

Research Article

Attenuation Relationships for the Horizontal Component of Peak Ground Acceleration (PGA_H) Using Gene Expression Programming (GEP) and Group Method of Data Handling (GMDH)

Mehrasa Ajam , Mohammad Shamekhi Amiri , and Hossein Pahlavan 

Faculty of Civil Engineering, Shahrood University of Technology, Shahrood, Iran

Correspondence should be addressed to Mohammad Shamekhi Amiri; shamekhi@shahroodut.ac.ir

Received 27 June 2022; Revised 6 December 2022; Accepted 13 April 2023; Published 4 December 2023

Academic Editor: Angelo Marcelo Tusset

Copyright © 2023 Mehrasa Ajam et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Earthquake is one of the natural disasters that has always influenced human life. It is not currently possible to predict exactly when and where an earthquake will occur, nor how large it will be. It is also impossible to prevent an earthquake. However, by designing seismic-resistant structures, the amount of financial losses and casualties can be reduced. This resistant design requires the use of earthquake risk analysis. By using the earthquake risk analysis, it will be possible to estimate the parameters of the strong ground motion, including acceleration, velocity, and displacement in each area. Estimating the parameters of strong ground motion will be possible just by obtaining the appropriate attenuation relationship. The aim of this paper is to present an appropriate attenuation relationship to estimate the horizontal component of the possible occurrence of peak ground acceleration in each region. Two methods were used to calculate attenuation relationship: gene expression programming (GEP) and group method of data handling (GMDH). In the first step, an up-to-date and comprehensive catalog consisting of 1,185 earthquake records that occurred around the world has been prepared. In the next step, the parameters of magnitude, hypocentral distance, and shear wave velocity of these records have been used as variables of the attenuation relationship. Then, the fitness function (f) was determined, and attenuation relationships were calculated using GEP and GMDH. The amount of fitness function (f) was obtained 766.12 from 1,000 in the GEP method and 767.77 from 1,000 in the GMDH method. The values of the fitness function, residuals and comparison plots showed a high-agreement between “the values predicted by the attenuation relationships” and “the actual values observed in the earthquakes.” Finally, according to the results of this research, it can be said that the use of GEP and GMDH methods has provided better results than the other similar researches. Also, the use of up-to-date records makes the results of this research more reliable than the previous researches.

1. Introduction

Human life has always been influenced by the natural disasters such as floods, storms, and earthquakes. These disasters have always threatened human life and property. The occurrence of these natural disasters will never stop and posterity will always be threatened. One of the most important and unpredictable events is an earthquake. Many earthquakes usually occur annually in the world. Some of them occur in vacant and barren areas, which are certainly not considered a threat to humans. Some earthquakes are so mild that they are not even felt by humans and are only recorded by the accelerometers. However, high-intensity

earthquakes may occur in residential environments and pose a constant threat to humans.

Japan, the United States, China, India, New Zealand, and Iran are considered as the seismic countries in the world. In Japan, from 1945 to 1995, 14 earthquakes with a magnitude greater than 6.5-Richter happened in 50 years that, left more than 8,000 victims. The Fukushima earthquake (2011) with a moment magnitude about 9, is one of the strongest earthquakes in Japan since 1900. The record for the largest earthquake in history with a moment magnitude 9.5 also belongs to Chile, which occurred in 1960 [1].

Earthquakes had been considered an unknown phenomenon several centuries ago. However, today scientific advances

made in the field of seismic geotechnical engineering and more knowledge of the internal structure of the earth have led to the understanding of the various phases of the earthquake and its effects on the earth's movement. Despite the achieved progress, it is impossible to predict its occurrence anywhere in the world, but it is possible to prepare for it by building seismic-resistant structures. The purpose of designing earthquake resistant structures is to construct structures that withstand seismic loads, do not undergo a lot of damage and rescue the inhabitants. To achieve this goal, it is necessary to predict and estimate the severity of possible earthquakes in each area following its various features. This goal is possible just through risk analysis. The prerequisite of risk analysis studies is to identify the parameters of strong ground motion for different areas [2].

The amount of strong ground motion is attributed to the magnitude, source to site distance, the source mechanism (type of faulting), geology of the region, surface topography, and dynamic properties of the material propagation [3]. Also, other parameters including soil nonlinear behavior, directivity, rupture propagation, and basin effects on ground motion can be even more effective than some of the above mentioned independent parameters [4].

In order to determine the impact of various earthquake parameters on a structure or a site located at a distance from the earthquake center, attenuation relationships are used.

Attenuation relationships are usually presented for important earthquake characteristics such as peak horizontal and vertical acceleration, peak ground velocity (PGV), elastic response spectra, and inelastic response spectra. The largest number of attenuation relationships is presented for peak ground acceleration (PGA) estimation.

The purpose of the present study is to present a new global attenuation relationship to estimate the horizontal component of peak ground acceleration (PGAH) using the two methods of gene expression programming (GEP) and group method of data handling (GMDH) and compare the results.

2. Overview of the Attenuation Relationships

Attenuation relations were first introduced in the 1960s by Neumann. Neumann [5] provided the attenuation model of PGA in 1954 for earthquakes in America. The only parameter involved was the distance of the accelerograph from the earthquake's center. In Newman's attenuation relationship, even the role of earthquake magnitude was ignored. Subsequently, models were developed from the 1970s to mid-1980s by Milne and Davenport [6], Esteva [7], Denham and Small [8], Donovan [9], Esteva and Villaverde [10], Orphal and Lahoud [11]. In these models, in addition to the distance of the accelerometer from the earthquake center, the magnitude of the earthquake was also included in the calculations. However, the strong role of soils in providing these relationships and the amount of damage was still ignored. Following the damage caused by the 1964 earthquakes in Niigata, Japan, and Alaska, progress was made in expanding the effective parameters in the attenuation relationships. Since 1976, a

new parameter indicating soil type was added to the attenuation relationship [12]. After that, many attenuation models were presented by various researchers for the different regions, some of them are explained below.

In 1981, Joyner and Boore [13] proposed a new attenuation relationship based on the North American earthquakes. The catalog consisted of earthquakes with surface wave magnitude (M_S) from 5 to 7.7 and records were taken from stations less than 370 km from the epicenter. Also, soil type has been effective in estimating acceleration as a coefficient for both soil and rock [13].

The Atkinson and Boore [14] attenuation model was also presented in 1990 based on North American earthquakes with a moment magnitude (M_W) between 5 and 7, which occurred at a distance of 10–100 km of earthquake accelerometers.

Sarma and Free [15] presented Equation 1 in 1995

$$\log(\text{PGA}) = -3.436 + 0.8532M - 0.0192M^2 - (0.9011 \times \log(R)) - 0.002R - 0.0316S, \quad (1)$$

where M , R , and PGA are the magnitude, the hypocentral distance (km) and PGA (cm/s^2), respectively. Also, the variable S in this relation is "zero" for the rock and "one" for the soil.

Following advances in technology and computer science in the early 21st century, intelligent methods have emerged and gradually replaced traditional methods. These changes led to the creation of more complex models in attenuation relationships. In 2007, Sobhaninejad et al. [16] used the intelligent genetic algorithm (GA) method to provide an attenuation relationship. Subsequently, Cabalar and Cevik [17] introduced a new attenuation relationship for Turkey in 2009 using the same method. In 2010, Ornthammarath et al. [18] also introduced a new relationship in which the shear wave velocity was also involved in this relationship and is shown in Equation (2).

$$\log(\text{PGA}) = -2.622 + 0.643M_W - 1.249 \log\left(\sqrt{r_{jb}^2 + 3.19^2}\right) + 0.344S_s, \quad (2)$$

where M_w is the moment magnitude, r_{jb} is Joyner–Boore distance (km), and PGA is the peak ground acceleration (cm/s^2). S_s that specifies the type of soil is selected according to the value of shear wave velocity as Equation 3.

$$\begin{aligned} 360(\text{m/s}) \leq V_{S_{30}} \leq 750(\text{m/s}) &\longrightarrow S_s = 1 \\ 750(\text{m/s}) < V_{S_{30}} &\longrightarrow S_s = 0. \end{aligned} \quad (3)$$

Other researchers who have used this method to present attenuation relationships including Kermani et al. [19], Jafarian et al. [20], Alavi and Gandomi [21], Ghodrati et al. [22, 23], and Gandomi et al. [24]. The relationship presented by Alavi and Gandomi [21] in 2011 is as Equation (4).

$$\begin{aligned} \ln(\text{PGA}) = & 6 - \ln(R_{\text{ClstD}}) + \frac{-1}{6} (\ln(R_{\text{ClstD}}))^2 \\ & + \frac{M_w}{6} \ln(R_{\text{ClstD}}) + \frac{6}{V_{s,30}} + (\ln(R_{\text{ClstD}}))^3 (V_{s,30})^{-1} \\ & + 64 \times \left[-8 - V_{s,30} + (5 \times V_{s,30} \sin(\lambda)) + \frac{V_{s,30}}{M_w - 7} \right]^{-1} \end{aligned} \quad (4)$$

In this relation, M_w is the moment magnitude, R_{ClstD} is the closest distance to the rupture surface (km), $V_{s,30}$ is the average shear-wave velocity over the top 30 m of the site (m/s), and PGA is the peak ground acceleration (cm/s^2).

