ORIGINAL ARTICLE

lajss.org
Latin American Journal of Solids and Structures

www.lajss.org

Peak Ground Acceleration Models Predictions Utilizing Two Metaheuristic Optimization Techniques

Mahir M. Hason[a], Alaa Hussein Al-Zuhairi[b], Ammar N. Hanoon[c], Ali A. Abdulhameed[c], Ahmed W. Al Zand[d]*, Imad S. Abbood[e]

[a]Disaster Information Management Centre, Ministry of Science and Technology, Baghdad 10071 Iraq. E-mail: mahir.mahmod@gmail.com
[b]Department of Civil Engineering, University of Baghdad, Baghdad 10071 Iraq. E-mail: alaalwn@coeng.uobaghdad.edu.iq
[c]Department of Reconstruction and Projects, University of Baghdad, Baghdad 10071, Iraq. E-mail: aliadel@uobaghdad.edu.iq, anh@uobaghdad.edu.iq
[d]Department of Civil Engineering, Universiti Kebangsaan Malaysia, Bangi 43600, Selangor, Malaysia. E-mail: ahmedzand@ukm.edu.my
[e]Engineering Affairs Department, Sunni Endowment Diwan, Baghdad 10071, Iraq. E-mail: imadshakirabbood@gmail.com

* Corresponding author

https://doi.org/10.1590/1679-78256940

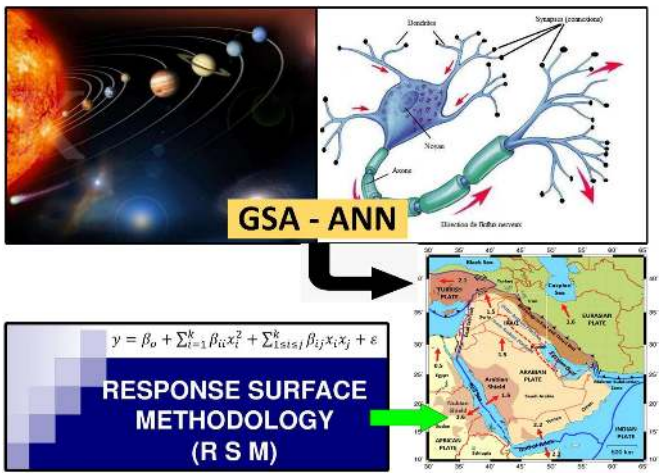
Abstract

Peak ground acceleration (PGA) is frequently used to describe ground motions accurately to defined the zone is critical for structural engineering design. This study developed a novel models for predicting the PGA using Artificial Neural Networks-Gravitational Search Algorithm (ANN-GSA) and Response Surface Methodology (RSM). This paper grants the prediction of PGA for the seismotectonic of Iraq, which is considered the earlier attempt in Iraqi region. The magnitude of the earthquake, the average shear-wave velocity, the focal depth, the distance between the station, and the earthquake source were used in this study. The proposed models are constructed using a database of 187 previous ground motion records, this dataset is also utilized to evaluate the effect of PGA's parameters. In general, the results demonstrate that the newly proposed models exhibit a high degree of correlation, perfect mean values, a low coefficient of variance, fewer errors, and an acceptable performance index value compared to actual PGA values. However, the composite ANN-GSA model performs better than the RSM model.

Keywords

Peak ground acceleration (PGA); Artificial neural network (ANN); Response Surface Methodology (RSM); Analyse factorial design; Gravitational Search Algorithm (GSA); Analysis of variance (ANOVA)

Graphical Abstract



Received January 07, 2022. In revised form April 19, 2022. Accepted May 03, 2022. Available online May 06, 2022.
https://doi.org/10.1590/1679-78256940

Latin American Journal of Solids and Structures. ISSN 1679-7825. Copyright © 2022. This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1 INTRODUCTION

Seismic hazard analysis is a critical step in the engineering phase. Seismological characteristics of earthquakes include their distance, magnitude, soil effects, and kind of faulting. The engineering parameters of an earthquake can be classified into two broad categories: 1) parameters in the response domain; and 2) parameters in the time domain. Pseudo-spectral acceleration (PSA) is a frequently used response domain parameter. Peak ground acceleration (PGA), peak ground velocity (PGV), and peak ground displacement (PGD) are the three major time-domain class parameters. Both of these categories could be used to evaluate the hazards associated with construction. It has been demonstrated that spectral parameters are more effective than time-domain parameters (Luco and Cornell 2007). However, time-domain parameters are more appropriate for applications due to their independence from the structures under consideration (Al-Zuhairi et al., 2021). As a result, PGV, PGD, and PGA are frequently used in seismic risk assessments.

PGA is a well-known earthquake engineering metric that can be utilized for structural analysis and risk assessment during a seismic event. This critical component can be approximated using various techniques, including physical modelling and on-site inspections (Alavi and Gandomi 2011). However, implementing such a method is inconvenient, time-consuming, costly, and frequently impossible (Gandomi et al., 2011). A method for analyzing PGA uses attenuation relationships, which are critical in seismic analysis. PGA is often defined using several independent variables, including the magnitude of the earthquake, the distance between the source and the site, the local site circumstances, and the features of the earthquake source (Güllü and Erçelebi 2007, Gandomi et al. 2011).

Due to the tremendous complexity and non-linearity of the PGA, it is not easy to establish a correlation between it and the predictors. Soft computing methods (SCTs) are widely utilized in engineering research to handle a variety of classification problems and anticipate a variety of issues (Hanoon et al. 2017a,b, Banyhussan et al. 2020), and more recently to predict ground motion characteristics (Alavi et al. 2011, Gandomi et al. 2011). SCTs are typically used to resolve complex numerical optimization problems and nonlinear systems. Numerous research problems in a variety of fields of science have been theoretically and analytically articulated utilizing soft computing techniques (Hanoon et al., 2021). Numerous classifications are incorporated in the SCT, for example, ANFIS (Adaptive Network-based Fuzzy Inference System), ANN (Artificial Neural Networks), SVM (Support Vector Machine), FL (Fuzzy Logic), and OA (Optimization Algorithms) (Jang and Topal 2014). Each category of soft computing also has a fine-grained set of algorithms; for example, the OA category includes the GSA (gravitational search algorithm), PSO (particle swarm optimization), GA (genetic algorithm), ABC (Artificial Bee Colony), ACO (Ant Colony Optimization), and DE (Differential Evolution). Pragmatic modelling and design using SCT are still hotly debated topics, particularly in engineering modelling. SCTs are based on experimental data rather than theoretical and/or analytical derivations, which is the primary distinction between SCTs and traditional models. SCTs are frequently complex and frequently unable to be expressed explicitly. As a result, they are best suited for inclusion as a more prominent computer program component, limiting their application (Hanoon et al., 2021). The artificial neural network (ANN) is the most widely used forecasting method in soft computing, having been successfully employed to address complicated pattern identification and analysis issues in a wide variety of domains, including earthquake engineering. ANNs have been widely used in recent years due to their superior pattern recognition capability, which is advantageous for various problems. The amount of hidden nodes in an ANN model is critical, as an overfitted model can result (Hanoon et al. 2021, Hason et al. 2021).

Using experimental and physical correlations, the prediction of model's equation based on site geology and event propagation has been developed. The regression analysis findings at a specific location with a massive amount of data are explained using empirical models based on mathematical procedures (Mahmod et al., 2017). Seismic wave models can also be utilized when there is insufficient data available to make an accurate determination. Response Surface Methodology (RSM) and Design of Experiments (DOE) have lately been used to create correlations between attenuation and other information processing methods (Hason et al. 2020). Research-based forecasting (RSM) relies on the results of experiments to predict future outcomes. In engineering, the RSM technique can be used to construct an acceptable model for establishing a relationship between the causes and the potential answers to a given issue (Hason et al., 2020).

The primary goal of optimization methods is to achieve values from a set of parameters that minimize and maximize objective functions under restrictions. The ANN-GSA and RSM are used in this study to create two models of peak ground acceleration. The proposed models are based on a large dataset of 187 events of a powerful earthquake that strikes Iraq (2004-2020). This research would primarily contribute to the expanding body of knowledge on PGA and seismic activity assessment, especially in Iraqi tectonic regions. Furthermore, this study also facilitates the usage of combined ANN-GSA and RSM methods for earthquake and hazard forecasting. To date, this is the first attempt to derive explicit and implicit PGA models for the Iraq tectonic region from a variety of parameters. However, there are many unanswered concerns about the extent to which these variables affect PGA. Thus, the factorial design of trials is used to examine the impact of various characteristics on PGA in Iraq. The current research is based on data from the international earthquake stations,

which is a limitation. The response (PGA) selected parameters (R_{Epi}, Mw, Vs30, and FD) that were considered and affected the PGA response are also subject to the constraints herein. A central composite design was used to model the interactions between the elements (CCD). According to the results, the Mw and Vs30 have the most significant impact on PGA, followed by FD, and the REpi has the most negligible impact on PGA.

2 STUDY AREA

Iraq is located within longitudinal (45°38'- 45°48') and latitudinal (28°5'-37°22') coordinates in south-western Asia (Hasan et al. 2014). The topography of Iraq is like a basin as it is surrounded on north and east by mountains, about 3,500 meters above sea level. Iraq has a total surface area of almost 437 thousand km^2 (Abbas et al. 2020, Abbas et al. 2020).

Figure 1 depicts the seismic and topographical conditions in the research region (Iraq) and the adjacent areas. After the collision of Anatolian plates with Iran's plates in the north and northeast of Iraq, the Bitlis-Zagros Fold and Thrust Belt were formed, which included a magnitude w.7.3 earthquake in November 2017 that killed 539 people, as well as thousands of others, in both countries, as it stretched from Turkey and Iraq to the Strait of Hormuz (Abdulnaby et al. 2014, Shafiqu and Sa'ur 2016). The seismicity of plate boundary is linked to a variety of limits in Aden Gulf and the Red Sea, which are both growing. The frontiers of Iraq, Iran, and Turkey, where the Zagros and Bitlis tectonic meeting zones intersect, are experiencing several earthquakes (Ghalib et al., 2006).

There are numerous faults in Mesopotamian Foredeep, including the fault of Badra-Amarah (Iraq's most seismically active fault), the fault of Euphrates (the boundary between Mesopotamian Foredeep and the Stable Platform) as well as the fault of Al-Refaee, Kut, and Hummar (north of Basra). For the most part, Iraq is divided into three main tectonic areas (Fouad and Sissakian 2011). Bitlis–Zagros Fold and Thrust Belt, Mesopotamian Foredeep, and Inner (stable) Arabian Plate, in order from northeast to southwest: 1) (Onur et al. 2017). Seismological and seismotectonic studies in Iraq have clearly shown that seismic activity varies from moderate to high in the northern and north-eastern areas and decreases in the southern and south-eastern regions. in Iraq (Abd Alridha and Jasem 2013).

