



HAL
open science

A Bayesian framework for estimating fragility curves based on seismic damage data and numerical simulations by adaptive neural networks

Zhiyi Wang, Irmela Zentner, Enrico Zio

► To cite this version:

Zhiyi Wang, Irmela Zentner, Enrico Zio. A Bayesian framework for estimating fragility curves based on seismic damage data and numerical simulations by adaptive neural networks. *Nuclear Engineering and Design*, 2018, 338, pp.232-246. 10.1016/j.nucengdes.2018.08.016 . hal-01988948

HAL Id: hal-01988948

<https://hal.science/hal-01988948>

Submitted on 8 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



A Bayesian framework for estimating fragility curves based on seismic damage data and numerical simulations by adaptive neural networks



Zhiyi Wang^{a,b,c,1}, Irmela Zentner^{a,b}, Enrico Zio^{c,d}

^a EDF Lab Saclay, France

^b Institute for Mechanical Sciences and Industrial Applications, UMR 9219 CNRS-EDF-CEA-ENSTA ParisTech, France

^c Chair on Systems Science and Energetic Challenge, European Foundation for New Energy of EDF, CentraleSupélec, Université Paris-Saclay, France

^d Energy Department, Politecnico di Milano, Italy

ABSTRACT

In seismic risk assessment, the fragility curve is used to estimate the reliability of structures and equipment under seismic loads. The shape of fragility curves is usually approximated by the cumulative distribution function of a lognormal distribution. The estimation of the parameters of the fragility curves requires gathering different sources of information and quantifying the uncertainties coming from these sources. This paper proposes a methodology for the computation of fragility curves for nuclear power plant equipment, based on a Bayesian updating framework that combines the results of numerical simulations and damage data. An artificial neural network is trained iteratively by optimizing its prediction uncertainties over the ground motion sample space, and it is used to conduct numerical simulations. The results of the numerical simulations provide a prior estimation of the seismic capacity of the equipment. The estimation of the uncertainty related to the equipment capacity is taken from the literature. Damage data, collected from the in situ observation and the database of the seismic qualification utility group (SQUG), are used to construct the likelihood function for the Bayesian updating. The posterior equipment capacity is evaluated by Markov chain Monte Carlo simulation and posterior fragility curves are, then, obtained. The main contributions of the work are: (i) proposal of an adaptive training algorithm of artificial neural networks to improve the design of experiments for finite element simulations; (ii) proposal of a two-step transformation method to construct the likelihood function with existing damage data from the SQUG database. The methodology is applied to compute the fragility curves of a low-voltage switchgear of a nuclear power plant, within the so-called KARISMA benchmark.

1. Introduction

Seismic probabilistic risk assessment (SPRA) is a widely applied approach to estimate the seismic risk of critical structures, such as nuclear power plants (NPPs). In the framework of SPRA, fragility analysis is conducted to evaluate the fragility curves, i.e. the conditional probabilities of failure of structures or components at given values of the seismic intensity measure (IM), e.g. the peak ground acceleration (PGA). The computation of fragility curves is typically realized by statistical analysis based on different sources of information, including expert judgments, numerical simulations, empirical damage data.

The safety factor method (Kennedy et al., 1980; EPRI, 1994), largely used in nuclear engineering, depends on safety margins determined from simplified structural analyses and experimental data. Uncertainties are evaluated from expertise of engineers or results of qualification tests. The safety margins and their associated uncertainties are used to assess the seismic capacity of structures and equipment. This method does not require numerical simulations. However, the safety margins determined from the simplified approaches can be

conservative, and thus, cannot offer a best estimate of the fragility curves.

Numerical simulations are commonly applied in the current practice of fragility analysis, e.g. by the finite element method (FEM) (Padgett and DesRoches, 2008; Ellingwood and Kinali, 2009; Zentner, 2010). Different sources of aleatory and epistemic uncertainties can be modeled and propagated through the numerical model. The conditional probability of failure can be computed either by pointwise Monte Carlo estimation or by assuming a parametric representation (e.g. lognormal) of fragility curves. The underlying parameters of the fragility curves are determined based on the results of the simulations. However, because of the high complexity of numerical models, the computational cost of the numerical analyses can be very high. One way to reduce the computational burden is to use fast-running statistical metamodelling. Various types of metamodelling have been tested and applied in fragility analysis, such as Kriging (Gidaris et al., 2015), artificial neural networks (ANNs) (Lagaros et al., 2009; Ferrario et al., 2017; Mangalathu et al., 2018; Wang et al., 2018), quadratic response surfaces (Towashiraporn, 2004), polynomial chaos expansion (Mai et al., 2016), among others.

E-mail address: zhiyi.wang@centralesupelec.fr (Z. Wang).

¹ 7 boulevard Gaspard Monge, 91120 Palaiseau, France.

Nevertheless, due to the fact that a numerical model cannot contain all the structural details and damage mechanisms, damage data of structures and equipment can be used for a more accurate computation of fragility curves.

Damage data have been also widely used for seismic fragility analysis. The damage data are obtained either from post-earthquake observations or from qualification tests. They are used to describe the performance of structures or equipment under real seismic excitations. Fragility analysis is, then, conducted by statistical analysis of the damage data. For example, fragility curves for European-type reinforced concrete buildings are determined in Rossetto and Elnashai (2003) with earthquake observational data. Using Italian seismic damage data, fragility curves for different building typologies are built in Rota et al. (2008) to provide a reliable estimate of the vulnerability of structures of different classes. In these works, the fragility curves are obtained directly by fitting the damage data into a cumulative distribution function (CDF) of the lognormal distribution. Other studies adopt Bayesian statistics to estimate the fragility parameters (Straub and Kiureghian, 2008; Gardoni et al., 2009). In the Bayesian framework, prior distributions of the fragility parameters are assumed. Then, damage data are used to build the likelihood function. The fragility parameters can be generated by Markov chain Monte Carlo simulation (MCMC) (Hastings, 1970), based on the posterior distributions obtained from Bayesian updating. The advantage of the Bayesian statistics method is that it yields a probability distribution of the parameter to be estimated (so the confidence intervals can be computed), rather than a single value for the estimation of the parameter. Representative examples of the application of Bayesian statistics in seismic risk assessment can be found in Singhal and Kiremidjian (1998), Koutsourelakis (2010), Jalayer et al. (2010), EPRI (2014), Jaiswal et al. (2011), Beck and Taflanidis (2013), Buratti et al. (2017), Noh et al. (2017), Jeon et al. (2017), EPRI (2017), among others.

The objective of this paper is to propose a framework to take into consideration both numerical simulation results and damage data in the computation of fragility curves. The methodology is divided into two parts: (i) estimation of the prior parameters with numerical simulations: to reduce the computational cost, an ANN metamodel is trained with an iterative active learning algorithm to substitute the computationally expensive FEM simulation; (ii) computation of the likelihood function with the damage data and execution of Bayesian updating to obtain the posterior distribution of the seismic capacity of the equipment. Different sources of uncertainties (aleatory and epistemic) are quantified and integrated in the computation of the fragility curves. Critical equipment of nuclear power plants are designed with high safety standards. The low probability of failure of the equipment may lead to a biased estimation of the fragility parameters, if these parameters are solely determined by the maximum likelihood estimation (Shinozuka et al., 2000). This requires assessing the confidence associated to the estimation, which can be achieved in the Bayesian framework.

This paper is organized as follows. Section 2 describes the global methodology to account for different sources of information in the computation of fragility curves. It consists of the determination of the prior fragility curves with numerical simulation results and the Bayesian updating with damage data. In Section 3, the proposed methodology is applied to evaluate the robustness of a low-voltage switchgear located in the Kashiwazaki-Kariwa nuclear power plant in Japan. Final conclusions are given in Section 4.

2. Description of the methodology

2.1. Seismic fragility curves

Fragility curves compute the conditional probability that the engineering demand parameter (EDP) exceeds a failure threshold, for a given seismic IM:

$$P_f(\alpha) = P(y > y_{\text{crit}} | \alpha) \quad (1)$$

where y is the EDP, such as inter-story drift, y_{crit} is the failure threshold and α represents the seismic IM. The lognormal fragility model proposed in Kennedy et al. (1980) and Huang et al. (2011) is often applied in practice. In the lognormal assumption, the shape of the fragility curve is approximated by the CDF of a lognormal distribution:

$$P_f(\alpha) = \Phi\left(\frac{\ln\alpha - \ln\hat{A}_m}{\beta_R}\right) \quad (2)$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution $N(0, 1)$, \hat{A}_m denotes the median capacity. The parameter β_R represents the aleatory uncertainty related to the inherent randomness. According to Kennedy et al. (1980) and Basu et al. (2017), an epistemic uncertainty β_U , resulting from the lack of knowledge of the structural capacity, should be also considered:

$$\hat{A}_m \sim \text{LogN}(A_m, \beta_U^2) \quad (3)$$

where A_m is the median of the lognormal distribution and LogN denotes a lognormal distribution. Consequently, the $\gamma \in [0, 1]$ non-exceedance confidence interval of the fragility curves can be computed (EPRI, 1994; Kwag et al., 2014; Zentner et al., 2017)

$$\tilde{P}_f(\alpha, \gamma) = \Phi\left(\frac{\ln\alpha - \ln A_m + \beta_U \Phi^{-1}(\gamma)}{\beta_R}\right) \quad (4)$$

Eq. (4) allows computing the high confidence low probability of failure (HCLPF), defined as the capacity where the probability of failure reaches 5% with 95% confidence:

$$A_{\text{HCLPF}} = A_m e^{-1.645(\beta_R + \beta_U)} \quad (5)$$

The mean fragility curve, which considers both aleatory and epistemic uncertainties, is defined by

$$P_f^{\text{mean}}(\alpha) = \Phi\left(\frac{\ln\alpha - \ln A_m}{\sqrt{\beta_R^2 + \beta_U^2}}\right) \quad (6)$$

The objective of the subsequent subsections is to describe the methodology to compute fragility curves for an equipment of interest located in a specific NPP structure, which is named ‘target structure’ in the sequel. The numerical model of the target structure is available. The general workflow of the proposed methodology is illustrated in Fig. 1. In this framework, reference values of β_U^{prior} and y_{crit} are obtained from the literature. To better explain the methodology, we start with the determination of prior fragility parameters based on numerical simulation results. Bayesian updating and MCMC are, then, executed with damage data to obtain the posterior curves. The assumptions made in this methodology are:

1. The fragility curves in this paper are all calculated under lognormal distribution assumptions, namely (i) the fragility curves are computed by the lognormal CDFs (Eq. (2)) and (ii) the epistemic uncertainty is considered lognormally distributed (Eq. (3)), in order to facilitate the application of the Bayesian theorem based on the damage data.
2. The seismic record-to-record variability is considered as the only source of aleatory uncertainty.
3. PGA is the IM parameter used to compute the fragility curves, since most damage data are provided with given values of PGA.
4. Without different specification, the PGA used in the fragility curve stands for the PGA value of the ground motion on the soil free surface in the proximity of the target structure.

2.2. Determination of prior fragility curve parameters with adaptive ANNs

Prior fragility curve parameters are determined based on the results

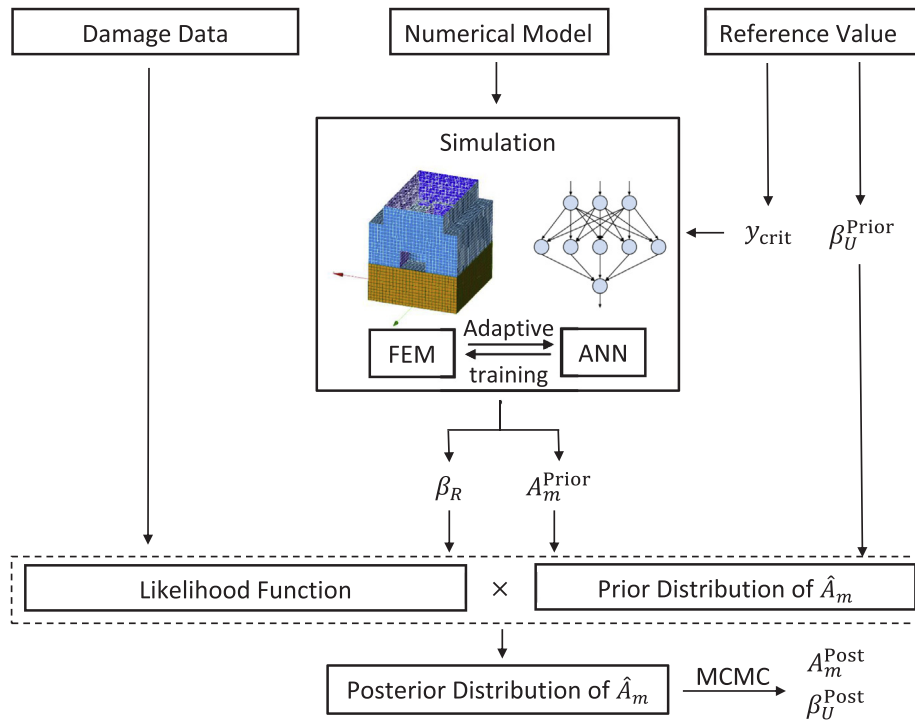


Fig. 1. The Bayesian framework for fragility analysis.

of numerical simulations. FEM is one of the most widely used numerical methods for structural analysis. However, in case of a complex structure, the large number of degrees of freedom of the numerical model makes the resolution process highly computationally expensive. In this case, metamodels, calibrated from the existing simulation results, can be used to substitute the mechanical model, in order to improve the computational efficiency. The ANN is adopted in this paper because of its excellent universal approximation capability (Reed and Marks, 1999; Bishop, 1995).