One of the most recent attenuation relations was offered by Kumar et al. [25]. The relationship that was determined for the northeastern Himalaya region is presented as Equation (5):

$$\log(\text{PGA}) = -1.497 + 0.3882M - 1.19\log(R + e^{0.2876M}), \quad (5)$$

where M is the magnitude, R is the hypocentral distance (km), and PGA is the peak ground acceleration (cm/s^2). This relationship is based on the earthquakes that occurred in the northeastern part of the Himalayas with a magnitude range of 4–6.8.

In 2018, Javan-Emrooz et al. [26] presented Equation (6):

$$\begin{aligned} \log(\text{PGA}) = & \log \left[\frac{M}{\sqrt{\log(e^{s^{a3}} + e^M - R - (1/F))}} \right] \\ & + \ln \left[\frac{M+1}{RM} \right] + \ln(\log(\sqrt{M})) + 3. \end{aligned} \quad (6)$$

In this relation, M and R are the moment magnitude and the hypocentral distance (km), respectively. Also, PGA is the peak ground acceleration in centimeters per square second (cm/s^2). This attenuation model is based on earthquakes in northern Iran with a range of moment magnitude (M_w) of 4.5–7.4 and a hypocentral distance of 2–100 km.

Also, Erken et al. [27] provided attenuation relation equations by using nonlinear regression analysis in Northwest Anatolia region for two different site conditions.

Karimi-Ghalehjouh and Mahinroosta [28] presented attenuation relationship of horizontal and vertical PGA using regression analysis fuzzy logic model in Iranian plateau. In this study, three different site conditions were considered: rock, stiff soil, and soft soil [28].

Shiuly et al. [29] in 2020 presented a new model for prediction of PGA by using artificial neural network (ANN) and GA in Himalayan region for earthquake data recorded in rock sites.

One of the most recent suggested relationships is the Abdelfattah et al. [30] model, which was presented in 2021 as Equation (7).

$$\log(\text{PGA}) = -1.36 + 0.85M_L - 0.85\log(r) - 0.005r, \quad (7)$$

where M_L is the local magnitude, R is the hypocentral distance (km), and PGA is the peak ground acceleration (cm/s^2).

Also, in 2021 Kumar et al. [31] offered an attenuation relationship using regression based on the earthquakes that occurred in the central Himalayas with a magnitude range of 5–6.8.

Pourzeynali and Khadivyan [32] have provided attenuation relationship of horizontal and vertical PGA using regression analysis for Alborz zone of Iran.

3. Data Collection

The first step to present the attenuation relationship is to prepare a catalog of occurred earthquakes in a sufficient number and with high accuracy. In the present study, considering that the final relationship is global, records from all over the world were used and were chosen completely randomly and with appropriate dispersion. These records are obtained from the PEER (Public Employees for Environmental Responsibility (<https://ngawest2.berkeley.edu/>)) website by using the possibility of filtering the information of records such as earthquake record number, earthquake name, earthquake recording station name, earthquake occurrence date, earthquake magnitude, fault distance, rupture radius, shear wave velocity, etc. which can be used from one or more of them. In this study, only the filters of magnitude and distance were used, then 1,218 records were found, which include earthquakes that occurred continuously from 1938 onwards.

As mentioned before, in this research, the horizontal component of the PGA is considered the target. Also, magnitude, distance, and the shear wave velocity were considered as variables of the seismic attenuation model. In the following, each of these parameters is described.

3.1. Magnitude (M_w). The magnitude of the earthquake indicates the amount of energy released by the earthquake and is directly related to it. Therefore, it will have the greatest impact on the amount of strong ground motion and it can be considered as the most important and effective parameter in determining the attenuation relationships. Magnitude has a direct effect on the horizontal component of PGA. The higher magnitude of earthquake occurrence will cause higher parameters of strong ground motion at a certain distance from the source of the earthquake.

Magnitude is expressed in different scales due to the variety of recorded seismic waves and the institutions and devices recording these earthquakes. The magnitude types are local magnitude (M_L), surface-wave magnitude (M_S), body-wave magnitude (m_b), and moment magnitude (M_w). Local magnitude (M_L) based on amplitude recorded by Wood Anderson seismograph, surface-wave magnitude (M_S) based on Riley wave amplitude, and body-wave magnitude (m_b) based on P-wave amplitude are measured. According to the saturation effects, although the energy released during an earthquake increases, the seismic properties do

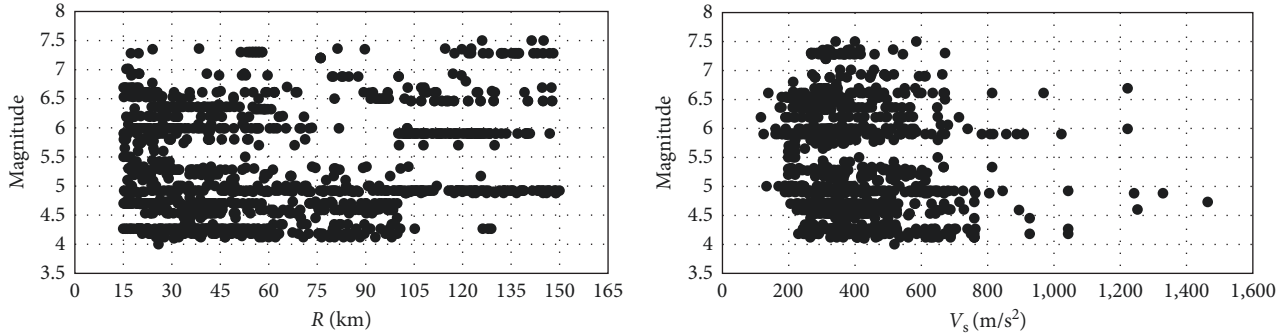


FIGURE 1: Magnitude distribution diagrams in terms of distance and shear wave velocity.

not necessarily increase at the same rate. Therefore, these three scales do not have high accuracy in large earthquakes [14]. The moment magnitude is independent of the characteristics of the earth's motion and can accurately define the magnitude of earthquakes. Therefore, in this research, moment magnitude has been used. To prepare the earthquake catalog in this research, earthquakes with a moment magnitude of 4–7.5 have been used because earthquakes with a magnitude of less than 4 are not considered dangerous for construction engineering structures.

3.2. Distance (R). The distance indicates the path that seismic waves pass from the source of the earthquake to the site. It is the second effective parameter in the attenuation relationships. The distance has an inverse effect on the parameters of the strong ground motion, so if the distance from the site to the epicenter is longer, the PGA will be smaller. The distance in seismology is expressed with different types, which types are: the shortest horizontal distance to the vertical rupture image (r_{jb}), the shortest distance to the rupture surface (r_{rup}), the shortest distance to the seismic rupture surface (r_{seis}), and distance from the epicenter (r_{hypo}) [23].

In this study, the path taken by the waves from the source of the earthquake to the desired site is the most effective parameter on the ground motion. The used catalog data have a hypocentral distance (r_{hypo}) from 15 to 150 km.

3.3. Shear Wave Velocity (V_s). Although most of the path, passed by the waves, is in the bedrock, these waves often pass through a layer of soil at the end of the path. This layer of soil can play a major and significant role in changing the frequency of waves reached to the surface of the ground. It can also cause the damping of some frequencies or intensify other frequencies and change the magnitude of the earthquake. Therefore, this is necessary to consider the ground type and tectonic conditions in providing attenuation relationships [3]. Substrate properties can be entered into computations in both qualitative and quantitative forms. If the effect of this parameter is considered qualitatively, based on the shear wave values, the ground type is divided into several categories and the effect of the ground type is entered as constant coefficients in the calculation of attenuation relationships. If the soil effect is considered quantitatively, the value of shear wave velocity will be used in the relations. According to the ability of the methods used in this research,

the effect of the soil type parameter has been applied quantitatively and using the value of shear wave velocity on the attenuation relationships. Obviously, this approach has increased the accuracy of calculations.