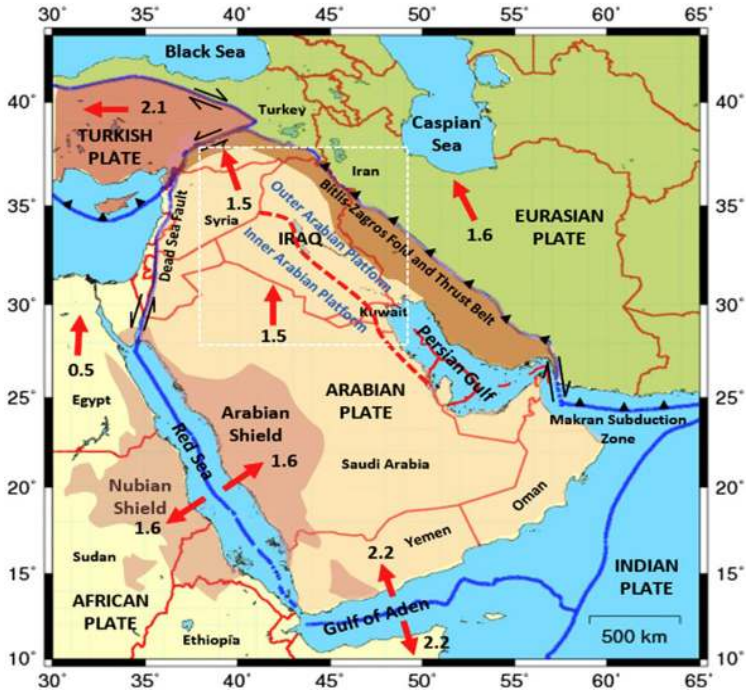


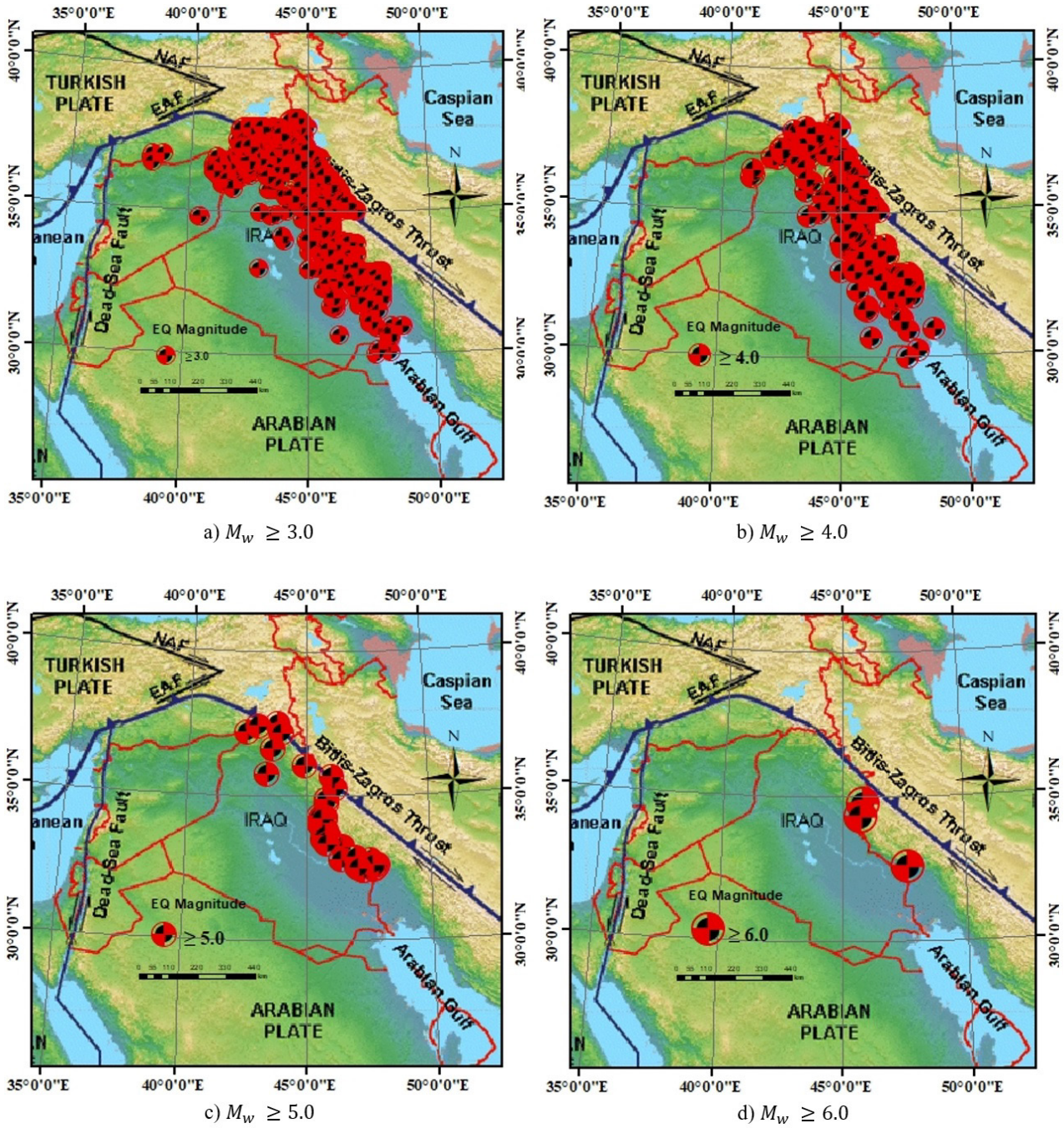
Figure 1: Seismo-geographical map of Iraq and the Arabian Peninsula (red arrows: plate motion in cm per year; blue lines: boundaries of the plate; red lines: country boundaries) is shown (Abdulnaby et al. 2020).

3 METHODS AND MATERIALS

3.1 A description of the database's structure

As illustrated in Figure 2, more than 1800 historical earthquakes with magnitudes ranging from 3.0 to 7.3 struck the research region. Table 1 shows that 187 ground motions occurred between 2004 and 2020, with a magnitude of 4.5 to 7.3, representing the mild, moderate, strong, and significant earthquakes in the research area and the surrounding area.

In order to build and verify the suggested models, the datasets were split into 150 records and 37 records, respectively. A wide range of moment magnitude earthquakes (M_w), average shear wave velocity (V_{s30}), focal depth (FD), and closest epicentre distance (R_{Epi}) are all included in the database. In order to determine the FD, we used worldwide datasets (NOAA, CSEMEMSC, IRIS, and the USGS), including the R_{Epi} and *Mw* values. Actual, $PGA_{act.}$ and V_{s30} on the other hand, were sourced from the GSHAP and USGS databases, respectively. The mapping and comparison were made using ArcGIS.



a) $M_w \geq 3.0$ b) $M_w \geq 4.0$

c) $M_w \geq 5.0$ d) $M_w \geq 6.0$

Figure 2: Earthquake description greater than 3.0 (Richter scale)

With the availability of input data, models derived using SCTs may be anticipated and used for further progress. When it comes to modelling processes, the data quantity is equally important, as it influences the model correctness concerning its intended form. Aside from that, the size of the sample and the parameter combinations affect the performance of an updated model based on these inputs. To further understand how seismic parameter details were included in the suggested models, Table 1 summarizes the input data.

Table 1 Total historical earthquakes records of tested data collected from 29 stations (2004-2020).

No	Station	Records	Range					Region name
			PGA (m/s2)	M_w (Richter)	V_{S30} (m/s)	FD (km)	R_{Epi} (km)	
1	AHWZ	4	2.972-3.74	4.7-5.0	199-713	10	138-189	Iraq–Iran border region
2	BTMT	1	3.172	4.8	847	10	200	Iraq-Turkey border region
3	CUKT	10	2.879-4.173	4.8-6.5	347-878	2-93	20-396	Iraq-Iran border region Iraq-Turkey border region
4	EPOS	1	3.06	4.7	366	10	1093	Iraq-Iran border region
5	GEVA	1	3.41	4.5	854	3	52	Iraq -Turkey-Syria border region
6	GNI	2	2.93-3.318	4.7-4.9	236-344	10	717-737	Iraq-Iran border region
7	HAKT	2	2.93-3.525	4.7-4.9	486-876	2-10	36-397	Iraq-Iran border region Iraq-Turkey border region
8	HKR	4	1.966-3.697	4.7-5.5	469-847	5-4	21-480	Iraq-Iran border region Iraq-Turkey border region Iraq-Turkey-Iran border region
9	IAZR	1	3.441	4.9	438	5	105	Iraq-Turkey-Iran border region
10	IBZA	1	3.697	5	681	10	205	Iraq-Iran border region
11	IDHR	28	1.747-4.445	4.7-7.3	268-900	2-47	50-321	Iraq-Iran border region Iraq
12	IDOB	2	3.129-3.616	4.8-5.9	505-817	10-40	116-142	Iraq-Iran border region
13	IGHG	32	2.018-4.303	4.7-5.8	310-900	2-46	59-168	Iraq-Iran border region
14	IKFM	53	0.922-4.569	4.7-5.9	220-810	2-52	82-261	Iraq-Iran border region
15	IKMR	11	3.229-4.241	5.2-6.2	392-653	10-50	113-132	Iraq-Iran border region
16	IKOM	2	2.163-3.506	4.8-4.9	604-768	10-40	168-177	Iraq-Iran border region
17	KBZ	1	3.62	4.7	700	10	1258	Iraq-Iran border region
18	KCHF	4	3.016-4.137	4.7-5.1	528-776	10-20	47-155	Iraq-Iran border region
19	MAHB	2	2.852-3.016	4.7-5.1	600-651	10	103-136	Iraq-Iran border region
20	MSL	3	2.991-3.102	5.1-5.2	273-479	10	46-78	Iraq
21	MYA	1	2.975	4.9	687	10	683	Iraq-Iran border region
22	SHGR	5	2.966-3.796	4.7-5.7	275-668	10-50	106-175	Iraq-Iran border region
23	SIRT	3	2.966-3.318	4.7-5.5	337-900	5-10	20-43	Iraq -Turkey-Syria border region
24	SLY	5	3.269-3.748	4.7-5.5	365-900	10-30	81-200	Iraq-Iran border region
25	SNGE	2	3.245-3.494	5.2-5.3	274-600	20-60	222-249	Iraq-Iran border region
26	VANB	1	2.057	4.7	587-900	80	408	Iraq-Iran border region
27	VANT	3	2.117-3.563	5.1-5.6	760-900	9-10	82-108	Iraq-Turkey border region
28	VNNZ	1	3.904	4.9	900	2	347	Iraq-Turkey-Iran border region
29	YOVA	1	4.017	5.0	509	10	173	Iraq-Iran border region
Total		187						

According to the recommendations, the best dataset-to-input variable ratio is 3 for model applicability and greater than 5 for additional safety (Frank and Todeschini 1994). Thus, of the 187 datasets, 150 datasets (80%) were utilized to create the models, while the remaining 37 datasets (20%) were used to verify the models. In the current study, 150/5= 30 and 37/5= 7.4 were used to construct and test the suggested model. There was a considerable difference between designing and checking datasets (i.e., 5).

3.2 A Computed Intelligence (CI)

Computed Intelligence (CI), which is also known as soft computing technique (SCT), is commonly utilized to determine nonlinear systems, complicated mathematical optimization queries, and non-differentiable problems (Jang and Topal 2014). One of the main objectives of the present paper is to precisely foretell the PGA of Iraq's tectonic regions employing the composite of ANN and GSA. The composite $ANN-GSA$ algorithm is generated according to four effective independent input variables ($M_w, V_{s30}, FD, R_{Epi}$) against the response of dependent PGA as follows:

$$PGA = f\left(M_w, V_{s30}, FD, R_{Epi}\right) \quad (1)$$

3.2.1 Artificial Neural Network (ANN)

ANN represents a data processing system, which has been produced as a comprehensive numerical form of rational natural nerves. In general, the main benefits of utilizing $ANNs$ are their ability to detect errors ideally and maintain training to develop their achievement when dealing with new learning data (Gholami et al., 2013). ANN techniques are capable to model the complicated numerical relationship between the input parameters (i.e. FD, R_{Epi}, M_w and V_{s30}) and target response (i.e. PGA). A BP neural network and the Levenberg-Marquardt (LM) training method were used in the current study to prepare, examine, and verify the results. The reason behind considering the LM training algorithm besides its activity and performance, since it provides fewer localization failure (Payal et al. 2014). Nevertheless, the LM algorithm needs a significant quantity of operating memory (Kukolj and Levi 2004).

