Available online at www.eccomasproceedia.org
Eccomas Proceedia COMPDYN (2017) 3163-3188

ECCOMAS
Proceedia

COMPDYN 2017
6th ECCOMAS Thematic Conference on
Computational Methods in Structural Dynamics and Earthquake Engineering
M. Papadrakakis, M. Fragiadakis (eds.)
Rhodes Island, Greece, 15–17 June 2017

APPROACH TO PREDICTION OF R/C BUILDINGS' SEISMIC DAMAGE AS PATTERN RECOGNITION PROBLEM USING ARTIFICIAL NEURAL NETWORKS

K.E. Morfidis[1], and K.G. Kostinakis[2]

[1] Assistant Researcher, Earthquake Planning and Protection Organization (EPPO-ITSAK)
Terma Dasylliou, 55535, Thessaloniki, Greece
e-mail: konmorf@gmail.com

[2] Post-Doc Researcher, Department of Civil Engineering, Aristotle University of Thessaloniki
Aristotle University campus, 54124, Thessaloniki, Greece
e-mail: kkostina@civil.auth.gr

Keywords: Seismic Damage Prediction, Artificial Neural Networks, Existing R/C Buildings, Structural Vulnerability Assessment, Seismic Response of Buildings, Pattern Recognition

Abstract. *In the present paper the investigation of the problem of reinforced concrete (r/c) buildings' seismic damage prediction is exhibited utilizing Artificial Neural Networks (ANN). More specifically, the problem is formulated and solved in terms of the Pattern Recognition Problem using Multilayer Feedforward Perceptron Networks (MFP). The networks are trained by the implementation of two training algorithms: the Scaled Conjugate Gradient algorithm (SCG) and the Resilient Back-Propagation algorithm (RP). The training data-set is created by means of Nonlinear Time History Analyses (NTHA) of 30 r/c buildings which are subjected to 65 earthquakes. The selected buildings have different heights, structural systems and structural eccentricities, and are designed on the basis of the suggestions of Eurocodes. The damage index which is used to describe the seismic damage state of buildings is the Maximum Interstorey Drift Ratio (MIDR). In the context of the present paper the influence of the number and the combination of seismic parameters which describe the level of impact of seismic excitations on the r/c buildings is also investigated. To this end, 8 different and widely used seismic parameters are utilized. Furthermore, the influence of the number of hidden layers, the number of neurons in the hidden layers, as well as the activation functions of neurons is also examined. The generalization abilities of the optimum configured ANNs are investigated through the assessment of their performance in the case of prediction of seismic damage state of the selected buildings subjected to 16 earthquakes different from the earthquakes which are used in the creation of the training data-set. The most significant conclusion that turned out is that the ANNs can reliably classify the r/c buildings into pre-defined damage classes if they are appropriately configured. More specifically, the parametric investigations prove that the most important factor for the effective ANNs' configuration is the activation functions of output layer's neurons (tansig function) as well as the number of the hidden layers (the utilization of two hidden layers leads to better results). As regards the examination of the optimum number and combination of seismic input parameters (using the "Stepwise" sensitivity analysis method) it is proved that are dependent on the utilized training algorithm.*

doi: 10.7712/120117.5636.17237

1 INTRODUCTION

The assessment of seismic damage level in case of r/c structures is at the core of the civil engineering research globally, since a large number of strong earthquakes that happened in the past have caused extensive structural damages and human losses. As a consequence, numerous researchers have published studies dealing with the methods that can be used in order to estimate the seismic vulnerability of existing buildings. The research studies can be classified into two general categories: a) studies dealing with the development of methods that can estimate the seismic performance of individual buildings and b) studies dealing with the development of methods that can rapidly assess the seismic vulnerability of groups of buildings with common structural characteristics. The research studies belonging to the first category resulted in analytical methodologies of assessment which are based on linear and nonlinear methods that have been adopted by modern seismic codes (e.g. [1, 2, 3]). Respectively, the researches belonging to the second category led to the development of methodologies that can accomplish the approximate assessment of the seismic vulnerability of buildings' groups with common structural characteristics (e.g. rapid screening methods; damage probability matrices; fragile curves, see for example [4, 5, 6]. These methods can be used in order to estimate the buildings' seismic vulnerability in a very short time comparing with the methods of the first category, with the price of losing accuracy. However, the merit of these approximate methods is important, since the use of them can lead to rapid and reliable decisions about the need for further investigation of buildings' seismic performance.

In the context of the direct (though approximate) assessment of the seismic vulnerability of buildings' groups with common structural characteristics, in the past 25 years many research studies have been conducted aiming to utilize the capacities of artificial intelligence, such as the Artificial Neural Networks (ANNs). The ANNs are complex computational structures, which attempt to mimic the operations of the biological (human or animal) brain and the central nervous system and are used for the solution of complex problems with the aid of computers using algorithms based on different philosophy than the conventional ones (see e.g. [7]). Although the ANNs have been developed on an idea that dates back to the 1940s [8], the first thorough investigation dealing with the use of them for the direct (in real time) estimation of the level of structural seismic damage was published in 1995 [9]. Since then, this approach has been the subject of numerous research studies (e.g. [10, 11]), which led to highly important and interesting results. These results designate the ability of ANNs to predict the potential seismic damage of buildings in an approximate but generally reliable way.

The aim of the present study is to further investigate the ability of ANNs to reliably predict the seismic damage level of r/c buildings. More specifically, the problem is formed and solved as a pattern recognition problem. Problems of this type can be solved reliably using ANNs. For the needs of the investigation, the training of the ANNs was achieved with the aid of a data set created using results from Nonlinear Time History Analyses of 30 r/c buildings with different heights, structural systems and structural eccentricities, subjected to 65 actual ground motions. The Maximum Interstorey Drift Ratio (MIDR), (e.g. [12]) was used as the damage index. As a consequence, 1950 (30 x 65) values of MIDR resulted from the analyses conducted. These values were categorized into five general damage levels, into which the investigated buildings were classified with the aid of the ANNs. The training of the ANNs was made by using both the Scaled Conjugate Gradient algorithm (SCG, [13]) and the Resilient Back-Propagation algorithm (RP, [14]). Moreover, the number and the combination of the seismic parameters leading to the strongest correlation between the analyses results and the results produced by the ANNs was investigated. To this end, 8 different and widely used seismic parameters were examined. Furthermore, the influence of the number of hidden layers, the number of neurons in the hidden layers, as well as the activation functions of neurons was also examined. The prediction ability of the trained ANNs was checked through the prediction they made in the case of future earthquakes and it was demonstrated that they can extract especially reliable (though approximate) results.

2 THE ARIFICIAL NEURAL NETWORKS

It is well-known that the Artificial Neural Networks (ANNs) are complex computational structures which are capable to solve problems using the general rules of the human brain functions (see e.g. [7]). Thus, using ANNs it is feasible to approximate the solution of problems such as pattern recognition, classification and function approximation problem.

The ANNs' function is based on the combined action of interconnected processing units that are called artificial neurons (Figure 1(a)). The artificial neuron receives input signals (x_1, x_2,…., x_m) and transform them to an output signal (y_k) through the use of an adder (which adds the products of the input signals by the respective synaptic weights (w_{k1}, w_{k2},…., w_{km}) of neuron's synapses), and the use of an activation function (which has as argument the u_k that results from the adder and transform it to the output signal y_k).

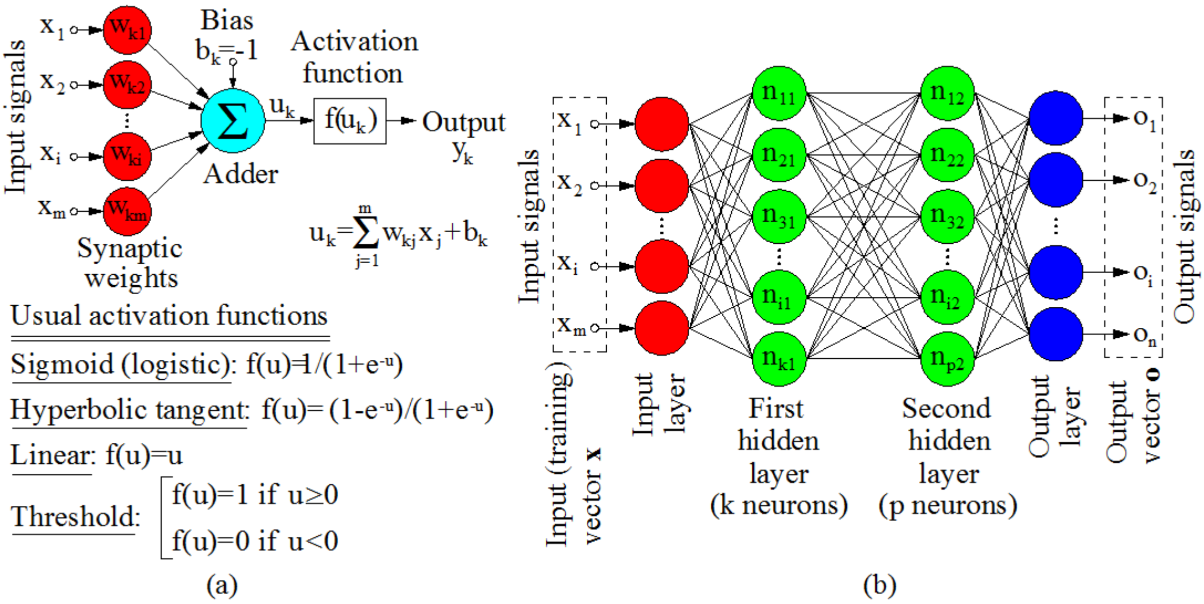


Figure 1: The artificial neuron (a), typical configuration of a multilayer feedforward perceptron ANN (b).

In Figure 1(b) the typical configuration of a Multilayer Feedforward Perceptron (MFP) type ANN with four layers of neurons (input layer, two hidden layers and output layer) is presented. MFP networks were selected for the analyses which were conducted in order to achieve the objective of the present investigation. The solution of problems through the utilization of ANNs is accomplished if they have been trained. The training of ANNs is achieved by procedures which are called training algorithms. These algorithms require a set of *n* input vectors **x** and the corresponding to them *n* output vectors **d** that called target vectors. The *n* pairs of vectors **x** and vectors **d** constitute the training data-set. During the training procedure, the values of the synaptic weights (w) are successively altered until the error vector **e**(**w**) [**e**(**w**)=**d**-**o**(**w**), where **o**(**w**) is the output vector] that is produced by the ANN is minimized.

3 FORMULATION OF THE SEISMIC DAMAGE PREDICTION PROBLEM AS A PATTERN RECOGNITION PROBLEM USING ANNs

As "pattern recognition" is defined the procedure of the detection and identification/classification of objects in specific categories (patterns). The solution of the pattern recognition problem is approached by the utilization of several methods (see e.g. [15]). One of these methods is based on the use of ANNs. In this section, it will be presented the procedure by which the problem of the r/c buildings' seismic damage prediction is formulated as a pattern recognition problem. The steps which are required for this formulation are illustrated in the flowchart of Figure 2. In this diagram the choices which were made for some of the problem's parameters in the current investigation are also presented. Detailed description of all selected parameters will be given in the following subsections.

