under tsunami loads. From this graph, it appears that once again flow depth is a poor predictor, this time of the building's failure mode.

The wide confidence intervals, particularly for RC and steel buildings, are also a source of concern. They indicate that despite the very large sample size, there is too much unexplained variation in the data to give us confidence in predicting the behaviour of

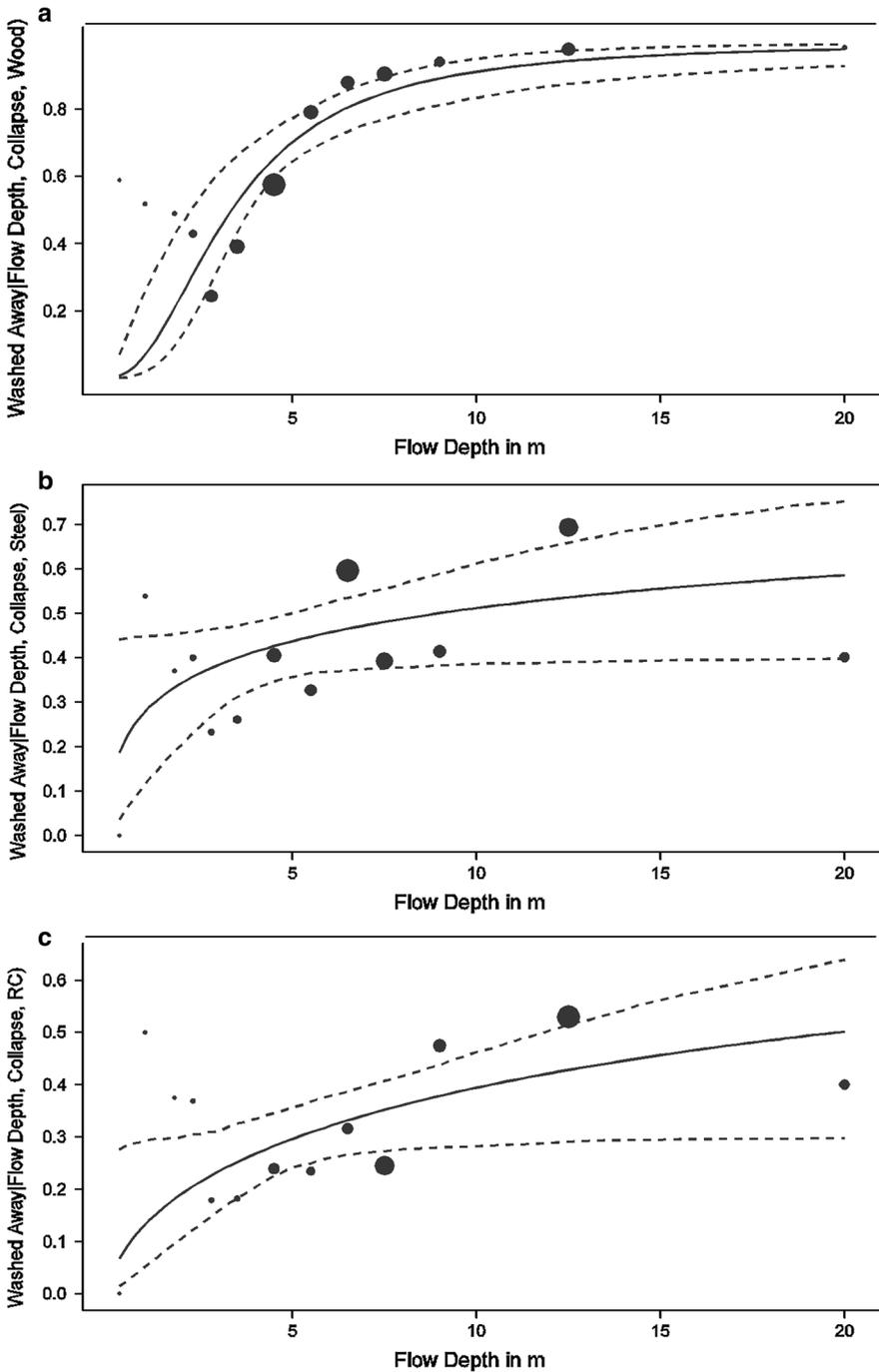


Fig. 9 Fragility corresponding to washed away wooden (a), steel (b) and RC (c) buildings with their corresponding 90 % confidence intervals constructed by fitting the binary model M2.3.2 to the damage data