3.4. Preparing the Final Catalog. In the first step, data have been collected and a preliminary catalog has been prepared to consist of magnitude, distance, and shear wave velocity of the 1,218 earthquake records. In the next step, all the records were processed using Seismosignal software and considering the appropriate correction frequency range of 0.2–30 [33]. Then the peak acceleration of each recorded earthquake at each station was taken. The important point in this section is that there are three components of acceleration in three orthogonal directions per record, but in the present study, whereas the study was done on the horizontal component of acceleration, the vertical component was set aside. Then for the horizontal components were considered to be the most critical case and for this purpose, square root of the sum of the squares (SRSS) of two orthogonal components was used.

After this stage, due to error possibility in recording and extracting the records, the records were corrected through sorting in an incremental magnitude. In the next step, the graph was drawn in terms of the horizontal component of PGA. Then the points of the graph that have relatively large dispersion were removed from the earthquake catalog. After this step, 1,185 records remain as the final records that are involved in the attenuation relationship. Magnitude distribution diagrams in terms of distance and shear wave velocity in Figure 1 and the number of records in terms of magnitude and distance in Figure 2 are shown.

4. Methods for Estimation of the Attenuation Relationships in This Research

4.1. Gene Expression Programming (GEP). GEP is a developed GA and genetic programming (GP) that was proposed by Ferreira [34] in 1999. It can be said that this algorithm, like GA and GP, is a GA that uses people as a population, choosing is based on the suitability and introduction of the genetic diversity using one or more genetic operators.

The main elements in GEP are just two elements, chromosomes and expression trees. The structure of chromosomes is organized in such a way that it allows the creation of multiple genes, and each of them is identified as a

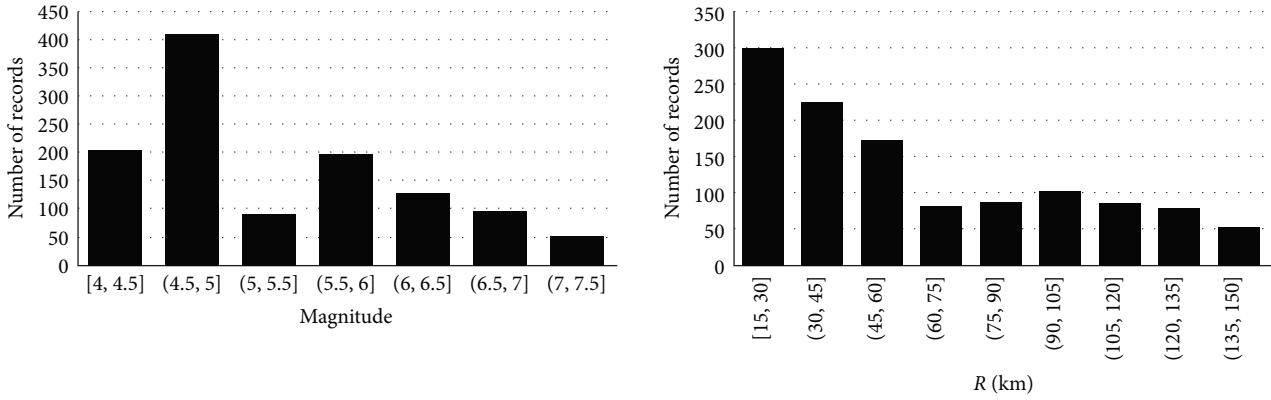


FIGURE 2: Number of records in terms of magnitude and distance.

subexpression tree. The structure of genes is also designed so that each gene has a head and a tail, and this structural and functional organization of the gene in the GEP always ensures the production of valid programs, and it is not important how much they change the depth of the chromosome. Expression trees are also genetic information encoded in chromosomes. As in nature, the process of decoding information is called translation, and this translation clearly indicates the type of code and set of rules. The genetic code is also very simple and shows the one-to-one relationship between chromosome symbols and functions or terminals. The rules are also very simple, they determine the organization consisting of functions and terminals in the development trees and the type of interaction between subexpression trees [35, 36].

Figure 3 shows the steps of the GEP and the operators defined in this algorithm are also shown in Table 1.

4.2. Group Method of Data Handling (GMDH). The GMDH is used to identify the behavior of nonlinear systems. Among all intelligent computational methods, the GMDH method is known as a self-organized system with a very complex nonlinear problem solving capacity [37, 38].

The GMDH was proposed by Ivakhnenko [39] in 1971. In this algorithm, by keeping the input variables in a flexible network, a connection is made between the input variables and the output, and the possibility of initial estimation of the output by regression equations including small subsets of input variables is provided. Each of these small subsets contains 2–3 independent variables [40, 41]. The main purpose of this method is to build a network based on a quadratic transfer function. The number of hidden layers and neurons, effective input variables, and the optimal structural model is automatically determined in this algorithm [42].

The communication between the input and output variables is done by the data management neural network in a group method which is a nonlinear function called Volterra series, which is as Equation (8).

$$\hat{y} = a_0 + \sum_{i=1}^m a_i x_i + \sum_{i=1}^m \sum_{j=1}^m a_{ij} x_i x_j + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m x_j a_{ijk} x_i x_j x_k + \dots \quad (8)$$

The Volterra series is processed as a quadratic polynomial by utilizing the following relation (Equation 9).

$$G(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i^2 + a_4 x_j^2 + a_5 x_i x_j. \quad (9)$$

The purpose of the data management algorithm is to find fixed and unknown coefficients a_i in a group method, which will be obtained by regression methods, based on obtaining the least squares of error for each pair of input variables x_i and x_j [43, 44].

5. Results

In this research, 80% of the records were considered as training records and the remaining 20% as testing records. The method of selecting these records is that the testing records are in the range of training records and allow comparison. As mentioned, in this study, the parameters of magnitude, hypocentral distance, and shear wave velocity are defined as variables and the PGA as a target function.

5.1. Results of Gene Expression Programming (GEP). Finally, after several trials, the best answer with the highest accuracy and least error was considered as the final answer. The final relationship chromosomes are shown in Figure 4. In this relation, $d_0 = M$, $d_1 = R$, $d_2 = V$, and the values of the coefficients c_0 and c_1 for each chromosome are given after that. Also, the links of chromosomes are sum operator (+). By replacing the variables, mathematical operators and the link operator (+) in the calculated equation, the global attenuation relationship for the PGA was obtained as Equation (10).

$$\log(\text{PGA}) = \frac{(-1.553V - R^2 - 9.095R)}{R \times \text{Exp}(M)} + \sqrt[6]{\text{Exp}\left(\frac{5.61 - R}{M - 1.37}\right)} + \sqrt[3]{9.999 - \sqrt[3]{V/M + M + 8.967}}. \quad (10)$$

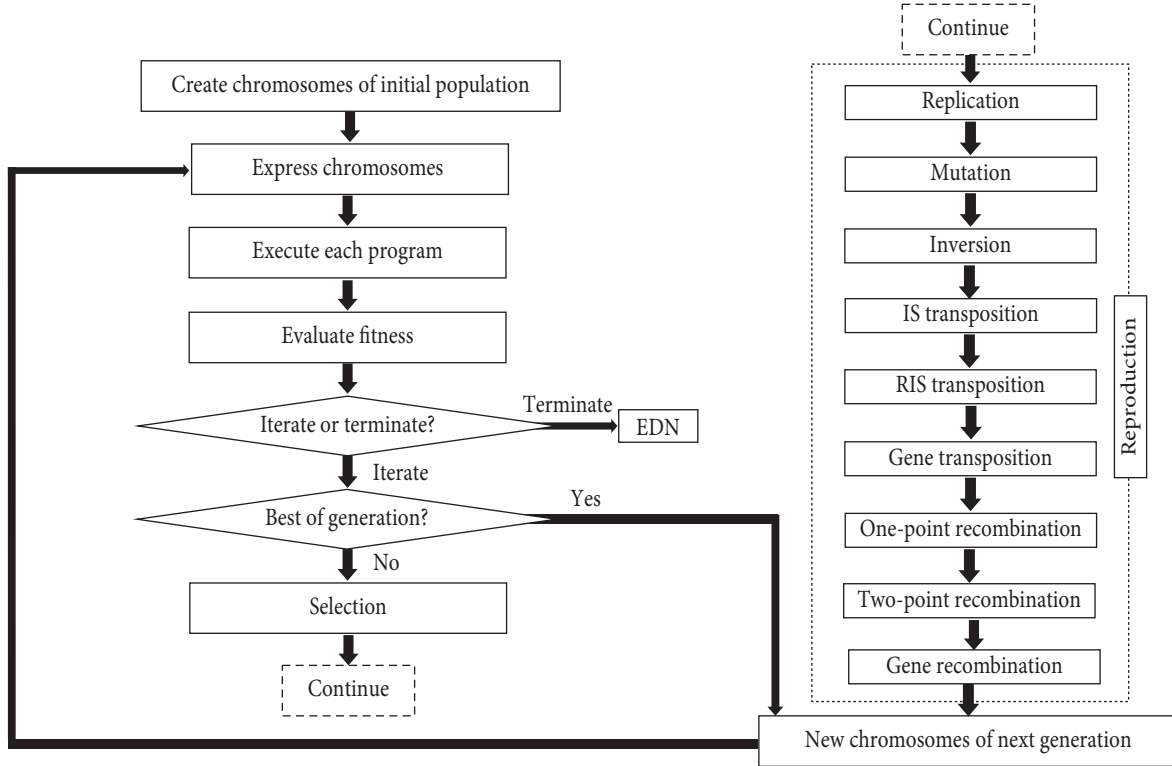


FIGURE 3: The flowchart of gene expression programming [36].

