SECOND EUROPEAN CONFERENCE ON EARTHQUAKE ENGINEERING AND SEISMOLOGY, ISTANBUL AUG. 25-29, 2014

Turkish Earthquake Foundation - Earthquake Engineering Committee
Prime Ministry, Disaster And Emergency Management Presidency

EPISTEMIC UNCERTAINTY ASSESSMENT USING TRAINED NEURAL NETWORKS

Dimitris G. GIOVANIS[1], Michalis FRAGIADAKIS[2] and Vissarion PAPADOPOULOS[3]

ABSTRACT

Incremental Dynamic Analysis (IDA) is a powerful method for seismic performance assessment method of structures. Recently IDA was shown also to be a very efficient tool for handling epistemic uncertainty. In the latter case IDA is nested within a Monte Carlo framework. Since the use of IDA with Monte Carlo simulation requires the execution of a vast number of nonlinear response history analyses, the calculation of performance statistics becomes time-consuming and therefore quicker alternatives are desirable. In order to reduce the computational effort, we propose the use of trained Artificial Neural Networks. Neural Networks, once trained, are able to provide inexpensive response estimates and combined with a Monte Carlo simulation can rapidly generate a large sample of IDA curves which subsequently allow to calculate useful response statistics of the buildings capacity due to epistemic uncertainty.

INTRODUCTION

The reliability assessment of structures under seismic loads is a topic that has drawn considerable attention over the last decades. The reliability assessment involves two main sources of uncertainty, ground motion uncertainty (also known as aleatoric) and the uncertainty owing to modeling assumptions, omissions or errors (also known as epistemic). For earthquake engineering applications, emphasis is primarily given on the evaluation of the seismic demand and capacity due to aleatory uncertainty. Recently, Vamvatsikos and Fragiadakis (2010) showed that epistemic uncertainty, is equally important to the record-to-record aleatoric variability, especially at late limit-states, as the structure approaches collapse.

A powerful method for seismic performance assessment method is Incremental Dynamic Analysis (IDA). IDA essentially requires subjecting a building's model to a set of ground motion records, each scaled to multiple levels of intensity, thus producing single-record capacity curves, which are then summarized to a median curve. The combination of IDA with reliability analysis methods such as Monte Carlo simulation (MCS) has been successfully implemented by several researchers (Dolsek 2009, Liel *et al.* 2009, Vamvatsikos and Fragiadakis 2010). Although such IDA-based methods are powerful, they necessitate the execution of a large number of nonlinear response history analyses and therefore are beyond the scope of many practical applications. Vamvatsikos and Fragiadakis (2010) also discussed the possibility of reducing the computational effort by adopting approximate, moment-estimating methods such as the Rosenblueth's point estimating method (PEM)

[1] Dr, National Technical University of Athens, moonwalkerd@hotmail.com
[2] Lecturer, National Technical University of Athens, mfrag@mail.ntua.gr
[3] Assistant Professor, National Technical University of Athens, vpapado@central.ntua.gr

(Rosenbleuth 1981) and the first-order, second-moment (FOSM) method (Baker and Cornell 2003). Such schemes manage to propagate uncertainty from the random parameters to the model using only a few IDA runs/simulations based on different versions of the model.

Artificial neural networks (ANN) have been also successfully implemented in probabilistic analysis problems where the seismic demand and capacity is evaluated using nonlinear analysis methods (Papadrakakis *et al.* (1996, 1998), Hurtado 2001, Papadopoulos *et al.* 2012). The major advantage of using a trained ANN in the core of a MCS process is that approximate results can be obtained requiring orders of magnitude less computational effort compared to the standard procedure. This is due to the associative memory properties featured by these artificial intelligence algorithms that allow them to become efficient surrogates to the numerical solver of the structural model which is repeatedly invoked in the MCS. For example, Lagaros and Fragiadakis (2007) used trained ANNs in order to obtain fragility curves of steel frames. The Neural Network was trained to predict the conditional interstorey drift demand using a vector of intensity measures.

In this work we show that artificial neural networks (ANN) can be used provide inexpensive estimates of the first moments of the demand. Once trained, ANN can rapidly generate a large sample of IDA curves with minimal computing effort. Therefore, the computing effort required is focused on the initial training of the network and is always of considerably smaller order of magnitude than the actual simulations. Thanks to the advent of powerful numerical libraries (ie. Matlab, NAG, IMSL) ANN methods are widely available to scientists and engineers and can be easily adopted, once the engineer becomes familiar with the underline concepts.

