Computation of Seismic Fragility Curves Using Artificial Neural Network Metamodels

Zhiyi Wang, Nicola Pedroni, Irmela Zentner, Enrico Zio

▶ **To cite this version:**

Zhiyi Wang, Nicola Pedroni, Irmela Zentner, Enrico Zio. Computation of Seismic Fragility Curves Using Artificial Neural Network Metamodels. 12th International Conference on Structural Safety and Reliability (ICOSSAR 2017), Aug 2017, Vienne, Austria. pp.1525-1534. hal-01716960

HAL Id: hal-01716960
https://hal.science/hal-01716960

Submitted on 9 Mar 2018

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.

Computation of Seismic Fragility Curves Using Artificial Neural Network Metamodels

Zhiyi Wang[a,b,c], Nicola Pedroni[c], Irmela Zentner[a,b], Enrico Zio[c,d]
[a]EDF R&D Lab Saclay, France
[b]IMSIA, UMR 9219 CNRS-EDF-CEA-ENSTA ParisTech, France
[c]Chair on Systems Science and Energetic Challenge, CentraleSupélec, Université Paris-Saclay, France
[d]Energy Department, Politecnico di Milano, Italy

Abstract: In earthquake engineering, the fragility curve is defined as the conditional probability of failure of a structure, or its critical components, at given values of seismic intensity measures (IMs). For simulation-based fragility curve estimations, this conditional probability of failure is usually computed with log-normal assumption. The artificial neural network (ANN) is used to improve the computational efficiency of the simulation-based fragility analysis. An ANN metamodel is built from 100 finite element soil-structure-interaction simulation results. The most relevant IMs are selected based on semi-partial correlation coefficients. The ANN metamodel is trained with the selected IMs, and a large number of Monte Carlo simulations are performed with this metamodel. Fragility curves are computed with both parametric (log-normal model) and non-parametric methods for the estimation of the risk of failure of an electrical cabinet in a reactor building studied in the framework of the KARISMA benchmark.

1 Introduction

In the seismic probabilistic risk assessment methodology, fragility curves are computed as the conditional probability of failure of a structure, or critical components, for a given value of a seismic intensity measure (IM) such as the peak ground acceleration (PGA). The total probability of failure of the plant is computed by the convolution of the fragility curves with the hazard curves, combined in fault tree and event tree analysis [4]. Fragility curves can be calculated by Monte Carlo simulation, to propagate the uncertainties in seismic ground motions, structural material properties, etc. These uncertainties are categorized into two groups [6]: aleatory uncertainties, which reflect the inherent randomness of variables or stochastic processes, and epistemic uncertainties due to lack of knowledge about the model and its parameters.
Mathematically, fragility curves are calculated as the conditional probability that the damage measure (DM) exceeds a critical threshold, for a given seismic IM:

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

where y is the DM, such as inter-story drift, y_{crit} is the failure threshold and α represents the seismic IM. This conditional probability can be evaluated pointwise for different α values. However, this requires a large set of numerical simulations with finite element method (FEM), and it is often difficult and computationally very expensive to perform such analyses due to the complexity of the mechanical model.
The log-normal fragility model was proposed in [6] to reduce the computational cost. Less

numerical simulations are required with this model because only two parameters, the median capacity A_m and the logarithmic standard deviation β, need to be determined. The estimation of A_m and β can be realized with linear regression or maximum likelihood estimation. However, the log-normal model assumes a constant β over the whole range of seismic IMs, which may not be appropriate in practice [1].

Another way to improve the computational efficiency consists in building a metamodel to calibrate the statistical relation IM-DM. Metamodels have been widely used in structural reliability, but have started to be applied to seismic fragility curve estimation only very recently [9]. This is mainly because the simplification of the continuous stochastic ground motion process by a small set of representative IMs cannot provide a satisfactory description of the randomness of the ground motion, and therefore cannot ensure the performance of the metamodels [12]. The simplest metamodel for seismic IM-DM relation is the power law function [3]. Other more complicated functional models were later proposed in [10]. Considering these works, a non-linear regression metamodel seems more suitable to adequately compute the non-linearity in the IM-DM relation.

In this study, we focus on the application of an artificial neural network (ANN) for the computation of fragility curves. This paper is organized as follows: after a brief recall of simulation-based fragility analysis methods and ANN, the methodology for ANN-based fragility curve estimation is presented and illustrated with a case study for Kashiwazaki-Kariwa nuclear power plant (NPP), studied in the context of the KARISMA benchmark [5]. Only randomness from seismic ground motions is considered here.