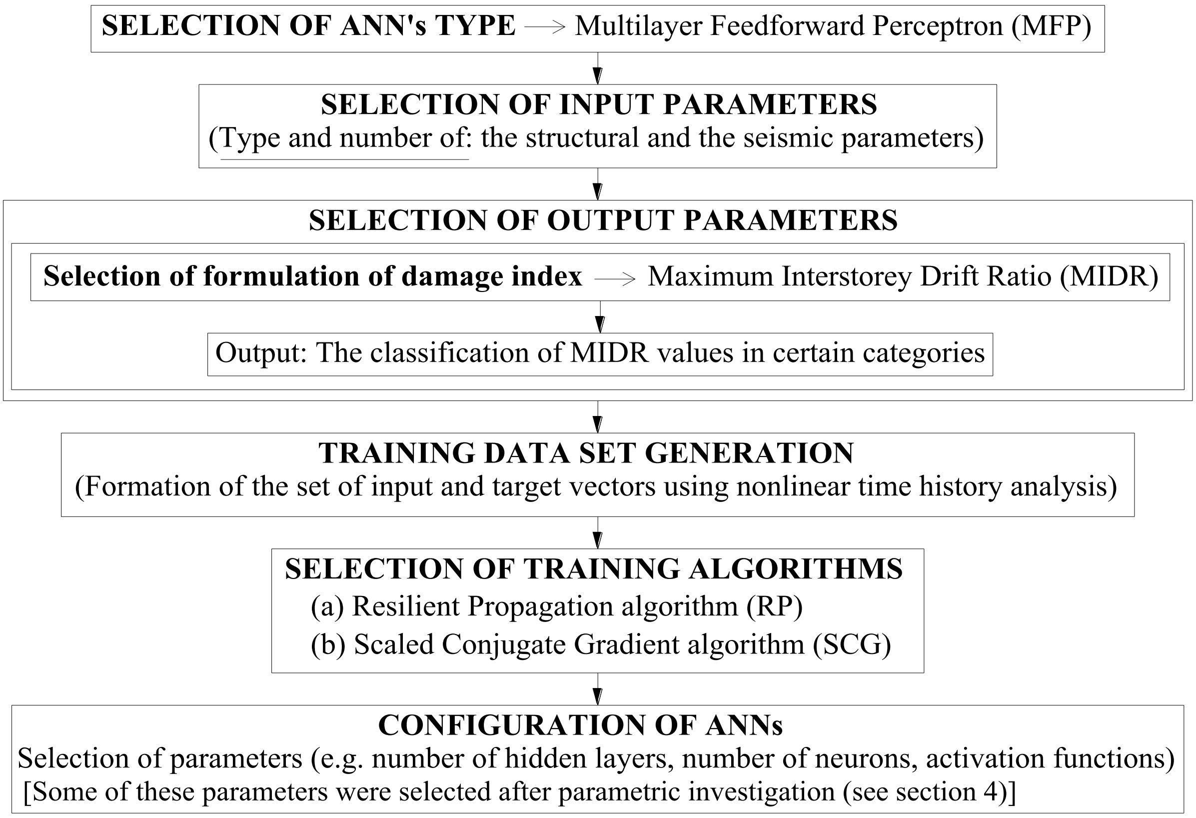


Figure 2: The procedure of the formulation of the pattern recognition problem.

3.1 Selection of the type of ANNs

As mentioned in the section 2, in the present investigation MFP networks were utilized. In the networks of this type, the neurons in any layer are connected to all neurons in the adjacent layer (Figure 1b). This choice was based on the fact that this type of ANNs was successfully used in many published investigations which are related to the subject of the present investigation (see e.g. [10, 11, 16, 17, 18]).

3.2 Selection of the input parameters

The ANNs are computational structures which are capable to handle and approach the solution of multiparametric problems. This feature gives the flexibility to select the number of the parameters (input parameters) through which a problem can be described. The parameters which describe the problem of the prediction of the seismic damage of r/c buildings can be grouped in two categories: the structural parameters and the seismic parameters.

The structural parameters are used for the description of the buildings' response on earthquake excitations. In the case of the vulnerability assessment of existing r/c buildings, the widespread methods such as the fragility curves method (see e.g. [5, 6]) utilize macroscopic characteristics which are related to the geometric morphology and the structural system. In the present study 4 structural parameters were selected: the total height of buildings H_{tot}, the ratio of the base shear that is received by r/c walls (if exist) along two perpendicular directions x and y (ratio n_{vx} and ratio n_{vy}), and the structural eccentricity e_0.

As regards the seismic parameters which are used to describe the seismic excitations and their impact to structures, there are many definitions which are resulted from the analysis of accelerograms records (see e.g. [19]). These parameters can be classified to: (a) seismic parameters determined from the time histories of the records; (b) seismic parameters determined from the response spectra of the records; (c) seismic parameters accounting for the earthquake's frequency content and (d) seismic parameters based on the earthquake's duration. For the investigation conducted in the present study, the 8 seismic parameters illustrated in Table 1 have been chosen.

Ground Motion Parameter	Calculation procedure	Category
Peak Ground Acceleration: **PGA**	max|a(t)|	Seismic parameters determined from the time histories of the records. [a(t), v(t), d(t) are the acceleration, velocity and displacement time history respectively]
Peak Ground Velocity: **PGV**	max|v(t)|	
Peak Ground Displacement: **PGD**	max|d(t)|	
Arias Intensity: $\mathbf{I_a}$	$I_a = (\pi / 2g)\cdot\int_0^{t_{tot}} [a(t)]^2 dt$	
Acceleration Spectrum Intensity: **ASI**	$ASI = \int_{0.1}^{0.5} S_a(\xi = 0.05, T)dT$	Seismic parameters determined from the response spectra of the records. (S_a is the acceleration spectrum, PSV is the pseudovelocity spectrum, ξ is the damping ratio)
Housner Intensity: **HI**	$HI = \int_{0.1}^{2.5} PSV(\xi = 0.05, T)dT$	
V_{max}/A_{max} (**PGV/ PGA**)	max|v(t)|/max|a(t)|	Seismic parameter accounting for the earthquake's frequency content.
Time of Uniform Duration: **TUD**	Special algorithm (e.g. [20])	Seismic parameter based on the earthquake's duration. [TUD is the total time during which the ground acceleration is larger than a given threshold value (usually 5% of PGA)]

Table 1: Examined ground motion parameters.

Thus, in the present study 12 input parameters (4 structural and 8 seismic) were selected. Therefore, the input vector of ANNs **x** has the general form which is given by the Eq. (1).

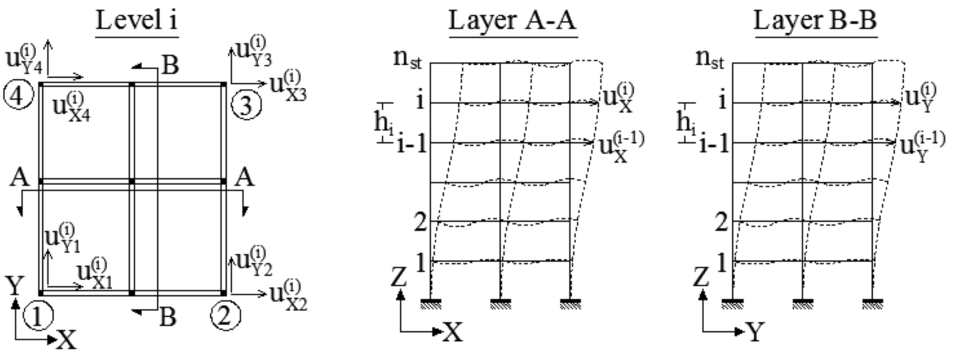
$$\mathbf{x} = \left[\mathbf{x}_{seism} \,|\, \mathbf{x}_{struct}\right]^T \tag{1}$$

$$\mathbf{x}_{seism} = \left[PGA \,¦\, PGV \,¦\, PGD \,¦\, I_a \,¦\, ASI \,¦\, HI \,¦\, PGV / PGA \,¦\, TUD\right]^T, \quad \mathbf{x}_{struct} = \left[H_{tot} \,¦\, e_0 \,¦\, n_{vx} \,¦\, n_{vy}\right]^T$$

However, it must be stressed that the dimension of the input vector **x** was not considered constant (12x1) in the analyses conducted in the present study. As will presented in the section 4, the number and the combination of the seismic parameters in the input vector **x** were investigated in order to accomplish the optimum performance of ANNs, i.e. the optimum prediction of the seismic damage. Thus, the dimension of the used input vectors was fluctuated between (5x1), i.e. 1 seismic parameter plus 4 structural parameters, and (12x1).

3.3 Selection of the output parameters

Due to the fact that the problem which is investigated in the present study is the r/c buildings' seismic damage prediction using ANNs, the first step in the procedure of the selection of the networks' output parameters is the choice of an appropriate damage index. As it is well-known, the damage indices are used for the numerical modeling of the damage level in the vulnerability assessment of structures, and can be grouped into categories based on whether they are local or global, deterministic or probabilistic, structural or financial (e.g. [21]). In the present study, the seismic damage of r/c buildings was expressed in terms of the Maximum Interstorey Drift Ratio (MIDR). The MIDR is a global and structural damage index which is generally considered as a reliable indicator of structural and nonstructural seismic damage of r/c buildings (e.g. [22]), and has been used for the evaluation of seismic response of structures (e.g. [23]). Indicative calculation of the MIDR value in the case of a n-storey building with a rectangular plan-view is shown in Figure 3.



$$\boxed{MIDR=\max\{\max\Delta u_X;\ \max\Delta u_Y\}}$$

$$\max\Delta u_X=\max\{\max\Delta u_X^{(1)},...\max\Delta u_X^{(i-1)},...\max\Delta u_X^{(i)},...\max\Delta u_X^{(n_{st})}\}$$

$$\max\Delta u_Y=\max\{\max\Delta u_Y^{(1)},...\max\Delta u_Y^{(i-1)},...\max\Delta u_Y^{(i)},...\max\Delta u_Y^{(n_{st})}\}$$

$$\max\Delta u_X^{(i)}=\max\{|\Delta u_{X1}^{(i)}|;\ |\Delta u_{X2}^{(i)}|;\ |\Delta u_{X3}^{(i)}|;\ |\Delta u_{X4}^{(i)}|\}$$

$$\max\Delta u_Y^{(i)}=\max\{|\Delta u_{Y1}^{(i)}|;\ |\Delta u_{Y2}^{(i)}|;\ |\Delta u_{Y3}^{(i)}|;\ |\Delta u_{Y4}^{(i)}|\}$$

$$\underline{\text{Point j}}: \ \Delta u_{Xj}^{(i)}=\max\left\{\frac{u_{Xj}^{(i)}(t)-u_{Xj}^{(i-1)}(t)}{h_i}\right\}_{t=0}^{t=t_{tot}};\ \Delta u_{Yj}^{(i)}=\max\left\{\frac{u_{Yj}^{(i)}(t)-u_{Yj}^{(i-1)}(t)}{h_i}\right\}_{t=0}^{t=t_{tot}}$$

(j=1-4)

(t_{tot}= total duration of the seismic excitation)

Figure 3: Indicative calculation of the MIDR value in the case of a n-storey building with rectangular plan-view.

The formulation of the investigated problem as a pattern recognition problem requires the definition of classes into which a r/c building can be classified on the basis of its damage level after an earthquake. To this end, in the present study five damage states were defined using specific limit values of the MIDR. These damage states (valid for r/c buildings) are presented in the Table 2 (see e.g. [24]).

MIDR (%)	<0.25	0.25-0.50	0.50-1.00	1.00-1.50	>1.50
Degree of damage	Null	Slight	Moderate	Heavy	Destruction

Table 2: Relation between MIDR and damage state.

