STRUCTURAL DAMAGE PREDICTION UNDER SEISMIC SEQUENCE USING NEURAL NETWORKS

Petros C. Lazaridis[1]**, Ioannis E. Kavvadias**[1]**, Konstantinos Demertzis** [1]**, Lazaros Iliadis**[1]**, Antonios Papaleonidas**[1]**, Lazaros K. Vasiliadis**[1]**, Anaxagoras Elenas**[1]

[1]Department of Civil Engineering, Democritus University of Thrace, Campus of Kimmeria, 67100 Xanthi, Greece
e-mail: {petrlaza1@civil, ikavvadi@civil, kdemertz@fmenr, liliadis@civil, papaleon@civil, lvasilia@civil, elenas@civil}.duth.gr

Abstract. *Advanced machine learning algorithms, such as neural networks, have the potential to be successfully applied to many areas of system modelling. Several studies have been already conducted on forecasting structural damage due to individual earthquakes, ignoring the influence of seismic sequences, using neural networks. In the present study, an ensemble neural network approach is applied to predict the final structural damage of an 8-storey reinforced concrete frame under real and artificial ground motion sequences. Successive earthquakes consisted of two seismic events are utilised. We considered 16 well-known ground motion intensity measures and the structural damage that occurred by the first earthquake as the features of the machine-learning problem, while the final structural damage was the target. After the first seismic events and after the seismic sequences, both actual values of damage indices are calculated through nonlinear time history analysis. The machine-learning model is trained using the dataset generated from artificial sequences. Finally, the predictive capacity of the fitted neural network is accessed using the natural seismic sequences as a test set.*

Keywords: Seismic Sequence, Neural Networks, Repeated Earthquakes, Structural Damage Prediction, Artificial Intelligence, Intensity Measures

1 INTRODUCTION

In recent years advanced Machine Learning (ML) algorithms, such as Artificial Neural Networks (ANN), has been successfully applied to many areas of system modelling. Their success is based on the thorough processing of data that captures the behaviour of a system. By detecting patterns in the collected data, valuable information can be extracted and predictions can be made that automate the decision-making process. That fact arrives ML as an advanced tool in modern engineering modelling.

The object of the present research is the application of neural networks to predict the structural damage under seismic sequences. The recent disaster held on March 2021 in Tyrnavos-Elassona region, Thessaly of Greece due to a pair of compatible magnitude (Mw=6.3, Mw=6.1) [1, 2] shallow earthquakes with more than 1800 damaged or non-serviceable buildings, indicate the critical role of seismic sequences in the final damage state of buildings again. In the field of seismic sequences have been conducted many kinds of research to access the contribution of repeated ground motions on the post-seismic performance state of a building so far, which concluded that aftershocks increase the ductility demands. Amadio et al. [3] studied the influence of repeated shocks on the response of nonlinear Single Degree of Freedom (SDOF) systems using different hysteretic models. Hatzigeorgiou and Beskos [4] conducted an exhausting parametric study on SDOF systems and proposed an empirical relation to calculate inelastic displacement ratio under repeated earthquakes. Hatzigeorgiou and Liolios [5] examined the nonlinear behaviour of Reinforced Concrete (RC) frames subjected to multiple shocks considering a set of eight frames that varied both at height regularity and dimensioning practice. Hatzivassiliou and Hatzigeorgiou [6] studied the accumulation of damage and ductility demands due to seismic sequence on three dimensional RC structures. Hosseinpour and Abdelnaby [7] studied the impact of different aspects, as earthquake direction, aftershock polarity and the influence of the vertical component, on the nonlinear response of RC frames under successive earthquakes. Also, more recently, Kavvadias et al. [8] and Zhou et al. [9] investigated the correlation between aftershock related Intensity Measures (IMs) and final structural damage indices.

Previous several studies have been conducted on forecasting of structural damage due to individual earthquakes, ignoring the influence of seismic sequences. De Latour and Omenzetter [10] investigated the efficiency of ANNs on the prediction of seismic damage on numerous RC frames, Alvanitopoulos et al. [11] also examined regular RC structures and, they incorporate fuzzy layers in ANN configuration (architecture). Later, Morfidis and Kostinakis [12] used feature selection methods in a dataset of 3 dimensional RC buildings to found the more damage correlated set of seismic IMs. More recently, the same authors [13] examined the effectiveness of ANNs on the prediction of damage of non-regular at height structures. Applications of Recurrent Neural Networks (RNNs) on Earthquake Engineering made by González et al. [14] and Mangalathu and Burton [15], also Zhang et al. [16] developed a LSTM (Long-Short Term Memory) network to predict structural response. Convolutional Neural Networks (CNNs) applied by Li et al. [17] and Oh et al. [18] for the same purpose.

In the present study, ANNs are applied, considering real and artificial sequences of double seismic shocks, to predict the amount of final structural damage on an already damaged 8-storey RC frame after a first shock. We use 16 well-known ground motion IMs and the established damage from the first earthquake as the features of the ML problem, while the final structural damage (after sequence) employed as the target. Both damage indices after the first seismic event and after the sequence are calculated through Nonlinear Time History Analysis (NTHA).

2 SELECTION OF PRIMITIVE DATA

In this section, the primitive data of the problem is described. Specifically, the selection of the ground motion records and the examined RC frame is presented.

2.1 Ground motion records

Both artificial and natural seismic sequences are considered to take into account, the phenomenon of multiple shocks. After the random combination from a suite of 318 acceleration records from several earthquakes, the set of artificial sequences determined. More specifically, every record of the above 318 records suite is combined randomly three times with another record from the same dataset. As a result, 954 pairs of first and second shock are constructed, which are utilised as a training set for the specific ML problem. As a test set, we employed 119 natural pairs of sequential shocks, which occurred from 1972 to 2020. The training and test sets are completely independent with no common record. Both sequential and individual records are selected and downloaded from ESM [19] and PEER NGA West [20] databases. The criteria according to the selection of natural sequences conducted are the time gap between the occurrences of successive shocks which are considered smaller than three months and that the sequential events are recorded by the same station. As a result of this process collected 119 pairs of successive ground motion acceleration records from 41 real seismic sequences, recorded by 65 stations. The occurrences time and Peak Ground Acceleration (PGA) data of the natural sequences listed in Table 1. Every pair concatenated in the same array to compose the sequential ground acceleration time history. In addition, an intermediate zero ceasing time gap of 20 seconds added to eliminate the overlap between the building oscillations due to successive earthquakes. This process is implemented in Python [21], and the new files containing the seismic sequence acceleration signals are exported.