In this subsection, the method to determine the parameters of prior fragility curves is explained. In order to improve the computational efficiency, ANNs are adopted in this paper to characterize the seismic IMs-EDP relation. With the adaptive enrichment, the quality of the training data is largely improved by reducing its scarcity in the design space. Therefore, less FEM numerical simulations are needed for the calibration of the ANN metamodel.

2.2.1. ANN training and prediction uncertainty

The objective of the application of adaptive ANNs is to establish a non-linear statistical regression model relating the seismic inputs and the EDP of interest:

$$\hat{y} = \hat{f}(IM_1, \dots, IM_k) \quad (7)$$

where the symbol $\hat{\cdot}$ used in this paper denotes all the computation results relevant to ANNs: the non-linear regression model \hat{f} constructed by ANNs and the EDP \hat{y} computed with ANNs.

A classical feed-forward ANN consists of activation functions (linear functions, or non-linear hyperbolic tangent functions) and a set of weighting parameters w adjusted to minimize a cost function. The activation functions are connected by the weighted links in a layered structure. There are three types of layers: (i) input layer, which feeds the variables from which the ANN model is constructed; (ii) hidden layers, being single or multiple, to add parameters and nonlinearity; (iii) output layer, which provides the results of the ANN. The cost function E computes the square error between the ANN predictions \hat{y} and the targets y (e.g. FEM simulation results), summed up over all training examples. The training of ANNs is typically realized by

gradient-based algorithms to find the optimal weighting parameters. The gradient vector $g = \frac{\partial E}{\partial w}$ can be computed efficiently by the back-propagation algorithm (Rumelhart et al., 1986; Bishop, 1995). One can refer to Bishop (1995) and Reed and Marks (1999) for detailed explanations on the basic theory about feed-forward ANNs. More details on the applied approach are also given in Wang et al. (2018).

The prediction intervals (PIs) of ANNs are estimated by the delta method (Chryssolouris et al., 1996; Zio, 2006). Assuming a normal distribution of the ANN training error, the standard deviation (Std) s of the ANN prediction is calculated (Rivals and Personnaz, 2000)

$$s = \sigma_{ANN} \sqrt{1 + \mathbf{h}_{test}^T (\mathbf{J}\mathbf{J}^T)^{-1} \mathbf{h}_{test}} \quad (8)$$

where σ_{ANN} is the Std of the ANN training errors. The Jacobian matrix \mathbf{J} is constructed by the gradient vectors $\mathbf{h}^i = \frac{\partial y^i}{\partial w}$ of the training examples; also \mathbf{h}^i can be computed based on the backward-propagation method (Bishop, 1995). In Eq. (8), \mathbf{h}_{test} is the \mathbf{h} vector for the test example, and the upper index T denotes the matrix transpose. One can refer to Rivals and Personnaz (2000) and Bishop (1995) for more details regarding the computation of \mathbf{h} and the delta method. An important property of the prediction uncertainty computed by Eq. (8) is that it shows large uncertainty at the locations where no enough training data are available. This is originated from the term $\mathbf{h}_{test}^T (\mathbf{J}\mathbf{J}^T)^{-1} \mathbf{h}_{test}$, in which the information of the training data is stored in the Jacobian matrix \mathbf{J} . One can refer to Fig. 5 of Wang et al. (2018) for an illustration example.

2.2.2. Adaptive ANN algorithm

An ANN adaptive learning algorithm is proposed in this paper to improve the design of experiments (DoEs) of FEM simulations. The algorithm is based on the prediction uncertainty of ANNs computed with the delta method. The algorithm adds iteratively in the training dataset the points at the zone of interest (e.g., at the location where the performance of the metamodel is limited, with a large prediction uncertainty). The iterative training of ANNs is stopped when a certain accuracy criterion is satisfied. The principle of such an algorithm has been used in Kriging for fragility analysis in Gidaris et al. (2015), but it is seldom used with ANNs. In fact, adaptive learning is widely used for Kriging metamodels, e.g. in Jones et al. (1998), Echard et al. (2011),

because the prediction uncertainty is directly provided in the output, which is not the case for other metamodels, such as ANNs or support vector machines. For this reason, the enrichment of new training samples in most adaptive training procedures applied to ANNs is not based on the prediction uncertainty. Rather, importance sampling, directional simulation or MCMC are typically used to create new samples in the area of interest for an enriched adaptive training (Hurtado and Alvarez, 2001; Papadopoulos et al., 2012; Pedroni and Zio, 2017). However, these approaches cannot be easily applied in fragility analysis, since it is difficult to generate or to find an earthquake motion, conditional on required values of multiple IMs (e.g. PGA and Irias intensity). An alternative is to generate a large population of initial samples and to enrich the DoEs with samples in the initial population according to an enrichment criterion. Such a strategy has been studied by Xiao et al. (2018) with ANNs, in which the computation of the prediction uncertainty is based on cross-validations, requiring retrainings of ANNs. In this paper, we quantify the uncertainty linked to ANN predictions with the delta method, which can be directly obtained once the ANN is trained, and integrate the prediction uncertainty in the enrichment criterion of the DoEs, to ensure the performance of the ANN on the whole input space, with less training data. The whole workflow of the adaptive ANN algorithm is summarized in Fig. 2. Combined with the FEM simulations, the ANN adaptive training algorithm is as follows:

1. Initialization of the DoEs. To prepare for the adaptive learning algorithm, a population X composed of N seismic signals should be generated. IMs are extracted for all N seismic ground motions. N_0 seismic motions are randomly selected from X to carry out FEM simulations. N_0 should be larger than the total number of the ANN parameters, including weights and biases.
2. Starting of the iterative ANN training. For iteration k ($k \geq 0$), the ANN is trained with N_k IMs and EDP pairs in the current DoEs. Here N_k is used to denote the number of IMs-EDP data in the DoEs at the iteration k . The ANN is suggested to be trained *in the log-log space*, i.e. with $\ln(\text{IMs})$ and $\ln(\text{EDP})$. ANN simulation is, then, carried out for every IMs set in the total population X . Predictions \hat{y}_k^i and the

associated prediction uncertainty s_k^i can be obtained with the ANN trained at the current iteration k .

3. Computation of the accuracy metric δ_k^i and the accuracy threshold δ_k^{crit} :

$$s_{\text{ref},k} = \frac{1}{N_k} \sum_{p=1}^{N_k} s_k^p, \text{ for } p \text{ in the training set } (p = 1, \dots, N_k) \quad (9)$$

$$\delta_k^i = \frac{|s_k^i - s_{\text{ref},k}|}{s_{\text{ref},k}}, \text{ for every } i \text{ in } X (i = 1, \dots, N) \quad (10)$$

$$\delta_k^{\text{crit}} = \max_p(\delta_k^p), \text{ for } p \text{ in the training set } (p = 1, \dots, N_k) \quad (11)$$

The quantity $s_{\text{ref},k}$ is the mean value of the prediction uncertainty of the training examples at the iteration k . It serves as the reference value to compute the accuracy metric. The accuracy metric δ_k^i calculates the relative deviation of s_k^i with respect to $s_{\text{ref},k}$. A large value of δ_k^i indicates a large prediction uncertainty s_k^i , so that the corresponding ANN prediction is less reliable. The accuracy threshold δ_k^{crit} is set as the maximal relative deviation of prediction uncertainty in the training dataset.

4. Verification of the ANN accuracy. The set of test samples is defined by all the samples in X which are not used to train the ANN. When δ_k^i of every test sample is smaller than the accuracy threshold, i.e. $\max(\delta_k^{\text{test}}) < \delta_k^{\text{crit}}$, it can be considered that the samples in the ANN training set are enough to cover the whole input space of the population X . So the ANN is accurate enough for X . The iterative training can be, therefore, stopped. Otherwise, a further enrichment of the training data is necessary.
5. Enrichment of the ANN training dataset. If the accuracy of the ANN is not satisfied, M samples in the test set with the largest δ_k^i values (i.e. with the largest prediction deviations) are selected to run the FEM simulations. The results of the FEM simulations are added to the ANN training data. Set $k = k + 1$, and go back to Step 2.

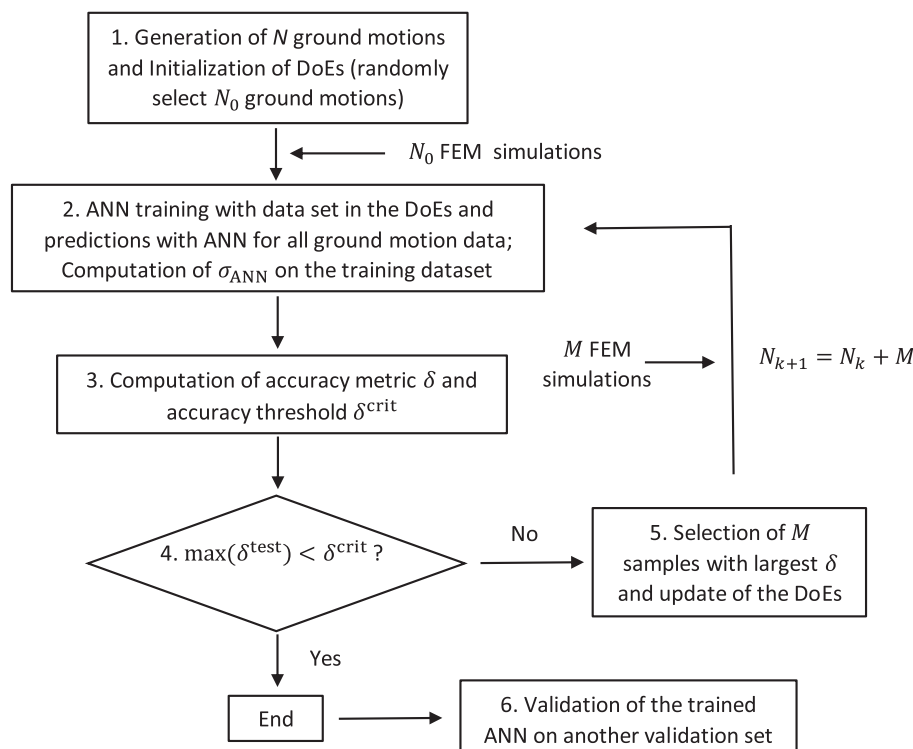


Fig. 2. Workflow of the adaptive training of ANNs.

- After being trained, the ANN is validated on another validation dataset, which is different from the training dataset.

Due to the property of the prediction uncertainty s , some ground motions with high intensities, which are often outside the validity domain of the ANNs in the first few iterations (so their prediction uncertainties are very high), can be also automatically selected by the adaptive training. With a reasonable number of FEM simulations, an ANN is obtained at the end of the adaptive training. Then, ANN simulations can be carried out for ground motions in the whole population X to predict the structural EDPs \hat{y} .

2.2.3. Determination of prior fragility parameters

Prior fragility parameters β_R and A_m^{prior} can be estimated from the ANN simulation results. The failure threshold y_{crit} , informed from the reference value in the literature, provides an estimation of A_m^{prior} . The Std of the ANN training error is integrated in the computation of fragilities to consider the metamodel uncertainty. Such an idea has been used in [Gidaris et al. \(2015\)](#) and [Wang et al. \(2018\)](#) to account for the metamodel error in the lognormal based fragility models. Since a set of IMs, instead of the whole ground motion, is used as the input of the ANN, there is a loss of information in the input. σ_{ANN} is used to quantify this loss of the ground motion randomness, which cannot be conveyed by the IM set and therefore cannot be captured by the ANN. More details concerning the inclusion of σ_{ANN} in the fragility analysis can be found in [Wang et al. \(2018\)](#). More precisely, the determination of A_m^{prior} and β_R is realized by:

- Linear regression of the data cloud (α, \hat{y}) in log-log space ([Cornell et al., 2002](#); [Ellingwood and Kinali, 2009](#); [Zentner et al., 2017](#)). In the application of this paper, α denotes the PGA.

$$\ln \hat{y} = c \ln \alpha + \ln b + \varepsilon \quad (12)$$

where b and c are regression parameters determined from the data cloud $(\ln \alpha, \ln \hat{y})$ and the residual ε follows a normal distribution $N(0, \sigma_{\text{RLIM}}^2)$.