TABLE 1: Operators defined in gene expression programming.

Name	Representation	Weight
Addition	+	4
Subtraction	-	4
Multiplication	*	4
Division	/	1
Square root	Sqrt	1
Exponential	Exp	1
Natural logarithm	Ln	1
x to the power of 2	X2	1
x to the power of 3	X3	1
Cube root	3Rt	1

In this relationship, M is the magnitude, R is the hypocentral distance (km), V is the shear wave velocity (m/s), and PGA is peak ground acceleration (cm/s^2). For better visualization of the results, PGA diagrams were drawn in terms of magnitude and hypocentral distance.

$$\begin{aligned}
 / - E / - d0 * d1 d1 c1 d2 c0 & \quad c0 = -1.553, c1 = 9.095 \\
 S 3 E / - + c0 d1 c1 d0 & \quad c0 = 5.61, c1 = -1.37 \\
 3 - c0 3 + / - d2 d0 d0 c1 & \quad c0 = 9.999, c1 = -8.967
 \end{aligned}$$

FIGURE 4: Expression trees of attenuation relationship of the peak ground acceleration.

Figure 5 shows the PGA variation curve in terms of hypocentral distance (15–150 km) for three magnitudes of 4.5, 5.5, and 6.5. In this diagram, the shear wave velocity is considered 400 m/s. As can be seen, the PGA decreases nonlinearly as it moves away from the epicenter. Figure 6 also shows the PGA variation curve in terms of magnitude (4–7.5) for five hypocentral distances of 20, 30, 60, 100, and 150 km. In this diagram, the shear wave velocity is considered 400 m/s. As can be seen, with increasing magnitude, the PGA increases nonlinearly.

5.2. Results of Group Method of Data Handling (GMDH). The final attenuation relationship obtained from the GMDH method is shown in Equation (11).

$$\begin{aligned}
 \log(\text{PGA}) = & -1788.07 + (3192.56\sqrt[3]{M}) + (-7.7484\sqrt[3]{R}) + (-1664.47\sqrt[3]{M^2}) \\
 & + (-6.97 \times 10^{-5} \times V\sqrt[3]{V}) + (5.715\sqrt[3]{M}\sqrt[3]{R}) + (170.807M\sqrt[3]{M}) \\
 & + (-10.447M^2) + (3.096 \times 10^{-6} \times RV) + (-0.505M\sqrt[3]{R}).
 \end{aligned} \tag{11}$$

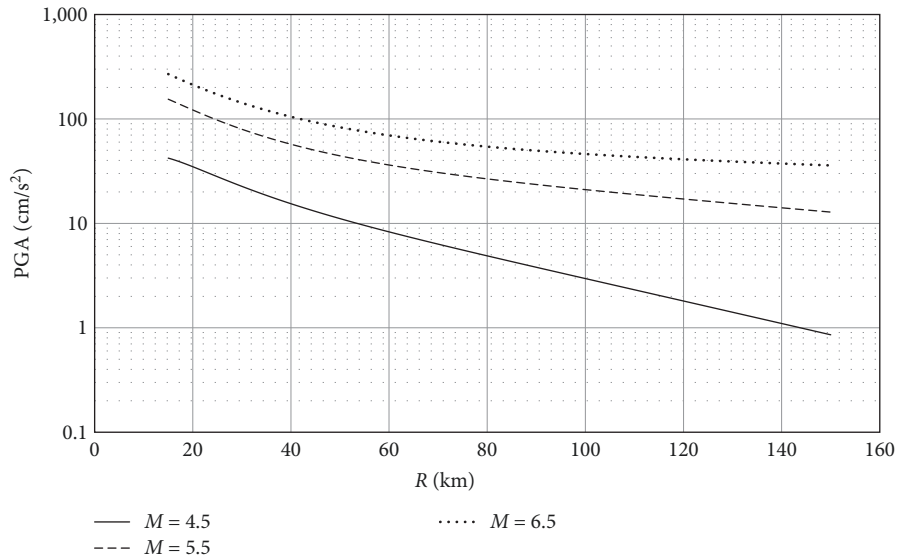


FIGURE 5: The peak ground acceleration variation curve in terms of hypocentral distance.

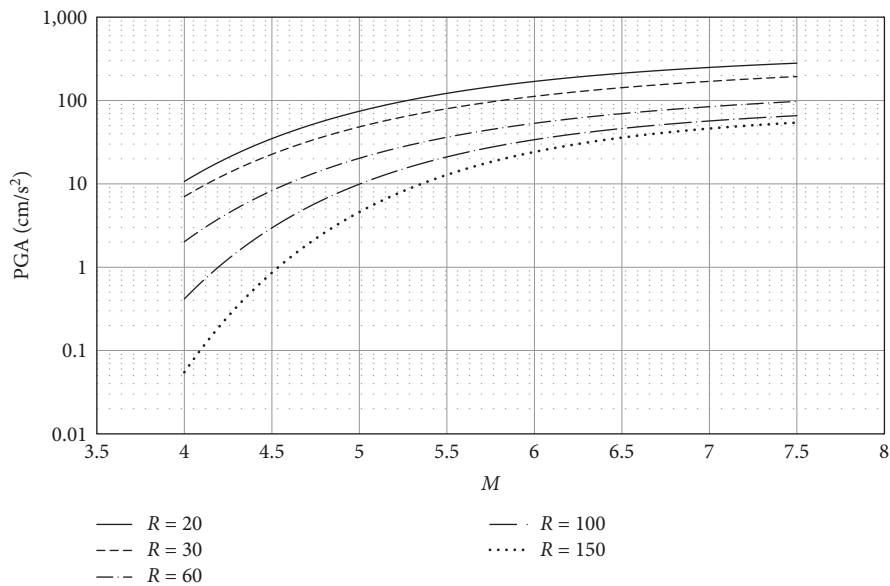


FIGURE 6: The peak ground acceleration variation curve in terms of magnitude.

In this relationship, M is the magnitude, R is the hypocentral distance (km), V is the shear wave velocity (m/s), and PGA is peak ground acceleration (cm/s^2).

Figures 7 and 8 show the changes in PGA with the attenuation relationship calculated by the GMDH method in terms of magnitude and hypocentral distance.

Figure 7 shows the PGA variation curve in terms of hypocentral distance (15–150 km) for 3 magnitudes of 4.5, 5.5, and 6.5. In this diagram, the shear wave velocity is considered 400 m/s. As can be seen, the PGA decreases nonlinearly as it moves away from the epicenter. Figure 8 also shows the PGA variation curve in terms of magnitude (4–7.5) for five hypocentral distances of 20, 30, 60, 100,

and 150 km. In this diagram, the shear wave velocity is considered 400 m/s. As can be seen, while magnitude increases, the PGA increases nonlinearly.

In Figures 5–8, it can be seen that in addition to the nonlinearity of the diagrams, the lines of each diagram do not follow a pattern. While in the past researches, usually the lines in each diagram had a single pattern. The reason, is due to the methods used in this research. In the past, a fixed formula was used to calculate attenuation relations, which only considered fixed coefficients as variables for different regions. But in the GEP and GMDH methods, due to the nonlinear changes of the chromosomes and the terms of the equation, the patterns are not assumed to be constant.

