Use of Artificial Neural Networks to predict strong ground motion duration of Interplate and Inslab Mexican Earthquakes for Soft and Firm Soils

R. Flores-Mendoza[1], J. U. Rodríguez-Alcántara[1], A. Pozos-Estrada[1*] and R. Gómez[1]

Resumen

Se desarrollan modelos de red neuronal artificial para predecir la duración del movimiento fuerte del terreno de eventos de subducción en suelos firme y blando. Para entrenar la red neuronal artificial se emplea una base de datos con un total de 3153 registros sísmicos con dos componentes horizontales para eventos de interplaca e intraslab. El método de componente principal es usado para realizar una reducción dimensional de los parámetros de entrada para desarrollar los modelos de red neuronal artificial. Los valores predichos de la duración del movimiento fuerte del terreno por la red neuronal entrenada son comparados con aquellos estimados con expresiones empíricas. En general, la duración del movimiento fuerte del terreno predicha con la red neuronal artificial sigue la misma tendencia que la calculada con las ecuaciones empíricas, aunque en algunos casos, ésta presenta cambios repentinos en su comportamiento. Por esta razón, es recomendado llevar a cabo varias verificaciones de los modelos entrenados de la red neuronal artificial antes de usarlos para más aplicaciones ingenieriles, por ejemplo, la simulación de registros sintéticos o la evaluación de índices sísmicos de daño.

Palabras Clave: Red neuronal artificial, duración del movimiento fuerte del terreno, eventos de subducción, expresiones empíricas y México.

———————————————————

Corresponding author at mdiazv@imp.mx
[1]Instituto de Ingeniería, Universidad Nacional Autónoma de México,
Ciudad Universitaria, 04510, Ciudad de México, México.

Abstract

Artificial neural network models are developed to predict strong ground motion duration of subduction events for soft and firm soils. To train the artificial neural network a database with a total of 3153 seismic records with two horizontal components for interplate and inslab earthquakes is employed. The principal component method is used to carry out a dimensionality reduction of the input parameters to develop the artificial neural network models. The predicted values of the strong ground motion duration trained by the artificial neural network models are compared with those estimated with empirical expressions. In general, the strong ground motion duration predicted with the artificial neural networks follows the same tendency of that calculated with the empirical equations, although in some cases, the strong ground motion duration predicted by using the artificial neural network models presents sudden changes in its behavior. For this reason, it is recommended to carry out several verifications of the trained artificial neural network models before using them for further engineering applications, for example the simulation of synthetic records or the evaluation of seismic damage indices.

Key words: Artificial neural network, strong ground motion duration, subduction events, empirical expressions and Mexico.

Introduction

An artificial neural network (ANN) is an effective tool that has been used to solve a great variety of engineering problems due to its flexibility to cope highly nonlinear problems. It is worth mentioning that ANN models as any other prediction technique have advantages and disadvantages (Pande and Shin, 2004). Some advantages of using ANN models include the storage information of the ANN with multidimensional inputs, the prediction of multiple outputs with a single ANN model, the ability to work with incomplete knowledge and machine learning. Perhaps two of the major drawbacks of ANN models are the loss of transparency and the unexpected behavior of the ANN model that may produce erroneous results.

The employ of ANNs in seismic engineering is vast, for example, García *et al.* (2007), used ANNs to estimate peak ground accelerations (PGA) for Mexican subduction earthquakes. Hong *et al.* (2012), employed 39 California earthquakes to predict pseudospectral accelerations (SA) and PGA. More recently, Pozos-Estrada *et al.* (2014) developed ANNs models to predict PGA and SA for Mexican inslab and interplate earthquakes. They showed that the predicted PGA and SA values by the trained ANN models, in general, follow a similar trend to those predicted by ground motion prediction equations (GMPEs). ANNs have also been used to estimate strong ground motion duration (SGMD) of earthquakes. Alcántara *et al.* (2014), developed ANNs to predict SGMD by using information compiled from the Mexican states of Puebla and Oaxaca. The prediction of SGMD has also been applied to other tectonic regions, for example Arjun and Kumar (2011), used ANN models to estimate the SGMD by using Japanese earthquake records, and more recently, Yaghmaei-Sabegh (2018), used a general regression neural network to estimate earthquake ground-motion duration recorded at the Iranian plateau.

The study of SGMD to estimate the structural damage has also been carried out by several researchers (Housner *et al.* 1952; Salmon *et al.*, 1992; Bommer and Martínez, 1999; Strasser and Bommer 2009; and Lindt and Goh, 2004). The general agreement among these studies is that the structural

damage not only depend on the maximum intensity or frequency content of the ground motion, but also on the SGMD. For the estimation of the seismic-induced structural response, the amplitude and frequency content of the ground motions are of paramount importance; however, if the cumulative structural damage or structural degradation of systems with hysteretic behavior is of interest, SGMD should be integrated as design parameter (Reinoso and Ordaz, 2001). More recently, Bhargavi and Raghukanth (2019), carried out a statistical analysis of several ground motion parameters, including SGMD, to rate damage potential of ground motion records.

The main objective of this study is to develop ANN models to estimate SGMD of interplate and inslab ground motion records of Mexican earthquakes from a broad network of stations. Ground motion records from 1985 up to 2017 are employed. Multilayer perceptron ANN models with back-propagation training were considered. The input parameters considered in the development of the ANN models include the moment magnitude (Mw), closest distance to the fault (R_C), focal depth (H), the vibration period of the soil (T), the seismic moment (M_0), as well as the strike (ϕ), dip (δ) and rake (λ) angles. The principal component method is used to carry out a dimensionality reduction of the input parameters to evaluate the ability of ANN models with different number of input parameters to predict the SGMD. The predicted SGMD values by the trained ANN models are compared with those estimated with empirical equations for comparison purposes.

Strong Ground Motion Database and SGMD Calculation

The strong ground motion database employed integrates information from different networks. A total of 3153 strong ground motion records, each one with two horizontal components from 71 earthquakes were used to develop the ANN models. Table 1 and 2 summarize the interplate and in-slab seismic events considered, respectively. The database for interplate includes 50 events with Mw from 5.0 to 8.1, while de database for intermediate-depth normal-faulting inslab events considers 21 with Mw within 5.1 and 7.1. It is noted that the seismic event occurred on September 7, 2017 was not included in the inslab database since it did not cause important damage in Mexico City (Pozos-Estrada *et al.*, 2019) and because the peak ground acceleration registered in Mexico City at lake-bed was below 4 Gal, which was the threshold recommended by Reinoso and Ordaz (2001) as selection criterion of records to calculate SGMD. However, the inslab event occurred on September 19, 2017, which was, together with the interplate event occurred on September 19, 1985, one of the deadliest for Mexico City (Singh *et al.*, 2018, Franke *et al.*, 2019) was included. A map showing the epicenters and recording stations considered is presented in Figure 1. Also in Figure 1, the geotechnical zones (i.e., firm, transition and lakebed) according to the Mexico City design code (2017) are presented. It is worth mentioning that the Mexico City design code (2017) includes rock in the firm soil zone. Figure 2 presents the distribution of Mw, H and M_0 with respect to R_C for the seismic events used.

For the analyses, only two type of soils were considered (soft and firm soil). The transition zone was included in the classification of soft soil on the basis that the dominant period of the soil at such zone is greater than 0.5 s, which is the limiting value used in the Mexican design code to separate firm soil from the rest, and that no evidence that duration is affected by amplification effects on strong ground motion (Singh and Ordaz, 1993). It is worth mentioning that the soft soil of Mexico City consists of lacustrine deposits with saturated clays and sand lenses, the transition soil is composed by alluvial deposits and the firm soil consists of basaltic and andesitic lava, ashes and epiclastic deposits (Marsal and Mazari, 1959; Flores-Estrella *et al.*, 2007). The firm soil of Mexico City can be classified

as class B according to NEHRP Recommended provisions for seismic regulations for new buildings and other structures (2004).

Table 1. Interplate events used in training the ANN models

Event No.	No. of Rec.	Date (dd/mm/yy)	M_w	Lat. °N	Long °W	H (km)	Institution*
1	9	19/09/85	8.1	18.081	102.94	15	CFE-GIEC, II**, IG**
2	22	21/09/85	7.6	18.021	101.48	15	CFE-GIEC, II**, IG**
3	8	29/10/85	5.4	17.583	102.64	20.3	CFE-GIEC, II**, IG**
4	19	30/04/86	7.0	18.024	103.06	20	CFE-GIEC, II**, IG**
5	8	05/05/86	5.5	17.765	102.80	19.9	CFE, II**, IG**
6	60	08/02/88	5.8	17.494	101.16	19.2	CFE, CIRES, FICA, II**, IG**
7	114	25/04/89	6.9	16.603	99.4	19	CFE, CIRES, FICA, II**, IG**
8	112	31/05/90	5.9	17.106	100.89	16	CENAPRED, CIRES, FICA, GIEC, II**, IG**
9	50	01/04/91	5.4	16.044	98.39	25.6	CENAPRED, CIRES, FICA, II**, IG**
10	22	31/03/92	5.1	17.233	101.30	11	CENAPRED, CFE, FICA, II**, IG**
11	74	15/05/93	5.5	16.47	98.72	15	CIRES, II**, IG**
12	200	24/10/93	6.6	16.54	98.98	5	CENAPRED, RIIS, GIEC, CIRES, II**, IG**
13	6	13/11/93	5.4	15.63	99.02	15	CIRES, II**, IG**
14	240	10/12/94	6.3	18.02	101.56	20	CENAPRED, RIIS, GIEC, CIRES, II**, IG**
15	149	14/09/95	7.3	16.31	98.88	22	CENAPRED, CIRES, RIIS, II**, IG**
16	138	09/10/95	7.3	18.74	104.67	5	CENAPRED, CIRES, GIEC, RIIS, II*, IG**
17	55	12/10/95	5.5	19.04	103.70	11	CENAPRED, CIRES, IG**
18	62	25/02/96	6.9	15.83	98.25	5	CENAPRED, CIRES, II**, IG**
19	186	15/07/96	6.5	17.45	101.16	20	CENAPRED, CIRES, GIEC, RIIS, II**, IG**
20	196	11/01/97	6.9	17.9	103.04	16	CPRED, CIRES, GIEC, RIIS, II**, IG**
21	18	21/01/97	5.4	16.44	98.15	18	CENAPRED, CIRES, II**, IG**
22	32	19/07/97	6.3	15.86	98.35	5	CENAPRED, CIRES, II**, IG**
23	18	16/12/97	5.9	15.7	99.04	16	CIRES, II*, IG**
24	22	22/12/97	5.0	17.14	101.24	5	CENAPRED, CIRES, II**, IG**
25	82	03/02/98	6.2	15.69	96.37	33	CENAPRED, CIRES, II**, IG**
26	12	05/07/98	5.0	16.83	100.12	5	CENAPRED, II**, IG**
27	246	30/09/99	7.5	15.95	97.03	16	CENAPRED, CIRES, II**, IG**, RIIS
28	149	09/08/00	6.1	17.99	102.66	16	CIRES, II**, IG**
29	48	08/10/01	5.4	16.94	100.14	10	II**, IG**
30	68	18/04/02	5.4	16.77	101.12	22	CIRES, II*, IG**
31	32	22/01/03	7.5	18.6	104.22	9	II**, IG**
32	67	01/01/04	5.6	17.34	101.42	6	CIRES, II**, IG**
33¥	186	13/04/07	6.3	17.09	100.44	41	CIRES, II**, IG**
34	46	06/11/07	5.6	17.08	100.14	9	II**, IG**
35	74	27/04/09	5.7	16.9	99.58	7	CIRES, II, IG
36	48	09/02/10	5.8	15.9	96.86	37	II, IG
37	124	30/06/10	6.0	16.22	98.03	8	CIRES, II, IG
38	176	20/03/12	7.4	16.251	98.521	16	CIRES, II, UAP, IG
39	162	02/04/12	6.0	16.27	98.47	10	CIRES, II, UAP, IG
40	17	11/04/12	6.4	17.9	103.06	16	CIRES, II, IG
41	48	22/09/12	5.4	16.23	98.30	10	CIRES, II

42	36	22/04/13	5.8	17.87	102.19	10	II
43	60	21/08/13	6.0	16.79	99.56	16	II
44	193	18/04/14	7.2	17.18	101.19	10	CIRES, II
45	190	08/05/14	6.4	17.11	100.87	17	CIRES, II
46	50	10/05/14	6.1	17.16	100.95	12	CIRES, II
47	46	11/10/14	5.6	15.97	95.61	10	II
48	58	23/11/15	5.8	16.86	98.94	10	II
49	78	08/05/16	6.0	16.25	97.98	35	II
50	38	27/06/16	5.7	16.2	97.93	20	II

Notes: *II: Institute of Engineering from UNAM (http://aplicaciones.iingen.unam.mx/AcelerogramasRSM/Consultas/Filtro.aspx); IG: Institute of Geophysics from UNAM (http://www2.ssn.unam.mx:8080/catalogo/); CENAPRED: National Center for Disaster Prevention (http://geografica.cenapred.unam.mx:8080/reporteSismosGobMX/BuscarAcelerograma); CIRES: Instrumentation and Seismic Record Center (http://www.cires.org.mx/racm_historico_es.php); RIIS: Interuniversity Network of Seismic Instrumentation; FICA: ICA Foundation; GIEC: Experimental Engineering and Control Management; UAP: Autonomous University of Puebla. ** Some of the records used were compiled by García (2005) and García et al. (2006). ¥Interplate event according to Franco *et al.*, (2007).