Using the five damage states which are defined in the Table 2 any r/c building which is subjected to an earthquake can be classified on the basis of its corresponding MIDR value. In terms of ANNs this means that any input vector **x** can be classified on the basis on its corresponding output vector **o** (or target vector **d** in the case in which the input vector is a training vector) in one of the five damage states of Table 2. Thus, the output vectors as well as the target vectors must have dimension (5x1). In other words, the number of outputs of ANNs in the present case is five. The general form of the output vectors is given (using an example) in Figure 4.

$$\mathbf{o}=\begin{bmatrix} o_1 \\ o_2 \\ o_3 \\ o_4 \\ o_5 \end{bmatrix} \begin{array}{l} \rightarrow \\ \rightarrow \\ \rightarrow \\ \rightarrow \\ \rightarrow \end{array} \begin{array}{c} \textbf{MIDR} \\ <0.25\% \\ (0.25\%\text{-}0.50\%] \\ \boxed{(0.50\%\text{-}1.00\%]} \\ (1.00\%\text{-}1.50\%] \\ >1.50\% \end{array} \xrightarrow[\text{MIDR=0.72\%}]{\text{Example:}} \mathbf{o}=\begin{bmatrix} 0 \\ 0 \\ \boxed{1} \\ 0 \\ 0 \end{bmatrix}$$

Figure 4: General form of output vectors **o**.

As it emerges from Figure 4, each element of a vector **o** represents one of the five classes/damage states of Table 2 and attains value equal to 1 if the corresponding MIDR belongs to the interval of values which define the specific damage state. Otherwise, it attains value equal to 0.

3.4 Training data-set generation

The procedure which is conducted for the generation of the training data-set is presented in the flowchart of the Figure 5. The steps of this procedure will be described subsequently.

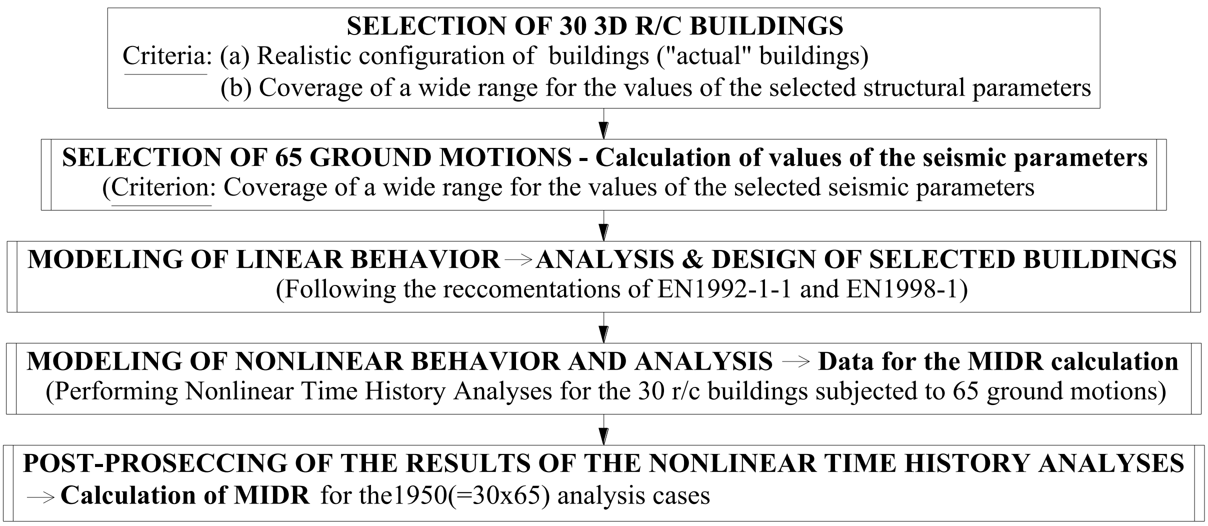


Figure 5: Procedure for the design and generation of the training data-set.

(a) Selection of 30 r/c buildings: The selected buildings (rectangular in plan and regular in elevation according to the criteria set by EN1998-1), differ in the total height H_{tot}, the structural eccentricity e_0 (=the distance between the mass center and the stiffness center of storeys), as well as the ratios of the base shear n_{vx} and n_{vy} that are received by the r/c walls (if they exist) along two perpendicular axes (axes x and y). The values of the above structural parameters for the selected buildings are given in Table 3.

No.	n_{st}	L_x (m)	L_y (m)	e_o (m)	n_{vx} (%)	n_{vy} (%)	No.	n_{st}	L_x (m)	L_y (m)	e_o(m)	n_{vx} (%)	n_{vy} (%)
1	3	13.5	10.0	0.0	0.0	0.0	16	3	13.0	9.0	0.98	0.0	0.0
2	5	20.0	14.0	0.0	0.0	0.0	17	5	17.5	10.0	2.58	0.0	0.0
3	7	20.0	14.0	0.0	0.0	0.0	18	7	17.5	10.0	2.39	0.0	0.0
4	3	15.0	10.0	0.0	73.0	76.0	19	3	13.5	9.0	4.65	52.0	46.0
5	5	19.0	16.2	0.0	77.0	80.0	20	5	16.0	14.5	4.19	43.0	42.0
6	7	19.0	16.2	0.0	57.0	64.0	21	7	16.0	14.5	3.79	37.0	36.0
7	3	15.0	15.0	0.0	41.0	41.0	22	3	13.5	9.0	2.23	47.0	0.0
8	5	21.2	18.7	0.0	46.0	50.0	23	5	16.0	14.5	2.65	38.0	0.0
9	7	21.2	18.7	0.0	43.0	46.0	24	7	16.0	14.5	2.49	35.0	0.0
10	3	17.0	12.5	0.0	43.0	0.0	25	3	14.5	9.0	3.53	64.0	0.0
11	5	20.2	15.2	0.0	41.0	0.0	26	5	14.0	16.0	3.01	0.0	69.0
12	7	20.2	15.2	0.0	38.0	0.0	27	7	14.0	16.0	3.01	0.0	65.0
13	3	15.0	10.0	0.0	77.0	0.0	28	3	13.5	10.0	6.73	64.0	58.0
14	5	20.2	15.2	0.0	68.0	0.0	29	5	16.5	16.5	6.29	65.0	72.0
15	7	20.2	15.2	0.0	51.0	0.0	30	7	16.5	16.5	5.96	59.0	67.0

Table 3: Values of structural parameters of the selected r/c buildings.

In Table 3 $e_0=(e_{0x}^2+e_{0y}^2)^{1/2}$, where e_{0x}, e_{0y} are the structural eccentricities along axes x and y.

(b) Selection of ground motions: A suite of 65 pairs of horizontal bidirectional earthquake ground motions obtained from the PEER [25] and the European strong-Motion database [26] was selected. The main criterion used for the selection of these excitations was to cover a large variety of realistic values for the 8 ground motion parameters considered in the present investigation (see Tables 1 and 4). The seismic parameters for each ground motion were determined as the geometric mean values of the parameters corresponding to the two horizontal components of each earthquake record.

Ground Motion Parameter	Units	Minimum Value	Maximum Value
PGA	%g	0.004	0.822
PGV	cm/sec	0.86	99.35
PGD	cm	0.36	60.19
I_a	m/sec	≈0.0	5.592
ASI	g·sec	0.003	0.633
HI	cm	3.94	317.6
PGV/PGA	sec	0. 036	0.336
TUD	sec	≈0.0	17.68

Table 4: Ranges of values of the selected seismic parameters.

(c) Modeling of linear behavior, analysis and design of selected buildings: The selected buildings (which were considered as Medium Ductility Class structures) were modeled, analyzed and designed based on the provisions of EN1992-1-1 [27] and EN1998-1 [28] (using the pro-

fessional analysis and design program RAF [29]). It must by stressed that in any case the choice of the dimensions of the r/c members' cross-sections, as well as of their reinforcement, was made while bearing in mind the optimum exploitation of steel (S500B) and concrete (C20/25).

(d) Nonlinear modeling and analysis (Nonlinear Time History Analyses) – Calculation of Damage Index MIDR: The buildings' nonlinear behaviour was modeled by means of lumped plasticity models at the column and beam ends, as well as at the base of the walls. The material inelasticity of the structural members was modeled with the aid of the Modified Takeda hysteresis rule [30]. The effects of the axial load-biaxial bending moments (P-M_1-M_2) interaction at column and wall hinges were taken into account by using the P-M_1-M_2 interaction diagram, which is implemented in the software adopted for the application of the analyses.

After the nonlinear modeling, the 30 selected r/c buildings were analyzed by Nonlinear Time History Analysis (NTHA) for each one of the 65 earthquake ground motion pairs. Thus, a total of 1950 NTHA (30 buildings x 65 earthquake records) were performed with the aid of the computer program Ruaumoko [31]. For each one of the 1950 analyses, the required data for the MIDR calculation were exported.

(e) Post-processing of the results of NTHA – Calculation of MIDR: The last step of the procedure for the training data-set generation (Figure 5) is the processing of the results of NTHA in order to calculate the MIDR values of the analyzed buildings. To this end, a computer code in Visual Basic was developed. This code calculates firstly the MIDR values following the procedure which is presented in Figure 3 and subsequently configures the target vectors **d** in the form which is given in Figure 4.

Thus, following the above procedure, 1950 training vectors **x** with dimension (12x1) which are described by Eq. (1) were created. Also, the corresponding 1950 target vectors **d** (dimension (5x1), Figure 4) were created.

3.5 Selection of training algorithms

Two training algorithms were adopted: the Resilient Backpropagation algorithm ("RP" algorithm, [14]) and the Scaled Conjugate Gradient algorithm ("SCG" algorithm, [13]). The RP algorithm constitutes a variation of methods which are based on variable learning rate and it is recommended for the quick and reliable solution of pattern recognition problems (see e.g. [32, 33]). The algorithm SCG belongs to the specific class of Conjugate Gradient (CG) methods, which can effectively handle large-scale problems. Furthermore, special features of the SCG algorithm are its fastness, as well as the fact that it does not contain user-dependent parameters.

3.6 Configuration of ANNs

The last step of the procedure for the formulation and solution of the seismic damage prediction problem as a pattern recognition problem using ANNs, is the configuration of the networks. More specifically, this step concerns the choice of the parameters which are required for the configuration of the used ANNs. These parameters are: (a) the number of the hidden layers; (b) the number of neurons in each hidden layer; (c) the activation functions of neurons; (d) the performance evaluation parameters; (e) the normalization functions of the input and output values and (f) the method for partitioning the data set in training, validation and testing subsets.

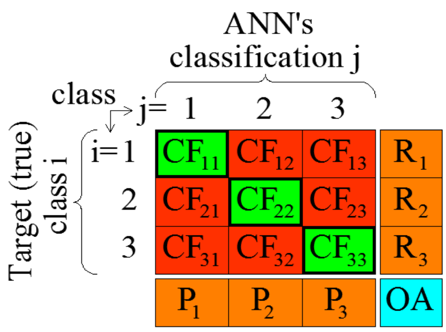
In the present study, specific choices for some of the aforementioned parameters were made, while for some others more than one choice were made. These choices are the following:

(a) *Number of the hidden layers*: ANNs with one or two hidden layers were selected.

(b) *Number of neurons in hidden layers*: The optimum number of neurons in hidden layers is not uniquely defined for all problems. Furthermore, there is no a direct method for its determination, so the trial and error method is always adopted. In the context of the present study, an investigation for the determination of the optimum number of neurons in hidden layers was conducted. More specifically, ANNs with a number of neurons in hidden layers that ranges between 10 and 60 were configured. Then, the number of neurons that led to the optimum results in each examined case was determined. In section 4, more details will be given.

(c) *Activation functions of neurons*: Two different types of activation functions for neurons were used: the sigmoid function (logistic - logsig) and the hyperbolic tangent function (tansig), (Figure 1(a)). The choice of using two activation functions (instead of using a single one) was made in order to investigate the optimum efficiency of the ANNs.