EPISTEMIC UNCERTAINTY ESTIMATION WITH INCREMENTAL DYNAMIC ANALYSIS

Incremental Dynamic Analysis (IDA) (Vamvatsikos and Cornell, 2002) offers thorough seismic demand and capacity prediction capability for the complete range of the model's response, from elastic to yielding, then to nonlinear inelastic and finally to global dynamic instability. As already discussed, IDA involves subjecting a structural model to one (or more) ground motion record(s), each scaled to multiple levels of intensity, thus producing curve(s) of the capacity, parameterized versus intensity level.

Every dynamic analysis is characterized by two scalars, an intensity measure (IM), which represents the scaling factor of the record, and an engineering demand parameter (EDP). For moderate-period structures with no near-fault activity, an appropriate choice for the IM is the 5%-damped first-mode spectral acceleration $S_a(T_1,5\%)$ while for the EDP we choose the maximum interstory drift θ_{max} due to the fact that we focus on deformation-sensitive structural and non-structural damage. Using the *hunt & fill* algorithm (Vamvatsikos and Cornell, 2004) we were able to capture each IDA curve with only 12 runs per record. Appropriate interpolation techniques allow the generation of a continuous IDA curve in the IM-EDP plane from the discrete points obtained by the dynamic analyses. Such results are in turn summarized to produce the median and the 16%, 84% percentile IDA curves that can accurately characterize the distribution of the seismic demand and capacity of the structure for the full range of ground motion intensities.

In order to evaluate the effect of uncertainties on the seismic performance of the structure, faced with the non-existence of a closed-form solution for the seismic response of a complex nonlinear model, there are few options that we can follow to estimate its variability. The most comprehensive, but also the most computationally expensive method is the Monte Carlo simulation (MCS). By sampling N times from the parameter distributions, Monte Carlo creates a population of N possible instances of the structure, and each will be analysed. Assuming that a sufficiently large number of structures have been sampled, we can reliably estimate the full distribution of the seismic performance of the structure. By performing IDA on every of the N samples we obtain $R \times N$ single-record IDA curves and N corresponding median IDAs , where R is the number of ground motion records used for every IDA simulation. These allow us to provide unbiased estimates of the mean and the variance of the median IDA curve due to the uncertainty in the parameters of the structure. Such dispersion caused by the uncertainty in the median capacity ΔS_a can be characterized by its β-value, β_U which can be calculated directly as the standard deviation of the natural logarithm of the estimates of the median capacities:

$$\Delta S_a = \mathbf{med}_j (S_{a,j}{}^{50\%}) \approx \overline{\ln S_{a,j}}^{50\%} \quad (1)$$

$$\beta_U = \sqrt{\frac{\sum_{j=1}^{N} (\ln S_{a,j}{}^{50\%} - \overline{\ln S_{a,j}}^{50\%})^2}{N-1}} \quad (2)$$

where $S_{a,j}^{50\%}$ are the estimates of the median S_a value of capacity for a given limit-state defined at a specific value of θ_{max} from each model realization and $\overline{\ln S_{a,j}}^{50\%}$ is the mean of the natural logarithm of the median S_a-values of capacity. Figure 1 shows 200 median IDA curves obtained by creating 200 instances of the structure with MCS and performing for each samples an Incremental Dynamic Analysis using 30 ground motion records. Therefore, Figure 1 has been produced after performing a total of 30×200 = 6000 IDAs. The mean IDA curve is also shown along with the mean plus and minus one standard deviation curves of $S_a(T_1,5\%)$ given θ_{max}.

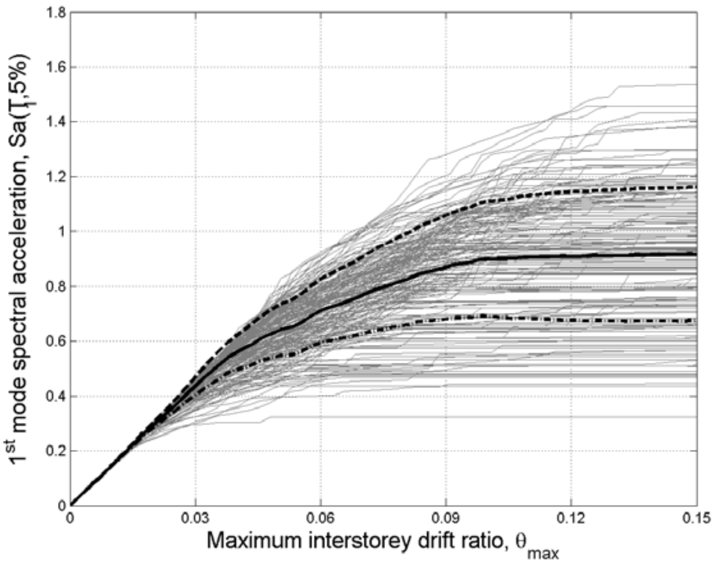


Figure 1. 200 median IDA curves shown against their mean and the mean plus (and minus) one standard deviation IDA curves.