Table 2. Inslab events used in training the ANN models.

Event No.	No. of Rec.	Date (dd/mm/yy)	M_w	Lat. °N	Long. °W	H (km)	Institution*
1¥	10	05/08/93	5.1	17.08	98.53	32	CENAPRED, II**, IG**
2¥	16	23/02/94	5.4	17.82	97.30	5	II**, IG**, CENAPRED
3¥	166	23/05/94	5.6	18.03	100.57	23	CENAPRED, CIRES, RIIS, II**, IG**
4¥	130	10/12/94	6.4	18.02	101.56	20	CIRES, II**, IG**
5¥	124	11/01/97	6.9	17.9	103.00	16	CIRES, II**, IG**
6¥	4	03/04/97	5.1	17.98	98.33	30	II**, IG**
7	144	22/05/97	6.0	18.41	101.81	59	CENAPRED, CIRES, RIIS, II**, IG**
8	38	20/04/98	5.1	18.37	101.21	66	CENAPRED, CIRES, RIIS, II**, IG**
9	254	15/06/99	6.5	18.18	97.51	69	CENAPRED, CIRES, RIIS, II**, IG**
10	172	21/06/99	5.8	17.99	101.72	54	CENAPRED, CIRES, II**, IG**
11	202	13/04/07	6.3	17.37	100.14	42	CIRES, II**, IG**
12	88	12/02/08	6.5	16.35	94.51	87	CIRES, II**, IG**
13	52	22/05/09	5.7	18.13	98.44	45	II**, IG**
14	40	21/07/00	5.4	18.09	98.97	48	II**, IG**
15	80	07/04/11	6.7	17.2	94.34	167	CIRES, II
16	210	11/12/11	6.5	17.89	99.84	58	CIRES, II, UAP
18	156	15/11/12	6.1	18.17	100.52	40	CIRES, II, UAP
19	50	29/07/14	6.4	17.7	95.63	117	CIRES, II
20	148	19/09/17	7.1	18.4	98.72	57	CIRES, II, UAP
21	68	23/09/17	6.1	16.48	94.90	75	CIRES

Notes: *II: Institute of Engineering from UNAM (http://aplicaciones.iingen.unam.mx/AcelerogramasRSM/Consultas/Filtro.aspx); IG: Institute of Geophysics from UNAM (http://www2.ssn.unam.mx:8080/catalogo/); CENAPRED: National Center for Disaster Prevention (http://geografica.cenapred.unam.mx:8080/reporteSismosGobMX/BuscarAcelerograma); CIRES: Instrumentation and Seismic Record Center (http://www.cires.org.mx/racm_historico_es.php); RIIS: Interuniversity Network of Seismic Instrumentation; FICA: ICA Foundation; GIEC: Experimental Engineering and Control Management; UAP: Autonomous University of Puebla. **Some of the records used were compiled by García (2005) and García et al., (2006). ¥Cortical events Lowry *et al.*, (2001).
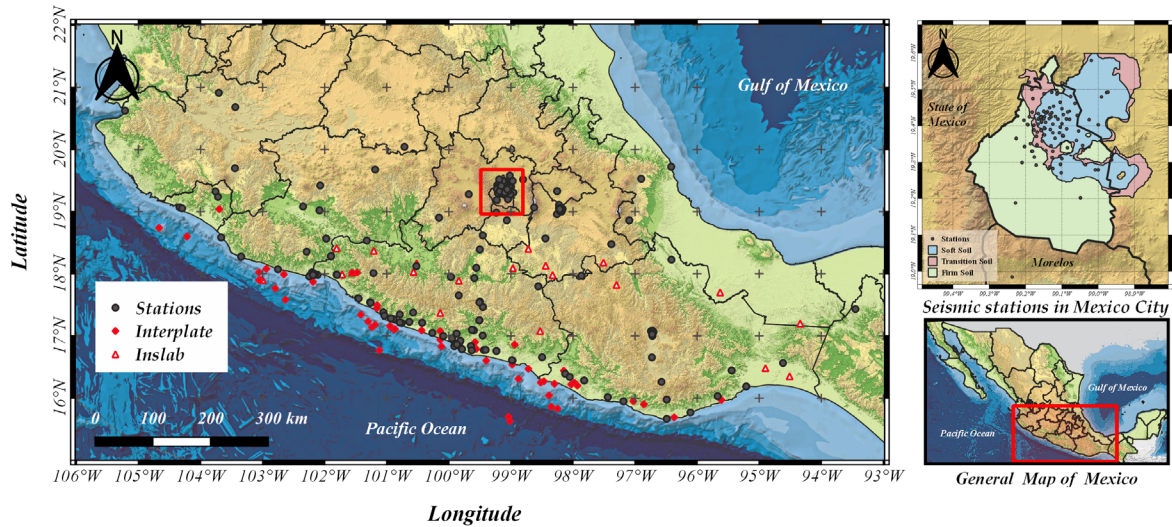
Figure 1. Location of events and recording stations employed.

The total number of records from interplate and inslab events were organized into three groups. The first group corresponds to seismic records registered at soft soil of Mexico City (SS-MC), the second group corresponds to seismic records registered at firm soil of Mexico City (FS-MC), and the third group includes the rest of records registered outside Mexico City at firm soil (i.e., rock) (FS-M). For the development of the ANN models and the empirical equations, a database with all horizontal components of motion without combining them was considered. Figure 3 presents a plot with the percentage of seismic records from interplate and inslab earthquakes used in each group as well as the type of soil. It is observed from Figure 3 that the distribution of records per group is similar between the interplate and inslab earthquakes, and that the records for sites with soft soil in Mexico City have the greatest percentage.

Data Processing

A program was developed in MATLAB (2019), to build a database of seismic records for each type of earthquake. The information considered for each record included the name of the file, station name, station coordinates, magnitude of the event, epicentral coordinates, date and hour of the seismic event, orientation of each sensor channel, peak ground acceleration and maximum seudoaceleration for each component for the soil period, strong ground motion duration and the acceleration values for each component. Bad quality records were discarded. A baseline correction of all the time histories of the records was carried out. Further, for events registered at firm soil (including rock) with Mw >6.5, a high-pass filter with a cut-off frequency of 0.05 Hz was used, for the rest events, a high-pass corner frequency 0.1 Hz was employed. This processing criterion was guided by the work of García *et al.* (2005) and García (2006). For events registered at soft soil, a band-pass filter with corner frequencies from 0.1 to 10 Hz was employed (Jaimes *et al.*, 2015).

SGMD Calculated from the Processed Data

The study of the SGMD has been carried out by several researchers in the past (Trifunac and Brady *et al.*, 1975; Trifunac and Westermo *et al.*, 1982). In this study the SGMD was calculated based on
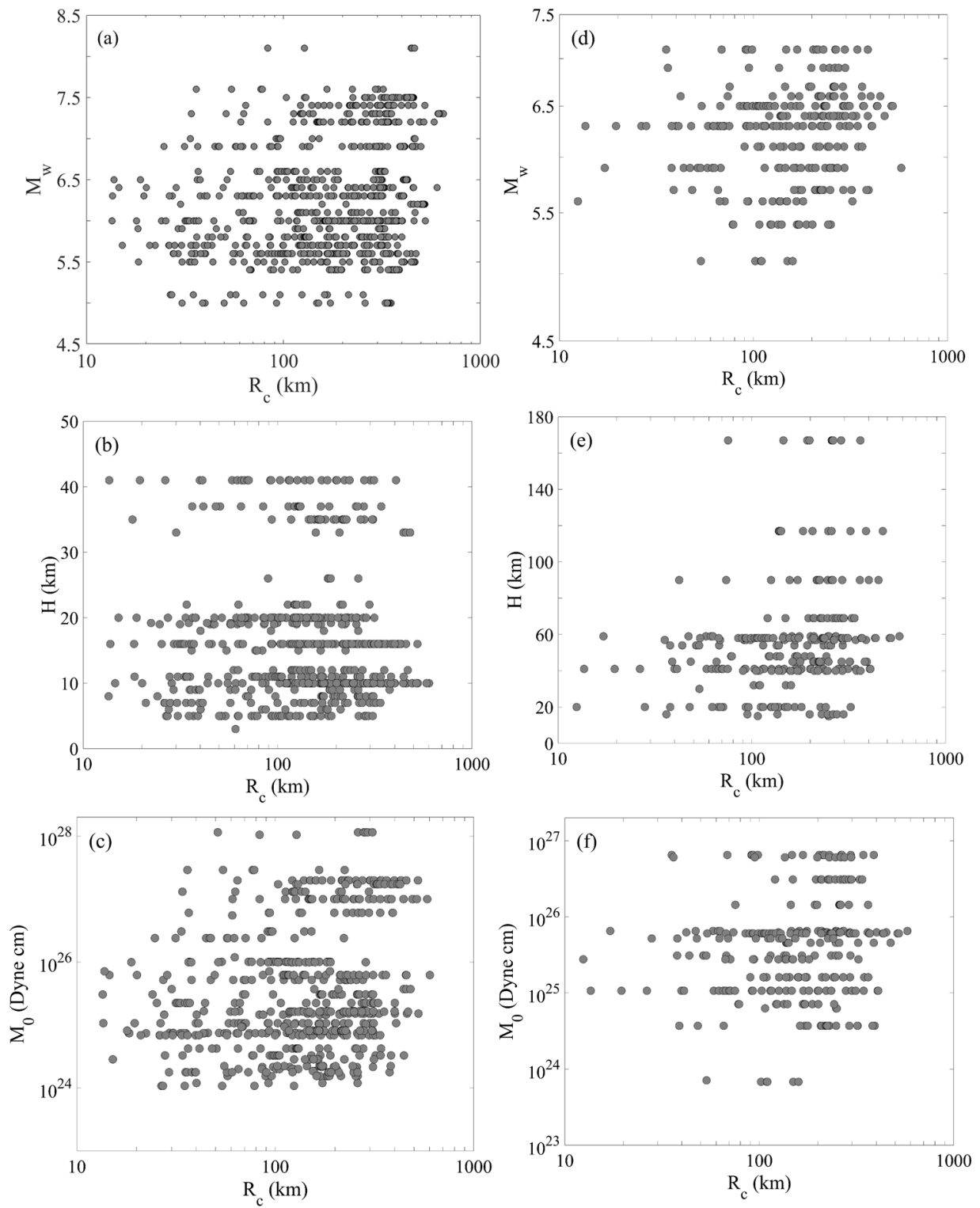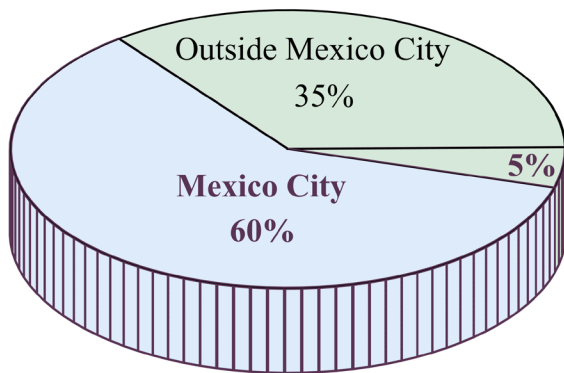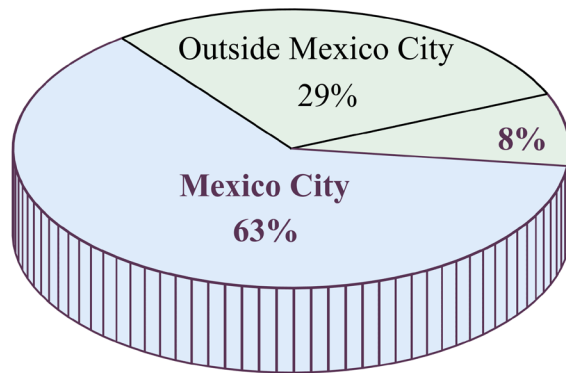
Figure 2. Distribution of Mw, H and M_0 with respect to R_c: (a), (b) and (c) for interplate events; (d), (e) and (f) for inslab events.