2 Simulation-based Fragility Analysis

2.1 Non-parametric Method

Monte Carlo (MC) simulation is the most fundamental approach for the computation of the fragility curves. In this method, N seismic records for the same value of the IM α are collected. Scaling of the signals has to be performed if the number of ground motions is not sufficient. Structural analyses for all N seismic motions are carried out, and the probability of failure with seismic level α is calculated as the number of signals that exceed the failure threshold y_{crit}, divided by the total number N of seismic motions:

$$P_{\text{MC}}(\alpha) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}[y_{\text{crit}} - y_i(\alpha) \leqslant 0] \tag{2}$$

where $\mathbb{1}[y_{\text{crit}} - y_i(\alpha) \leqslant 0]$ is the indicator function, which equals 1 if $y_{\text{crit}} - y_i(\alpha) \leqslant 0$, otherwise it equals 0. One inconvenience of the MC estimation method is that a large number of simulations N may be needed for good accuracy of the MC estimation, especially when the probability of failure to be estimated is small. Therefore, the pointwise evaluation can be very time-consuming, in particular with a complex structure model.

2.2 Parametric Methods

Parametric methods adopt a log-normal assumption of the capacity A of the structures:

$$A = A_m \varepsilon \tag{3}$$

where A_m denotes the median capacity of the structure, ε is a log-normal random variable (RV) with median 1 and a constant logarithmic standard deviation β, which is independent of the range of the seismic IMs. In this way, the conditional probability of failure is calculated by

$$P_f(\alpha) = \Phi\left(\frac{\ln \alpha - \ln A_m}{\beta}\right) \tag{4}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. One can also distinguish the aleatory uncertainty β_R from the epistemic uncertainty β_U. In this way, the total uncertainty is computed by

$$\beta^2 = \beta_R^2 + \beta_U^2 \tag{5}$$

The epistemic uncertainty β_U defines a family of confidence intervals of the fragility curves. The $\gamma\,(\in [0,1])$ non-exceedance confidence interval is calculated by

$$\tilde{P}_f(\alpha,\gamma) = \Phi\left(\frac{\ln\alpha - \ln A_m + \beta_U \Phi^{-1}(\gamma)}{\beta_R}\right) \tag{6}$$

Parameters A_m and β can be determined with linear regression or maximum likelihood estimation.

Linear Regression The linear regression (LR) method is based on the IM-DM model proposed in [3]:

$$\ln y = c\ln\alpha + \ln b + \varepsilon \tag{7}$$

where b and c are regression parameters, and the residual ε following the normal distribution $N(0,\sigma)$. The parameters b, c and σ can be determined from the LR results of data sets ($\ln\alpha_i$, $\ln y_i$); A_m, β in Eq. (4) are then computed as $A_m = \sqrt[c]{y_{\text{crit}}/b}$ and $\beta = \sigma/c$ respectively.

Maximum Likelihood Estimation Different from linear regression, the maximum likelihood estimation (MLE) models the structural responses with binary Bernoulli variables. For a numerical experiment i, $x_i = 1$ if the failure occurs, otherwise $x_i = 0$. Therefore, the likelihood function is written [8]

$$L = \prod_{i=1}^{N} [P_f(\alpha_i)]^{x_i} [1 - P_f(\alpha_i)]^{1-x_i} \tag{8}$$

The values of A_m and β are determined by maximizing the likelihood function, which is equivalent to minimizing $-\ln L$:

$$[A_m, \beta] = \arg\min_{A_m,\beta}(-\ln L) \tag{9}$$

3 Artificial Neural Networks

Inspired by biological neural networks in human brains, ANNs have been widely applied as non-linear regression models since the mid 1980s. The structure of a classic, three-layer, feed-forward ANN is illustrated in Figure 1. Mathematically, this ANN consists of activation functions (linear and non-linear sigmoid functions φ) and a set of weighting parameters $\mathbf{w}$ adjusted to minimize the difference between the ANN predictions $\hat{y}$ and the targets y:

$$\mathbf{w}^* = \arg\min_{\mathbf{w}}(E(\mathbf{x};\mathbf{w})) \tag{10}$$

$$E(\mathbf{x};\mathbf{w}) = \frac{1}{2}\sum_{i=1}^{N}(\hat{y}^i(\mathbf{x};\mathbf{w}) - y^i)^2 \tag{11}$$

where N denotes the total number of ANN training examples, and $\mathbf{x}$ is the ANN input vector. The ANN is trained based on the gradient vector $\mathbf{g}$ (the derivative of $E(\mathbf{x};\mathbf{w})$ with respect to the weights $\mathbf{w}$) computed by the back-propagation algorithm [2]. In order to prevent overfitting, the available data set is divided into training, validation and testing subsets. Cross-validation and early stopping can be applied during the ANN training to obtain the optimal weights.