2.2 Reinforced concrete structure

Existing buildings designed and constructed without earthquake provisions comprise the majority of structures both in Greece and worldwide, which causes particular concern about their response to a potential earthquake. As an example, in this study we examined an 8-storey planar regular RC frame (Figure 1) designed only for gravity loads by Hatzigeorgiou and Liolios [5], with mean compressive strength of concrete equal to 28 MPa and steel grade S500s. The finite element simulation of the frame held in IDARC 2D [22], using spread plasticity concept and three-parameter Park hysteretic model [23]. Every floor considered to have only one horizontal degree of freedom to into account the huge plane stiffness of RC slab as a rigid diaphragm. Sparsely placed stirrups with poor anchor details are assumed in order to be in accordance with obsolete design codes. Thus, a nonlinear deformation-stress model for concrete without confinement is adopted. As a result, the corresponding concrete curve defined by the initial modulus of elasticity ($E_0 = 31.42\,GPa$), the strain at the maximum stress ($\varepsilon_{c0} = 2\%_0$), the ultimate strain in compression ($\varepsilon_{cu} = 3.5\%_0$), the stress at tension cracking $\sigma_t = 0.0022\,GPa$, and the slope of the post-peak falling branch ($E_{fb} = -6.2\,GPa$). Also, the bilinear curve with hardening employed for steel with yield and ultimate strengths equal to 550 MPa and 660 MPa respectively and the corresponding strains equal to 2.75$\%_0$ and 45$\%_0$, according to Eurocode-2 [24] provisions. The initial elastic fundamental period of the structure is equal to 1.27 seconds.

2.3 Technicalities

The generation of IDARC 2D input files and the extraction of responses from output files implemented through GNU Octave [25, 26] code. Using the same high-level programming language we automated the multiple analysis process and computed the IMs, which are described in the next section, for the overall dataset of first and second seismic shocks.

Region	1st Sh. Date	2nd Sh. Date	Station Code / Name	Stream	PGA 1st $\frac{cm}{s^2}$	PGA 2nd $\frac{cm}{s^2}$
Ancona	1972-06-14	1972-06-21	ANP	N-S	216.17	402.88
				E-W	206.15	208.0
Friuli	1976-09-11	1976-09-15	BUI	N-S	229.04	108.53
				E-W	106.41	91.37
			GMN	N-S	322.02	318.58
				E-W	293.6	632.0
Northwestern Balkan Peninsula	1979-04-15	1979-04-15	PETO	E-W	298.18	87.4
		1979-05-24	BAR	N-S	364.69	197.33
				E-W	353.28	262.19
			HRZ	N-S	211.32	65.68
				E-W	249.49	75.05
			ULO	N-S	277.13	32.9
				E-W	232.34	30.13
Imperial Valley	1979-10-15	1979-10-15	Holtville Post Office	315	217.24	250.07
Mammoth Lakes	1980-05-25	1980-05-25	Convict Creek	90	411.49	364.47
Irpinia	1980-11-23	1980-11-24	BGI	N-S	126.96	30.71
				E-W	185.88	33.32
			STR	N-S	220.44	18.37
				E-W	314.41	31.41
Greece	1981-02-24	1981-02-25	KORA	Trans	290.82	119.32
				Logn	235.91	118.72
Coalinga	1983-07-22	1983-07-25	Elm (Old CHP)	90	509.27	664.53
				0	334.97	472.1
Southern Greece	1986-09-13	1986-09-15	KAL1	Trans	264.6	138.13
				Logn	227.84	232.78
			KALA	Trans	290.91	149.6
				Logn	212.21	327.68
Spitak	1988-01-07	1988-01-07	GUK	N-S	178.23	142.0
				E-W	178.81	97.63
Western Caucasus	1989-01-08	1989-01-08	NAB	N-S	188.92	192.91
				E-W	202.05	213.12
Greece	1990-05-17	1995-06-15	AIGA	Trans	198.1	510.79
				Logn	116.29	488.87
Western Caucasus	1991-05-03	1991-05-03	SAMB	N-S	347.84	204.3
				E-W	494.46	120.01
Turkey	1992-03-13	1992-03-15	AI 178 ERC MET	N-S	403.55	32.15
				E-W	478.36	39.11
Southern Greece	1993-03-26	1993-03-26	PYR1	Trans	218.77	117.3
				Logn	107.23	98.62
Northridge	1994-01-17	1994-01-17	Moorpark - Fire Sta	90	189.38	136.86
				180	286.27	180.75
			Pacoima Kagel Canyon	360	424.6	52.45
		1994-03-20	Rinaldi Receiving Sta	228	857.45	519.21
			Sepulveda Hospital	270	738.22	100.1
			Sylmar - Olive Med	90	593.39	178.32
Umbria Marche	1997-09-26	1997-09-26	CLF	N-S	271.52	193.84
				E-W	251.78	223.46
			NCR	N-S	387.53	492.33
				E-W	251.0	415.1
Southern Greece	1997-10-13	1997-11-18	KRN1	Trans	117.43	70.44
				Logn	116.52	90.77
Bovec	1998-04-12	1998-08-31	FAGG	N-S	24.08	23.0
				E-W	23.32	26.16
Azores Islands	1998-07-09	1998-07-11	HOR	N-S	397.33	80.57
				E-W	362.48	90.37
Izmit	1999-08-17	1999-11-12	ARC	N-S	206.54	7.39

Region	1st Sh. Date	2nd Sh. Date	Station Code / Name	Stream	PGA 1st $\frac{cm}{s^2}$	PGA 2nd $\frac{cm}{s^2}$
			ATK	E-W	130.42	7.55
				N-S	100.31	16.25
			DHM	E-W	164.3	16.01
				N-S	88.28	17.03
			FAT	E-W	82.92	17.24
				N-S	178.45	34.32
			KMP	E-W	158.58	23.77
				N-S	100.59	14.43
			ZYT	E-W	125.13	17.21
				N-S	116.96	21.41
Greece	1999-09-07	1999-09-07	ATHENS-SEPOLIA B	E-W	107.05	28.88
				Trans	317.98	58.14
				Logn	335.1	69.79
Chi-Chi Taiwan	1999-09-20	1999-09-20	TCU071	N-S	638.91	375.02
				E-W	518.53	189.69
			TCU129	N-S	390.42	612.17
				E-W	929.28	985.77
			TCU078	N-S	301.13	379.95
				E-W	438.81	261.21
			TCU079	N-S	416.3	614.35
				E-W	580.9	761.01
Duzce	1999-11-12	1999-11-12	AI 010 BOL	E-W	804.95	59.27
Turkey	2003-05-01	2003-05-01	AI 049 BNG	N-S	509.25	144.45
				E-W	286.03	67.56
L Aquila	2009-04-06	2009-04-07	AQK	N-S	346.9	80.4
				E-W	323.84	88.45
			AQV	N-S	535.39	144.13
				E-W	644.47	127.22
			AVZ	N-S	67.71	21.44
		2009-04-09	AQA	N-S	434.0	56.85
Darfield New Zealand	2010-09-03	2011-02-21	Botanical Gardens	S01W	186.4	443.68
				N89W	152.43	542.17
			Cashmere High School	S80E	246.32	342.41
			Cathedral College	N26W	190.79	376.79
				N64E	228.91	469.24
			Christchurch Hospital	N01W	205.41	339.41
				S89W	149.4	356.74
Emilia	2012-05-20	2012-05-29	MRN	N-S	258.89	288.73
				E-W	257.32	218.65
Northern Italy	2012-06-03	2012-06-12	T0827	N-S	481.11	574.06
				E-W	258.35	230.4
Central Italy	2016-08-24	2016-08-24	AQK	E-W	49.54	10.59
		2016-08-26	AMT	N-S	368.51	329.9
				E-W	851.09	318.84
	2016-10-26	2016-10-26	CMI	N-S	334.88	302.66
				E-W	707.04	638.53
			CNE	E-W	545.76	527.19
		2016-10-30	ACC	N-S	49.29	384.83
				E-W	88.35	426.01
			CIT	N-S	51.02	209.24
				E-W	90.49	319.58
			CLO	N-S	189.45	571.62
				E-W	180.02	418.76
			CNE	N-S	373.26	288.38
			MMO	N-S	164.82	185.15
				E-W	166.99	185.39
			NOR	E-W	210.87	305.85
	2016-10-30	2016-10-31	T1213	N-S	850.26	182.16
				E-W	779.54	207.93
	2017-01-18	2017-01-18	PCB	N-S	575.2	550.63
				E-W	400.57	380.9
Near The Coast Of Western Turkey	2019-08-08	2020-10-30	GMLD	N-S	441.94	882.26
				E-W	660.9	748.75

