



Empirical tsunami fragility modelling for hierarchical damage levels

Fatemeh Jalayer^{1,2}, Hossein Ebrahimian², Konstantinos Trevelopoulos², and Brendon Bradley³

¹Institute for Risk and Disaster Reduction, University College London, Gower Street, London WC1E 6BT, UK

²Department of Structures for Engineering and Architecture, University of Naples Federico II, Naples 80125, Italy

³Department of Civil and Natural Resources Engineering, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

Correspondence: Fatemeh Jalayer (fatemeh.jalayer@unina.it)

Received: 13 April 2022 – Discussion started: 20 April 2022

Revised: 30 November 2022 – Accepted: 24 January 2023 – Published: 2 March 2023

Abstract. The present work proposes a simulation-based Bayesian method for parameter estimation and fragility model selection for mutually exclusive and collectively exhaustive (MECE) damage states. This method uses an adaptive Markov chain Monte Carlo simulation (MCMC) based on likelihood estimation using point-wise intensity values. It identifies the simplest model that fits the data best, among the set of viable fragility models considered. The proposed methodology is demonstrated for empirical fragility assessments for two different tsunami events and different classes of buildings with varying numbers of observed damage and flow depth data pairs. As case studies, observed pairs of data for flow depth and the corresponding damage level from the South Pacific tsunami on 29 September 2009 and the Sulawesi–Palu tsunami on 28 September 2018 are used. Damage data related to a total of five different building classes are analysed. It is shown that the proposed methodology is stable and efficient for data sets with a very low number of damage versus intensity data pairs and cases in which observed data are missing for some of the damage levels.

are seismic fragility (Rota et al., 2009; Rosti et al., 2021), tsunami fragility (Koshimura et al., 2009a; Reese et al., 2011; a comprehensive review can be found in Charvet et al., 2017), flooding fragility (Wing et al., 2020) and debris flow fragility curves (Eidsvig et al., 2014). Empirical fragility modelling is greatly affected by how the damage and intensity parameters are defined. Mutually exclusive and collectively exhaustive (MECE, see next section for the definition) damage states are quite common in the literature as discrete physical damage states. The MECE condition is necessary for damage states in most probabilistic risk formulations, leading to the mean rate of exceeding loss (e.g. Behrens et al., 2021).

Tsunami fragility curves usually employ the tsunami flow depth as the measure of intensity; although different studies also use other measures like current velocity (e.g. De Risi et al., 2017b; Charvet et al., 2015). Charvet et al. (2015) demonstrate that the flow depth may cease to be an appropriate measure of intensity for higher damage states, and other parameters such as the current velocity, debris impact and scour can become increasingly more important. De Risi et al. (2017b) developed bivariate tsunami fragilities, which account for the interaction between the two intensity measures of tsunami flow depth and current velocity.

Early procedures for empirical tsunami fragility curves used data binning to represent the intensity. For example, Koshimura et al. (2009b) binned the observations by the intensity measure, i.e. the flow depth; however, the latest procedures have mostly used point-wise intensity estimates instead.

Fragility curves for MECE damage states are distinguished by their nicely “laminar” shape; in other words, the curves should not intersect. When fitting empirical fragility

1 Introduction

Fragility models express the probability of exceeding certain damage thresholds for a given level of intensity for a specific class of buildings or infrastructure. Empirical fragility curves are models derived from observed pairs of damage and intensity data for buildings and infrastructure usually collected, acquired and even partially simulated in the aftermath of disastrous events. Some examples of empirical fragility models

curves to observed damage data, this condition is not satisfied automatically. For example, fragility curves are usually fitted for individual damage states separately, and they are filtered afterwards to remove the crossing fragility curves (e.g. Miano et al., 2020), or ordered (“parallel”) fragility models are used from the start (Charvet et al., 2014; Lahcene et al., 2021). Charvet et al. (2014) and De Risi (2017a) also used partially ordered models to derive fragility curves for MECE damage states. They used the multinomial probability distribution to model the probability of being in any of MECE damage states based on binned intensity representation. De Risi et al. (2017a) used Bayesian inference to derive the model parameters for an ensemble of fragility curves.

Empirical tsunami fragility curves are usually constructed using generalised linear models based on probit, logit or the complementary loglog link functions (Charvet et al., 2014; Lahcene et al., 2021). As far as the assessment of the goodness of fit, model comparison and selection are concerned, approaches based on the likelihood ratio and Akaike information criterion, (e.g. Charvet et al., 2014; Lahcene et al., 2021) and on k -fold cross validation have also been used (Chua et al., 2021). For estimating confidence intervals for empirical tsunami fragility curves, bootstrap resampling has been commonly used (Charvet et al., 2014; Lahcene et al., 2021; Chua et al., 2021).

The present paper presents a simulation-based Bayesian method for inference and model class selection for the ensemble modelling of the tsunami fragility curves for MECE damage states for a given class of buildings. By fitting the (positive definite) fragility link function to the conditional probability of being in a certain damage state, given that building is not in any of the preceding states, the method ensures that the fragility curves do not cross (i.e. they are “hierarchical” as in De Risi et al., 2017a). The method uses adaptive Markov chain Monte Carlo simulation (MCMC, Beck and Au, 2002), based on likelihood estimation using pointwise intensity values, to infer the ensemble of the fragility model parameters. Alternative link functions are compared based on log evidence, which considers both the average goodness of fit (based on log likelihood) and the model parsimony (based on relative entropy). This way, among the set of viable models considered, it identifies the simplest model that fits the data best. By “simplest model”, we mean the model having maximum relative entropy (measured using the Kullback–Leibler distance; Kullback and Leibler, 1951) with respect to the data. This usually means the model has a small number of parameters.

2 Methodology

2.1 Definitions of intensity and damage parameters

The intensity measure, IM, (or simply “intensity”, e.g. the tsunami flow depth) refers to a parameter used to convey in-

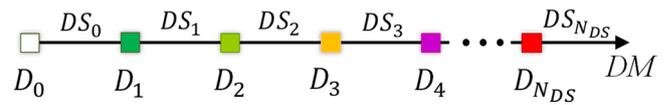


Figure 1. Graphical representation of damage levels D_j and damage states DS_j , where $j = 0 : N_{DS}$.

formation about an event from the hazard level to the fragility level – it is an intermediate variable. The damage parameter, D , is a discrete random variable, and the vector of damage levels is expressed as $\{D_j, j = 0 : N_{DS}\}$, where D_j as the j th damage level (threshold) and N_{DS} as the total number of damage levels considered (depending on the damage scale being used and on the type of hazard, e.g. earthquake, tsunami, debris flow). Normally, D_0 denotes the no-damage threshold, while $D_{N_{DS}}$ defines the total collapse or being totally washed away. Let us assume that DS_j is the j th damage state defined by the logical statement that the damage D is between the two damage thresholds D_j and D_{j+1} ; i.e. D is equal to or greater than D_j and smaller than D_{j+1} as follows (see also Fig. 1 for a graphical representation of the above expressions):

$$DS_j \equiv (D \geq D_j) \cdot (D < D_{j+1}), \quad (1)$$

where (\cdot) denotes the logical product and is read as “AND”. Obviously, for the last damage state, we have $DS_{N_{DS}} \equiv D \geq D_{N_{DS}}$.

Damage states $\{DS_0, DS_1, \dots, DS_{N_{DS}}\}$ are mutually exclusive and collectively exhaustive (MECE) if and only if $P(DS_i \cdot DS_j | IM) = 0$ (if $i \neq j, j = 0 : N_{DS}$) and $\sum_{j=0}^{N_{DS}} P(DS_j | IM) = 1$; (\cdot) denotes the logical product and is read as “AND”. In simple words, the damage states are MECE if being in one damage state excludes all others and if all the damage states together cover the entire range of possibilities in terms of damage. The ensemble of MECE damage states $DS_j, j = 0 : N_{DS}$ is usually referred to as the damage scale (e.g. the EMS98; Grünthal, 1998).

The proposed methodology herein is also applicable to fragility assessment in cases where observed damage data are not available for some of the damage levels. Let **index** be the vector of j values ($j = 0 : N_{DS}$), indicating damage levels N_j for which observed data are available (j values are in ascending order). The new damage scale formed as $\{DS_{index(1)}, DS_{index(2)}, \dots, DS_{index(N)}\}$, where N is the length of vector **index**, is also MECE. It is noteworthy that the number of fragility curves derived in this case is going to be equal to $N - 1$. In the following, for simplicity and without loss of generality, we have assumed that observed data are available for all damage levels, i.e. **index** = $\{0 : N_{DS}\}$, that is, $N_{DS} = N - 1$. However, the proposed methodology is also applicable to the modified damage scale formed by damage level indices in vector **index** ($1 : N$). We will later see examples of such applications in the case studies.

2.2 Fragility modelling using generalised regression models

The generalised regression models (GLMs) are more suitable for empirical fragility assessment with respect to the standard regression models. This is mainly because the dependent variable in the case of the generalised regression models is a Bernoulli binary variable (i.e. only two possible values: 0 or 1). Bernoulli variables are particularly useful in order to detect whether a specific damage level is exceeded or not (only two possibilities). In the following, fragility assessment based on GLMs is briefly described.

The term $P(DS_j | IM)$ denotes the probability of being in damage state DS_j for a given intensity level IM . Based on N_{DS} damage thresholds, this conditional probability $P(DS_j | IM)$ can be read (see Eq. 1) as the probability that $(D \geq D_j)$ and $(D < D_{j+1})$ and can be estimated as follows (see Appendix A for the derivation):

$$\begin{aligned}
 P(DS_j | IM) &= P[(D \geq D_j) \cdot (D < D_{j+1}) | IM] \\
 &= \begin{cases} P(D \geq D_j | IM) - P(D \geq D_{j+1} | IM) & \text{for } 0 \leq j < N_{DS} \\ P(D \geq D_j | IM) & \text{for } j = N_{DS} \end{cases}, \quad (2)
 \end{aligned}$$

where $P(D \geq D_j | IM)$ is the *fragility function* for damage level D_j .

For each damage threshold, fragility can be obtained for a desired building class considering that the damage data provide Bernoulli variables (binary values) of whether the considered damage level was exceeded or not for given IM levels. For damage threshold D_j , all buildings with an observed damage level $D < D_j$ will have a probability equal to zero, while those with $D \geq D_j$ will have an assigned probability equal to one. In other words, for building i and damage state j , the Bernoulli variable Y_{ij} indicates whether building i is in damage state j

$$Y_{ij} = \begin{cases} 1 & \text{if building } i \text{ exceeds } D_j \\ & \text{with probability } P(D \geq D_j | IM_i) \\ 0 & \text{if building } i \text{ does not exceed } D_j \\ & \text{with probability } 1 - P(D < D_j | IM_i) \end{cases}, \quad (3)$$

where IM_i is the intensity evaluated at the location of building i . A Bernoulli variable is defined by one parameter, which is $P(D \geq D_j | IM_i)$ herein. This latter is usually linked to a linear logarithmic predictor in the form

$$l_{ij} = \alpha_{0,j} + \alpha_{1,j} \ln IM_i, \quad (4)$$

where $\alpha_{0,j}$ and $\alpha_{1,j}$ are regression constants for damage level j . We have employed generalised linear regression (e.g. Agresti, 2012) with different link functions “logit”, “probit” and “cloglog” to define probability function π_{ij} as the fol-

lowing:

$$\pi_{ij} = \pi_j(IM_i) = \begin{cases} (1 + \exp(-l_{ij}))^{-1} & \text{logit} \\ \Phi(l_{ij}) & \text{probit} \\ 1 - \exp(-\exp(l_{ij})) & \text{cloglog} \end{cases}. \quad (5)$$

The *logit* link function is equivalent to presenting $\pi_j(IM)$ with a logistic regression function. The *probit* link function is equivalent to a lognormal cumulative distribution function for $\pi_j(IM)$. In the *cloglog* (complementary log–log) transformation, the link function at the location of building i can be expressed as $l_{ij} = \ln[-\ln(1 - \pi_{ij})]$. It is noted that the generalised linear regression based on maximum likelihood estimation (MLE) is available in many statistical software packages (e.g. MathWorks, Python, R).

In the following, we have referred to the general methodology of fitting fragility model to data – one damage state at a time – the “*Basic method*”. In the *Basic method*, the probability of exceeding damage level j is equal to the probability function defined in Eq. (5); that is, $\pi_{ij} = P(D \geq D_j | IM_i)$. This method for empirical fragility curve parameter estimation is addressed in detail in the section “*Results*”, under the “*MLE-Basic*” method. The fragility curves obtained under the “*MLE-Basic*” method could potentially cross, leading to the ill condition that $P(DS_j | IM) < 0$. To overcome this, a *hierarchical fragility modelling approach* has been adopted like that in De Risi et al. (2017a).

2.3 Hierarchical fragility modelling

Equation (2) for $0 \leq j < N_{DS}$, and given IM_i , can also be written as follows using the product rule in probability:

$$\begin{aligned}
 P(DS_j | IM_i) &= P[(D < D_{j+1}) \cdot (D \geq D_j) | IM_i] \\
 &= [1 - P(D \geq D_{j+1} | D \geq D_j, IM_i)] \\
 &\cdot P(D \geq D_j | IM_i). \quad (6)
 \end{aligned}$$

The term $P(D \geq D_{j+1} | D \geq D_j, IM_i)$ embedded in Eq. (6) denotes the conditional probability that the damage exceeds the damage threshold D_{j+1} knowing that it has already exceeded the previous damage level D_j given IM_i . By making $\pi_{ij} = P(D \geq D_{j+1} | D \geq D_j, IM_i)$ (see Eq. 5, which is positive definite), we ensure that the fragility curve of a lower damage level will not fall below the fragility curve of the subsequent damage threshold (the ill condition of $P(DS_j | IM) < 0$ does not take place). Hence, Eq. (6) can be expanded as follows (see Appendix B for derivation):

$$\begin{aligned}
 P(DS_j | IM_i) &= \begin{cases} (1 - \pi_{ij}) \cdot \left[1 - \sum_{k=0}^{j-1} P(DS_k | IM_i) \right] & \text{for } j \geq 1 \\ 1 - \pi_{i0} \triangleq P(D < D_1 | IM_i) & \text{for } j = 0 \end{cases}. \quad (7)
 \end{aligned}$$

In this way, the fragility curves are constructed in a hierarchical manner by first constructing the “*fragility incre-*

ments” $P(DS_j | IM_i)$. Note that for the last damage state $DS_{N_{DS}}$, the probability $P(DS_{N_{DS}} | IM_i)$, which is also equal to the fragility of the ultimate damage threshold $D_{N_{DS}}$, i.e. $P(D \geq D_{N_{DS}} | IM)$ (see Eq. 2), can be estimated by satisfying the collectively exhaustive (CE) condition

$$P(DS_{N_{DS}} | IM_i) = P(D \geq D_{N_{DS}} | IM_i) = 1 - \sum_{j=0}^{N_{DS}-1} P(DS_j | IM_i). \tag{8}$$

Accordingly, the fragility of other damage levels $P(D \geq D_j | IM_i)$, where $0 < j < N_{DS}$, can be obtained from Eq. (2) by starting from the fragility of the higher threshold $P(D \geq D_{j+1} | IM)$ and adding successively $P(DS_j | IM)$ (see Eq. 7) as follows:

$$P(D \geq D_j | IM_i) = P(DS_j | IM_i) + P(D \geq D_{j+1} | IM_i) \text{ for } 0 < j < N_{DS}. \tag{9}$$

As a result, the set of hierarchical fragility models based on Eq. (9) has $2 \times N_{DS}$ model parameters with the vector $\theta = [\{\alpha_{0,j}, \alpha_{1,j}\}, j = 0 : N_{DS} - 1]$. Obviously, with reference to Eq. (8), no model parameter is required for the last damage level, which is derived by satisfying the CE condition. The vector θ of the proposed hierarchical fragility models can be defined by two different approaches

1. *MLE method.* A generalised linear regression model (as explained in previous section) is used for the conditional fragility term $\pi_{ij} = P(D \geq D_{j+1} | D \geq D_j, IM)$ for the j th damage state DS_j (see Eq. 7, $0 \leq j < N_{DS}$). Herein, we need to work with partial damage data so that all buildings in DS_j (with an observed damage $D_j \leq D < D_{j+1}$) will be assigned a probability equal to 0, while those in higher damage states (with $D \geq D_{j+1}$) will be assigned a probability equal to 1 (i.e. in order to model the conditioning on $D \geq D_j$, the domain of possible damage levels is reduced to $D \geq D_j$).
2. *Bayesian model class selection (BMCS).* The Bayesian inference for model updating is employed to obtain the joint distribution of the model parameters.