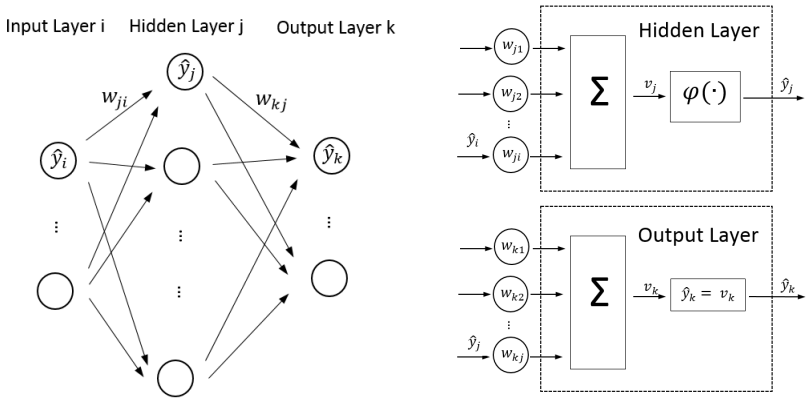

Figure 1: Illustration of a three-layer ANN

As any regression model, there exist uncertainties in the ANN predictions. The delta method is adopted in this study for the computation of the prediction intervals (PIs), because of its computational efficiency. Assuming that the ANN training error is normally distributed, this method relies on the linear Taylor expansion of the ANN model and the estimation of the PIs of the corresponding linear model [7]. This method does not calculate strictly accurate PIs (because the Hessian matrix is approximated by the product of the Jacobian matrices.), but it is fast and gives satisfactory estimations of the PIs in practical applications. Mathematically, the PIs are computed with ANN training errors and the gradient vector $\partial\hat{y}/\partial\mathbf{w}$:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(\hat{y}^i - y^i)^2}{N-Q} \tag{12}$$

$$\mathbf{h}^i = \frac{\partial\hat{y}^i}{\partial\mathbf{w}} = \frac{\partial\hat{y}^i}{\partial E(\mathbf{x};\mathbf{w})}\frac{\partial E(\mathbf{x};\mathbf{w})}{\partial\mathbf{w}} = \frac{1}{\hat{y}^i - y^i}\frac{\partial E(\mathbf{x};\mathbf{w})}{\partial\mathbf{w}} = \frac{\mathbf{g}}{\hat{y}^i - y^i} \tag{13}$$

where Q denotes the dimension of the ANN weighting parameters, $\mathbf{g}$ can be calculated by the back-propagation algorithm. The Jacobian matrix $\mathbf{J}$ of the ANN training data is, hence, constructed as

$$\mathbf{J} = \begin{bmatrix}\mathbf{h}^1 & \mathbf{h}^2 & \cdots & \mathbf{h}^i & \cdots & \mathbf{h}^N\end{bmatrix} \tag{14}$$

where $\mathbf{J}$ is a $Q \times N$ matrix, with $\mathbf{h}^i$ a $Q \times 1$ sub-matrix. Consequently, the $(1-\gamma)$ PIs are calculated as

$$\hat{y} \pm q_{\gamma/2,N-Q}\sigma\sqrt{1+\mathbf{h}^T(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{h}} \tag{15}$$

where $q_{\gamma/2,N-Q}$ is the $(1-\gamma/2)$ quantile of a student distribution with $N-Q$ degrees of freedom and the upper index T denotes the matrix transpose.

4 Application of ANN to KARISMA Benchmark

4.1 Kashiwazaki-Kariwa FEM Model

In 2007, the Japanese Kashiwazaki-Kariwa NPP was affected by the Niigataken-Chuetsu-Oki (NCO) earthquake with the magnitude $M_w = 6.6$ and the epicenter distance 16 km. In this study, we are interested in the reliability of an electrical cabinet located on the fifth floor of the Unit 7 of the NPP. The FEM simulations are carried out with the Code_Aster, a finite element analysis open-source software developed by EDF group. The finite element model for the Unit 7 consists of 92 000 degrees of freedom with 10 700 nodes and 15 600 elements, including bar, beam, and different shell elements. The constitutive law of the materials is considered as linear. The NPP

model is embedded 23 meters in the soil (Figure 2), which is accounted for in the soil-structure interaction (SSI) analysis.
Given the scenario of the NCO earthquake, 100 triples of 3D synthetic ground motions are generated at the bedrock and used for the uncertainty propagation. In order to obtain enough failure counts for the fragility analysis, the synthetic seismic motions at the bedrock are scaled with a factor of three. The seismic signals are then reconvoluted onto the free surface using a 1D column of soil. The degradation of the soil during the earthquake is considered by the equivalent linear method [11] based on the 1D column. 100 triples of ground motions and 100 degraded soil profiles are thus obtained from the equivalent linear method results. The impedances of the

