Proceedings of the 9th U.S. National and 10th Canadian Conference on Earthquake Engineering
Compte Rendu de la 9ième Conférence Nationale Américaine et
10ième Conférence Canadienne de Génie Parasismique
July 25-29, 2010, Toronto, Ontario, Canada • Paper No 715

CLASSIFICATION OF EARTHQUAKE DAMAGES IN BUILDINGS USING A GENETIC ALGORITHM PROCEDURE

A. Elenas[1], P. Alvanitopoulos[2] and I. Andreadis[3]

ABSTRACT

In this paper efficient seismic damage classification procedures based on artificial neural networks (ANN) and neuro-fuzzy (NF) systems combined with a genetic algorithm (GA) are reported. The proposed method uses a set of artificial accelerograms in order to examine damages on buildings, expressed by several damages indices. With the use of seismic accelerograms, a set of twenty seismic parameters have been extracted for the seismic excitation description. In this approach the GA algorithm was used to find the optimal feature subset of the seismic parameters that minimizes the computational cost and maximizes the classification performance. The proposed methods have been applied to a six-story reinforced concrete frame structure. The results indicate that the use of the GA was able to classify the damages in the examined building with classification rates up to 100%, while the mean correct classification rate is 91.54%.

Introduction

Seismic accelerograms are records of ground acceleration versus time during earthquakes that cannot be described analytically. However, several seismic parameters have been presented in the literature during the last decades that can be used to express the intensity of a seismic excitation and to simplify its description. Post-seismic field observations and numerical investigations have indicated the interdependency between the seismic parameters and the damage status of buildings after earthquakes. The latter can be expressed by proper damage indices [DiPasquale 1989, Park 1985].

For the present classification system the proposed algorithm consists of two processing stages. First a set of artificial accelerograms have been used to describe the earthquake ground motion. Then, a set of twenty parameters have been extracted from seismic signals to express the damage potential of earthquakes. In addition, several damage indices have been used to estimate the earthquake damages in structures. Previous works prove that there is a correlation between the damage indices and the aforementioned intensity parameters [Elenas 2000, Elenas 2001].

[1]Assoc. Professor, Dept. of Civil Engineering, Democritus University of Thrace, Xanthi, 67100, Greece
[2]PhD candidate, Dept. of Electrical & Computer Engineering, Democritus University of Thrace, Xanthi, GR-67100, Greece
[3]Professor, Dept. of Electrical & Computer Engineering, Democritus University of Thrace, Xanthi, 67100, Greece

At the second stage of processing a genetic algorithm (GA) has been used to reduce the number of the seismic parameters and find the subset that maximizes the classification rates. The GA starts the feature extraction process using an initial population of individuals (combination of seismic parameters) and after a specific number of generations produce an optimal single solution. In order to select the optimum representation of seismic signals different kinds of classifiers have been used. Previous studies [Alvanitopoulos 2008] proposed artificial neural networks and artificial neuro-fuzzy inference systems for the classification of seismic damages. The classification accuracy of these systems has been used to evaluate the fitness value of the individuals.

The last part of the research was the investigation of the classification performance. The classifiers have been trained and simulated using the optimal subset of the intensity parameters. Classification results testify the effectiveness of this method.

The proposed methodology is a simple tool for the evaluation of the post-seismic damage status of buildings in form of damage degree, avoiding complicated and time consuming nonlinear dynamic analyses. The procedure can be used by the public administration for the evaluation of damage scenarios of separate important buildings, building blocks, or even of whole regions. This is important for the adequate post-seismic distribution of financial and other resources in the case of severe earthquakes. Another application of the proposed technique is its implementation on a micro-chip in combination with an accelerograph and a signal transmit unit for the direct and in real time evaluation of the damage status of a building after an earthquake. Thus, the administrative officials would be informed about possible damage extension of important buildings (hospitals, fire stations, bridges, etc.) or even of whole regions, immediately after an earthquake. In the latter case the proposed model after the training phase, is far easier to implement on a micro-chip than the complicated nonlinear dynamic procedure.