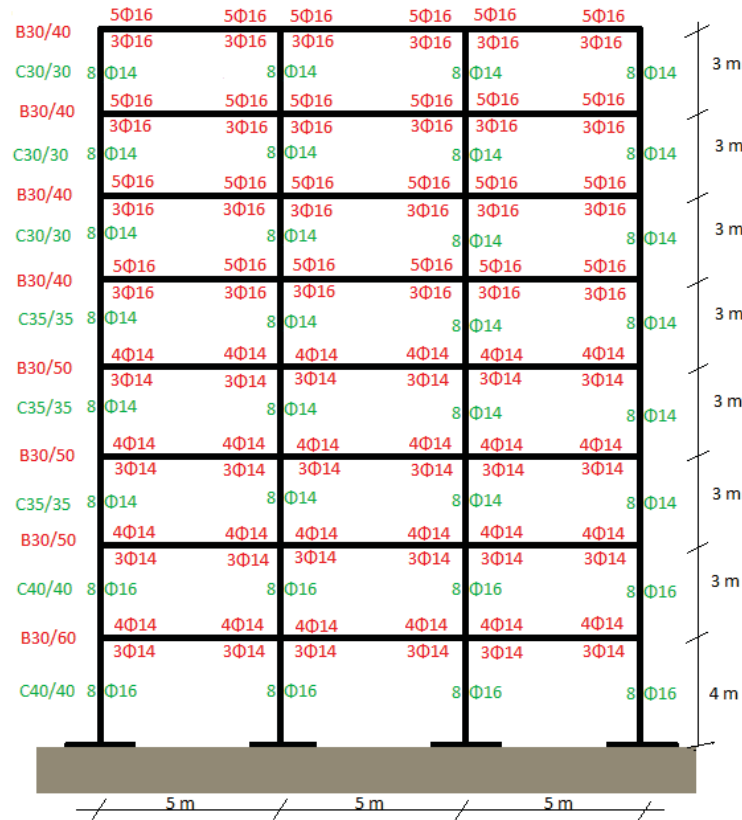Table 1: Seismic Metadata for Natural Sequences (Test Set).

Figure 1: Reinforced Concrete Frame designed by Hatzigeorgiou and Liolios [5].

3 FEATURES AND TARGET OF THE ML PROBLEM

This work examined a reinforced concrete planar frame subjected to both artificial and natural sequential acceleration records. The ground motion intensity is evaluated, through scalar measures which are proposed in the literature. The structural damage is accessed in terms of well-known overall damage indices. These features are presented and explained below.

3.1 Ground motion IMs

As previous studies in the prediction of seismic damage amount, we use as features well-known ground motion IMs. The simplest of these are the maximum absolute values of ground accelerations ($a_g(t)$), velocities ($v_g(t)$), and displacements ($d_g(t)$) signals, which referenced as PGA, PGV, and PGD [27], respectively. Widely used IMs are the Intensity Arias (I_A) [28] and the Cumulative Absolute Velocity (CAV) [29]. Based on the evolution of Intensity Arias over time the Husid Diagram (H_d) [30] is constructed from which duration parameters calculated.

An inherent feature of signals is the frequency content which varies dynamically against the time in the case of ground motion records. However, it is often quantified using the equivalent frequency PGA/PGV [27] as if it was a sinusoid signal. Another quantity that has been proposed and related to the frequency is the Potential Destructiveness Measure after Araya and Saragoni (I_{AS}) [31], resulting from the zero crossings number of the acceleration signal (u_o) per unit of time.

Various definitions have been given in the past for the strong motion duration of a seismic excitation, in order to identify the time interval of the signal in which the vast amount of its total intensity is released. In this work, the strong motion durations defined by Trifunac and

Num	Name	Expression	Ref.	Num	Name	Expression	Ref.
1	PGA	$max\lvert a_g(t)\rvert$	[27]	9	SMD_{ROG}	$t(H_d = 97.5\%) - t(H_d = 2.5\%)$	[33]
2	PGV	$max\lvert v_g(t)\rvert$	[27]	10	SMD_{Bolt}	$\Delta t(a > 0.05g)$	[34]
3	PGD	$max\lvert d_g(t)\rvert$	[27]	11	P_{90}	$\frac{I_A(H_d=95\%) - I_A(H_d=5\%)}{SMD_{TB}}$	[27]
4	I_A	$\frac{\pi}{2g}\int_0^{tend} a_g^2(t)dt$	[28]	12	a_{rms}	$\sqrt{\frac{1}{SMD_{TB}}\int_{t5\%}^{t95\%} a(t)^2 dt}$	[27]
5	CAV	$\int_o^{tend}\lvert a(t)\rvert dt$	[27]	13	I_c	$a_{rms}^{1.5}\cdot SMD_{TB}^{0.5}$	[27]
6	PGA/PGV	$\frac{PGA}{PGV}$	[27]	14	I_{FVF}	$PGV\cdot SMD_{TB}^{0.25}$	[35]
7	I_{AS}	$\frac{I_A}{u_o^2}$	[31]	15	I_{RG}	$PGD\cdot SMD_{TB}^{\frac{1}{3}}$	[36]
8	SMD_{TB}	$t(H_d = 95\%) - t(H_d = 5\%)$	[32]	16	SI_H	$\int_{0.1}^{2.5} PSV(T, \xi = 0.05)dT$	[37]