- Computation of the conditional probability of failure, considering the Std of ANN training errors σ_{ANN} .

$$P_f(\alpha) = \Phi \left(\frac{\ln b \alpha^c - \ln y_{\text{crit}}}{\sqrt{\sigma_{\text{RLIM}}^2 + \sigma_{\text{ANN}}^2}} \right) \quad (13)$$

- Reformulating Eq. (13) for coherence with Eq. (2).

$$P_f(\alpha) = \Phi \left(\frac{c \ln \alpha - (\ln y_{\text{crit}} - \ln b)}{\sqrt{\sigma_{\text{RLIM}}^2 + \sigma_{\text{ANN}}^2}} \right) = \Phi \left(\frac{\ln \alpha - (\ln y_{\text{crit}} - \ln b)/c}{\sqrt{\sigma_{\text{RLIM}}^2 + \sigma_{\text{ANN}}^2/c}} \right) \quad (14)$$

Therefore, $A_m^{\text{prior}} = \sqrt[3]{y_{\text{crit}}/b}$ and $\beta_R = \sqrt{\sigma_{\text{RLIM}}^2 + \sigma_{\text{ANN}}^2}/c$.

The prior value of β_U of the equipment capacity is chosen in agreement with the literature, such as ([EPRI, 2014](#)). With the computed value of A_m^{prior} , the prior distribution of \hat{A}_m is determined: $f^{\text{prior}}(\hat{A}_m) \sim \text{LogN}(A_m^{\text{prior}}, (\beta_U^{\text{prior}})^2)$.

2.3. Bayesian updating of fragility curves with damage data

2.3.1. Damage database

The damage data \mathbf{z} used in this study are taken from the seismic qualification utility group (SQUG) database. The SQUG database ([EPRI, 2016](#)), built by the Electric Power Research Institute (EPRI), gathers seismic experience data related to seismic capacity of equipment in industrial facilities (not limited to NPPs) ([Starck and Thomas, 1990](#)). The data in the SQUG database are mostly obtained from post-earthquake inspections of equipment in these industrial facilities. 32

earthquakes from 1971 to 2010 are registered in the SQUG database with most of them taking place in the USA. Some strong earthquakes in Chile, Japan, Turkey, etc are also included. The equipment in the SQUG database is divided into 20 conventional classes, including switchgears, batteries, motor control centers. A list of the 20 equipment classes is summarized in [Starck and Thomas \(1990\)](#).

For the data collected in the SQUG database, each observation contains the information: (i) equipment description (size, manufacturer, etc); (ii) the earthquake and the PGA; (iii) the industrial facility where the equipment is located; (iv) the elevation h of the equipment in the facility structure; (v) the description of the performance of the equipment after the earthquake. It has to be mentioned that no details on the supporting structures are provided in the database, so that the FEM models for the structures in the SQUG database are in general not available. The integration of the damage data in the Bayesian updating depends also on these supporting structures. The method to construct the likelihood function with the damage data is explained in detail in Section 2.3.3. In our study, the damage data for the low-voltage switchgear are collected from the SQUG database. They are used in the Section 3 for the Bayesian updating of the fragility curves.

2.3.2. Bayesian framework in fragility analysis

Given the damage data \mathbf{z} , the posterior distribution of \hat{A}_m can be computed by the Bayes' theorem:

$$f^{\text{post}}(\hat{A}_m | \mathbf{z}) = k L(\mathbf{z} | \hat{A}_m) f^{\text{prior}}(\hat{A}_m) \quad (15)$$

where $L(\mathbf{z} | \hat{A}_m)$ is the likelihood function determined by the observed data, and k is a constant to normalize the posterior distribution. Every observational data vector \mathbf{z}^i has two components: the PGA value α^i of the seismic excitation and the damage state x^i of the equipment of interest after the earthquake. This latter x^i is modeled as a binary Bernoulli variable: $x^i = 0$ if no failure occurs and $x^i = 1$ if the equipment fails. According to [Shinozuka et al. \(2000\)](#), the likelihood function with given \mathbf{z} is written as:

$$\begin{aligned} L(\mathbf{z} | \hat{A}_m) &= \prod_{i=1}^{n_{\text{obs}}} [P_f(\alpha^i)]^{x^i} [1 - P_f(\alpha^i)]^{1-x^i} \\ &= \prod_{i=1}^{n_{\text{obs}}} \left[\Phi \left(\frac{\ln \alpha^i - \ln \hat{A}_m}{\beta_R} \right) \right]^{x^i} \left[1 - \Phi \left(\frac{\ln \alpha^i - \ln \hat{A}_m}{\beta_R} \right) \right]^{1-x^i} \end{aligned} \quad (16)$$

where n_{obs} is the number of the empirical data. Substituting Eq. (16) into Eq. (15), one can obtain the expression of the posterior distribution of \hat{A}_m :

$$f^{\text{post}}(\hat{A}_m | \mathbf{z}) \propto \left(\prod_{i=1}^{n_{\text{obs}}} [P_f(\alpha^i)]^{x^i} [1 - P_f(\alpha^i)]^{1-x^i} \right) f^{\text{prior}}(\hat{A}_m) \quad (17)$$

Knowing $L(\mathbf{z} | \hat{A}_m)$ and $f^{\text{prior}}(\hat{A}_m)$, the MCMC simulation allows sampling efficiently the posterior distribution without computing explicitly the constant k of Eq. (15) ([Hastings, 1970](#)). Therefore, the essential part of the Bayesian updating is to determine the parameters in $L(\mathbf{z} | \hat{A}_m)$ to construct the likelihood function.

It appears that the computation of $L(\mathbf{z} | \hat{A}_m)$ is straightforward. However, different kinds of PGA values can be provided in the database. Before computing the likelihood function, one has to ensure that the PGA values to plug in Eq. (17) describe the free surface ground motions near the target equipment in which the equipment is located.

2.3.3. Determination of the likelihood function

Two groups of earthquake observational damage data are discussed in this study:

- In-situ earthquake observational data $\mathbf{z}_t = (\alpha_t, x)$ of the equipment of interest in the *target* structure, where α_t denotes the PGA level recorded on the free surface near the target structure (Let us recall

that the FEM model of the target structure is available).

- SQUG earthquake observational data $\mathbf{z}_s = (\alpha_s, x)$ of a similar equipment positioned in other civil structures in the database, named SQUG structures in this paper. The quantity α_s denotes the PGA level recorded on the free surface near the SQUG structures.

The total workflow of the determination of the likelihood function with damage data is summarized in Fig. 3.

Likelihood function for \mathbf{z}_t . Recalling that the PGA values in Eq. (17) should be α_t , data of the first category can be directly inserted into Eq. (17).

$$L(\mathbf{z}_t | \hat{A}_m) = \prod_{i=1}^{n_t} \left[\Phi \left(\frac{\ln \alpha_t^i - \ln \hat{A}_m}{\beta_R} \right) \right]^{x^i} \left[1 - \Phi \left(\frac{\ln \alpha_t^i - \ln \hat{A}_m}{\beta_R} \right) \right]^{1-x^i} \quad (18)$$

where n_t is the number of the in-situ observational data.

However, the use of the SQUG data is not straightforward. The purpose of the subsequent part is to propose a method to integrate SQUG data in Eq. (17), i.e. a method to transform \mathbf{z}_s into \mathbf{z}_t with also the quantification of the associated uncertainty in the transformation. The essential idea of the assumption is that the damage state of the equipment after the earthquake depends on the PGA value of the floor acceleration.

Likelihood Function for \mathbf{z}_s . The transformation from \mathbf{z}_s into \mathbf{z}_t consists of two steps:

- Compute the PGA of the floor acceleration of the SQUG structure given the PGA on the free soil surface.
- Considering that the equipment is positioned in the target structure with the same floor acceleration, compute the PGA of the free surface ground motion of the target structure.

We start with the first step of the transformation. In general, the numerical model of the SQUG structure is difficult to obtain, and only the elevation h of the location of the equipment is provided in SQUG data. Without any information on the SQUG structure, the simple amplification factor model used in EPRI (2014) is adopted in this study:

$$\alpha_{fl} = \lambda(h) \alpha_s \quad (19)$$

where $\lambda(h)$ is the amplification factor, which is a function of the elevation. α_{fl} denotes the PGA of the floor acceleration. The quantity $\lambda(h)$ contains a median value $\bar{\lambda}(h)$ and a lognormal uncertainty $\tilde{\varepsilon}$: $\lambda(h) = \bar{\lambda}(h) \tilde{\varepsilon}$ with $\tilde{\varepsilon} \sim \text{LogN}(1, \tilde{\beta}^2)$. Here, a linear relation is selected for $\bar{\lambda}(h)$:

$$\bar{\lambda}(h) = c_h h + b_h \quad (20)$$

The parameters c_h and b_h are determined according to the amplification factor values used in EPRI (2014):

$$\bar{\lambda} = 1 \text{ if } h = 0 \quad (21)$$

$$\bar{\lambda} = 1.5 \text{ if } h = 12.192\text{m} \quad (22)$$

So far, the floor acceleration α_{fl} and its associated uncertainty $\tilde{\varepsilon}$ have been computed. The second step of the transformation is explained in what follows. The transformation of α_{fl} to α_t is, in fact, the transformation of the floor PGA of the target structure into the free surface PGA. This transformation can be realized with a statistical model established from the FEM simulation results, which are used to train the adaptive ANN.

From the FEM simulation results of Section 2.2, PGA values of the floor accelerations α_f^{FEM} and the free surface ground motions α_g^{FEM} of the target structure can be extracted. Let us assume that both PGA values are lognormally distributed. This assumption is checked later in our specific case study. The lognormal assumption allows building a bi-variate normal distribution of $\ln \alpha_f^{\text{FEM}}$ and $\ln \alpha_g^{\text{FEM}}$. The marginal distributions of $\ln \alpha_f^{\text{FEM}}$ and $\ln \alpha_g^{\text{FEM}}$ are denoted by $N(\ln \mu_f^{\text{FEM}}, (\sigma_f^{\text{FEM}})^2)$ and $N(\ln \mu_g^{\text{FEM}}, (\sigma_g^{\text{FEM}})^2)$, respectively, with the correlation coefficient ρ . According to the property of the conditional bi-variate normal distribution, it can be shown that the median value of the transformed free surface PGA $\alpha_{fl \rightarrow t}$ and its uncertainty $\beta_{fl \rightarrow t}$ are calculated by

$$\ln \alpha_{fl \rightarrow t} = \ln \mu_g^{\text{FEM}} + \rho (\ln \alpha_{fl} - \ln \mu_f^{\text{FEM}}) \frac{\sigma_g^{\text{FEM}}}{\sigma_f^{\text{FEM}}} \quad (23)$$

$$\beta_{fl \rightarrow t}^2 = (1 - \rho^2) (\sigma_g^{\text{FEM}})^2 \quad (24)$$

Combining Eqs. (19), (23), (24) and considering the property of the normal distribution, one can show that the median value of the transformed PGA $\ln \alpha_{s \rightarrow t}$ on the free surface of the target structure is calculated by

$$\ln \alpha_{s \rightarrow t} = \ln \mu_g^{\text{FEM}} + \rho [\ln(\bar{\lambda}(h) \alpha_s) - \ln \mu_f^{\text{FEM}}] \frac{\sigma_g^{\text{FEM}}}{\sigma_f^{\text{FEM}}} \quad (25)$$

and its related uncertainty is

$$\beta_{s \rightarrow t}^2 = \left(\frac{\sigma_g^{\text{FEM}}}{\sigma_f^{\text{FEM}}} \rho \right)^2 \tilde{\beta}^2 + \beta_{fl \rightarrow t}^2 = \left(\frac{\sigma_g^{\text{FEM}}}{\sigma_f^{\text{FEM}}} \rho \right)^2 \tilde{\beta}^2 + (1 - \rho^2) (\sigma_g^{\text{FEM}})^2 \quad (26)$$

In the end, the likelihood function for data \mathbf{z}_s is derived:

$$L(\mathbf{z}_s | \hat{A}_m) = \prod_{i=1}^{n_s} \left[\Phi \left(\frac{\ln \alpha_{s \rightarrow t}^i - \ln \hat{A}_m}{\sqrt{\beta_R^2 + \beta_{s \rightarrow t}^2}} \right) \right]^{x^i} \left[1 - \Phi \left(\frac{\ln \alpha_{s \rightarrow t}^i - \ln \hat{A}_m}{\sqrt{\beta_R^2 + \beta_{s \rightarrow t}^2}} \right) \right]^{1-x^i} \quad (27)$$

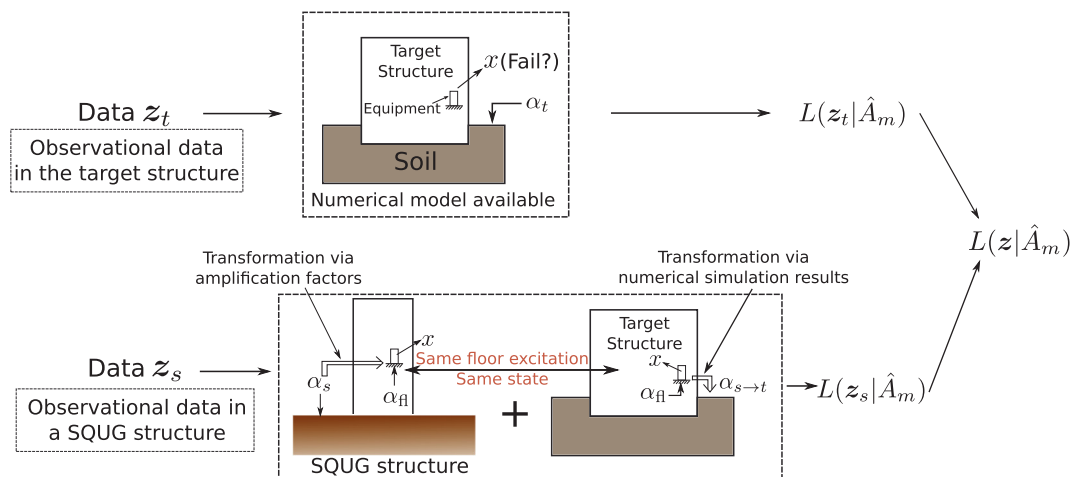


Fig. 3. The workflow to compute the likelihood function.