Genetic Algorithms

Genetic algorithm is an adaptive search and an optimization model which have been inspired from the principles of natural evolution [Sivanandam 2008]. Genetic algorithms were first introduced in the early 1970s by John Holland. They are able to exploit the information from the acceptable solutions and select the optimal one.

The implementation of a GA starts with the generation of the initial population of the candidate solutions. Usually the selection of the first population is random. GA is an iterative process which modifies the current population by selecting individuals to be parents and uses them to produce the children for the next generation. A more detailed description of the GA procedure is given below:

Generation of the initial population: Assume a population $p = \{x_1, \dots, x_N\}$ with N individuals. Each member of the population is a string of length L. Each string x_i is referred to as a chromosome.

Fitness function: Every chromosome is evaluated and assigned a fitness function $f(x_i)$. The lower the fitness is the better is the chance for this chromosome or its descendants to survive.

Selection process: Here, chromosomes equal to the number N (number of individuals in initial population) are selected for reproduction. All chromosomes have the opportunity, to be picked for reproduction, proportional to theirs fitness value. Selection is done "with replacement" meaning that the same chromosome can be selected more than once to become a parent.

Crossover: The next step of the GA is the crossover operation. All the parent couples, which are selected from previous step, are able to produce two new chromosomes according to a pre-specified, usually high, probability (p_c). The crossover action chooses a single point and the two parents exchange genetic material by cutting and swapping their string after the single point. If no crossover takes place then the two new chromosomes (offsprings) are identical copies of their parents. Fig. 1(a) shows a crossover operation with a single point on the fifth bit.
Mutation: The probability of mutation is usually small. For the binary chromosomes produced by crossover step mutation just change the bit 1 to 0 and 0 to 1 (Fig. 1(b)). The mutation procedure prevents the algorithm to be trapped in a local minimum.

GA moves from generation to generation and terminates until a stopping criterion is met. The maximum number of generations or a threshold in fitness value, may be used in order to find the optimize solution.

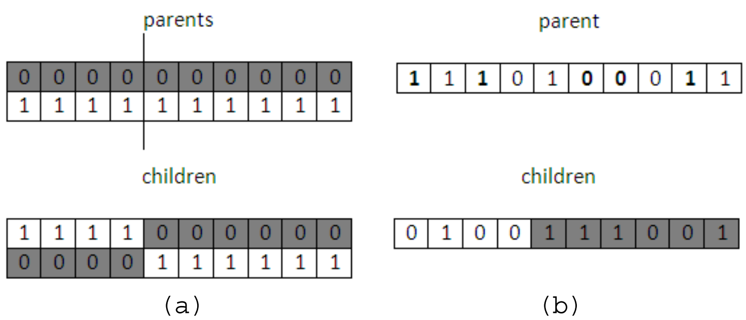

Figure 1. A crossover operator (a) and a mutation operator (b).

Proposed Method

Seismic Parameters

Accelerogram records describe the seismic excitation explicitly. However, due to their random nature it is very difficult to exploit their similarities. Therefore, a set of twenty seismic parameters have been used to represent the seismic excitation. In detail the following parameters are considered: the peak ground acceleration (PGA) a_{max}, the peak ground velocity (PGV) v_{max}, the term a_{max}/v_{max}, the Arias intensity (AI), the root mean square acceleration (RMS_a), the strong motion duration as defined by Trifunac and Brady (SMD_{TB}), the seismic power $P_{0.90}$, the spectral intensities after Housner (SI_H), after Kappos (SI_K) and after Martinez (SI_M), the effective peak ground acceleration (EPA) and its maximum value (EPA_{max}), the spectral total seismic energy input E_{inp}, the cumulative absolute velocity (CAV), the seismic damage potential after Araya and Saragoni (DP_{AS}), the central period (CP), the spectral acceleration (SA), the spectral velocity (SV), the spectral displacement (SD) and the seismic intensity as defined by Fajfar, Vidic and Fischinger (I_{FVF}). The used spectral values are calculated for the period 0.88 s. It corresponds to the eigenperiod of the examined reinforced concrete frame. The definitions of the particular parameters are indicated in the literature [Andreadis 2007] and are not here repeated. The calculation of the parameters took place via a computer-aided processing of the accelerograms. In our method the GA attempts to find the optimum representation of the seismic accelerograms (minimum number of required seismic intensity parameters) in order to produce the best results.