(d) *Performance evaluation parameters*: The performance evaluation parameters are indices which are used for the assessment of ANNs' prediction abilities. In the case of solution of a pattern recognition problem the most useful tools for the evaluation of ANNs are the Confusion Matrices - CM (see e.g. [32, 34]). The general form of a CM (for a three-class problem) is presented in Figure 6.



CF_{ij} = the number of the input vectors whose true class is i and were classified by the ANN to class j

R_i (Recall) $= CF_{ii} / \sum_{j=1}^{3} CF_{ij}$

(the percentage of input vectors of true class i which were correctly classified by the ANN to class i)

P_j (Precision) $= CF_{jj} / \sum_{i=1}^{3} CF_{ij}$

(the percentage of input vectors which were classified by the ANN into class j, whose true class is j)

OA (Overall Accuracy) $= \sum_{i=1}^{3} CF_{ii} / N$

(the percentage of input vectors that correctly classified by the ANN)

(N=total number of input and target vectors)

Figure 6: General form of a confusion matrix for a three-class problem

On the basis of CMs three types of metrics for ANNs' prediction accuracy are defined, namely the "Recall" index, the "Precision" index and the "Overall Accuracy" index (Figure 6). In the present study, the "Overall Accuracy" or (OA) index was mainly used. However, for the evaluation of the several configurations of the ANNs which were examined the corresponding CMs are also presented and evaluated in section 4.

(e) *Normalization functions for the input and target vectors' elements*: The utilization of functions which normalize the values of the elements of input vectors **x** before these vectors introduced to ANNs is considered as necessary in order to optimize the training (e.g. [35]). The same transformation is also required for the elements of the target vectors **d**. A function, through which the elements of input and the target vectors of the data set attain values in the range [-1,1], was selected in the present study [32].

(f) *Partition of the data-set*: The partition of the data-set in three sub-sets, namely the training, the validation and the testing sub-set is recommended in order to ensure good generalization of networks and to avoid the overfitting (e.g. [36]). In the present study, the partition of the data-set in training, validation and testing sub-sets was done in any case using the ratio 70%/15%/15% respectively. The training and target vectors, which consist of the three sub-sets, were chosen randomly [32].

4 PARAMETRIC INVESTIGATION FOR THE OPTIMUM PERFORMANCE OF ANNS

In this section the results of the parametric analyses which were conducted for the investigation of optimum ANNs' seismic damage prediction is exhibited. More specifically, the aim of these analyses was the investigation of the influence of:

(a) the number and the combination of seismic input parameters,
(b) the type of the neurons' activation functions,
(c) the training algorithms,
(d) the number of neurons in hidden layers, and
(e) the number of hidden layers

to the ANNs' performance.

This investigation was conducted in two parts. In the first part ANNs with one hidden layer were utilized. In this part the influence of the factors (a)-(d) were investigated. In the second part, ANNs with two hidden layers were used. It must be noted that in the second part of investigation, specific conclusions of the first part (such as conclusions about the influence of the number and the combination of seismic input parameters as well as the influence of the type of the neurons' activation functions) were taken into consideration in order to reduce the number of required training procedures, which is significantly more in the case of ANNs with two hidden layers. For example, in the case of ANNs with one hidden layer, the number of different combinations of activation functions' types in the hidden and output layer is 4 (i.e. logsig/logsig, tansig/tansig, logsig/tansig and tansig/logsig). The corresponding number in the case of ANNs with two hidden layers is 8. If we also take into consideration the fact that the investigation of the optimum number of neurons required training procedures for ANNs with different number of neurons in each hidden layer (10÷60 as it will be presented in the following subsections), it is clear that the total number of the required calculations in the case of ANNs with two hidden layers is significantly greater.

4.1 Investigation of optimum performance of ANNs with one hidden layer

The investigation of the optimum performance of ANNs with one hidden layer was conducted in three stages. In each stage, different number of seismic input parameters was considered:

(a) *Stage 1*: Each one of the 8 seismic input parameters of Table 1, and the 4 considered structural parameters were introduced into the input vectors **x** (Eq. 1). Therefore, eight different types of input vectors were formed (with dimension (5x1)).

(b) *Stage 2*: The 8 seismic input parameters of Table 1 were introduced simultaneously with the 4 structural parameters into the input vectors **x**. Therefore, one type of input vectors was formed (with dimension (12x1)).

(c) *Stage 3*: Using the two versions of "Stepwise" sensitivity analysis method [37, 38, 39] the number and the combination of seismic input parameters which lead to the optimum

ANNs' predictions for the seismic damage state of r/c buildings was investigated. Thus, in this stage input vectors with dimension between (5x1) and (12x1) were formed.

The Stages 1 and 2 were utilized for the detection of the optimum combination of neurons' activation functions in the hidden and the output layer between the 4 possible combinations. Thus, the analyses of the Stage 3 were conducted only for the optimum combination of these functions, but for both of the two training algorithms which were considered in the present study (RP and SCG). For the training of the investigated ANNs in Stages 1 and 2 a specific procedure was performed in order to detect the number of neurons in hidden layer which leads to the best predictions. This procedure is described in Figure 7.

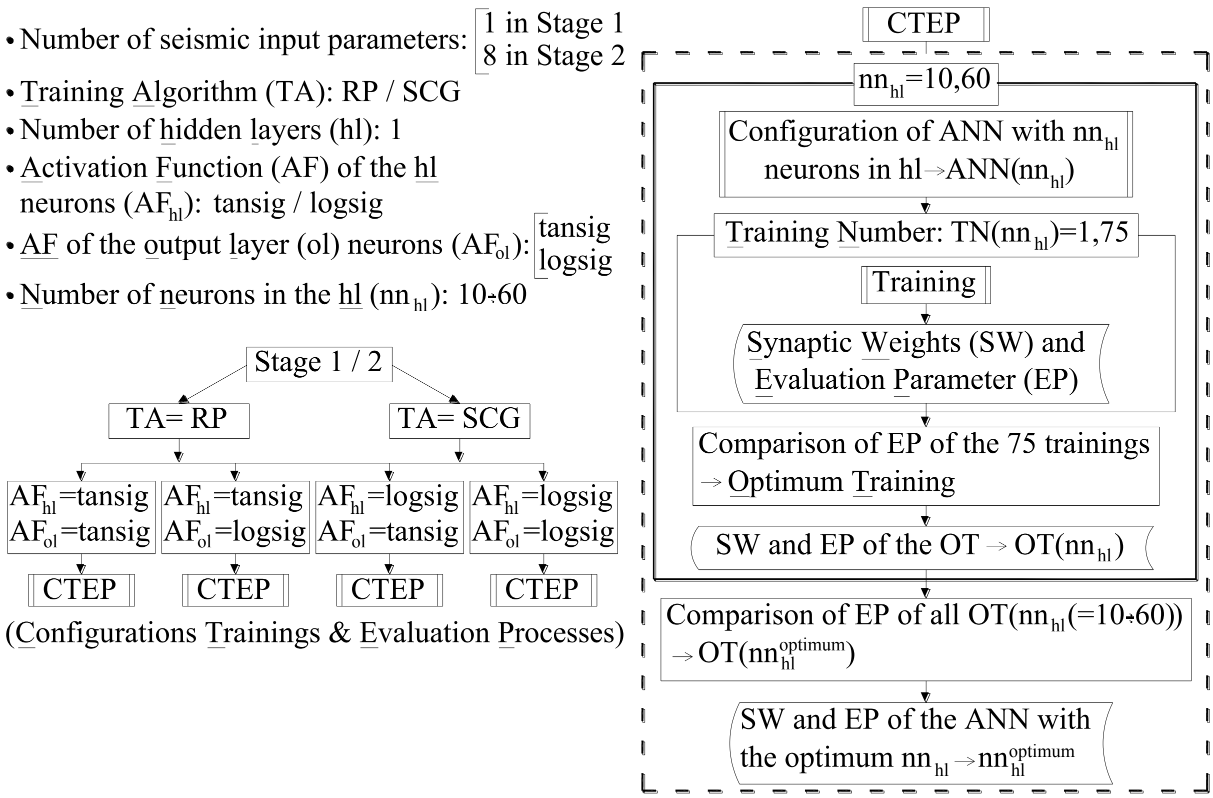


Figure 7: Calculations' procedure of Stages 1 and 2

It must be noted that the Evaluation Parameter (EP) in Figure 7 is the Overall Accuracy (OA) index which is defined in Figure 6. It must also be stressed that the Evaluation Parameter OA was calculated in any case on the basis of the samples which belong to the testing sub-set. This set, is used to control the generalization ability of the trained ANNs. Thus, the results of the current investigation have wider validity because they are exported from a data sub-set which is comprised of training and target vectors that are not used to optimize the values of ANNs' synaptic weights, but is comprised of unseen vectors. As turns out from the Figure 7, for each ANN which was configured, 75 training procedures were conducted. This is due to the fact that differences in the performance of the networks are caused by the initial values of the synaptic weights and biases (see e.g. [11]), and also due to the random composition of the three sub-sets of the data set (training, validation and testing data sub-sets). The best trained network (optimum trained network) for each one of the different ANN architectures was finally adopted in the procedure.

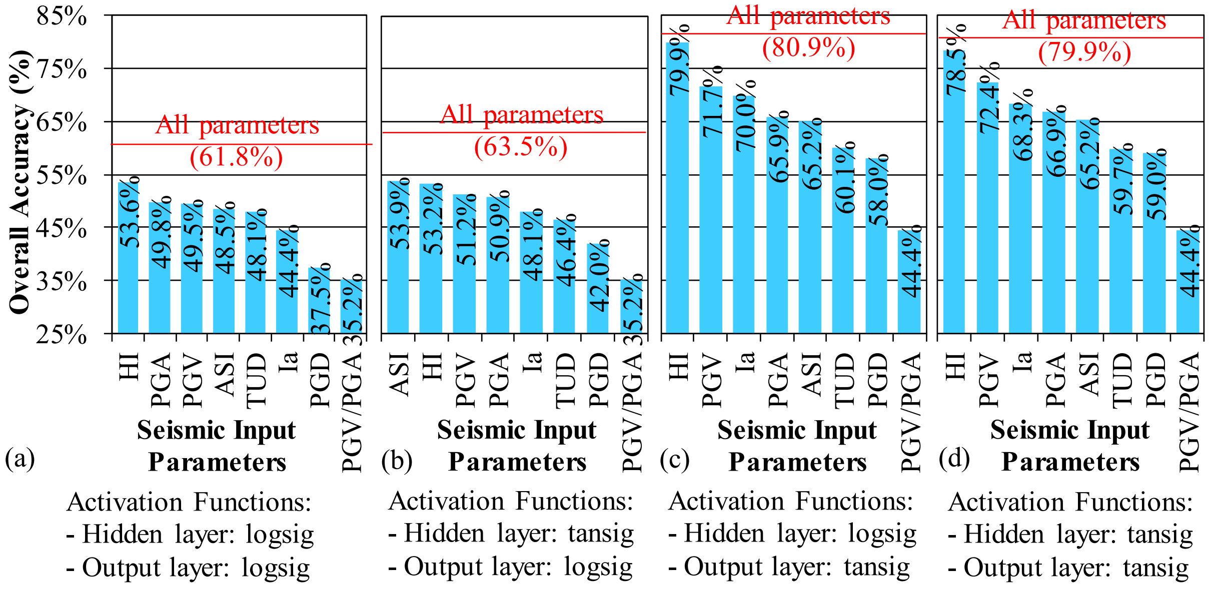


Figure 8: Overall Accuracy which is extracted from ANNs with one hidden layer and one seismic input parameter trained using the RP algorithm