Detailed discussion about these two approaches, namely “MLE” and “BMCS”, for parameter estimation of empirical fragility curves are provided in Sect. 3.

2.4 Bayesian model class selection (BMCS) and parameter inference using adaptive MCMC

We use the Bayesian model class selection (BMCS) herein to identify the best link model to use in the generalised linear regression scheme. However, the procedure is general and can be applied to a more diverse pool of candidate fragility models. BMCS (or model comparison) is essentially

Bayesian updating at the model class level to make comparisons among candidate model classes given the observed data (e.g. Beck and Yuen, 2004; Muto and Beck, 2008). Given a set of N_M candidate model classes $\{M_k, k = 1 : N_M\}$, and in the presence of the data D , the posterior probability of the k th model class, denoted as $P(M_k | D)$ can be written as follows:

$$P(M_k | D) = \frac{p(D | M_k) P(M_k)}{\sum_{k=1}^{N_M} p(D | M_k) P(M_k)}. \tag{10}$$

In lieu of any initial preferences about the prior $P(M_k)$, one can assign equal weights to each model; thus, $P(M_k) = 1/N_M$. Hence, the probability of a model class is dominated by the likelihood of $p(D|M_k)$ (a.k.a. *evidence*). It is important to note that p herein stands for the probability density function (PDF). Here, data vector $D = \{(IM, DS)_i, i = 1 : N_{CL}\}$ defines the observed intensity and damage data for N_{CL} buildings surveyed for class CL. Let us define the vector of model parameters θ_k for model class M_k as $\theta_k = [\{\alpha_{0,j}, \alpha_{1,j}\}_k, j = 0 : N_{DS} - 1]$. We use the Bayes theorem to write the “evidence” $p(D | M_k)$ provided by data D for model M_k as follows:

$$p(D | M_k) = \frac{p(D | \theta_k, M_k) p(\theta_k | M_k)}{p(\theta_k | D, M_k)}. \tag{11}$$

It can be shown (see Appendix C; Muto and Beck, 2008) that the logarithm of the evidence (called *log-evidence*) $\ln[p(D | M_k)]$ can be written as

$$\begin{aligned} \ln[p(D | M_k)] &= \underbrace{\int_{\Omega_{\theta_k}} \ln[p(D | \theta_k, M_k)] p(\theta_k | D, M_k) d\theta_k}_{\text{Term 1}} \\ &\quad - \underbrace{\int_{\Omega_{\theta_k}} \ln\left[\frac{p(\theta_k | D, M_k)}{p(\theta_k | M_k)}\right] p(\theta_k | D, M_k) d\theta_k}_{\text{Term 2}}, \end{aligned} \tag{12}$$

where Ω_{θ_k} is the domain of θ_k , and $p(D | \theta_k, M_k)$ is the likelihood function conditioned on model class M_k . “Term 1” denotes the posterior mean of the log likelihood, which is a measure of the average data fit to model M_k . “Term 2” is the relative entropy (Kullback and Leibler, 1951; Cover and Thomas, 1991) between the prior $p(\theta_k | M_k)$ and the posterior $p(\theta_k | D, M_k)$ of θ_k given model M_k , which is a measure of the distance between the two PDFs. The latter Term 2 measures quantitatively the amount of information (on average) that is “gained” about θ_k from the observed data D . It is interesting that Term 2 in the log-evidence expression penalises for model complexity; i.e. if the model extracts more information from data (which is a sign of being a complex model

with more model parameters), the log-evidence reduces. The exponential of the log-evidence, $p(\mathbf{D} | \mathbb{M}_k)$, is going to be implemented directly in Eq. (10) to provide the probability attributed to the model class \mathbb{M}_k . More details on how to estimate the two terms in Eq. (12) are provided in Sect. 3.

The likelihood $p(\mathbf{D} | \boldsymbol{\theta}_k, \mathbb{M}_k)$ can be derived, based on point-wise intensity information, as the likelihood of $n_{CL,j}$ buildings being in damage state DS_j (considering that $\sum_{j=0}^{N_{DS}} n_{CL,j} = N_{CL}$), according to data \mathbf{D} defined before

$$p(\mathbf{D} | \boldsymbol{\theta}_k, \mathbb{M}_k) = \prod_{j=0}^{N_{DS}} \prod_{i=1}^{n_{CL,j}} P(DS_j | IM_i). \quad (13)$$

The posterior distribution $p(\boldsymbol{\theta}_k | \mathbf{D}, \mathbb{M}_k)$ can be found based on Bayesian inference

$$\underbrace{p(\boldsymbol{\theta}_k | \mathbf{D}, \mathbb{M}_k)}_{\text{posterior}} = \frac{p(\mathbf{D} | \boldsymbol{\theta}_k, \mathbb{M}_k) p(\boldsymbol{\theta}_k | \mathbb{M}_k)}{\int_{\Omega_{\boldsymbol{\theta}_k}} p(\mathbf{D} | \boldsymbol{\theta}_k, \mathbb{M}_k) p(\boldsymbol{\theta}_k | \mathbb{M}_k) d\boldsymbol{\theta}_k} \\ = C^{-1} \underbrace{p(\mathbf{D} | \boldsymbol{\theta}_k, \mathbb{M}_k)}_{\text{likelihood}} \underbrace{p(\boldsymbol{\theta}_k | \mathbb{M}_k)}_{\text{prior}}, \quad (14)$$

where C^{-1} is a normalising constant. In lieu of additional information (or preferences), the prior distribution, $p(\boldsymbol{\theta}_k | \mathbb{M}_k)$, can be estimated as the product of marginal normal PDFs for each model parameter, i.e. a multivariate normal distribution with zero correlation between the pairs of model parameters $\boldsymbol{\theta}_k$ (see Appendix D). More detail about an efficient prior joint PDF is provided in Sect. 3. To sample from the posterior distribution $p(\boldsymbol{\theta}_k | \mathbf{D}, \mathbb{M}_k)$ in Eq. (14), an *adaptive* MCMC simulation routine (see Appendix E) is employed. MCMC is particularly useful for drawing samples from the target posterior, while it is known up to a scaling constant C^{-1} (see Beck and Au, 2002); thus, in Eq. (14), we only need unnormalised PDFs to feed the MCMC procedure. The MCMC routine herein employs the Metropolis–Hastings (MH) algorithm (Metropolis et al., 1953; Hasting, 1970) to generate samples from the target joint posterior PDF $p(\boldsymbol{\theta}_k | \mathbf{D}, \mathbb{M}_k)$.

2.5 Calculating the hierarchical fragilities and the corresponding confidence intervals based on the vector of model parameters $\boldsymbol{\theta}_k$

For each realisation of the vector of model parameters $\boldsymbol{\theta}_k$, the corresponding set of hierarchical fragility curves can be derived based on the procedure described in the previous sections. Since we have realisations of the model parameters drawn from the joint PDF $p(\boldsymbol{\theta}_k | \mathbf{D}, \mathbb{M}_k)$ (based on samples drawn from adaptive MCMC procedure, see also Appendix E), we can use the concept of robust fragility (RF) proposed in Jalayer et al. (2017) (see also Jalayer et al., 2015; Jalayer and Erahimian, 2020) to derive confidence intervals for the fragility curves. RF is defined as the expected value for a prescribed fragility model considering the joint probability distribution for the fragility model parameters $\boldsymbol{\theta}_k$. The

RF herein can be expressed as

$$P(D \geq D_j | IM, \mathbf{D}, \mathbb{M}_k) \\ = \int_{\Omega_{\boldsymbol{\theta}_k}} P(D \geq D_j | IM, \boldsymbol{\theta}_k) p(\boldsymbol{\theta}_k | \mathbf{D}, \mathbb{M}_k) d\boldsymbol{\theta}_k \\ = \mathbb{E}_{\boldsymbol{\theta}_k | \mathbf{D}, \mathbb{M}_k} [P(D \geq D_j | IM, \boldsymbol{\theta}_k)] \quad , \quad (15)$$

where $P(D \geq D_j | IM, \boldsymbol{\theta}_k)$ is the fragility given the model parameters $\boldsymbol{\theta}_k$ associated with the model \mathbb{M}_k (it has been assumed that once conditioned on fragility model parameters $\boldsymbol{\theta}_k$, the fragility becomes independent of data \mathbf{D}); $\mathbb{E}_{\boldsymbol{\theta}_k | \mathbf{D}, \mathbb{M}_k}$ is the expected value over the vector of fragility parameters $\boldsymbol{\theta}_k$ for model \mathbb{M}_k . The integral in Eq. (15) can be solved numerically by employing Monte Carlo simulation with N_d generated samples from the vector $\boldsymbol{\theta}_k$ as follows:

$$P(D \geq D_j | IM, \mathbf{D}, \mathbb{M}_k) \cong \frac{1}{N_d} \sum_{l=1}^{N_d} P(D \geq D_j | IM, \boldsymbol{\theta}_{k,l}), \quad (16)$$

where $P(D \geq D_j | IM, \boldsymbol{\theta}_{k,l})$ is the fragility given the l th realisation ($l = 1 : N_d$) of the model parameters $\boldsymbol{\theta}_k$ for model \mathbb{M}_k . Based on the definition represented in Eqs. (15) and (16), the variance $\sigma_{\boldsymbol{\theta}_k | \mathbf{D}, \mathbb{M}_k}^2$, which can be used to estimate a confidence interval for the fragility considering the uncertainty in the estimation of $\boldsymbol{\theta}_k$, is calculated as follows:

$$\sigma_{\boldsymbol{\theta}_k | \mathbf{D}, \mathbb{M}_k}^2 [P(D \geq D_j | IM, \boldsymbol{\theta}_k)] \\ = \underbrace{\mathbb{E}_{\boldsymbol{\theta}_k | \mathbf{D}, \mathbb{M}_k} [P(D \geq D_j | IM, \boldsymbol{\theta}_k)^2]}_{\cong 1/N_d \sum_{l=1}^{N_d} P(D \geq D_j | IM, \boldsymbol{\theta}_{k,l})^2} \\ - \underbrace{(\mathbb{E}_{\boldsymbol{\theta}_k | \mathbf{D}, \mathbb{M}_k} [P(D \geq D_j | IM, \boldsymbol{\theta}_k)])}_{= P(D \geq D_j | IM, \mathbf{D}, \mathbb{M}_k)}^2 \quad (\text{Eq. 16}). \quad (17)$$

The empirical fragilities derived through the hierarchical fragility procedure are not necessarily attributed to a lognormal distribution. Hence, we have derived equivalent lognormal statistics (i.e. the median and dispersion) for the resulting fragility curves. The median intensity, η_{IMC} , for a given damage level is calculated as the IM corresponding to 50 % probability on the fragility curve. The logarithmic standard deviation (dispersion) of the equivalent lognormal fragility curve at the onset of damage threshold, β_{IMC} , is estimated as half of the logarithmic distance between the IMs corresponding to the probabilities of 16 % (IM_C^{16}) and the 84 % (IM_C^{84}) on the fragility curve; thus, the dispersion can be estimated as $\beta_{IMC} = 0.50 \times \ln(IM_C^{84} / IM_C^{16})$. The overall effect of epistemic uncertainties (due to the uncertainty in the fragility model parameters and reflecting the effect of limited sample size) on the median of the empirical fragility curve is considered through (logarithmic) intensity-based standard deviation denoted as β_{UF} (see Jalayer et al., 2020). β_{UF} can be estimated as half of the (natural) logarithmic distance (along

Table 1. The classification of damage thresholds (the damage scale) used for the 2009 South Pacific tsunami (from Reese et al., 2011).

Damage level	Damage level description	
D_0	None	no damage
D_1	Light	non-structural damage
D_2	Minor	significant non-structural damage, minor structural damage
D_3	Moderate	significant structural and non-structural damage
D_4	Severe	irreparable structural damage, will require demolition
D_5	Collapse	complete structural collapse

the IM axis) between the median intensities (i.e. 50 % probability) of the RF's derived with 16 % (denoted as IM^{84}) and 84 % (IM^{16}) confidence levels, respectively; i.e. $\beta_{UF} = 0.50 \times \ln(IM^{84}/IM^{16})$. The RF and its confidence band, the sample fragilities $\theta_{k,l}$ (where $l = 1 : N_d$), the equivalent log-normal parameters η_{IM_C} and β_{IM_C} of the RF, the epistemic uncertainty β_{UF} , and finally the intensities IM^{16} and IM^{84} are shown in Figs. 2d to 6d in the following Sect. 3.

3 Results

3.1 Case study 1: the 2009 South Pacific tsunami

The central South Pacific region-wide tsunami was triggered by an unprecedented earthquake doublet (M_w 8.1 and M_w 8.0) on 29 September 2009, between about 17:48 and 17:50 UTC (Goff and Dominey-Howes, 2011). The tsunami seriously impacted numerous locations in the central South Pacific. Herein, the damage data related to the brick masonry residential (1 storey) and Timber residential buildings associated with the reconnaissance survey sites of American Samoa and the Samoan Islands were utilised. Based on the observed damage regarding different indicators (see Reese et al., 2011 for more details on damage observation), each structure was assigned a damage state between (DS_0 and DS_5). The original data documented in Reese et al. (2011) reporting the tsunami flow depth and the attributed damage state to each surveyed building can be found on the site of the European Tsunami Risk Service (<https://eurotsunamirisk.org/datasets/>, last access: 30 November 2022). The five damage levels ($N_{DS} = 5$) and a description of the indicators leading to the classification of the damage states are given in Table 1 based on Reese et al. (2011).

3.2 Case study 2: the 2018 Sulawesi–Palu tsunami

On Friday 28 September 2018, at 18:02 local time, a shallow strike-slip earthquake of moment magnitude M_w 7.5 occurred near Palu City, Central Sulawesi, Indonesia followed by submarine landslides, a tsunami, and massive liquefaction caused substantial damage (Muhari et al., 2018; Raffiana et al., 2022). In Sulawesi, more than 3300 fatalities and missing people, 4400 serious injuries and 170 000 people were displaced by earthquake, tsunami, landslides, liquefaction or

Table 2. The classification of damage thresholds (the damage scale) used for the 2018 Sulawesi–Palu tsunami (from Paulik et al., 2019).