In the ANN, prior preparation, examination, and verification of the data, many data must be determined first, i.e, the variables, number of hidden zones, rate of learning, and output number. Based on the selected variables in this research, which is represented by four parameters (M_w, V_{s30}, FD, and R_{Epi}), the number of zones to build the ANN design is four to determine the response PGA of the tectonic study area. These zones included: (i) input zone; (ii) hidden zone; and (iii) output zone, as displayed in Figure 3. The input zone comprised four parameters, M_w V_{s30}, FD and R_{Epi} values. The input variables were addressed versus the powers of the individual connections W_{ij} and summated per every neuron of the hidden zone to measure the output y_k in the output zone. Tan-Sigmoidal activation functions and the linear activation functions were utilized in the hidden and output zones, respectively, to include all varieties of the response values (PGA). The neurons number in the hidden zone and training degree was chosen according to the GSA algorithm, which determines the most suitable number of neurons and training degree values of ANN in the hidden zone, to achieve the best optimum solution. Hence, the performance of ANN can be modified. This GSA-algorithm comprises ANN to create a new algorithm called the composite $ANN - GSA$ with a minimum of PGA predictions.

The algorithm of BP (backpropagation) in multi-zone feed-forward networks is the appropriate standard algorithm based on the mathematical design of the training complex nonlinear connections. This performance index of the BP algorithm is called MSE (least mean square error), which can be determined by calculating the difference between the objective and the network outputs (Eq. 2).

$$MSE = \frac{1}{n}\sum_{i=1}^{ns}(t_i - o_i)^2 \quad (2)$$

Where n, $t_i, and\ o_i$ represents the number of learning configurations, the objectives output, and the networks output, respectively.

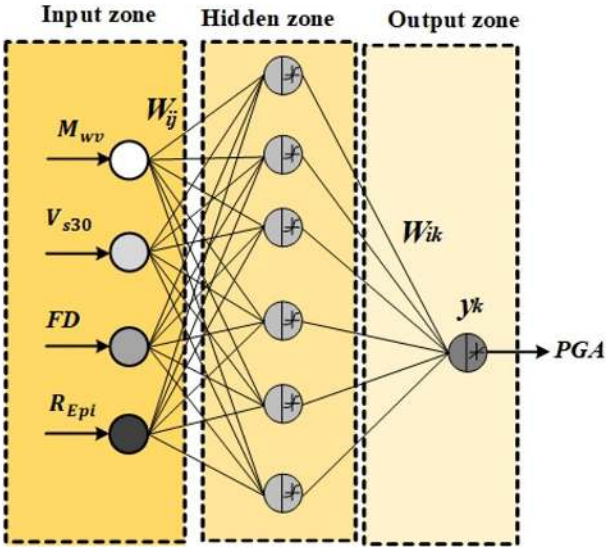


Figure 3: A diagram design of the neural network

3.2.2 GSA technique

There are several excellent ways for finding a rational software solution, but one of the finest is the Heuristic algorithm (Rashedi et al., 2009). GSA is a contemporary heuristic algorithm based on Newton's gravitation and motion equations. An acceptable variable and sufficient input parameter values are necessary for an ANN to reflect the correlation between inputs and output. Consequently, the challenge concerns the number of hidden nodes and learning rate values used in an ANN algorithm. Hence, developing the hybridized version of soft computing models has overcome the previous drawbacks. In this study, these problems can be addressed using the performance of the gravitational search algorithm (GSA) algorithm through searching the optimal values. PSO algorithm is one of the distinguished optimization approaches introduced in the literature within the soft computing implementation for related structural engineering problems (Yaseen et al., 2018, Chen et al. 2018, Kaveh and Talatahari 2012). In order to determine the optimal ANN variables (number of neurons in every hidden layer and training rate), the GSA (heuristic algorithm) is merged with ANN. According to the suggested algorithm, agents' performance is defined by the number of targets they have been identified as being.

Objects are examined in the method, and their mass is used to gauge performance. Thanks to Newton's equations of motion and gravity (Schutz 2003). Illustration of the mass effect concerning other masses as shown in Figure 4. The heavier masses in this figure are linked to good outcomes and travel slower than the lighter masses. Taking this step will allow the algorithm to be more easily attacked.

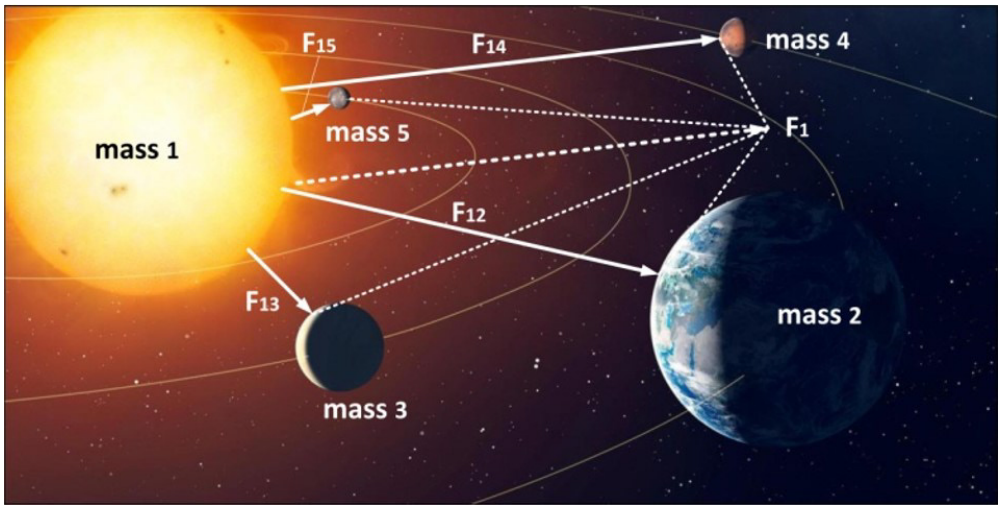


Figure 4: Effect of a mass on other masses

Eq. (3) defines the ith agent location.

$$X_i = \left(x_i^1, x_i^2, \ldots., x_i^d, \ldots\ldots, x_i^n\right) i = \ 1, 2, 3, \ldots\ldots, P \tag{3}$$

The expressions x_i^d, n, and P represent the location of the dth dimension, the search space, and the agent's number, respectively. Besides, the gravitational force affecting on the ith target according to the jth target can be formulated as (Rashedi et al. 2009):

$$F_{ij}^d(t) = G(t) \times \left(\frac{M_j(t) \times M_i(t)}{R_{ij}(t) + \varepsilon}\right) \times \left(x_j^d(t) - x_i^d(t)\right) \tag{4}$$

$$R_{ij} = \left\| x_i(t) + x_j(t) \right\|_2 \tag{5}$$

$$G(t) = G_0 \times e^{-\alpha t / Iter_{max}} \tag{6}$$

Where M_i is the jth agent's mass, and the gravitational constant at time t is $G(t)$, $R_{ij}(t)$ is the Euclidean distance between the ith and jth agents. G_0 is the gravitational constants initial value, and, $Iter_{max}$ is the maximum number of iterations (the total age of the system). When applied to the ith agent, the total force is denoted as:

$$F_i^d(t) = \sum_{j=1}^{kbest} rand_j \times F_{ij}^d(t)\ j \neq i \tag{7}$$

Where $kbest$ is the set of the first k agents with the best fitness (objective function) value and is evaluated in such a way that it decreases linearly with time (Rashedi et al. 2009) and its value becomes 2% of the initial number of agents at the last iteration. where, rand-j. is a random number in the interval (0, 1). According to Newton's rule of motion, the ith agent's acceleration, position, and velocity are given in Eqs. (8-10) at the tth repetition and position. $fit_i(t)$ represents the ith agents fitness value at the tth repetition of the, i^{th} repetition. The suggested algorithm stages are shown in Figure 5. According to the GSA algorithm, all masses are presumed to be equal. A more significant inertia mass improves search accuracy since the agents movement is slow. In contrast, an enormous gravitational mass attracts more agents, resulting in a quicker convergence rate.

$$a_i^d(t) = \frac{F_i^d(t)}{M_i(t)} \tag{8}$$

$$v_i^d(t+1) = rand_i \times v_i^d(t) + a_i^d(t) \tag{9}$$

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1) \tag{10}$$

$$M_i(t) = \frac{m_i(t)}{\sum_{j=1}^{P} m_j(t)} \tag{11}$$

$$m_i(t) = \frac{fit_i(t) - worst(t)}{best(t) - worst(t)} \tag{12}$$

$$best(t) = min\big(fit_i(t)\big)\ i = 1, 2, \dots\dots, P \tag{13}$$

$$worst(t) = max\big(fit_i(t)\big)\ i = 1, 2, \dots\dots, P \tag{14}$$

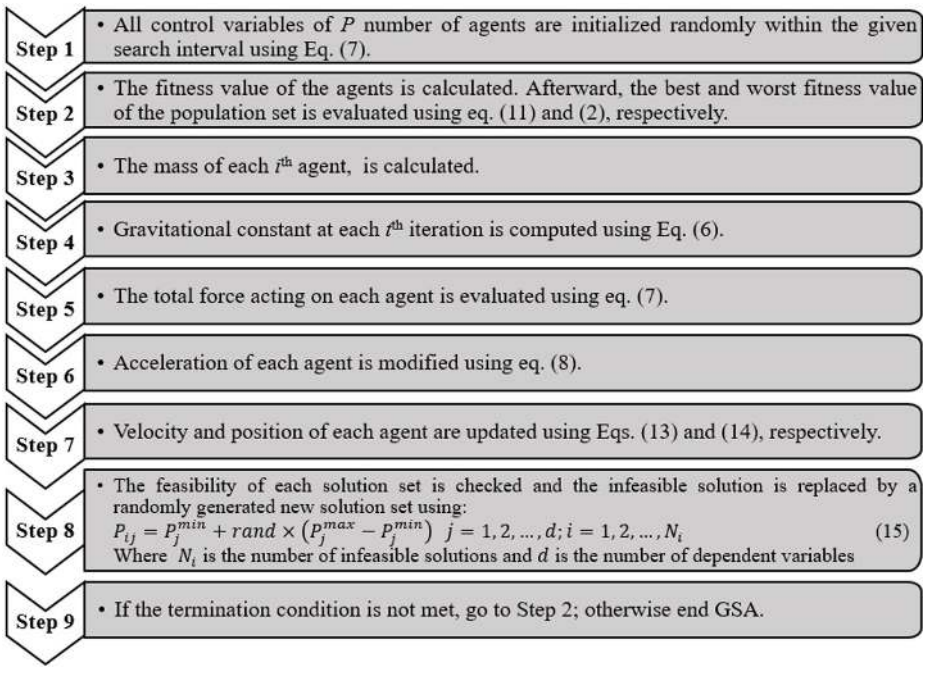


Figure 5: Flowchart steps of the GSA algorithm.

3.2.3 Composite ANN-GSA development

$ANNs$ are prepared from a dataset named as the preparing or training data. Through the preparing trial, the network's measurements are optimized. Besides, the preparation system includes two significant actions: initialization

and optimization. A solid and effective preparation procedure requires both processes of initialization and optimization (Alavi and Gandomi 2011). The flow chart in Figure 6 illustrates the soft computing techniques (SCTs) used in this work.

The mean absolute error (MAE), represented by Eq. (16), is the accurate tool between the actual and proposed models. (Hanoon et al. 2016):

$$MAE = \frac{|\sum(PGA_{act.,i} - PGA_{pred.,i})|}{n} \tag{16}$$

where n, $PGA_{act.,i}$ and $PGA_{pred.,i}$ represent the total number of constructed records the actual and predicted values of the PGA, respectively.