Table 2: Mathematical Expressions of IMs.

| | PGA | PGV | PGD | I_A | CAV | $\frac{PGA}{PGV}$ | I_{AS} | SMD | | | P_{90} | a_{rms} | I_c | I_{FVF} | I_{RG} | SI_H |
| | | | | | | | | TB | ROG | $Bolt$ | | | | | | |
	$\frac{cm}{s^2}$	$\frac{cm}{s}$	cm	$\frac{cm}{s}$	$\frac{cm}{s}$	s^{-1}	$\frac{cm}{s}$	s	s	s	$\frac{cm}{s^2}$	$\frac{cm}{s^2}$	$\frac{cm^{1.5}}{s^{2.5}}$	$cm\cdot s^{-0.75}$	$cm\cdot s^{\frac{1}{3}}$	cm
mean	305.0	29.5	55.5	135.6	747.7	13.1	3.5	13.6	17.0	11.9	16.3	77.1	2320.5	53.2	152.2	91.5
std	254.0	24.9	132.0	190.4	519.3	8.8	5.2	10.0	11.2	10.1	26.8	65.3	2516.4	42.8	417.5	73.5
min	32.7	1.2	0.1	0.7	28.0	1.7	0.0	0.6	1.9	0.0	0.1	7.4	55.3	1.4	0.1	2.1
25%	124.7	10.5	2.7	20.6	325.2	7.2	0.3	6.1	8.6	4.8	1.5	30.9	639.9	18.6	5.6	31.5
50%	218.8	23.8	10.5	81.1	693.1	11.3	1.5	10.4	13.6	10.2	4.7	54.2	1544.1	43.2	21.5	76.5
75%	414.2	41.5	40.5	157.2	1008.6	16.6	4.9	18.0	22.8	14.8	16.9	102.8	2916.0	77.3	98.2	134.0
max	1465.2	130.4	1314.2	1332.4	3119.3	75.5	41.5	46.8	55.5	58.6	131.1	286.3	13323.4	201.8	4625.5	387.1

Table 3: Descriptive Statistics of 2nd Seismic Shock IMs (Train Set-837 examples).

Brady (SMD_{TB}) [32], and by Reinoso, Ordaz and Guerrero (SMD_{ROG}) [33] are assumed. Both of these are based on the time evolution of Arias Intensity according to Husid Diagram. Also, the bracketed duration as described by Bolt (SMD_{Bolt}) [34], which is defined by the first and last exceedance of the 5 percent of g, is employed.

Combining the above parameters results in more complex measures such as Power P_{90} [27], a_{rms} [27], Characteristic Intensity (I_c) [27], the potential damage measure according to Fajfar, Vidic and Fischinger (I_{FVF}) [35] and the IM after Riddell and Garcia (I_{RG}) [36].

Also, we did not use seismic parameters that depend on the fundamental period, like individual spectral values of the structure due to the elongation of the initial elastic period before the start of the second seismic event. Instead, we employed Housner Intensity [37] (SI_H) which accumulate spectral pseudo-velocities to a constant range of possible eigenperiods and presents a high correlation with the structural damage [38, 12, 39]. The calculation of spectrums for the above purpose implemented using OpenSeismoMatlab [40].

All of the mathematical expressions for employed IMs summarised in Table 2. Tables 3 and 4 present the descriptive statistics of the IMs corresponding to the second seismic shocks of the training and test datasets, respectively. We observed that the training set covers the major range of the test set values. From the initially 954 artificially generated sequences, finally we hold 837 of them, which the NTHAs did not present convergence problems.

	PGA	PGV	PGD	I_A	CAV	$\dfrac{PGA}{PGV}$	I_{AS}	SMD TB	ROG	$Bolt$	P_{90}	a_{rms}	I_c	I_{FVF}	I_{RG}	SI_H
	$\dfrac{cm}{s^2}$	$\dfrac{cm}{s}$	cm	$\dfrac{cm}{s}$	$\dfrac{cm}{s}$	s^{-1}	$\dfrac{cm}{s}$	s	s	s	$\dfrac{cm}{s^2}$	$\dfrac{cm}{s^2}$	$\dfrac{cm^{1.5}}{s^{2.5}}$	$cm \cdot s^{-0.75}$	$cm \cdot s^{\frac{1}{3}}$	cm
mean	239.2	19.4	15.8	91.4	448.5	16.0	2.5	9.2	11.9	5.2	14.6	69.0	1690.4	31.5	34.1	62.8
std	212.3	20.9	42.4	184.8	571.7	10.0	5.8	7.7	9.3	7.1	26.1	66.3	2240.4	37.3	99.2	75.7
min	7.4	0.8	0.0	0.2	29.3	2.9	0.0	0.5	0.7	0.0	0.0	1.8	12.5	1.4	0.1	1.5
25%	68.7	4.5	0.7	3.4	133.3	9.2	0.1	3.9	5.5	0.2	0.5	17.7	189.3	7.8	1.3	12.5
50%	185.4	10.8	2.6	15.4	234.9	13.3	0.2	7.0	9.7	2.5	2.8	42.2	616.8	16.2	4.6	31.7
75%	369.7	30.1	9.4	99.5	588.2	20.6	1.2	11.0	14.3	7.8	18.9	108.7	2527.0	42.0	20.1	80.7
max	985.8	91.8	277.5	1165.6	3354.8	53.0	32.3	37.4	44.2	42.7	170.7	326.2	11393.9	185.6	737.3	304.0

Table 4: Descriptive Statistics of 2nd Seismic Shock IMs (Test Set-119 examples).

3.2 Damage indicators

In the present section, the damage indices considered in this study are described. The structural damage is used both as an input feature to take into account the existing damage due to the first seismic shock and as a target of the ML problem to describe the accumulated damage after the occurrence of the second earthquake.

3.2.1 Damage index after Park and Ang (DI_{PA})

The damage index, proposed by Park and Ang [41], results from summing the contributions of the maximum flexural responses and the hysteretic energy consumed in the plastic hinges areas and is calculated by Equation (1) modified by Kunnath et al. (Equation (2)) [42]. The total damage index [22] is calculated as a weighted average of the sub-factors, weighted by the percentages of the total energy consumed by each member of the construction, according to Equation (3). The overall damage index $DI_{G,PA}$ as close to zero as possible implies a complete damage-free structural system with an elastic response but when tends to unit characterises a near collapse structure. For abbreviation reason in the next sections $DI_{G,PA}$ referred to simple as DI_{PA}.

$$DI_{PA} = \frac{\delta_m}{\delta_u} + \frac{\beta}{Q_y \delta_u} \int dE \tag{1}$$

$$DI_{PA,component} = \frac{\theta_m - \theta_r}{\theta_u - \theta_r} + \frac{\beta}{\theta_u M_y} E_h \tag{2}$$

$$DI_{G,PA} = \frac{\sum E_i DI_{PA,component}}{\sum E_i} \tag{3}$$

where δ_m is the maximum element displacement response, δ_u is the ultimate element displacement, β is the model constant parameter for strength deterioration proposed by Park et al. [43], $\int dE$ is the cumulative hysteretic energy consumed by the element during its response, Q_y is the yield strength of the element, θ_m is the maximum element rotation during the time history response, θ_u is the ultimate capacity of the element, θ_r is the recoverable element rotation during unloading.