Genetic Algorithm Implementation

A GA was used to find the optimal feature set in order to produce the best classification accuracy of the proposed classifiers. First several subsets of seismic parameters have been examined. The classifiers have been trained according to these features. The fitness function of these subsets has been evaluated and the optimal set of seismic parameters has been extracted.

Let L=20 (twenty seismic parameters) as the number of feature descriptors. Assume a population of N individuals. In this research a population size of N=20 individuals has been used. A chromosome of L genes is an individual which represents the subset of seismic parameters. In the initial population $p=\{x_1,...,x_N\}$ the first sample x_1 has all the genes equal to 1. The genes were allowed to take either a 0 or 1 as a value. A value of 1 implied that the corresponding parameter would be included in the feature subset. The parameter would be excluded from the feature subset if its gene value was set to 0. Here, the negative classification accuracy of the classifiers is equal to the fitness function of the subset. We use the negative classification accuracy because the algorithm selects as elitist individuals the subsets which have the lowest fitness value. The GA was allowed to run for a maximum of 100 generations.

The GA creates three types of children to the next generation. The first type is the elite children. These are the best individuals in the previous generation which are guaranteed to survive to the next generation. In this approach the elite children parameter was set to 2. Besides elite children, the algorithm creates the crossover and mutation children. The crossover operation recombines the best genes from different individuals in order to produce a superior child. After the crossover the mutation step was used to search through a larger search area in order to find the best solution. In each generation 80% of the individuals in the population excluding the elites were created through the crossover operation and the remaining 20% were generated through mutation. Using these parameters in our algorithm it is clear that for a population equal to 20 there are 2 elite children from the previous generation, 14 crossover and 4 mutation children.

The Neural Network Classifier

In this study an artificial neural network (ANN) was used for the classification of seismic damages. This network consists of one input layer, a hidden layer of 17 neurons and one output layer. The proposed classifier is a supervised feed-forward ANN with hyperbolic tangent sigmoid activation function. The first layer presents the inputs of the network. The number of the inputs to the first layer is not fixed. All the individuals from the GA are passed through the ANN in order to estimate the classification accuracy of them and their fitness function. Each time the inputs are equal to the number of genes which their value is set to 1. The number of output units is fixed to four, since four are the categories of possible damages. During the training of the ANN a set of representative vector samples have been used. Then the ANN was simulated using the entire set of seismic signals in order to evaluate the classification performance. Each time a seismic signal was represented in ANN with the set of seismic parameters according to the GA individuals. During the supervised training process whenever a training vector appears, the output of the neuron, which represent the class, where the input belongs, is set to 1 and all the rest outputs are set to 0. The training algorithm for the network is the Levenberg/Marquardt Back Propagation.