It is worth mentioning that the interpretations of β_R and $\beta_{s \rightarrow t}$ are different: the former represents the record-to-record aleatory uncertainty when the ground motion time history is characterized by a scalar PGA value, whereas the latter expresses the uncertainty of the transformed PGA value due to the underlying statistical modeling.

Consequently, the total likelihood function for the two categories of data is computed by

$$L(\mathbf{z}|\hat{A}_m) = L(\mathbf{z}_s|\hat{A}_m)L(\mathbf{z}_t|\hat{A}_m) \tag{28}$$

3. Case study: Application to KARISMA benchmark

This section is dedicated to apply the proposed methodology to an industrial case study. Moreover, a sensitivity analysis is conducted at the end with respect to some uncertain parameters. The equipment of interest is a low-voltage switchgear (LVSG) in the Kashiwazaki-Kariwa NPP (K-K NPP). In NPPs, the LVSG is a combination of electrical control units such as circuit breaks and relays, etc, whose function is to ensure and protect the performance of 480V-AC (alternative current) electrical systems. K-K NPP experienced the strong Niigataken-Chuetsu-Oki (NCO) earthquake with magnitude $M_w = 6.6$ in 2007. In this context, the Kashiwazaki-Kariwa Research Initiative for Seismic Margin Assessment (KARISMA) benchmark was organized by the International Atomic Energy Agency (IAEA). The objective of this benchmark is to compare seismic responses calculated by numerical simulations to registered responses of K-K NPP Unit 7 (IAEA, 2013). In addition, a post-earthquake inspection was carried out in order to check the performance of the equipment in K-K NPP after the earthquake (EPRI, 2007).

3.1. KARISMA numerical model

The FEM model of the K-K NPP Unit 7 is shown in Fig. 4. The model consists of 92,000 degrees of freedom with 10,700 nodes and 15,600 elements, including bars, beams, and different shell elements. The constitutive law of the materials is considered linear. The NPP model is embedded 23 meters in the soil, which is accounted for in the soil-structure-interaction (SSI) analysis. The LVSG of interest is located on the -1 floor of the K-K NPP building. The structural analyses are carried out with the Code_Aster, an open-source FEM software developed by Electricity of France (Code_Aster), while the soil part is solved with MISS based on the boundary element method (BEM) using MISS3D (Clouteau, 2005) available with Code_Aster via Salome_Meca platform (Salome_Meca).

Two strong earthquake scenarios, which have affected the area of Kashiwazaki and Kariwa, are considered in this study: (i) the 2007 NCO earthquake scenario with $M_w = 6.6$ and source-to-site distance $r = 16$ km; (ii) the 2004 Chuetsu earthquake with $M_w = 6.8$ and source-

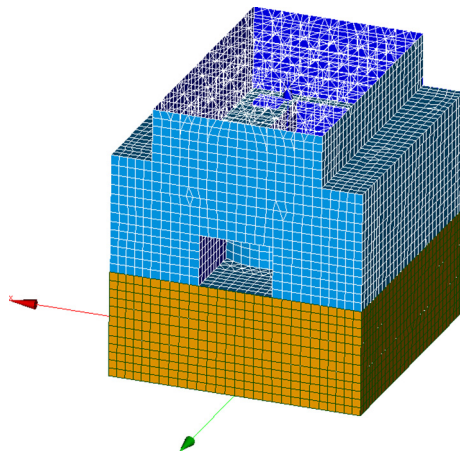


Fig. 4. FEM model of the K-K NPP Unit 7.

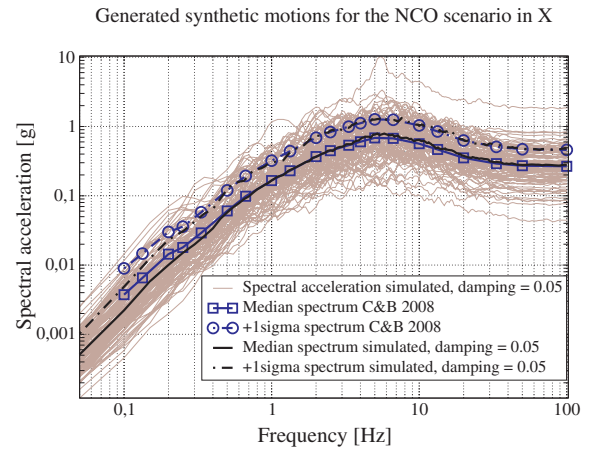


Fig. 5. Generated motions for NCO earthquake scenario and comparison with the spectra predicted by C&B 2008.

to-site distance $r = 29$ km. Given the NCO and Chuetsu scenarios, the generation of the synthetic ground motions at the bedrock with $V_{S30} = 720$ m/s is based on the median and 1σ spectra given by the Campbell-Bozorgnia 2008 (C&B 2008) ground motion prediction equations (Campbell and Bozorgnia, 2008). 250 triplets of 3D synthetic ground motions are generated for each scenario (so 500 triplets in total) and used for the uncertainty propagation. The generated 3D ground motions of each scenario have the median and 84% percentile in agreement with the spectra provided by C&B 2008, using the operator GENE_ACCE_SEISME in Code_Aster (Zentner, 2014). A correlation coefficient of 0.2 is assumed for the two horizontal components (in X and Y), according to Zentner et al. (2017). The vertical component is not correlated to the horizontal ones. A vertical-to-horizontal ratio of $\frac{2}{3}$ is applied in the generation procedure, as suggested by Newmark and Hall (1978). The generated motions in X direction for the NCO scenario are shown in Fig. 5. To obtain enough failure counts for the fragility analysis, the synthetic seismic motions at the bedrock are scaled with a factor of two for the NCO scenario and a factor of three for the Chuetsu scenario.

500 bedrock motions are convoluted on the free surface based 1D soil column reconvolution with the equivalent linear method (ELM) (Yoshida et al., 2002). Meanwhile, 500 degraded soil profiles are obtained. In order to reduce the computational cost, the impedances of the soil and the seismic forces have not been computed for each soil profile using BEM. The 3D seismic signals at the bedrock are regrouped into four soil classes according to their PGA values: (i) $PGA \in [0, 0.3g]$; (ii) $PGA \in [0.3g, 0.5g]$; (iii) $PGA \in [0.5g, 1.0g]$; (iv) $PGA \in [1.0g, +\infty)$. The degraded soil profiles are averaged within each class and 4 soil profiles are obtained to represent four different degradation levels. The 500 ground motions on the free surface, as well as the impedances and seismic forces calculated from the four soil profiles, are used as inputs of the SSI analyses to compute the floor accelerations of the K-K NPP.

In this paper, the failure is described by the non-operational state of the LVSG after the earthquake. Reparation of the equipment is necessary. It can be caused by the fact that (i) relays or breakers cannot return to their operational state after the earthquake or (ii) structural damage has occurred to the equipment, for example anchorage failure at its base (EPRI, 1991). The capacity of the switchgear is given by the average floor spectral acceleration in 5-9Hz, which covers the first natural frequency of the LVSG. The maximum value of the floor spectral accelerations in the two horizontal directions, averaged over 5-9Hz is defined as the EDP y :

$$y = \frac{1}{9} \max_{-5, i=X, Y} \int_{5\text{Hz}}^{9\text{Hz}} S_{a,i}^e(f) df \tag{29}$$

where $S_{a,i}^e$ denotes the floor spectral acceleration of the LVSG in the i -th

direction. A value of 1.8g is selected for y_{crit} according to EPRI (1991), in which the failure threshold of the floor spectral acceleration $S_{a,i}^e(f)$ of the LVSG is a constant value 1.8g for the frequency range [3Hz, 16Hz] (so its average for the frequency range [5Hz, 9Hz] is also 1.8g).

3.2. Prior fragility parameters

An ANN is trained iteratively with the algorithm explained in Section 2.2.2. The IMs of the 500 convoluted seismic motions on the free surface of the K-K NPP can be extracted. Three IMs are used as the inputs of the ANN: (i) PGA, which is widely used in fragility analysis; (ii) ASA (Biasio et al., 2015), the average spectral acceleration in 5-9Hz; (iii) PGV (peak ground velocity), a classical IM for the mid-frequency range. The geometric means of IMs in the two horizontal directions are used as scalar IMs for 3D ground motions. The number of the neurons in the hidden layer of the ANN is 4. In this way, the architecture of the ANN is determined: 3 input parameters (PGA, ASA, PGV), 4 hidden layer nodes and 1 output parameter (the EDP defined by Eq. (29)). $N_0 = 30$ seismic motions from the total 500 signals are randomly selected for the initialization of the adaptive ANN training. $M = 4$ data are added in the DoE in every iteration. The ANN is trained in *log-log* space, i.e. with $\ln(\text{IMs})$ as inputs and $\ln(\text{EDP})$ as outputs.

The adaptive training of the ANN is stopped after 62 calls of FEM simulations. To visualize the DoEs determined by the proposed ANN algorithm, we plot the data cloud in PGA-EDP space in Fig. 6. The PGA used in the horizontal axis is PGA on the free surface. The convergence curve of the adaptive ANN training is shown in Fig. 7. In this figure, the maximum value of δ of the test dataset is plotted against the iteration number k . At iteration 8, the stopping condition is satisfied and the ANN iterative training is stopped.

In order to show the advantage of the adaptive algorithm, 62 seismic motions are randomly selected from the total 500 signals. FEM simulations are conducted for the 62 randomly selected signals and the corresponding data cloud is plotted in PGA-EDP space in Fig. 8. From Fig. 6 and Fig. 8, one can clearly observe the improvement of the DoEs with the ANN adaptive training algorithm. The data in Fig. 6 are better distributed in the PGA-EDP space than the data in Fig. 8: too many data are concentrated in low PGA range in Fig. 8, with only one point exceeding the failure threshold. On the contrary, more failures are contained in the dataset obtained by the ANN adaptive training, which is more convenient for the accuracy of the fragility analysis.

The training results of the ANN are shown in Fig. 9a. The ANN predictions of the training dataset are compared to the real FEM results (target output) in *log space*. From Fig. 9a, one can conclude that the training results are satisfactory. Most of the results in the ‘prediction-target’ space are located in the proximity of the dashed diagonal line.

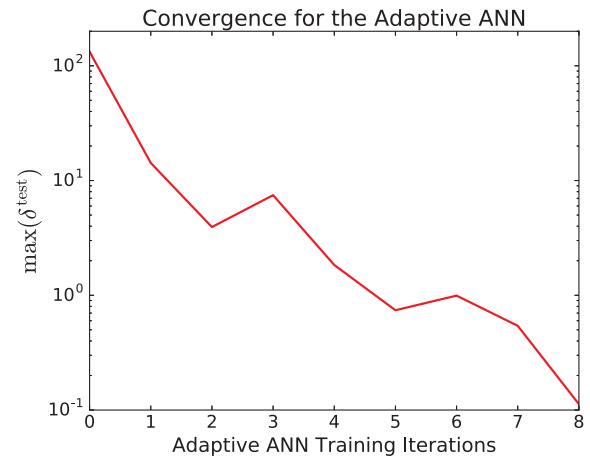


Fig. 7. Convergence curve for the adaptive ANN training.

Another 60 ground motions, which are different from the training dataset determined by the adaptive algorithm, are selected to validate the constructed ANN model. FEM and ANN simulations are performed for the 60 validation seismic signals, respectively. The validation results are plotted in Fig. 9b: the validation results are also satisfactory.

A total of 500 ANN simulations are conducted for the whole ground motion population with the constructed ANN metamodel: 500 pairs of PGA-EDP are, then, obtained. The values of β_R and A_m^{prior} are estimated from the ANN simulations results with the method explained in Section 2.2.3, with $A_m^{prior} = 2.46g$ and $\beta_R = 0.145$. According to EPRI (2009, 2014), a reasonable estimation of β_U^{prior} concerning the uncertainty of the equipment capacity is 0.4. Consequently, the prior distribution of \hat{A}_m follows $LogN(2.46g, 0.4^2)$.