3.2.2 Damage index after DiPasquale and Cakmak (DI_{DC})

During high-intensity seismic events, it is known that the cross-sections in plastic hinges areas of a building can be severely cracked or even present steel yielding, resulting in a stiffness reduction and consequently degradation of the overall stiffness of the structure. Therefore, it is expected that the building will become more flexible and its fundamental period increase. On the above mentioned increasing of the fundamental period due to inelastic dynamic response of a building, based the DiPasquale and Cakmak [44] damage index, calculated according to (4):

$$DI_{DC} = 1 - \frac{T_{0_{initial}}}{T_{0_{equivalent}}}$$ (4)

$T_{0_{initial}}$: The fundamental period before the starting of analysis.
$T_{0_{equivalent}}$: The fundamental period at the end of the analysis.

3.2.3 Max Inter-Storey Drift Ratio (MIDR)

The Max Inter-Storey Drift Ratio [45] during the seismic response of the construction can be used as a measure of the performance it has shown against the earthquake it has hit. The increase in the relative displacements results in an increase of the required floor ductility with similar effects for the rotation and curvature ductility in the areas of the plastic hinges of each component and their cross-sections respectively. So estimating the magnitude of generalised displacements such as MIDR can indirectly aggregate local damages and quantify the total degree of structural damage define by (5).

$$MIDR = \frac{|u|_{max}}{h} 100\%$$ (5)

$|u|_{max}$: Max absolute interstorey deformation recorded during the time history response
h: Storey height

3.2.4 Residual Max Inter-Storey Drift Ratio (RMIDR)

Residual displacements cause an important role in the post-seismic performance state of a building to indicate the established damage. Also, consist a critical variable for the repairability of seismic damaged structures [45]. We included it in the problem and evaluate it as the residual maximum inter-storey drift ratio in the post-seismic state, after the end of building oscillation.

4 APPLICATION OF NEURAL NETWORKS

The most conventional used Neural Network configuration is the Multi-Layer Perceptron (MLP) [46], which is considered as the simpler learning architecture, consisted from of a few linear layers which are interconnected successively. Each of them receive input from the previous layer, which is multiplied matricialy by appropriate weights, subsequently a vector of biases is added. Finally, in every element of the previously defined vector, a nonlinear function is applied to produce the activations of the following layer. This forward process continues until the regression is completed by concluding with the target value from the output layer. The final output values are compared with the real ones, and the regression error is calculated using an

appropriate loss function. The weights of all layers are updated one by one using Stochastic Gradient Descent to minimise the cost function. Data scientists, based on their assumptions and experience, tune their numerous parameters by relating them to the specific problems they need to solve, making use of the available training datasets. Thus, the application of MLP to realistic problems is still challenging and specialised for each problem.

4.1 Gaussian kernel functions

Since the problem under consideration essentially involves the prediction of values based on similar training data samples, kernel function modelling was chosen, which is an appropriate solution as it allows for accurate modelling of data with complex dependencies between input features and target data. They are essentially representations of the input vectors x in $\Re$, which have specific form and properties and are highly generalisable. The appropriate transformation of the input vectors in order to achieve a more easily generalisable solution to the regression problem first requires the transformation of the input space into a transformed feature space with a higher dimension and with a nonlinear representation. The transfer of the problem in a higher dimensional feature space is achieved by the kernel method. In this work, the Gaussian Radial Basis Function kernel (GRBFk) was used, which is appropriate in cases where the identification or manipulation of real values, which tend to cluster around an average value, is required.

4.2 Extreme learning machines

In this paper, a holistic Single-hidden Layer Feed Forward Network (ShLFFN) shallow architecture approach with N neurons in the hidden layer, randomly chosen input weights and random values of bias constants on the hidden layer neurons, while the output weights are computed by a single matrix multiplication, which automates and optimally solves the problem, was used. These ShLFFN architectures, called Extreme Learning Machines (ELMs) [47], used in complex problems, due to their ability to access any continuous function. Also, the networks in question can learn N samples with accuracy, while their learning speed can be even thousands of times faster than conventional Feed Forward Network Networks, as their training is not based on time-consuming repetitive processes such as the algorithm. Back Propagation, which changes the weights of the neural network by estimating the quadratic error between the target vectors and the actual network outputs, for all training samples, which are entered into the network in a random serial manner and for many repetitions (epochs)

Driven by the observation that the first stage of the algorithm can be considered as a process of mapping the training vectors into a higher dimensional vector space, in the hidden layer we chose a sufficiently large number of neurons (at least twice the input parameters) so that the algorithm has the property of a universal approximator and can train ShLFFN efficiently. More specifically, the input data are mapped to a random $\tilde{N}$-dimensional space with a discrete training set N where (x_i, t_i) with $i \in [1, N]$ $x_i \in \Re^d$ and $t_i \in \Re^c$. The output of the network presented in Equation (6).

$$f_L(x) = \sum_{i=1}^{\tilde{N}} \beta_i h_i(x) = \mathbf{h}(x)\beta, \, i \in [i, N] \tag{6}$$

where $\beta = [\beta_1, \ldots, \beta_{\tilde{N}}]^T$ is the weight matrix between hidden and output neurons $h(x) = [g(x_1), \ldots, g(x_2)]$ is the output of hidden neurons for the input x, and $g(x_i)$ is the output of ith hidden node. Based on N length training set $\{(x_i, t_i)\}_{i=1}^{N}$, the ELM can solve the learning problem $\mathbf{H}\beta = \mathbf{T}$, where $\mathbf{T} = [t_1, \ldots, t_N]^T$ are the target labels and the output matrix of the hidden layer

output presented in Equation (7).

$$\mathbf{H}(\mathbf{W}, \mathbf{b}, \mathbf{X}) = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_1 + b_{\tilde{N}}) \\ \vdots & \ddots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \cdots & g(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_N + b_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}} \tag{7}$$

Before the training, the weight matrix $\mathbf{W}$ and the bias vectors $\mathbf{b}$ initialised randomly in the interval $[-1, 1]$, given $\mathbf{w}_j = [w_{j1}, w_{j2}, \ldots, w_{jm}]^T$ and $\beta_j = [\beta_{j1}, \beta_{j2}, \ldots, \beta_{jm}]^T$ .The output weight matrix of the hidden layer $\mathbf{H}$ came up from Equation (8), given the training data and the activation function g.

$$\mathbf{H} = g(\mathbf{W}\mathbf{x} + \mathbf{b}) \tag{8}$$

The hidden layer (Figure 2) is responsible for transforming the input data into a different representation. The transformation is achieved in two steps:

- First Step. The input data is input to the hidden layer using the weights and corresponding values of the input layer.