Damage level	Damage level description	
D_0	None	no damage
D_1	Repairable	partial damage, repairable
D_2	Unrepairable	partial damage, unrepairable
D_3	Complete	complete structural collapse

building collapse, or combinations of these hazards (Paulik et al., 2019; Mas et al., 2020). Herein, the damage data related to the non-engineered unreinforced clay brick masonry (1 and 2 storey) and non-engineered light timber buildings located in Palu City were utilised. Based on the observed damage (see Paulik et al., 2018 for details on damage observation), each structure was assigned a damage state between (DS_0 and DS_3). The original data reporting the tsunami flow depth and the attributed damage state to each surveyed building can be found as the Supplement to Paulik et al. (2018). The three damage levels ($N_{DS} = 3$) and a description of the indicators leading to the classification of the damage state are given in Table 2.

3.3 The building classes

Table 3 illustrates the building classes, for which fragility curves are obtained based on the proposed procedure and based on the databases related to the two tsunami events described above. The taxonomy used for describing the building class matches the original description used in the raw databases. The number of data points available for different building classes showcases both classes with large number of data available, e.g. brick masonry 1 storey (South Pacific) and non-engineered brick masonry 1 storey (Sulawesi), and classes with few data points available, e.g. timber residential (South Pacific) and non-engineered masonry 2 storeys and timber (Sulawesi). The fifth column in the table shows the proportion of the number of damage levels for which observed data are available N (see Sect. 2.1) to the total number of damage levels in the corresponding damage scales, namely $N_{DS} + 1 = 6$ for South Pacific and $N_{DS} + 1 = 4$ for Sulawesi tsunami events (to include level 0). If the ratio is equal to unity, it indicates that data are available for all the damage levels from 0 to N_{DS} . Note that the number of fragility curves derived is going to be equal to $N - 1$, that is, equal to the number of damage levels for which observed damage data are available minus one.

3.4 The different model classes

For each building class considered, we have considered the set of candidate models consisting of the fragility models resulting from the three alternative link functions used in the generalised linear regression in Eq. (5). That is, \mathbb{M}_1 refers to hierarchical fragility modelling based on “logit”, \mathbb{M}_2 refers

Table 3. The building classes.

Building class	Tsunami event	Number of data	$N/(N_{DS} + 1)$, <i>index</i>
1 Brick masonry residential, 1 storey	South Pacific 2009	120	6/6, <i>index</i> = {0, 1, 2, 3, 4, 5}
2 Timber residential		23	4/6, <i>index</i> = {2, 3, 4, 5}
1 Non-engineered masonry, unreinforced with clay brick, 1 storey	Sulawesi–Palu 2018	279	3/4, <i>index</i> = {0, 1, 2}
2 Non-engineered masonry, unreinforced with clay brick, 2 storeys		37	3/4, <i>index</i> = {0, 1, 2}
3 Non-engineered light timber		14	3/4, <i>index</i> = {1, 2, 3}

to hierarchical fragility modelling based on “probit”, and \mathbb{M}_3 refers to hierarchical fragility modelling based on “cloglog”. For each model, both the MLE method using the MATLAB generalised regression toolbox and the BMCS using the procedure described in the previous section are implemented.

3.5 Fragility modelling using MLE

The first step towards fragility assessment by employing the MLE method (see Sect. 2.3) is to define the vector of model parameters $\theta = \{\alpha_{0,j}, \alpha_{1,j}\}$, where $j = \mathbf{index}(1 : (N - 1))$ and where vector *index* is defined in Sect. 2.1 as the vector of damage level indices (in ascending order) for which observed damage data are available and N is the length of *index*. To accomplish this, the model parameters $\{\alpha_{0,j}, \alpha_{1,j}\}$ are obtained by fitting the link functions in Eq. (5) to conditional fragility $P(D \geq D_{j+1} | D \geq D_j, \text{IM})$ according to Eq. (6). Herein, we have used MATLAB as a statistical software package (developed by MathWorks) to estimate the maximum likelihood of model parameters $\{\alpha_{0,j}, \alpha_{1,j}\}$ using the following MATLAB command: `glmfit(log(xj), yj, ‘binomial’, ‘link’, ‘model’)`. The ‘model’ will be either ‘logit’, ‘probit’, or ‘comploglog’. For each damage level D_{j+1} , $\mathbf{index}(1) \leq j < \mathbf{index}(N)$, the vector x_j is the IM’s for which the condition $D \geq D_j$ is satisfied; y_j is the column vector containing one-to-one probability assignment to the IM data in x_j , with zero (= 0.0) assigned to those data corresponding to DS_j ($D_j \leq D < D_{j+1}$) and one (= 1.0) to those related to higher damage states (with $D \geq D_{j+1}$).

The vectors defining the MLE of the model parameters, θ_{MLE} , are presented in Table 4 for each of the building classes listed in Table 3 and for each of the three models \mathbb{M}_1 , \mathbb{M}_2 , and \mathbb{M}_3 defined in Sect. 3.4. Given the model parameter θ_{MLE} , the damage state probability $P(DS_j | \text{IM})$ can be estimated based on the recursive Eqs. (7) and (8). Then, the fragility for the ultimate damage level is calculated first based on Eq. (8). For the lower damage thresholds, the empirical fragility is derived based on Eq. (9). The resulting hierarchical fragility curves by employing the direct fragility assessment given θ_{MLE} , i.e. $P(D \geq D_j | \text{IM}, \theta_{MLE})$

for $j = \mathbf{index}(2 : N)$, are shown later in the next section by comparison with those obtained from the BMCS method.

3.6 Fragility modelling using BMCS

In the first step, the model parameters are estimated for each model class separately. For each model class \mathbb{M}_k , the model parameters θ_k are estimated through the adaptive MCMC method described in detail in Appendix E, which yields the posterior distribution in Eq. (14). With reference to Eq. (14), the prior joint PDF $p(\theta_k | \mathbb{M}_k)$ is a multivariate normal PDF with zero correlation between the pairs of model parameters (see Appendix D). The vector of the mean values, μ_θ is set to be the MLE tabulated in Table 4 (= θ_{MLE} related to \mathbb{M}_k). We have attributed a high value for the coefficient of variation (more than 3.20 herein) for each of the model parameters. Appendix F illustrates the histograms representing the drawn samples from the joint posterior PDF $p(\theta_k | D, \mathbb{M}_k)$ for a selected building class.

The RF curves derived from the hierarchical fragility curves (see Sect. 2.5) and the corresponding plus/minus 1 standard deviation ($\pm 1\sigma$) intervals from Eqs. (16) and (17) are also plotted in Figs. 2a to 6a corresponding to Classes 1–2 South Pacific tsunami and Classes 1–3 Sulawesi–Palu for one of the model classes \mathbb{M}_k ($k \in \{1, 2, 3\}$) and for the damage thresholds D_j , $j = \mathbf{index}(2 : N)$. The colours of the hierarchical robust fragility curves, labelled as D_j -BMCS, match closely with those shown in Tables 1 and 2. The corresponding $\pm 1\sigma$ confidence interval curve, which reflects the uncertainty in the model parameters, is shown as a light grey area with different colour intensities. Figures 2b to 6b compare the hierarchical robust fragility and its confidence interval with the result of hierarchical fragility assessment based on maximum likelihood estimation (see previous Sect. 3.5) labelled as D_j -MLE. The fragility curves are shown with similar colours (and darker intensity) and with the same line type (and half of the thickness) of the corresponding robust fragility curves. The first observation is that the results of MLE-based fragilities and the BMCS-based fragilities are quite close in all damage thresholds (as expected, see Jalayer and Erahimian, 2020). Moreover, the BMCS provides also the confidence bands for the fragility curves, which can-

Table 4. The model parameters θ_{MLE} .

	Building class	Model class	$\alpha_{0,0}$	$\alpha_{1,0}$	$\alpha_{0,1}$	$\alpha_{1,1}$	$\alpha_{0,2}$	$\alpha_{1,2}$	$\alpha_{0,3}$	$\alpha_{1,3}$	$\alpha_{0,4}$	$\alpha_{1,4}$
South Pacific tsunami 2009	1	M_1	5.242	4.190	3.900	4.255	-1.175	4.805	-1.345	2.887	-1.994	2.917
		M_2	2.742	2.190	2.007	2.221	-0.670	2.804	-0.803	1.745	-1.157	1.733
		M_3	2.079	2.011	1.322	1.850	-1.268	3.057	-1.366	1.961	-1.981	2.218
	2	M_1					1.127	1.512	2.484	0.771	-2.846	7.708
		M_2					0.657	0.909	1.390	0.426	-1.575	4.316
		M_3					0.251	0.862	0.883	0.355	-2.141	4.648
Sulawesi–Palu tsunami 2018	1	M_1	6.059	4.355	-0.630	2.909						
		M_2	3.264	2.340	-0.371	1.709						
		M_3	2.498	2.088	-0.907	2.056						
	2	M_1	3.556	3.672	-2.486	5.126						
		M_2	2.077	2.144	-1.464	3.036						
		M_3	1.664	2.239	-2.451	4.217						
	3	M_1			0.466	1.375	0.474	1.195				
		M_2			0.295	0.847	0.296	0.774				
		M_3			-0.041	1.068	-0.076	0.835				

not be directly provided by the MLE method. To showcase an individual fragility curve, Figs. 2c to 6c illustrate the empirical fragility curves associated with the l th realisation of the vector of model parameters $\theta_{k,l}$ for model class M_k (where l is defined on each figure separately), i.e. $P(D \geq D_j | IM, \theta_{k,l})$ (see Sect. 2.5). Figures 2d to 6d illustrate the robust fragility curve associated with the ultimate damage threshold $D_{index(N)}$, together with all the sample fragilities. The intensity values for which the damage level is not exceeded are shown with blue circles having the probability equal to zero. Other IMs that lead to the exceedance of the damage level are shown with red circles with a probability equal to one. Figures 2d to 6d also illustrate all the fragility parameters described in Sect. 2.5 including the equivalent lognormal parameters η_{IM_C} and β_{IM_C} ; the epistemic uncertainty in the empirical fragility assessment β_{UF} ; and also the intensities IM_C^{16} , IM_C^{84} , IM^{84} , and IM^{16} (the latter two are IMs at the median, i.e. 50 % probability, from the $RF \pm 1$ standard deviation, respectively). For all five building classes considered (see Table 3), η_{IM_C} , β_{IM_C} , and β_{UF} are tabulated in Table 5 for all damage thresholds associated to model classes M_1 to M_3 .

3.7 Model selection

With reference to Eq. (12), the *log-evidence* $\ln [p(D|M_k)]$ can be estimated by subtracting Term 2 from Term 1. Term 1 denotes the posterior mean of the log likelihood, and Term 2 is the relative entropy between the prior and the posterior. Within the BMCS method, these two terms are readily computable.

Given the samples generated from the joint posterior PDFs θ_k , Term 1 (the average data fit) can be seen as the expected

value of the log likelihood over the vector of fragility parameters θ_k given the model M_k , i.e. $E_{\theta_k|D, M_k} (\ln [p(D|M_k)])$. Term 2 (the relative entropy) is calculated as the expected value of information gain or entropy between the two PDFs posterior and prior over the vector θ given the model M_k , i.e. $E_{\theta_k|D, M_k} (\ln [p(\theta_k|D, M_k) / p(\theta_k|M_k)])$. It is noted that based on Jensen’s inequality, the mean information gain (relative entropy) of the posterior compared to the prior is always non-negative (see e.g. Jalayer et al., 2012; Ebrahimian and Jalayer, 2021). Hence, Term 2 should always be positive. Herein, $p(\theta_k|D, M_k)$ is constructed by an adaptive kernel density function (see Eq. E5, Appendix E) as the weighted sum (average) of Gaussian PDFs centred among the samples θ_k given model M_k ($k = 1 : 3$). The prior $p(\theta_k|M_k)$ is a multivariate normal PDF, respectively, with the mean and covariance described previously for each model (see Eq. D1 in Appendix D). Table 6 shows the results for model class selection for all five building classes considered. The last column illustrates the posterior probability (weight) of the model $P(M_k|D)$ according to Eq. (10) assuming that the prior $P(M_k) = \frac{1}{3}$ (where $k = 1 : 3$). The best model for each building class is shown in bold.

For instance, for Class 1 (masonry residential) for the South Pacific tsunami, model class M_3 (using a complementary log–log “cloglog” transformation of π_{ij} to the linear logarithmic space, see Eq. 5) is preferred, since it has an overall larger difference between data fit and mean information gain, which leads to a higher log-evidence. The posterior weights (last column of Table 6, see also Eq. 10) of 6 %, 11 % and 83 % are stabilised through different runs of the BMCS method. It should be noted that in Figs. 2 to 6 we reported directly the fragility results for the “best” fragility model class

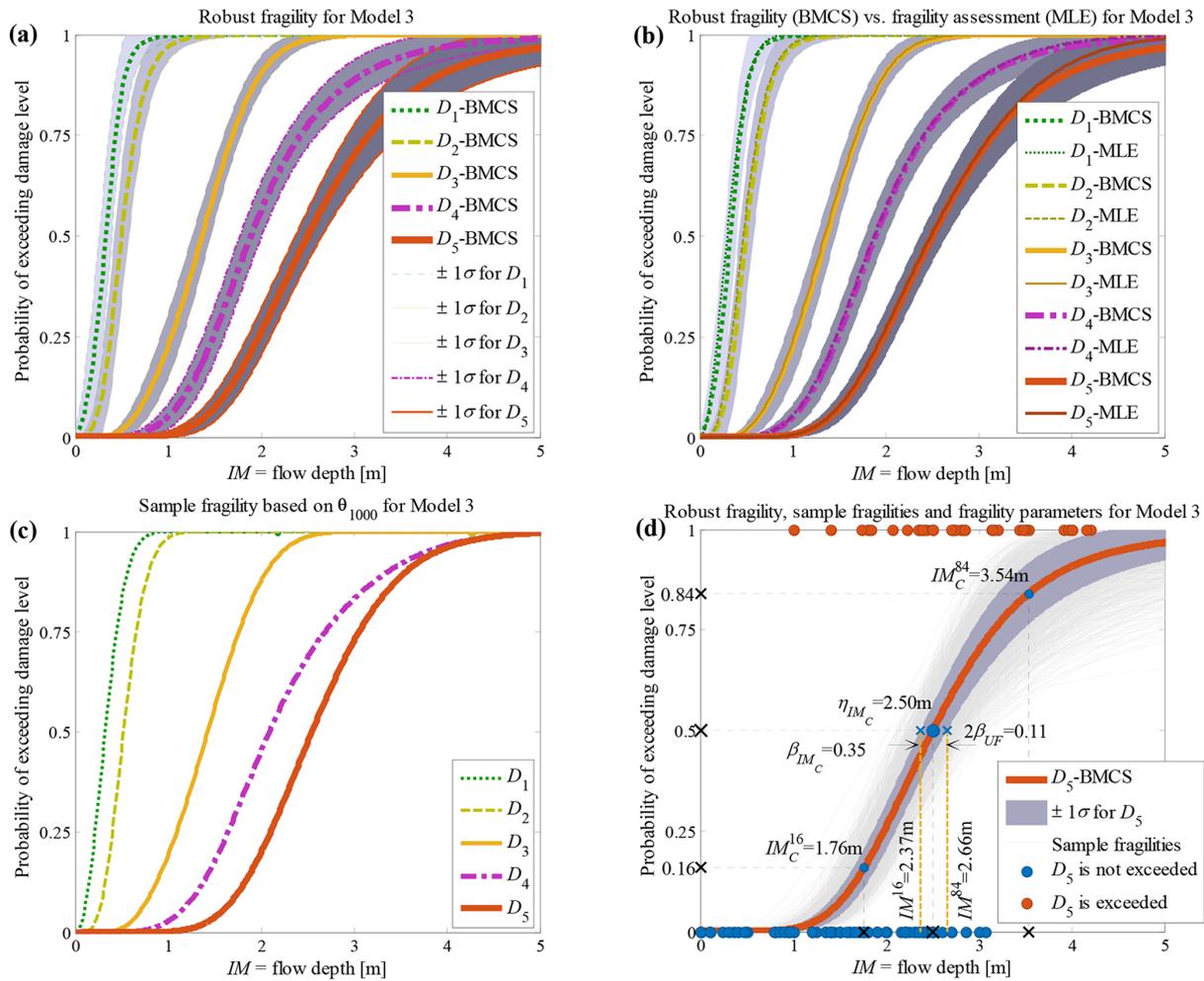


Figure 2. Building class 1 (brick masonry residential, 1 storey) of the South Pacific 2009 tsunami considering fragility model class \mathbb{M}_3 (a) hierarchical robust fragility curves and their ± 1 standard deviation confidence intervals; (b) comparison between hierarchical robust fragility curves and their confidence band (based on BMCS method) and fragility assessment based on MLE method; (c) the fragility curves $P(D \geq D_j | IM, \theta_{3,1000})$, where $1 \leq j \leq 5$ associated with the 1000th realisation of the model parameters, $\theta_{3,1000}$ ($k = 3$ associated to model \mathbb{M}_3 , $l = 1000$); and (d) RF associated with the damage threshold D_5 , together with all the sample fragilities and the equivalent lognormal fragility parameters.