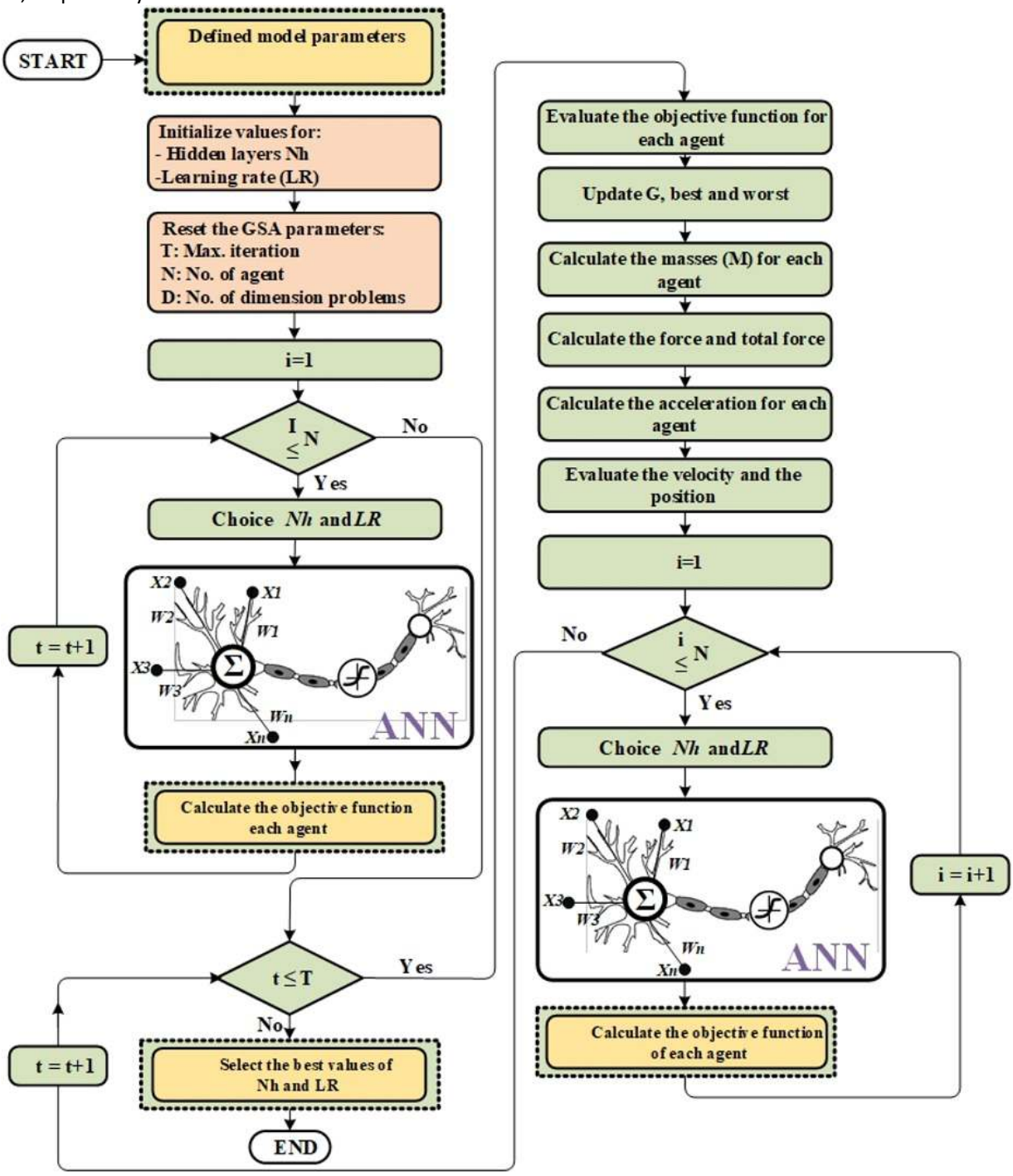


Figure 6: Schematic diagram of the proposed composite $GSA-ANN$ algorithm.

3.2.4 Data pre-processing for composite ANN-GSA technique

A dataset of 150 points, representing 80% of the total records (187 points), is processed as learning data. On the other hand, the remaining dataset of 20% (37 points) verifies the composite ANN-GSA model.

Where the total data were divided into two main parts, the first one (80% of the total data), verification of these data was carried out by dividing it into three main parts, which were as follows:

1. During the training phase, 70% of the data was utilized.
2. 15% of the data was used during the testing phase.
3. 15% of the data was used during the validation phase. External data were used to verify further the model's accuracy in the second main part of the data, which accounts for 20% of the total data and whose values are not included in the first part.

In implementing any ANN model, the main problem is the selection of the datasets correlated to the issue during examination. To achieve stability during the analysis of this study, all input and output datasets were normalized between 0 to 1, by applying Eq. (17), before training the network by utilizing the max-min normalization criterion. The normalization of the data record prioritizes all of the computation parameters equally.

$$x_n = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{17}$$

Where x, x_n, x_{min}, and, x_{max} are the actual parameter value, the normalized value of the specific parameter, the min. and max. values of the database respectively.

3.3 Statistical model designation by RSM

To construct a new statistical model, numerous phases should be considered: (i) deriving the final model according to available datasets; and (ii) carrying out a parametric analysis according to the principles of engineering and problem physics. The first phase is constructed by the Response Surface Methodology (RSM). The second phase focuses on engineering principles and must be carried out based on a parametric analysis by an engineer that acknowledges the issue being modelled. This study tries mainly to analyze the effect of individual parameters on PGA.

Second-order polynomial or quadratic models can model and evaluate problems using RSM, a statistical and mathematical technique (Box and Draper 1987). An optimal output value can be found by exposing the solution to various factors using this technique (Montgomery 2017). The surface response methodology's central composite design (CCD) technique becomes extremely flexible whenever the intended parameters' preliminary lower and upper limitations have been surpassed. Using the CCD process, it is still possible to obtain the best possible values for parameters that appear well outside the initial fixed range. This study aims to learn more about what goes into calculating the PGA. To determine how design parameters interact, the CCD approach of RSM was utilized in conjunction with the DOE method (a statistical approach). This method does check the impact of parameters on the selected response (Antony 2014).

It is possible to discover the ideal combination of factors and their relationship, which is impossible in standard optimization techniques through factorial designs. Aside from that, mathematical models can be generated using these designs as a starting point. For the RSM, Eq. (18) represents a second-order polynomial model, which is commonly used to assess the impact of many variables on a response based on the datasets collected:

$$y = \beta_o + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{1 \le i \le j}^{k} \beta_{ij} x_i x_j + \varepsilon \tag{18}$$

Where y denotes the estimated response, x_i and x_j are coded variables, β_o denotes the constant, β_i denotes the linear coefficient, β_{ii} denotes the quadratic coefficient, and, β_{ij} denotes the interactive coefficient (Montgomery 2017). The most significant characteristics illustrating PGA behaviour were picked after conducting a trial study and a literature review (Gandomi et al. 2016). As a result, the formulation of PGA must take into account the link between the response and the specified parameters as described before in Eq. 1 ($PGA = f\left(M_w, V_{s30}, FD, R_{Epi}\right)$). The RSM mechanism is depicted in Figure 7.

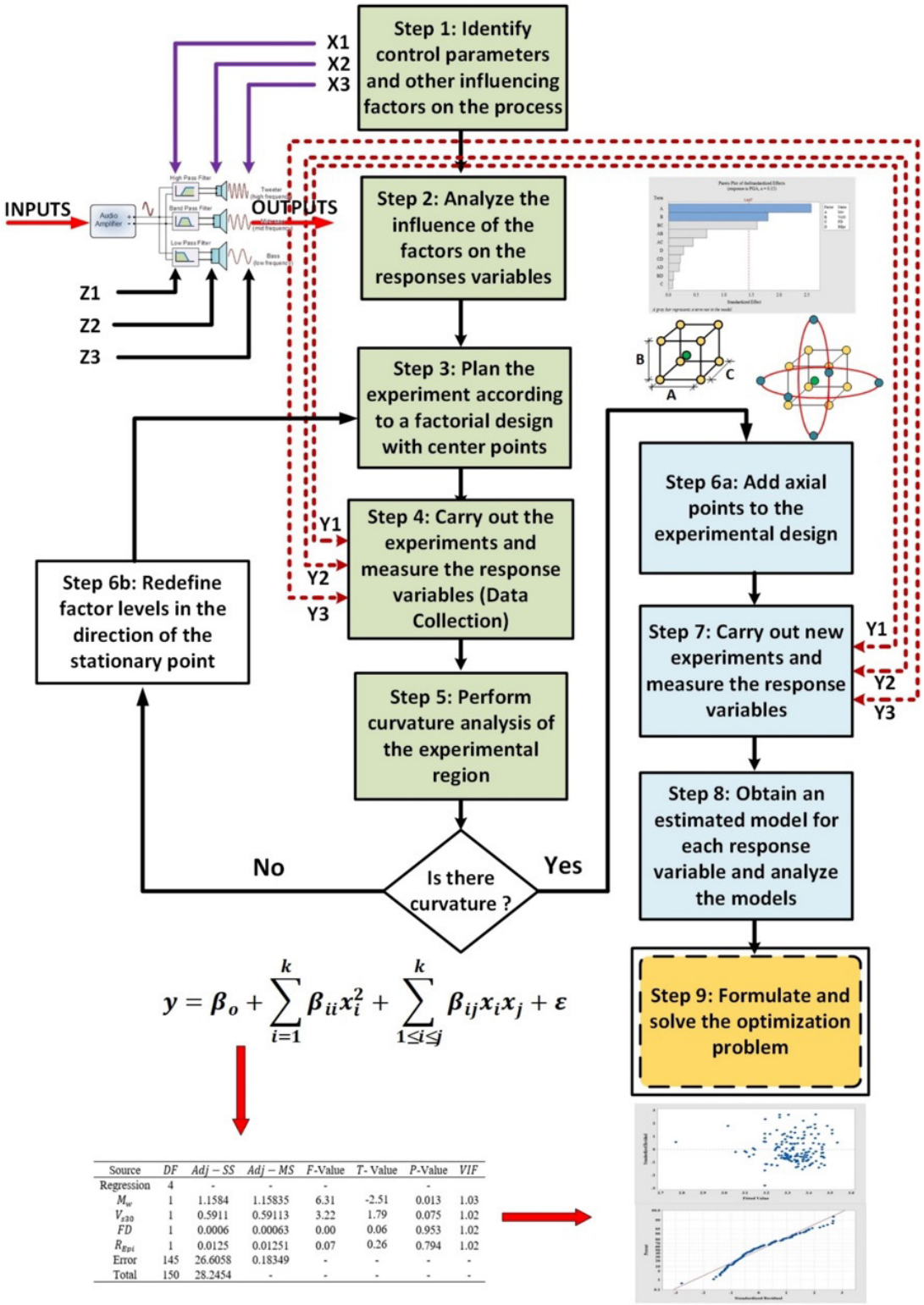


Figure 7: Flowchart of the RSM technique.

4 RESULTS AND DISCUSSION

The results of the proposed PGA models in terms of $ANN - GSA$ algorithm (composite $PGA_{ANN-GSA}$) and RSM (PGA_{RSM}) are presented and discussed in this section. Followed by the comparison between the proposed models relative to the actual values ($PGA_{act.}$). Validation and verification are conducted to investigate the degree of accuracy of the implicit and explicit models. Finally, the impact of the independent variables on the dependent response (PGA) is carried out to know which parameter has a significant influence during the selection of the ground motion components.

4.1 Composite ANN-GSA Algorithm model ($PGA_{ANN-GSA}$)

It should be noted that a specific algorithm to produce an exact outcome for all optimization problems is not exist. However, the GSA algorithm was performed throughout the parameter settings with different population sizes (60, 80, and 100) to enable the algorithm to choose the best population that accomplished the minimum accurate function, as shown in Table 2. To achieve the minimum errors between the actual and predicted PGAs, the preparation operation of ANN was replicated numerous times utilizing a high number of iterations (i.e., 1000).

