A New Peak-Ground-Acceleration Prediction Model by Using Genetic Optimization Techniques for Iranian Plateau Database

Sahar Rahpeyma [1], Alireza Azarbakht [2*], and Mehdi Mousavi [3]

1. Graduate Student, Dept. of Civil Engineering, Faculty of Engineering, Arak University, Iran
2. Associate Professor, Dept. of Civil Engineering, Faculty of Engineering, Arak University, Iran,
* Corresponding Author; email: a-azarbakht@araku.ac.ir
3. Associate Professor, Dept. of Civil Engineering, Faculty of Engineering, Arak University, Iran

Received: 12/06/2013
Accepted: 21/05/2014

ABSTRACT

Keywords:
Peak Ground Acceleration (PGA); Genetic Algorithm (GA); Genetic Programming (GP); Re-sampling; Iranian plateau

The main scope of this paper is to study an application of evolutionary algorithms, i.e. genetic algorithm and genetic programming (GP), for obtaining Peak Ground Acceleration (PGA) prediction model in the case of Iranian database. The proposed GP model is compared with a set of existing attenuation relationships via several traditional and modern mathematical and statistical methods. A new re-sampling approach is also introduced to assess the stability of the chosen models. The obtained model shows clearly more consistency with the local data in comparison with the other selected models.

1. Introduction

Prediction of ground motion intensity measures is one of the most important parts of Seismic Hazard Analysis (SHA). Generally, SHA, can be performed either probabilistically (PSHA) or deterministically (DSHA) [1], in both of which Ground Motion Prediction Equations (GMPEs) are inherent parts.

Researchers have been developed two basic different methodologies, i.e. empirical and physical relationships, for attaining prediction equation models according to site geology and distribution of events. Empirical models, which are based on mathematical methods, describe the observations by means of regression analysis on a specific site with abundant data set. On the other hand, physical models, which describe seismic wave's generation and propagation, are used in a specific site with lack of observations. Recently, beside two mentioned approaches, methods of information processing known as soft computing techniques, such as Evolutionary Algorithms (EA), have been used in order to obtain attenuation relationships as a modern approach [2-5]. Evolutionary algorithms, specifically genetic programming (GP) and genetic algorithm (GA), are optimization techniques based on the rules of natural selection [6-7]. Although, using GP and GA methods does not reduce the uncertainties; however, there is more complicated interaction among the observation and prediction values [3].

The main aim of this study is to derive a new Peak-Ground-Acceleration (PGA) attenuation model via GP and GA methods based on an Iranian database. The incorporation of information-theoretic method [8] and re-sampling analysis has been proposed here in order to improve the fitness functions of GP and GA. Afterwards, the new PGA attenuation model is compared with a set of available prediction models.

2. Genetic Algorithm and Genetic Programming

Genetic algorithm, as the most popular member of evolutionary algorithms, is the best known intellectual optimization technique based on the principles of natural selection and survival of the fittest [7, 9, 10], initially developed by John Holland at the University of Michigan in 1975.

Genetic programming, as a particular form of genetic algorithm, was introduced by John Koza in 1992. In this subset of evolutionary algorithm, the absolute solution, without any explicit programming, is obtained by using the concepts of genetic algorithm and the parse trees (tree structures). The development of the initial population in GP, consisting of functions and terminals, is accomplished by means of biological selection and reproduction [6].

The following three main steps are key elements within any Genetic programming procedure [6]:

1. Generate an initial population of random compositions of functions and terminals.
2. Repeat (below) steps 2.1 and 2.2 until the establishment of the program's suitable and final condition:
   2.1. Executes each program and assigns a fitness value to it according to the fitness function.
   2.2. Create a new population of computer programs by means of the genetic operators (reproduction, mutation, and cross-over).
- Reproduction: Copy the best existing programs in the new population.
- Mutation: Select an existing program, change a node of the individual randomly and move the program to the new population.
- Cross-over: Select two programs and change one branch with another randomly and move the two produced programs to the new population.
3. Select the best computer program that has been appeared in any generation.