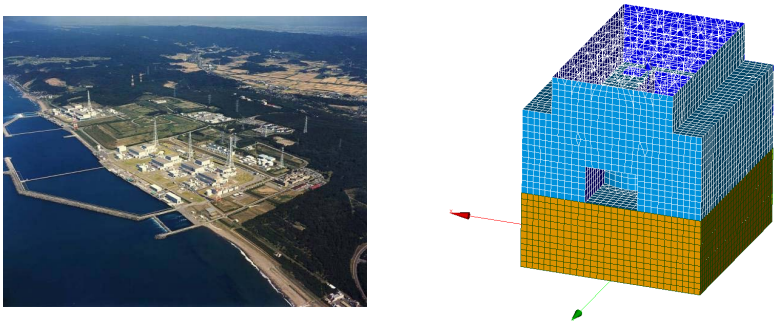

Figure 2: Kashiwazaki-Kariwa nuclear power plant (left) [5] and numerical model of Unit 7 (right)

soil and the seismic forces should have been computed for each soil profiles. However, the high complexity of the embedded foundation makes it hard to achieve: it takes 24 hours to compute the impedances and the seismic forces for one soil profile. A compromise is made between the computational cost and the accuracy. The 3D seismic signals at the bedrock are regrouped into 4 classes according to their PGA values: [0, 0.5g), [0.5g, 1.0g), [1.0g, 1.5g), [1.5g, +∞). The degraded soil profiles are averaged within each class, and 4 soil profiles are obtained to represent four different degradation levels. The SSI analyses are performed with the 100 ground motions on the free surface, and the impedances and seismic forces calculated from the 4 soil profiles.
The anchorage failure of the electrical cabinet is considered in this study. The capacity is given by the spectral acceleration around 4Hz, the natural frequency of the cabinet. The maximum value of the floor spectral accelerations in horizontal directions is defined as the DM y.

$$y = \max_{i=X,Y} \int_{3.5}^{4.5} S^{e}_{a,i}(f)\mathrm{d}f \tag{16}$$

where $S^{e}_{a,i}$ denotes the spectral acceleration of the electrical equipment in the i direction (i can be direction X or Y).

4.2 Selection of Relevant IMs

After the 100 damage measures have been obtained from FEM simulations, these results, combined with the IMs of the ground motions on the free surface, are available for the construction and the training of the ANN metamodel. In this study, seismic IMs are used as inputs of the metamodel. Before the training of the ANN, it is important to select the IMs that are the most relevant to the structural damage measure. This step, named feature selection, is crucial in the metamodel construction phase, because the redundant information carried by the IMs reduces the accuracy of the ANN model [2]. In addition, a large input dimension increases the number of the ANN weights to be determined, which requires more training cases to avoid the overfitting of the network. Based on experience, 8 preliminary IMs (Table 1) are chosen as candidates for

the feature selection. The geometric means of IMs in horizontal directions are used as scalar IMs for 3D ground motions. The intensity measures are assumed to follow log-normal distributions.

Table 1: Definitions of classic seismic intensity measures

Intensity Measure	Definition	Comment
PGA	$\max\|a(t)\|$	$a(t)$: seismic acceleration
PGV	$\max\|v(t)\|$	$v(t)$: seismic velocity
PGD	$\max\|u(t)\|$	$u(t)$: seismic displacement
$\mathrm{PS}_a(f_0)$	Spectral acceleration (damping 5%)	f_0=4Hz: natural frequency
ASA	$\int_{3.5}^{4.5}\mathrm{PS}_a(f)\mathrm{d}f$	f: frequency
T_p	$\arg\max_T \mathrm{PS}_a(\frac{1}{T})$	$T=1/f$
CAV	$\int_0^{t_{\max}}\|a(t)\|\mathrm{d}t$	$t_{\max}$: total seismic duration
I_A	$\frac{\pi}{2g}\int_0^{t_{\max}}a(t)^2\mathrm{d}t$	$g=9.81\mathrm{m/s}^2$