MATLAB software was used to run the GSA algorithm according to the selected population sizes to obtain the Nh (number of Neuron in the hidden layer) and LR (Learning rate) as shown in Table 3. The mean absolute error (MAE), which represents the objective function, of the composite $ANN-GSA$ for various population sizes is demonstrated in Figure 8. This figure demonstrates that the optimal solution for the GSA is represented by the population size of 100 since it reaches the lowest MAE of about 125 iterations compared to the population size of 60 and 80, which needs more time of about 160 and 275 iterations respectively to reach the minimum error. According to the results of the proposed composite $ANN-GSA$, the ANN was run based on the variables that performed the minimum MAE of population size—resulting in a high calculation accuracy of proposed PGA model.

Table 2 shows the GSA and Neuron parameters in each hidden zone and the ANN learning rate based on population size.

Parameters			Value
Count of inputs			4
No. of hidden zones			1
Count of outputs			1
Error locating the desired target			1e-05
Iteration			1,000
G_0 (Gravitational constant)			1
α (Constant term)			0.2
Population size	60	Nh	18
		LR	0.2743
	80	Nh	15
		LR	0.8794
	100	Nh	16
		LR	0.7011

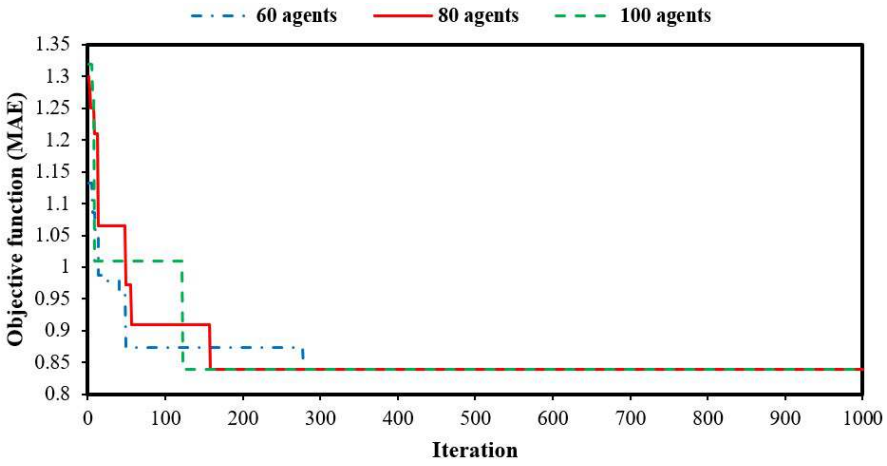


Figure 8: Objective function versus iteration.

4.2 ANN training and validation

To choose the best values of neurons in every Nh and LR, the GSA is utilized to optimize the ANN running, which operated using the input parameters (FD, R_{Epi}, M_w and V_{s30}) values and output response of the actual $PGA_{act.}$ for Iraq's tectonic regions. The potential of different ANN structures. Figure 9 depicts the ANN structure with four input variables, sixteen hidden neurons in the hidden layer, and a single output variable. This structure was evaluated to determine the optimized circumstances of PGA for Iraq's tectonic region.

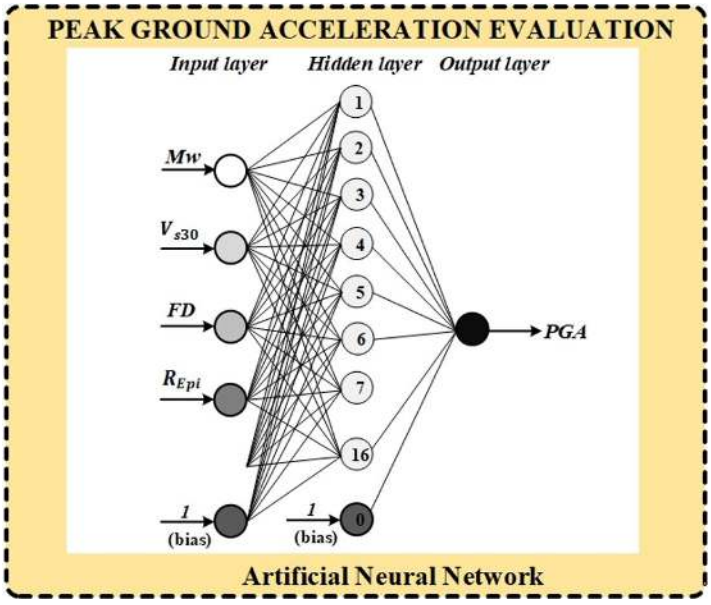


Figure 9: The structure of the developed ANN for the proposed PGA model.

In the ANN, the learning dataset of the composite $ANN - GSA$ is divided into three categories: training data, testing data, and validation data, as shown in the regression graph of Figure 10. The x-axis (target) and y-axis (output) are represented by the actual and predicted PGA. It is seen that the statistical results of the correlation coefficient (R) for the training, testing, and validation records have a good correlation between the actual and predicted PGA that produced by $ANN - GSA$ algorithm.

Ideally, the most residual error values (differences between target and output) are remarkably less, located near the zero lines as revealed in the error histogram with 20 bins of Figure 11. The $LMBP$-algorithm (Levenberg-Marquardt backpropagation) performance throughout the training development is plotted in Figure 12 using MATLAB R2019b. The difference of MSE (mean square error) by training epochs is drawn, where the best validation performance is 0.897 at epoch 2.

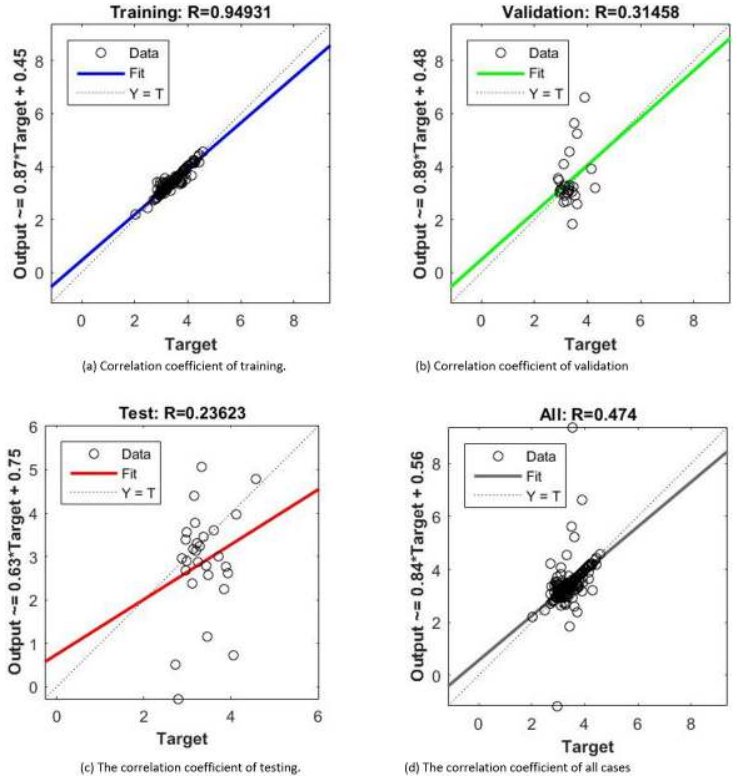


Figure 10: Plot of the regression analysis ($LMBP$-algorithm) for the proposed model (composite $ANN - GSA$)

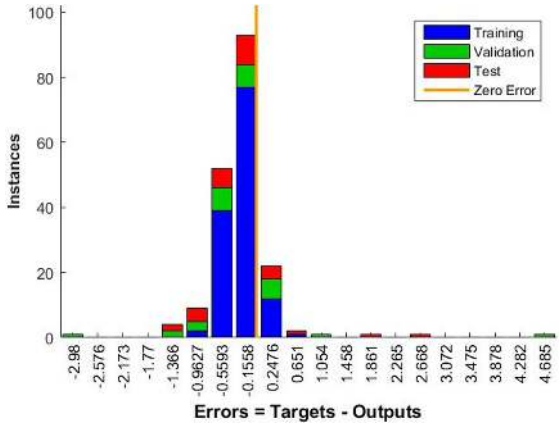


Figure 11: Residual histogram of the proposed model (composite $ANN - GSA$)

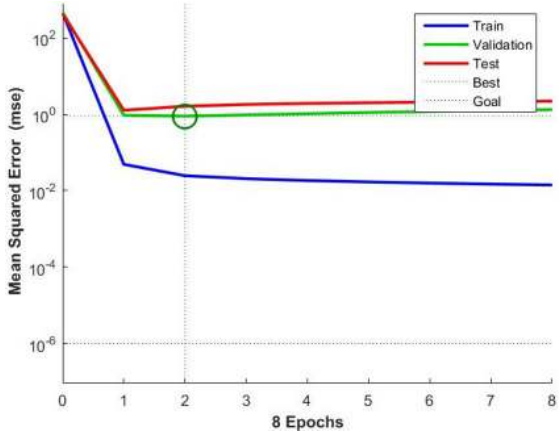


Figure 12: $LMBP$-algorithm performance.

4.3 PGA model using RSM (PGA_{RSM})

To capture the PGA response against other variables (M_w, V_{s30}, FD and R_{Epi}), response surface methodology (RSM) is employed using Minitab software. With $ANOVA$ (analysis of variance) and model summary, results are demonstrated in Tables 3. Moreover, T-value and P-value are two vital factors in evaluating the correlation between variables. T and P are connected (Winship and Zhuo 2020). They go hand in hand, similar to Tweedledee and Tweedledum. When T-value is close to 0 (negative or positive values), the more probable there is not a notable variation between variables, as shown in Table 3. The P -value is utilized to determine evidence strength in the data provided. Generally, the lower P -value, the greater the sample evidence for a significant correlation (Chaubey 1993). By convention, a P-value higher than 5% is called not statistically significant and vice versa (Altman and Bland 1995). It should be noted from Table 3 that the P-value (for M_w) is < 0.05, which means the factor is more significant in PGA findings.

A realistic measure of the degree of multi-collinearity in a regression is the Variance Inflation Factor (VIF), which is a term used to quantify how strongly two or more predictors in the regression are associated. For more information (Robinson and Schumacker 2009). The VIF is, therefore, an essential part of examining interaction effects in multiple regression. According to Table 3, VIF values are around 1, the regressions have a good shape and are multi-collinearity.

Table 3 Analysis of Variance ($ANOVA$).