(a) Interplate events

(b) Inslab events

Figure 3. Distribution of percentage of seismic records at soft and firm soils.

the accumulation of energy along time of the strong ground motion records. Arias intensity (Arias, 1970) has been widely used to relate the SGMD with the acceleration time history energy, although it has also been used to study the damage patterns and principal direction of seismic excitations (Arias, 1996; Hong and Goda, 2007; and Hong *et al.*, 2009). In this study the SGMD is assessed based on the Arias intensity, defined as

$$I_A = \frac{2\pi}{g} \int_0^{t_0} a^2(t)dt , \tag{1}$$

where $a(t)$ is the acceleration time history, t_0 is the total duration of the strong ground motion and g is the acceleration due to gravity. Several procedures have been reported in the literature to determine the SGMD (Bommer and Martínez, 1999), based on lower and upper bounds of duration related to I_A. According to Reinoso and Ordaz (2001), SGMD for Mexican earthquakes can be obtained based on the duration of the strong ground motion between 2.5 and 97.5% of I_A, which is useful for engineering problems. These limits are adopted in the present study, since the records employed in the databases also include most of those used in Reinoso and Ordaz (2001).

ARTIFICIAL NEURAL NETWORK MODELING AND TRAINING

The ANN architecture with multiple hidden layers and neurons used in this study is shown in Figure 4, where three main layers can be identified: input, hidden and output layer. The flow of information starts from the input layer, this information is weighted to optimize the mapping between the input and the hidden layer(s), and finally transferred it into output value(s). The information transferred from the hidden layer(s) to the output layer is affected by biases that modified the output of the neuron. If an ANN model with a single output neuron and two hidden layers is considered, the mathematical expressions that relate the output neuron in the output layer with the neurons in input and hidden layers are given by,
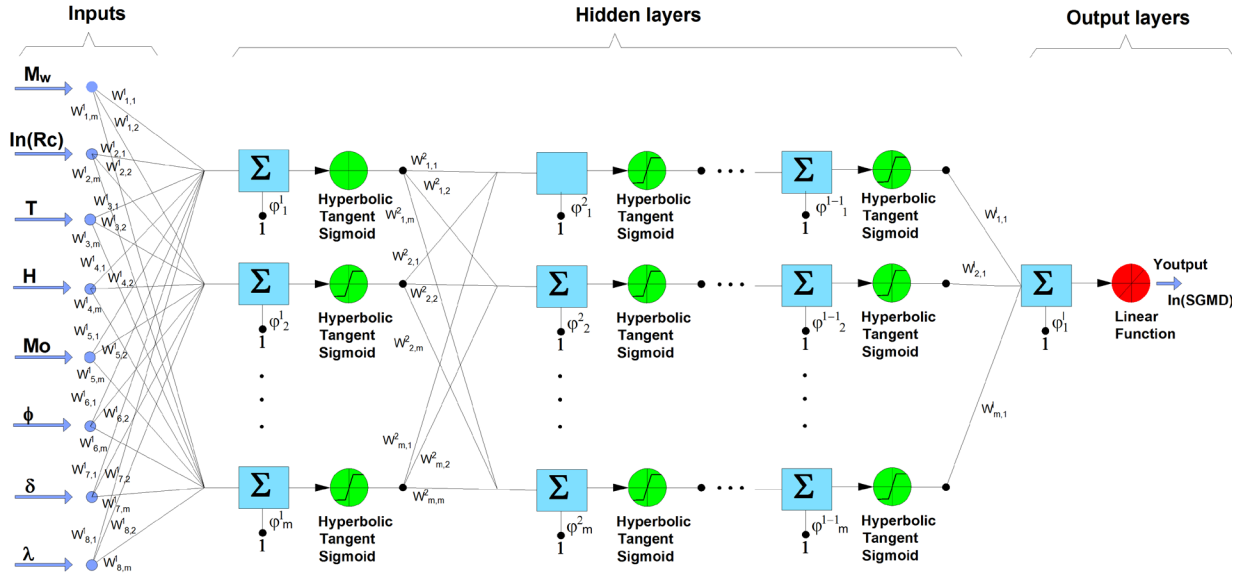
Figure 4. ANN architecture with multiple hidden layers and neurons.

$$y_{output} = f_3\left[\sum_{k=1}^{m}[w_3]_{k,1}\,O_{HL1-HL2} + (\phi_3)_1\right], \tag{2a}$$

$$O_{HL1-HL2} = f_2\left[\sum_{j=1}^{m}\left([w_2]_{j,k}\,O_{HL1-HL2} + (\phi_2)_k\right)\right], \tag{2b}$$

$$O_{IL-HL1} = f_1\left[\sum_{i=1}^{n}\left([w_1]_{i,j}\,x_i + (\phi_1)_j\right)\right], \tag{2c}$$

where y_{output} is the value of the output neuron, O_{IL-HL1} is the outcome obtained when the information given in the input layer has passed through the first hidden layer, $O_{HL1-HL2}$ is the outcome obtained when the information from the first hidden layer has passed through the second hidden layer, m is the total number of neurons in the hidden layers, n is the total number of neurons in the input layer, $f_1(\,)$, $f_2(\,)$ and $f_3(\,)$ are activation functions between the input and the first hidden layer, the first and the second hidden layer and between the second hidden layer and the output layer, respectively; $[w_1]_{i,j}$, $[w_2]_{j,k}$ and $[w_3]_{k,1}$ are the weights that map the information between the input and the first hidden layer, between the first and second hidden layer and between the second hidden layer and output layer, respectively; $(\phi_1)_j$, $(\phi_2)_k$ and $(\phi_3)_1$ are the biases associated with the hidden and output layers and x_i is the i-th neuron in the input layer.

The optimization of the weights and biases of the ANN model is carried out during the training process. Although a variety of algorithms is available in the literature (Swingler, 1996; Principe and Euliano, 1999; and Haykin, 1999), the back-propagation algorithm (Rumelhart *et al.*, 1986) is one of the most popular. The back-propagation algorithm employs a predefined error function, which is minimized to evaluate the weights and biases. The back-propagation algorithm is adopted in the present study to train the ANN models. Aside from the back-propagation algorithm, in the last

decade the constant improvement of Machine Learning techniques has allowed the development of deep learning that employs the deep neural network. The use of deep neural networks has gained much attention due to their ability to solve complex problems. The use of deep neural networks is outside of the scope of the present study.

ESTIMATION OF STRONG GROUND MOTION DURATION USING ANN

APPLICATION OF PRINCIPAL COMPONENT ANALYSIS TO IDENTIFY THE INPUTS OF THE ANNS MODELS

Dimensionality reduction methods have been widely used to reduce the number of input parameters to develop ANN models (Yuce *et al.*, 2014). One of the classical dimensionality reduction methods is the principal component analysis (PCA), which transforms a set of observations of correlated variables into a set of principal components, which are linearly uncorrelated. Based on the identified principal components, the amount of total variance contributed by each component is assed to select a reduced number of principal components which cumulative variance is within predefined acceptable values. Once a reduced number of principal components is selected, the relative importance of each input parameter to a particular component is evaluated by using correlation coefficients that relate the reduced set of principal components and input parameters. The reduced number of input parameters is selected based on predefined thresholds of correlation coefficients.

To illustrate the use of PCA in the dimensionality reduction of the input neurons for the development of the ANN models, we use the data for inslab events for firm soil for places outside Mexico. To proceed with the calculation of the principal components, the correlation coefficient matrix between the input variables is given in Table 3.

The correlation coefficient matrix is then decomposed by using the singular value decomposition to calculate the amount of total variance contributed by each principal component. Table 4 summarizes the percentage of variance associated with each principal component and its corresponding eigenvalue.

It is noted that there are different criteria to select the reduced number of principal components. According to Lovric (2011), the reduced number of principal components can be selected when they account for a cumulative variance within 70 to 90%. A simpler criterion, which is adopted in this study, is to select those principal components whose eigenvalues are greater than one. Based on the afore-mentioned criterion, from Table 4, the reduced number of principal components is equal to 4.

Table 3. Correlation coefficient matrix for input variables for inslab events for firm soil for places outside Mexico.

	R_C	M_w	T	H	M_0	ϕ	δ	λ
R_C	1	0.220	0.347	0.242	0.124	0.215	-0.028	-0.101
M_w	0.220	1	0.087	0.397	0.659	0.123	0.384	-0.002
T	0.347	0.087	1	0.119	0.033	0.099	-0.056	-0.015
H	0.242	0.397	0.119	1	0.046	0.457	0.200	-0.284
M_0	0.124	0.659	0.033	0.046	1	0.160	0.124	-0.210
ϕ	0.215	0.123	0.099	0.457	0.160	1	-0.308	-0.053
δ	-0.028	0.384	-0.056	0.200	0.124	-0.308	1	0.326
λ	-0.101	-0.002	-0.015	-0.284	-0.210	-0.053	0.326	1

Table 4. Percentage of variance associated with each principal component and its corresponding eigenvalue.

Principal component	1	2	3	4	5	6	7	8
Eigenvalue	2.24	1.64	1.19	1.03	0.87	0.63	0.29	0.11
Variance (%)	27.99	20.46	14.84	12.90	10.89	7.82	3.66	1.41
Cumulative variance (%)	27.99	48.46	63.30	76.20	87.10	94.92	98.59	100

By using the first 4 principal components, the correlation coefficients that relate the reduced set of principal components and input parameters is presented in Table 5.

It is observed in Table 5 that there are considerable variation in the correlation coefficients. To identify the reduced input parameters, we use two thresholds of correlation coefficients (Moore *et al.*, 2013), the first one equals 0.7 and the second one equals 0.55. It is noted that strong relation between variables is associated with correlation coefficients greater than 0.7, while moderate relation is associated with correlation coefficients greater than 0.55. Table 6 presents the reduced number of input parameters when the correlation thresholds are adopted. Also in Table 6, the reduced number of input parameters for the rest of the database are summarized. It is observed from Table 6 that the number of input parameters with strong correlation for interplate events ranges from 2 to 3, while for inslab it varies from 2 to 6. For the input parameters with strong correlation, the greater number of input parameters are related to the SS-MC case. If moderate correlation is considered, the number of input parameters is within 6 to 7 and 5 to 7 for interplate and inslab events, respectively. Based on the PCA results, ANN models are developed by using the identified input parameters based on the type of relation (i.e., strong and moderate relation). Furthermore, for the sake of comparison, ANN models by using the complete set of input parameters, referred to as 'all inputs', was included. One more ANN model with predefined input parameters is considered, this model includes Mw, natural logarithmic of R_c, H and T as input neurons and is referred to as 'original case'. This last ANN model is considered as much of the information available in applications of engineering include those parameters. It should be pointed out that all the ANN models employ the natural logarithmic of R_C instead of the R_C. The output layer of the ANN models consists of a single neuron that represents the natural logarithmic of the SGMD for a considered earthquake and soil type.

The activation functions as well as the setups used during the training process are presented in Table 7. Some authors have suggested rules to identify the number of hidden neurons and their relation

Table 5. Correlation coefficients that relate the reduced set of principal components and input parameters.

Input parameter	1	2	3	4
R_C	0.531	-0.261	0.466	-0.255
M_w	0.784	0.476	-0.114	-0.069
T	0.321	-0.254	0.638	-0.414
H	0.688	-0.153	0.046	0.582
M_0	0.635	0.289	-0.437	-0.424
ϕ	0.495	-0.511	-0.017	0.420
δ	0.193	0.823	0.242	0.225
λ	-0.284	0.483	0.546	0.211

Table 6. Reduced number of input parameters after applying the PCA method.