(i.e. the one that maximises log evidence) identified based on the procedure described here.

3.8 The “basic” (MLE-Basic) method: fitting data to one damage state at a time

In the basic method (see Sect. 2.2), the fragility $P(D \geq D_j | IM)$ is obtained by using a generalised linear regression model according to Eq. (5) with logit, probit or cloglog link function fitted to the damage data (\mathbb{M}_k where $k = 1 : 3$). With reference to the MLE method described previously, the vector x_j herein is the IM associated with all damage data (and not partial, as in the hierarchical fragility method described in Sect. 3.3), and y_j is the column vector of one-to-one probability assignment to the IM data in x_j , with zero (= 0) assigned to those data with an observed dam-

age threshold $D < D_j$ and one (= 1) to those with $D \geq D_j$. Thus, for the empirical fragility associated with the damage threshold D_j , and based on the model \mathbb{M}_k , there are two model parameters to be defined, namely $\theta_{MLE-Basic} = \{\alpha_0, \alpha_1\}_k$. As noted previously, there might be conditions (depending on the quantity of the observed damage data) where a part of the fragility of damage threshold D_j lies below the fragility of the higher damage level D_{j+1} , indicating that $P(DS_j | IM) < 0$. This is due to the fact that in the traditional method there is no explicit requirement to satisfy $P(DS_j | IM) > 0$ as compared to the proposed method. The MLE of model parameters $\{\alpha_0, \alpha_1\}$ for the damage levels D_j , $j = index (2 : N)$ associated with the building classes in Table 3 are presented in Table 7.

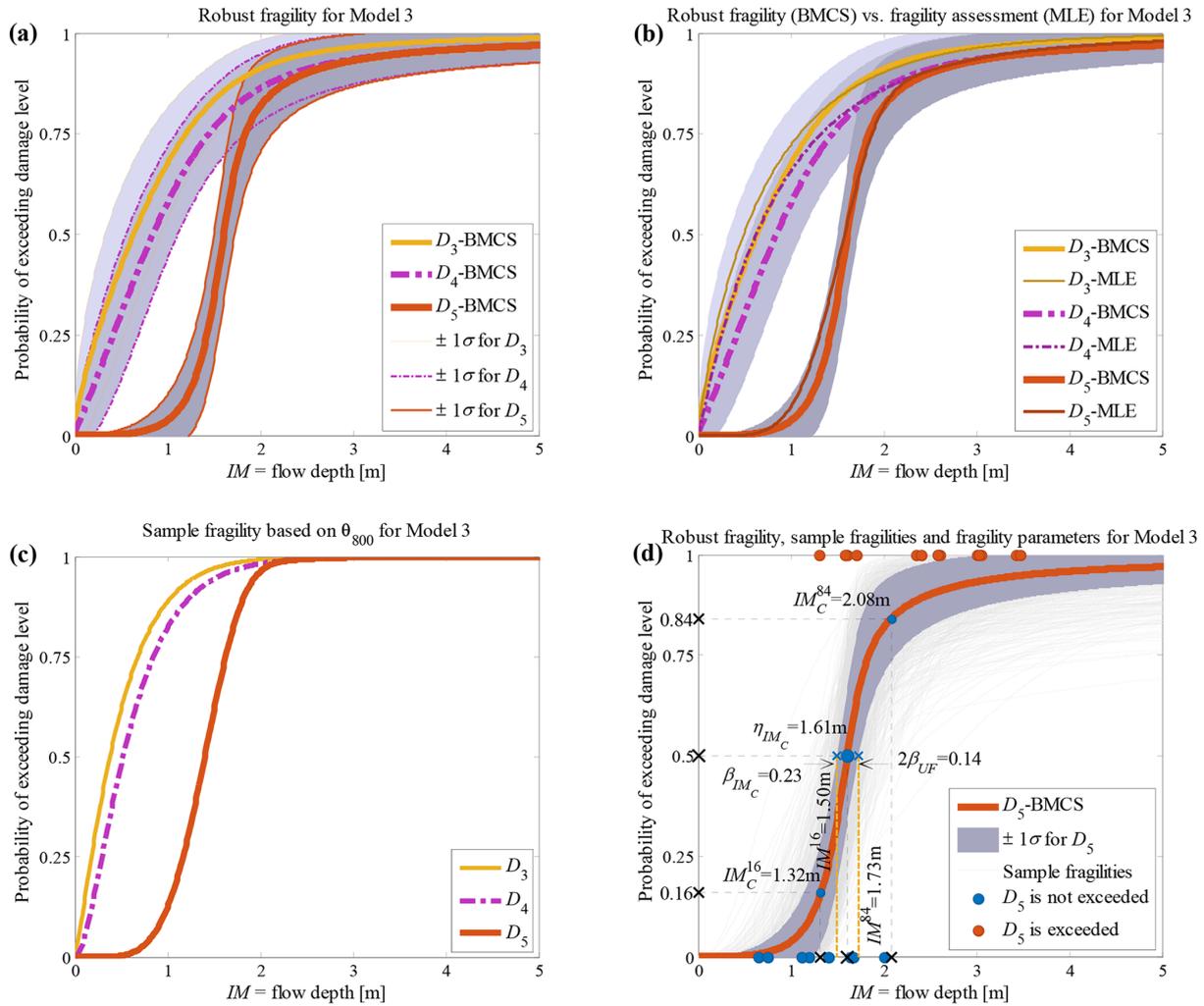


Figure 3. Building class 2 (timber residential) of the South Pacific 2009 tsunami considering fragility model class \mathbb{M}_3 (a) hierarchical robust fragility curves and their ± 1 standard deviation confidence intervals; (b) comparison between hierarchical robust fragility curves and their confidence band (based on BMCS method) and fragility assessment (based on MLE method); (c) the fragility curves $P(D \geq D_j | IM, \theta_{3,800})$ where $3 \leq j \leq 5$ associated with the 800th realisation of the model parameters, $\theta_{3,800}$ ($k = 3$ associated to model \mathbb{M}_3 , $l = 800$); and (d) robust fragility associated with the damage threshold D_5 , together with all the sample fragilities and the equivalent lognormal fragility parameters.

Figure 7 compares the fragility assessment obtained based on MLE-based hierarchical fragility modelling (see also the MLE-based curves in Figs. 2b to 6b) with the result of the fragility assessment by employing the MLE-Basic method for the “best” model class \mathbb{M}_k ($k \in \{1, 2, 3\}$) identified according to the procedure outlined in the previous section. It is noted that the fragility functional form is different between the two methods. MLE-based fragility assessment given \mathbb{M}_k uses Eqs. (7) to (9) to construct a hierarchical fragility curve given that the conditional fragility term $\pi_{ij} = P(D \geq D_{j+1} | D \geq D_j, IM_i)$, $j = \text{index}(1 : N - 1)$ has one of the functional forms in Eq. (5). However, the fragility assessment using the MLE-Basic method employs directly one of the expressions in Eq. (5) (corresponding to \mathbb{M}_k , $k = 1 : 3$) to derive the fragility curve $\pi_{ij} = P(D \geq D_j | IM_i)$ (based

on the whole damage data) and $j = \text{index}(2 : N)$. This difference manifests itself in Fig. 7a (for brick masonry residential, Class 1, South Pacific tsunami) as the deviation between the two fragility models for higher damage thresholds D_4 and D_5 . The deviations between the fragility curves are particularly noticeable at higher IM values (with exceedance probability $> 50\%$); however, their medians are quite similar. In Fig. 7b for (timber residential buildings, Class 2, South Pacific tsunami) and 7e (light informal timber buildings, Class 3, Sulawesi–Palu tsunami), we can observe that fragility curves intersect in the case of the MLE-Basic fragility assessment. However, they do not intersect for hierarchical fragility curves. The intersection points of the consequent damage states D_{j+1} with D_j when using the MLE-

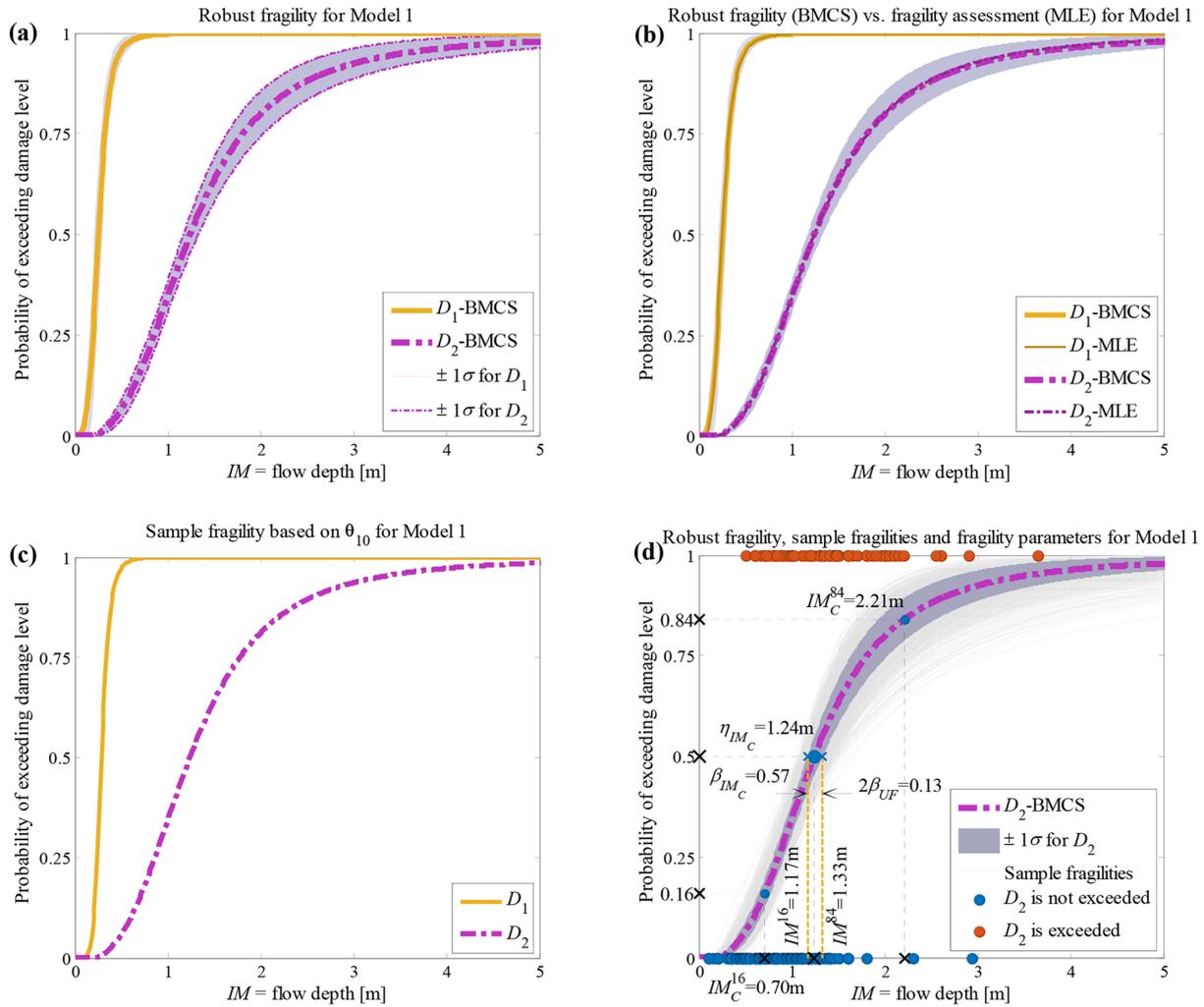


Figure 4. Building class 1 (non-engineered masonry, unreinforced with clay brick, 1 storey) of the Sulawesi–Palu 2018 tsunami considering fragility model class \mathbb{M}_1 (a) robust fragility curves (RF) and their ± 1 standard deviation confidence intervals; (b) comparison between hierarchical robust fragility and its confidence band (based on BMCS method) and fragility assessment based on MLE method; (c) the fragility curves $P(D \geq D_j | IM, \theta_{1,10})$, where $1 \leq j \leq 2$ associated with the 10th realisation of the model parameters, $\theta_{1,10}$ ($k = 1$ associated to model \mathbb{M}_1 , $l = 10$); and (d) robust fragility curve associated with the damage threshold D_2 , together with all the sample fragilities and the equivalent lognormal fragility parameters.

Basic fragility estimation method are shown with colour stars on each figure.

Table 8 reports the fragility assessment parameters of the MLE and MLE-Basic methods for the damage thresholds $D_{index(2)}$ to $D_{index(N)}$ with the equivalent lognormal parameters η_{IM_C} and β_{IM_C} (explained in Sect. 2.5) for \mathbb{M}_1 to \mathbb{M}_3 for all five classes considered. The medians are almost identical among the four models while there are higher dispersion estimates for the MLE method derived by hierarchical fragility modelling.