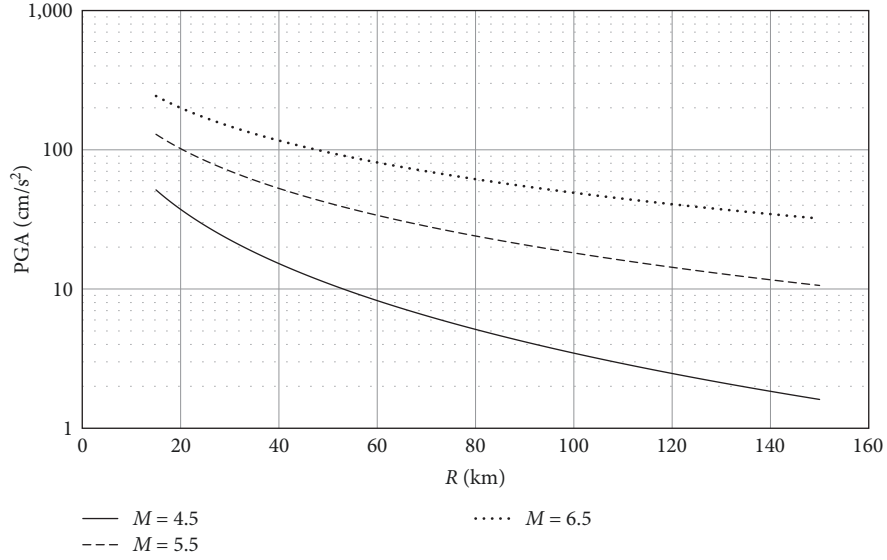


FIGURE 7: The peak ground acceleration variation curve in terms of hypocentral distance for three magnitudes of 4.5, 5.5, and 6.5.

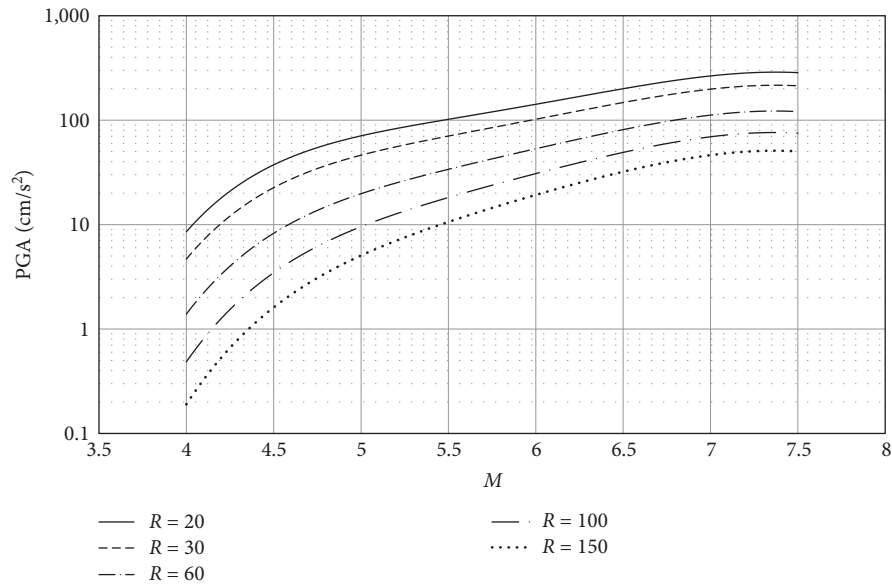


FIGURE 8: The peak ground acceleration variation curve in terms of magnitude for five hypocentral distances of 20, 30, 60, 100, and 150 km.

5.3. Verification and Comparison of Results. As mentioned before, the attenuation relationships obtained from GEP and GMDH methods have high accuracy. The accuracy of each of the relationships presented in this research can be measured with the parameters defined below.

One of these parameters is the root-mean-square error (RMSE), which can be evaluated and calculated by Equation (12).

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\text{PGA}_{\text{actual}_i} - \text{PGA}_{\text{model}_i})^2}. \quad (12)$$

In the above relation, $\text{PGA}_{\text{actual}}$ is the PGAH values of the target and the $\text{PGA}_{\text{model}}$ is the PGAH values obtained from

the final attenuation relationship presented in this article. It is clear when the predicted values and the target values are equal, the RMSE will be zero, and as this number increases, it indicates more error and less compliance between the target value and the predicted value.

Now by using the RMSE value, the value of the fitness function (f) is calculated from Equation (13).

$$f = \frac{1,000}{1 + \text{RMSE}}. \quad (13)$$

As a result, from the above equation, the amplitude of the changes in the fit function is 0–1,000. The maximum value is 1,000 for zero error and shows the high proportion of the

TABLE 2: Values of RMSE, f , R^2 , and SD for the recent studies and present study.

Standard deviation of residuals	Coefficient of determination (R^2)	Fitness function (f)	Root mean square error (RMSE)		Method
0.307	0.77	765.103	0.307	Training results	Present Study (GEP)
0.298	0.79	770.209	0.298	Testing results	
0.3024	0.777	767.425	0.303	Training results	Present Study (GMDH)
0.2988	0.788	769.15	0.3	Testing results	
0.385	0.689	610.037	0.639	Training results	Sarma and Free [15]
0.388	0.697	607.147	0.647	Testing results	
0.396	0.685	645.29	0.549	Training results	Ornthammarath et al. [18]
0.399	0.694	642.508	0.556	Testing results	
0.391	0.63	712.277	0.404	Training results	Alavi and Gandomi [21]
0.385	0.668	714.677	0.399	Testing results	
0.361	0.697	732.24	0.366	Training results	Kumar et al. [25]
0.359	0.716	733.956	0.362	Testing results	
0.308	0.657	760.668	0.315	Training results	Javan-Emrooz et al. [26]
0.306	0.658	760.52	0.314	Testing results	
0.27	0.496	685.5	0.4	Training results	Abdelfattah et al. [30]
0.267	0.502	683.7	0.46	Testing results	

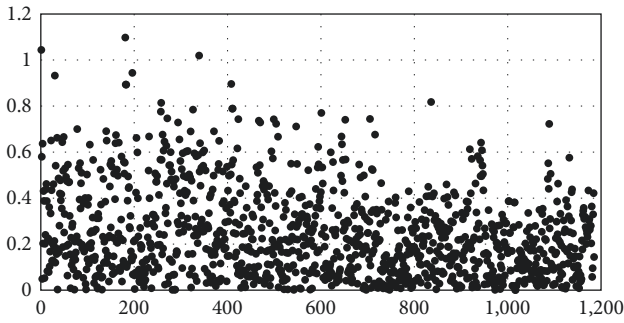


FIGURE 9: Residual trends of inter and intraevent error in the GEP method.

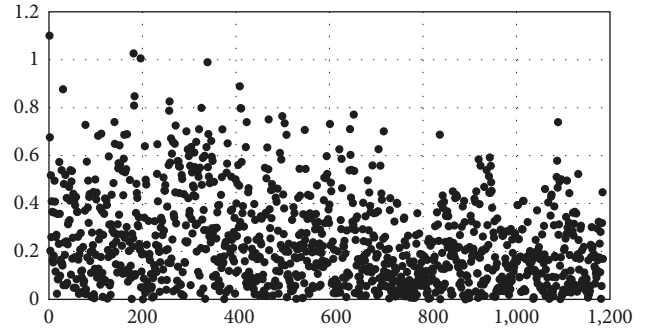


FIGURE 10: Residual trends of inter and intraevent error in the GMDH method.

predicted values with measured values from the previous earthquakes that are the most possible ideal state. The lower value of the fitness function shows a more inadequate matching.

Another parameter that was used to calculate the accuracy is the coefficient of determination (R^2), which can be determined by Equation (14).

$$R^2 = \left(\frac{\sum_{i=1}^n (PGA_{actual_i} - \overline{PGA_{actual}}) (PGA_{model_i} - \overline{PGA_{model}})}{\sqrt{\sum_{i=1}^n (PGA_{actual_i} - \overline{PGA_{actual}})^2 \sum_{i=1}^n (PGA_{model_i} - \overline{PGA_{model}})^2}} \right)^2. \quad (14)$$

The range of changes of this parameter is 0–1. The closer this number is to one, the greater the match between the predicted values and the target, and overall, there is a higher accuracy in the results.