The training steps in batch mode using the Levenberg/Marquardt Back Propagation algorithm are as follows:

1. First present all inputs to the network and compute the corresponding outputs and errors.
2. Compute the Jacobian matrix, **J**, where x is weight and biases of the network.
3. Solve the LM weight update equation.
   $x_{k+1} = x_k - [\mathbf{J}^T\mathbf{J} + \mu\mathbf{I}]^{-1}\mathbf{J}^T\mathbf{e}$ (1)
   where, x_k is calculated weight, x_{k+1} is weight value in the next step, μ is training parameter, $\mathbf{J}^T\mathbf{J}$ is Hessian matrix, **I** is the identity matrix and $\mathbf{J}^T\mathbf{e}$ is the gradient (where **e** is the network error vector).
4. Recompute the error using $x + \Delta x$. If this new error is smaller than that computed in step 1, then reduce the training parameter μ by μ- let $x = x + \Delta x$, and go back to step 3 μ+ and μ- are predefined values set by the user.

The Neuro-Fuzzy Classifier

The last intelligent system for the classification of damages in structures is a neuro-fuzzy technique. This system combines the fuzzy set theory and the ANNs. The neuro-fuzzy system has six layers. The first layer is the input layer where the inputs correspond to the subset of seismic parameters that represent the individual of the GA. The second layer implements the fuzzification process. Each neuron in the second layer represents a membership function. The key point of this step is the fuzzyfication of the inputs using four membership functions for each intensity parameter. Membership functions equations are presented below.

$$m_{f1}(x, 0.1416, 0) = e^{\frac{-(x-0)^2}{2*(0.1416)^2}} \quad (2)$$

$$m_{f2}(x, 0.1416, 0.3359) = e^{\frac{-(x-0.335)^2}{2*(0.1416)^2}} \quad (3)$$

$$m_{f3}(x, 0.1416, 0.667) = e^{\frac{-(x-0.667)^2}{2*(0.1416)^2}} \quad (4)$$

$$m_{f4}(x, 0.1416, 1) = e^{\frac{-(x-1)^2}{2*(0.1416)^2}} \quad (5)$$

The next layer of the classifier consists of fuzzy rules. Each neuron corresponds to a fuzzy rule. The number of rules is related with the volume of the training samples (training accelerograms). The last three layers compose an embedded ANN. The inputs of the embedded ANN are the firing strength of rules according to the input seismic sample and the output declares the winning class.

Application

The design of the examined reinforced concrete frame structure, shown in Fig. 2, is in agreement with the rules of the recent Eurocodes for structural concrete and aseismic structures, EC2 and EC8, respectively. The application is a sixth floor structure with a total height 19 m. The ground floor has a 4 m height and all subsequent floors 3 m. The cross-sections of the beams are considered as T-beams with 30 cm width, 20 cm slab thickness, 60 cm total beam height and 1.45 m effective slab width. The distances between each frame of the structure is equal to 6 m. The eigenperiod of the frame was 0.88 s. In addition to the seismic loads, the following loads have been taken into consideration: self-weight, seismic loads, snow, wind and live loads. The structure has been considered as an "importance class II, ductility class M"-structure, with a subsoil category B, according to the design rules of the EC8 Eurocode.

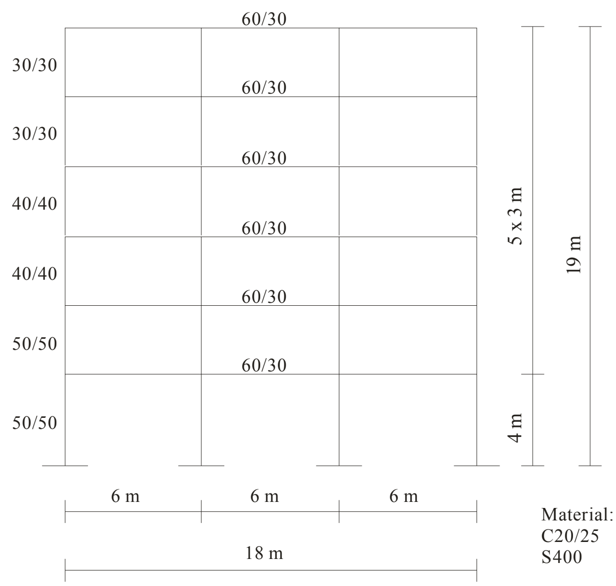

Figure 2. Reinforced concrete frame structure.