4 Discussion

The results outlined in this section show a fragility assessment for two different data sets corresponding to observed damaged in the aftermath of South Pacific and Sulawesi–Palu tsunami events. We have demonstrated the versatility of the proposed workflow and tool for hierarchical fragility assessment both for cases in which a large number of data points are available (e.g. Class 1, brick masonry residential, South Pacific tsunami; Class 1, one-storey non-engineered masonry, Sulawesi–Palu tsunami) and cases where very few data points are available (e.g. Class 2, timber residential, South Pacific tsunami; Class 3, non-engineered light timber, Sulawesi–Palu tsunami). Moreover, we demonstrated how

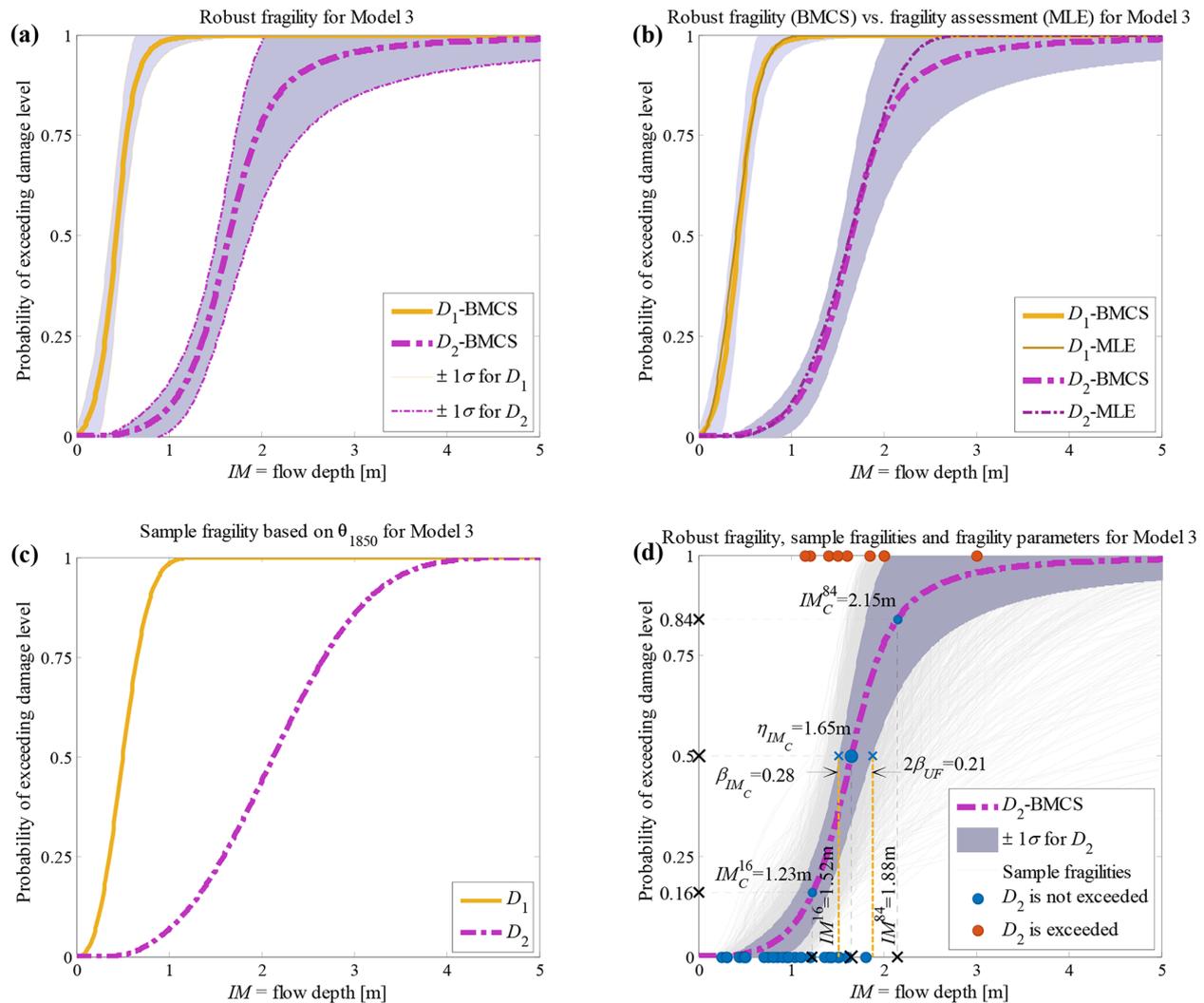


Figure 5. Building class 2 (non-engineered masonry, unreinforced with clay brick, 2 storey) of the Sulawesi–Palu 2018 tsunami considering fragility model class \mathbb{M}_3 (a) hierarchical robust fragility curves and their ± 1 standard deviation confidence intervals; (b) comparison between robust fragility and its confidence band (based on BMCS method) and fragility assessment based on MLE method; (c) the fragility curves $P(D \geq D_j | IM, \theta_{3,1850})$, where $1 \leq j \leq 2$ associated with the 1850th realisation of the model parameters, $\theta_{3,1850}$ ($k = 3$ associated to model \mathbb{M}_3 , $l = 1850$); and (d) robust fragility curve associated with the damage threshold D_2 , together with all the sample fragilities and the equivalent lognormal fragility parameters.

the proposed workflow avoids crossing fragility curves (e.g. Class 2, timber residential, South Pacific tsunami; Class 3, non-engineered light timber, Sulawesi–Palu tsunami). The results illustrated for the five building classes demonstrate that the proposed workflow for hierarchical fragility assessment can be applied in cases in which data points are not available for all the damage levels within the damage scale.

5 Conclusion

An integrated procedure based on Bayesian model class selection (BMCS) for empirical hierarchical fragility modelling for a class of buildings or infrastructure is presented.

This procedure is applicable to fragility modelling for any type of hazard as long as the damage scale consists of mutually exclusive and collectively exhaustive (MECE) damage states, and the observed damage data points are independent. This simulation-based procedure can (1) perform hierarchical fragility modelling for MECE damage states, (2) estimate the confidence interval for the resulting fragility curves, and (3) select the simplest model that fits the data best (i.e. maximises log evidence) amongst a suite of candidate fragility models (herein, alternative link functions for generalised linear regression are considered). The proposed procedure is demonstrated for empirical fragility assessment based on observed damage data to masonry residential (1 storey) and

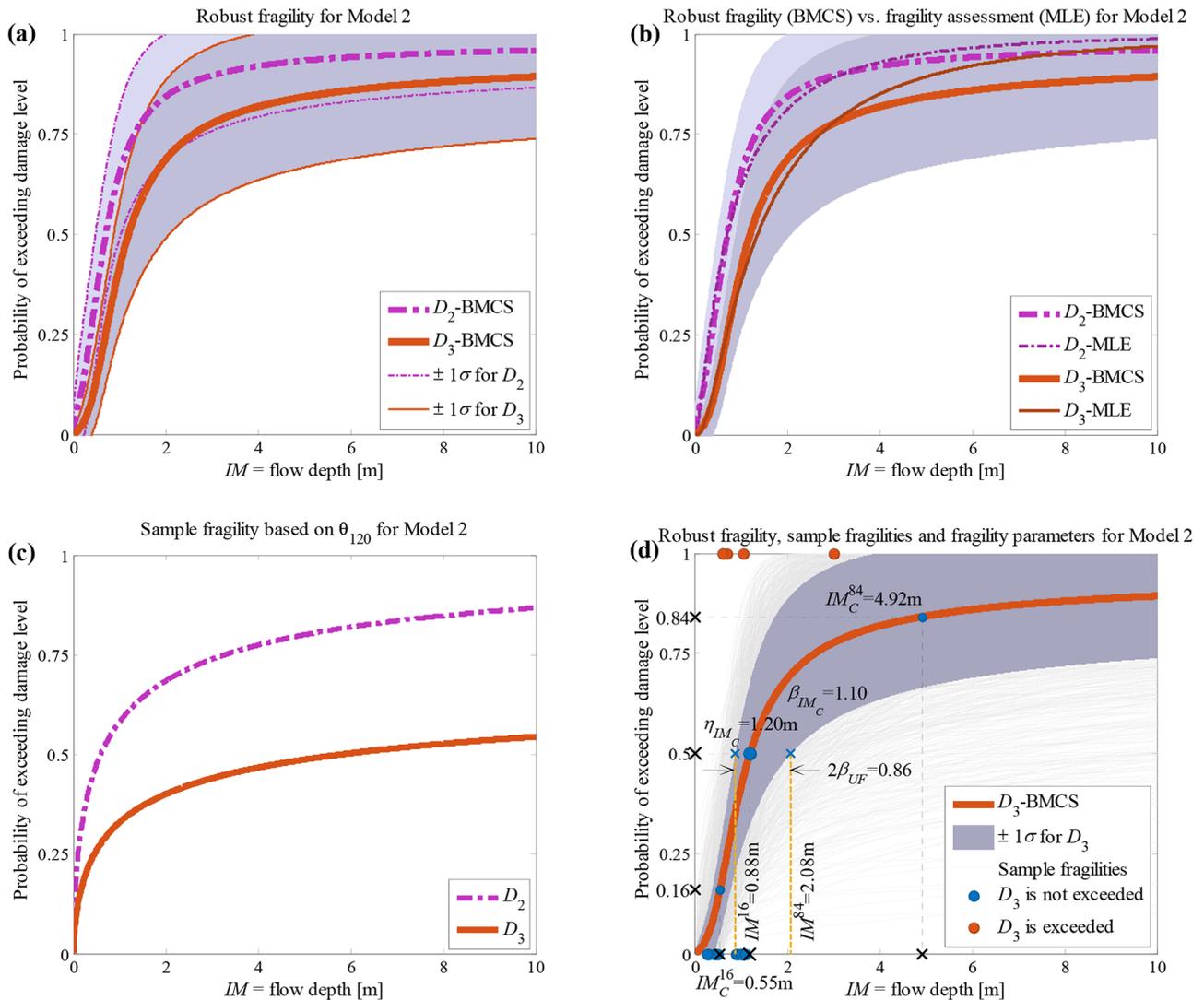


Figure 6. Building class 3 (non-engineered light timber) of the Sulawesi–Palu 2018 tsunami considering fragility model class M_2 (a) hierarchical robust fragility curves and their ± 1 standard deviation confidence intervals; (b) comparison between robust fragility and its confidence band (based on BMCS method) and fragility assessment based on MLE method; (c) the fragility curves $P(D \geq D_j | IM, \theta_{2,120})$, where $2 \leq j \leq 3$ associated with the 120th realisation of the model parameters, $\theta_{2,120}$ ($k = 2$ associated to model M_2 , $l = 120$); and (d) robust fragility curve associated with the damage threshold D_3 , together with all the sample fragilities and the equivalent lognormal fragility parameters.

timber residential buildings due to the 2009 South Pacific tsunami in American Samoa and the Samoan Islands and non-engineered masonry buildings (1 and 2 storeys) and non-engineered light timber buildings due to the 2018 Sulawesi–Palu tsunami. It is observed that

- For each model class, the same set of simulation realisations is used to estimate the fragility parameters, the confidence band and the log evidence. The latter, which consists of two terms depicting the goodness of fit and the information gain between posterior distribution resulting from the observed data and the prior distribu-

tion, is used to compare the candidate fragility models to identify the model that maximises the evidence.

- Hierarchical fragility assessment can be done also based the maximum likelihood estimation (MLE) and the available statistical toolboxes (e.g. MATLAB’s generalised linear model). For each damage level, the reference domain should be the subset of data that exceeds the consecutive lower damage level, instead of taking the entire set of data points as reference domain. Note that the basic fragility estimation (“MLE-Basic”, non-hierarchical fragility model) fits the damage data for

Table 5. The equivalent lognormal parameters and the epistemic uncertainty in the RF assessment for all the building classes, damage thresholds, and model classes M_1 to M_3 .

	Building class	Damage level	Model 1 (M_1)			Model 2 (M_2)			Model 3 (M_3)		
			η_{IMC} (m)	β_{IMC}	β_{UF}	η_{IMC} (m)	β_{IMC}	β_{UF}	η_{IMC} (m)	β_{IMC}	β_{UF}
South Pacific tsunami 2009	1	D_1	0.29	0.40	0.20	0.30	0.46	0.22	0.33	0.51	0.22
		D_2	0.43	0.35	0.15	0.46	0.37	0.15	0.50	0.40	0.15
		D_3	1.28	0.35	0.08	1.28	0.35	0.07	1.37	0.37	0.07
		D_4	1.80	0.45	0.07	1.81	0.42	0.07	1.89	0.37	0.06
		D_5	2.49	0.47	0.07	2.48	0.47	0.07	2.50	0.35	0.06
	2	D_3	0.64	1.08	0.58	0.63	1.20	0.64	0.63	1.26	0.53
D_4		0.73	1.01	0.48	0.75	1.01	0.46	0.84	0.98	0.35	
D_5		1.52	0.25	0.08	1.54	0.26	0.08	1.61	0.23	0.07	
Sulawesi–Palu tsunami 2018	1	D_1	0.25	0.39	0.15	0.26	0.43	0.15	0.27	0.56	0.18
		D_2	1.24	0.57	0.06	1.24	0.59	0.07	1.31	0.58	0.06
	2	D_1	0.39	0.44	0.20	0.38	0.45	0.19	0.43	0.43	0.17
		D_2	1.60	0.31	0.11	1.59	0.32	0.12	1.65	0.28	0.11
	3	D_2	0.74	1.01	0.34	0.71	1.03	0.38	0.81	0.85	0.29
		D_3	1.16	0.98	0.39	1.20	1.10	0.43	1.29	0.95	0.31

Table 6. Bayesian model class selection results for empirical fragility models. The best model for each building class is shown in bold.

	Building class	Model class	Term 1: Average data fit	Term 2: Information gain	Log-evidence	Posterior probability of the model
South Pacific tsunami 2009	1	M_1	-124.4561	23.6853	-148.1414	0.055
		M_2	-123.4659	23.9566	-147.4224	0.113
		M_3	-120.6454	24.7810	-145.4264	0.832
	2	M_1	-20.4791	9.7549	-30.2340	0.315
		M_2	-19.9106	10.4660	-30.3766	0.273
		M_3	-19.7565	10.2117	-29.9682	0.411
Sulawesi–Palu tsunami 2018	1	M_1	-161.9565	9.5690	-171.5255	0.445
		M_2	-161.2320	10.5660	-171.7979	0.339
		M_3	-161.8821	10.3673	-172.2494	0.216
	2	M_1	-23.3696	6.8292	-30.1987	0.213
		M_2	-22.7429	7.2697	-30.0126	0.257
		M_3	-22.4307	6.8551	-29.2858	0.531
3	M_1	-15.8034	4.2741	-20.0775	0.210	
	M_2	-15.1575	3.9226	-19.0802	0.570	
	M_3	-14.6294	5.4015	-20.0309	0.220	

each damage level at a time. In other words, the reference domain is set to all damage data.

needs to have data available for at least two damage levels.

- The procedure is also applicable to cases in which observed data are available only for a subset of the damage levels within the damage scale. The number of fragility curves is going to be equal to the total number of damage levels for which data are available minus one. This means, in order to have at least one fragility curve, one

- Although the resulting fragility curves are not lognormal (strictly speaking), equivalent statistics work quite well in showing the fragility curves (median and logarithmic dispersion) and the corresponding epistemic uncertainty (logarithmic dispersion).

Table 7. The Model parameters $\theta_{MLE-Basic}$.

	Building class	Model class	$D \geq D_1$		$D \geq D_2$		$D \geq D_3$		$D \geq D_4$		$D \geq D_5$	
			α_0	α_1								
South Pacific tsunami 2009	1	M_1	5.242	4.190	3.655	4.556	-1.221	4.884	-2.666	4.213	-4.271	4.652
		M_2	2.742	2.190	1.946	2.486	-0.695	2.846	-1.506	2.425	-2.293	2.515
		M_3	2.079	2.011	1.347	2.361	-1.319	3.139	-2.389	3.009	-3.919	3.806
	2	M_1					1.127	1.512	0.813	1.392	-2.931	6.055
		M_2					0.657	0.909	0.472	0.873	-1.747	3.597
		M_3					0.251	0.862	0.051	0.914	-2.416	3.922
South Pacific tsunami 2009	1	M_1	6.059	4.355	-0.638	2.948						
		M_2	3.264	2.340	-0.375	1.728						
		M_3	2.498	2.088	-0.918	2.088						
	2	M_1	3.556	3.672	-2.500	5.160						
		M_2	2.077	2.144	-1.465	3.039						
		M_3	1.664	2.239	-2.473	4.268						
	3	M_1			0.466	1.375	-0.501	1.843				
		M_2			0.295	0.847	-0.286	1.141				
		M_3			-0.041	1.068	-0.787	1.565				

Table 8. Comparison between fragility assessment based on MLE method (by hierarchical fragility modelling) and the MLE-Basic method for damage thresholds $D_{index(2)}$ to $D_{index(N)}$.