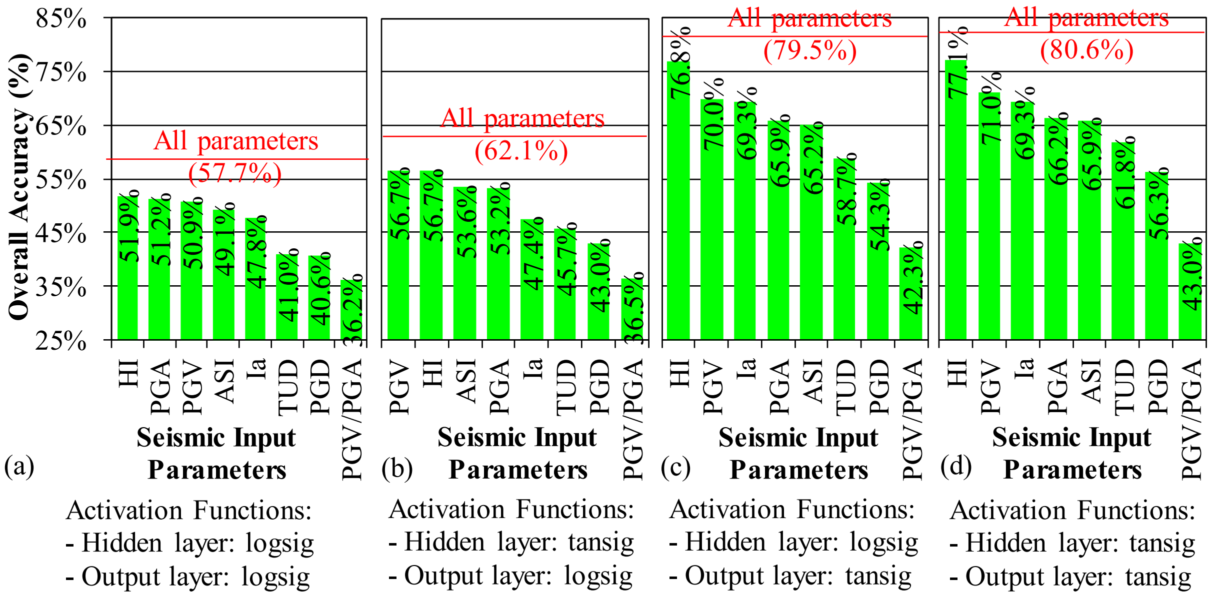


Figure 9: Overall Accuracy which is extracted from ANNs with one hidden layer and one seismic input parameter trained using the SCG algorithm

Figures 8 and 9 present the results of analyses of Stages 1 and 2. From these Figures, the following conclusions can be drawn:

- The ANNs in which the tansig function is used in the output layer extract higher values for OA index than the values which are extracted from ANNs with logsig function in this layer. This conclusion is valid either in cases in which one of the studied seismic input parameter is used or in the case in which all studied seismic input parameters are introduced to ANNs (horizontal red lines in Figures 8 and 9). For example, when the HI parameter is used as seismic input parameter the values of the OA index which are exported from ANNs with tansig function in their output layer fluctuate between 76.8% and 79.9%, whereas the corresponding fluctuation range in the case of the utilization of logsig function in the output layer is 51.9%-56.7%.

- The training algorithm has significantly less influence to the ANNs' prediction ability than the type of the output layer's activation function. For example, when all the studied seismic input parameters are introduced to networks the value of the OA index is equal to 80.9% in the case of utilization of RP algorithm and equal to 79.5% in the case of utilization of SCG algorithm for the training of ANNs with logsig function in the hidden layer and tansig function in the output layer (Figures 8(c) and 9(c)). Similar differences are observed for all the other combinations of activation functions in the hidden and the output layer.
- The ranking of the seismic input parameters concerning their importance (i.e. concerning the value of OA index which is extracted from ANNs when only one of them is imported to input vectors) is identical and independent of the training algorithm and of the type of hidden layer's activation functions only in the case of utilization of function tansig in the output layer (Figures 8(c), 8(d), 9(c), 9(d)).
- In almost all cases which are presented in Figures 8 and 9 the most effective seismic input parameter is the Housner Intensity (HI). Using this parameter as the only seismic input in the input vectors, values between 76.8% and 79.9% for the OA index are achieved in the case of utilization of the tansig activation function in the output layer. The corresponding values in the case of utilization of the logsig function range between 51.9% and 56.7%. On the contrary, the less effective seismic input parameters in all studied cases are the ratio PGV/PGA and the PGD. It must also be noted that the difference between the OA values which are extracted using the most effective seismic parameter and the corresponding values which are extracted using the less effective seismic parameter is almost double in the case in which the tansig function is used in the networks' output layer over the case in which the logsig function is used.
- The level of effectiveness of the HI parameter as the only seismic input parameter in input vectors approaches the corresponding level which is achieved when all seismic parameters are introduced in input vectors, mainly in the case of the utilization of tansig function in the output layer. As turns out from the Figures 8(c), 8(d), 9(c), 9(d) the OA values which are extracted using only the HI parameter in the input vectors are only 1.3%-4.2% smaller than the corresponding values which are extracted when all seismic parameters are introduced. Similar level of divergence is not observed if the logsig function is used in the output layer.

As mentioned above, in the Stage 3 of the investigation of one hidden layer ANNs' performance concerning to the seismic damage prediction, the optimum number and combination of seismic input parameters (i.e. the number and the combination of seismic input parameters which lead to the optimum ANNs' predictions for the seismic damage state) was investigated. To this end the two versions of the "Stepwise" sensitivity analysis method i.e. the "Forward Stepwise Method" (FSM), and the "Backward Stepwise Method" (BSM) were utilized. The "Stepwise" method is widely used for the investigation of importance of the parameters which affect multi-parametric problems with unknown mathematical formulation. The Stepwise Method consists of adding (FSM) or rejecting (BSM) step-by-step one input parameter of the network [37]. Thus, at each step, a new ANN is configured and is trained by the use of the available training data-set. The performance of the new network of each step is assessed (i.e. the OA index is calculated). On the basis of the changes of the performance parameter's values, the input parameters are ranked according to their importance. The sequence of FSM's and BSM's steps is presented in details by Morfidis and Kostinakis [40]. It must also be noted that the required procedure at each step of the Stepwise method (FSM or BSM) is the procedure which is described in the Figure 7. Thus, it is obvious that the amount of calculations which are required for the complete implementation of the FSM and the BSM is significantly great. In order to make this statement clear it can be noted that in the present study (in which the influence of 8 seismic input parameters was examined) the full implementation of FSM and BSM required 8 steps each [40]. In each of these steps the procedure of Figure 7 was implemented. For this reason, in the present study the full implementation of the FSM and BSM was conducted only for ANNs with tansig function in the hidden and the output layer, but for the both of two training algorithms which were utilized (RP and SCG).

Figure 10 illustrates the results of the full implementation of the “Stepwise” method. More specifically, the maximum values of the OA index for each number of seismic input parameters between 1 and 8 in input vectors are presented. These values are the maximum OA index values which are exported from all the examined combinations of each method’s step. In Figure 10 the numbers of neurons in the hidden layer of ANNs which export the maximum OA index values are also presented.

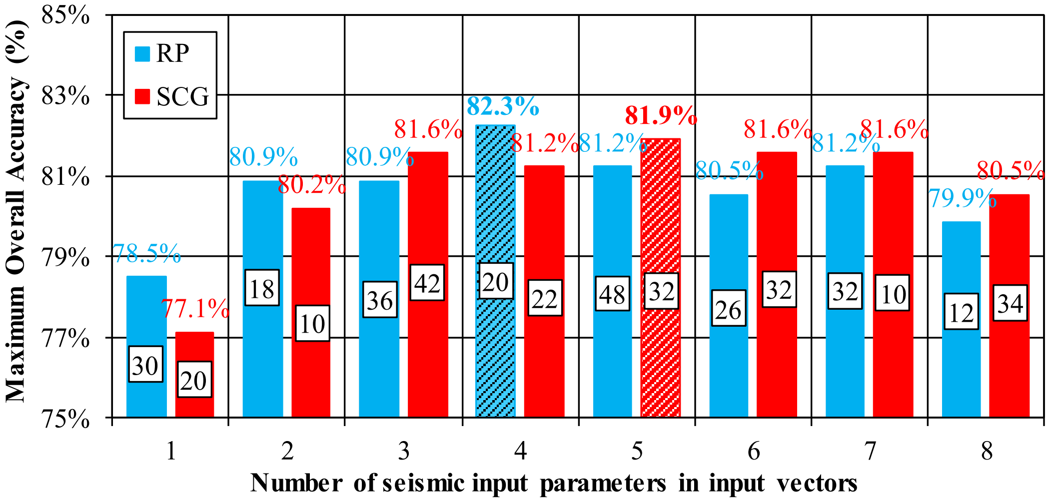


Figure 10: Maximum values of the Overall Accuracy index from the examined combinations of seismic input parameters in the context of the implementation of the “Stepwise” method

The main conclusions which arise from the Figure 10 are the following:

- With the exception of the case of only one seismic input parameter, in all other cases the maximum value of OA index fluctuates within a narrow range of values (2.39% in the case of utilization of RP algorithm and 1.71% in the case of utilization of SCG algorithm).
- The best performance of ANNs in the case of using the RP training algorithm is accomplished when 4 seismic parameters are introduced in input vectors (OA=82.3%). The corresponding number of seismic parameters in the case of use of the SCG algorithm is 5 (OA=81.9%).
- The optimum number of neurons in the hidden layer is altered randomly. This conclusion substantiates the nonexistence of a direct method for its calculation, and the recourse to the trial and error procedure.
- The two utilized training algorithms lead to results which do not differ significantly. Thus, the factor “training algorithm” is not considered as a critical factor in the present study.

In the Table 5 the optimum combinations of seismic input parameters for every number of seismic parameters which are introduced in input vectors are presented.

	RP algorithm	SCG algorithm
1	HI	HI
2	HI-PGD	HI-PGV
3	HI-PGD-PGA	HI-PGV-PGD
4	HI-PGD-PGA-ASI	HI-PGV-PGD-PGA
5	HI-PGD-PGA-ASI-PGV/PGA	HI-PGV-PGD-PGA-PGV/PGA
6	HI-PGD-PGV-ASI-PGV/PGA-Ia	HI-PGV-PGD-PGA-PGV/PGA-Ia
7	HI-PGD-PGA-PGV-ASI-PGV/PGA-TUD	HI-PGV-PGD-PGA-PGV/PGA-Ia-ASI

Table 5: Optimum combinations for each number of seismic input parameters – "Stepwise" method

From the combined study of the Figure 10 and the Table 5 it is concluded that the optimum combination of seismic parameters is consisted of the seismic parameters HI-PGD-PGA-ASI in the case of utilization of the RP algorithm, and of the seismic parameters HI-PGV-PGD-PGA-PGV/PGA in the case of utilization of the SCG algorithm.