A filter approach for the feature selection is applied in this study, with a forward selection algorithm driven by the semi-partial correlation coefficients (SPCCs). As all the IMs are correlated, SPCC is used as a relevance measure to eliminate the redundant information. For this purpose, as shown in Figure 3, the non-selected IMs (X_2) are projected into the orthogonal space of the selected ones (X_1). The SPCC calculates, in fact, the cosine value of the angle between the projection and the DM. The forward selection is executed in the log-log space to benefit from the properties of the normal distributions. The orthogonal projections can be easily realized by means of Cholesky decomposition. The forward selection algorithm adopted here is as follows:

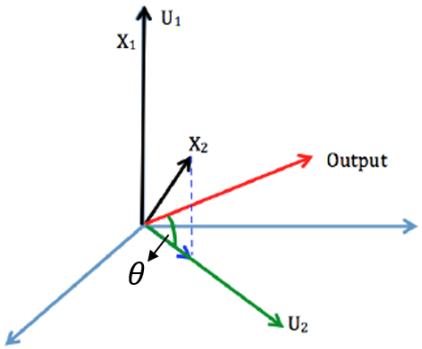

Figure 3: SPCC: $\cos(\theta)$

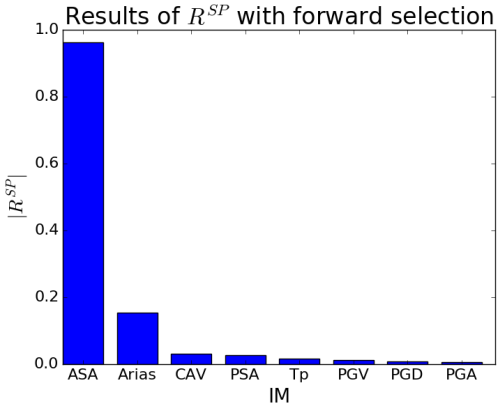

Figure 4: Results of forward selection

1. Define the input and the output of the algorithm: they are respectively the feature set $S_0=\{X_1,\cdots,X_8\}$ (8 IMs) and the ordered feature set S^*, and initialize the output set $S_0^*=\emptyset$.

2. Begin the iteration i (i starts from 0): for each feature X_j in S_i, compute the relevance measure SPCC between X_j and the Output Y, conditional to S_i^*, and select the feature with the largest SPCC value:

$$j^*=\arg\max_j \mathrm{SPCC}(X_j,Y|S_i^*) \tag{17}$$

$$R_i^{\mathrm{SP}}=\mathrm{SPCC}(X_{j^*},Y|S_i^*) \tag{18}$$

As $S_0^*=\emptyset$ initially, for $i=0$, SPCC is actually the correlation coefficient.

3. Subtract the selected X_{j^*} from the feature set: $S_i\ominus X_{j^*}\rightarrow S_{i+1}$, and add X_{j^*} into the output set: $S_i^*\oplus X_{j^*}\rightarrow S_{i+1}^*$.

4. Set $i = i + 1$ and return to Step 2 until all the IMs are selected in S^*.

From the forward selection result (Figure 4), ASA and I_A are selected as the relevant features because the R^{SP} for the other IMs are less than 0.05, so that they can be regarded as non-influential if ASA and I_A have already been considered.

4.3 ANN Training

The ANN is trained with the selected seismic intensity measures, i.e. ASA and I_A. 100 input-output data sets obtained from the FEM simulation results are used. Cross-validation and early stopping are applied in order to prevent overfitting. The 100 data sets are divided into two groups: 80 data sets for the 4-fold cross-validation and 20 data sets for testing. The 80 data sets