Tsunami	Class	Damage level	Model 1 (M_1)				Model 2 (M_2)				Model 3 (M_3)			
			MLE method		MLE-Basic method		MLE method		MLE-Basic method		MLE method		MLE-Basic method	
			η_{IM_C} (m)	β_{IM_C}										
South Pacific 2009	1	D_1	0.29	0.40	0.29	0.40	0.29	0.46	0.29	0.46	0.30	0.59	0.30	0.59
		D_2	0.43	0.35	0.45	0.37	0.45	0.38	0.46	0.40	0.47	0.44	0.48	0.50
		D_3	1.28	0.34	1.28	0.34	1.27	0.35	1.28	0.35	1.34	0.38	1.35	0.38
		D_4	1.82	0.43	1.88	0.40	1.82	0.42	1.86	0.41	1.88	0.38	1.96	0.39
		D_5	2.50	0.46	2.50	0.36	2.47	0.44	2.49	0.40	2.49	0.34	2.54	0.31
South Pacific 2009	2	D_3	0.47	1.10	0.47	1.10	0.49	1.10	0.49	1.10	0.49	1.37	0.49	1.37
		D_4	0.56	1.12	0.56	1.20	0.59	1.11	0.58	1.15	0.62	1.25	0.63	1.29
		D_5	1.54	0.30	1.62	0.28	1.55	0.30	1.63	0.28	1.58	0.28	1.69	0.30
Sulawesi-Palu 2018	1	D_1	0.25	0.38	0.25	0.38	0.25	0.43	0.25	0.43	0.25	0.57	0.25	0.57
		D_2	1.24	0.57	1.24	0.57	1.24	0.58	1.24	0.58	1.30	0.57	1.30	0.57
	2	D_1	0.38	0.45	0.38	0.45	0.38	0.47	0.38	0.47	0.40	0.53	0.40	0.53
		D_2	1.63	0.32	1.62	0.32	1.62	0.33	1.62	0.33	1.64	0.28	1.64	0.28
	3	D_2	0.71	1.21	0.71	1.21	0.71	1.18	0.71	1.18	0.74	1.11	0.74	1.11
		D_3	1.38	1.10	1.31	0.91	1.36	1.02	1.29	0.88	1.32	0.84	1.31	0.76

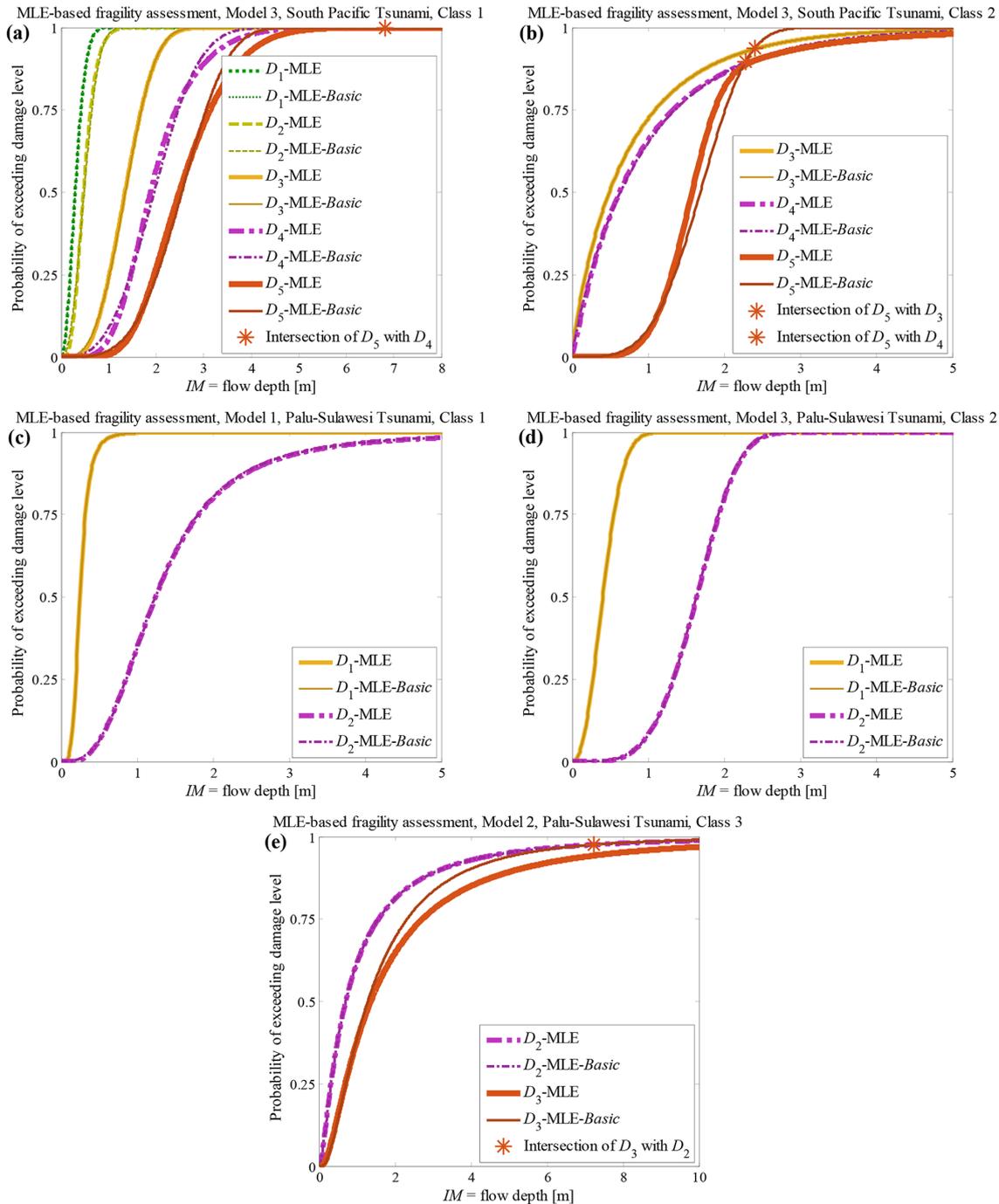


Figure 7. Comparison between the fragility assessment by MLE-based hierarchical fragility modelling and MLE-Basic fragility assessment given (a) South Pacific 2009 tsunami, Class 1, M_3 ; (b) South Pacific 2009 tsunami, Class 2, M_3 ; (c) Sulawesi–Palu 2018 tsunami, Class 1, M_1 ; (d) Sulawesi–Palu 2018 tsunami Class 2, M_3 ; and (e) Sulawesi–Palu 2018 tsunami Class 3, M_2 .

– The proposed BMCS method and the one based on MLE lead to essentially the same set of parameter estimates for hierarchical fragility estimation. However, the latter does not readily lead to the confidence band and log evidence.

– Using the basic method for fragility estimation (MLE-Basic) leads to results that are slightly different from the hierarchical fragility curves. The difference grows for higher damage levels. It is important to note that following the basic method MLE-Basic led to ill-conditioned results (i.e. fragility curves crossing) in some of the

cases (Class 2 for the South Pacific tsunami and Class 3 for the Sulawesi–Palu tsunami, both Timber constructions) studied in this work.

The major improvement offered by this method is in providing a tool that can fit fragility curves to a set of hierarchical levels of damage or loss in an ensemble manner. This method, which starts from prescribed fragility models and explicitly ensures the hierarchical relation between the damage levels, is very robust in cases where few data points are available and/or where data are missing for some of the damage levels. This tool provides confidence bands for the fragility curves and performs model selection among a set of viable link functions for generalised regression. It is important to note that the proposed method is in general applicable to hierarchical vulnerability modelling for human or economic loss levels and to different types of hazards if (1) the defined levels are mutually exclusive and collectively exhaustive, and (2) a suitable intensity measure (IM) can be identified.

Appendix A: The derivation of Eq. (2)

The probability of being in damage state DS_j for a given intensity measure IM can be estimated as follows:

$$\begin{aligned}
 P(DS_j | IM) &= P[(D \geq D_j) \cdot (D < D_{j+1}) | IM] \\
 &= 1 - P[\overline{(D \geq D_j)} \cdot \overline{(D < D_{j+1})} | IM] \\
 &= 1 - P[\overline{(D \geq D_j)} + \overline{(D < D_{j+1})} | IM] \\
 &= 1 - \underbrace{P[(D < D_j) + (D \geq D_{j+1}) | IM]}_{ME \triangleq P(D < D_j | IM) + P(D \geq D_{j+1} | IM)} \\
 &= 1 - P(D < D_j | IM) - P(D \geq D_{j+1} | IM) \\
 &= P(D \geq D_j | IM) - P(D \geq D_{j+1} | IM) \quad , \quad (A1)
 \end{aligned}$$

where the upper-bar sign stands for the logical negation and is read as “NOT”, and (+) defines the logical sum and is read as “OR”. The above derivation is based on the *rule of sum* in probability and considers the fact that the two statements $D < D_j$ and $D \geq D_{j+1}$ are mutually exclusive (ME); thus, the probability of their logical sum is the sum of their probabilities.

Appendix B: The derivation of Eq. (7)

The probability of being in damage state DS_j (where $j \geq 1$) given the intensity measure evaluated at the location of building i , denoted as IM_i , based on Eq. (6) can be expanded in a

recursive format as follows:

$$\begin{aligned}
 P(DS_j | IM_i) &= \underbrace{[1 - P(D \geq D_{j+1} | D \geq D_j, IM_i)]}_{1 - \pi_{ij}} \\
 &\cdot [1 - P(D < D_j | IM_i)] \\
 &= (1 - \pi_{ij}) \cdot [1 - P((D < D_j) \cdot (D \geq D_{j-1}) + \dots \\
 &\quad + (D < D_1) \cdot (D \geq D_0) | IM_i)] \\
 &= (1 - \pi_{ij}) \cdot \left[1 - \sum_{k=0}^{j-1} P((D < D_{k+1}) \cdot (D \geq D_k) | IM_i) \right] \\
 &= (1 - \pi_{ij}) \cdot \left[1 - \sum_{k=0}^{j-1} P(DS_k | IM) \right] \quad , \quad (B1)
 \end{aligned}$$

where (+) defines the logical sum and is read as “OR”. The above derivation is based on the rule of sum in probability and considers the fact that the recursive statements in the second term expressed generally as $(D < D_{k+1}) \cdot (D \geq D_k)$, where $0 \leq k \leq j - 1$, are ME; hence, the probability of their logical sum is the sum of their probabilities. It is important to note that in case where $j = 0$, the above equation can be written as

$$P(DS_0 | IM_i) = (1 - \pi_{i0}) \triangleq P(D < D_1 | IM_i). \quad (B2)$$

Appendix C: The derivation of log-evidence in Eq. (13)

From an information-based point of view, the logarithm of the evidence (log-evidence), denoted as $\ln[p(\mathbf{D} | \mathbb{M}_k)]$, can provide a quantitative measure of the amount of information as evidence of model \mathbb{M}_k . Moreover, the posterior PDF $p(\boldsymbol{\theta}_k | \mathbf{D}, \mathbb{M}_k)$ (see Eq. 14) over the domain of the model parameters $\boldsymbol{\Omega}_\theta$ given the k th model is equal to unity. Thus, $\ln[p(\mathbf{D} | \mathbb{M}_k)]$ can be written as follows:

$$\begin{aligned}
 \ln[p(\mathbf{D} | \mathbb{M}_k)] &= \ln[p(\mathbf{D} | \mathbb{M}_k)] \\
 &\cdot \underbrace{\int_{\boldsymbol{\Omega}_\theta} p(\boldsymbol{\theta}_k | \mathbf{D}, \mathbb{M}_k) d\boldsymbol{\theta}_k}_{=1.0} \quad (C1)
 \end{aligned}$$

Since the log-evidence is independent of $\boldsymbol{\theta}$, we can bring it inside the integral and do some simple manipulation (also

using the relation in Eq. 11) as follows:

$$\begin{aligned}
 \ln [p(\mathbf{D}|\mathbb{M}_k)] &= \int_{\Omega_{\theta_k}} \ln [p(\mathbf{D}|\mathbb{M}_k)] \cdot p(\theta_k|\mathbf{D}, \mathbb{M}_k) d\theta_k \\
 &= \int_{\Omega_{\theta}} \ln \left[\frac{p(\mathbf{D}|\theta_k, \mathbb{M}_k) p(\theta_k|\mathbb{M}_k)}{p(\theta_k|\mathbf{D}, \mathbb{M}_k)} \right] \cdot p(\theta_k|\mathbf{D}, \mathbb{M}_k) d\theta_k \\
 &= \int_{\Omega_{\theta}} \ln \left[\frac{p(\mathbf{D}|\theta_k, \mathbb{M}_k)}{p(\theta_k|\mathbf{D}, \mathbb{M}_k) / p(\theta_k|\mathbb{M}_k)} \right] \cdot p(\theta_k|\mathbf{D}, \mathbb{M}_k) d\theta_k \\
 &= \underbrace{\int_{\Omega_{\theta}} \ln [p(\mathbf{D}|\theta_k, \mathbb{M}_k)] \cdot p(\theta_k|\mathbf{D}, \mathbb{M}_k) d\theta_k}_{\text{Term 1}} \\
 &\quad - \underbrace{\int_{\Omega_{\theta}} \ln \left[\frac{p(\theta_k|\mathbf{D}, \mathbb{M}_k)}{p(\theta_k|\mathbb{M}_k)} \right] \cdot p(\theta_k|\mathbf{D}, \mathbb{M}_k) d\theta_k}_{\text{Term 2}}. \tag{C2}
 \end{aligned}$$

Appendix D: Multivariate normal distribution and generating dependent Gaussian variables

Let us assume that the vector of parameters for the k th model is set to θ ; i.e. $\theta = \theta_k$. A multivariate normal PDF can be expressed as follows:

$$p(\theta) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{S}|}} \exp \left(-\frac{1}{2} (\theta - \mu_{\theta})^T \mathbf{S}^{-1} (\theta - \mu_{\theta}) \right), \tag{D1}$$

where n is the number of components (uncertain parameters) of vector $\theta = \{\theta_i, i = 1 : n\}$; μ_{θ} is the vector of the mean value of θ ; and \mathbf{S} is the covariance matrix. The positive definite matrix $\mathbf{S}_{n \times n}$ can be factorised based on Cholesky decomposition as $\mathbf{S} = \mathbf{L}\mathbf{L}^T$, where $\mathbf{L}_{n \times n}$ is a lower triangular matrix (i.e. for all $j > i$, $L_{ij} = 0$ where L_{ij} denotes the (i, j) -entry of the matrix \mathbf{L}). A Gaussian vector $\theta_{n \times 1}$ with mean μ_{θ} and covariance \mathbf{S} can be generated as follows:

$$\theta = \mu_{\theta} + \mathbf{L}\mathbf{Z}, \tag{D2}$$

where $\mathbf{Z}_{n \times 1}$ is a vector of standard Gaussian *i.i.d.* random variables with zero mean $\mathbf{0}_{n \times n}$ and a covariance equal to the identity matrix $\mathbf{I}_{n \times n}$. To verify the properties of θ , we know that with reference to Eq. (D2), it should have a mean equal to μ_{θ} and a covariance matrix equal to \mathbf{S} . The expectation of θ , denoted as $E(\theta)$, can be estimated as

$$E(\theta) = E(\mu_{\theta} + \mathbf{L}\mathbf{Z}) = E(\mu_{\theta}) + \underbrace{\mathbf{L}E(\mathbf{Z})}_{=\mathbf{0}_{n \times 1}} = \mu_{\theta}. \tag{D3}$$

The covariance matrix of θ can be written as

$$\begin{aligned}
 E[(\theta - \mu_{\theta})(\theta - \mu_{\theta})^T] &= E(\mathbf{L}\mathbf{Z}\mathbf{Z}^T\mathbf{L}^T) \\
 &= \underbrace{\mathbf{L}E(\mathbf{Z}\mathbf{Z}^T)}_{=\mathbf{I}_{n \times n}}\mathbf{L}^T = \mathbf{L}\mathbf{L}^T = \mathbf{S}. \tag{D4}
 \end{aligned}$$

Thus, the vector θ can be written according to Eq. (D2).