3. Ground Motion Database

Iran is located in the middle part of the Alpine-Himalayan seismotectonic belt and is known as one of the most seismicity active regions in the world. Several comprehensive studies on geological characteristics and the seismicity nature of this region have been carried out [11-13]. Researchers usually classify Iranian plateau into two major seismic zones i.e. the Central Iran and the Zagros [14]. Most of the seismic activities are concentrated in the Zagros region, and less seismic activity is observed in the central Iran and other regions [12].

The data set used in this study, as seen in Table (1), consists of 179 strong ground motion records of 36 earthquake events with moment magnitude (Mw) ranging from 5.0 to 7.4 and distance ranging less than 200 km occurred between 1978 and 2008. The total 179 records have been extracted from the Iran Strong Motion Network (according to the

Table 1. Ground motion database for Iranian plateau (date is listed as YYYY/MM/DD).

No	Date	M_w	*FD	**N	***Zone
1	1978/09/16	7.4	10	4	1
2	1979/11/27	7.1	10	7	1
3	1990/06/20	7.3	12	2	1
4	1994/06/20	5.8	9	7	2
5	1997/02/04	6.5	8	1	1
6	1997/02/28	6	9	3	1
7	1997/05/10	7.2	13	7	1
8	1998/03/14	6.6	5	2	1
9	1999/08/21	5	25	3	2
10	1999/05/06	6.2	7	5	2
11	1999/05/06	5.7	10	3	2
12	1999/10/31	5.2	15	4	2
13	2002/04/24	5.4	25	6	2
14	2002/06/22	6.4	10	12	1
15	2002/12/24	5.2	20	6	2
16	2003/07/10	5.8	10	4	2
17	2003/07/10	5.7	15	4	2
18	2003/08/21	5.9	20	3	1
19	2003/11/28	5	25	3	2
20	2003/12/26	6.5	3	6	1
21	2004/05/28	6.3	27	5	1
22	2004/10/07	5.6	30	9	1
23	2005/01/10	5.3	32	8	1
24	2005/02/22	6.3	10	6	1
25	2005/11/27	5.9	12	6	2
26	2006/03/30	5.1	20	8	2
27	2006/03/31	6.1	12	9	2
28	2006/03/31	5.1	26	6	2
29	2006/06/28	5.8	12	4	2
30	2008/05/05	5.2	12	3	2
31	2008/09/10	6.1	12	5	2
32	2008/09/11	5.2	7	3	2
33	2008/09/17	5.2	12	3	2
34	2008/12/07	5.4	12	4	2
35	2008/12/08	5.1	12	4	2
36	2008/12/09	5	14	3	2

*FD: FD is focal depth (km)

**N: Number of records for each earthquake.

***Zone: 1 and 2 refer to the Central Iran Zone and Zagros Zone, respectively.

Building and Housing Research Center, BHRC website, last accessed December 2012). Most of the Iranian earthquake events are reverse-thrust, strike-slip, or a combination of these two mechanisms [15].

Figure (1), exhibits the distribution of magnitude versus distance, with displaying different site types. The site classification in this study is the same as the one defined in the Iranian code of standard seismic resistant design of buildings, Standard No. 2800 [16], which includes four classes. Site categories I and II ($V_{S30} \geq 375$ m/s) were combined together and considered as the rock site, and categories III and IV ($V_{S30} < 375$ m/s) were combined together and named the soil site.

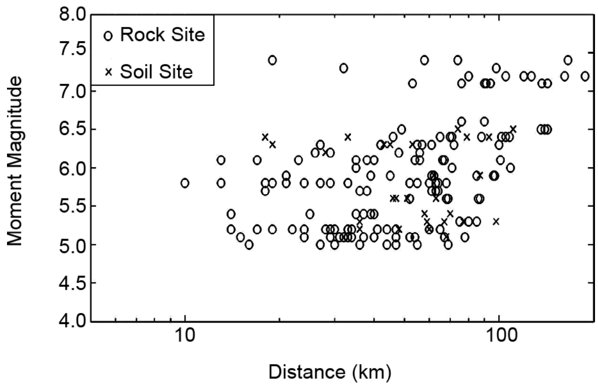

Figure 1. Distribution in terms of magnitude, distance and site classification of study accelerograms recorded by the BHRC in Iran.