Another parameter used in this research to measure accuracy is the standard deviation (SD) of residuals. Using Equation (15) the SD of residuals can be calculated.

$$SD = \sqrt{\frac{\sum_{i=1}^n \left(|PGA_{actual_i} - PGA_{model_i}| - \overline{|PGA_{actual_i} - PGA_{model_i}|} \right)^2}{n - 1}}. \quad (15)$$

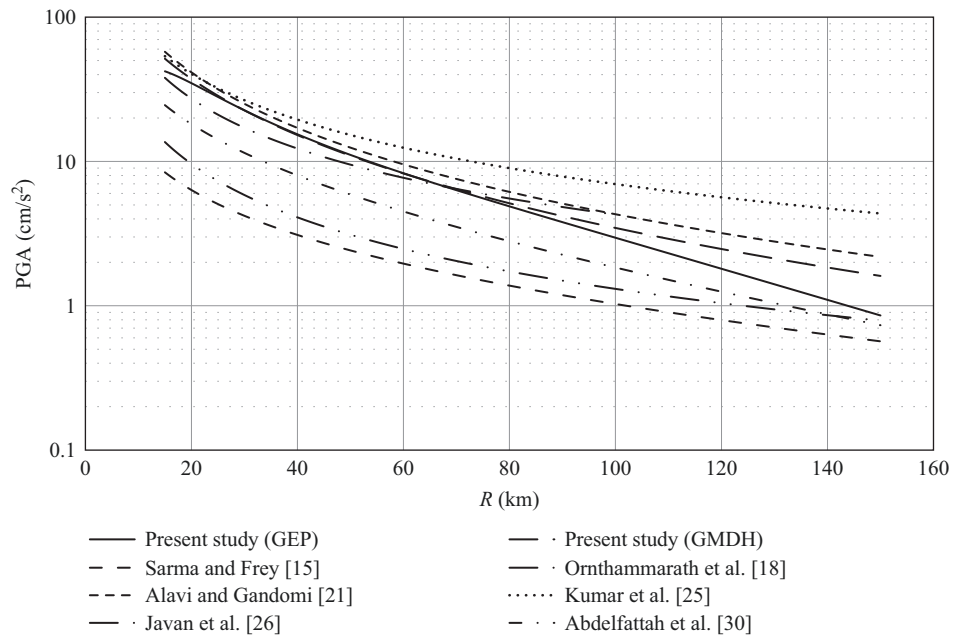


FIGURE 11: The PGAH change curve in terms of hypocentral distance for earthquakes with magnitudes of 4.5 and shear wave velocities of 400 m/s.

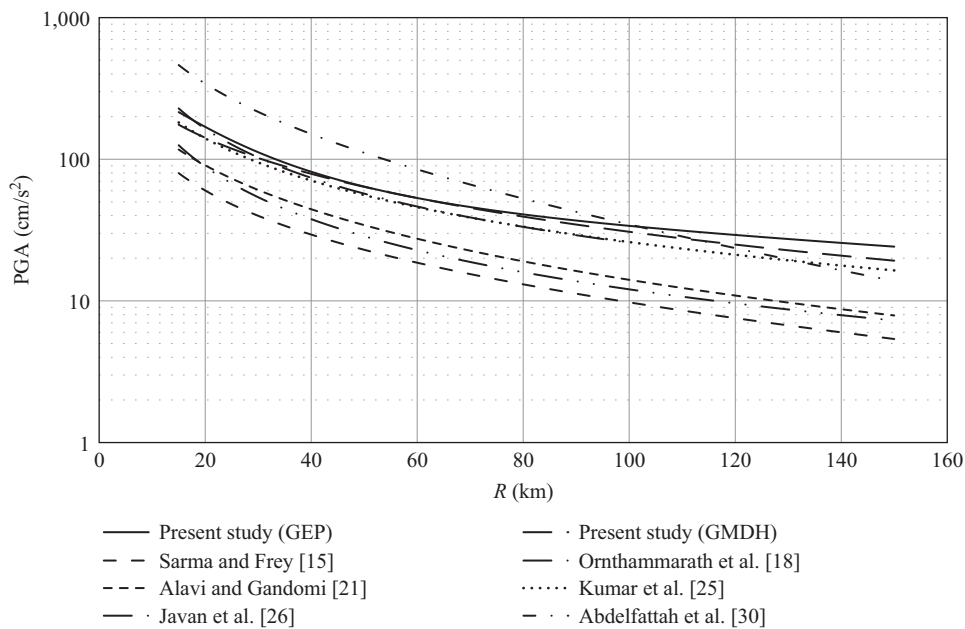


FIGURE 12: The PGAH change curve in terms of hypocentral distance for earthquakes with magnitudes of 6 and shear wave velocities of 400 m/s.

For each attenuation relationship obtained from GEP and GMDH methods, the parameters of RMSE, Fitness function, Coefficient of determination, and SD of residuals were calculated and presented in Table 2.

On the other hand, to show the agreement between target and predicted values, it is possible to present a diagram for the difference values between these parameters. These scatter

diagrams for GEP and GMDH methods are drawn in Figures 9 and 10, respectively.

Finally, in this section, the relationships presented in this study are compared with some relationships presented in the past. For comparison, it was tried to use attenuation relationships calculated by the other methods. The relationships selected for comparison are: Sarma and Frey [15], Ornthammarath et al.

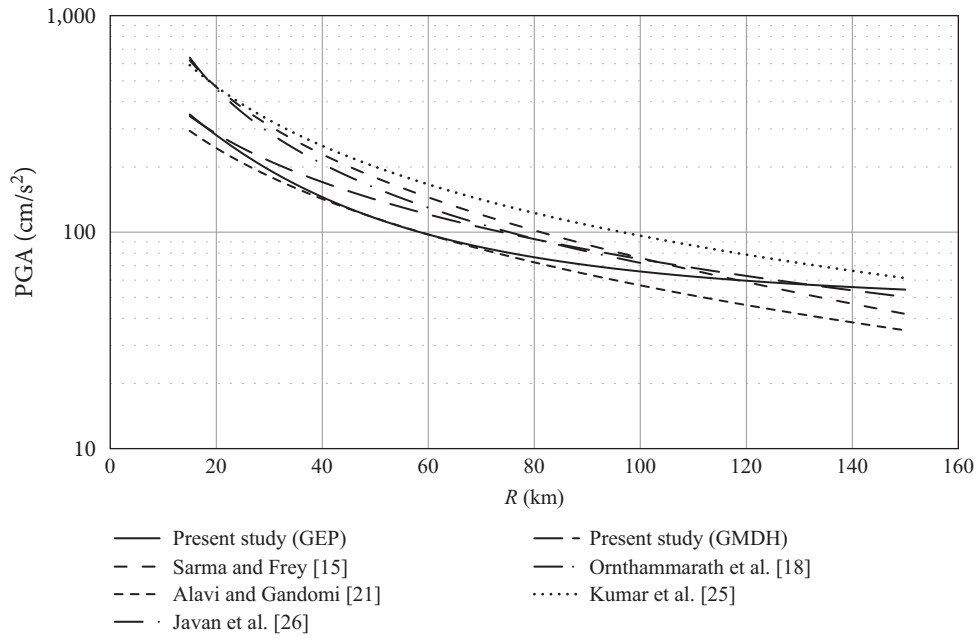


FIGURE 13: The PGAH change curve in terms of hypocentral distance for earthquakes with magnitudes of 7.5 and shear wave velocities of 400 m/s.

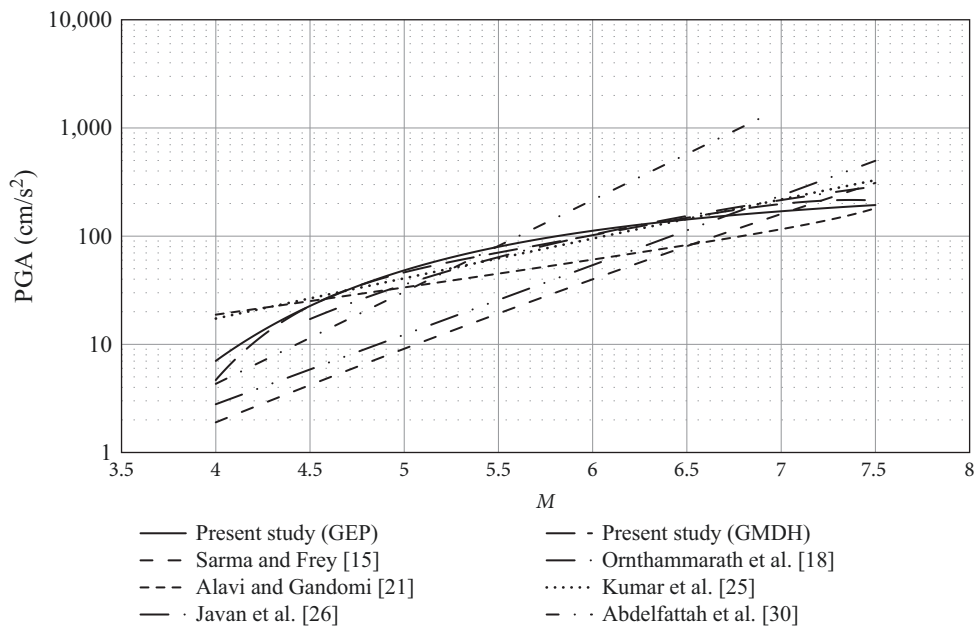


FIGURE 14: The PGAH change curve in terms of magnitude for a hypocentral distance of 30 km and shear wave velocities of 400 m/s.