Appendix E: Adaptive MCMC scheme

E1 MCMC procedure

The MCMC simulation scheme has a Markovian nature where the transition from the current state to a new state is done by using a conditional transition function that is conditioned on the current (last) state. Let us assume that the vector of parameters for the k th model is set to θ ; i.e. $\theta = \theta_k$. To generate $(i + 1)$ th sample θ_{i+1} from the current i th sample θ_i based on MH routine, the following procedure is adopted herein

- Simulate a *candidate* sample θ^* from a *proposal* distribution $q(\theta|\theta_i)$. It is important to note that there are no specific restrictions about the choice of $q(\cdot)$ apart from the fact that it should be possible to calculate both $q(\theta_{i+1}|\theta_i)$ and $q(\theta_i|\theta_{i+1})$.
- Calculate the acceptance probability $\min(1, r)$, where r is defined as follows (it is to note that the following Eq. (E1) is written in the general format for brevity compared to Eq. (14) of the paper, and we have used θ instead of θ_k and dropped the conditioning on \mathbb{M}_k ; hence, when we write the i th sample θ_i it is actually the i th sample drawn from θ_k and “ k ” is dropped for brevity):

$$\begin{aligned}
 r &= \frac{p(\theta^*|\mathbf{D})}{p(\theta_i|\mathbf{D})} \cdot \frac{q(\theta_i|\theta^*)}{q(\theta^*|\theta_i)} \\
 &= \left(\underbrace{\frac{p(\mathbf{D}|\theta^*)}{p(\mathbf{D}|\theta_i)}}_{\text{likelihood ratio}} \cdot \underbrace{\frac{p(\theta^*)}{p(\theta_i)}}_{\text{prior ratio}} \right) \cdot \underbrace{\frac{q(\theta_i|\theta^*)}{q(\theta^*|\theta_i)}}_{\text{proposal ratio}}. \tag{E1}
 \end{aligned}$$

- Generate u from a uniform distribution between $(0, 1)$, $u \sim \text{uniform}(0, 1)$.
- If $u \leq \min(1, r) \rightarrow$ set $\theta_{i+1} = \theta^*$ (*accept the candidate* state to be taken as the *next* state of the Markov chain); else set $\theta_{i+1} = \theta_i$ (the *current* state is taken as the *next* state).

Estimating the likelihood in the arithmetic scale based on Eq. (E1) may encounter instability as $p(\mathbf{D}|\theta)$ may become very small; thus, the likelihood ratio becomes indeterminate. To avoid this numerical instability, it is desirable to substitute the likelihood ratio in Eq. (E1) with $\exp(\ln(p(\mathbf{D}|\theta^*)) - \ln(p(\mathbf{D}|\theta_i)))$ if the ratio becomes indeterminate or zero.

With reference to Eq. (E1), samples from the posterior can be drawn based on MH algorithm without any need to define the normalising C^{-1} coefficient according to Eq. (14). Equation (E1) always accepts a candidate if the new proposal is

more likely under the target posterior distribution than the old state. Therefore, the sampler will move towards the regions of the state space where the target posterior function has high density.

The choice of the proposal distribution q is very important. The ratio $q(\theta_i|\theta^*)/q(\theta^*|\theta_i)$ corrects for any asymmetries in the proposal distribution. Intuitively, if $q(\theta^*|\theta_i) = p(\theta^*|\mathbf{D})$, the candidate state is always accepted (with $r = 1$); thus, the closer q is to the target posterior PDF, the better the acceptance rate and the faster the convergence. This is not a trivial task as information about the important region $p(\theta|\mathbf{D})$ is not available. If the proposal distribution q is *non-adaptive*, it means that the information of the current sample θ_i is not used to explore the important region of the target posterior distribution $p(\theta|\mathbf{D})$; thus, we can say that $q(\theta^*|\theta_i) = q(\theta^*)$. Therefore, it is more desirable to choose an *adaptive* proposal distribution which depends on the current sample (Beck and Au, 2002). Having the proposal PDF q centred around the current sample renders the MH algorithm like a local random walk that adaptively leads to the generation of the target PDF. However, if the Markov chain starts from a point that is not close to the region of the significant probability density of $p(\theta|\mathbf{D})$, the chance of generating a candidate state θ^* will become extremely small (and we will face high rejection of candidate samples). Therefore, most of the samples will be repeated. To solve this problem, Beck and Au (2002) introduce a sequence of PDFs that bridge the gap between the prior PDF and the target posterior PDF. This issue will be further explored hereafter under the adaptive MCMC. Finally, it can mathematically be shown that (see Beck and Au, 2002) if the current sample θ_i is distributed as $p(\cdot|\mathbf{D})$, the next sample θ_{i+1} is also distributed as $p(\cdot|\mathbf{D})$.

E2 Adaptive Metropolis–Hastings algorithm (adaptive MCMC)

The adaptive MH algorithm (Beck and Au, 2002) introduces a sequence of intermediate candidate evolutionary PDFs that resemble more and more the target PDF. Let $\{p_1, p_2, \dots, p_{N_{\text{chain}}}\}$ be the sequence (chain) of PDFs leading to $p(\theta|\mathbf{D}) = p_{N_{\text{chain}}}$, where N_{chain} is the number of chains and each chain contains N_d samples (as indicated subsequently). The following adaptive simulation-based procedure is employed:

Step 1. Simulate N_d samples $\{\theta_1, \theta_2, \dots, \theta_{N_d}\}^{(1)}$, where the superscript (1) denotes the first simulation level or the first chain (nc = 1, where nc denotes the chain number/simulation level), with the target PDF p_1 as the first sequence of samples. Instead of accepting or rejecting a proposal for θ involving all its components simultaneously (called *block-wise* updating scheme), it might be computationally simpler and more efficient for the first chain to make proposals for individual components of θ one at a time (called *component-wise* updating approach). In the *block-wise* updating, the proposal distribution has the same dimension as the target distribu-

tion. For instance, if the model parameters involve n uncertain parameters (e.g. the vector of model parameters $\theta_{n \times 1}$ in this paper has $n = 2(N - 1)$ variables for each of the three models $\mathbb{M}_1, \mathbb{M}_2$, and \mathbb{M}_3), we design an n -dimensional proposal distribution and either accept or reject the candidate state (with all n variables) as a block. The block-wise updating approach can be associated with high rejection rates. This may cause problem when we want to generate the first sequence of samples (first chain). Therefore, we have utilised the more stable component-wise updating for the first chain. We start from the first variable and generate a candidate state based on a proposal distribution for this individual component and finally accept or reject it based on MH algorithm. Note that in this stage we have varied the current component and kept the other variables in vector θ constant. Then, we move to the next components one by one and do the same procedure while taking into account the previous (updated) components. Therefore, what happens in the current step is conditional on the updated parameters in the previous steps.

Step 2. Construct a kernel density function $\kappa^{(1)}$ as the weighted sum (average) of n -dimensional Gaussian PDFs centred among the samples $\{\theta_1, \theta_2, \dots, \theta_{N_d}\}^{(1)}$, with the covariance matrix $\mathbf{S}^{(1)}$ of the samples $\theta_i^{(1)}$ and the weights associated to each sample as w_i , where $i = 1 : N_d$ as follows (see Ang et al., 1992; Au and Beck, 1999):

$$\kappa^{(1)}(\theta) = \frac{1}{N_{\text{seed}}} \sum_{i=1}^{N_{\text{seed}}} \frac{1}{w_i^n \sqrt{(2\pi)^n |\mathbf{S}^{(1)}|}} \cdot \exp\left(-\frac{1}{2w_i^2} (\theta - \theta_i^{(1)})^T (\mathbf{S}^{(1)})^{-1} (\theta - \theta_i^{(1)})\right). \quad (\text{E2})$$

The kernel density $\kappa^{(1)}$ constructed in Eq. (E2) approximates p_1 . The kernel function κ can be viewed as a PDF consisting of bumps at θ_i , where width w_i controls the common size of the bumps. Therefore, a large value of w_i tends to over-smooth the kernel density, while a small value may cause noise-shaped bumps. In view of this, w_i can be assumed to have a fixed width ($= w$), or alternatively the *adaptive kernel* estimate can be employed (Ang et al., 1992; Au and Beck, 1999) that is defined for each sample θ_i , $i = 1 : N_d$. The adaptive kernel has better convergence and smoothing properties over the fixed-width kernel estimate. The fixed width w is estimated as follows (Epanechnikov, 1969):

$$w = \left(\frac{4}{(n + 2)N_{\text{dist}}}\right)^{\frac{1}{n+4}}, \quad (\text{E3})$$

where N_{dist} is the number of samples ($N_{\text{dist}} \leq N_d$). For one-dimensional problems ($n = 1$), this leads to the well-known fixed-width value of $[(4/3)N_{\text{seed}}]^{1/5}$. In the adaptive kernel method, the idea is to vary the shape of each bump so that a larger width (flatter bump) is used in regions of lower probability density. Following the general strategy used in the past

(see Ang et al., 1992; Au and Beck, 1999), the adaptive band width w_i for the i th sample θ_i can be written as $w_i = w\lambda_i$, where the local bandwidth factor λ_i can be estimated as follows:

$$\lambda_i = \left(\kappa(\theta_i) / \left(\prod_{j=1}^{N_{\text{seed}}} \kappa(\theta_j) \right)^{\frac{1}{N_{\text{seed}}}} \right)^{-\omega}, \quad (\text{E4})$$

where $0 \leq \omega \leq 1.0$ is the sensitivity factor, and $\kappa(\theta_i)$ is calculated based on Eq. (E2) where $\theta = \theta_i$, with the choice of fixed-width w in Eq. (E3). The denominator in Eq. (E4) is a geometric mean of the kernel estimator at all N_d points. The value of $\omega = 0.50$ is employed herein as also suggested by other research endeavours (Abramson, 1982; Ang et al., 1992; Au and Beck, 1999). It is numerically more stable to estimate the denominator in Eq. (E4) as $\prod_{j=1}^{N_{\text{seed}}} [\kappa(\theta_j)^{1/N_{\text{seed}}}]$.

Step 3. Simulate N_d Markov chain samples $\{\theta_1, \theta_2, \dots, \theta_{N_d}\}^{(2)}$ with the target PDF p_2 as the second simulation level ($nc = 2$). We use $\kappa^{(1)}$ as the proposal distribution $q(\cdot)$ in Eq. (E1) in this stage to generate the second chain of samples. To generally simulate sample θ from the kernel $\kappa^{(nc)}$ (where $nc = 1 : N_{\text{chain}}$), we generate a discrete random index from the vector $[1, 2, \dots, N_d]$ with the corresponding weights $[w_1, w_2, \dots, w_{N_{\text{seed}}}]$ using an inverse transformation sampling; if index = j , then generate θ from the Gaussian PDF κ_j , where

$$\begin{aligned} \kappa_j(\theta) &= \frac{1}{(w\lambda_j)^n \sqrt{(2\pi)^n |\mathbf{S}^{(nc)}|}} \\ &\cdot \exp\left(-\frac{1}{2(w\lambda_j)^2} (\theta - \theta_j)^T (\mathbf{S}^{(nc)})^{-1} (\theta - \theta_j)\right) \\ &= \frac{1}{\sqrt{(2\pi)^n |\mathbf{S}_j^{(nc)}|}} \\ &\cdot \exp\left(-\frac{1}{2} (\theta - \theta_j)^T (\mathbf{S}_j^{(nc)})^{-1} (\theta - \theta_j)\right), \quad (\text{E5}) \end{aligned}$$

where $\mathbf{S}_j^{(nc)} = w_j^2 \mathbf{S}^{(nc)}$, where $\mathbf{S}^{(nc)}$ is the covariance matrix of the samples $\{\theta_1, \theta_2, \dots, \theta_{N_d}\}^{(nc)}$. Appendix D shows how a sample θ can be drawn from the Gaussian PDF κ_j . From this sequence on, the MCMC updating is done in a block-wise manner as we generate a candidate θ and accept/reject it as a block. The second chain of samples $\{\theta_1, \theta_2, \dots, \theta_{N_d}\}^{(2)}$ are then used to construct the kernel density $\kappa^{(2)}$ based on Eq. (E2).

Step 4. In general, $\kappa^{(nc)}$ is used as the proposal distribution in order to move from the nc th simulation level (which approximates p_{nc}) into $(nc + 1)$ th chain (with target PDF p_{nc+1}). This will continue until the N_{chain} th simulation level where Markov chain samples are simulated for the target updated $p(\theta|\mathbf{D}) = p_{N_{\text{chain}}}$.

Appendix F: MCMC samples for each model

The adaptive MCMC procedure for drawing samples from the model parameters from the joint posterior PDF $p(\theta_k|\mathbf{D}\mathbb{M}_k)$ is carried out by considering $N_{\text{chain}} = 6$ chains (simulation levels) and $N_d = 2000$ samples per each chain (see Appendix E). In the first simulation level (first chain, $nc = 1$), for which a component-wise updating approach is employed (see Appendix E, Step 1 for the description of component-wise and block-wise updating), the first 20 samples are not considered in order to reduce the initial transient effect of the Markov chain. The proposal distribution (see Eq. E1) for each component is assumed to be a normal distribution with a coefficient of variation $\text{COV} = 0.30$ herein. In addition, the prior ratio according to Eq. (E1) will become the ratio of two normal distributions for each component one at a time. In the next simulation levels (i.e. $nc = 2 : 6$), the adaptive kernel estimate (Eq. E2) is employed; i.e. the MCMC updating is performed in a block-wise manner. There will be N_d Markov chain samples generated within the each chain, denoted as $\{\theta_{k,1}, \theta_{k,2}, \dots, \theta_{k,N_d}\}^{(nc)}$, where $nc = 2 : N_{\text{chain}} (= 6)$. The N_d samples of the last chain ($nc = 6$) will be used as the fragility model parameters, as discussed in Sect. 3.6. It is important to note that the likelihood $p(\mathbf{D}|\theta_k, \mathbb{M}_k)$ (used in calculating the acceptance probability within the MCMC procedure in Eq. E1) is estimated according to Eq. (13).