Monte Carlo simulation can be further improved by replacing crude random sampling from the population with the Latin Hypercube sampling (LHS) algorithm (Dolsek 2009, Vamvatsikos and Fragiadakis 2010). The IDA curves can also be used to estimate the variability caused by the parameter uncertainties in the median capacity for every limit-state. Since the use of IDA in the framework of MCS requires the execution of a vast number of nonlinear responses history analyses the calculation of performance statistics becomes an extremely time-consuming procedure and thus it is of no practical interest. A simpler alternative to performing Monte Carlo simulation is the use of moment-estimation methods to approximate the variability in the IDA results, specifically the point estimate method (PEM) and the first-order-second-moment method (FOSM) (Vamvatsikos and Fragiadakis 2010).

NEURAL NETWORKS

Artifial neural networks (ANN) is a biologically inspired intelligence technique, since they are composed of elements that perform in a manner analogous to the elementary functions of a biological neuron. In most cases a ANN is an adaptive algorithm that changes its structure based on external or internal information that flows through the network during the learning phase. The artificial neuron (Figure 1) processes the information (input signal [$\mathbf{x}_1$, $\mathbf{x}_2$, ... , $\mathbf{x}_\mathbf{p}$]) and predicts an output vector $\mathbf{y_k}$, which depends on a set of connections **w**, known as synaptic weights. Furthermore, neural networks can be viewed as highly nonlinear functions with the basic the form $F(\mathbf{x}, \mathbf{w}) = \mathbf{y}$.

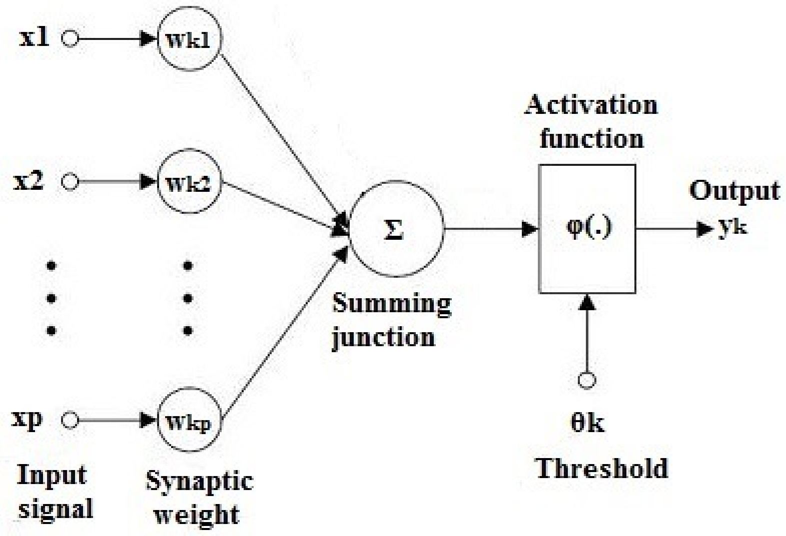


Figure 2 Schematic plot of an artificial neuron.

One of the most crucial matters in the use of an ANN is the learning procedure which is necessary in order for the network to operate. There are two kinds of learning procedures, the supervised learning and the unsupervised learning. In the case of supervised learning we train the ANN by using a pair of input/output data, while in the second case the ANN learns automatically from its enviroment. Basically, the learning procedure tries to adjust the synaptic weights in order to have a mapping that fits well the training set. The training procedure of an ANN can be considered as a general function optimization problem, with the adjustable parameters being the weights **w** the network.

The major advantage of a trained ANN over a conventional numerical procedure is that it leads to results that can be produced rapidly, requiring orders of magnitude less computational effort than performing the actual analysis. As mentioned before, in the case of supervised learning the training set of an ANN is composed by input-target pairs. If a set of weight values **w** is assigned to the connections of the network, a mapping $F(\mathbf{x},\mathbf{w})$ is defined between the input vector **x** and the output vector **y.** The quality of this mapping, with respect to the training set, is measured with the aid of an error function:

$$E_D(\mathbf{w}) = \sum_m \frac{1}{2} e^2 \tag{3}$$

where e are the network errors composed from the difference of the desired response and the actual system output. A learning algorithm tries to determine the values of the weights **w** in order to achieve the correct response for every input vector applied to the network by minimizing the value of E_D. The numerical minimization algorithm used for the training, generates a sequence of weight parameters **w** through an iterative procedure.