[18], Kumar et al. [25], and Abdelfattah et al. [30], were presented in 1995, 2010, 2017 and 2021, respectively. Also, the relationships of Alavi and Gandomi [21] and Javan et al. [26], which were obtained using intelligent algorithms. For comparison, the values of parameters RMSE, Fitness function, Coefficient of determination, and SD of residuals were calculated for each of these relationships and presented in Table 2. As seen from Table 2, the values of Fitness function and Coefficient of determination in the present study have the highest values and

the values of RMSE and SD of residuals have the lowest values compared to the other methods. From this comparison, it can be concluded the methods used in this research are appropriate and highly accurate.

In the continuation of the comparison with other attenuation relationships, comparative diagrams were drawn. The PGAH change curve in terms of hypocentral distance of 15–150 km for earthquakes with magnitudes of 4.5, 6, and 7.5 and shear wave velocities of 400 m/s, for the results of the

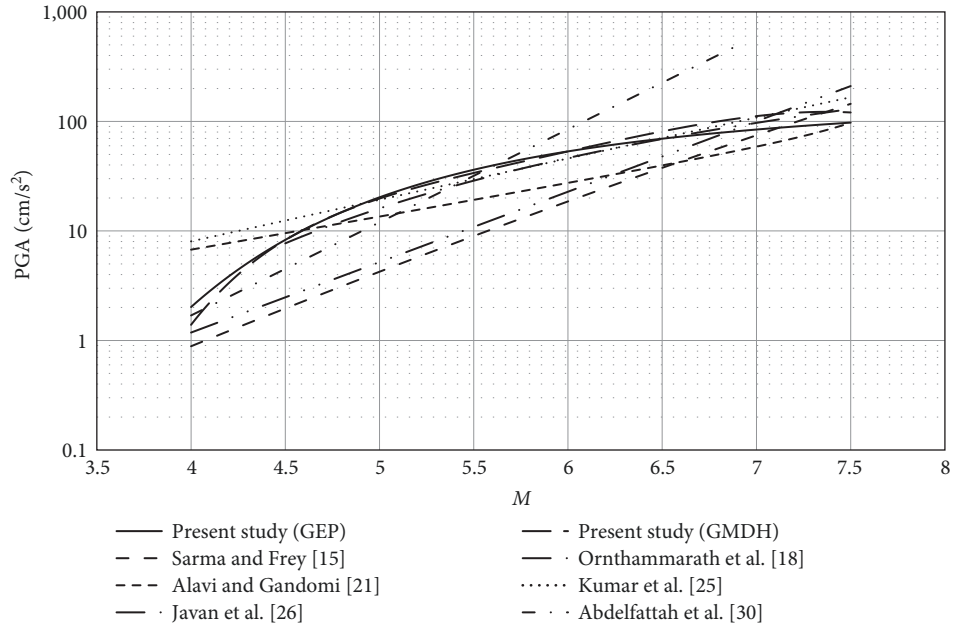


FIGURE 15: The PGAH change curve in terms of magnitude for a hypocentral distance of 60 km and shear wave velocities of 400 m/s.

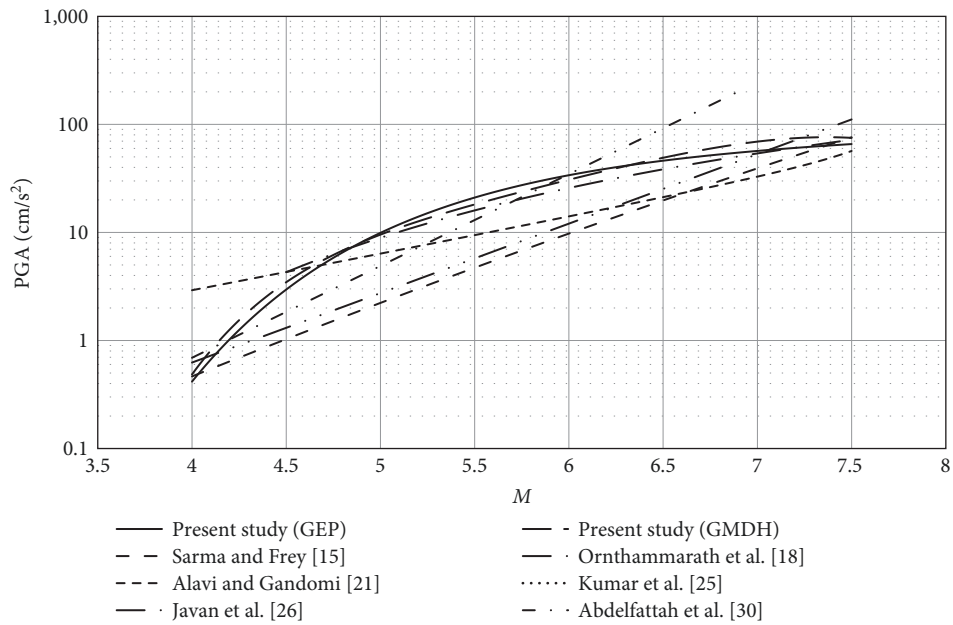


FIGURE 16: The PGAH change curve in terms of magnitude for a hypocentral distance of 100 km and shear wave velocities of 400 m/s.

present study and the mentioned researches, are shown in Figures 11–13. As a result, there is an acceptable correlation between the results of the present study and the previous studies.

Also, the horizontal component of the PGA change curve in terms of magnitude is 4–7.5, for a hypocentral distance of 30, 60, and 100 km and shear wave velocities of 400 m/s, belong to the results of the present study and the mentioned researches are shown in Figures 14–16. As it is clear from the graphs, the values obtained from the relationships presented

in this research are in the range of other researches and have a good match.

6. Conclusion

Every year, new earthquakes occur, accelerometers are installed in new areas and the accelerometer network expands, and in general, the earthquake catalog becomes more complete and accurate. On the other hand, providing new methods for

regression has caused that new attenuation relationships are always needed in earthquake risk analysis.

In this study, correcting the existing records and deleting some of them, 1,185 earthquake records were used to provide the attenuation relationship of the horizontal component of PGA. The records used have a moment magnitude of 4–7.5 and a hypocentral distance of 15–150 km. The parameters of moment magnitude, hypocentral distance, and shear wave velocity were considered as variables and the square root of the sum of the squares of peak ground horizontal acceleration in two horizontal orthogonal directions was treated as an objective function. Then new attenuation relationships were extracted by the GEP and the GMDH.

One of the innovations of this research is the addition of the shear wave velocity parameter as a variable in the attenuation relationship, which was rarely seen in the past researches.

The most important advantage of using GEP and GMDH methods in this research is that a predetermined pattern was not used to calculate relationships. As seen, in past researches, PGA changes in terms of magnitude were assumed to be linear, but in this research, it was shown that these changes are nonlinear. Also, changes in terms of hypocentral distance are not assumed to have a fixed pattern and it is observed that the pattern also changes in different intervals. For this reason, as it is clear from the comparisons, the accuracy of these methods is higher and shows a lower amount of error than the previous researches.