3.3. Determination of the likelihood function

The LVSG damage data can be divided into two groups: one in-situ observation z_i for K-K NPP and 78 post-earthquake inspection data z_s for the LVSG in the SQUG structures. Regarding the in-situ observation, the LVSG in the K-K NPP Unit 7 was not damaged after the NCO earthquake with $PGA = 0.69g$ near the Unit 7. As no detailed information on the performance of the K-K NPP equipment has been found after the 2004 Chuetsu earthquake in the SQUG database, the in-situ observational data contain only the one from the 2007 NCO earthquake. On the other hand, the total number of SQUG damage data for the LVSG is 78, with only one failure observed in the El Centro Steam Plant after the 1979 Imperial Valley Earthquake with local PGA

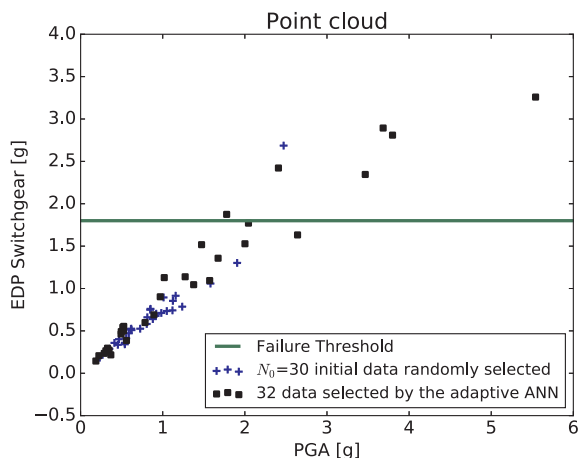


Fig. 6. Data determined by the adaptive algorithm from 500 seismic ground motions plotted in PGA-EDP space.

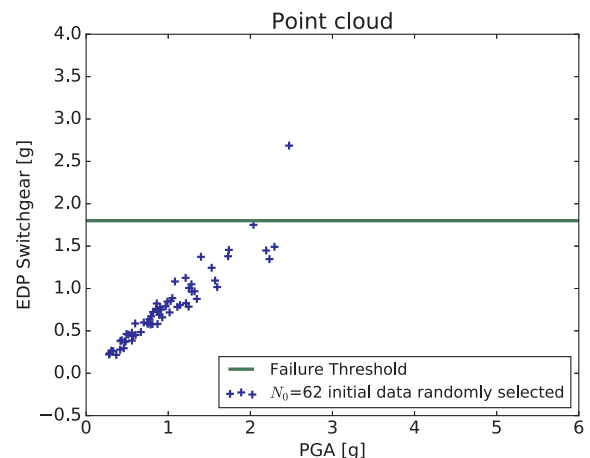


Fig. 8. Data randomly selected from 500 seismic ground motions plotted in PGA-EDP space.

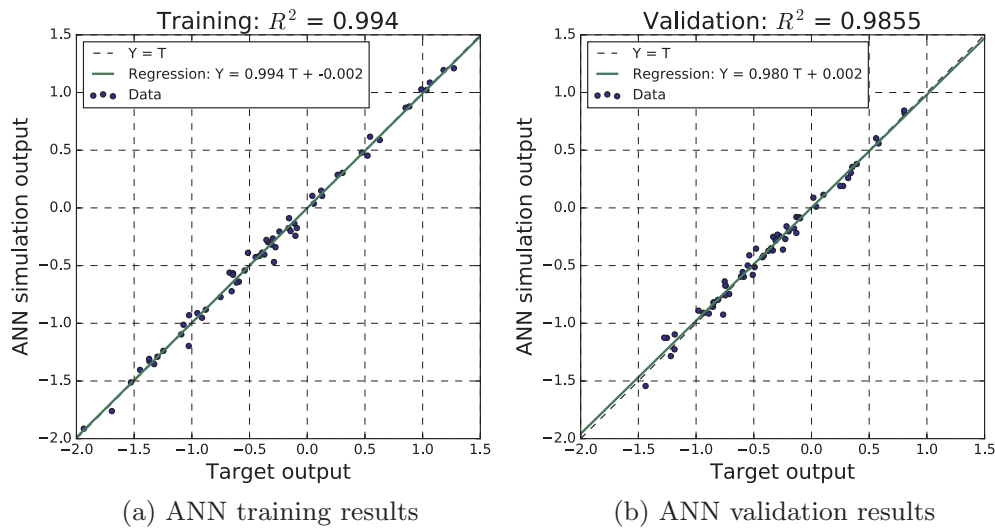


Fig. 9. Comparison of ANN results with FEM results. The comparison is conducted for $\ln\hat{y}$ and $\ln y$.

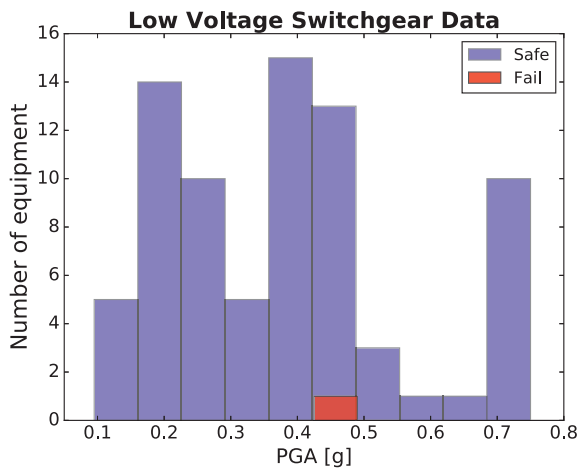


Fig. 10. PGA values of the SQUG data for the LVSG.

value of 0.43g. After the earthquake, it has been noticed that circuit breakers of the LVSG had refused to close. However, according to the inspection report, it is not clear that the failure of the LVSG is caused by the earthquake. It can be also due to the corrosion in the mechanical linkages, which is not earthquake-related. Therefore in the present paper, we set $x^i = 0.5$ for this potential failure, meaning that with a probability of 50% the LVSG failed during the Imperial Valley Earthquake. The local PGA values measured at different industrial structures are plotted in Fig. 10. A summary of the SQUG data for the low voltage switchgear is given in Table 1.

The likelihood function $L(\mathbf{z}_s | \hat{A}_m)$ for the K-K NPP in-situ observation is straightforward with Eq. (18). We focus on the computation of $L(\mathbf{z}_s | \hat{A}_m)$ in what follows.

We follow the two-step method described in Section 2.3.3 to calculate $L(\mathbf{z}_s | \hat{A}_m)$:

1. Step 1: computation of the PGA value of the floor acceleration of the SQUG structures with the amplification factor model, given the PGA on the soil free surface. The median values $\bar{\lambda}(h)$ for the amplification factors can be obtained by Eqs. (20)–(22) with the elevation values h provided in the database. The uncertainty $\tilde{\beta}$ of the amplification factors is assumed to be 0.2, so that the true values of the amplification factors $\lambda(h)$ have a probability of 95% to lie approximately between $\frac{2}{3}\bar{\lambda}(h)$ and $1.5\bar{\lambda}(h)$. Therefore, the median PGA

Table 1

Summary of the SQUG data for the LVSG

Earthquake	Number of the inspected LVSGs	Number of failures
1971 San Fernando Earthquake	9	0
1973 Point Mugu Earthquake	1	0
1975 Ferndale Earthquake	1	0
1979 Imperial Valley Earthquake	5	0.5
1983 Coalinga Earthquake	1	0
1984 Morgan Hill Earthquake	1	0
1985 Chile Earthquake	4	0
1985 Mexico Earthquake	1	0
1986 Adak Earthquake	2	0
1986 Chalfant Valley Earthquake	1	0
1987 Bay of Plenty Earthquake	3	0
1987 Superstition Hills Earthquake	1	0
1987 Whitter Earthquake	7	0
1989 Loma Prieta Earthquake	7	0
1992 Cape Mendocino Earthquake	2	0
1992 Landers/Big Bear Earthquake	3	0
1993 Guam Earthquake	3	0
1994 Northridge Earthquake	19	0
1995 Manzanillo Earthquake	4	0
1999 Kocaeli Turkey Earthquake	1	0
2010 Baja California Earthquake	2	0

Table 2

Summary of parameters used in the transformation of PGA values.

c_h	b_h	$\tilde{\beta}$	μ_g^{FEM}	σ_g^{FEM}	μ_j^{FEM}	σ_j^{FEM}	ρ
0.041	1	0.2	0.846	0.746	0.354	0.743	0.924

value of the floor acceleration and its uncertainty can be determined.

2. Step 2: transformation of the PGA of the floor acceleration to the K-K NPP free surface PGA with the bi-variate normal distribution model established from the 62 FEM simulation results.

First, the lognormal assumption of the marginal distributions of the PGA values is checked for both floor accelerations and free field

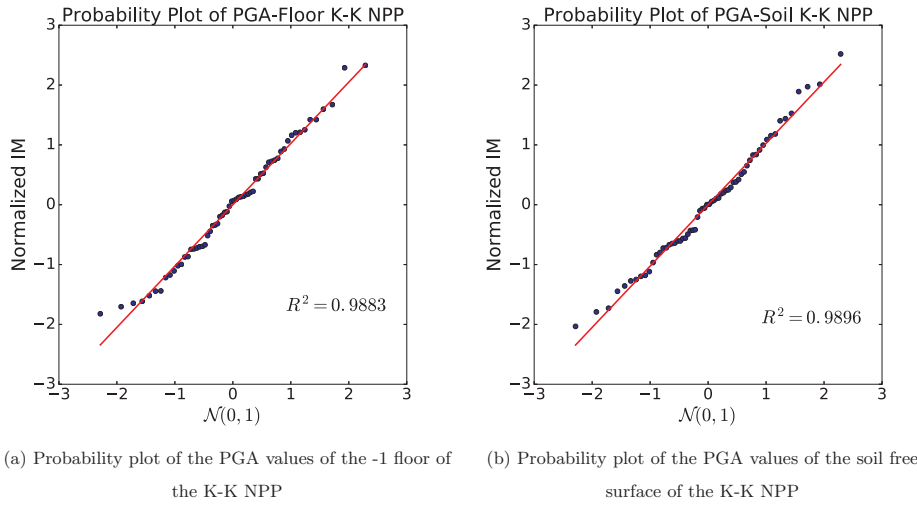


Fig. 11. Probability plot of the normalized PGA values of the -1 floor and the normalized PGA values on the soil free surface of the K-K NPP.

accelerations of the K-K NPP. The values of PGAs are obtained from 62 FEM simulation results. The medians μ and logarithmic Stds σ of the assumed lognormal distributions are computed and listed in Table 2. The $\ln(\text{PGA})$ values are normalized with respect to the medians and Stds and compared with $N(0, 1)$ in Fig. 11 to verify the lognormal assumption.

From the results of the probability plots, it can be concluded that the lognormal assumption for both α_g^{FEM} and α_f^{FEM} can be considered acceptable. Additionally, from Table 2 it can be observed that the median of the soil PGA μ_g^{FEM} is larger than the median of the floor PGA μ_f^{FEM} : this is due to the fact that the LVSG is located on the -1 floor (7.2m below the ground level) in the K-K NPP.

Furthermore, the transformed PGA values $\alpha_{s \rightarrow t}$ on the K-K NPP free surface and the relevant uncertainty $\beta_{s \rightarrow t}$ due to the transformation are computed with Eqs. (25) and (26). The transformed median PGA values $\alpha_{s \rightarrow t}$ are plotted in Fig. 12. We can see an increase of the PGA values after the transformation process due to the low elevation of the target LVSG in the K-K NPP. In Fig. 12, a linear tendency can be observed for some $(\alpha_{s \rightarrow t}, \alpha_s)$ data values. This is because their corresponding LVSGs have the same elevations h (in particular $h = 0$), so that their amplification factors in the transformation step 1 are the same. In addition, as $\beta_{s \rightarrow t}$ calculated with Eq. (26) is independent of α_s , it stays the same for all 78 \mathbf{z}_s data, with $\beta_{s \rightarrow t} = 0.299$.

In the end, the transformed $\alpha_{s \rightarrow t}$ are plugged into Eq. (27) to compute the likelihood function $L(\mathbf{z}_s | \hat{A}_m)$. We also justify the

application of the fractional x^i value 0.5 for the potential failure case of El Centro steam plant. It can be regarded as two realizations of earthquake observations, with one failure and one survival. Then, the likelihood function established by the two realizations should be normalized to one observation by the square-root operation:

$$L(\mathbf{z}_{\text{El-Centro}} | \hat{A}_m) = \sqrt{[P_f(\alpha_{\text{El-Centro}})] [1 - P_f(\alpha_{\text{El-Centro}})]} = [P_f(\alpha_{\text{El-Centro}})]^{0.5} [1 - P_f(\alpha_{\text{El-Centro}})]^{0.5} \quad (30)$$

The same procedure is also used by EPRI (2017). As a result, the total likelihood function is computed: $L(\mathbf{z} | \hat{A}_m) = L(\mathbf{z}_t | \hat{A}_m) L(\mathbf{z}_s | \hat{A}_m)$

3.4. Posterior fragility curve

The posterior distribution $f^{\text{post}}(\hat{A}_m | \mathbf{z})$ is obtained based on the prior distribution and the likelihood function, which are calculated in Section 3.2 and Section 3.3, respectively. MCMC is used to generate 10,000 samples from $f^{\text{post}}(\hat{A}_m | \mathbf{z})$. A lognormal distribution is approximated for $f^{\text{post}}(\hat{A}_m | \mathbf{z})$ with the median and logarithmic Std of the generated 10,000 samples. The MCMC sampling of $f^{\text{post}}(\hat{A}_m | \mathbf{z})$ and the approximated lognormal distribution are shown in Fig. 13. The posterior distribution of \hat{A}_m has the median $A_m^{\text{post}} = 2.70g$ and the associated uncertainty $\beta_U^{\text{post}} = 0.176$. The comparison of fragility parameters of the LVSG before and after Bayesian updating is reported in Table 3.