Earthquake Type	Case	Type of relation	
		Strong	Moderate
Interplate	SS-MC	M_w, R_C, δ	M_w, R_C, T, M_0, δ, λ
	FS-MC	M_w, M_0	M_w, R_C, H, M_0, δ, ϕ
	FS-M	M_w, λ	M_w, R_C, H, T, M_0, δ, λ
Inslab	SS-MC	M_w, R_C, H, M_0, δ, λ	M_w, R_C, T, M_0, δ, λ, ϕ
	FS-MC	M_w, δ	M_w, R_C, H, M_0, δ, λ, ϕ
	FS-M	M_w, δ	M_w, H, T, M_0, δ

with the hidden layers (Masters, 1993; Swingler, 1996; Berry and Linoff, 1997), yet no closed-form expression is available to indicate the total number of hidden layers and neurons, a trial and error scheme is adopted to determine the structure of the ANN model (Shahin *et al.*, 2004).

During the trial and error process, ANN models with one and two hidden layers with 3 and up to 50 hidden neurons in each hidden layer were considered. The former was carried out to avoid potential overfitting of the ANN model.

It is noted that different metrics as the mean square error (MSE) or the mean absolute error (MAE) are available to evaluate the performance of the ANN models, each metric has advantages and disadvantages depending of the problem at hand. According to Twomey and Smith (1995), there is no consensus as to which measure should be reported, and thus, comparisons among techniques and results of different researchers are practically impossible. In this study we adopted the MSE to evaluate the performance of the ANN models.

To evaluate the impact of the number of hidden neurons and layers on the trained ANN models by using both: the samples used for training and those not employed for training, the average MSE based on 300 trials (Pozos-Estrada *et al.*, 2014), is presented in Figure 5. It is observed in Figure 5 that, in general, when the data used for training is employed, the average MSE decreases as the number of hidden neurons increases. It is also observed from Figure 5 that the average MSE when the samples used for testing are employed is greater than that observed when the samples used for training are considered. This can be explained by noting that the trained ANN models are tested with input parameters that are different than those used during the training process. This trend is

Table 7. Activation functions and training setups.

Layer	Activation function
Input to hidden layer	Tan-Sigmoid
Hidden to hidden layer	
Hidden to output layer	Linear
Training and testing	$f(x) = \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$
Training data	80% of the complete database (randomly selected)
Testing data	20% of the complete database not selected for training
Error function	Mean square error (MSE)
Minimization algorithm	Levenberg-Marquardt (Marquardt (1963)), Press et al. (1992))

observed for the ANN models with different number of input neurons (i.e., all inputs, strong relation, moderate relation and original case) with one and two hidden layers, irrespective of the type of earthquake or soil considered. It is also observed from Figure 5 that the lowest average MSE associated with the testing stage is obtained for ANN models with 3 to 50 hidden neurons, and that the best ANN models for interplate and inslab are associated with 1HL and 2HL, respectively. Based on
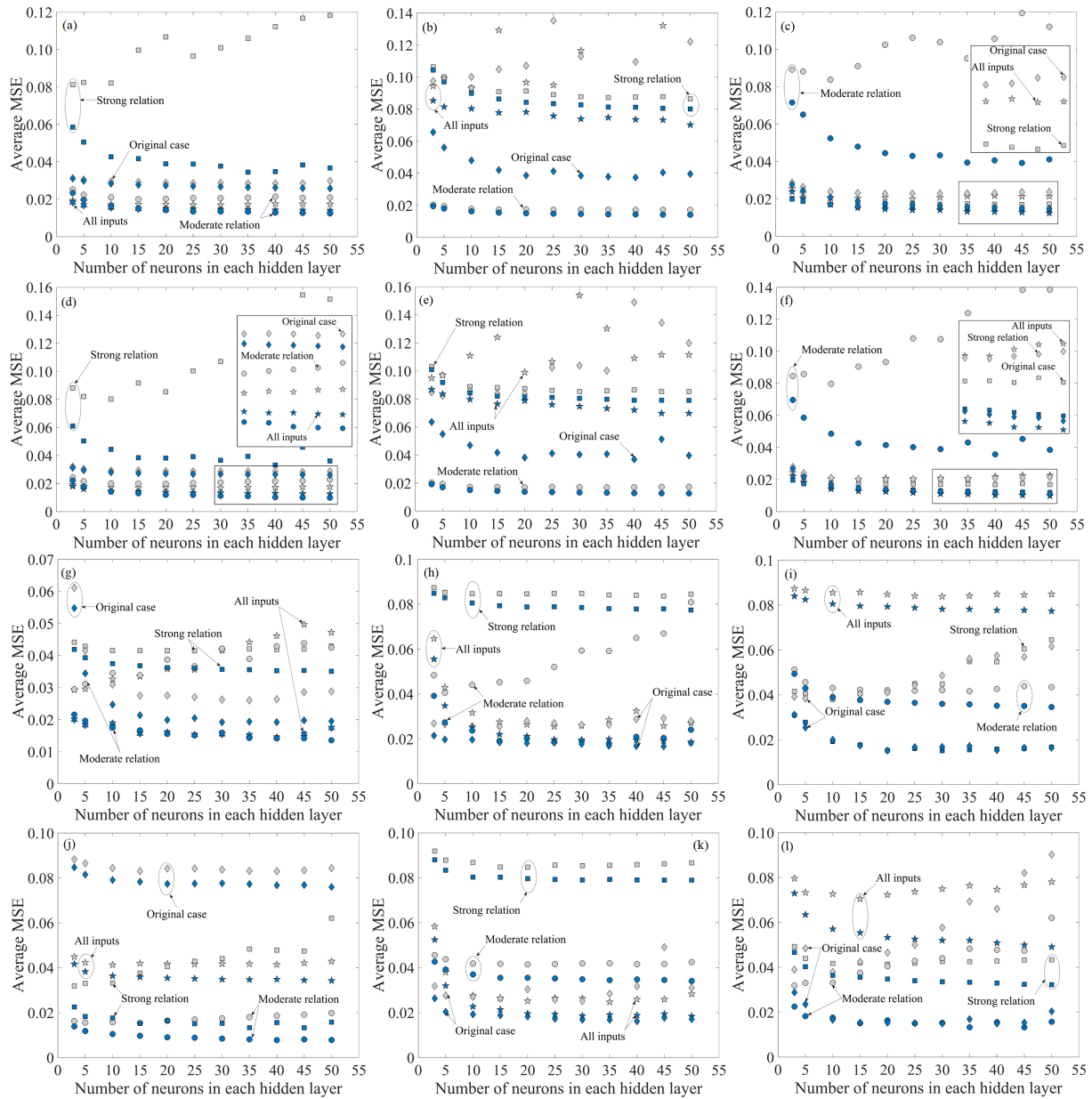

Figure 5. Average MSE for the developed ANN models by using the samples employed for training (in blue) and testing (in gray). Interplate events with one hidden layer: (a) Mexico City soft soil, (b) Mexico City firm soil, (c) Outside Mexico City firm soil. Interplate events with two hidden layers: (d) Mexico City soft soil, (e) Mexico City firm soil, (f) Outside Mexico City firm soil. Inslab events with one hidden layer: (g) Mexico City soft soil, (h) Mexico City firm soil, (i) Outside Mexico City firm soil. Inslab events with two hidden layers: (j) Mexico City soft soil, (k) Mexico City firm soil, (i) Outside Mexico City firm soil.

these observations, the parameters of the optimum ANN models (those associated with the minimum average MSE) are summarized in Table 8. It is observed in Table 8 that the number of hidden layers for the cases with different number of input neurons (i.e., all inputs, strong relation, moderate relation and original case) for interplate events ranges from 3 to 30, while that for inslab ranges from 3 to 50. This observation indicates that the selection of the inputs neurons is of paramount importance. Based on the identified optimum models with different number of input neurons, a further selection was carried out to employ a single ANN model (i.e., a single ANN model among all inputs, strong relation, moderate relation and original case) to predict the SGMD for SS-MC, FS-MC and FS-M. The criterion used to identify a single ANN model was based on selecting the ANN model which exhibits the best prediction behavior under several case scenarios. The ANN models selected to predict the SGMD for SS-MC, FS-MC and FS-M are indicated in Table 8 in italics. The obtained weights and biases for the selected trained ANN models for interplate and inslab, summarized in Table 8, are presented in Appendix A.

Comparison Between the Observed and the Predicted SGMD by Using Trained ANN

Figure 6 presents a comparison between the observed and the predicted SGMD by using the trained ANN models with the datasets used for training and testing to those obtained from the actual records. It is observed from Figure 6 that there is good agreement between the observed and predicted values in most of the cases, with a correlation coefficient, ρ, ranging from 0.50 to 0.92 when the dataset used for testing is considered, and from 0.74 to 0.96 when the dataset used for training is employed. It is also observed from Figure 6 that the highest ρ values are associated the ANN models for inslab.

Table 8. Summary of the minimum MSE of ANN models for interplate and inslab events.

Soil type	Case	Interplate		Inslab	
		MSE	ANN model	MSE	ANN model
SS-MC	All inputs	0.013	*3N-1HL*	0.032	*15N-2HL*
	Strong relation	0.024	5N-1HL	0.065	15N-2HL
	Moderate relation	0.013	30N-1HL	0.032	45N-2HL
	Original case	0.014	5N-1HL	0.031	50N-2HL
		0.014	3N-1HL*	0.031	30N-2HL*
FS-MC	All inputs	0.028	*3N-1HL*	0.010	50N-2HL
		-	-	0.032	15N-2HL*
	Strong relation	0.034	20N-1HL	0.011	5N-2HL
	Moderate relation	0.020	3N-1HL	0.011	3N-2HL
	Original case	0.036	5N-1HL*	0.015	*3N-2HL*
FS-M	All inputs	0.013	3N-1HL	0.010	*5N-2HL*
		-	-	0.009	10N-2HL*
	Strong relation	0.064	3N-1HL	0.054	30N-2HL
	Moderate relation	0.014	3N-1HL	0.046	15N-2HL
	Original case	0.017	*20N-1HL*	0.016	30N-2HL