Data Availability

The data used in this research have been uploaded in the supplemental files section.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] K. Satake and B. F. Atwater, "Long-term perspectives on giant earthquakes and tsunamis at subduction zones," *Annual Review of Earth and Planetary Sciences*, vol. 35, no. 1, pp. 349–374, 2007.
- [2] M. Ajam, M. S. Amiri, and H. Pahlavan, "Attenuation relationship for the horizontal component of peak ground acceleration (PGA) using gene expression programming (GEP)," in *Presented at the Eighth International Conference on Seismology and Earthquake Engineering (SEE8)*, Tehran, Iran, 11–13 November 2019.
- [3] S. L. Kramer, *Geotechnical Earthquake Engineering*, Pearson Education India, 1996.
- [4] P. G. Somerville and R. W. Graves, "Characterization of earthquake strong ground motion," in *Landslide Tsunamis: Recent Findings and Research Directions*, J. P. Bardet, F. Imamura, C. E. Synolakis, E. A. Okal, and H. L. Davies, Eds., Pageoph Topical, pp. 1811–1828, Springer, Birkhäuser, Basel, 2003.
- [5] F. Neumann, *Earthquake Intensity and Related Ground Motion*, University of Washington Press Seattle, Washington, 1954.
- [6] W. G. Milne and A. G. Davenport, "Distribution of earthquake risk in Canada," *Bulletin of the Seismological Society of America*, vol. 59, no. 2, pp. 729–754, 1969.
- [7] L. Esteva, *Seismic Risk and Seismic Design Decisions*. 1970, Massachusetts Inst, Tech., Cambridge. Univ. of Mexico, Mexico City.
- [8] D. Denham and G. R. Small, "Strong motion data centre: bureau of mineral resources, Canada," *Bulletin of the New Zealand Society for Earthquake Engineering*, vol. 4, no. 1, pp. 15–30, 1971.
- [9] N. C. Donovan, *A Statistical Evaluation of Strong Motion Data: Including the February 9, 1971*, San Fernando Earthquake, Dames & Moore San Francisco, CA, USA, 1973.
- [10] L. Esteva and R. Villaverde, "Seismic risk, design spectra and structural reliability," *The 5th World Conference on Earthquake Engineering*, vol. 2, pp. 2586–2596, 1973.
- [11] D. L. Orphal and J. A. Lahoud, "Prediction of peak ground motion from earthquakes," *Bulletin of the Seismological Society of America*, vol. 64, no. 5, pp. 1563–1574, 1974.
- [12] M. Trifunac, "Preliminary analysis of the peaks of strong earthquake ground motion—dependence of peaks on earthquake magnitude, epicentral distance, and recording site conditions," *Bulletin of the Seismological Society of America*, vol. 66, no. 1, pp. 189–219, 1976.
- [13] W. B. Joyner and D. M. Boore, "Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake," *Bulletin of the Seismological Society of America*, vol. 71, no. 6, pp. 2011–2038, 1981.
- [14] G. M. Atkinson and D. M. Boore, "Recent trends in ground motion and spectral response relations for North America," *Earthquake Spectra*, vol. 6, no. 1, pp. 15–35, 1990.
- [15] S. Sarma and M. Free, "The comparison of attenuation relationships for peak horizontal acceleration in intraplate regions," in *Proceedings of the Fifth Pacific Conference on Earthquake Engineering*, 1995.
- [16] G. Sobhaninejad, A. Noorzad, and A. Ansari, "Genetic algorithm (GA): A new approach in estimating strong ground motion attenuation relations," in *Proceedings of the Fourth International Conference on Earthquake Geotechnical Engineering*, 2007.
- [17] A. F. Cabalar and A. Cevik, "Genetic programming-based attenuation relationship: an application of recent earthquakes in Turkey," *Computers & Geosciences*, vol. 35, no. 9, pp. 1884–1896, 2009.
- [18] T. Ornthammarath, J. Douglas, R. Sigbjörnsson, and C. G. Lai, "Assessment of strong ground motion variability in Iceland," in *Proceedings of Fourteenth European Conference on Earthquake Engineering*, Citeseer, 2010.
- [19] E. Kermani, Y. Jafarian, and M. H. Baziar, "A new attenuation model for peak ground acceleration in soil sites using genetic programming," in *2nd International Conference on New Developments in Soil Mechanics and Geotechnical Engineering*, Near East University, Nicosia, North Cyprus, 2009.
- [20] Y. Jafarian, E. Kermani, and M. H. Baziar, "Empirical predictive model for the v_{max}/a_{max} ratio of strong ground motions using genetic programming," *Computers & Geosciences*, vol. 36, no. 12, pp. 1523–1531, 2010.
- [21] A. H. Alavi and A. H. Gandomi, "Prediction of principal ground-motion parameters using a hybrid method coupling artificial neural networks and simulated annealing," *Computers & Structures*, vol. 89, no. 23–24, pp. 2176–2194, 2011.
- [22] G. Ghodrati Amiri, M. S. Amiri, and Z. Tabrizian, "Ground motion prediction equations (GMPEs) for elastic response spectra

- in the Iranian plateau using gene expression programming (GEP)," *Journal of Intelligent & Fuzzy Systems*, vol. 26, no. 6, pp. 2825–2839, 2014.
- [23] G. G. Amiri, M. S. Amiri, P. Namiarnian, and A. Emadzadeh, "Attenuation relationship for inelastic earthquake spectra in Iran," *Proceedings of the Institution of Civil Engineers - Structures and Buildings*, vol. 168, no. 5, pp. 313–325, 2015.
- [24] A. H. Gandomi, A. H. Alavi, M. Mousavi, and S. M. Tabatabaei, "A hybrid computational approach to derive new ground-motion prediction equations," *Engineering Applications of Artificial Intelligence*, vol. 24, no. 4, pp. 717–732, 2011.
- [25] A. Kumar, H. Mittal, R. Kumar, and R. S. Ahluwalia, "Empirical attenuation relationship for peak ground horizontal acceleration for North–East Himalaya," *Vietnam Journal of Earth Sciences*, vol. 39, no. 1, pp. 47–57, 2017.
- [26] H. Javan-Emrooz, M. Eskandari-Ghadi, and N. Mirzaei, "Prediction equations for horizontal and vertical PGA, PGV, and PGD in northern Iran using prefix gene expression programming," *Bulletin of the Seismological Society of America*, vol. 108, no. 4, pp. 2305–2332, 2018.
- [27] A. Erken, G. Şengül Nomaler, and Z. Gündüz, "The development of attenuation relationship for Northwest Anatolia region," *Arabian Journal of Geosciences*, vol. 11, no. 2, pp. 1–15, 2018.
- [28] B. Karimi Ghalehjough and R. Mahinroosta, "Peak ground acceleration prediction by fuzzy logic modeling for Iranian plateau," *Acta Geophysica*, vol. 68, no. 1, pp. 75–89, 2020.
- [29] A. Shiuly, N. Roy, and R. B. Sahu, "Prediction of peak ground acceleration for Himalayan region using artificial neural network and genetic algorithm," *Arabian Journal of Geosciences*, vol. 13, no. 5, pp. 1–10, 2020.
- [30] A. K. Abdelfattah, A. Al-Amri, K. Abdelrahman, M. Fnais, and S. Qaysi, "Attenuation relationships of peak ground motions in the Jazan Region," *Arabian Journal of Geosciences*, vol. 14, no. 3, pp. 1–9, 2021.
- [31] P. Kumar, B. P. Chamoli, A. Kumar, and A. Gairola, "Attenuation relationship for peak horizontal acceleration of strong ground motion of Uttarakhand region of central Himalayas," *Journal of Earthquake Engineering*, vol. 25, no. 12, pp. 2537–2554, 2021.
- [32] S. Pourzeynali and A. Khadivyan, "Attenuation relationships for horizontal and vertical peak and spectral accelerations for Alborz zone of Iran," *Journal of Earthquake Engineering*, vol. 26, no. 14, pp. 7469–7485, 2022.
- [33] A. G. Ghodrati, M. Zahedi, S. Gholami, and A. Mahdavian, "Appropriate frequency band for correcting Iranian accelerograms in different site conditions," 2004.
- [34] C. Ferreira, "Gene expression programming: a new adaptive algorithm for solving problems," 2001.
- [35] C. Ferreira, "Gene expression programming in problem solving," in *Soft Computing and Industry*, pp. 635–653, Springer, 2002.
- [36] C. Ferreira, *Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence*, Vol. 21, Springer, 2006.
- [37] H. S. Hwang, "Fuzzy GMDH-type neural network model and its application to forecasting of mobile communication," *Computers & Industrial Engineering*, vol. 50, no. 4, pp. 450–457, 2006.
- [38] N. Amanifard, N. Nariman-Zadeh, M. H. Farahani, and A. Khalkhali, "Modelling of multiple short-length-scale stall cells in an axial compressor using evolved GMDH neural networks," *Energy Conversion and Management*, vol. 49, no. 10, pp. 2588–2594, 2008.
- [39] A. G. Ivakhnenko, "Polynomial theory of complex systems," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-1, no. 4, pp. 364–378, 1971.
- [40] S. J. Farlow, "The GMDH algorithm," in *Self-Organizing Methods in Modeling: GMDH Type Algorithms*, vol. 54, p. 350, CRC Press, 1984.
- [41] Y. Pachepsky, W. Rawls, D. Giménez, and J. P. C. Watt, "Use of soil penetration resistance and group method of data handling to improve soil water retention estimates," *Soil and Tillage Research*, vol. 49, no. 1-2, pp. 117–126, 1998.
- [42] I. Ebtehaj, H. Bonakdari, A. H. Zaji, H. Azimi, and F. Khoshbin, "GMDH-type neural network approach for modeling the discharge coefficient of rectangular sharp-crested side weirs," *Engineering Science and Technology, an International Journal*, vol. 18, no. 4, pp. 746–757, 2015.
- [43] H. Iba, H. deGaris, and T. Sato, "A numerical approach to genetic programming for system identification," *Evolutionary Computation*, vol. 3, no. 4, pp. 417–452, 1995.
- [44] N. Nariman-Zadeh, A. Darvizeh, A. Jamali, and A. Moeini, "Evolutionary design of generalized polynomial neural networks for modelling and prediction of explosive forming process," *Journal of Materials Processing Technology*, vol. 164-165, pp. 1561–1571, 2005.