After the design procedure of the reinforced concrete frame structure, a nonlinear dynamic analysis evaluates the structural seismic response, using the computer program IDARC [Valles 1996]. A three-parameter Park model specifies the hysteretic behavior of beams and columns at each member end. This hysteretic model incorporates stiffness degradation, strength deterioration, slip-lock and a tri-linear monotonic envelope. Experimental results of cyclic force-deformation characteristics of typical components of the studied structure, specifies the parameter values of the above degrading model. This study uses the nominal parameter for stiffness degradation and the focus is on the overall damage indices. This is due to the fact, that these parameters summarizes all the existing damages on columns and beams in a single value, which can be easily interrelated to single value seismic parameters. In this study the damage index of Park/Ang [Park 1985], the damage index of DiPasquale/Çakmak [DiPasquale 1989] and the maximum inter-storey drift ratio (MISDR) have been used. In addition, the MISDR has been used to express the architectural damages. The damage indices have been evaluated for 450 created synthetic accelerograms.

Results

After the nonlinear dynamic analysis of the structure, for the entire set of artificial accelerograms, the global damage indices of Park/Ang (DI_{PA}), of DiPasquale/Çakmak (DI_{DC}) and the MISDR, have been calculated. The latter damage index have been used to express structural ($MISDR_{struc}$) and architectural damages ($MISDR_{arch}$). According to the DIs, the damages were classified in four clusters. These correspond to the damage degrees low-medium-large-total. In this context, the damage degrees denote undamaged or minor damage-repairable damage-irreparable damage-partial or total collapse of the building, respectively. Table 1 classifies the structural and architectural damage degrees according to the used damage indices [Gunturi 1992].

Table 1. Damage degree classification.

Global Damage Indices	**Damage Degree**			
	Low	**Medium**	**Large**	**Total**
DI_{PA} or DI_{DC} [-]	≤ 0.3	0.3 < DI ≤ 0.6	0.6 < DI ≤ 0.8	DI > 0.8
$MISDR_{struc}$ [%]	≤ 0.5	0.5< MISDR ≤ 1.5	1.5< MISDR ≤ 2.5	> 2.5
$MISDR_{arch}$ [%]	≤ 0.5	0.5< MISDR ≤ 1.2	1.5< MISDR ≤ 2.7	> 1.7

In the present investigation a total set of 450 artificial accelerograms have been used. The representation of the artificial accelerograms has been examined using different subsets (individuals) of the twenty intensity parameters. Each individual is a 1x20 matrix. Due to the bit string type of individuals the total number of the possible candidate solution is 2^{20}. In this research a GA with a population of 20 individuals was employed and run for a maximum number of 100 generations. This means that the GA searches for the optimal feature selection and tests up to 2000 possible solutions. Using only the selection process in the GA without the crossover and mutation step it will create a negative effect on the convergence. On the other hand, using mutation alone is similar to a random search. The GA has been used one time for each one of the DIs. Two types of classifiers have been used in order to estimate the fitness function of the GA.

Table 2 shows the classification rates for the architectural and the structural damage degree, based on the MISDR and on the global damage indices of Park/Ang (DI_{PA}) and DiPasquale/Çakmak (DI_{DC}). Thus, the used procedure, the number of training samples (NOTS), 450 or 300 alternatively, the number of non-trained samples (NONTS), these samples were not used during the training process, the number of essential parameters (NOEP), the number of well recognized samples (NOWRS) and the percentage of well recognized samples (POWRS) are provided. It must be noticed that the number of samples used for the classification process, was 450 (NOTS+NONTS) in all the cases. Here, it is recognized that the use of Genetic Algorithms improved the percentage of well recognized samples (POWRS) in most cases. In the cases where the use of GA provided a lesser POWRS in comparison with the pure ANN procedure, the reduction was lesser than the percentage reduction of the number of essential parameters (NOEP). In addition, the use of genetic Algorithms reduces NOEP in all the cases. Thus, it can be concluded that ANN and NF procedures, combined with Genetic Algorithms provide high damage classification rates (up to 100%). It is obvious that the proposed method can be applied to any other structure type, to alternative damage indices and/or different seismic parameters, on the condition that the proposed methodology steps are applied.