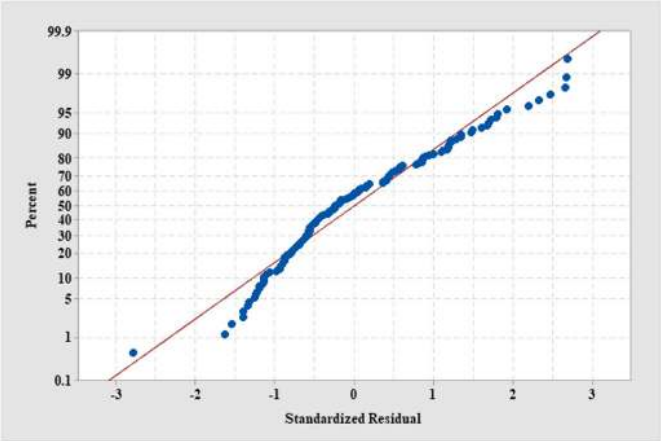
Source	*DF*	*Adj-SS*	*Adj-MS*	*F-Value*	*T- Value*	*P-Value*	*VIF*
Regression	4	-	-	-		-	
M_w	1	1.1584	1.15835	6.31	-2.51	0.013	1.03
V_s30	1	0.5911	0.59113	3.22	1.79	0.075	1.02
FD	1	0.0006	0.00063	0.00	0.06	0.953	1.02
R_Epi	1	0.0125	0.01251	0.07	0.26	0.794	1.02
Error	145	26.6058	0.18349	-	-	-	-
Total	150	28.2454	-	-	-	-	-

As we mentioned previously, out of the 187 datasets, 80% of them (i.e., 150 datasets) are adopted for the explicit model fabrication processes, and the rest 20% (37 datasets) are considered for verifying the final model. Fit Factorial multi-linear regression model is conducted to find the best correlation between dependent response and independent parameters (PGA versus M_w, V_{s30}, FD and R_{Epi}). The explicit equation is formulated to be:

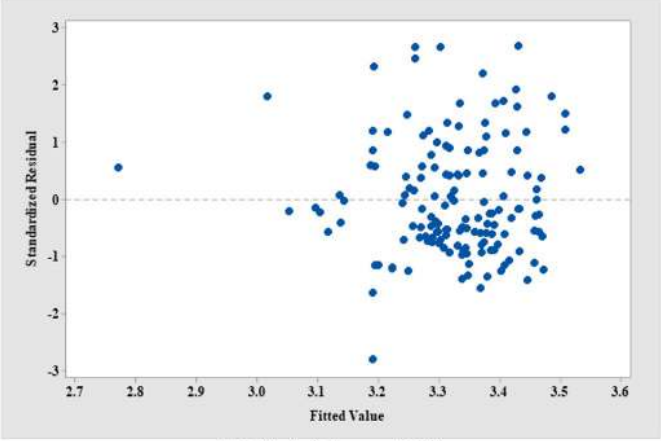
$$PGA_{DOE} = 4.287 - 0.2308\, M_w + 0.000321\, V_{S30} + 0.00016\, FD + 000056\, R_{Epi} \quad (19)$$

Following the building of the model, it is essential to assess its statistical trustworthiness. From a statistical point of view, one condition has to be settled by a model for its reliability fits approved. This condition is normalized residuals. The expression residual denotes the variation between an experimentally determined value achieved by the model.

The PGA Fit multi-linear regression analysis findings are shown in Figure 13 (Minitab). The residuals are shown in the outline for demonstration purposes. Figures showing the distribution of differences between predicted and observed values show few outliers in the data set. The PGA Fit multi-linear regression analysis findings are shown in Fig. 13 (Minitab). These graphs afford good evidence to confirm the use of the suggested PGA model. The reliability is verified since the circumstance of model authenticity has been performed.



(a) Normal Probability Plot (response is PGA)



(b) Residual vs fits (response is PGA)

Figure 13: Residual output.

4.4 Comparative study

Datasets of 20% (37 records) of the overall datasets (187 records) are employed to evaluate the proposed PGA models by $ANN - GSA$ algorithm and RSM. Those data have not been used in the model building process. The comparisons are conducted between the prediction models. The outcomes achieved from the proposed models by the verification records are displayed in this section. Besides, the outcomes show that the proposed composite model ($PGA_{GSA-ANN}$) exhibits better than the model produced by RSM (PGA_{RSM}).

The standard deviation (SD) was detected as the data variance was measured. The fewer SD results in less data variance, and conversely. As a result, the variance coefficient (CoV) measures the actual amount of relative variation and

reflects the correctness between output and input data. According to [Pimentel-Gomes 2000], a CoV value of less than 10% indicates great precision, whereas values of 20–30% indicate low precision, and more than 30% indicate low precision. Table 4 demonstrates that both the proposed models have reasonable CoV-values. It was found that the CoV values for the two models ($PGA_{GSA-ANN}$ and PGA_{RSM}) were 8.739 percent and 10.362%, respectively, with great accuracy in determining the target values. Furthermore, an excellent value adjacent to 1.0 of the mean-values for both models were obtained (1.061 and 1.07). The RSM shows accuracy slightly as compared with the GSAANN technique.

Table 4 The comparative details of actual and predicted values.

No.	M_w	V_{S30}	FD	R_{Epi}	PGA_{act}.	$PGA_{ANN-GSA}$	PGA_{RSM}	$\frac{PGA_{act}}{PGA_{ANN-GSA}}$	$\frac{PGA_{act}}{PGA_{RSM}}$
1	4.7	466	10	108	3.129	3.356	3.367	0.932	0.929
2	5.6	310	10	97	2.887	3.192	3.101	0.904	0.931
3	4.8	285	40	141	3.494	3.235	3.292	1.080	1.061
4	5.2	553	30	115	4.224	3.435	3.280	1.230	1.288
5	4.8	594	10	98	3.172	3.399	3.384	0.933	0.937
6	5.0	365	93	330	3.46	3.360	3.289	1.030	1.052
7	5.1	555	10	127	3.127	3.282	3.302	0.953	0.947
8	4.9	876	2	36	3.245	2.957	3.430	1.097	0.946
9	5.2	548	10	100	3.129	3.368	3.274	0.929	0.956
10	4.9	719	20	128	3.324	3.554	3.404	0.935	0.977
11	4.7	854	10	117	3.41	3.394	3.492	1.005	0.976
12	5.1	789	10	82	3.305	3.226	3.374	1.025	0.979
13	5.8	400	10	101	3.011	3.128	3.084	0.963	0.976
14	5.1	723	10	82	3.324	3.293	3.353	1.009	0.991
15	5.2	751	10	86	3.335	3.240	3.339	1.029	0.999
16	4.7	255	52	138	3.483	3.295	3.308	1.057	1.053
17	5.0	545	10	79	3.339	3.435	3.320	0.972	1.006
18	4.9	426	10	107	3.32	3.367	3.307	0.986	1.004
19	5.3	668	50	148	3.55	3.058	3.298	1.161	1.076
20	5.0	279	40	140	3.424	3.181	3.242	1.076	1.056
21	4.7	734	40	83	3.71	3.734	3.457	0.994	1.073
22	6.2	392	10	126	3.091	2.726	2.991	1.134	1.034
23	4.9	651	46	90	3.711	3.338	3.384	1.112	1.097
24	4.8	847	8	21	3.641	3.461	3.461	1.052	1.052
25	5.1	900	47	321	3.835	3.646	3.429	1.052	1.118
26	5.1	760	9	82	3.587	3.211	3.365	1.117	1.066
27	4.7	610	45	60	3.868	3.420	3.416	1.131	1.132
28	4.9	380	20	445	3.616	3.603	3.306	1.004	1.094
29	5.0	270	10	140	3.527	3.337	3.235	1.057	1.090
30	4.9	337	10	113	3.616	3.393	3.279	1.066	1.103
31	5.3	257	20	168	3.598	3.251	3.162	1.107	1.138
32	4.8	387	10	112	3.796	3.359	3.318	1.130	1.144
33	4.7	533	10	85	3.927	3.447	3.388	1.139	1.159
34	5.0	623	10	68	3.932	3.442	3.333	1.142	1.180
35	4.9	689	10	132	4.137	3.249	3.393	1.273	1.219
36	4.8	688	27	71	4.569	3.819	3.415	1.196	1.338
37	5.2	288	16	68	4.445	3.626	3.190	1.226	1.393
					Mean Value (MV)			1.061	1.070
					Standard Deviation (SD)			0.093	0.111
					Coefficient of Variation (CoV) (%)			8.739	10.362

The correlation coefficient (R) is an inadequate indicator of a model's predicting effectiveness since (R) is not sensitive to output values multiplied by a constant (Gandomi et al. 2011). Accordingly, another recommended function should be employed to assess the proposed models' performance. The suggested models' performance may be assessed by comparing their PIs (performance indices) with $RRMSEs$ side to side with R (Eq. 20).

$$PI = \frac{RRMSE}{R+1} \tag{20}$$

$$RRMSE = \frac{1}{\left|\overline{PGA}_{av.,act.,i}\right|}\sqrt{\frac{\sum_{i=1}^{n}\left(PGA_{act.,i}-PGA_{pred.,i}\right)^2}{n}} \tag{21}$$

$$R = \frac{\sum_{i=1}^{n}\left(PGA_{act.,i}-\overline{PGA}_{av.,act.,i}\right)\left(PGA_{pred.,i}-\overline{PGA}_{av.,pred.,i}\right)}{\sqrt{\sum_{i=1}^{n}\left(PGA_{act.,i}-\overline{PGA}_{av.,act.,i}\right)^2\sum_{i=1}^{n}\left(PGA_{pred.,i}-\overline{PGA}_{av.,pred.,i}\right)^2}} \tag{22}$$

Where $PGA_{act.,i}$ refers to the actual value, $PGA_{pred.,i}$ refers to the predicted outputs, $\overline{PGA}_{av.,act.,i}$ refers to average actual values, and $\overline{PGA}_{av.,act.,i}$ refers to the average predicted outputs for n number of samples.

The value of PI is located between 0 and $+\infty$. The relationship between R and $RRMSE$ values is an inverse relationship. Thus, to obtain a perfect prediction performance, the PI value should be closed to 0, which means higher R ith lower $RRMSE$. As shown in Table 5, the PI values are too small of about 0.0613 and 0.0862 for $PGA_{ANN-GSA}$ and PGA_{RSM} respectively, which indicate that the proposed models predict perfectly the experiment values of PGA_{actual}. According to Table 5, the statistical performance of the composite model produced by $ANN-GSA$ algorithm exhibit better outcomes relative to the proposed model generated by the RSM approach in terms of R, $RRMSE$, PI, mean, and CoV.

Table 5 Statistical comparison performances of the developed PGA models.

Function	Proposed model	
	$PGA_{ANN-GSA}$	PGA_{RSM}
$RRMSE$ (%)	0.0967	0.1121
R	0.5770	0.3005
PI	0.0613	0.0862
MV	1.061	1.070
SD	0.093	0.111
CoV (%)	8.739	10.362

As shown in Figure 14, the high (PGA-act./PGA-pred.) ratio histogram frequency illustrates the high accuracy of the suggested equations in prediction, which means that a model has a respectable level of predictive accuracy if the ratio of actual to projected values is equal to (1.0).

In addition, ARE may be a preferred method for assessing the model prediction capabilities of the relative error distributions (Bagheri et al., 2012). From the percentage formula, ARE could be predicted as below:

$$ARE\% = \left|\frac{PGA_{act.}-PGA_{pred.}}{PGA_{act.}}\right| \times 100 \tag{23}$$

Superlatively, the frequency should decrease with every ARE% increasing. It is seen distinctly from Figure 15, where the proposed model has the lowest ARE of the highest frequency (less than 10%) and the largest ARE of the lowest frequency (more than 17.5%). Therefore, the two predicted PGA models have a very satisfactory error distribution.

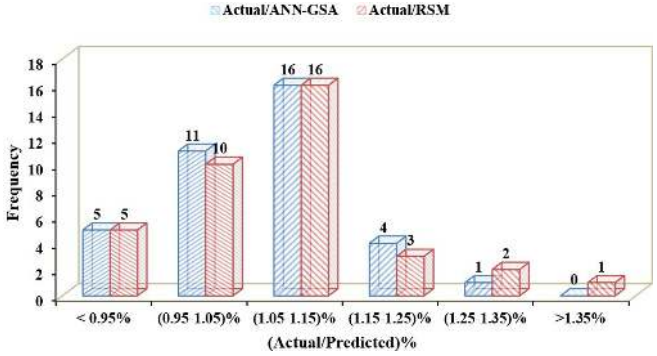


Figure 14: Frequency of the ratio between actual over predicted values.