TRAINING OF THE NEURAL NETWORK

In order to use the ANN to obtain estimates of the $S_a(T_1,5\%)$ building capacities for a given limit-state $\theta_{\max}$, we must first provide an appropriate set of input/output data for training the network. The input vector **x** contains the random parameters and also their perturabations of the around their mean. If the number of random variables describing the epistemic uncertainty of the system is K, the size of the input vector **x** will be $2K+1$. Each of the $2K+1$ elements of the input vector $\boldsymbol{x}_i$ contain a shift of the random variable around its mean value, plus (or minus) one standard deviation, while the remaing variables are set equal to their mean.

$$\mathbf{x} = [\mathbf{x}_0, \mathbf{x}_1, ..., \mathbf{x}_{2K}]^T \tag{4}$$

where,

$$\begin{aligned} \mathbf{x}_0 &= \{\mu_1, \mu_2, \cdots, \mu_K\}^T \\ \mathbf{x}_1 &= \{\mu_1 + \sigma_1, \mu_2, \cdots, \mu_K\}^T \\ \mathbf{x}_2 &= \{\mu_1 - \sigma_1, \mu_2, \cdots, \mu_K\}^T \\ &\vdots \\ \mathbf{x}_{2K} &= \{\mu_1, \mu_2, \cdots, \mu_K - \sigma_K\}^T \end{aligned} \tag{5}$$

where μ_i, $i = 1, \dots, K$ is the mean value of each random variable and σ_i is the corresponding standard deviation. The vector $\mathbf{x}_0$ represents the base-case in which all random variables are set equal to their mean value. The output vector for the ANN training set, will contain the median $S_{a,50\%}$ IDA capacities given θ_{max}. Therefore, for every training vector $\mathbf{x}_i$ we perform IDA analysis to obtain, the median $S_{a,50\%}$ capacities, conditional on θ_{max}. The output vector of the ANN will be of the form:

$$\mathbf{y} = \left[S^0_{a,50\%}, S^1_{a,50\%}, S^2_{a,50\%}, \dots \right] \tag{6}$$

In order to obtain accurate estimates of the seismic demand using trained neural networks, we must define the suitable neural network architecture and select proper training sets ginving emphasis on the selection of the input/output (I/O) pair. In the proposed methodology as I/O pair we use the **x/y** vectors presented above.

The training of the network has an important influence on the approximation quality of a neural network. The training for a feed-forward multi-layered network is called back-propagation (Riedmiller and Braun 1993) where the network operation is executed reversely for the training sets and the calculated input values are compared to the given values. Depending on a given learning rate the calculated error is corrected and the same procedure is repeated until the network reproduces the training values optimally. If no convergence is achieved, generally the training is terminated after a given number of training loops called epochs. In this work the Levenberg-Marquardt (LM) algorithm (Levenberg 1944, Marquardt 1963) is used since it has a fast and stable convergence. Another important point for a sufficient network approximation is the design of the network architecture. Depending on the number of available training samples the number of neurons in the hidden layers has to be chosen in such a way that over fitting is avoided. This phenomenon occurs if the number of hidden nodes is too large for the number of training samples.

$$(n+2)M + 1 < m \tag{7}$$

where n is the number of input values, M is the number of hidden neurons for a network with single hidden layer and m the number of training samples. After the ANN is trained we can use it to get a large number of estimations of the IDA curve without computational cost. This is achieved in a way similar to Monte Caro simulation where each random parameter value is sampled from its probability distribution using a random number generator. The efficiency of the sampling can be further improved using Latin Hypercube sampling instead of crude Monte Carlo sampling.

NUMERICAL EXAMPLE

The structure selected is a nine-story steel moment-resisting frame with a single-story basement (Figure 3) that has been designed for Los Angeles, following the 1997 NEHRP provisions. A centerline model with nonlinear beam–column connections was formed allowing for plastic hinge formation at the beam ends while the columns are assumed to remain elastic. The structural model also includes P-Δ effects while the internal gravity frames have been directly incorporated (Figure 3). The fundamental period of the reference frame is T_1=2.35 sec and accounts for approximately 84% of the total mass. A suite of thirty ordinary ground motion records representing a scenario earthquake was used. These belong to a bin of relatively large magnitudes of 6.5-6.9 and moderate distances, all recorded on firm soil and bearing no marks of directivity.