Fig. 10 Diagnostics plot for model M2.3.2

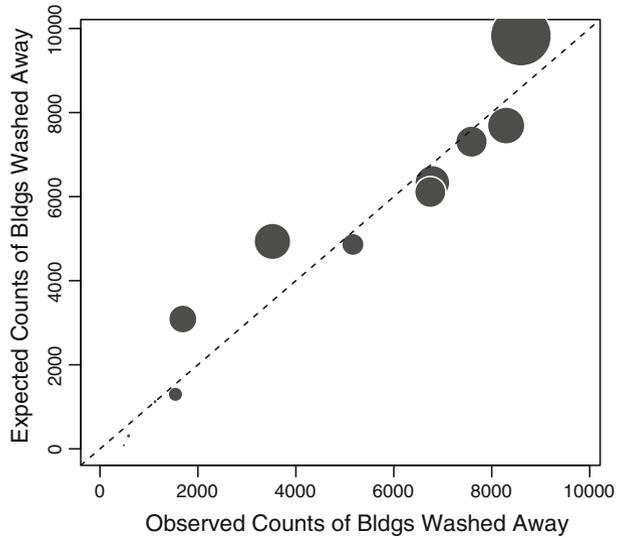


Table 6 Mean probability of a building to be washed away by the tsunami, according to the building class, and upper and lower bounds values given by confidence intervals

Class	<i>P</i> (DS = washed away/2 m)		
	Mean	95 %	5 %
Wood	13 %	12.5 %	14 %
Steel	26 %	20 %	33 %
RC	7.5 %	3.5 %	14.5 %

buildings in a future tsunami. It is clear once more that more explanatory variables and a more detailed database are necessary in order to obtain reliable predictions of tsunami damage probability.

In addition, as highlighted in Sect. 4.1, it is likely that prior earthquake damage may have affected the structure prior to tsunami inundation (for instance, earthquake induced soil liquefaction was widespread in Chiba prefecture, and evidence of this phenomenon was observed a few hundreds of metres from the shoreline nearby Sendai airport (Kazama 2011)). In this particular case, it is highly possible that a building would collapse due to ground shaking (the Great East Japan Earthquake had a magnitude of Mw9.0) and its remains to be carried away by the flow—therefore, the actual damage level would reflect mainly the effects of ground shaking.

5 Conclusions

This study applied regression based on GLM to the extensive, aggregated dataset of buildings damaged by the 2011 Great east Japan tsunami in order to understand the factors that influence tsunami-induced damage and provide tools to the risk assessment community. From a theoretical standpoint, these statistical methods provide a more realistic representation of the available data compared with the ones previously used for similar purposes in the tsunami engineering field. Indeed, they intrinsically allow for the fragility

functions to be bounded between 0 and 1, allow for the complete data to be used (without aggregation and/or unnecessary dismissal), automatically weight each data point according to the number of buildings, and most importantly take into account the discrete, and as applicable the binary, categorical (partially ordered) or ordinal (ordered) nature of the response variable.

The damage scale originally designed by the data provider (MLIT) proved to be inadequately defined for the purpose of performing the required analysis, as a satisfactory ordinal scale should represent consecutive levels that are mutually exclusive, and correspond to an increase in tsunami intensity. Therefore, an improved scale was employed, by aggregating the damage levels that essentially represented the same damage state (i.e. “collapse”), but different failure modes, i.e. DS5 and DS6. The proposed damage scale transformed the seven-state scale given my MLIT into a six-state scale. Modelling the particular failure mode of a building (i.e. “washed away”) may be of interest for future studies that wish to take into account the impact of debris generation, or conditional probability of damage. Therefore, we chose to perform here a two-stage analysis, in order to be able to estimate damage states for the whole population of buildings, and also estimate the probability that a particular building will fail by being washed away.