Table 2. Classification results using intelligent techniques.

DI	Case	Procedure	NOTS	NONTS	NOEP	NOWRS	POWRS
$MISDR_{arch}$	1	ANN	450	0	20	442	98.2
	2	GA-ANN	450	0	13	450	100
	3	ANN	300	150	20	422	93.7
	4	GA-ANN	300	150	10	449	99.7
	5	NF	450	0	20	441	98
	6	GA-NF	450	0	12	450	100
	7	NF	300	150	20	416	92.4
	8	GA-NF	300	150	12	422	93.7
$MISDR_{struc}$	9	ANN	450	0	20	406	90.2
	10	GA-ANN	450	0	13	417	92.6
	11	ANN	300	150	20	383	85.1
	12	GA-ANN	300	150	14	390	86.6
	13	NF	450	0	20	406	90.2
	14	GA-NF	450	0	13	415	90.2
	15	NF	300	150	20	386	85.7
	16	GA-NF	300	150	11	389	86.4
DI_{PA}	17	ANN	450	0	20	446	99.1
	18	GA-ANN	450	0	13	408	90.6
	19	ANN	300	150	20	385	85.5
	20	GA-ANN	300	150	12	392	87.1
	21	NF	450	0	20	448	99.5
	22	GA-NF	450	0	13	410	91.1
	23	NF	300	150	20	404	89.7
	24	GA-NF	300	150	14	403	89.5
DI_{DC}	25	ANN	450	0	20	412	91.5
	26	GA-ANN	450	0	13	410	91.1
	27	ANN	300	150	20	390	86.6
	28	GA-ANN	300	150	15	396	88.0
	29	NF	450	0	20	422	93.7
	30	GA-NF	450	0	13	404	89.7
	31	NF	300	150	20	386	85.7
	32	GA-NF	300	150	11	397	88.2

Table 3 shows explicitly the essential parameters, in all the cases where the use of Genetic Algorithms reduces the number of the initial used twenty seismic parameters. The initial 20 seismic parameters, were reduced in a range between 15 (case 28) and 10 (case 4). That means a reduction between 25% and 50%. In addition, it is recognized that the most important parameters are the spectral displacement (SD) and the seismic damage potential after Araya and Saragoni (DP_{AS}). The first mentioned parameter appears in 13, while the second one appears in 12, of 16 maximum possible, cases.

Table 3. Essential intensity parameters.

Seismic Parameter	**Case**																
	2	**4**	**6**	**8**	**10**	**12**	**14**	**16**	**18**	**20**	**22**	**24**	**26**	**28**	**30**	**32**	**Sum**
PGA		X		X	X	X		X		X		X		X		X	9
Arias			X	X	X	X			X		X	X	X	X	X		10
PGV	X	X		X	X	X	X			X		X	X		X	X	11
PGA/PGV	X	X	X			X		X		X	X				X	X	9
CAV	X			X	X		X	X	X	X		X	X	X	X		11
RMS			X	X		X	X	X	X		X			X	X	X	10
SMD_{TB}	X		X	X	X	X				X	X	X	X	X		X	11
$P_{0.90}$	X			X		X	X	X	X	X	X	X	X		X		11
SI_H	X	X	X	X			X		X			X	X	X		X	10
SD	X		X	X	X	X	X	X	X	X	X		X	X	X		13
SV	X				X		X				X	X		X		X	7
SA	X	X			X	X		X	X	X		X	X	X	X		11
E_{inp}		X	X		X	X		X	X	X	X			X	X		10
SI_K		X	X	X			X		X		X	X				X	8
EPA				X	X	X			X			X	X	X	X	X	9
EPA_{max}	X		X		X	X	X	X		X	X		X	X	X		11
I_{FVF}	X		X				X		X		X	X	X	X			8
DP_{AS}	X	X	X		X	X	X	X		X		X		X	X	X	12
CP		X			X	X	X	X	X	X	X		X	X	X		11
SI_M	X	X	X	X			X		X		X	X	X			X	10
Sum	13	10	12	12	13	14	13	11	13	12	13	14	13	15	13	11	