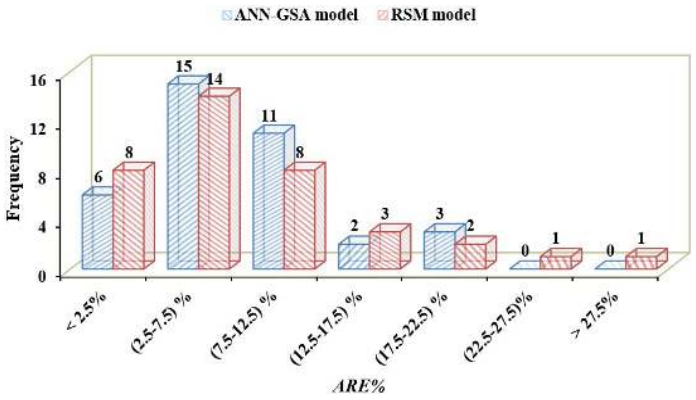
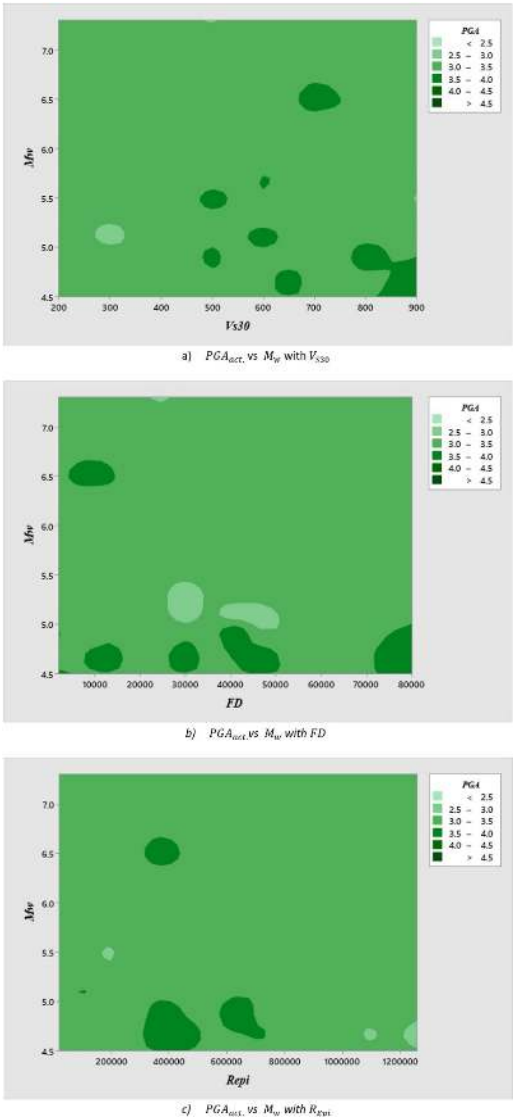


Figure 15: Distribution of the absolute relative error (ARE) for proposed models ($PGA_{ANN-GSA}$ and PGA_{RSM}).

4.5 Parametric analyses

To evaluate the impact of individual parameters on $PGA_{act.}$, Figure 16 (a-c) show the PGA as a function of a pair of parameters involving: M_w with V_{S30}, FD, and R_{Epi}, respectively. It can be observed that the increases in the amounts of individual variables (V_{S30} and FD) up to a specific range leading to an increase in PGA values, as depicted in Figure 16. By contrast, decreasing the $PGA_{act.}$ to less than 2.5 m/s^2 with increasing the R_{Epi}. Thus, the tested parametric analyses could directly enhance the assessment of $PGA_{act.}$ through selecting the suitable parameters.



a) $PGA_{act.}$ vs M_w with V_{S30}

b) $PGA_{act.}$ vs M_w with FD

c) $PGA_{act.}$ vs M_w with R_{Epi}

Figure 16: Contour plots.

The results of the current study were examined further in terms of the interaction plot for $PGA_{act.}$ response, as demonstrated in Figure 17. Different levels of one element may have different responses at different levels of another. As a result, there is an interplay between the various variables. Any point where the lines cross denotes a relationship between two variables. Non-parallel lines in the interaction plots (Figure 17 (a)) reveal a significant interaction between Mw and VS30. Similar outcomes have been indicated FD (as depicted in Figure 17 (b)). By contrast, slightly weak interaction between the REpi and the response, as presented in Figure 17 (c).

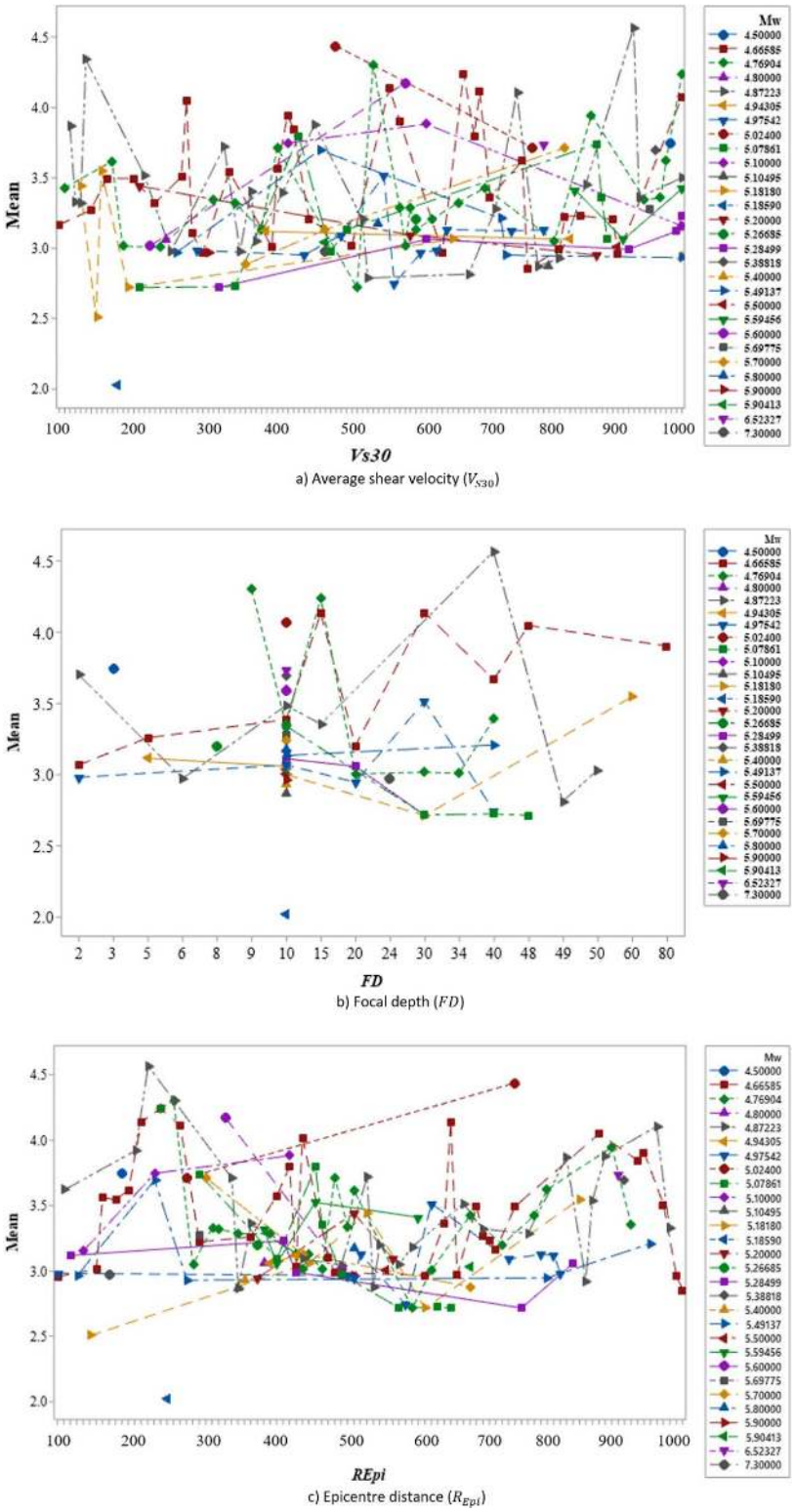
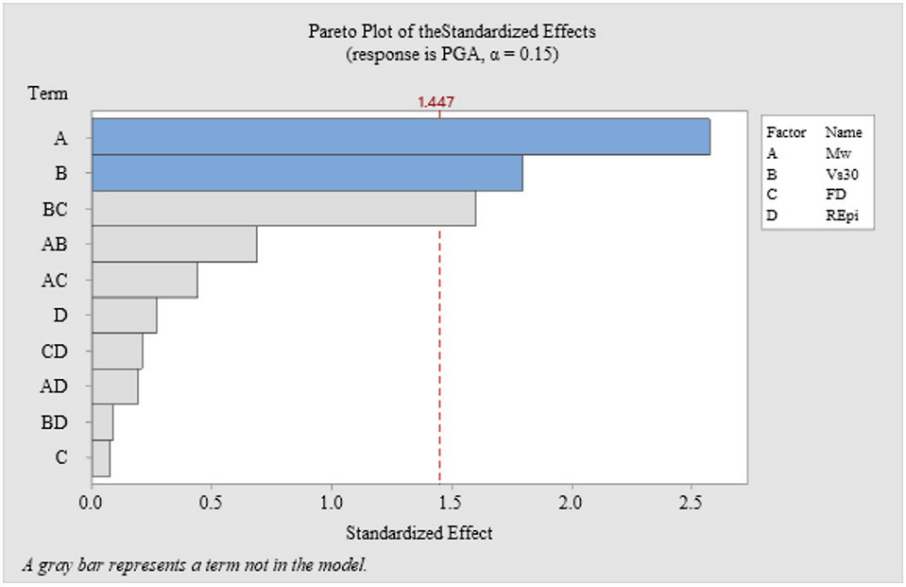


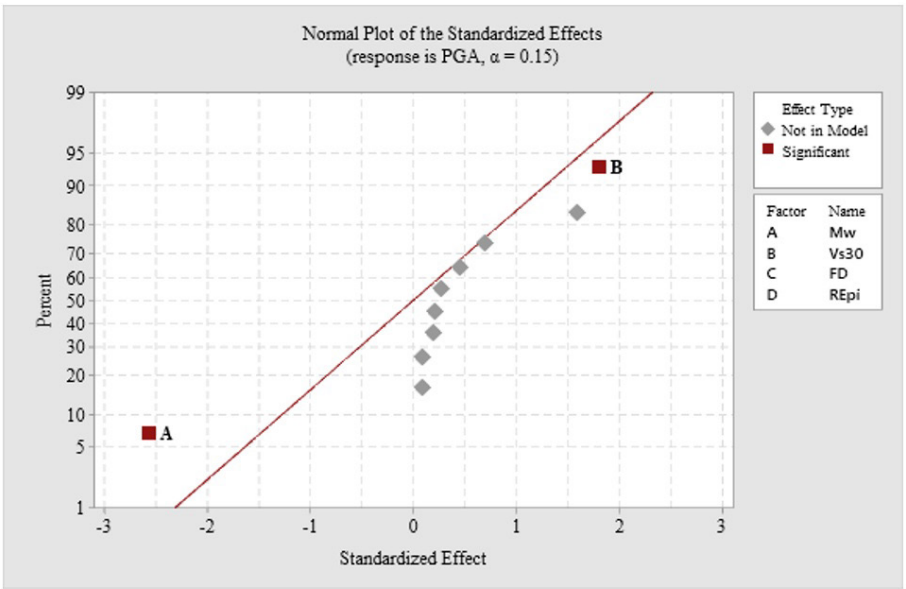
Figure 17: Plot of the interaction between M_w and the other parameters of the fitted response ($PGA_{act.}$).

4.6 Screening analysis

Screening analysis or analysis of a factorial design is of extreme significance for selecting the necessary input parameters. Screening analysis offers a valuable method for assessing and evaluating the contributions of every predictive parameter to the response. To achieve that, Minitab software is utilized to conduct the screening analysis by inputting the actual values of the response against other parameters. The findings of the screening analysis are shown in Figure18, employing a standardized diagram. This figure demonstrates that the significant parameters impacting the peak ground acceleration (PGA) value by order of the earthquake magnitude (M_w) and the average shear velocity (V_{S30}) followed byFD and R_{Epi}.