4. A New PGA Prediction Model

GPLAB, which is used in this paper, is a genetic programming which is written by Sara Silva in 2007 [17]. This toolbox is an operational and practical application for different types of users. Recently, some researchers have used this toolbox for obtaining predictive equations [2, 18].

For using GPLAB, the database is divided into the training set (80% of the data set) and the testing set (20% of the data set), chosen randomly (uniformly distributed).

The programs in GPLAB (tree structures), are initialized with one of the three accessible initializing methods "Full, Grow, and Ramped Half-and-Half" [6]. In this study, initial population is produced based on Ramped Half-and-Half method. In the standard procedure, an equal number of individuals is initialized for each depth between two and the initial tree depth value [17]. The population of trees resulting from this initialization method is very diverse, with balanced and unbalanced trees of bseveral different depths [17].

One of the important features of GPLAB is some appropriate restrictions on tree's depth or size to avoid bloat that is a phenomenon consisting of an excessive code growth without any corresponding improvement in the fitness [6, 17].

In GPLAB, parents are selected for reproduction according to four usual sampling methods [6, 17]. In this paper, Lexictour sampling approach was used for selecting parents. In this approach, a random number of individuals are chosen from the population and the best of them is chosen [17]. Table (2) indicates important parameters used for running GPLAB in the current paper.

Table 2. The adoptive parameters for GPLAB.

Function Set	$+ \; - \; \div \; \times \; \sqrt[n]{\times}$ *Power* (X^n) $\quad n = 1, 2, 3$
Terminals	{*Mw*, *R*, VS_{30}, Ramped-half-and-half}
Initial Population	Ramped nit (Ramped-half-and-half)
Sampling Method	Lexictour
Operators	Mutation, Cross-Over
Elitism	Total elitism
Total Data	179
Training Data	145
Testing Data	34
End Point	Number of Generations
Population Size	800
Generation	150
Fixed Level	2 (1 = depth, 2 = nodes)
Real Max Level	40

The GP fitness function based on information-theoretic method [8] is proposed in order to quantitatively assess the predictive models. In this study, by using the average sample log likelihood definition, the LLH criterion is defined as written in Eq. (1) and GPLAB minimized it.

$$LLH = -\frac{1}{N}\sum_{i=1}^{N}\log_2(g(x_i)) \qquad (1)$$

The Expression Tree (ET) of LnPGA, obtained from genetic programming (GPLAB), is shown in Figure (2), where X_1, X_2, and X_3 are moment magnitude, distance measure, and shear wave velocity, respectively.