Notes: SS-MC = Soft soil Mexico City; FS-MC = Firm soil Mexico City; FS-M = Firm soil for places outside Mexico City; N = Neurons; HN = Hidden neurons; * ANN models employed in Figure 9.
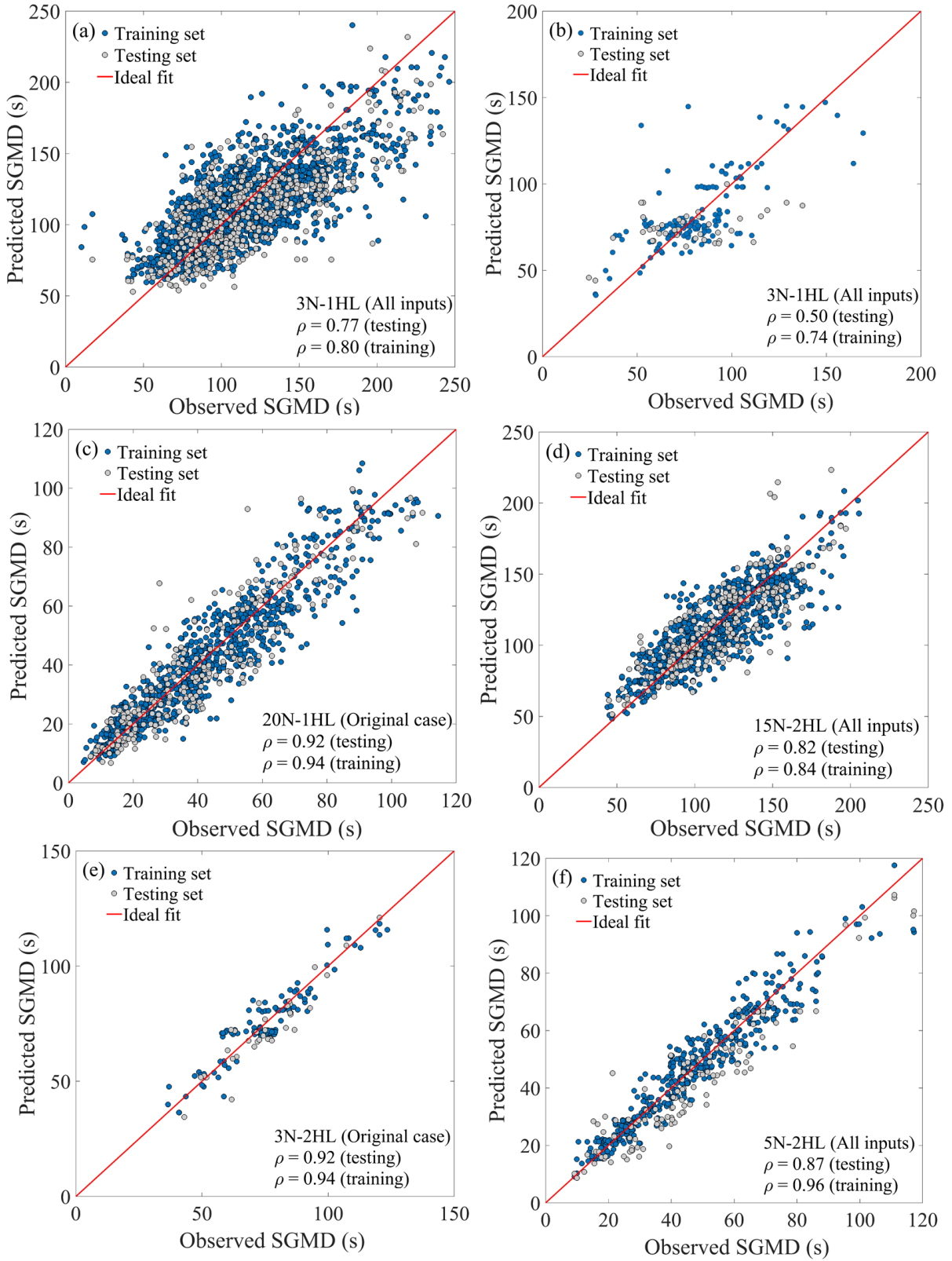
Figure 6. Comparison of the predicted SGMD with the trained ANN models and the observed values. Interplate events: (a) Mexico City soft soil, (b) Mexico City firm soil, (c) Outside Mexico City hard soil. Inslab events: (d) Mexico City soft soil, (e) Mexico City firm soil, (f) Outside Mexico City firm soil.

Among all the ANN models developed, those that predict the SGMD for sites outside Mexico City present the highest ρ values for interplate and inslab events. Further, the ANN model developed for soft soil in Mexico City for interplate events is similar to the ANN model for inslab events. On the other hand, the ANN model developed for hard soil in Mexico City for inslab events showed better predicting results that the ANN model for interplate events.

Comparison of the Predicted SGMD Using Trained ANN and Empirical Equations

Empirical Models to Estimate SGMD for Interplate and Inslab Events

The functional form of the empirical equation adopted in this study for soft and hard soils is the one proposed in Reinoso and Ordaz (2001), which is shown in Table 9, where D denotes the SGMD, in s, M_w is the moment of magnitude, R_C is the closest distance to the fault, in km, and T is the soil period, in s. In this study the model coefficients were estimated using a least-squares regression algorithm, implemented in MATLAB (2019). It is noted that that more accurate regression methods are available in the literature (Joyner and Boore, 1993; Boore *et al.*, 1997; Bommer *et al.*, 2010) and that the use of the traditional least-squares method is an oversimplification; however, for simplicity, and since the main focus of this work is on the development and discussion of the ANN models to predict the SGMD, we adopted the traditional least-squares method.

The estimated coefficients for each type of earthquake and soil are summarized in Table 9. Figure 7 presents a comparison of the predicted SGMD with the fitted empirical models and the observed values for interplate and inslab events for soft and firm soils. Similar conclusions to those draw from Figure 6 are applicable to Figure 7, except that the correlation coefficients between the observed and predicted values are in general slightly smaller than those obtained when the ANN models are used. This difference is related to the number of data used to fit the empirical models, which is greater than that employed to test the ANN models.

Comparison Between Predicted SGMD Using Trained ANN and Empirical Equations

Figure 8 presents a comparison of the variation of the SGMD predicted with the trained ANN models and that obtained with the empirical equations as a function of M_w and modal values of R_c, T, H,

Table 9. Coefficients of the empirical equation.

$D = c_1 \exp(M_w) + (c_2 M_w + c_3)R_C + (c_4 M_w + c_5)(T + c_6) + \varepsilon$				
Interplate events				
Case	c_1	c_2	c_3	c_4	c_5
SS-MC	0.0237	-0.0212	0.3063	6.345	-25.013
FS-MC	0.0332	0.0035	0.1528	-	-
FS-M	0.0160	-0.0090	0.2361	-	-
Inslab events				
Case	c_1	c_2	c_3	c_4	c_5
SS-MC	0.0684	-0.0852	0.6722	-2.6447	38.11
FS-MC	0.0501	-0.0931	0.764	-	-
FS-M	0.027	-0.0233	0.3278	-	-

Notes: c_i, i = 1,2,3,4,5, are regression coefficients; c_6 = 0.5; T is set equal to 0.5 s for firm soils (Reinoso E. Ordaz M (2001), NTC-SISMO (2017)); ε is the error term.
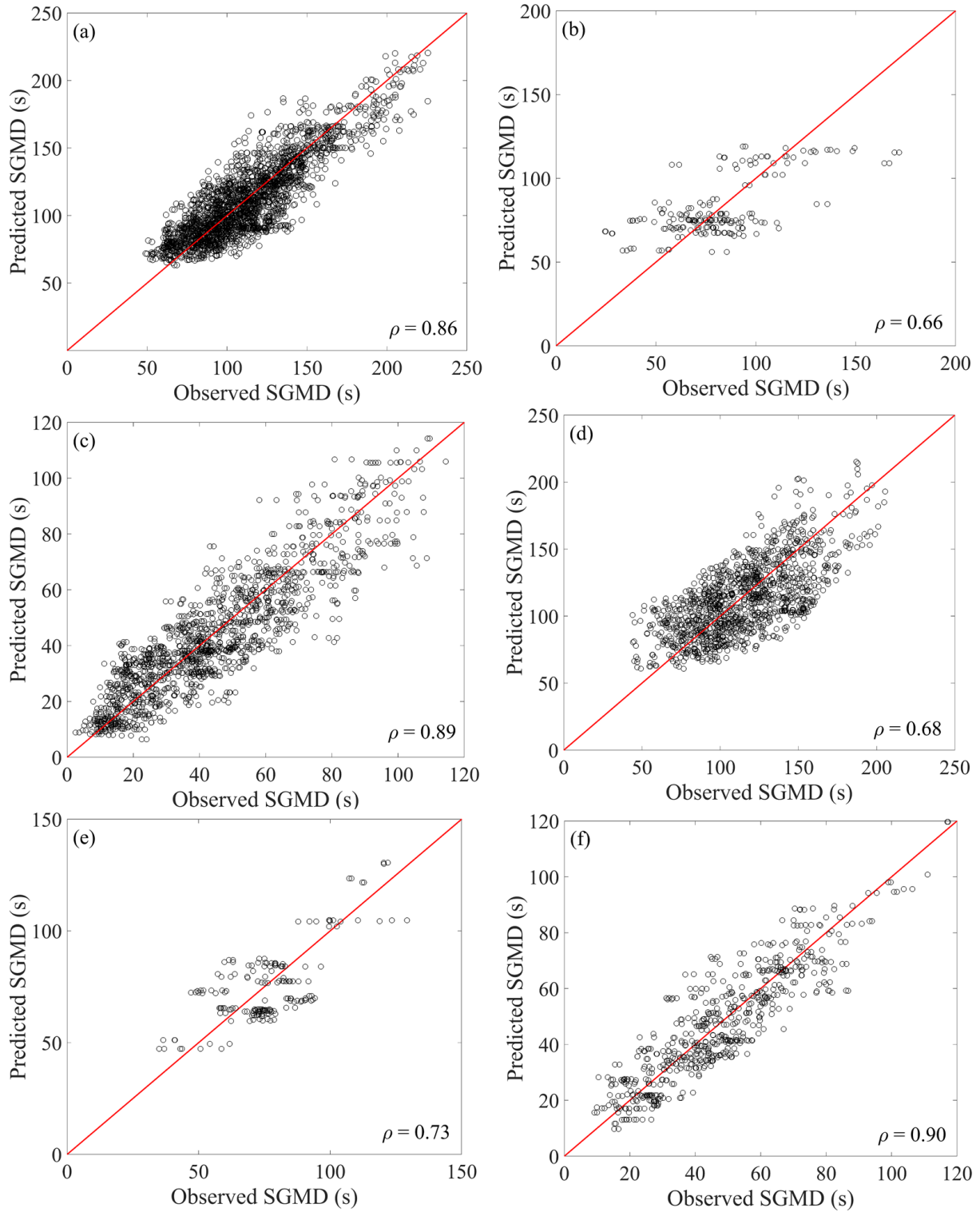
Figure 7. Comparison of the predicted SGMD with the fitted empirical models and the observed values. Interplate events: (a) Mexico City soft soil, (b) Mexico City firm soil, (c) Outside Mexico City firm soil. Inslab events: (d) Mexico City soft soil, (e) Mexico City firm soil, (f) Outside Mexico City firm soil.

M_0, ϕ, δ and λ for soft soil and modal values of Rc, H, M0, ϕ, δ and λ and T = 0.5 (s) for firm soil. It is noted that the use of H, M_0, ϕ, δ and λ are input parameters of the trained ANN models; however, they were not considered as input parameter in the empirical expressions. In Figure 8, the empirical model developed in Reinoso and Ordaz (2001), is also included for comparison purposes. It is observed from Figure 8, that in general, the predicted values by using the ANN models follow those predicted by the developed empirical equations. It is also observed that, for interplate events, the SGMD predicted with the trained ANN models oscillates within those predicted with the empirical equations. It is further observed that, for inslab events, the SGMD predicted with the trained ANN models follows the trend of the SGMD predicted with the empirical equation developed in this study. For case 8c that corresponds to the SGMD for inslab events for firm soil for places outside Mexico, the SGMD predicted by using the ANN model presents a bump for M_W within 5.5 and 6.5. The rest of the cases showed a smoother behavior. In practically all the cases presented in Figure 8, the model proposed in Reinoso and Ordaz (2001), predicted smaller SGMD values with respect to those predicted by the trained ANN models and the empirical equations developed in this study. The differences between the SGMD values predicted by the ANN models and those with the empirical equation given in Reinoso and Ordaz (2001), was also observed in the study by Alcántara *et al.* (2014).

Figure 9 presents a comparison of the variation of the SGMD predicted with the trained ANN models and that obtained with the empirical equations as a function of R_C and selected values of Mw and modal values of T, H, M_0, ϕ, δ and λ. Similar conclusions to those drawn for Figure 8 are applicable to Figure 9, except that a better behavior of the predictions made with the ANN models are observed for both inteplate and inslab events. It is noted that the ANN models used to predict the SGMD for Figure 8, were also employed for Figure 9; however, it was observed that better predictions could be obtained if ANN models with different number of neurons were used. The ANN models employed in Figure 9 are also reported in Table 8. The previous observation indicates that the trained ANN models for the considered records are not very robust because the trained models with almost identical mean square errors do not always lead to the same predicted SGMD.