Stage 1 consisted in deriving fragility functions for five discrete and ordered levels of damage—from “no damage” to “collapse”, and Stage 2 focussed on deriving fragility functions using a binary approach (i.e. the probability of “collapsed” buildings to be “washed away”). A step by step approach, progressing from the simplest model (one explanatory variable: *flow depth*) to the most complex model (two variables, *flow depth* and building class with interactions between the variables) showed that in both stages of analysis, the most complex model displays the best performance. This indicates that both variables are important in estimating tsunami damage and that the impact of flow depth on the rate of damage probability exceedance is simultaneously influenced by the building construction type. The present findings are consistent with previous studies for they show the relatively greater structural vulnerability of wooden structures in comparison with RC and steel buildings. The model appears to be an efficient classifier for small damage states (ds_1 and ds_2), which reflects the obviously dominant role of flow depth in estimating the probability of flooding. However, despite the choice of such statistical tools, the diagnostics (absolute goodness-of-fit) also reveal that even the most complex model does not result in a fully satisfactory fit for high damage states. More specifically, the presence of a slight trend along the perfect predictions line for moderate-to-high damage states indicates that the systematic component fails to capture some aspect of the behaviour of the data. While it could be tempting to use a nonparametric model as a last resort, the presence of such trend demonstrates that some parameter(s) which have been omitted in the analysis are also significant predictors of tsunami damage. Thus, the correct avenue for model improvement would be the inclusion of missing information. The inefficiency of expressing tsunami intensity in terms of flow depth in order to estimate damage, especially for higher damage states, can be explained by examining the physical processes of tsunami inundation, which show the effect of other tsunami parameters (i.e. flow velocity, debris—mass and type), and are expected to be increasingly important for high damage states. In addition, the effect of ground shaking and soil liquefaction from the preceding earthquake is likely to have dramatically enhanced damage—thus introducing bias, which is particularly apparent in Stage 2 of the analysis. Finally, the high level of aggregation of the database (by range of flow depths, and by location) may hide some important information that cannot be captured by the model. This is especially true in the range of smaller flow depths. Data such as flow velocity, debris impact or scour are difficult to measure and/or

quantify in the field, but access to non-aggregated databases can be gained in the future, and numerical simulations may provide useful flow velocity estimates in order to improve the results. The level of detail in the data plays an important role; thus, not all databases are informative, even though the statistical tools applied here have rigorously exploited the information available in the database. In addition, the definition of damage states for tsunami surveys may need to be revised, in order for the scale to be directly usable for GLM analysis and to avoid misclassifications in future field surveys. Provided the availability of high-quality data, GLM have the potential to provide reliable predictions of tsunami-induced damage at the locations under investigation, thus should supersede previous models in the derivation of future fragility functions.

Acknowledgments The authors would like to acknowledge the invaluable advice of Pr Richard Chandler, from the UCL Department of Statistical Science. The work of Ingrid Charvet has been supported by the University College London/Engineering and Physical Sciences Research Council Knowledge Transfer Scheme, and the Japan Society for the Promotion of Science. The work of Ioanna Ioannou and Tiziana Rossetto is supported by the EPSRC “Challenging Risk” project (EP/K022377/1). The work of all UCL authors has also been supported by EPICentre (EP/F012179/1). Ingrid Charvet and Anawat Suppasri are part of the Willis Research Network, under the Pan-Asian Oceanian Tsunami Risk Modeling and Mapping project. The contribution of Fumihiko Imamura has been supported by the Japan Society for the Promotion of Science.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Charvet I, Suppasri A, Imamura F, Rossetto T (2013) Comparison between linear least squares and GLM regression for fragility functions: example of the 2011 Japan tsunami. In: Proceedings of international sessions in coastal engineering, JSCE, vol 4, November 13–15, 2013, Fukuoka, Japan
- Chock GYK, Robertson I, Riggs HR (2011) Tsunami structural design provisions for a new update of building codes and performance-based engineering. In: Proceedings of solutions to coastal disasters 2011, June 26–29 2011, Anchorage, Alaska
- Dias WPS, Yapa HD, Peiris LMN (2009) Tsunami vulnerability functions from field surveys and Monte Carlo simulation. *Civ Eng Environ Syst* 26(2):181–194
- Federal Emergency Management (FEMA) (2008) Guidelines for design of structures for vertical evacuation from tsunamis. FEMA report 646. <http://www.fema.gov/media-library/assets/documents/14708?id=3463>
- Forbes C, Evans M, Hastings N, Peacock B (2011) Statistical distributions. Wiley series in probability and statistics, 4th edn. Wiley, London
- Gokon H, Koshimura S, Matsuoka M, Namegaya Y (2011) Developing tsunami fragility curves due to the 2009 tsunami disaster in American Samoa. In: Proceedings of the coastal engineering conference (JSCE), Morioka, 9–11 November 2011 (in Japanese)
- Green SB (1991) How many subjects does it take to do a regression analysis? *Multivar Behav Res* 26(3):499–510
- Kazama M (2011) Geotechnical challenges to reconstruction and geotechnical disaster by the 2011 Tohoku region Pacific Ocean earthquake and tsunami, technical conference proceedings of the great east Japan earthquake—damages and lessons from the great earthquake and the great tsunami 2011, pp 41–66. <http://www.tohoku-geo-ne.jp/earthquake/img/41.pdf>
- Koshimura S, Oie T, Yanagisawa H, Imamura F (2009) Developing fragility functions for tsunami damage estimation using numerical model and post-tsunami data from Banda Aceh, Indonesia. *Coast Eng J* 51(3):243–273
- Leone F, Lavigne F, Paris R, Denain JC, Vinet F (2011) A spatial analysis of the December 26th, 2004 tsunami-induced damages: lessons learned for a better risk assessment integrating buildings vulnerability. *Appl Geogr* 31:363–375