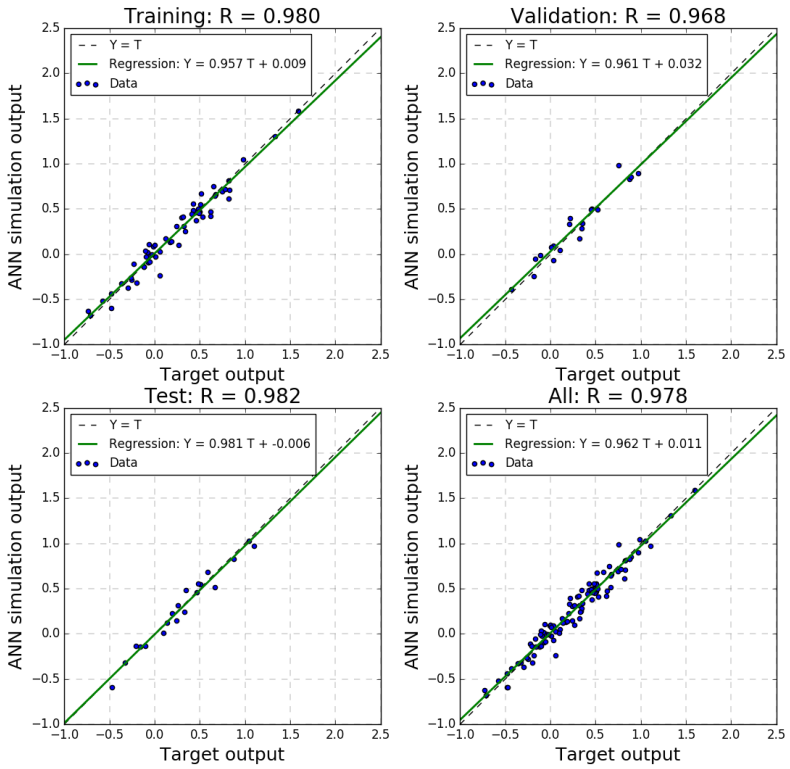

Figure 5: ANN training results

are evenly divided into 4 folds, with 20 in each fold. When the cross-validation is executed, one fold is chosen as the validation set, and the other three folds are used to train the ANN. The training is stopped when the validation error reaches its minimum. This procedure is repeated 4 times to cover all the 4 folds as the validation set. The ANN is trained in log-log space. The ANN with the smallest 4-fold cross-validation error is chosen for the final metamodel. The number of neurons in the hidden layer used in this study is 4. The results of the ANN training, as well as the point clouds of the ANN outputs $\hat{y}$ of the 20 test data are shown in Figure 5 and Figure 6. The prediction intervals are computed with the delta method described in Section 3. The point clouds are plotted for ASA, which is the most relevant IM in this study. From Figure 5, one can conclude that the training results are satisfactory. Most of the results in the 'prediction-target' space are located in the neighborhood of the diagonal line (dashed line). The ANN prediction results for the 20 test data sets in Figure 6 show that the global prediction quality is satisfactory: the ANN predictions remain globally coherent with the FEM results. In fact, with a regression model like ANN, it is somehow not possible to obtain the exact prediction results as the targets. In addition, one phenomenon to emphasize is that the variability, more precisely, the dispersion

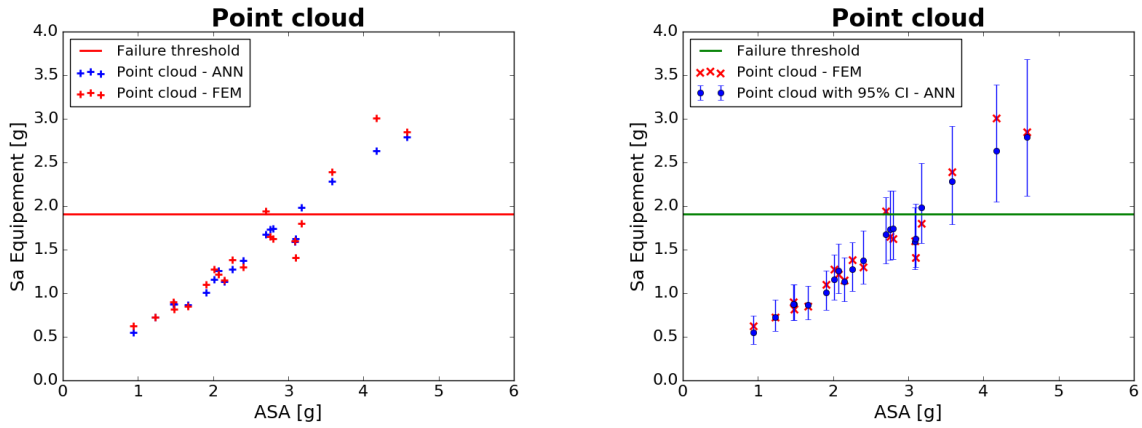

Figure 6: ANN test point cloud (left) and associated prediction intervals (right)

of the ANN predictions (blue points) is reduced compared to the target results (red points). The reduction of the variability on the ANN predictions will reduce the uncertainty β of the fragility curve. Therefore, the uncertainty of the ANN predictions, should be taken into account to obtain a correct β, in particular for the fragility curves determined with ASA.