The posterior median and mean fragility curves are computed with Eqs. (2) and (6), respectively. The 0.05 and 0.95 non-exceedance

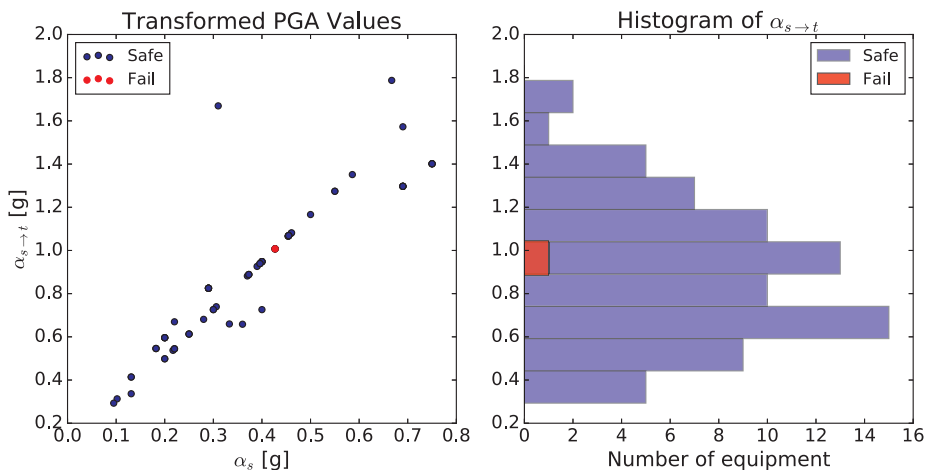


Fig. 12. PGA values $\alpha_{s \rightarrow t}$ after the transformation into K-K NPP free surface.

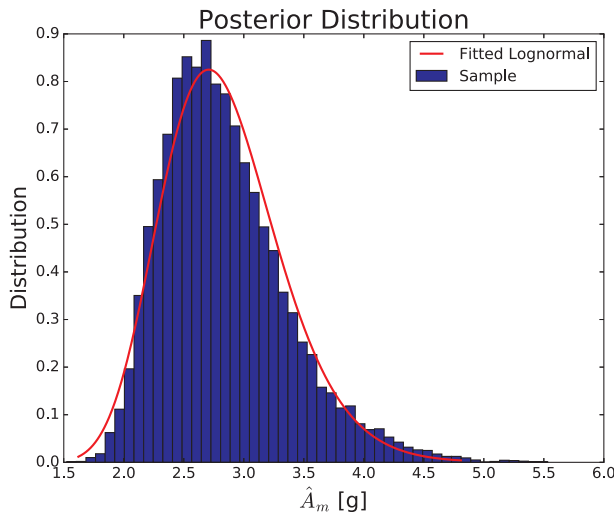


Fig. 13. MCMC sampling of the posterior distribution of \hat{A}_m .

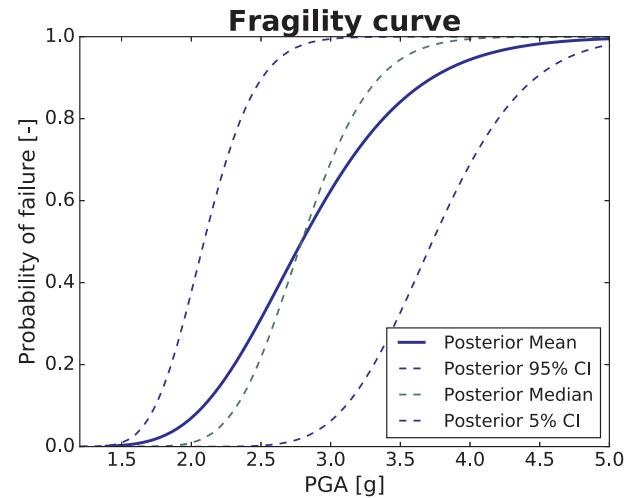


Fig. 14. Posterior fragility curves and its confidence intervals.

confidence intervals are calculated with Eq. (4). The computed fragility curves and the associated confidence intervals are shown in Fig. 14. Compared to the prior fragility parameters, the median capacity A_m increases after Bayesian updating, because few failure cases (only one) have been observed in the post-earthquake inspection for the LVSG. The epistemic uncertainty β_U is reduced due to the supplementary information from the observational data. Relatively large confidence bounds have been shown in Fig. 14, since despite the Bayesian updating, the value of β_U^{post} (0.176) is still larger than the aleatory uncertainty $\beta_R = 0.145$. Nevertheless, the HCLPF capacity, computed by Eq. (5), is largely increased after the Bayesian updating, mainly due to the reduction of the epistemic uncertainty.

Furthermore, the influence of the selected value of β_U^{prior} on the posterior fragility curve is investigated. Different values of β_U^{prior} varying from 0.1 to 0.4 are taken for the prior \hat{A}_m distributions. With the same SQUG observational data, the likelihood function remains the same. Posterior distributions of \hat{A}_m are computed and plotted in Fig. 15. It can be observed from Fig. 15 that $\beta_U^{prior} = 0.4$ reveals in fact a relatively large uncertainty of the median capacity compared to the likelihood function $L(z|\hat{A}_m)$ provided by the observational data. Therefore, the contribution of $L(z|\hat{A}_m)$ to the posterior distribution is dominant if $\beta_U^{prior} = 0.4$. On the other hand, if one is very certain about the median capacity estimated from the numerical simulations (i.e. $\beta_U^{prior} = 0.1$), the influence of $L(z|\hat{A}_m)$ on the posterior distribution is not that evident: the posterior median increases slightly in comparison with the prior median, whereas the value of β_U is hardly modified. For $\beta_U^{prior} = 0.2$ and 0.3, the posterior distribution is a trade-off between the prior distribution and the likelihood function, which is a reasonable outcome from the Bayes' theorem.

Finally, we study the influence of the uncertain observational data of the El-Centro steam plant. Other than the observational value $x^i = 0.5$ applied in the previous sections, the likelihood function is also computed with $x^i = 0$ and $x^i = 1$ for the El-Centro steam plant. The posterior distributions with different levels of uncertainty on the El-Centro observational data are plotted in Fig. 16. From the results of Fig. 16, a complete failure of the LVSG $x^i = 1$ in the El-Centro steam plant makes decrease the posterior \hat{A}_m , since the equipment is more

fragile according to the observational results. On the contrary, the LVSG is more resistant if $x^i = 0$ for the El-Centro steam plant, so that an increase of the posterior \hat{A}_m can be observed.

3.5. Discussion

The proposed Bayesian framework has been applied to the K-K NPP to compute the fragility curve of a low-voltage switchgear. Some assumptions made in the methodology and in the application are discussed as follows:

1. As a first application of the proposed Bayesian framework, the assumption of linear material constitutive law has been applied in the FEM simulation of the K-K NPP model, since the building reveals to be very rigid and response remains mainly linear under the NCO earthquake (IAEA, 2013). However, it should be noted that the results computed from the linear material assumption do not provide best estimates of the structural responses for the higher load levels. A further step is to apply the proposed method to a nonlinear structure model but with a smaller number of degrees of freedom.
2. The derivation of the likelihood function using the two-step transformation depends on the assumptions of normality or lognormality. These assumptions are applied so that an analytical form of the final likelihood function can be derived. In real applications, if synthetic ground motions are generated from an earthquake scenario, the IMs at the free surface can be considered lognormally distributed. Nevertheless, when the ground motions, which are used to perform FEM simulations, are selected by the adaptive training, the initial lognormal distribution is disturbed. Therefore, one should be careful to check the lognormality of the free surface PGA and the floor PGA from FEM simulations for the joint lognormal model in the second stage of the PGA transformation. Despite this fact, it should be mentioned that, theoretically, any analytical statistical distributions can be used to compute the transformed free surface PGA value and to quantify the uncertainty in the two-step transformation. However, no analytical solution exists and Monte-Carlo simulations are needed to calculate the transformed PGA value and its uncertainty. In addition, the value of $P_f(\alpha)$ cannot be computed from a CDF of a normal distribution, if the uncertainty of the two-step transformation is not assumed lognormally distributed. Again, in this case, the only way to evaluate this $P_f(\alpha)$ is the Monte-Carlo simulation combined with the Nataf transformation (Eq. (7.2.2) in Ditlevsen and Madsen (2005)).
3. It is worth mentioning that the likelihood function computed from Eq. (27) is, in fact, the ‘mean likelihood’ considering the PGA transformation uncertainty $\beta_{s \rightarrow t}$, analogue to the capacity in the

Table 3
Posterior fragility parameters for the LVSG in K-K NPP and comparison to the prior parameters

β_R	A_m^{prior}	β_U^{prior}	A_{HCLPF}^{prior}	A_m^{post}	β_U^{post}	A_{HCLPF}^{post}
0.145	2.46g	0.4	1.00g	2.70g	0.176	1.59g

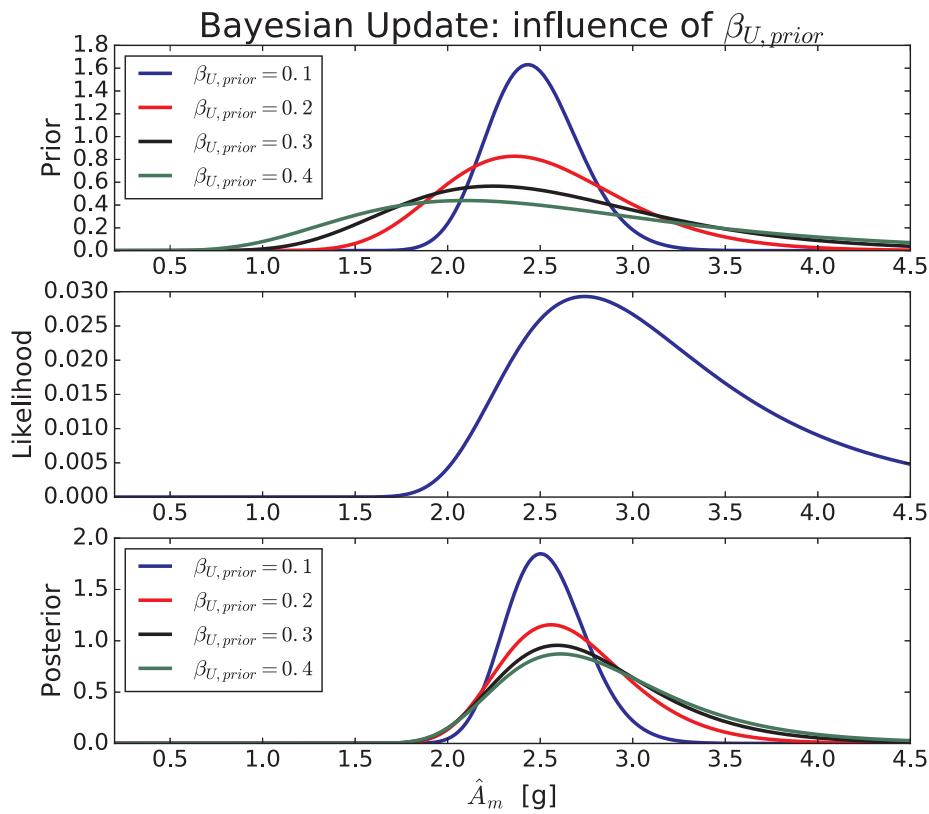


Fig. 15. Influence of the choice of β_U^{prior} on the posterior fragility curve. Upper: prior distributions of \hat{A}_m with different β_U^{prior} . Middle: likelihood function. Lower: posterior distributions of \hat{A}_m with different β_U^{prior} .