Figure F1 illustrates the histograms representing the drawn samples from the joint posterior PDFs corresponding to the sampled model parameters $\{\theta_{k,1}, \theta_{k,2}, \dots, \theta_{k,N_d}\}^{(6)}$ corresponding to the brick masonry residential, Class 1 South Pacific tsunami classes \mathbb{M}_k ($k \in \{1, 2, 3\}$) shown in Fig. 2. The marginal normal prior PDFs are also shown with dashed orange-coloured lines. The statistics of the samples, mean and confidence interval (CI) between 2nd and 98th percentiles for the posterior of model parameters θ_k are shown on the figures associated to each parameter. It is expected to have the mean values of the marginal posterior samples close to and comparable with those obtained by the MLE in Table 4.

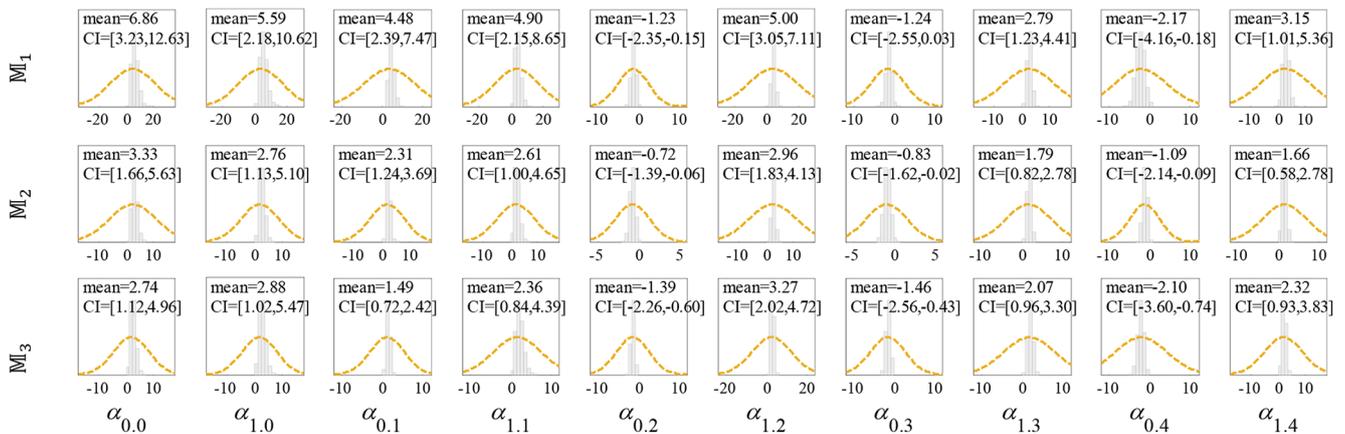


Figure F1. Distribution of the fragility model parameters $\theta_k = [\{\alpha_{0,j}, \alpha_{1,j}\}_k, j = 0 : 4]$ based on model class \mathbb{M}_k (where $k = 1 : 3$) for Class 1, brick masonry residential, South Pacific tsunami, by employing an adaptive MCMC procedure including samples drawn from the joint posterior PDF with their statistics (mean and COV) and the marginal normal priors (subfigures show the posterior statistics).

Code and data availability. The code implementing the methodology and data used to produce the results in this article are available at the following URL: https://github.com/eurotsunamirisk/computeFrag/tree/main/Code_ver2 (last access: 30 November 2022; Ebrahimian and Jalayer, 2022).

Author contributions. FJ designed and coordinated this research. HE performed the simulations and developed the fragility functions. KT cured the availability of codes and software on the European Tsunami Risk Service (ETriS). BB provided precious insights on the damage data gathered for American Samoa and the Samoan Islands in the aftermath of the 2009 South Pacific tsunami (documented in Reese et al., 2011). All authors have contributed to the drafting of the paper. The first two authors contributed in an equal manner to the drafting of the paper.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Special issue statement. This article is part of the special issue “Tsunamis: from source processes to coastal hazard and warning”. It is not associated with a conference.

Acknowledgements. The authors would like to acknowledge partial support from Horizon Europe Project Geo-INQUIRE. Geo-INQUIRE is funded by the European Commission under project no. 101058518 within the HORIZON-INFRA-2021-SERV-01 call. The authors are grateful to the anonymous reviewer and Carmine Galasso for their insightful and constructive comments.

The authors gratefully acknowledge partial support by PRIN-2017 MATISSE (*Methodologies for the Assessment of anthropogenic environmental hazard: Induced Seismicity by Sub-surface geo-resources Exploitation*) project no. 20177EPPN2 funded by the Italian Ministry of Education and Research.

Financial support. This research has been partially supported by PRIN-2017 MATISSE (*Methodologies for the Assessment of anthropogenic environmental hazard: Induced Seismicity by Sub-surface geo-resources Exploitation*) project no. 20177EPPN2 funded by the Italian Ministry of Education and Research.

Review statement. This paper was edited by Animesh Gain and reviewed by Carmine Galasso and one anonymous referee.

References

Abramson, I. S.: On bandwidth variation in kernel estimates—a square root law, *Ann. Stat.*, 10, 1217–1223, 1982.
 Agresti, A.: *Categorical Data Analysis*, 3rd edn., Wiley, New Jersey, ISBN: 978-0-470-46363-5, 2012.
 Ang, G. L., Ang, A. H.-S., and Tang, W. H.: Optimal importance sampling density estimator, *J. Eng. Mech.*, 118, 1146–1163, 1992.
 Au, S. K. and Beck, J. L.: A new adaptive importance sampling scheme, *Struct. Saf.*, 21, 135–158, 1999.
 Beck, J. L. and Au, S. K.: Bayesian updating of structural models and reliability using Markov Chain Monte Carlo simulation, *J. Eng. Mech.*, 128, 380–391, 2002.

- Beck, J. L. and Yuen, K.-V.: Model selection using response measurements: Bayesian probabilistic approach, *J. Eng. Mech.*, 130, 192–203, 2004.
- Behrens, J., Løvholm, F., Jalayer, F., Lorito, S., Salgado-Gálvez, M. A., Sørensen, M., Abadie, S., Aguirre-Ayerbe, I., Aniel-Quiroga, I., Babeyko, A., Baiguera, M., Basili, R., Belliazzi, S., Grezio, A., Johnson, K., Murphy, S., Paris, R., Rafliana, I., De Risi, R., Rossetto, T., Selva, J., Taroni, M., Del Zoppo, M., Armigliato, A., Bureš, V., Cech, P., Cecioni, C., Christodoulides, P., Davies, G., Dias, F., Bayraktar, H. B., González, M., Gritsevich, M., Guillas, S., Harbitz, C. B., Kânoğlu, U., Macías, J., Papadopoulos, G. A., Polet, J., Romano, F., Salamon, A., Scala, A., Stepinac, M., Tappin, D. R., Thio, H. K., Tonini, R., Triantafyllou, I., Ulrich, T., Varini, E., Volpe, M., and Vyhmeister, E.: Probabilistic Tsunami Hazard and Risk Analysis: A Review of Research Gaps, *Front. Earth Sci.*, 9, 628772, <https://doi.org/10.3389/feart.2021.628772>, 2021.
- Cover, T. M. and Thomas, J. A.: *Elements of information theory*, Wiley, New York, ISBN: 0-471-06259-6, 1991.
- Charvet, I., Ioannou, I., Rossetto, T., Suppasri, A., and Imamura, F.: Empirical fragility assessment of buildings affected by the 2011 Great East Japan tsunami using improved statistical models, *Nat. Hazards*, 73, 951–973, 2014.
- Charvet, I., Suppasri, A., Kimura, H., Sugawara, D., and Imamura, F.: A multivariate generalized linear tsunami fragility model for Kesenuma City based on maximum flow depths, velocities and debris impact, with evaluation of predictive accuracy, *Nat. Hazards*, 79, 2073–2099, 2015.
- Charvet, I., Macabuag, J., and Rossetto, T.: Estimating tsunami-induced building damage through fragility functions: critical review and research needs, *Front. Built Environ.*, 3, 36, <https://doi.org/10.3389/fbuil.2017.00036>, 2017.
- Chua, C. T., Switzer, A. D., Suppasri, A., Li, L., Pakoksung, K., Lallemand, D., Jenkins, S. F., Charvet, I., Chua, T., Cheong, A., and Winspear, N.: Tsunami damage to ports: cataloguing damage to create fragility functions from the 2011 Tohoku event, *Nat. Hazards Earth Syst. Sci.*, 21, 1887–1908, <https://doi.org/10.5194/nhess-21-1887-2021>, 2021.
- De Risi, R., Goda, K., Mori, N., and Yasuda T.: Bayesian tsunami fragility modelling considering input data uncertainty, *Stoch. Env. Res. Risk A.*, 31, 1253–1269, 2017a.
- De Risi, R., Goda, K., Yasuda, T., and Mori, N.: Is flow velocity important in tsunami empirical fragility modeling?, *Earth-Sci. Rev.*, 166, 64–82, 2017b.
- Ebrahimian, H. and Jalayer, F.: Selection of seismic intensity measures for prescribed limit states using alternative nonlinear dynamic analysis methods, *Earthq. Eng. Struct. D.*, 50, 1235–1250, 2021.
- Ebrahimian, H. and Jalayer, F.: ComputeFrag, Version 2, provided by the European Tsunami Risk Service (ETRiS), GitHub [code], https://github.com/eurotsunamirisk/computeFrag/tree/main/Code_ver2, last access: 30 November 2022.
- Eidsvig, U. M. K., Papatoma-Köhle, M., Du, J., Glade, T., and Vangelsten, B. V.: Quantification of model uncertainty in debris flow vulnerability assessment, *Eng. Geol.*, 181, 15–26, 2014.
- Epanechnikov, V. A.: Nonparametric estimation of a multidimensional probability density, *Theor. Probab. Appl.*, 14, 153–158, 1969.
- Goff, J. and Dominey-Howes, D.: The 2009 South Pacific Tsunami, *Earth-Sci. Rev.*, 107, v–vii, 2011.
- Grünthal, G.: *European macroseismic scale 1998*, European Seismological Commission (ESC), ISBN N2-87977-008-4, 1998.
- Hastings, W. K.: Monte-Carlo sampling methods using Markov chains and their applications, *Biometrika*, 57, 97–109, 1970.
- Jalayer, F. and Ebrahimian, H.: Seismic reliability assessment and the non-ergodicity in the modelling parameter uncertainties, *Earthq. Eng. Struct. D.*, 49, 434–457, 2020.
- Jalayer, F., Beck, J., and Zareian, F.: Analyzing the sufficiency of alternative scalar and vector intensity measures of ground shaking based on information theory, *J. Eng. Mech.*, 138, 307–316, 2012.
- Jalayer, F., De Risi, R., and Manfredi, G.: Bayesian Cloud Analysis: efficient structural fragility assessment using linear regression, *B. Earthq. Eng.*, 13, 1183–1203, 2015.
- Jalayer, F., Ebrahimian, H., Miano, A., Manfredi, G., and Sezen, H.: Analytical fragility assessment using unscaled ground motion records, *Earthq. Eng. Struct. D.*, 46, 2639–2663, 2017.
- Jalayer, F., Ebrahimian, H., and Miano, A.: Intensity-based demand and capacity factor design: a visual format for safety-checking, *Earthq. Spectra*, 36, 1952–1975, 2020.
- Koshimura, S., Namegaya, Y., and Yanagisawa, H.: Tsunami Fragility—A New Measure to Identify Tsunami Damage, *J. Disaster Res.*, 4, 479–488, 2009a.
- Koshimura, S., Oie, T., Yanagisawa, H., and Imamura, F.: Developing fragility functions for tsunami damage estimation using numerical model and post-tsunami data from Banda Aceh, Indonesia, *Coast. Eng. J.*, 51, 243–273, 2009b.
- Kullback, S. and Leibler, R. A.: On information and sufficiency, *Ann. Math. Stat.*, 22, 79–86, 1951.
- Lahcene, E., Ioannou, I., Suppasri, A., Pakoksung, K., Paulik, R., Syamsidik, S., Bouchette, F., and Imamura, F.: Characteristics of building fragility curves for seismic and non-seismic tsunamis: case studies of the 2018 Sunda Strait, 2018 Sulawesi–Palu, and 2004 Indian Ocean tsunamis, *Nat. Hazards Earth Syst. Sci.*, 21, 2313–2344, <https://doi.org/10.5194/nhess-21-2313-2021>, 2021.
- Mas, E., Paulik, R., Pakoksung, K., Adriano, B., Moya, L., Suppasri, A., Muhari, A., Khomarudin, R., Yokoya, N., Matsuoka, M., and Koshimura, S.: Characteristics of tsunami fragility functions developed using different sources of damage data from the 2018 Sulawesi earthquake and tsunami, *Pure Appl. Geophys.*, 177, 2437–2455, 2020.
- Miano, A., Jalayer, F., Forte, G., and Santo, A.: Empirical fragility assessment using conditional GMPE-based ground shaking fields: Application to damage data for 2016 Amatrice Earthquake, *B. Earthq. Eng.*, 18, 6629–6659, 2020.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E.: Equations of state calculations by fast computing machines, *J. Chem. Phys.*, 21, 1087–1092, 1953.
- Muhari, A., Imamura, F., Arikawa, T., Hakim, A. R., and Afriyanto, B.: Solving the puzzle of the September 2018 Palu, Indonesia, tsunami mystery: clues from the tsunami waveform and the initial field survey data, *J. Disaster Res.*, 13, sc20181108, <https://doi.org/10.20965/jdr.2018.sc20181108>, 2018.
- Muto, M. and Beck, J. L.: Bayesian updating and model class selection for hysteretic structural models using stochastic simulation, *J. Vib. Control*, 14, 7–34, 2008.

- Paulik, R., Gusman, A., Williams, J. H., Pratama, G. M., Lin, S.-I., Prawirabhakti, A., Sulendra, K., Zachari, M. Y., Fortuna, Z. E. D., Layuk, N. B. P., and Suwarni, N. W. I.: Tsunami hazard and built environment damage observations from Palu city after the September 28 2018 Sulawesi earthquake and tsunami, *Pure Appl. Geophys.*, 176, 3305–3321, 2019.
- Rafliana, I., Jalayer, F., Cerase, A., Cugliari, L., Baiguera, M., Salmanidou, D., Necmioğlu, Ö., Ayerbe, I. A., Lorito, S., Fraser, S., Løvholt, F., Babeyko, A., Salgado-Gálvez, M. A., Selva, J., De Risi, R., Sørensen, M. B., Behrens, J., Aniel-Quiroga, I., Del Zoppo, M., Belliazzi, S., Pranantyo, I. R., Amato, A., and Hancilar, U.: Tsunami risk communication and management: Contemporary gaps and challenges, *Int. J. Disast. Risk Re.*, 70, 102771, <https://doi.org/10.1016/j.ijdr.2021.102771>, 2022.
- Reese, S., Bradley, B. A., Bind, J., Smart, G., Power, W., and Sturman, J.: Empirical building fragilities from observed damage in the 2009 South Pacific tsunami, *Earth-Sci. Rev.*, 107, 156–173, 2011.
- Rosti, A., Del Gaudio, C., Rota, M., Ricci, P., Di Ludovico, M., Penna, A., and Verderame, G. M.: Empirical fragility curves for Italian residential RC buildings, *B. Earthq. Eng.*, 19, 3165–3183, 2021.
- Rota, M., Penna, A., and Strobbia, C. L.: Processing Italian damage data to derive typological fragility curves, *Soil Dyn. Earthq. Eng.*, 28, 933–947, 2008.
- Wing, O. E., Pinter, N., Bates, P. D., and Kousky, C.: New insights into US flood vulnerability revealed from flood insurance big data, *Nat. Commun.*, 11, 1–10, 2020.