To further compare the trained ANN models and the empirical equations, a probabilistic characterization of the error, defined as the difference of the logarithmic of the predicted values of SGMD by using the trained ANN models or the empirical equations for the results presented in Figure 8, and the logarithmic of the observed values is carried out. Figure 10 presents a comparison of the calculated errors in Normal probability paper. It is observed from Figure 10 that the error could be modeled as a normal variate. This was also verified with the Kolmogorov-Smirnov goodness-of-fit test (Benjamin and Cornell, 1970), which indicates that the normality hypothesis could not be rejected at a significance level of at least 1% for both type of seismic events. The mean and standard deviation of the errors presented in Figure 10 are summarized in Table 10. It is observed from Table 10 that the trained ANN models and the empirical equations are only slightly biased and that the statistics of the error for the developed ANN models are similar to those of the empirical equations. Similar results were observed for the ANN models and the empirical equations presented in Figure 9 and for that reason they are not shown.
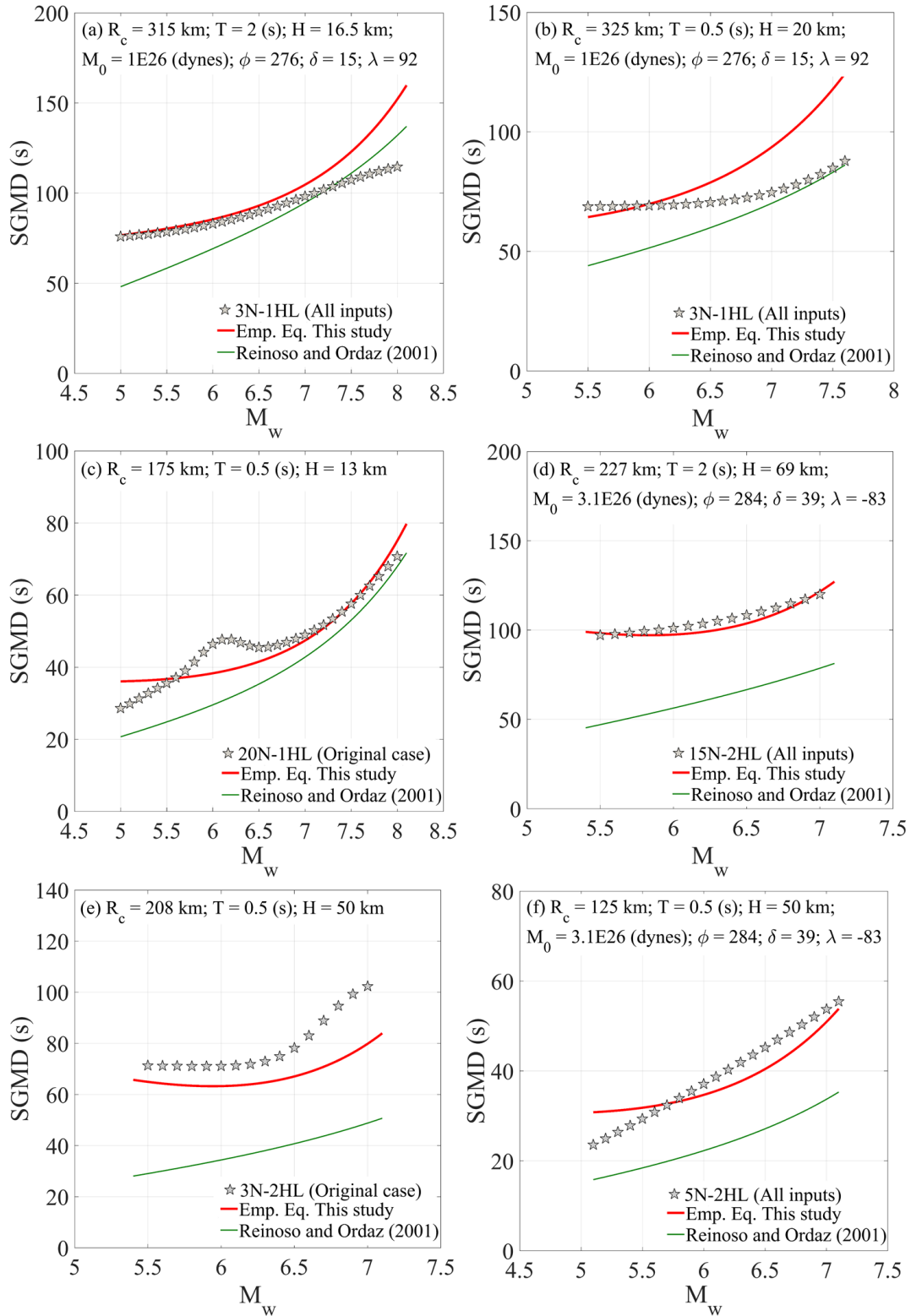
Figure 8. Comparison of SGMD predicted by the trained ANN models and by the empirical models with variation in Mw. Interplate events: (a) Mexico City soft soil, (b) Mexico City firm soil, (c) Outside Mexico City firm soil. Inslab events: (d) Mexico City soft soil, (e) Mexico City firm soil, (f) Outside Mexico City firm soil.
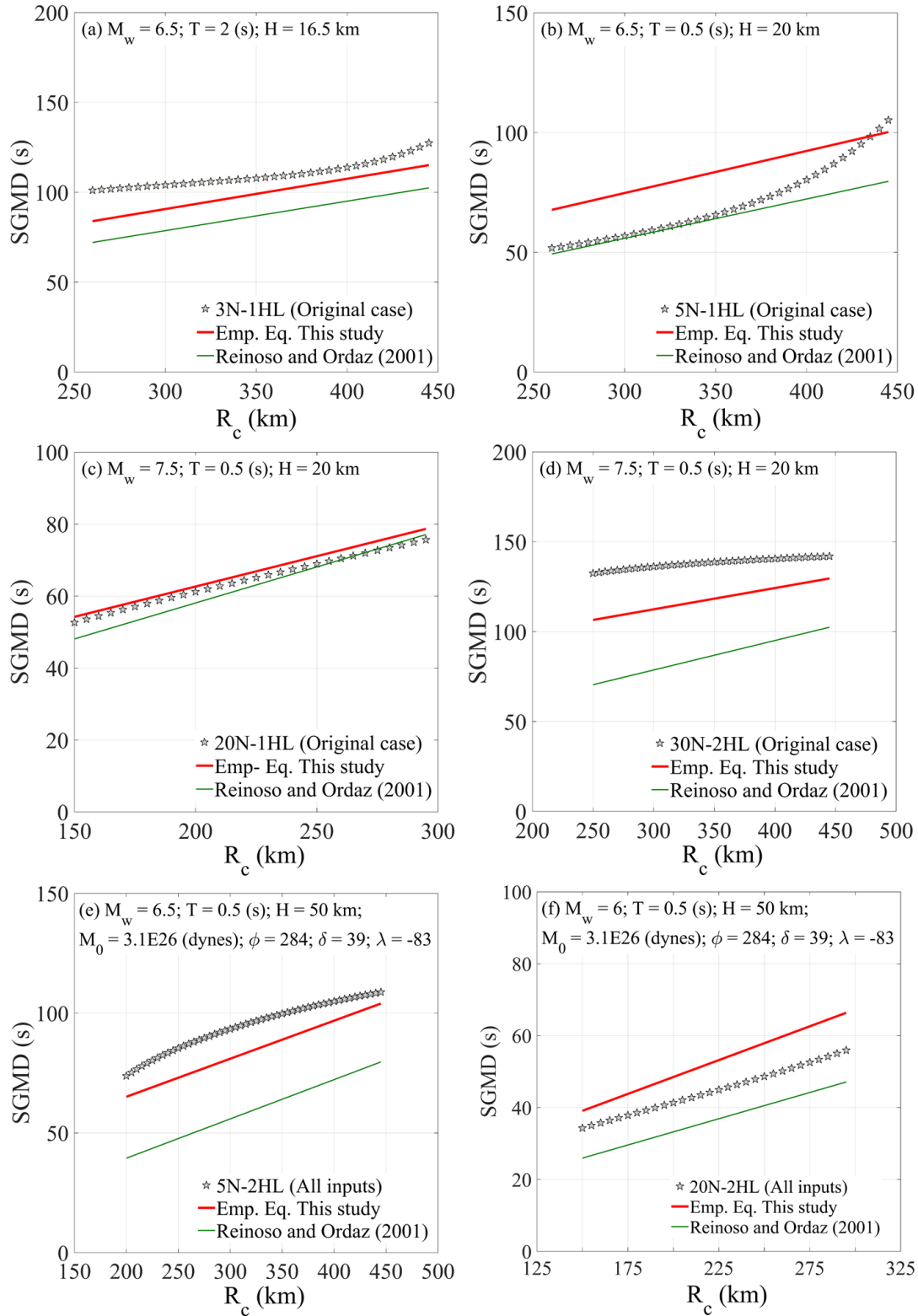
Figure 9. Comparison of SGMD predicted by the trained ANN models and by the empirical models with variation in R_C. Interplate events: (a) Mexico City soft soil, (b) Mexico City firm soil, (c) Outside Mexico City firm soil. Inslab events: (d) Mexico City soft soil, (e) Mexico City firm soil, (f) Outside Mexico City firm soil.
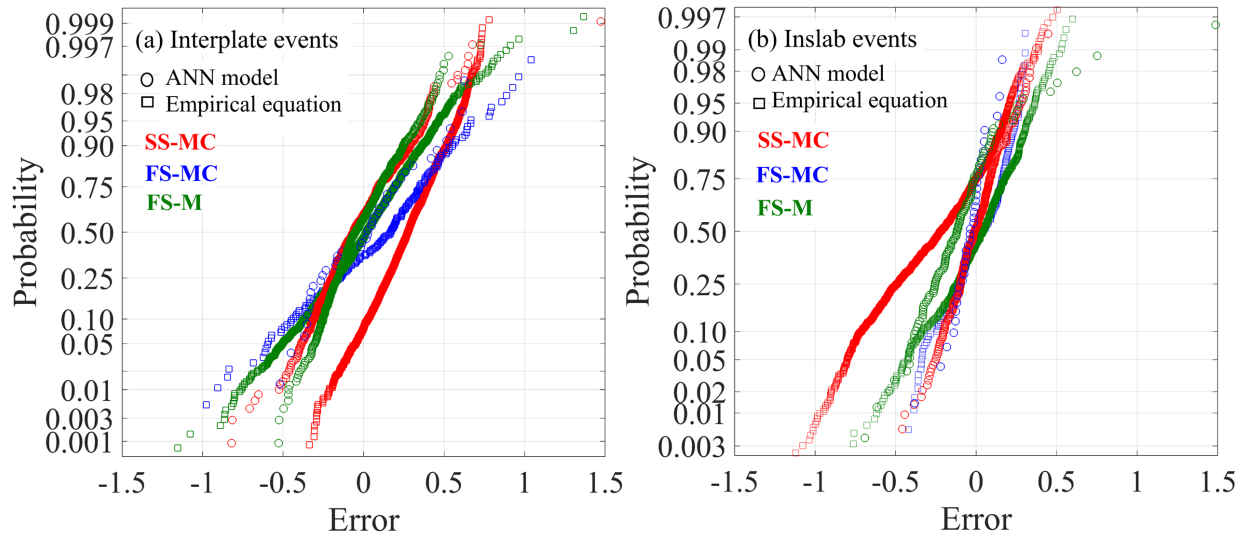
Figure 10. Normal probability paper of the error.

Table 10. Statistics of the error.

Type of earthquake	Type of soil	ANN models			Empirical equations	
		Case	Mean	Std. Dev.	Mean	Std. Dev.
Interplate	SS-MC	All inputs	-3.98E-02	0.24	2.70E-01	0.19
	FS-MC	All inputs	1.07E-02	0.31	1.01E-01	0.38
	FS-M	Original case	-2.03E-02	0.21	6.39E-03	0.3
Inslab	SS-MC	All inputs	-5.50E-03	0.16	-2.51E-01	0.32
	FS-MC	Original case	-4.14E-02	0.10	1.22E-02	0.18
	FS-M	All inputs	-1.05E-01	0.26	1.58E-02	0.24

CONCLUSIONS

Mexican records from interplate and inslab events were employed to develop Artificial Neural Network models to predict the strong ground motion duration. The principal component method was used to carry out a dimensionality reduction of the input parameters to develop the artificial neural network models. Several ANN architectures were tested. For the training of the ANN models, the input layer considered four different cases (i.e., all inputs, strong relation, moderate relation and original case), while the logarithmic of the SGMD is used to represent the output neuron.

The model tested considered up to 50 hidden neurons with one and two hidden layers. Additionally, new regression coefficients to fit empirical equations to estimate the strong ground motion duration were also obtained. The main observations that can be drawn from the analysis results are:

1. The analyses results indicated that the best prediction of the SGMD is obtained with ANN models with one hidden layer and 3, 5 and 20 hidden neurons when interplate events are considered; however, when inslab events are considered, the ANN models with two hidden layers and 3, 5, 15, 20 an 30 hidden neurons provide the best predictions.

2. The best ANN models for interplate and inslab events are associated with the cases referred to as all inputs and original case, which considered 8 and 4 input neurons, respectively. The later indicates that the selection of the inputs neurons is of paramount importance and that the ANN models could improve their prediction ability if the number of input neurons is increased.