- Second Step. The data is transformed based on a nonlinear transformation function.

Finally, the output weights β could compute from Equation (9).

$$\beta = \mathbf{H}^* \mathbf{T} \tag{9}$$

where $\mathbf{H}^*$ is the Moore-Penrose generalised inverse matrix for $\mathbf{H}$.

4.3   Bootstrap sampling and ensemble technique

Bootstrap sampling [48] is adopted so that different subsets of the dataset can train the model in the most fruitful way, to increase the efficiency of the method. Bootstrap sampling is the process of applying progressively larger random samples until the accuracy of the neural network is improved. Every sample is used to train an individual model and the results of each model are summarised by "voting", i.e. for any input vector, each classifier predicts the output variable and finally, for every variable the value with the most "votes" is chosen. This methodology, which belongs to the Ensemble methods, is called Bagging and has many advantages, such as reducing covariance. The Ensemble technique [49] is used due to the fact that in the specific problem, the prediction results show multivariability, which can be attributed to the sensitivity of the correlation models to the data and the complex relationship that describes them. Also, an important advantage gained by the proposed system is focused on the fact that it offers more efficient and stable prediction model, since the overall behaviour of multiple MLP ANNs is less noisy than a single one, and in any case reduces the overall risk of a particularly inaccurate values. Finally, in order to evaluate the reliability of the implemented bagging method, the variance of the expected error is checked to be around the mean error value.

5   RESULTS

This section presents both the results from NTHAs in terms of overall seismic damage indices and the neural network testing in terms of correlations between real and predicted values. In the case of structural damage indices we present both the results of the first shocks NTHAs adopted between the inputs of the ML problem (Figure 2) and the corresponding after the second shocks which employed as targets of the problem.
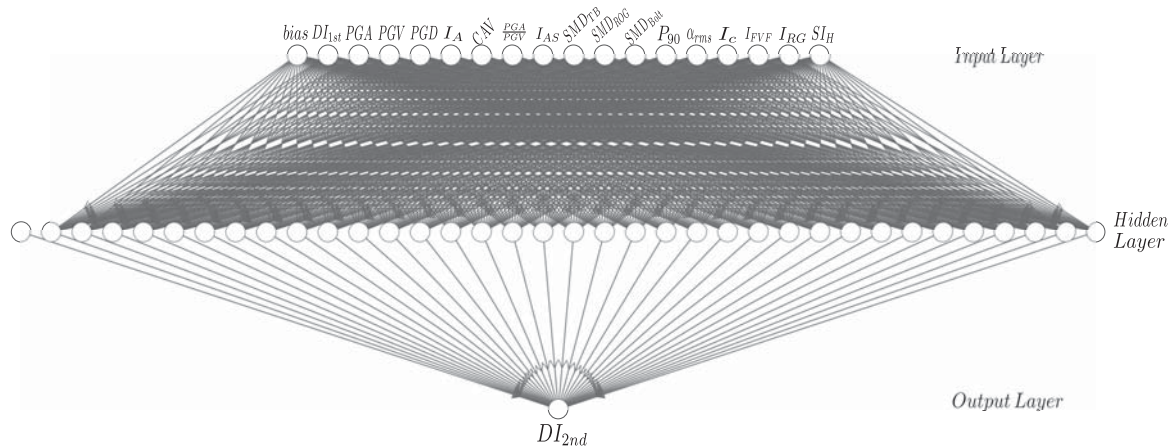
Figure 2: An ELM neural network architecture uses as input arguments both damage index from the first seismic event and the IMs of the second event to predict the final damage index (after second event).

5.1 Results of NTHAs

The results of NTHAs represented in terms of damage indices after the occurrence of the first shocks and after the sequences (Figure 3). As we compare the distributions of damage indices after the first and second shock, we see a clear increase after the second shock, both in the training and test set. The increase is higher in the case of DI_{DC}. The maximum value of this damage index seems to be constant around 0.7 both in the training and test set. The other damage indices have a smaller increase but a clear offset of the upper-end value both to training and test set as we shown in Figures 3a and 3b respectively.

5.2 Results of ANN testing

Since the model described in Section 4, trained using the dataset generated from artificial sequences, several correlation coefficients calculated between predicted and actual values of the targets on the test set (natural sequences) to evaluate the effectiveness of the fitted ML model. The results of three correlation coefficients (Pearson, Kendall and Spearman) and two error metrics, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), for the final fitted model are presented in Table 5. We can observe that the damage indices with the best predictive behaviour are the $MIDR$ and DI_{PA} which display the higher values of Pearson, Kendall and Spearman coefficients. The relation between predicted and the actual values of DI_{DC} looks to be quite linear since the values of Pearson and Spearman coefficients are similar and quite high.

	Pearson	Kendall	Spearman	MAE	RMSE
DI_{PA}	0.9104	0.7956	0.9366	0.0574	0.1059
DI_{DC}	0.8875	0.7140	0.8898	0.0648	0.0867
$MIDR$	0.8960	0.8277	0.9582	0.3542	0.7933
$RMIDR$	0.8316	0.7488	0.9146	0.2956	0.6870