4.4 Computation of Fragility Curves with ANN

After being trained, the ANN can be used to carried out fast-running simulations. For this goal, a large number of seismic indicators have to be generated to represent the seismic motions. IMs are assumed to follow log-normal distributions according to seismological models such as ground motion prediction equations (GMPE). With statistical results from the GMPE, seismic IMs can be generated directly as the inputs of the ANN. One advantage of using the IMs as the inputs is that no synthetic seismic accelerations are in need with this approach.
In this study, the following statistical properties of the log-normal distributions of ASA and I_A are obtained from the 100 groups of seismic signals on the free surface (Table 2). As ASA is the most relevant IM, the fragility curves in this study are plotted with ASA.

Table 2: Statistics of ASA and I_A on the free surface

IM	Median	Logarithmic standard deviation	Correlation coefficient (ASA-I_A)
ASA [g]	2.28	0.417	0.846
I_A [m/s]	13.13	0.842	

With the large number of simulation results provided by the ANN, both parametric and non-parametric methods can be applied for the computation of fragility curves. This also allows confirming the validity of the log-normal assumption in fragility analysis. Moreover, it is important to consider the ANN prediction uncertainty, because the regression of the ANN metamodel reduces the variability of the DMs, and thus β of the fragility curves. The residual uncertainty of the ANN β_{Resi} should be integrated additionally in order to obtain a correct β. In fact, there exist two sources of this residual uncertainty β_{Resi}. From the fact that the selected features can not totally represent the variability of ground motions, there exist some missing variables characterizing the earthquake motions, which are not identified and not used as the inputs of the ANN. On the other hand, it is also because the ANN cannot represent exactly the real model. Both sources are considered as epistemic uncertainties in this study.

Log-normal Based Fragility Analysis 10 000 ASA-I_A samples are generated with the statistics in Table 2 for the computation of the fragility curves with log-normal assumption. The gen-

eration of ASA-I_A can be realized by applying the Cholesky decomposition on the covariance matrix. 10 000 ANN simulations are run with these samples, which provides the 2D response surface in Figure 7. Based on the 10 000 IMs-DM, the fragility curves are computed with LR and MLE. A constant residual uncertainty is assumed for the log-normal based methods, which is coherent with the constant β in the log-normal assumption. The β_{Resi} is computed as follows:

$$\ln y = \ln \hat{y} + \varepsilon_{Resi} = \ln b + c \ln \alpha + \varepsilon + \varepsilon_{Resi} \tag{19}$$

where ε_{Resi} is the ANN regression error. The second equality in Eq. (19) is due to the application of Eq. (7) to $\hat{y}$. Both ε and ε_{Resi} follow normal distributions $N(0,\sigma)$ and $N(0,\sigma_{Resi})$, respectively. Therefore, β_{Resi} is defined as σ_{Resi}/c and β_R is calculated from the data $(\alpha,\hat{y})$ with parametric methods. The total β is computed by $\beta = \sqrt{\beta_R^2 + \beta_{Resi}^2}$. Confidence intervals of fragility curves are computed with Eq. (6), replacing β_U by β_{Resi}.

Non-parametric Fragility Analysis The pointwise fragility analysis can be carried out by conditional sampling for a given ASA. In this study, at every given ASA, 10 000 I_A values are generated. In the pointwise method, the residual uncertainty depends on the seismic intensity level. The integration of the residual uncertainty in the pointwise fragility analysis can be realized with a second MC level: the ANN residual can be sampled given the results of the delta method (Eq. (15)), and the residual is added to the ANN prediction $\hat{y}$ for counting the failure number. Confidence intervals are computed with the PI intervals of the ANN. The final