- Mas E, Koshimura S, Suppasri A, Matsuoka M, Matsuyama M, Yoshii T, Jimenez C, Yamazaki F, Imamura F (2012) Developing tsunami fragility curves using remote sensing and survey data of the 2010 Chilean Tsunami in Dichato. *Nat Hazards Earth Syst Sci* 12:2689–2697
- McCullagh P, Nelder JA (1989) *Generalized linear models*, 2nd edn. Chapman & Hall, London
- Ministry of Land, Infrastructure and Transport (MLIT) (2012) Disaster status survey result of the great east Japan earthquake (first report) related documents. http://www.mlit.go.jp/toshi/city_plan/crd_city_plan_tk_000005.html
- Murao O, Nakazato H (2010) Vulnerability Functions for buildings based on damage survey data in Sri Lanka after the 2004 Indian Ocean Tsunami. In: *Proceedings of international conference on sustainable built environment (ICSBE 2010)*, December 13–14 2010, Kandy
- Peiris N (2006) Vulnerability functions for tsunami loss estimation. In: *Proceedings of 1st European conference on earthquake engineering and seismology*, no. 1121
- Reese S, Cousins WJ, Power WL, Palmer NG, Tejakusuma IG, Nugrahadi S (2007) Tsunami vulnerability of buildings and people in South Java—field observations after the July 2006 Java tsunami. *Nat Hazards Earth Syst Sci* 7:573–589
- 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(2011):156–173
- Rossetto T, Ioannou I, Grant DN (2014) Guidelines for empirical vulnerability assessment. GEM technical report. GEM Foundation, Pavia
- Ruangrassamee A, Yanagisawa H, Foytong P, Lukkunaprasit P, Koshimura S, Imamura F (2006) Investigation of tsunami-induced damage and fragility of buildings in Thailand after the December 2004 Indian Ocean Tsunami. *Earthq Spectra* 22:377–401
- Suppasri A, Koshimura S, Imamura F (2011) Developing tsunami fragility curves based on the satellite remote sensing and the numerical modeling of the 2004 Indian Ocean tsunami in Thailand. *Natral Hazards Earth Syst Sci* 11(1):173–189
- Suppasri A, Mas E, Koshimura S, Imai K, Harada K, Imamura F (2012) Developing tsunami fragility curves from the surveyed data of the 2011 Great East Japan tsunami in Sendai and Ishinomaki Plains. *Coast Eng J* 54(1) (Special Anniversary Issue on the 2011 Tohoku Earthquake Tsunami)
- Suppasri A, Charvet I, Imai K, Imamura F (2013a) Fragility curves based on data from the 2011 Great East Japan tsunami in Ishinomaki city with discussion of parameters influencing building damage. *Earthq Spectra*. doi: [10.1193/053013EQS138M](https://doi.org/10.1193/053013EQS138M)
- 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
- Yamamoto Y, Takanashi H, Hettiarachchi S, Samarawickrama S (2006) Verification of the destruction mechanism of structures in Sri Lanka and Thailand due to the Indian Ocean tsunami. *Coast Eng J* 48(2):117–145