After obtaining the initial predictive model, by GP,

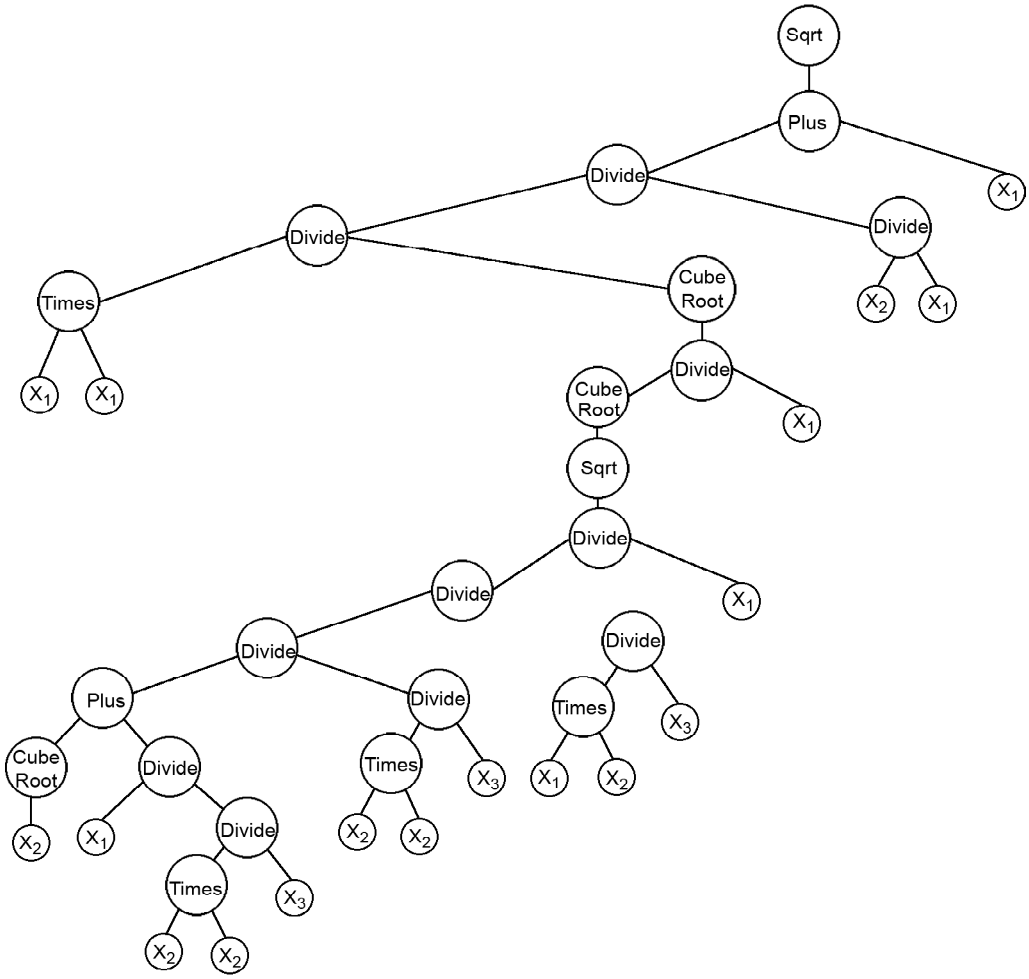

Figure 2. ET for the proposed GP model (X_1, X_2, and X_3 are moment magnitude, distance measure, and shear wave velocity, respectively).

in order to reduce the bias toward different earthquake parameters and likewise for reducing the sensitivity of the initial attenuation model to the considered database, the GA fitness function is defined according to Eq. (2) as a combination of LLH criterion and re-sampling analysis.

$$Fitness\ Function = \omega_1 \times LLH_{Training\ Data} + \omega_2 \times \sum_{n}^{130,145 \atop 90,110} \frac{\sum_{i=1}^{Nd} \sum_{j}^{M,R,V_{S30}} \left|1 - PV_j^{S_i^n}\right|}{3 \times N_d} \tag{2}$$

where N_d is the uniformly distributed random databases (in this study, N_d =100), S_i^n is the i^{th} sample with n records, S_i^n is the residual's p-value of S_i^n versus j^{th} parameter, M is the earthquake moment magnitude, R is the distance measure, V_{S30} is the shear wave velocity, ω_1 = 0.25 and ω_2 = 0.125 are the weighting constants based on authors judgment.

The final form of LnPGA is shown in Eq. (3) and Table (3) shows the result of the coefficients achieved by GA.

$$Ln(PGA) = Sqrt(a_1 \frac{M_w^{a_2}}{R^{a_3}} \times \frac{a_4}{V_{S30}^{a_5}\left(a_6 R^{a_7} + a_8 M_w{}^{a_9} \times V_{S30}{}^{a_{10}}\right)^{a_{11}}} + a_{12} M_w{}^{a_{13}}) + \sigma \tag{3}$$

M_w, R, and V_{S30} denote, respectively, the earthquake moment magnitude, earthquake distance measure and shear-wave velocity.