Table 5: Results of correlation between actual and predicted values on the test set.
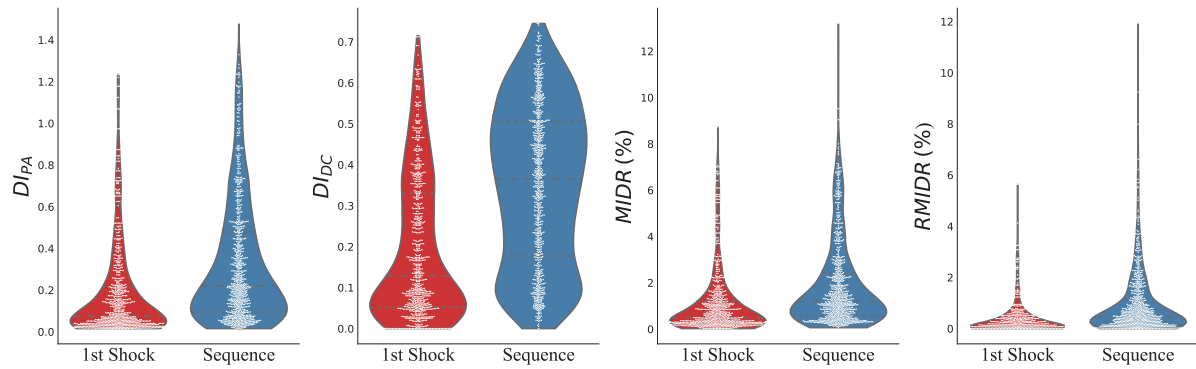
(a) Violin plots of damage indicators distributions of the training set (837 artificial sequences).
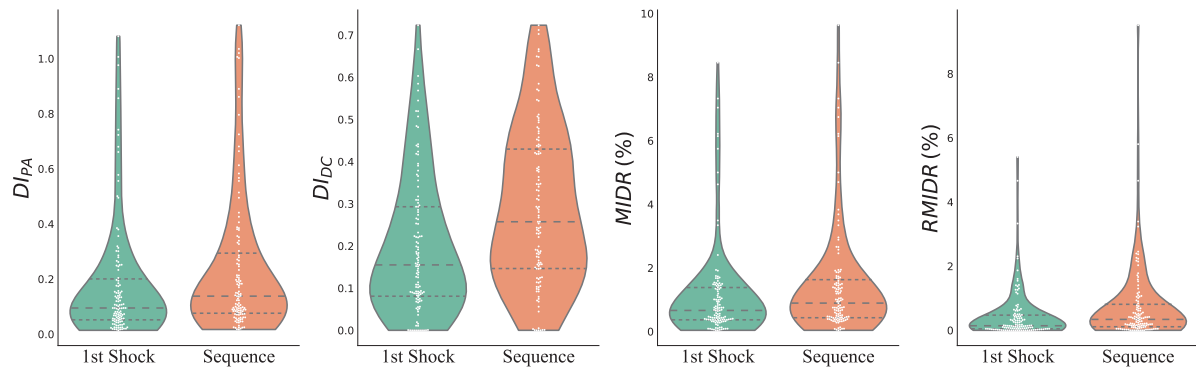
(b) Violin plots of damage indicators distributions of the test set (119 natural seismic sequences).

Figure 3: Violin plots of damage indicators distributions.

6 CONCLUSIONS

In the present study, we used an ensemble neural network methodology to predict the final damage of an 8-storey RC frame under successive earthquakes. For this purpose, we employed a training dataset of 837 artificial seismic sequences and a test set of 119 natural seismic sequences. Finally, the main conclusions are summarised as follows:

- In both training and test sets, the structural damage under the sequences is increased in comparison with the corresponding damage that occurred after the initial shocks.

- In general, the predictive capacity of the configured Ensemble Neural Network seems to be quite satisfactory.

- The Ensemble Neural Network presents a higher performance in forecasting the final structural damage in terms of DI_{PA} (damage index after Park-Ang) and MIDR (Maximum Inter-storey Drift Ratio).

For future machine learning applications in structural damage prediction due to successive earthquakes enriched dataset of natural seismic sequences may be considered. Moreover, a wide range of structural systems should be examined in order to incorporate structural attributes as ML problem input features.

REFERENCES

[1] C. Karakostas, N. Klimis, V. Lekidis, K. Makra, B. Margaris, K. Morfidis, C. Papaioannou, E. Rovithis, T. Salonikios, D. Sotiriadis, and N. Theodoulidis, Thessaly Earthquakes M6.3, 03/03/2021 and M6.1, 04/03/2021 – Preliminary Report, tech. rep., Research Unit ITSAK, EPPO and Department of Civil Engineering, DUTh, Mar. 2021.

[2] E. Lekkas, K. Agorastos, S. Mavroulis, H. Kranis, E. Skourtsos, P. Carydis, M. Gogou, K.N. Katsetsiadou, G.A. Papadopoulos, I. Triantafyllou, A. Agalos, M. Sotiris, E. Stamati, D. Psarris, G. Kaviris, V. Kapetanidis, P. Papadimitriou, A. Karakonstantis, I. Spingos, V. Kouskouna, I. Kassaras, K. Pavlou, N. Voulgaris, M. Mavrouli, S. Pavlides, A. Chatzipetros, S. Sboras, E. Kremastas, A. Chatziioannou, A.A. Kiratzi, C. Papazachos, N. Chatzis, V. Karakostas, E. Papadimitriou, I. Koukouvelas, K. Nikolakopoulos, A. Kyriou, D. Apostolopoulos, V. Zygouri, S. Verroios, A. Belesis, I. Tsentzos, P. Krassakis, K. Lymperopoulos, A. Karavias, D. Bafi, T. Gatsios, M. Karatzia, I. Gkougkoustamos, T. Falaras, I. Parcharidis, G. Papathanassiou, C. Evangelidis, V. Karastathis, A. Tselentis, A. Ganas, V. Tsironi, I. Karasante, S. Valkaniotis, D. Galanakis, G. Konstantopoulou, N. Papadopoulos, A. Kourou, M. Manousaki, and T. Thoma, The early March 2021 Thessaly (Greece) earthquake sequence, *Newsletter of Environmental, Disaster and Crises Management Strategies*.

[3] C. Amadio, M. Fragiacomo, and S. Rajgelj, The effects of repeated earthquake ground motions on the non-linear response of SDOF systems, *Earthquake Engineering & Structural Dynamics*, **32**(2), 291–308, 2003.

[4] G.D. Hatzigeorgiou and D.E. Beskos, Inelastic displacement ratios for SDOF structures subjected to repeated earthquakes, *Engineering Structures*, **31**(11), 2744–2755, 2009.

[5] G.D. Hatzigeorgiou and A.A. Liolios, Nonlinear behaviour of RC frames under repeated strong ground motions, *Soil Dynamics and Earthquake Engineering*, **30**(10), 1010–1025, 2010.

[6] M. Hatzivassiliou and G.D. Hatzigeorgiou, Seismic sequence effects on three-dimensional reinforced concrete buildings, *Soil Dynamics and Earthquake Engineering*, **72**, 77–88, 2015.

[7] F. Hosseinpour and A. Abdelnaby, Effect of different aspects of multiple earthquakes on the nonlinear behavior of RC structures, *Soil Dynamics and Earthquake Engineering*, **92**, 706–725, 2017.

[8] I.E. Kavvadias, P.Z. Rovithis, L.K. Vasiliadis, and A. Elenas, Effect of the aftershock intensity characteristics on the seismic response of RC frame buildings, in *Proceedings of the 16th European Conference on Earthquake Engineering, Thessaloniki, Greece*, 18–21, 2018.

[9] Z. Zhou, X. Yu, and D. Lu, Identifying Optimal Intensity Measures for Predicting Damage Potential of Mainshock–Aftershock Sequences, *Applied Sciences*, **10**(19), 6795, 2020.

[10] O.R. De Lautour and P. Omenzetter, Prediction of seismic-induced structural damage using artificial neural networks, *Engineering Structures*, **31**(2), 600–606, 2009.

[11] P. Alvanitopoulos, I. Andreadis, and A. Elenas, Neuro–fuzzy techniques for the classification of earthquake damages in buildings, *Measurement*, **43**(6), 797–809, 2010.

[12] K. Morfidis and 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.