3. The number of neurons per hidden layer of the ANN models that presented the smallest average MSE during the testing process were within 3 to 50.

4. In general, the predicted values by using the ANN models follow those predicted by the developed empirical equations. This indicates that the ANN models represent a good alternative to the empirical equations in some applications if one does not have to understand the causality to apply the ANN model.

5. In some cases, the SGMD predicted by using the ANN models presented physically unrealistic trends in its behavior. For this reason, caution is warranted when the model is extrapolated and it is recommended to carry out several verifications of the trained ANN models before using them for further engineering applications, for example the simulation of synthetic records or the evaluation of damage indices.

References

Alcántara L., Garcia S., Ovando E., Macias C., 2014, Neural estimation of strong ground motion duration. *Geofísica Internacional*, 56 (3): 221-239. https://doi.org/10.1016/S0016-7169(14)71502-8.

Arias A., 1970, A measure of earthquake intensity, Seismic Design of Nuclear Power Plants. *R. Hansen*, Editor, MIT Press, Cambridge, Mass., pp 438-483.

Arias A., 1996, Local directivity of strong ground motion. *In Proc. 11th World Conference on Earthquake Eng.*, Acapulco, Mexico, Paper No. 1240.

Arjun C.R., Kumar A., 2011, Neural network estimation of duration of strong ground motion using Japanese earthquake records. *Soil Dynamics and Earthquake Engineering*, 31, 866–872. https://doi.org/10.1016/j.soildyn.2011.01.001.

Benjamin J.R., Cornell C.A., 1970, Probability, Statistics and Decision for Civil Engineers. McGraw-Hill, New York.

Bhargavi P., Raghukanth S.T.G., 2019, Rating damage potential of ground motion records. *Earthquake Engineering and Engineering Vibration*, 18, 233-254. https://doi.org/10.1007/s11803-019-0501-1.

Bommer J.J., Martínez-Pereira A., 1999, The effective duration of earthquake strong motion. *Journal of Earthquake Engineering*, 3(2): 127–172. https://doi.org/10.1080/13632469909350343

Bommer, J.J., Douglas, J., Scherbaum, F., Cotton, F., Bungum, H., Fah. D. (2010), On the selection of ground-motion prediction equations for seismic Hazard analysis, *Seism. Res. Lett.*, 81(5) 783 – 793. https://doi.org/10.1785/gssrl.81.5.783.

Boore, D.M., Joyner, W.B., and Fumal, T.E. (1997). Equations for estimating horizontal response spectra and peak acceleration from western North America, *Seism. Res. Lett.*, 68(1) 128 – 153. https://doi.org/10.1785/gssrl.68.1.128.

BSSC, NEHRP Recommended provisions for seismic regulations for new buildings and other structures, FEMA 450, 2004 Edition.

CMT 2013, Centroid Moment Tensor Catalog, https://www.globalcmt.org/CMTsearch.html.

Cruz J., Arredondo C., Jaimez M., 2020, New source duration relationships for Mexican earthquakes: Practical application to stochastic summation methods. *Pure Appl. Geophys.* 177, 4775-4796.

Flores-Estrella H., Yussim S., Lomnitz C., 2007, Seismic response of the Mexico City Basin: A review of twenty years of research. *Natural Hazards*, 40(2): 357-372. https://doi.org/10.1007/s11069-006-0034-6.

Franco I., Cruz C., Estrada C., 2007, Reporte preliminar del sismo del 13 de abril de 2007, Guerrero, México. *Boletín de la Sociedad Geológica Mexicana*. 59(1): 71-81.

Franke K.W., Candia G., Mayoral J.M., Wood C.M., Montgomery J., Hutchinson T., Morales-Velez A., 2019, Observed building damage patterns and foundation performance in Mexico City following the 2017 M7.1 Puebla-Mexico City earthquake. *Soil Dynamics and Earthquake Engineering*, 125, 105708. https://doi.org/10.1016/j.soildyn.2019.105708

García D., 2006, Estimación de parámetros del movimiento fuerte del suelo para terremotos interplaca e intraslab en México Central. PhD Tesis, Universidad Complutense de Madrid, Madrid, España. (In Spanish).

García D., Singh S.K., Herraiz M., Ordaz M., Pacheco J.F., 2004, Inslab earthquakes of Central Mexico: Q, source spectra, and stress drop. *Bull. Seism. Soc. Am.*, 94, 3; 789-802.

García D., Singh S.K., Herraiz M., Ordaz M., Pacheco J.F., 2005, Inslab earthquakes of Central Mexico: Peak ground-motion parameters and response spectra. *Bull. Seism. Soc. Am.*, 95, 2272-2282. https://doi.org/10.1785/0120050072.

García D., Singh S.K., Herraiz M., Ordaz M., Pacheco J.F., Cruz J. H., 2009, Influence of subduction zone structure on coastal and inland attenuation in Mexico. *Geophys. J. Int.*, 179, 215-230. https://doi: 10.1111/j.1365-246X.2009.04243.x.

Garcia J. D., 2006, Estimación de parámetros del movimiento fuerte del suelo para terremotos interplaca e intraslab en México central. Facultad de ciencias físicas, Universidad complutense de Madrid, Tesis de doctorado.

García S.R., Romo M.P., Mayoral J.M., 2007, Estimation of peak ground accelerations for Mexican subduction zone earthquakes using neural networks. *Geofísica Internacional*, 46, 1: 51 – 63.

Haykin S., 1999, Neural Networks: A Comprehensive Foundation. Prentice-Hall, Englewood Cliffs, NJ, USA, 2nd edition, 156–255.

Hong H.P., Goda K., 2007, Orientation-dependent ground-motion measure for seismic-hazard assessment. *Bull. Seism. Soc. Am.*, 97(5):1525-1538. https://doi.org/10.1785/0120060194.

Hong H.P., Liu T.J., Lee C.S., 2012, Observation on the application of artificial neural network for predicting ground motion measures. *Earthq Sci.*, 25, 161–175. https://doi.org/10.1007/s11589-012-0843-5.

Hong H.P., Pozos-Estrada A., Gomez R., 2009, Orientation effect on ground motion measure for Mexican subduction earthquakes. *Earthquake Engineering and Engineering Vibration*, 8(1): 1-16. https://doi.org/10.1007/s11803-009-8155-z.

Housner G.W., 1952, Intensity of ground motion during strong earthquakes. Calif. Inst. of Tech., Earthquake Eng. Res. Lab., Pasadena. . http://resolver.caltech.edu/CaltechEERL:1952.EERL.1952.001

Jaimes M.A., Ramirez A., Reinoso E., 2015, Ground-Motion Prediction Model From Intermediate-Depth Intraslab Earthquakes at the Hill and Lake-Bed Zone of Mexico City. *Journal of Earthquake Engineering*, 19, 1260-1278. https://doi.org/10.1080/13632469.2015.1025926.

Joyner, W.B. and Boore, D.M. (1993), Methods for regression analysis of strong-motion data, *Bull. Seism. Soc. Am.*, 83(2), 469-487. https://doi.org/10.1785/BSSA0830020469.

Kamatchi P., Rajasankar J., Ramana G.V., Nagpal A.K., 2010, A neural network based methodology to predict site-specific spectral acceleration values. *Earthquake Engineering and Engineering Vibration*, 9(4): 459-472. https://doi.org/10.1007/s11803-010-0041-1

Lindt J.W., Goh G., 2004, Earthquake Duration Effect on Structural Reliability. *Journal of Structural Engineering*, 130(5): 821-826. https://doi.org/10.1061/(ASCE)0733-9445(2004)130:5(821)

Lovric, M. 2011. International Encyclopedia of Statistical Science. Springer, Berlin, Heidelberg, p.p. 636. DOIhttps://doi.org/10.1007/978-3-642-04898-2

Lowry A.R., Larson K.M., Kostoglodov V., Bilham R., 2001. Transient fault slip in Guerrero, southern Mexico. *Geophys. Res. Lett.*, 28, 3753–3756. https://doi.org/10.1029/2001GL013238

Marquardt D.W., 1963, An algorithm for least squares estimation of non-linear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431-41. https://doi.org/10.1137/0111030

Marsal R.J., Masari M., 1959, El subsuelo de la ciudad de México, Contribución al primer Congreso Panamericano de Mécanica de Suelos e ingeniería de Cimentaciones. Facultad de Ingeniería, UNAM, México D.F. (In Spanish).

MATLAB 2019a, The MathWorks, (2019).

Moore, D. S., Notz, W., Fligner, M. A. 2013. The basic practice of statistics, New York: W.H. Freeman and Co. p.p. 654.

Normas Técnicas Complementarias para Diseño por Sismo, Reglamento de Construcciones para el Distrito Federal, Gaceta Oficial del Departamento del Distrito Federal, México. (2004). (In Spanish).

Normas Técnicas Complementarias para Diseño por Sismo, Reglamento de Construcciones para la Ciudad de México, Gaceta Oficial de la Ciudad de México (2017), México. (In Spanish).

Pande, G. N. and Shin, H.S. (2004). *Proceedings of ICE Civil Engineering 157*, February 2004 Pages 39–42 Paper 13242.

Pozos-Estrada A., Chávez M.M., Jaimes M.Á., Arnau O., Guerrero H., 2019, Damages observed in locations of Oaxaca due to the Tehuantepec Mw8.2 earthquake, Mexico. *Natural Hazards*, 97(2): 623-641. https://doi.org/10.1007/s11069-019-03662-9.

Pozos-Estrada A., Gómez R., Hong H.P., 2014, Use of neuronal network to predict the peak ground accelerations and pseudo spectral accelerations for Mexican inslab and interplate earthquake. *Geofísica Internacional*, 53(1): 39-57. https://doi.org/10.1016/S0016-7169(14)71489-8.

Press W.H., Teukolsky S.A., Vetterling A.W.T., Flannery B.P., 1992, Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press, New York. 683–688.

Principe C., Euliano N.R., Lefebvre W.C., 1999, Neural and Adaptive Systems: Fundamentals through Simulation. John Wiley and Sons, New York, NY, USA.

Reinoso E., Ordaz M., 2001, Duration of strong ground motion during Mexican earthquakes in terms of magnitude, distance to the rupture area and dominant site period. *Earthquake engineering and structural dynamics*, 30, 653-673. https://doi.org/10.1002/eqe.28.

Rodríguez P. Q., 2016, Seismic source parameters of normal faulting inslab earthquake in central Mexico. *Pure Appl. Geophys.*, 173, 2587-2619.

Rumelhart D.E., Hinton G.E., Williams R.J., 1986, Learning Internal Representation by Error Backpropagation, in Parallel Distributed Processing: Explorations Microstructure of Cognition, MIT Press, Cambridge, Mass, USA,1, 318–362.

Salmon M.W., Short S.A., Kennedy R.P., 1992, Strong motion duration and earthquake magnitude relationships. Lawrence Livermore National Laboratory, https://dx.doi.org/10.2172/67453

Shahin M., Maier H.R., Jaksa M.B., 2004, Data division for developing neural networks applied to geotechnical engineering. *Journal of Computing in Civil Engineering*, 18(2): 105 – 114. https://doi.org/10.1061/(ASCE)0887-3801(2004)18:2(105).

Singh S.K., Ordaz M., 1993, On the origin of long coda observed in the lake-bed strong-motion records of Mexico City. *Bull. Seism. Soc. Am.*, 83, 1298-1306.

Singh S.K., Reinoso E., Arroyo D., Ordaz M., Cruz-Atienza V., Pérez-Campos X., Iglesias A., Hjörleifsdóttir V., 2017, Deadly Intraslab Mexico Earthquake of 19 September 2017 (Mw 7.1): Ground Motion and Damage Pattern in Mexico City. *Seism. Res. Lett.*, 89(6) 2193 – 2203. https://doi.org/10.1785/0220180159.

Singh, S. K., Pacheco, J. F., Ordaz, M., Kostoglodov, V. 2000, Source Time Function and Duration of Mexican Earthquakes, *Bull. Seism. Soc. Am.*, 90(2), 468-482. https://doi:10.1785/0119990081.

Strasser F.O., Bommer J.J., 2009, Review: Strong Ground Motions—Have We Seen the Worst?, *Bull. Seism. Soc. Am.*, 99(5): 2613 2637. https://doi.org/10.1785/0120080300

Trifunac M.D., Brady A.G., 1975, A study on the duration of strong earthquake ground motion. *Bull. Seism. Soc. Am.*, 65(3) 581-626.