(a) Pareto Plot



(b) Normal Plot

Figure 18: The findings of the screening analysis

5 CONCLUSION

This research assesses and investigates PGA (peak ground acceleration) of Iraq's tectonic zones. Based on data from 2004-2020, a database of 187 historical records was analyzed to construct and forecast implicit and explicit models utilizing soft computing techniques (such as the composite- $ANN-GSA$ algorithm). This study would significantly contribute to the growing literature on PGA and seismic activity assessment. Besides, the contribution of this paper also

includes facilitating artificial intelligence for earthquake and hazard predictions. The following are the main findings of this study:

- The collected datasets can be utilized to construct a novel model for forecasting peak ground acceleration, especially for the Iraq tectonic region.
- Soft computing algorithms and Response Surface Methodology (RSM) are effective and practical tools and efficient techniques in engineering problems to provide an optimized solution with sufficient accuracy utilizing various parameters, which can be easily employed for predicting PGA values.
- The implicit and explicit models produced by this study are considered the first novel attempt to predict the PGA values for the Iraqi tectonic zone. Besides, models could be efficiently employed to provide estimates for the PGA values in a spreadsheet or hand calculations.
- The statistical examination resulted in a mean, SD, and CoV values of 1.061, 0.093, and 8.739%, respectively for composite model ($PGA_{GSA-ANN}$) and of about 1.07, 0.111, and 10.36%, consecutively for RSM model (PGA_{RSM}). Besides, minimum values are obtained for the absolute relative error ($ARE\%$) for both models. These statistical results indicate good solid accuracy and compatibility of the forecasts made by the soft computing algorithms model. The statistical results show that the RSM approach displays accuracy slightly compared to the GSAANN technique.
- The M_w and the V_{S30}. which reflect the $PGA's$ the most influential factor, and the standardized plot was determined to be in accord.
- Rarely do we come across data that can be utilized to predict the maximum PGA. As a result, further data should be used to assess and refine the proposed peak ground acceleration models and investigate a broader range of parameters. Consequently

One of the most prominent limitations facing the current paper is the use of data within the Iraq region. Therefore, the current study suggests generalizing the current models to a wider range that includes data for other countries. Further improvements may be made to obtain more accurate results, including employing the parameters used in this study for the related domain. Besides, extra data with a wide range of duration and parameters for Iraq's tectonic zones is recommended to compare with the finding obtained herein since this research is the first attempt to propose the Iraqi PGA formula.

Acknowledgements

The authors would like to express their gratitude for the support provided by their related organizations.

Author's Contributions: Conceptualization, Resources and Writing - original draft; Mahir M. Hason, Alaa Hussein Al-Zuhairi, and Ali A. Abdulhameed; Investigation, Writing – review & editing; Ammar N. Hanoon, Ali A. Abdulhameed and Imad S. Abbood; Supervision; Alaa Hussein Al-Zuhairi, Ammar N. Hanoon, and Ahmed W. Al Zand; Data curation and Visualization, Mahir M. Hason and Imad S. Abbood, Project administration; Ahmed W. Al Zand
Funding: The author(s) received no financial support for this research article, authorship, or publication.

Conflicts of Interest: The authors declare no conflict of interest.

Editor: Pablo Andrés Muñoz Rojas

References

Abbas, M. R., Hason, M. M., Ahmad, B. B., and Abbas, T. R. (2020). Surface roughness distribution map for Iraq using satellite data and GIS techniques. Arabian Journal of Geosciences, 13(17), 839.

Abd Alridha N and Jasem NA, (2013). Seismicity evaluation of central and southern Iraq. Iraqi Journal of Science. 54(4), 911-918.

Abdulnaby W, Mahdi H, Numan NMS and Al-Shukri H, (2014). Seismotectonics of the Bitlis–Zagros Fold and Thrust Belt in Northern Iraq and Surrounding Regions from Moment Tensor Analysis. Pure and Applied Geophysics. 171: (7), 1237-1250.

Abdulnaby W, Onur T, Gök R, Shakir AM, Mahdi H, Al-Shukri H, Numan NM, Abd NA, Chlaib HK and Ameen TH (2020) Probabilistic seismic hazard assessment for Iraq. Journal of Seismology.

Alavi AH and Gandomi AH, (2011). Prediction of principal ground-motion parameters using a hybrid method coupling artificial neural networks and simulated annealing. Computers and Structures. 89: (23-24), 2176-2194.

Alavi AH, Gandomi AH, Modaresnezhad M and Mousavi M, (2011). New ground-motion prediction equations using multi expression prograMming. Journal of Earthquake Engineering. 15: (4), 511-536.

Altman DG and Bland JM, (1995). Statistics notes: Absence of evidence is not evidence of absence. Bmj. 311: (7003), 485.

Al-Zuhairi, A. H., Al-Ahmed, A. H. A., Hanoon, A. N., & Abdulhameed, A. A. (2021). Structural behavior of reinforced hybrid concrete columns under biaxial loading. Latin American Journal of Solids and Structures, 18(6).

Antony J (2014). Design of experiments for engineers and scientists, Elsevier.

Bagheri M, Bagheri M, Gandomi AH and Golbraikh A, (2012). Simple yet accurate prediction method for sublimation enthalpies of organic contaminants using their molecular structure. Thermochimica acta. 543, 96-106.

Banyhussan QS, Hanoon AN, Al-Dahawi A, Yıldırım G and Abdulhameed AA, (2020). Development of gravitational search algorithm model for predicting packing density of cementitious pastes. Journal of Building Engineering. 27, 100946.

Box GE and Draper NR (1987). Empirical model-building and response surfaces, John Wiley & Sons.

Chaubey YP, (1993). Resampling-based multiple testing: Examples and methods for p-value adjustment. Technometrics. 35: (4), 450-451.

Chen, X.L., Fu, J.P., Yao, J.L. and Gan, J.F., (2018). Prediction of shear strength for squat RC walls using a hybrid ANN–PSO model. Engineering with Computers, 34(2), 367-383

Fouad SF and Sissakian VK, (2011). Tectonic and structural evolution of the Mesopotamia Plain. Iraqi Bulletin of Geology and Mining (4), 33-46.

Frank IE and Todeschini R, (1994). The Data Analysis Handbook. Amsterdam, Elsevier.

Gandomi AH, Alavi AH, Mousavi M and Tabatabaei SM, (2011). A hybrid computational approach to derive new ground-motion prediction equations. Engineering Applications of Artificial Intelligence. 24: (4), 717-732.

Gandomi M, Soltanpour M, Zolfaghari MR and Gandomi AH, (2016). Prediction of peak ground acceleration of Iran's tectonic regions using a hybrid soft computing technique. Geoscience Frontiers. 7: (1), 75-82.

Ghalib HA, Aleqabi GI, Ali BS, Saleh BI, Mahmood DS, Gupta IN, Wagner RA, Shore PJ, Mahmood A and Abdullah S, (2006). Seismic characteristics of northern Iraq and surrounding regions. Proceedings of the 28th Seismic Research Review: Ground-Based Nuclear Explosion Monitoring Technologies, 40-48.

Gholami M, Cai N and Brennan R, (2013). An artificial neural network approach to the problem of wireless sensors network localization. Robotics and Computer-Integrated Manufacturing. 29: (1), 96-109.

Güllü H and Erçelebi E, (2007). A neural network approach for attenuation relationships: An application using strong ground motion data from Turkey. Engineering Geology. 93: (3-4), 65-81.

Hanoon AN, Al Zand AW and Yaseen ZM, (2021). Designing new hybrid artificial intelligence model for CFST beam flexural performance prediction. Engineering with Computers. 1-27.

Hanoon AN, Jaafar M, Hejazi F and Abdul Aziz FN, (2016). Energy absorption evaluation of reinforced concrete beams under various loading rates based on particle swarm optimization technique. Engineering Optimization, 1-19.

Hanoon AN, Jaafar M, Hejazi F and Abdul Aziz FN, (2017a). Energy absorption evaluation of reinforced concrete beams under various loading rates based on particle swarm optimization technique. Engineering Optimization. 49: (9), 1483-1501.

Hanoon AN, Jaafar M, Hejazi F and Aziz FNA, (2017b). Strut-and-tie model for externally bonded CFRP-strengthened reinforced concrete deep beams based on particle swarm optimization algorithm: CFRP debonding and rupture. Construction and Building Materials. 147, 428-447.

Hasan G, Al Kubaisy MA, Nahhas FH, Ali AA, Othman N and Hason MM, (2014). Sulfur Dioxide (SO^ sub 2^) Monitoring Over Kirkuk City Using Remote Sensing Data. Journal of Civil & Environmental Engineering. 4(5), 1.

Hason, M. M., Hanoon, A. N., & Abdulhameed, A. A. (2021). Particle swarm optimization technique based prediction of peak ground acceleration of Iraq's tectonic regions. Journal of King Saud University - Engineering Sciences.

Hason MM, Hanoon AN, Al Zand AW, Abdulhameed AA and Al-Sulttani AO, (2020). Torsional Strengthening of Reinforced Concrete Beams with Externally-Bonded Fibre Reinforced Polymer: An Energy Absorption Evaluation. Civil Engineering Journal. 6, 69-85.

Jang H and Topal E, (2014). A review of soft computing technology applications in several mining problems. Applied Soft Computing. 22, 638-651.

Kaveh, A. and Talatahari, S., (2012). A hybrid CSS and PSO algorithm for optimal design of structures. Structural Engineering and Mechanics, 42: (6), 783-797.

Kukolj D and Levi E, (2004). Identification of complex systems based on neural and Takagi-Sugeno fuzzy model. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics). 34: (1), 272-282.

Luco N and Cornell CA, (2007). Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. Earthquake Spectra. 23: (2), 357-392.

Mahmod HM, Aznieta AFN and Gatea SJ, (2017). Evaluation of rubberized fibre mortar exposed to elevated temperature using destructive and non-destructive testing. KSCE Journal of Civil Engineering. 21: (4), 1347-1358.

Montgomery DC (2017). Design and analysis of experiments, John wiley & sons.

Onur T, Gök R, Abdulnaby W, Mahdi H, Numan NM, Al-Shukri H, Shakir AM, Chlaib HK, Ameen TH and Abd NA, (2017). A Comprehensive Earthquake Catalogue for Iraq in Terms of Moment Magnitude. Seismological Research Letters. 88: (3), 798-811.

Payal A, Rai CS and Reddy BVR, (2014). Artificial Neural Networks for developing localization framework in Wireless Sensor Networks. 2014 International Conference on Data Mining and Intelligent Computing (ICDMIC).

Pimentel-Gomes F, (2000). Course of experimental statistics. Piracicaba: FEALQ. 15.

Rashedi E, Nezamabadi-pour H and Saryazdi S, (2009). GSA: A Gravitational Search Algorithm. Information Sciences. 179: (13), 2232-2248.

Robinson C and Schumacker RE, (2009). Interaction effects: centering, variance inflation factor, and interpretation issues. Multiple linear regression viewpoints. 35: (1), 6-11.

Shafiqu, Q. S. A.-D. M., & Sa'ur, R. H. (2016). Data Base for Dynamic Soil Properties of Seismic Active Zones in Iraq. Journal of Engineering, 22(7), 1–18.

Schutz B, (2003). Gravity from the Ground Up: An Introductory Guide to Gravity and General Relativity. Cambridge, Cambridge University Press.

Winship C and Zhuo X, (2020). Interpreting t-Statistics under Publication Bias: Rough Rules of Thumb. Journal of Quantitative Criminology. 36: (2), 329-346.

Yaseen, Z.M., Tran, M.T., Kim, S., Bakhshpoori, T. and Deo, R.C., (2018). Shear strength prediction of steel fiber reinforced concrete beam using hybrid intelligence models: a new approach. Engineering Structures, 177, 244-255.