[13] K. Kostinakis and K. Morfidis, Application of Artificial Neural Networks for the Assessment of the Seismic Damage of Buildings with Irregular Infills' Distribution, in *Seismic Behaviour and Design of Irregular and Complex Civil Structures III*, 291–306, Springer, 2020.

[14] J. González, W. Yu, and L. Telesca, Earthquake Magnitude Prediction Using Recurrent Neural Networks, in *Multidisciplinary Digital Publishing Institute Proceedings*, **24**(1), 22, 2019.

[15] S. Mangalathu and H.V. Burton, Deep learning-based classification of earthquake-impacted buildings using textual damage descriptions, *International Journal of Disaster Risk Reduction*, **36**, 101111, 2019.

[16] R. Zhang, Z. Chen, S. Chen, J. Zheng, O. Büyüköztürk, and H. Sun, Deep long short-term memory networks for nonlinear structural seismic response prediction, *Computers & Structures*, **220**, 55–68, 2019.

[17] J. Li, Z. He, and X. Zhao, A data-driven building's seismic response estimation method using a deep convolutional neural network, *IEEE Access*, 1, Mar 2021.

[18] B.K. Oh, Y. Park, and H.S. Park, Seismic response prediction method for building structures using convolutional neural network, *Structural Control and Health Monitoring*, **27**(5), e2519, 2020.

[19] L. Luzi, G. Lanzano, C. Felicetta, M. D'Amico, E. Russo, S. Sgobba, F. Pacor, and OR-FEUS Working Group 5, Engineering Strong Motion Database (ESM) (Version 2.0), *Istituto Nazionale di Geofisica e Vulcanologia (INGV), Rome, Italy*, 2020.

[20] T.D. Ancheta, R.B. Darragh, J.P. Stewart, E. Seyhan, W.J. Silva, B.S. Chiou, K.E. Wooddell, R.W. Graves, A.R. Kottke, D.M. Boore, *et al.*, Peer NGA-West2 database, tech. rep., Pacific Earthquake Engineering Research Center Berkeley, CA, 2013.

[21] G. Rossum, Python reference manual, 1995.

[22] R. Valles, A.M. Reinhorn, S.K. Kunnath, C. Li, and A. Madan, IDARC2D version 4.0: A computer program for the inelastic damage analysis of buildings, tech. rep., US National Center for Earthquake Engineering Research (NCEER), 1996.

[23] Y.J. Park, A.M. Reinhorn, and S.K. Kunnath, *IDARC: Inelastic Damage Analysis of Reinforced Concrete Frame–Shear–Wall Structures*, National Center for Earthquake Engineering Research Buffalo, NY, 1987.

[24] CEN, EN 1992-1-1 Eurocode 2: Design of concrete structures - Part 1-1: General rules and rules for buildings, *Brussels: European Committee for Standardization*, 2005.

[25] J.W. Eaton, GNU Octave and reproducible research, *Journal of Process Control*, **22**(8), 1433–1438, 2012.

[26] J.W. Eaton, D. Bateman, S. Hauberg, and R. Wehbring, *GNU Octave version 6.2.0 manual: a high-level interactive language for numerical computations*, 2021.

[27] S.L. Kramer, *Geotechnical earthquake engineering*, Pearson Education India, 1996.

[28] A. Arias, A measure of earthquake intensity. Seismic Design for Nuclear Power Plants, *Massachusetts Institute of Technology*, 1970.

[29] A. EPRI, Criterion for determining exceedance of the operating basis earthquake. Rapport NP-5930 2848-16, *Electric Power Research Institute, USA*, 1988.

[30] R. Husid, Características de terremotos. Análisis general, *Revista IDIEM*, **8**(1), ág–21, 1969.

[31] R. Araya, Earthquake accelerogram destructiveness potential factor, in *Proc. 8th World Conference on Earthquake Engineering, 1985. 7*, **11**, 835–843, 1985.

[32] M.D. Trifunac and A.G. Brady, A study on the duration of strong earthquake ground motion, *Bulletin of the Seismological Society of America*, **65**(3), 581–626, 1975.

[33] E. Reinoso, M. Ordaz, and R. Guerrero, Influence of strong ground-motion duration in seismic design of structures, in *Proceedings, 12th World Conference on Earthquake Engineering*, **1151**, 2000.

[34] B.A. Bolt, Duration of strong ground motion, in *5th World Conference on Earthquake Engineering*, **292**, 25–29, 1973.

[35] P. Fajfar, T. Vidic, and M. Fischinger, A measure of earthquake motion capacity to damage medium-period structures, *Soil Dynamics and Earthquake Engineering*, **9**(5), 236–242, 1990.

[36] R. Riddell and J.E. Garcia, Hysteretic energy spectrum and damage control, *Earthquake Engineering & Structural Dynamics*, **30**(12), 1791–1816, 2001.

[37] G. Housner, Spectrum intensities of strong-motion earthquakes, 1952.

[38] A. Masi, M. Vona, and M. Mucciarelli, Selection of Natural and Synthetic Accelerograms for Seismic Vulnerability Studies on Reinforced Concrete Frames, *Journal of Structural Engineering*, **137**(3), 367–378, 2011.

[39] P.C. Lazaridis, I.E. Kavvadias, and L.K. Vasiliadis, Correlation between Seismic Parameters and Damage Indices of Reinforced Concrete Structures, 2019.

[40] G. Papazafeiropoulos and V. Plevris, OpenSeismoMatlab: A new open-source software for strong ground motion data processing, *Heliyon*, **4**(9), e00784, 2018.

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

[42] S.K. Kunnath, A.M. Reinhorn, and R. Lobo, IDARC Version 3.0: A program for the inelastic damage analysis of reinforced concrete structures, tech. rep., US National Center for Earthquake Engineering Research (NCEER), 1992.

[43] Y.J. Park, A.H. Ang, and Y.K. Wen, Damage-limiting aseismic design of buildings, *Earthquake Spectra*, **3**(1), 1–26, 1987.

[44] E. DiPasquale and A. Çakmak, Detection of seismic structural damage using parameter-based global damage indices, *Probabilistic Engineering Mechanics*, **5**(2), 60–65, 1990.

[45] FEMA, P58–1 Seismic performance assessment of buildings (volume 1-Methodology), *Federal Emergency Management Agency, Washington*, **1**, 2012.

[46] M. Minsky and S.A. Papert, *Perceptrons: An introduction to computational geometry*, MIT Press, 2017.

[47] G.B. Huang, Q.Y. Zhu, and C.K. Siew, Extreme learning machine: theory and applications, *Neurocomputing*, **70**(1-3), 489–501, 2006.

[48] T.H. Lee, A. Ullah, and R. Wang, Bootstrap aggregating and random forest, in *Macroeconomic Forecasting in the Era of Big Data*, 389–429, Springer, 2020.

[49] L.K. Hansen and P. Salamon, Neural network ensembles, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**(10), 993–1001, 1990.