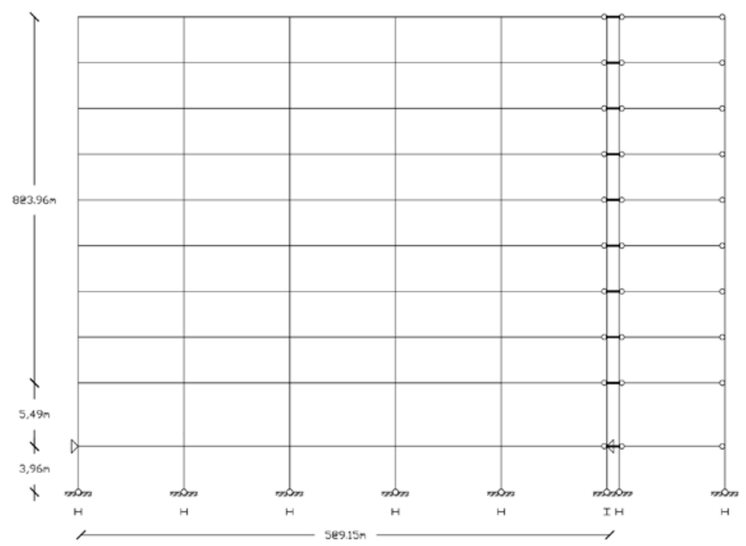


Figure 3 The LA9 steel moment-resisting frame.

The beam-hinges are modeled as rotational springs with a quadrilinear moment-rotation backbone (Figure 4) that is symmetric for positive and negative rotations. The backbone hardens after a yield moment of a_{My} times the nominal, having a non-negative slope of a_h up to a normalized rotation (or rotational ductility) μ_c where the negative stiffness segment starts. The drop, at a slope of a_c, is arrested by the residual plateau appearing at normalized height r that abruptly ends at the ultimate rotational ductility μ_u. Thus, in order to completely describe the backbone of the monotonic envelope of the hinge moment-rotation relationship, the aforementioned six parameters are necessary.

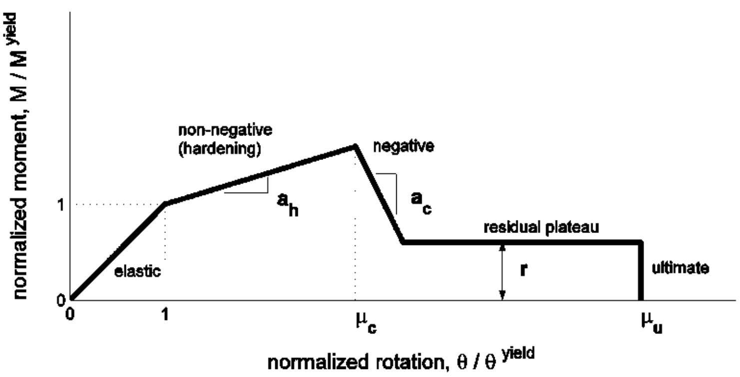


Figure 4 The moment-rotation relationship of the beam point-hinge in normalised coordinates.

The variability of the quadrilinear backbone curve properties is considered to be the only source of epistemic uncertainty and thus, the six parameters mentioned above are considered to be random variables, independently normally distributed with mean and coefficient of variation (c.o.v.) as shown in Table 1. The mean values represent the best estimates of the backbone parameters, while the c.o.v values were assumed because there is no explicit guidance in the literature. Thus, we used a c.o.v equal to 40% for all the parameters except for the yield moment where 20% was used instead. To avoid assigning the random parameters with values with no physical meaning, e.g. $a_h > 1$ or $r < 0$ their distribution is appropriately truncated within 1.5 standard deviations as shown in Table 1. The plastic hinge properties are assumed to be varying simultaneously for every frame connection.

Table 1 Random parameters and their statistics

	Mean (μ)	**c.o.v**	**Lower bound**	**Upper bound**
a_{My}	1.0	0.2	0.70	1.30
a_h	0.1	0.4	0.04	0.16
μ_c	3.0	0.4	1.20	4.80
a_c	-0.5	0.4	-0.80	-0.20
r	0.5	0.4	0.20	0.80
μ_u	6.0	0.4	2.40	9.60