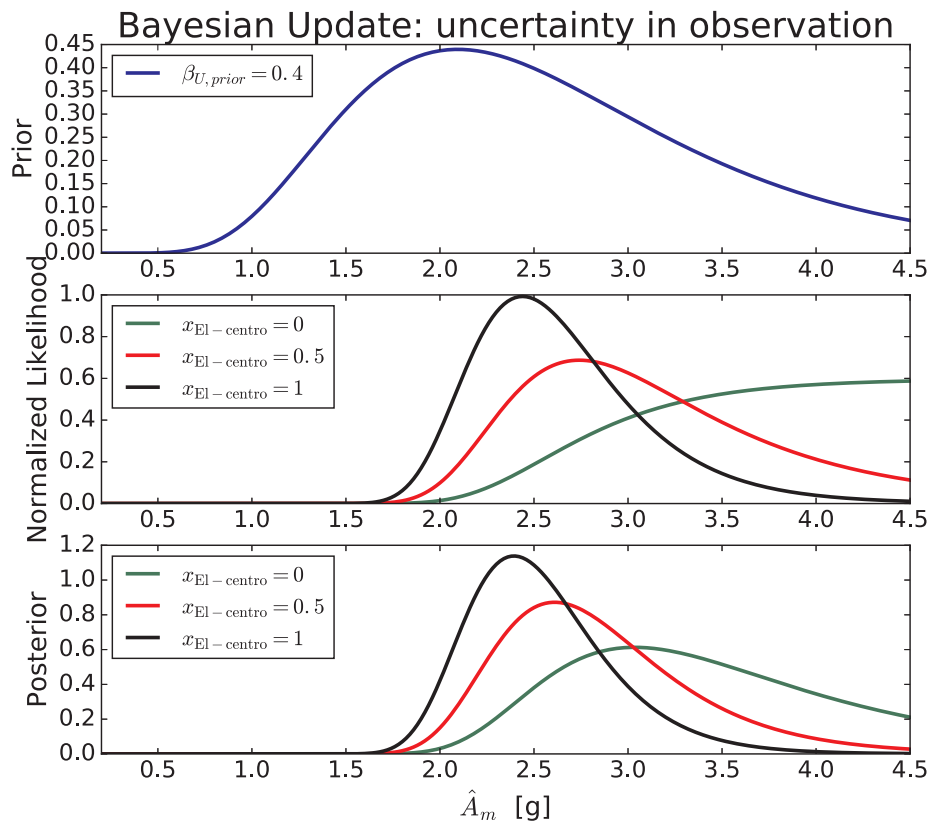


Fig. 16. Influence of observational data of El-Centro steam plant. Upper: prior distributions of \hat{A}_m . Middle: likelihood function with different x^l values for the El-Centro steam plant. Lower: posterior distributions of \hat{A}_m .

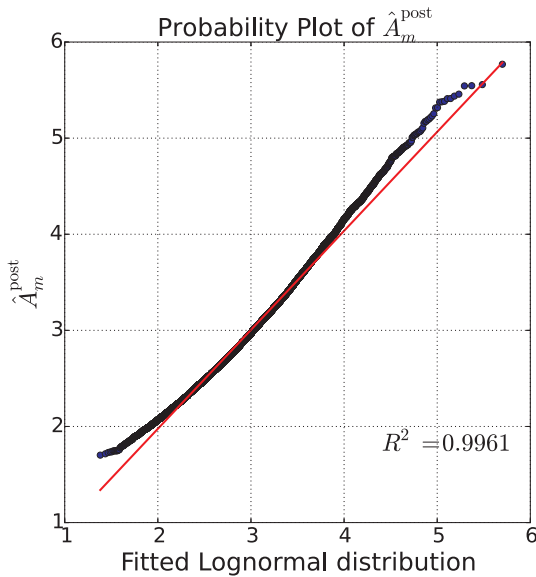


Fig. 17. Probability plot of the true posterior distribution against the fitted lognormal distribution.

mean fragility curve of Eq. (6). In a similar way of defining the non-exceedance confidence interval of fragility curves (Eq. (4)), the confidence interval of the likelihood function \tilde{L} can also be obtained:

$$\begin{aligned} \tilde{L}(\mathbf{z}_s, \gamma' | \hat{A}_m) &= \prod_{i=1}^{n_s} \left[\Phi\left(\frac{\ln \alpha_{s \rightarrow t}^i + \beta_{s \rightarrow t} \Phi^{-1}(\gamma') - \ln \hat{A}_m}{\beta_R}\right) \right]^{x^i} \\ &\quad \left[1 - \Phi\left(\frac{\ln \alpha_{s \rightarrow t}^i + \beta_{s \rightarrow t} \Phi^{-1}(\gamma') - \ln \hat{A}_m}{\beta_R}\right) \right]^{1-x^i} \end{aligned} \quad (31)$$

where $\gamma' \in [0, 1]$ is the confidence level related to the transformed PGA value. This latter modeling leads to the definition of a two-level confidence interval of the posterior fragility curves. The first level is linked to the uncertainty β_U of the capacity, whereas the second

level is associated to the uncertainty $\beta_{s \rightarrow t}$ of the PGA transformation. Further explorations are necessary for this two-level confidence interval modeling.

- In the Bayesian updating, as the prior lognormal distribution of the capacity is not a conjugate prior of the likelihood function computed by Eq. (28), the posterior distribution $f^{\text{post}}(\hat{A}_m | \mathbf{z})$ cannot remain lognormal. A lognormal fit is, thus, applied to $f^{\text{post}}(\hat{A}_m | \mathbf{z})$, to be able to provide the values of A_m^{post} and β_U^{post} , which are widely used in engineering practice. In order to quantify the quality and the influence of the lognormal fit, firstly, the probability plot of the true posterior distribution against the fitted lognormal distribution is plotted. The probability plot, shown in Fig. 17, compares the quantiles between the true posterior and the fitted lognormal distribution. It can be observed that the quality of the fitting, in terms of the quantile, is acceptable for \hat{A}_m^{post} between 2g and 4g. But the fitting quality is not very satisfactory at the two ends of the distribution. Furthermore, we investigate the influence of the lognormal fitting on the final fragility curves and the confidence intervals. The posterior fragility curves and the confidence intervals can be computed with the true posterior distribution. The computation of the confidence intervals is achieved by identifying the quantiles of the true distribution of \hat{A}_m^{post} :

$$\tilde{P}_f^{\text{True}}(\alpha, \gamma) = \Phi\left(\frac{\ln \alpha - \ln \tilde{A}_m^{\text{True}}(\gamma)}{\beta_R}\right) \quad (32)$$

where $\tilde{A}_m^{\text{True}}(\gamma)$ denotes the $(1-\gamma)$ quantile of the true \hat{A}_m^{post} . As for the mean fragility curve, no analytical form can be derived when the lognormality of \hat{A}_m^{post} is not satisfied. Monte-Carlo simulation has to be used to sample the aleatory uncertainty term ε_R . The computation of the mean fragility curve with the true posterior distribution is as follows:

- Sample N^{post} aleatory uncertainty term ε_R , following $\text{LogN}(1, \beta_R^2)$, where N^{post} is the number of the MCMC samples of \hat{A}_m^{post} .
- Evaluate the probability of failure with a given value of α :

$$P_f^{\text{mean, true}}(\alpha) = \frac{1}{N^{\text{post}}} \mathbf{I}[\alpha > \hat{A}_m^{\text{post}} \varepsilon_R] \quad (33)$$

where $\mathbf{I}[\cdot]$ is the indicator function.

- Select another value of α and return to the step (i).

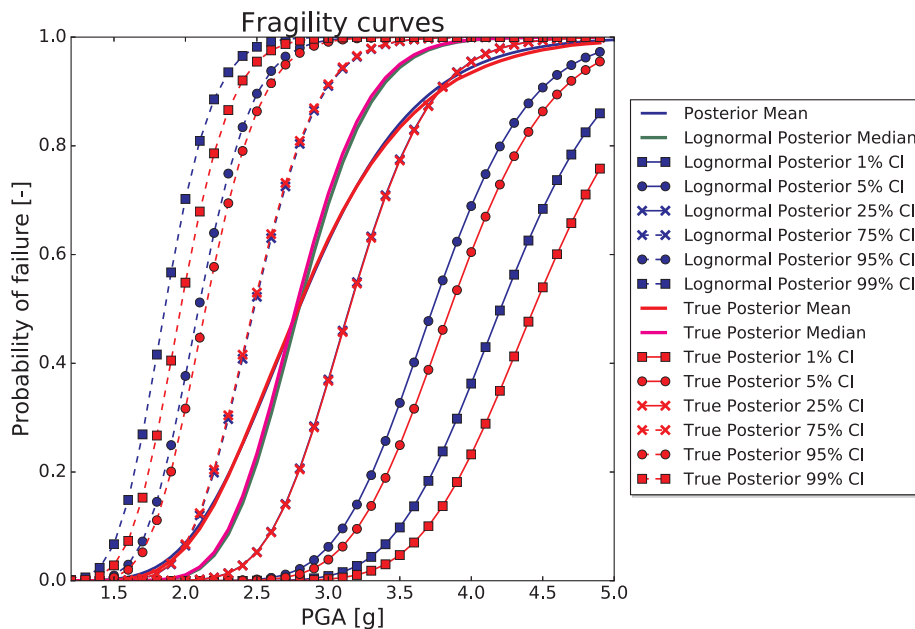


Fig. 18. Comparison of fragility curves computed by the true posterior distribution and the fitted lognormal distribution.

The posterior fragility curves computed by Eqs. (32) and (33), shown in Fig. 18, are compared to the ones obtained with the lognormal fitting. It can be observed that the fragility curves do not show evident differences when γ is between 25% and 75%. However, when γ becomes very low (<5%) or very high (>95%), discrepancies start to appear. This phenomenon is due to the difference of the quantiles at the two ends of the two distributions. Nevertheless, good coherence is obtained for the mean fragility curves: the influence of the lognormal fitting on the mean fragility curve can be neglected in the considered case study.

4. Conclusion

In this paper, a Bayesian updating framework is proposed for considering different sources of information, including numerical simulations, damage data and reference critical values informed from the literature, in the computations of seismic fragility curves. In the framework, the results from numerical simulations are used to determine the prior parameters of the fragility curves. Damage data are, then, integrated to compute the likelihood function for the Bayesian updating. Finally, MCMC is applied to sample the posterior distribution of the updated equipment capacity.

The main contributions of the work are: (i) An ANN adaptive training algorithm is proposed for a more intelligent experimental design to conduct FEM simulations. The adaptive training is based on the prediction uncertainty computed by the delta method. Relying on a relatively large initial population of ground motions, the adaptive ANN aims to select a representative subset of ground motions, which can ensure the performance of the ANN over the whole population; (ii) A method to construct the likelihood function is proposed to deal with existing damage data from the SQUG database. For post-earthquake observational data, the computation of the likelihood function is achieved by estimating an amplification factor and by assuming a joint lognormal distribution between floor PGA values and free surface PGA values. This latter lognormal assumption allows providing an analytical form of the final likelihood function.

The methodology is applied to evaluate the fragility curve of a low-voltage switchgear in a Japanese nuclear power plant Kashiwazaki-Kariwa. The application of the adaptive ANN training provides an improved design of experiments, in which more failure cases have appeared in the FEM simulation results. Then, the construction of the likelihood function with SQUG damage data is realized by the proposed two-stage PGA transformation. The uncertainty related to the transformation is also quantified and integrated in the computation of the likelihood function. Compared to the prior fragility parameters, the posterior capacity of the low-voltage switchgear has increased due to few observed failure cases, whereas the epistemic uncertainty is largely reduced with the additional information from the damage data. These two aspects give rise to a higher value of the high confidence low probability of failure capacity of the studied low-voltage switchgear. The sensitivity analysis has shown that, in this case study, the contribution of the likelihood function to the posterior estimations is predominant, if the assumed value of β_U^{prior} is larger than 0.2.

Acknowledgement

The authors want to thank Hugo Jadot for his comments and suggestions to this work. We would like also to thank the anonymous reviewers for their valuable comments and suggestions. The financial support from the project NARSIS H2020 (New Approach to Reactor Safety ImprovementS, Horizon 2020) with the grant agreement No. 755439 is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.nucengdes.2018.08.016>.