Conclusions

Alternative intelligent techniques for the classification of seismic damage degrees based on seismic parameters are presented. The focus is on the use of Genetic Algorithms, to improve ANN and Neuro-Fuzzy procedures. The proposed procedures have been applied to a six-story reinforced concrete frame structure designed in accordance to the rules of the EC2 and EC8 Eurocodes for reinforced concrete and antiseismic structures, respectively. A set of 20 intensity parameters provided by artificial accelerograms have been applied for the training and for the classification phase of the models. In addition, the global damage indices of Park/Ang, of DiPasquale/Çakmak and the MISDR, have been used. The latter DI has been used to express structural and architectural damages. Four degrees have been used (low-medium-large-total) for the classification of the seismic damage. The numerical results show a significant reduction of the essential seismic parameters (from 25% up to 50%) using Genetic Algorithms in combination with ANN and Neuro-Fuzzy procedures. The mean correct classification result of all the used intelligent techniques is 91.54%. Thus, these results lead to the conclusion that the proposed techniques are fast and confident numerical tools for the estimation of the structural and architectural post-seismic damage status of buildings.

References

Alvanitopoulos, P., I. Andreadis, A. Elenas, 2008. A New Algorithm for the Classification of Earthquake damages in Structures, *CD ROM Proceedings of the IASTED Int. Conf. on Signal Processing, Pattern Recognition and Applications*, Paper No. 599-062, Innsbruck, Austria.

Andreadis, I. and Y. Tsiftzis, A. Elenas, 2007. Intelligent Seismic Acceleration Signal Processing for Structural Damage Classification, *IEEE Transactions on Instrumentation and Measurement* 56 (5), 1555-1564.

DiPasquale, E. and A.S. Çakmak, 1989. On the relation between local and global damage indices, *Technical Report NCEER-89-0034*, State University of New York, Buffalo.

Elenas, A., 2000. Correlation Between Seismic Acceleration Parameters and Overall Structural Damage Indices of Buildings, *Soil Dynamics and Earthquake Engineering* 20 (1), 93-100.

Elenas, A. and K. Meskouris, 2001. Correlation study Between Seismic Acceleration Parameters and Damage Indices of Structures, *Engineering Structures* 23 (6), 698-704.

Gunturi, S.K.V. and H.C. Shah, 1992. Building specific damage estimation. *Proceedings of the Tenth World Conference on Earthquake Engineering*, Balkema, Rotterdam, Netherland, 6001–6006.

Park,Y.J. and A.H.S. Ang, 1985. Mechanistic seismic damage model for reinforced concrete, *Journal of Structural Engineering* 111 (4), 722-739.

Sivanandam, S.N. and S.N. Deepa, 2008. *Introduction to Genetic Algorithms*, Springer Verlag, Berlin, Germany.

Tsiftzis, I., I. Andreadis, A. Elenas, 2006. A Fuzzy System for Seismic Signal classification, *IEE Proc. Vision, Image & Signal Processing* 153 (2), 109-114.

Valles, R., A.M. Reinhorn, S.K. Kunnath, C. Li, Madan A., 1996. IDARC 2D Version 4.0: A program for the Inelastic Damage Analysis of Buildings. *Technical Report NCEER-96-0010*, State University of New York, Buffalo.