As the intensity measure (IM) we have chosen the 5%-damped, first-mode spectral acceleration $S_a(T_1,5\%)$, to represent the seismic intensity while the engineering demand parameter (EDP) adopted is the maximum interstorey drift, θ_{max}, that previous research has shown to be a good measure of structural damage. In Figure 5 we can see the $2K+1=13$ IDA curves used for the ANN training, where each curve corresponds to a vector representing the epistemic uncertainty of the LA9 steel frame. The trainning vector contains the six parameters of Table 1 and Figure 4, thus $\mathbf{x}_0=[a_{My}, a_h, \mu_c, a_c, r, \mu_u]$. Thus, according to Table 1, for the case study examined we obtain: $\mathbf{x}_0=[1.0\ 0.1\ 3.0\ -0.5\ 0.5\ 6.0]^T$, $\mathbf{x}_1=[1.2\ 0.1\ 3.0\ -0.5\ 0.5\ 6.0]^T$, $\mathbf{x}_2=[0.8\ 0.1\ 3.0\ -0.5\ 0.5\ 6.0]^T$, etc. For every row of the $\mathbf{x}$ vector we perform IDA analysis, to obtaine trhe medina IDA curve. For the example here considered we performed in total 13 IDAs. Figure 5 shows the 13 IDA curves used as the training set of our example. The solid line is the median basecase IDA, while the dashed lines denote the mean plus and minus one standard deviation curves of the first random variable (a_{My}). Therefore, the process of producing an approximate IDA curve with an ANN using only 13 samples for its efficient training involves initially generating the 13 realizations of the LA9 frame which correspond to the perturbations of the random parameters and subjected it to non-linear dynamic analysis in order to estimate the corresponding $S_a(T_1,5\%)$ capacities for every value of θ_{max}.

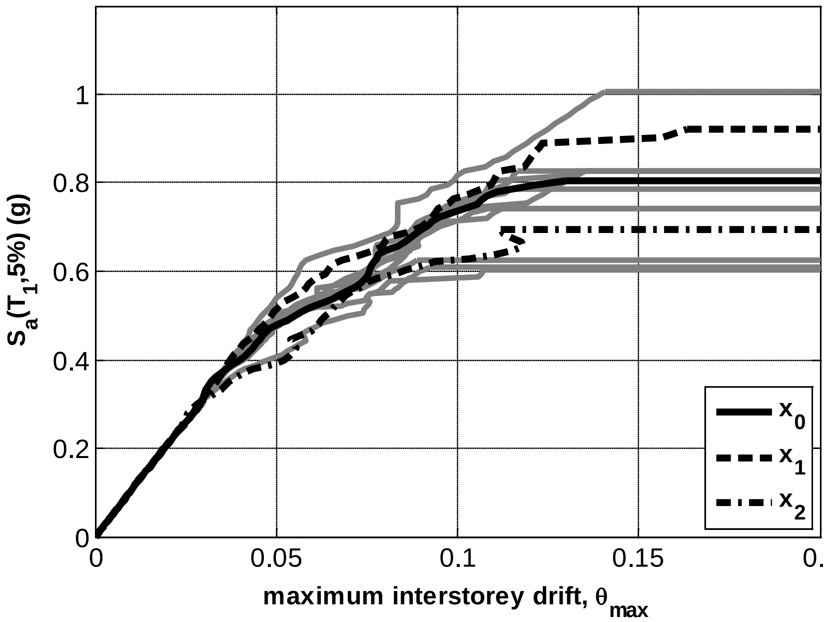


Figure 5 The 2K+1 = 13 $S_{a,50\%}$-capacities, for a given θ_{max} (IDA curves) corresponding to the13 realizations used for training the NN in the case of the LA9 builiding. The black lines refer to the basecase IDA (solid line) and to the first random parameter (dashed lines).

Once the input vector **x** and the output vector **y** are obtained, we choose the appropriate architecture for the neural network and we train it, under the condition that the prediction error is less than a threshold value (~5%). The trained network is then used to quickly obtain a sufficiently large sample of IDA curves, which can be post-processed to provide the required response statistics. Figure 6 shows 1000 IDA curves obtained with the ANN for 1000 realizations of the steel frame here considerd. According to Figure 6 the ultimate $S_a(T_1,5\%)$ capacities of the frame vary between 0.2g and 1.2g.