References

- Basu, P.C., Ravindra, M.K., Mihara, Y., 2017. Component fragility for use in PSA of nuclear power plant. *Nucl. Eng. Des.* 323, 209–227. <https://doi.org/10.1016/j.nucengdes.2016.10.018>.
- Beck, J.L., Taflanidis, A.A., 2013. Prior and posterior robust stochastic predictions for dynamical systems using probability logic. *Int. J. Uncertainty Quantif.* 3, 271–288. <https://doi.org/10.1615/Int.J.UncertaintyQuantification.2012003641>.
- Biasio, M.D., Grange, S., Dufour, F., Allain, F., Petre-Lazar, I., 2015. Intensity measures for probabilistic assessment of non-structural components acceleration demand. *Earthquake Eng. Struct. Dyn.* 44, 2261–2280. <https://doi.org/10.1002/eqe.2582>.
- Bishop, C.M., 1995. *Neural Networks for Pattern Recognition*. Oxford University Press.
- Buratti, N., Minghini, F., Ongaretto, E., Savoia, M., Tullini, N., 2017. Empirical seismic fragility for the precast RC industrial buildings damaged by the 2012 Emilia (Italy) earthquakes. *Earthquake Eng. Struct. Dyn.* 48, 2317–2335. <https://doi.org/10.1002/eqe.2906>.
- Campbell, K.W., Bozorgnia, Y., 2008. NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s. *Earthquake Spectra* 24, 139–171. <https://doi.org/10.1193/1.2857546>.
- Chryssolouris, G., Lee, M., Ramsey, A., 1996. Confidence interval prediction for neural network models. *IEEE Trans. Neural Netw.* 7, 229–232. <https://doi.org/10.1109/72.478409>.
- Clouteau, D. Miss 6.5: Manuel Utilisateur. Lab MSSMat, CentraleSupélec; 2005. A software in earthquake engineering and structural dynamics, http://old-www.ecp.fr/mssmat/moyens/moyens_techniques_logiciels/miss/presentation_de_miss.
- CodeAster, Code Analyses des Structures et Thermomécanique pour des Etudes et des Recherches. EDF Group. Opensource Finite Element code, <http://www.code-aster.org>.
- Cornell, C.A., Jalayer, F., Hamburger, R.O., Foutch, D.A., 2002. Probabilistic basis for 2000 SAC federal emergency management agency steel moment frame guidelines. *J. Struct. Eng.* 128, 526–533. [https://doi.org/10.1061/\(ASCE\)0733-9445\(2002\)128:4\(526\)](https://doi.org/10.1061/(ASCE)0733-9445(2002)128:4(526)).
- Ditlevsen, O., Madsen, H., 2005. *Structural Reliability Methods*.
- Echard, B., Gayton, N., Lemaire, M., 2011. AK-MCS: an active learning reliability method combining Kriging and Monte Carlo simulation. *Struct. Saf.* 33, 145–154. <https://doi.org/10.1016/j.strusafe.2011.01.002>.
- Ellingwood, B.R., Kinali, K., 2009. Quantifying and communicating uncertainty in seismic risk assessment. *Struct. Saf.* 31, 179–187. <https://doi.org/10.1016/j.strusafe.2008.06.001>.
- EPRI, 1991. *Generic Seismic Ruggedness of Power Plant Equipment (Revision 1)*. Technical Report; Electric Power Research Institute EPRI, Palo Alto, CA.
- EPRI, 1994. *Methodology for developing seismic fragilities*. Technical Report; Electric Power Research Institute EPRI, Palo Alto, CA. Report TR-103959.
- EPRI, 2007. EPRI independent peer review of the TEPCO seismic shutdown and evaluation of the Kashiwazaki-Kariwa nuclear power plants. Technical Report; Electric Power Research Institute EPRI, Palo Alto, CA. Report TR-1016317.
- EPRI, *Seismic Fragility Applications Guide Update*. Technical Report; Electric Power Research Institute EPRI, Palo Alto, CA; 2009. Report TR-1019200.
- EPRI, 2014. *Assessment of the Use of Experience Data to Develop Seismic Fragilities*. Technical Report; Electric Power Research Institute EPRI, Palo Alto, CA. Report 3002002933.
- EPRI, SQUG seismic experience database. Electric Power Research Institute EPRI, Palo Alto, CA; 2016. <http://esqug.epri.com/>.
- EPRI, 2017. *Updated Equipment Seismic Capacities from Experience Data for Use in the Fragility Calculations*. Technical Report; Electric Power Research Institute EPRI, Palo Alto, CA. Report 3002011627.
- Ferrario, E., Pedroni, N., Zio, E., Lopez-Caballero, F., 2017. Bootstrapped artificial neural networks for the seismic analysis of structural systems. *Struct. Saf.* 67, 70–84. <https://doi.org/10.1016/j.strusafe.2017.03.003>.
- Gardoni, P., Mosalam, K.M., Kiureghian, A.D., 2009. Probabilistic seismic demand models and fragility estimates for RC bridges. *J. Earthquake Eng.* 79–106. <https://doi.org/10.1080/13632460309350474>.
- Gidaris, I., Taflanidis, A.A., Mavroeidis, G.P., 2015. Kriging metamodeling in seismic risk assessment based on stochastic ground motion models. *Earthquake Eng. Struct. Dyn.* 44, 2377–2399. <https://doi.org/10.1002/eqe.2586>.
- Hastings, W.K., 1970. Monte carlo sampling methods using Markov chains and their applications. *Biometrika* 57, 97–105. <https://doi.org/10.2307/2334940>.
- Huang, Y.N., Whittaker, A., Luco, N., 2011. A probabilistic seismic risk assessment procedure for nuclear power plant: (i) Methodology. *Nucl. Eng. Des.* 241, 3966–4003. <https://doi.org/10.1016/j.nucengdes.2011.06.051>.
- Hurtado, J.E., Alvarez, D.A., 2001. Neural-network-based reliability analysis: a comparative study. *Comput. Methods Appl. Mech. Eng.* 191, 113–132. [https://doi.org/10.1016/S0045-7825\(01\)00248-1](https://doi.org/10.1016/S0045-7825(01)00248-1).
- IAEA, 2013. *Review of Seismic Evaluation Methodologies for Nuclear Power Plants Based on a Benchmark Exercise*. Technical Report; International Atomic Energy Agency.
- Jaiswal, K., Wald, D., D'Ayala, D., 2011. Developing empirical collapse fragility functions for global building types. *Earthquake Spectra* 27, 775–795. <https://doi.org/10.1193/1.3606398>.
- Jalayer, F., Iervolino, I., Manfredi, G., 2010. Structural modeling uncertainties and their influence on seismic assessment of existing RC structures. *Struct. Saf.* 32, 220–228. <https://doi.org/10.1016/j.strusafe.2010.02.004>.
- Jeon, J.S., Mangalathu, S., Song, J., DesRoches, R., 2017. Parameterized seismic fragility curves for curved multi-frame concrete box-girder bridges using Bayesian parameter estimation. *J. Earthquake Eng.* <https://doi.org/10.1080/13632469.2017.1342291>.

- Jones, D.R., Schonlau, M., Welch, W.J., 1998. Efficient global optimization of expensive black-box functions. *J. Global Optimiz.* 13, 455–492. <https://doi.org/10.1023/A:1008306431147>.
- Kennedy, R., Cornell, C., Campell, R., Kaplan, S., Perla, H., 1980. Probabilistic seismic safety study of an existing nuclear power plant. *Nucl. Eng. Des.* 59, 315–338. [https://doi.org/10.1016/0029-5493\(80\)90203-4](https://doi.org/10.1016/0029-5493(80)90203-4).
- Koutsourelakis, P.S., 2010. Assessing structural vulnerability against earthquakes using multi-dimensional fragility surfaces: A Bayesian framework. *Probab. Eng. Mech.* 25, 49–60. <https://doi.org/10.1016/j.probengmech.2009.05.005>.
- Kwag, S., Lee, J.M., Oh, J., Ryu, J.S., 2014. Development of system design and seismic performance evaluation for reactor pool working platform of a research reactor. *Nucl. Eng. Des.* 266, 199–213. <https://doi.org/10.1016/j.nucengdes.2013.10.025>.
- Lagaros, N.D., Tsompanakis, Y., Psarropoulos, P.N., Georgopoulos, E.C., 2009. Computationally efficient seismic fragility analysis of geostructures. *Comput. Struct.* 87, 1195–1203. <https://doi.org/10.1016/j.compstruc.2008.12.001>.
- Mai, C.V., Spiridonakos, M.D., Chatzi, E.N., Sudret, B., 2016. Surrogate modelling for stochastic dynamical systems by combining narx models and polynomial chaos expansions. *Int. J. Uncertainty Quant.* 6, 419–430. <https://doi.org/10.1615/Int.J.UncertaintyQuantification.v6.i4>.
- Mangalathu, S., Heo, G., Jeon, J.S., 2018. Artificial neural network based multi-dimensional fragility development of skewed concrete bridge classes. *Eng. Struct.* 162, 166–176. <https://doi.org/10.1016/j.engstruct.2018.01.053>.
- Newmark, N.M., Hall, W.J., 1978. Development of criteria for seismic review of selected nuclear power plants. Technical Report; Nuclear Regulatory Commission. NUREG/CR-0098.
- Noh, H.Y., Kiremidjian, A., Ceferino, L., So, E., 2017. Bayesian updating of earthquake vulnerability functions with application to mortality rates. *Earthquake Spectra* 33, 1173–1189. <https://doi.org/10.1193/081216EQS133M>.
- Padgett, J.E., DesRoches, R., 2008. Methodology for the development of analytical fragility curves for retrofitted bridges. *Earthquake Eng. Struct. Dyn.* 37, 1157–1174. <https://doi.org/10.1002/eqe.801>.
- Papadopoulos, V., Giovanis, D.G., Lagaros, N.D., Papadrakakis, M., 2012. Accelerated subset simulation with neural networks for reliability analysis. *Comput. Methods Appl. Mech. Eng.* 223–234, 70–80. <https://doi.org/10.1016/j.cma.2012.02.013>.
- Pedroni, N., Zio, E., 2017. An adaptive Metamodel-based subset importance sampling approach for the assessment of the functional failure probability of a thermal-hydraulic passive system. *Appl. Math. Model.* <https://doi.org/10.1016/j.apm.2017.04.003>.
- Reed, R.D., Marks, R.J., 1999. *Neural Smoothing*. MIT Press.
- Rivals, I., Personnaz, L., 2000. Construction of confidence intervals for neural networks based on least squares estimation. *Neural Netw.* 13, 463–484. [https://doi.org/10.1016/S0893-6080\(99\)00080-5](https://doi.org/10.1016/S0893-6080(99)00080-5).
- Rossetto, T., Elnashai, A., 2003. Derivation of vulnerability functions for european-type RC structures based on observational data. *Eng. Struct.* 25, 1241–1263. [https://doi.org/10.1016/S0141-0296\(03\)00060-9](https://doi.org/10.1016/S0141-0296(03)00060-9).
- Rota, M., Pennab, A., Strobba, C.L., 2008. Processing Italian damage data to derive typological fragility curves. *Soil Dyn. Earthquake Eng.* 28, 933–947. <https://doi.org/10.1016/j.soildyn.2007.10.010>.
- Rumelhart, D., Hinton, G., Williams, R., 1986. *Learning Internal Representations by Error Propagation in Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. The MIT Press.
- SalomeMeca, Opensource platform for numerical simulation including pre- and post-processing. EDF Group. <https://www.code-aster.org/spip.php?article294>.
- Shinozuka, M., Feng, M.Q., Lee, J., Naganuma, T., 2000. Statistical analysis of fragility curves. *J. Eng. Mech.* 126, 1224–1231. [https://doi.org/10.1061/\(ASCE\)0733-9399\(2000\)126:12\(1224\)](https://doi.org/10.1061/(ASCE)0733-9399(2000)126:12(1224)).
- Singhal, A., Kiremidjian, A.S., 1998. Bayesian updating of fragilities with application to RC frames. *J. Struct. Eng.* 124, 922–929. [https://doi.org/10.1061/\(ASCE\)0733-9445\(1998\)124:8\(922\)](https://doi.org/10.1061/(ASCE)0733-9445(1998)124:8(922)).
- Starck, R.G., Thomas, G.G., 1990. Overview of SQUG generic implementation procedure (GIP). *Nucl. Eng. Des.* 123, 225–231. [https://doi.org/10.1016/0029-5493\(90\)90241-0](https://doi.org/10.1016/0029-5493(90)90241-0).
- Straub, D., Kiureghian, A.D., 2008. Improved seismic fragility modeling from empirical data. *Struct. Saf.* 30, 320–336. <https://doi.org/10.1016/j.strusafe.2007.05.004>.
- Towashiraporn, P., 2004. *Building Seismic Fragilities using Response Surface Metamodels*. Georgia Institute of Technology (Ph.D. thesis).
- Wang, Z., Pedroni, N., Zentner, I., Zio, E., 2018. Seismic fragility analysis with artificial neural networks: Application to nuclear power plant equipment. *Eng. Struct.* 162, 213–225. <https://doi.org/10.1016/j.engstruct.2018.02.024>.
- Xiao, N.C., Zuo, M.J., Zhou, C., 2018. A new adaptive sequential sampling method to construct surrogate models for efficient reliability analysis. *Reliab. Eng. Syst. Safe* 169, 330–338. <https://doi.org/10.1016/j.res.2017.09.008>.
- Yoshida, N., Kobayashi, S., Suetomi, I., Miura, K., 2002. Equivalent linear method considering frequency dependent characteristics of stiffness and damping. *Soil Dyn. Earthquake Eng.* 22, 205–222. [https://doi.org/10.1016/S0267-7261\(02\)00011-8](https://doi.org/10.1016/S0267-7261(02)00011-8).
- Zentner, I., 2010. Numerical computation of fragility curves for NPP equipment. *Nucl. Eng. Des.* 240, 1614–1621. <https://doi.org/10.1016/j.nucengdes.2010.02.030>.
- Zentner, I., 2014. A procedure for simulating synthetic accelerograms compatible with correlated and conditional probabilistic response spectra. *Soil Dyn. Earthquake Eng.* 63, 226–233. <https://doi.org/10.1016/j.soildyn.2014.03.012>.
- Zentner, I., Gündel, M., Bonfils, N., 2017. Fragility analysis methods: Review of existing approaches and application. *Nucl. Eng. Des.* 323, 245–258. <https://doi.org/10.1016/j.nucengdes.2016.12.021>.
- Zio, E., 2006. A study of the bootstrap method for estimating the accuracy of artificial neural networks in predicting nuclear transient processes. *IEEE Trans. Nucl. Sci.* 53, 1460–1478. <https://doi.org/10.1109/TNS.2006.871662>.