(a)

ANN's classification (True (target) class in rows)

	1	2	3	4	5	R
1	**42**	4	1	0	0	89.4%
2	5	**20**	4	0	0	69.0%
3	0	15	**55**	14	3	63.2%
4	0	0	8	**24**	2	70.6%
5	0	0	0	7	**89**	92.7%
P	89.4%	51.3%	80.9%	53.3%	94.7%	***78.5%***

RP algorithm-tansig/tansig - 30 neurons in the hidden layer

ANN's classification (True (target) class in rows)

	1	2	3	4	5	R
1	**37**	4	0	0	0	90.2%
2	3	**23**	10	3	0	59.0%
3	0	5	**54**	8	3	77.1%
4	0	0	11	**36**	12	61.0%
5	0	0	1	8	**75**	89.3%
P	92.5%	71.9%	71.1%	65.5%	83.3%	***76.8%***

SCG algorithm-tansig/tansig -20 neurons in the hidden layer

(b)

ANN's classification (True (target) class in rows)

	1	2	3	4	5	R
1	**40**	9	0	0	0	81.6%
2	6	**22**	10	0	0	57.9%
3	0	6	**47**	7	1	77.0%
4	0	0	5	**39**	7	76.5%
5	0	0	0	8	**86**	91.5%
P	87.0%	59.5%	75.8%	72.2%	91.5%	***79.9%***

RP algorithm-tansig/tansig - 12 neurons in the hidden layer

ANN's classification (True (target) class in rows)

	1	2	3	4	5	R
1	**43**	10	0	0	0	81.1%
2	6	**34**	3	0	0	79.1%
3	0	0	**53**	7	2	85.5%
4	0	0	12	**26**	8	56.5%
5	0	0	0	9	**80**	89.9%
P	87.8%	77.3%	77.9%	61.9%	88.9%	***80.5%***

SCG algorithm-tansig/tansig -34 neurons in the hidden layer

(c)

ANN's classification (True (target) class in rows)

	1	2	3	4	5	R
1	**44**	6	0	0	0	88.0%
2	3	**26**	5	0	0	76.5%
3	0	5	**36**	8	2	70.6%
4	0	0	10	**42**	10	67.7%
5	0	0	0	3	**93**	96.9%
P	93.6%	70.3%	70.6%	79.2%	88.6%	***82.3%***

RP algorithm-tansig/tansig 20 neurons in the hidden layer

ANN's classification (True (target) class in rows)

	1	2	3	4	5	R
1	**40**	5	0	0	0	88.9%
2	2	**19**	7	0	0	67.9%
3	0	6	**54**	9	1	77.1%
4	0	2	7	**30**	4	69.8%
5	0	0	0	10	**97**	90.7%
P	95.2%	59.4%	79.4%	61.2%	95.1%	***81.9%***

SCG algorithm-tansig/tansig-32 neurons in the hidden layer

Figure 11: CMs of classifications made by the best trained (single-hidden layered) ANNs with: (a) one seismic input parameter (HI parameter), (b) eight seismic input parameters, (c) the optimum combination of the seismic input parameters

The evaluation of the performance of ANNs using the OA index is a procedure which leads directly to results (i.e. direct comparison between the results which are exported from various ANNs' configuration by means of only one index). However, more detailed study of the evaluation of the performance of ANNs in the pattern recognition problem's solution is accomplished through the study of the corresponding CMs and their related metrics (Figure 6). In Figure 11 the CMs of the classifications made by the best trained ANNs with one hidden layer are illustrated. In the blue colored cells of these matrices the OA index values which have been used above for the evaluation of the performance of ANNs are placed. The main characteristic of CMs is the information which they give about not only the total accuracy of predictions (this is given by the OA index) but also about the percentages of correct classifications of samples to the classes in which the examined objects can be classified. To this end the indices R (Recall index) and P (Precision index) are defined (Figure 6). Valuable information about the quality of the predictions of ANNs is also given by the configuration of CMs. More specifically, when the vast majority of the non-zero elements of a CM are located about the main diagonal means that the ANN achieves an acceptable classification. For example, if all the non-zero elements are located in the cells of the main diagonal and in the adjacent cells means that the objects are classified into correct classes and into classes which are adjacent to them (i.e. if the correct class for an object is the class i, the ANN classifies this object in the class i-1 or in class i+1). Therefore, all predictions/classifications about the expected damage states which are made by the ANN in this case are correct or nearby to the correct ones.

The main conclusions which can be drawn from the Figure 11 are the following:

- The difference of the quality of classification which is achieved by the ANNs trained using the RP algorithm and by the ANNs trained using the SCG algorithm is least and not clear.
- The vast majority of classifications in all cases which are presented in the Figure 11 are in the correct classes (green colored cells) or in classes adjacent to them (yellow colored cells). The classifications in categories which are not adjacent to the correct category (grey colored cells) are generally very few, especially in the cases in which all seismic parameters or their optimum combination are introduced to input vectors (Figures 11(b), 11(c)). For example, in the case of utilizing the RP algorithm, 4 classifications are made in classes not adjacent to correct ones when one seismic parameter (the HI parameter) was introduced in input vectors (Figure 11(a)). When the eight examined seismic parameters are introduced to input vectors, the corresponding number of classifications is 1 (Figure 11(b)).
- The percentages of correct classifications in each one of the 5 damage classes (values of indices R and P) are generally high. These percentages are significantly high in the case of classes 1 and 5 (with very little exceptions, they are greater than 85%). On the contrary relatively lower percentages of correct classifications in the classes 2, 3 and 4 are achieved, especially in the case of introduction of 1 seismic parameter (HI parameter) to input vectors.
- The quality of classification from the best trained ANNs with one hidden layer can be rated as acceptable, because, besides the achieved high percentages of correct classifications (OA=80%-82%), the approach which is accomplished for the damage state of the overall testing sample is very satisfactory.

4.2 Investigation of the influence of the second hidden layer to ANNs' performance

In the subsection 4.1 the results of the investigation for the optimum configuration of ANNs with one hidden layer (first part of the parametric investigation for the optimum performance of ANNs) were presented. As it was mentioned in the introduction of the section 4, the second part of the parametric investigation concerns the performance of networks with two hidden layers. For reasons which were emphasized at that point, the investigation of ANNs with two hidden layers was based on the basic conclusions which were resulted from the corresponding investigation of the single layered ANNs. Thus, networks with tansig functions in the two hidden and the output layer were examined. Due to the fact that no clear conclusion was extracted about the more effective training algorithm between RP and SCG, both algorithms were utilized in the second part of investigation. Furthermore, the optimum combinations of the number of neurons in the two hidden layers were also examined. To this end, networks with all possible combinations of numbers of neurons in the two hidden layers between 10 and 60 were configured. Finally, as regards the seismic input parameters, three cases which correspond to the three stages of the investigation of the single layered ANNs were examined. The only difference is that at the 3rd stage it was not conducted analysis with the "Stepwise" method. Therefore, the results (i.e. the optimum seismic parameter combinations) of the "Stepwise" method which was conducted for the single layered networks were utilized. In the Figure 12 the procedure of the second part of the parametric investigation is illustrated.

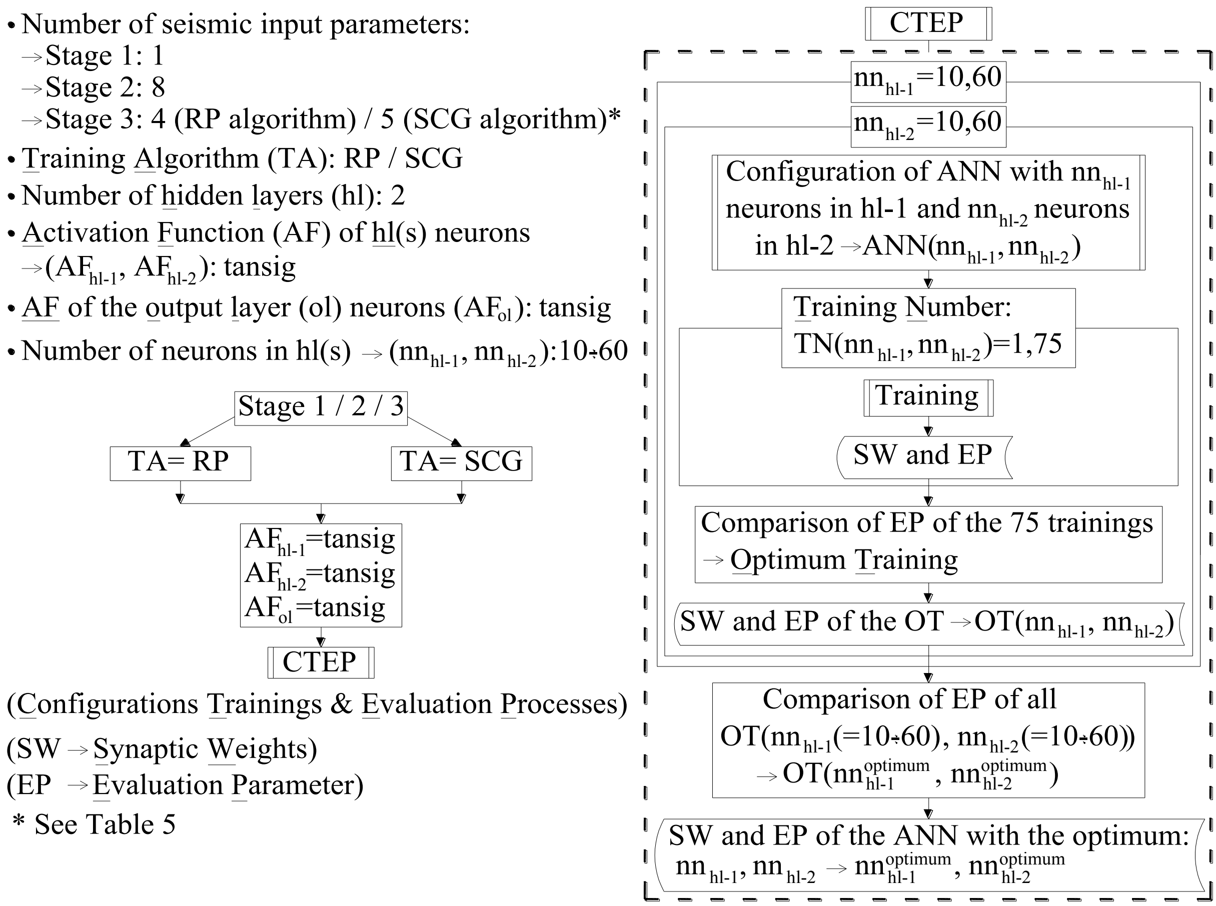


Figure 12: Calculations' procedure of the second part of parametric investigation (ANNs with two hidden layers)

Figure 13 illustrates the ranking of the examined seismic input parameters on the basis of their importance when ANNs with one and two hidden layers are utilized. In this Figure the numbers of neurons in hidden layers which export the maximum OA index values (the notation n/m states: n neurons in the first hidden layer and m neurons in the second hidden layer) are also presented.