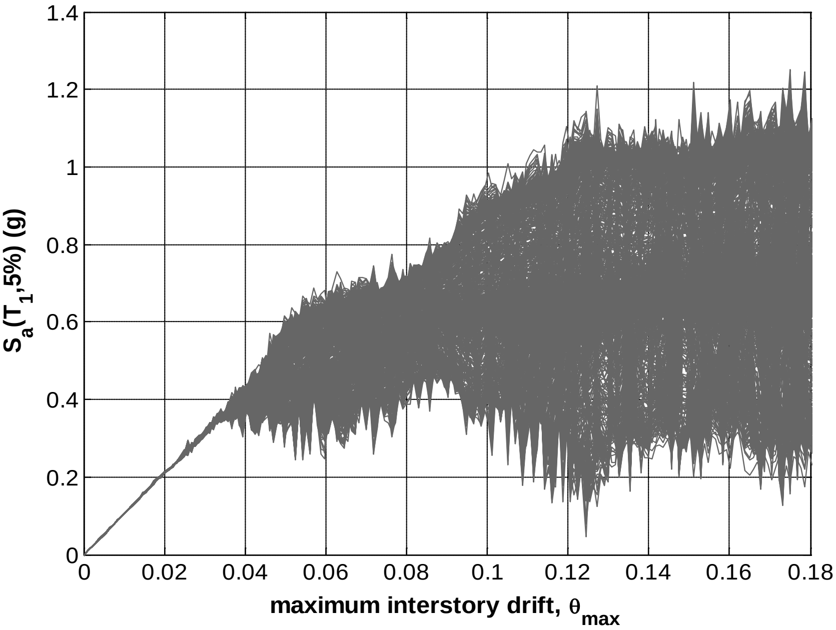
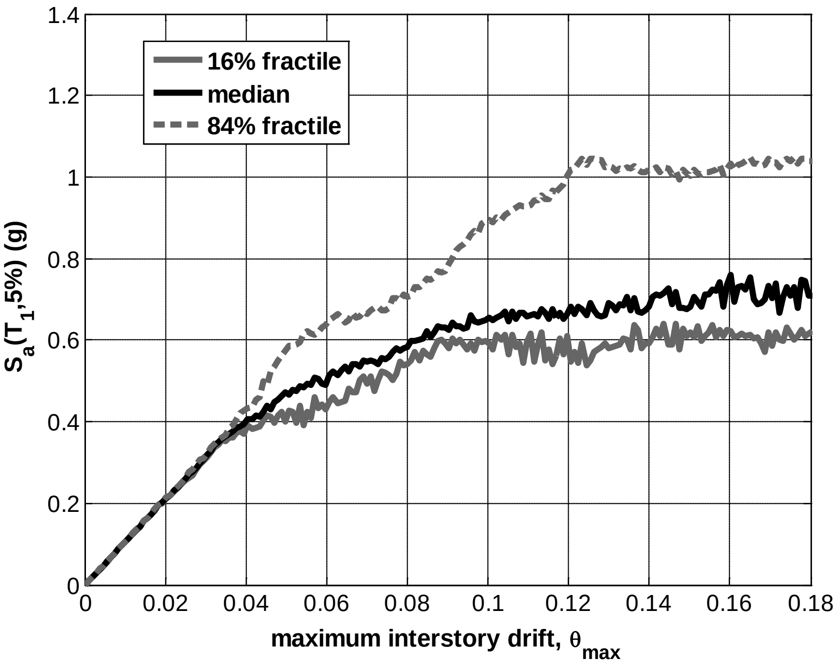


Figure 6 1000 IDA curves obtained with the ANN.



.

Figure 7 The mean value of the 1000 ANN-IDA curves along with the 16% and 84% fractiles

Figure 7 shows the response statistics obtained after post-processing the IDA curves obtained with the aid of the ANN. The median (50% fractile) curve provides a ‘central’ capacity curve, while the 16 and 84% percentiles give a measure of the dispersion around the median. Figure 8 provides a comparison of the median ANN-based IDAs with the base-case which is the IDA curve obtained when random variables are set equal to their mean value. The basecase is mereley a first-order estimate of the median curve. Moreover, the samed figure also shows the curve obtained using Rosenbleuth's point estimation method (PEM) method and the first order second moment (FOSM), as discused in Vamvatsikos and Fragiadakis (2010). These methods also simple alternatives to performing Monte Carlo simulation and provide estimates of the response. Using functional approximations or moment-matching, such schemes manage to propagate uncertainty from the parameters to the final results using only a few IDA runs. In this Figure (8), the median capacity curve obtained with the ANN is very close to the basecase where FOSM and PEM underestimate the capacity beyond θmax = 0.07.