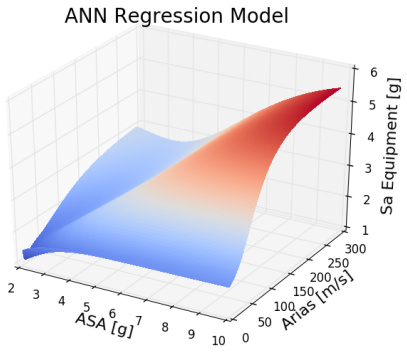

Figure 7: ANN response surface

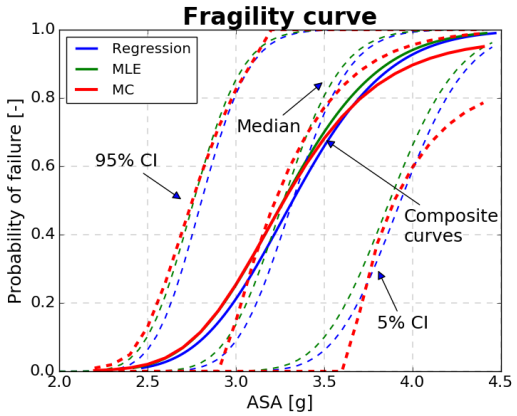

Figure 8: Fragility curves with ANN

fragility curves are plotted in Figure 8. Fragility curves are computed with three methods: linear regression method (blue curves), maximum likelihood estimation (greed curves) and direct Monte Carlo estimation (red curves). The curves drawn with solid lines ('composite curves') are the mean fragility curves that consider both the uncertainties from IMs and the residual uncertainty. Recall that in our study only the earthquake randomness is modeled. The dashed lines are families of the confidence intervals computed with three methods. In this study, the residual uncertainty, considered as the epistemic uncertainty, is the source of the families of confidence intervals. Regarding the composite fragility curves, although there exist some differences between the log-normal based fragility curves and the MC estimation, the log-normal assumption performs well in this study. In addition, at high ASA level, the confidence intervals computed with pointwise estimation are larger than the ones with the log-normal based methods. This is due to the fact that a non-constant epistemic uncertainty has been applied for the MC curves: it increases with the ASA level, while a constant epistemic uncertainty has been used for the methods with the log-normal assumption.

5 Conclusion

The methodology of the application of ANN metamodels to the computation of fragility curves has been explored in this study. The ANN metamodel is utilized to build a statistical relation between the seismic intensity measures and the structural response. Once built and trained, the ANN metamodel allows carrying out a large number of simulations for both parametric and non-parametric fragility analyses, at negligible computational cost. Based on FEM simulation results, this methodology consists of three parts:

1. Selection of the most relevant seismic intensity measure features and training of the ANN.
2. Seismic intensity sampling and simulation with ANN.
3. Computation of fragility curves and their confidence intervals, with both parametric (log-normal assumption) and non-parametric methods.

This methodology has been applied to an industrial case study, i.e. Kashiwazaki-Kariwa nuclear power plant in Japan to evaluate the robustness of an electrical cabinet. The mean fragility curve and a family of confidence intervals have been computed.

Future work will address the adaptive learning of the ANN, which can improve the quality of the training data set, while requiring less FEM simulations to obtain an ANN with accuracy.

References

[1] J. W. Baker. "Probabilistic structural response assessment using vector-valued intensity measures". In: *Earthquake Engineering and Structural Dynamics* 36 (2007), pp. 1861–1883.

[2] C. M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 2005.

[3] C. A. Cornell et al. "Probabilistic Basis for 2000 SAC Federal Emergency Management Agency Steel Moment Frame Guidelines". In: *Journal of Structural Engineering* 128 (2002), pp. 526–533.

[4] EPRI. *Methodology for developing seismic fragilities*. Tech. rep. Electric Power Research Institute, 1994.

[5] IAEA. *Review of Seismic Evaluation Methodologies for Nuclear Power Plants Based on a Benchmark Exercise*. Tech. rep. International Atomic Energy Agency, 2013.

[6] R. Kennedy et al. "Probabilistic seismic safety study of an existing nuclear power plant". In: *Nuclear Engineering and Design* 59 (1980), pp. 315–338.

[7] I. Rivals and L. Personnaz. "Construction of confidence intervals for neural networks based on least squares estimation". In: *Neural Networks* 13 (2000), pp. 463–484.

[8] M. Shinozuka et al. "Statistical Analysis of Fragility Curves". In: *Journal of Engineering Mechanics* 126 (2000), pp. 1224–1231.

[9] B. Sudret and C. V. Mai. "Computing seismic fragility curves using polynomial chaos expansions". In: *Proceedings of the 11th ICOSSAR conference*. 2013.

[10] H. Xu and P. Gardoni. "Probabilistic capacity and seismic demand models and fragility estimates for reinforced concrete buildings based on three-dimensional analyses". In: *Engineering Structures* 112 (2016), pp. 200–214.

[11] N. Yoshida et al. "Equivalent linear method considering frequency dependent characteristics of stiffness and damping". In: *Soil Dynamics and Earthquake Engineering* 22 (2002), pp. 205–222.

[12] I. Zentner and E. Borgonovo. "Construction of variance-based metamodels for probabilistic seismic analysis and fragility assessment". In: *Georisk* 8.3 (2014), pp. 202–216.