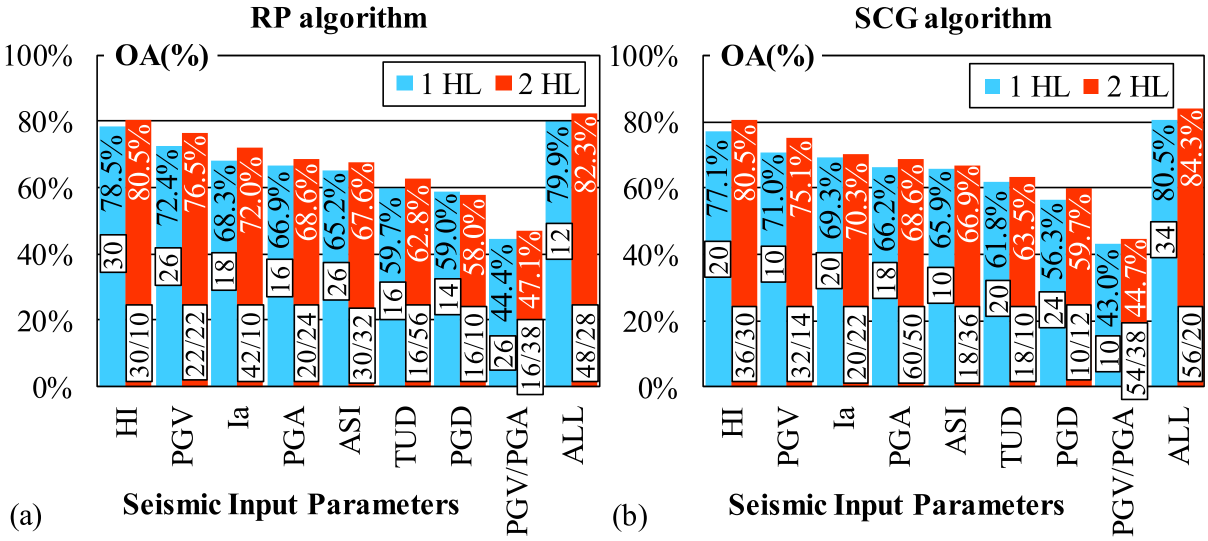


Figure 13: Ranking of the examined seismic input parameter which is resulted from ANNs with one and two hidden layers

The main conclusions which are extracted from the Figure 13 are the following:

- The ranking of the examined seismic parameters which is extracted using single layered ANNs or ANNs with two hidden layers is the same. This conclusion is valid in the case of networks with tansig function in all neurons' layers (hidden and output), and for both of two utilized algorithms.
- The use of two hidden layers increases the effectiveness of networks but not significantly. For example, the value of the OA index in the case in which one seismic parameter is introduced to input vectors (HI parameter) is 78.5% when a single layered network is used, while when a network with two hidden layers is utilized the corresponding value is 80.5% (Figure 13(a)). Hence, the use of a second hidden layer leads to increase of the OA value about 2.5% in the case of use of RP algorithm. The corresponding increase when the SCG algorithm is utilized is about 4.0% (Figure 13(b)). When all the examined seismic parameters are introduced to input vectors, the increase of the OA value due to the introduction of the second hidden layer is about 3.0% in the case of utilization of the RP algorithm and about 4.5% in the case of utilization of the SCG algorithm.
- The level at which the second hidden layer improves the classification quality does not depend on the utilized training algorithm, because in some cases the increase of the OA value is greater when the RP algorithm is used (for example when the parameter Ia or the parameter TUD introduced to input vectors), while in some others is greater when the SCG algorithm is utilized (for example if the parameter HI is introduced to the input vectors).

In Figure 14 the CMs of the classifications made by the best trained ANNs with two hidden layers are illustrated. The combined study of the Figures 11 and 14 indicates that, while the introduction of the second hidden layer does not increase significantly the value of the OA index, it improves the percentages of correct classifications in damage classes 2, 3, 4 which are not high in the case of utilization of single layered networks. More specifically, the relatively low values of the Recall and the Precision indices for these damage classes (for example about 50%-60% when the HI parameter is introduced in input vectors) when single layered networks are used, are increased up to 60% in the case of utilization of ANNs with two hidden layers. Moreover, in the case of introduction of all the examined seismic parameters in input vectors the most of values of the Recall and the Precision indices are greater than 75%.

(a)

ANN's classification (True (target) class × ANN's classification)

True (target) class	1	2	3	4	5	R
1	**42**	7	1	0	0	84.0%
2	2	**28**	3	0	0	84.8%
3	0	10	**56**	12	1	70.9%
4	0	0	5	**31**	9	68.9%
5	0	0	0	7	**79**	91.9%
P	95.5%	62.2%	86.2%	62.0%	88.8%	***80.5%***

RP algorithm - activation functions: tansig/tansig/tansig
30/10 neurons in hidden layers

ANN's classification (True (target) class × ANN's classification)

True (target) class	1	2	3	4	5	R
1	**38**	2	0	0	0	95.0%
2	5	**18**	5	0	0	64.3%
3	0	6	**53**	9	1	76.8%
4	0	0	10	**32**	8	64.0%
5	0	0	0	11	**95**	89.6%
P	88.4%	69.2%	77.9%	61.5%	91.3%	***80.5%***

SCG algorithm - activation functions: tansig/tansig/tansig
36/30 neurons in hidden layers

(b)

ANN's classification (True (target) class × ANN's classification)

True (target) class	1	2	3	4	5	R
1	**35**	5	1	0	0	85.4%
2	5	**39**	3	0	0	83.0%
3	0	4	**46**	8	0	79.3%
4	0	0	12	**36**	5	67.9%
5	0	0	0	9	**85**	90.4%
P	87.5%	81.3%	74.2%	67.9%	94.4%	***82.3%***

RP algorithm - activation functions: tansig/tansig/tansig
48/28 neurons in hidden layers

ANN's classification (True (target) class × ANN's classification)

True (target) class	1	2	3	4	5	R
1	**42**	3	0	0	0	93.3%
2	8	**32**	3	0	0	74.4%
3	0	8	**55**	5	1	79.7%
4	0	0	3	**47**	10	78.3%
5	0	0	0	5	**71**	93.4%
P	84.0%	74.4%	90.2%	82.5%	86.6%	***84.3%***

SCG algorithm - activation functions: tansig/tansig/tansig
56/20 neurons in hidden layers

(c)

ANN's classification (True (target) class × ANN's classification)

True (target) class	1	2	3	4	5	R
1	**38**	4	0	0	0	90.5%
2	0	**35**	6	0	0	85.4%
3	0	9	**55**	10	1	73.3%
4	0	0	9	**33**	6	68.8%
5	0	0	0	7	**80**	92.0%
P	100%	72.9%	78.6%	66.0%	92.0%	***82.3%***

RP algorithm - activation functions: tansig/tansig/tansig
10/34 neurons in hidden layers

ANN's classification (True (target) class × ANN's classification)

True (target) class	1	2	3	4	5	R
1	**40**	3	0	0	0	93.0%
2	1	**30**	6	0	0	81.1%
3	0	3	**52**	11	0	78.8%
4	0	0	11	**35**	9	63.6%
5	0	0	0	6	**86**	93.5%
P	97.6%	83.3%	75.4%	67.3%	90.5%	***82.9%***

SCG algorithm - activation functions: tansig/tansig/tansig
14/38 neurons in hidden layers

Figure 14: CMs of classifications made by the best trained (two-hidden layered) ANNs with: (a) one seismic input parameter (HI parameter), (b) eight seismic input parameters, (c) the optimum combination of the seismic input parameters

Finally, it must be noted that the number of classifications in classes which are not adjacent to the correct ones is least in the case of networks with two hidden layer. More specifically, in the case of introduction of HI in input vectors this number is 2 when the RP algorithm is used and 1 when the selected algorithm is the SCG. The corresponding number in the case of single layered networks is 4 and 7 respectively (Figure 11(a)). Additionally, in all cases which are illustrated in Figures 14(b), (c) the number of classifications to classes not adjacent to the correct ones is equal to 1 or 0.

In the Table 6 the configuration of the best trained networks which were used for the evaluation of ANNs' seismic damage predictions accuracy in the case of earthquakes that are not included to the training data-set is illustrated. The results of this evaluation are presented in the following section.

ANN's name	Number of hidden layers	Number of neurons in hidden layers	Activation functions	Training algorithm	Number of seismic input parameters
N1-RP-20	1	20	tan/tan	RP	4*
N1-SCG-32	1	32	tan/tan	SCG	5*
N2-RP-48/28	2	48/28	tan/tan/tan	RP	8
N2-SCG-56/20	2	56/20	tan/tan/tan	SCG	8

* The optimum combination of seismic input parameters of Table 5

Table 6: The configuration parameters of best trained ANNs

5 EVALUATION OF THE BEST TRAINED ANNS' SEISMIC DAMAGE PREDICTIONS FOR UNKOWN EARTHQUAKES

In the previous section were presented the investigation which was conducted for the definition of optimum configurations of ANNs' using as criterion the maximization of the percentage of the r/c buildings' correct classification into predefined seismic damage classes (Table 2). The ANNs' predictions accuracy was also evaluated using also the other metrics which are defined on the basis of the CMs (Figure 6). In the Table 6 the ANNs which achieved the best predictions/classifications of the seismic damage of the testing sub-set's buildings are illustrated.

In the current section the prediction ability of ANNs presented in Table 6 is further investigated using a set of samples which are not included to the training data-set. To this end, additional time history analyses were conducted using the 30 selected r/c buildings (Table 3) and 16 pairs of horizontal bidirectional ground motions (Table 7) different from the 65 ground motions which were used for the generation of the training data-set. Thus, a new testing set of 480(=30x16) samples was generated.

Ground Motion Parameter	Units	Minimum Value	Maximum Value
PGA	%g	0.014	1.20
PGV	cm/sec	1.14	123.44
PGD	cm	0.13	57.73
I_a	m/sec	0.003	11.07
ASI	g·sec	0.015	0.98
HI	cm	3.47	471.17
PGV/PGA	sec	0.04	0.29
TUD	sec	≈0.0	14.57

Table 7: Ranges of the selected seismic parameters' values corresponding to the 16 testing earthquakes

Then, the best configured ANNs (Table 6) were used to classify the 480 samples of the new testing set to the seismic damage classes of the Table 2. Similarly to the case of the investigation of the optimum ANNs' configuration, the OA index, as well as the other metrics which are related to the CMs, were utilized for the evaluation of the networks' performance in the current case.

In the Figure 15 the CMs of the classifications made by the best configured ANNs for the samples of the new testing set are illustrated.

(a) ANN's classification

True (target) class	1	2	3	4	5	R
1	**110**	2	0	0	0	98.2%
2	16	**43**	40	0	0	43.4%
3	0	7	**43**	21	5	56.6%
4	0	0	7	**13**	10	43.3%
5	0	0	0	11	**152**	93.3%
P	87.3%	82.7%	47.8%	28.9%	91.0%	***75.2%***

RP -tan/tan - 20 neurons in the hl - 4 seismic inputs

(b) ANN's classification

True (target) class	1	2	3	4	5	R
1	**115**	6	2	0	0	93.5%
2	11	**39**	40	1	0	42.9%
3	0	7	**38**	19	3	56.7%
4	0	0	10	**13**	6	44.8%
5	0	0	0	12	**158**	92.9%
P	91.3%	75.0%	42.2%	28.9%	94.6%	***75.6%***

SCG - tan/tan -32 neurons in the hl - 5 seismic inputs

(c) ANN's classification

True (target) class	1	2	3	4	5	R
1	**115**	9	0	0	0	92.7%
2	11	**38**	40	0	0	42.7%
3	0	5	**42**	16	3	63.6%
4	0	0	8	**16**	4	57.1%
5	0	0	0	13	**160**	92.5%
P	91.3%	73.1%	46.7%	35.6%	95.8%	***77.3%***

RP - tan/tan/tan - 48/28 neurons in hls - 8 seismic inputs

(d) ANN's classification

True (target) class	1	2	3	4	5	R
1	**108**	2	0	0	0	98.2%
2	18	**45**	32	0	0	47.4%
3	0	5	**49**	19	3	64.5%
4	0	0	9	**11**	6	42.3%
5	0	0	0	15	**158**	91.3%
P	85.7%	86.5%	54.4%	24.4%	94.6%	***77.3%***

SCG - tan/tan/tan -56/20 neurons in hls -8 seismic inputs

Figure 15: CMs of classifications made by the best trained ANNs for the samples of the new testing set: (a) N1-RP-20, (b) N1-SCG-32, (c) N2-RP-48/28, (d) N2-SCG-56/20

From the Figure 15, the following main conclusions can be drawn:

- The percentages of the correct classifications (i.e. the values of the OA index) for the samples of the new testing set are lower than the corresponding percentages which are extracted from the examined networks for the samples of the training data-set's testing sub-set. For example, the value of the OA index which is extracted from the network N1-RP-20 is 82.3% (Figure 11(c)) for the samples of the training data-set's testing sub-set, whereas the corresponding value for the samples of the new testing set is 75.2% (reduction 8.6%). Similar reductions arise for the other examined networks. This decrease of the values of the OA index is not significant if we take into consideration the fact that the input vectors of the new testing set include values of seismic parameters which are unknown to the used networks.
- Greater reductions than those of the OA index arise for the indices R and P especially for the damage classes 2, 3, 4, whereas the corresponding values of these parameters for the classes 1 and 5 remain in the same level. This conclusion is valid for all the examined networks and it is independent of the choice of the number of hidden layers and the utilized training algorithm.
- The general configuration of the CMs of the Figure 15 is similar to the configuration of the CMs which correspond to the classification of the testing sub-set's samples as regards the location of the non-zero elements. As mentioned in the section 4, the CMs in which the vast majority of the non-zero elements are located about the main diagonal (green and yellow colored cells) correspond to classifications which can be evaluated as acceptable. Therefore, the examined networks are capable to give a reliable vision of the expected damage state even for unknown earthquakes on the basis of their existing training. This conclusion is more valid in the case of networks with two hidden layers (only three samples are classified into classes not adjacent to the correct ones – see the grey colored cells in the Figures 15(c), (d)).