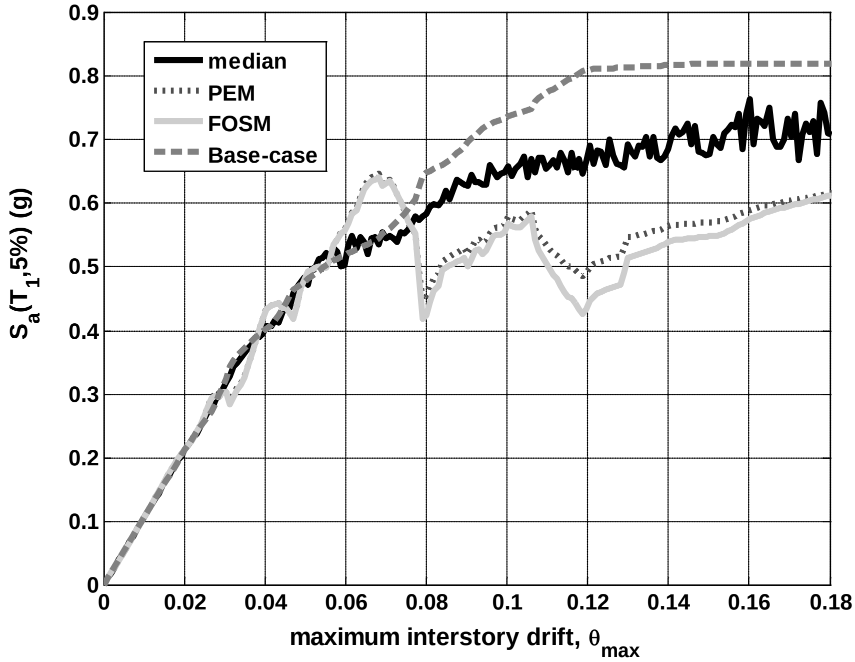


Figure 8 The mean value of the ANN-IDA curves along with the FOSM, PEM and base-case IDAs

CONCLUSIONS

A neural network–based procedure is proposed for obtaining inexpensive estimates of IDA curves considering epistemic uncertainty. The computing effort involved in Monte Carlo simulation , together with the need for an accurate reliability analysis procedure, motivated the use of neural networks. The computating effort is excessive because of the required sample size and the large number of nonlinear response history analyses required for a single IDA. The use of ANN substantially reduces the sample size and leads to close estimates of the median IDA and the corresponding dispersion. Neural networks are trained using a set of perturbations of the *K* random variables around their mean value. 2*K*+1 realizations of the LA9 steel frame are generated by shifting each random parameter plus or minus one standard deviation while all the other remain equal to their mean value. A good estimation of the IDA curves depends on the training of the neural network. The Levenber-Marquardt algorith was used for this purpose along with a proper architecture. The proposed methodology, compared to other simplified approaches (e.g., FOSM, PEM), offers more accurate approach of the IDA curve.

ACKNOWLEDGMENTS

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National

Strategic Reference Framework (NSRF). Research Funding Program: THALES. Investing in knowledge society through the European Social Fund.

REFERENCES

Baker J W, Cornell C A (2003) "Uncertainty specification and propagation for loss estimation using FOSM methods", *PEER Report 2003/07*, Pacific Earthquake Engineering Research Center, University of California, Berkeley.

Dolsek M (2009) "Incremental dynamic analysis with consideration of modeling uncertainties", *Earthquake Engineering and Structural Dynamics*, 38: 805–825

Hurtado J E (2001) "Neural network in stochastic mechanics", *Arch. Comput. Methods Eng.*, 8:303-342.

Lagaros N D and Fragiadakis M (2007) "Fragility Assessment of Steel Frames Using Neural Networks", *Earthquake Spectra*, 23(4): 735-752.

Levenberg K (1944) "A method for the solution of certain problems in least squares", *Quarterly of Applied Mathematics*, 5:164-168.

Liel A, Haselton C, Deierlein G, Baker J (2009) "Incorporating modeling uncertainties in the assessment of seismic collapse risk of buildings", *Structural Safety*, 31(2): 197–211.

Marquardt D (1963) "An algorithm for least-squares estimation of nonlinear parameters", *SIAM Journal on Applied Mathematics*, 11(2):431-441.

Papadopoulos V, Giovanis D G, Lagaros N D, Papadrakakis M (2012) "Accelerated subset simulation with neural networks for reliability analysis", *Comput. Methods Appl. Mech. Engrg*., 223-224:70-80.

Papadrakakis M, Lagaros N D, Tsompanakis Y (1998) "Structural optimization using evolution strategies and neural networks", *Comput. Methods Appl. Mech. Engrg*., 156:309-333.

Papadrakakis M, Papadopoulos V, Lagaros N D (1996) "Structural reliability analysis of elastic-plastic structures using neural networks and Monte Carlo simulation", *Comput. Methods Appl. Mech. Eng*., 136:145-163.

Riedmiller M, Braun H (1993) "A direct adaptive method for faster back-propagation learning: the RPROP algorithm", *IEEE International Conference on Neural Networks (ICNN)*, San Francisco, 586-591.

Rosenblueth E (1981) "Point estimates for probability", *Applied Mathematical Modeling*, 5: 329-335.

Vamvatsikos D, Cornell CA (2002) "Incremental dynamic analysis", *Earthquake Engineering and Structural Dynamics*, 31(3): 491-514.

Vamvatsikos D, Cornell CA (2004) "Applied incremental dynamic analysis", *Earthquake Spectra*, 20(2): 523-553.

Vamvatsikos D, Fragiadakis M. (2010) "Incremental Dynamic Analysis for Seismic Performance Uncertainty Estimation", *Earthquake Engineering and Structural Dynamics*, 39(20): 141-163.