Trifunac M.D., Westermo B.D., 1982, Duration of strong earthquake shaking. *Soil Dynamics and Earthquake Engineering*, 1(3) 117–121. https://doi.org/10.1016/0261-7277(82)90002-X.

Twomey J.M., Smith, A.E. 1995. Performance measures, consistency, and power for artificial neural network models. *Mathl. Comput. Modelling*, Vol: 21(1/2), pp. 243-258.

Yaghmaei-Sabegh S., 2018, Earthquake ground-motion duration estimation using general regression neural network. *Scientia Iranica*, 25(5), 2425–2439. https://dx.doi.org/10.24200/sci.2017.4217

Yuce, B., Mastrocinque, E., Packianather, M. S., Pham D., Lambiase, A., Fruggiero, F. (2014). Neural network design and feature selection using principal component analysis and Taguchi method for identifying wood veneer defects. *Production and manufacturing research*, Vol. 2(1), 291-308. Doi: 10.1080/21693277.2014.892442

APPENDIX A. WEIGHTS AND BIASES FOR THE TRAINED MODELS

The coefficients of the ANN models summarized in Table 8 in italics are given in the following table.

Table A1. Trained ANN models for interplate events.

Case	Weights			Biases	
	$[W_1]_{i,j}$		$[W_2]_{j,k}$	$[\phi_1]_j$	$[\phi_2]_k$

SS-MC

$[W_1]_{i,j}$:
$$\begin{bmatrix} -0.839 & 1.011 & 0.971 & -0.008 & -1.890 & -0.592 & 0.365 \\ 0.455 & -1.239 & -0.313 & -0.717 & 0.796 & 1.647 & 0.098 \\ -0.584 & 0.991 & 0.456 & -1.158 & -0.380 & -9.768 & -1.484 \end{bmatrix}$$

$[W_2]_{j,k}^T = \begin{bmatrix} -0.614 \\ 0.999 \\ -1.442 \end{bmatrix}$ $[\phi_1]_j = \begin{bmatrix} 1.830 \\ 1.348 \\ -1.086 \end{bmatrix}$ $[\phi_2]_k = \begin{bmatrix} -0.693 \end{bmatrix}$

FS-MC

$[W_1]_{i,j}$:
$$\begin{bmatrix} 0.265 & -0.259 & 0.245 & -0.219 & -1.224 & 0.591 & 0.550 \\ 1.231 & 0.070 & 0.032 & 1.209 & 0.303 & 0.346 & -0.617 \\ 1.447 & 0.440 & -0.205 & 1.669 & 0.151 & 0.604 & 0.383 \end{bmatrix}$$

$[W_2]_{j,k}^T = \begin{bmatrix} -0.278 \\ -0.447 \\ 0.385 \end{bmatrix}$ $[\phi_1]_j = \begin{bmatrix} -1.608 \\ -0.336 \\ 1.765 \end{bmatrix}$ $[\phi_2]_k = \begin{bmatrix} -0.344 \end{bmatrix}$

FS-M

$[W_1]_{i,j}$:
$$\begin{bmatrix} -2.460 & 1.128 & -1.598 & 0.679 \\ -1.169 & 1.676 & -2.565 & -0.437 \\ 0.125 & 2.481 & -0.627 & 0.674 \\ 1.532 & -1.421 & -0.608 & 1.836 \\ 1.039 & 0.037 & -2.611 & 0.449 \\ 0.182 & -2.277 & 0.936 & -1.718 \\ -2.107 & 0.495 & -0.785 & 1.543 \\ -2.949 & 1.083 & 1.394 & 0.851 \\ 1.833 & -1.087 & 1.115 & -1.125 \\ -2.168 & -0.932 & -0.587 & -0.039 \\ 0.166 & 2.162 & -0.441 & 0.812 \\ -3.503 & -1.550 & -0.533 & -0.057 \\ 0.751 & -0.498 & 0.358 & -2.794 \\ -0.213 & 1.587 & 2.650 & 0.335 \\ 1.931 & 1.533 & 0.941 & -1.366 \\ 0.703 & -0.945 & 2.787 & 0.533 \\ -3.326 & -0.361 & -1.196 & 0.465 \\ -2.907 & -0.053 & -1.008 & 0.658 \\ -1.962 & -1.879 & 0.812 & 1.068 \\ 0.505 & 0.771 & 1.969 & 0.885 \end{bmatrix}$$

$[W_2]_{j,k}^T = \begin{bmatrix} -0.944 \\ 1.075 \\ 0.261 \\ 0.000 \\ -0.125 \\ -0.241 \\ 0.448 \\ 0.162 \\ 0.495 \\ -1.193 \\ 0.495 \\ 0.784 \\ 0.161 \\ -0.177 \\ 0.124 \\ -0.117 \\ 1.064 \\ -0.935 \\ 0.535 \\ 0.168 \end{bmatrix}$ $[\phi_1]_j = \begin{bmatrix} 2.484 \\ 2.975 \\ -2.181 \\ -2.023 \\ -1.952 \\ -1.489 \\ 1.363 \\ 0.523 \\ -0.394 \\ -0.505 \\ -0.536 \\ -0.354 \\ -1.315 \\ 0.855 \\ 1.120 \\ 0.683 \\ -2.630 \\ -2.448 \\ -2.600 \\ 3.408 \end{bmatrix}$ $[\phi_2]_k = \begin{bmatrix} 0.136 \end{bmatrix}$

Table A2. Trained ANN models for inslab events.

Case	Weights			Biases		
	$[W_1]_{ij}$	$[W_2]_{jk}$	$[W_3]_{kj}$	$[\phi_1]_k$	$[\phi_2]_k$	$[\phi_3]_l$

SS-MC

$[W_1]_{ij}$

```
-0.594  0.392 -1.109  0.757  0.672 -0.829  0.123  0.599
 0.640  0.142 -0.738  0.437  0.871 -0.585  0.620  0.906
-0.333  0.828  1.272 -1.303 -0.376 -0.534  0.398  0.681
 0.536  0.728  0.462 -1.067  0.400 -0.343  1.378  0.498
 0.138  0.043 -0.249 -0.514  0.373 -0.514 -0.001  1.034
 0.237  0.779  1.362 -0.687  0.616 -0.291  0.121 -0.376
 0.643 -0.940 -0.500 -0.486  0.868  0.564  0.383  0.157
 0.420 -0.085  1.128  0.607  0.359 -0.719 -0.223 -0.763
-0.158 -0.958  0.239 -0.764 -0.774  0.986  0.351 -0.031
 0.560 -0.922 -0.065  0.533  0.608  0.363  0.524 -0.793
-0.821 -0.462  0.948  0.397 -0.191  0.897  0.337 -0.532
 0.341  0.247 -0.324 -0.831  0.622  0.585  0.677 -0.981
 1.507 -1.054  0.689  0.032  0.943 -0.509 -0.140 -0.140
-0.633 -0.905 -0.332  0.040 -0.403 -0.436 -0.439  0.418
-0.707  0.739 -0.700  0.256  0.721  0.095  0.693  0.151
```

$[W_2]_{jk}$

```
-0.296  0.417 -0.019 -0.523 -0.108 -0.077  0.486  0.587  0.384 -0.142  0.853 -0.527  0.437 -0.446 -0.642
-0.485 -0.313 -0.569 -0.141  0.613  0.390  0.267 -0.726  0.127 -0.473  0.414 -0.474 -0.122  0.502  0.025
-0.522  0.372 -0.366 -0.471  0.142 -0.073  0.576 -0.891  0.522 -0.092  0.657 -0.716 -0.441 -0.111  0.588
 0.139  0.522  0.981  0.444  0.180 -0.771  0.297 -0.407 -0.049 -0.524 -0.303  0.584 -0.393 -0.223
 0.391 -0.076  0.345 -0.885  0.270  0.528  0.813 -0.141  0.301  0.061  0.672 -0.712  0.132 -0.367
-0.210  0.217  0.983  0.430  0.329 -0.849  0.215  0.472  0.534  0.002  0.080  0.303
 0.279  0.105 -0.419  0.417  0.066 -0.254 -0.231  0.451  0.625 -0.899 -0.956 -0.121  0.265  0.665  0.586
 0.132  0.614 -0.393 -0.035  0.529  0.021  0.422 -0.344 -0.301 -0.605  0.324  0.146  0.327  0.528
-0.463  0.191  0.339  0.219 -0.304  0.260 -0.687  0.240 -0.505 -0.268  0.605 -0.018  0.534  0.228 -0.533
-0.115 -0.764  0.454 -0.156 -0.514  0.813  0.024 -0.070  0.103  0.252 -0.460  0.468 -1.160 -0.288  0.355
 0.740  0.210  0.739  0.120 -0.181  0.741 -0.067  0.509  0.250 -0.105  0.026  0.306 -0.026  0.243 -0.925
 0.602 -0.674 -0.900  0.040 -0.645  0.814  0.375 -0.159 -0.529  0.147 -0.374  0.289  0.431  0.414 -0.162
-0.294  0.236  0.698  0.641 -0.460  0.281 -0.023 -0.028 -0.319  0.808  0.390  0.217  0.122 -0.632
-0.437  0.352 -0.612 -0.390 -0.372 -0.391  0.121 -0.237 -0.060  0.736 -0.004  0.076  0.634 -0.449  0.685
-0.545  0.348  0.364  0.391  0.368 -0.032 -0.351 -0.256  0.099 -0.436  0.233 -0.515 -0.495 -0.455  0.238
```

$[W_3]_{kj}^{\mathrm{T}}$

```
 0.635
 0.044
-0.652
 0.809
-0.472
-0.713
-0.872
 0.341
-0.064
 0.923
-0.652
-0.999
 0.313
 0.095
-0.004
```

$[\phi_1]_k$

```
 1.755
-1.708
 1.022
-0.995
-0.719
-0.375
-0.094
 0.275
-0.188
 0.413
-0.759
 1.157
 1.229
-1.699
-1.969
```

$[\phi_2]_k$

```
 1.694
 1.276
 1.091
-0.990
-0.741
 0.522
-0.197
 0.003
-0.271
-0.488
 0.861
 0.987
-1.166
-1.398
-1.798
```

$[\phi_3]_l$

```
0.770
```

FS-MC

$[W_1]_{ij}$

```
 1.079 -1.056  0.830 -0.093 -0.792 -0.215  1.328  0.986
 0.026 -1.370  0.584 -0.091 -0.091 -0.395  0.659  0.938
-1.103  1.314 -1.258  0.181  0.492  1.342 -1.064  0.234
```

$[W_2]_{jk}$

```
 0.000 -1.105 -2.284
-1.594 -0.911  0.233
 1.669 -1.603  1.891
```

$[W_3]_{kj}^{\mathrm{T}}$

```
 1.061
-0.653
 1.032
```

$[\phi_1]_k$

```
-1.475
 0.522
-2.212
```

$[\phi_2]_k$

```
-1.417
 0.072
 1.520
```

$[\phi_3]_l$

```
0.667
```

FS-M

$[W_1]_{ij}$

```
 0.540 -3.950  1.335 -0.837 -0.419 -0.677  1.435  0.278
 0.156  0.403  0.135 -0.015 -0.072 -0.075  0.067 -0.019
-0.944 -2.716  2.295 -0.692  0.496 -1.447 -1.252 -2.580
 2.182  3.261 -2.212  0.813 -0.183  2.232 -0.173  0.959
-0.294 -1.913  0.601 -0.782  0.497  1.827  2.501 -0.006
```

$[W_2]_{jk}$

```
 1.453 -1.713 -1.065 -0.690  2.435
-0.304  0.699  0.204 -0.425  0.297
-0.893  1.304 -1.802  0.760  0.616
-1.843  0.751  1.287  0.765  0.812
 1.013 -1.205  0.243  1.581  1.613
```

$[W_3]_{kj}^{\mathrm{T}}$

```
-0.457
 2.043
 0.119
-0.677
-1.227
```

$[\phi_1]_k$

```
 2.029
 0.160
-0.668
 2.556
-1.283
```

$[\phi_2]_k$

```
-2.161
 0.187
-0.630
-0.209
 2.070
```

$[\phi_3]_l$

```
1.137
```