Finally, it must be stressed that the aforementioned reduction of the percentages of the overall correct classifications (OA index) as well as the reduction of the quality of classifications to the classes 2, 3, 4 can be attributed to the fact that the selection of the 16 testing earthquakes was done on purpose, in such a way in order to have values for seismic parameters which are out of the bounds of the corresponding parameters' values of the 65 seismic excitations which are used for the networks' training (see Table 4 and the red colored numbers in Table 7).

6 CONCLUSIONS

In the present paper it was investigated the ability of Artificial Neural Networks (ANN) to successfully approach the solution of the reinforced concrete (r/c) buildings' seismic damage prediction problem when it is formulated in terms of a pattern recognition problem. To this end, Multilayer Feedforward Perceptron Networks (MFP) were utilized. Two training algorithms, namely the Scaled Conjugate Gradient algorithm (SCG) and the Resilient Back-Propagation algorithm (RP), were used for the training of ANNs. The required training dataset was generated using results from Nonlinear Time History Analyses (NTHA) of 30 r/c buildings with different heights, structural systems and structural eccentricities, subjected to 65 actual ground motions. The selected r/c buildings were designed following the provisions of Eurocodes. In total, 1950 NTHA were conducted. From each one of them the damage index formulated in terms of Maximum Interstorey Drift Ratio (MIDR) was calculated. In order to investigate the optimum ANNs' performance several networks with different configurations were examined. More specifically, ANNs with different number of hidden layers (1 or 2), different number of neurons in the hidden layers (between 10 and 60), as well as different activation functions of neurons (sigmoid function (logsig) or hyperbolic tangent function (tansig)) were formed. Furthermore, the number and the combination of the seismic input parameters which lead to the optimum ANNs' predictions was investigated. To this end, 8 different and widely used seismic parameters were examined using the "Stepwise" sensitivity analysis method. Finally, the prediction ability of the best configured ANNs was further investigated using scenarios of future earthquakes. In the investigation procedure the ANNs' performance was evaluated in any case using the percentages of correct classifications of buildings into pre-defined damage classes (Overall Accuracy (OA) index) as well as the corresponding Confusion Matrices (CM). The basic conclusions that turned out from the present investigation are the following:

- The utilization of the tansig function in the output layer's neurons leads in all examined cases to much more exact predictions than those which are extracted when the function logsig is used.
- The used training algorithm affects the performance of ANNs but no clear conclusion is extracted about the more effective training algorithm between RP and SCG. Thus, the factor "training algorithm" is not considered as a critical factor in the present study.
- When only one seismic parameter is introduced to input vectors, the Housner Intensity (HI) parameter is the most effective one in almost all examined cases. This conclusion is valid in any case if the tansig function is used for the neurons of the output layer.
- The ranking of the seismic input parameters concerning their importance (i.e. concerning the value of OA) is identical and independent of the training algorithm, of the type of hidden layer's activation functions and of the number of hidden layers only in the case of utilization of function tansig in the output layer.
- The optimum number and combination of the seismic parameters depends on the used training algorithm, and on the activation functions of the hidden and output layer.

- The utilization of second hidden layer improves the quality of ANNs' predictions especially when the percentages of correct classifications into individual classes instead of the overall percentages of correct classifications are examined.
- The vast majority of classifications which are made by the ANNs correspond to the correct classes or to classes which are adjacent to the correct ones. This conclusion is valid in both cases of ANNs with one or two hidden layers. It also valid for the samples of the training data-set's testing sub-set, as well as for the samples of the data-set which is formed considering earthquakes unknown to the trained ANNs. Therefore, the examined networks are capable to give a reliable vision of the expected damage state even for unknown earthquakes on the basis of their existing training.

REFERENCES

[1] FEMA-356, *Prestandard and Commentary for the Seismic Rehabilitation of Buildings.* Federal Emergency Management Agency, Washington DC, 2000.

[2] EN1998-3, *Design of structures for earthquake resistance – Part 3: Assessment and retrofitting of buildings*. European Committee for Standardization, Brussels, 2005.

[3] ASCE/SEI 41-06, *Seismic Rehabilitation of Existing Buildings*. American Society of Civil Engineers, ASCE, Virginia, 2009.

[4] FEMA-154, *Rapid visual screening of buildings for potential seismic hazards: A Handbook (FEMA-154)*. 2nd Edition, Washington DC, 2002.

[5] ATC, *Earthquake Damage Evaluation Data for California, ATC-13 Report*. Redwood City, CA: Applied Technology Council, 1985.

[6] A.J. Kappos, G. Panagopoulos, Ch. Panagiotopoulos, Gr. Penelis, A Hybrid Method for the vulnerability assessment of R/C and URM buildings. *Bulletin of Earthquake Engineering*, **4(4)**, 391-413, 2006.

[7] S. Haykin, *Neural Networks and Learning Machines, 3rd Edition*, Prentice Hall, 2009.

[8] M. McCulloch, W. Pitts, A logical calculus and the ideas immanent in the nervous activity. *Bulletin of Mathematical Biophysics*, **5**, 115-133, 1943.

[9] G. Molas, F. Yamazaki, Neural Networks for quick Earthquake Damage Estimation. *Earthquake Engineering and Structural Dynamics*, **24**, 505-516, 1995.

[10] N. Lagaros, M. Fragiadakis M, Fragility assessment of steel frames using neural networks. *Earthquake Spectra*, **23(4)**, 735-752, 2007.

[11] O.R. Lautour, P. Omenzetter, Prediction of seismic-induced structural damage using artificial neural networks. *Engineering Structures*, **31**, 600-606, 2009.

[12] S.L. Dimova, P. Negro, Seismic assessment of an industrial frame structure designed according to Eurocodes. Part 2: Capacity and vulnerability. *Engineering Structures*, **27(5)**, 724–735, 2005.

[13] M.F. Moller, A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, **6**, 525-533, 1993.

[14] M. Riedmiller, H. Braun, A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm. *Proceedings of IEEE*, San Francisco, 586-591, 1993.

[15] S. Asht, R. Dass, Pattern Recognition Techniques: A Review. *International Journal of Computer Science and Telecommunications*, **3(8)**, 25-29, 2012.

[16] B. Erkus, Utilization of Artificial Neural Networks in Building Damage Prediction. *MSc Thesis*. Middle East Technical University, Ankara, 1999.

[17] M.H. Arslan, An evaluation of effective design parameters on earthquake performance of RC buildings using neural networks. *Engineering Structures*, **32(7)**, 1888-1898, 2010.

[18] T.K. Šipoš, V. Sigmund, M. Hadzima-Nyarko, Earthquake performance of infilled frames using neural networks and experimental database. *Engineering Structures*, **51**, 113-127, 2013.

[19] S.L. Kramer, *Geotechnical Earthquake Engineering*. Prentice-Hall, 1996.

[20] SeismoSoft, *SeismoSignal v.5.1.0.*, 2015: www.seismosoft.com

[21] A.J. Kappos, Seismic damage indices for RC buildings: evaluation of concepts and procedures. *Construction Research Communications Limited ISSN 1365-0556*, 78-87, 1997.

[22] F. Naeim, *The seismic design handbook, 2nd edition*. Kluwer Academic, Boston, 2011.

[23] A. Elenas, K. Meskouris, Correlation study between seismic acceleration parameters and damage indices of structure. *Engineering Structures*, **23**, 698–704, 2001.

[24] A. Masi, M. Vona, M. Mucciarelli, Selection of Natural and Synthetic Accelerograms for Seismic Vulnerability Studies on Reinforced Concrete Frames. *J. Struct. Eng.*, **137**, 367-378, 2011.

[25] PEER (Pacific Earthquake Engineering Research Centre), Strong Motion Database, 2003. http://peer.berkeley.edu/smcat/

[26] European Strong-Motion Database,2003,http://www.isesd.hi.is/ESD Local/frameset.htm

[27] EN1992-1-1, *Design of Concrete Structures, Part 1-1: General rules and rules for buildings*. European Committee for Standardization, Brussels, 2005.

[28] EN1998-1, *Design of structures for earthquake resistance – Part 1: General rules, seismic actions and rules for buildings*. European Committee for Standardization, Brussels, 2005.

[29] T.O.L. - Engineering Software House, *RAF Version 4.4: Structural Analysis and Design Software*. Iraklion, Crete, Greece, 2014.

[30] A. Otani, Inelastic Analysis of RC frame structures. *Journal of the Structural Division, ASCE*, **100(7)**, 1433–1449, 1974.

[31] A.J. Carr, *Ruaumoko – A program for inelastic time-history analysis: Program manual. Department of Civil Engineering*. University of Canterbury, New Zealand, 2006.

[32] Matlab, *Neural networks toolbox user's guide*, 2013.

[33] H.M. Mushgil, H.A. Alani, L.E. George, Comparison between Resilient and Standard Back Propagation Algorithms Efficiency in Pattern Recognition. *International Journal of Scientific & Engineering Research*, **6(3)**, 773-778, 2015.

[34] T. Fawcett, An introduction to ROC analysis. *Pattern Recognition Letters*, **27**, 861–874, 2006.

[35] M.Y. Rafiq, G. Bugmann, D.J. Easterbrook, Neural network design for engineering applications. *Computers and Structures*, **79**, 1541-1552, 2001.

[36] M.T. Hagan, H.B. Demuth, M.H. Beale, *Neural Network Design*. MA: PWS Publishing, Boston, 1996.

[37] M. Gevrey, I. Dimopoulos, S. Lek, Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecological Modelling*, **160**, 249-264, 2003.

[38] G.R. Balls, D. Palmer-Brown, G.E. Sanders, Investigating microclimatic influences on ozone injury in clover (Trifolium subterraneum) using artificial neural networks. *New Phytologist*, **132**, 271-280, 1996.

[39] J.D. Olden, M.K. Joy, R.G. Death, An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecological Modelling*, **178**, 389–397, 2004.

[40] K. Morfidis, K. Kostinakis, Seismic parameters' combinations for the optimum prediction of the damage state of R/C buildings using neural networks. *Advances in Engineering Software*, **106**, 1-16, 2017.