Table 3. The constant coefficients obtained by GA from final optimized model.

a_1	1.00	a_8	1.00
a_2	3.44	a_9	1.00
a_3	0.72	a_{10}	1.00
a_4	1.00	a_{11}	0.056
a_5	0.11	a_{12}	1.00
a_6	1.00	a_{13}	1.00
a_7	2.33	σ_{LnY}	0.9276

5. Selected Ground Motion Prediction Equations

According to different researches on seismotectonic nature of Iranian Plateau, it has been demonstrated that all of the earthquake events occurred in this region are shallow and intra-plate events [19]. Furthermore, meaningful correlation between this kind of events from different regions, including Turkey and California is reported by researchers [20]. Based on these facts, the candidate ground motion prediction equations are consisted of the following three classes:

- The local GMPEs which have been developed based on the Iranian datasets.
- The regional GMPEs corresponding to Europe and Middle East datasets.
- The global GMPEs developed in the Next Generation Attenuation (NGA) project.

In 2008, the NGA project, which was initiated by the Pacific Earthquake Engineering Research center (PEER) [21], has published five new models through a comprehensive and highly interactive research program, for shallow crustal earthquakes in the Western North of America to predict Peak Ground Acceleration (PGA), Peak Ground Velocity (PGV), and 5% damped response spectra for periods ranging between 0.01 to 10 seconds. These relationships are Abrahamson and Silva [22], Boore and Atkinson [23], Campbell and Bozorgnia [24], Chiou and Youngs [25], and Idriss [26]. It should be noted that Idriss [26] model only includes rock sites (assumed to be sites with V_{S30} 450 m/s), in which this significant difference isolates the model from the other models because it can only be applied to rock sites. Therefore, this model is excluded in this paper for further investigations. These models consist of different parameters, e.g. terms of influence the nonlinear site, sediment depth, hanging wall effects, source parameters, and etc. The NGA database used to develop the NGA GMPEs is relatively large i.e. 3551 recordings from 173 earthquakes (A few Iranian events are also included in this database).

Some recent papers have presented a number of suggestions as criteria that can be used to select GMPEs [27]. Four significant points are particularly considered in this study;

- The models superseded by a more recent publication are excluded.
- The models which lack either in non-linear magnitude dependence or magnitude-dependent decay with distance are avoided [27-28]. This issue should be met just by empirical models, not by physical model.
- The models which use inappropriate definitions for explanatory variables, such as M_L or R_{epi}, or models with the site effects without consideration of V_{S30} are excluded.
- The coefficients of the model were not determined by a method that accounts for inter-event and intra-event components of variability. In other words, models must be derived using one- or two-stage maximum likelihood approaches or the random effects approach.

Here, the selected ground motion models are briefly described as follows:

5.1. Saffari et al. [15] (Setal12)

The model has been developed for predicting PGA, PGV, and acceleration response spectra with 5% damping based on a subset of Iran database (78 earthquakes and 351 records). This model includes Moment magnitude, distance measure, fault mechanism, site class, and zone as seismic parameters.

5.2. Zafarani et al. [29] (Zetal12)

Zetal12 is a physical GMPE relationship which was developed by using the Specific Barrier Model (SBM). An Iranian data set consists of 171 strong-motion records from 24 earthquakes for Zagros region was used to obtain this model.

5.3. Ghodrati et al. [30] (Getal07)

Getal07 has been developed for predicting PGA, PGV, and Effective Peak Acceleration (EPA) for Zagros, Alborz, and Central-Iran seismotectonic regions. The data set includes 89 earthquakes and 307 records. In this model, surface wave magnitude (Ms) as moment measure has been used.

According to the third criterion recommended by Bommer et al. [27], this model should be excluded, nonetheless it is kept in this stage of the study to show the inconsistency of this model in comparison with the other models.

5.4. Akkar and Bommer [31] (AB10)

This model can be used for the prediction of PGA, PGV, and response spectral ordinates in the Europe, the Middle East and the Mediterranean. They used a subset of 532 strong-motion records from 131 events in these regions.

5.5. Akkar and Cagnan [32] (AC10)

This model is proposed for predicting PGA, PGV, and acceleration response spectra with 5% damping for periods ranging from 0.03 to 2 seconds by means of Turkish ground-motion database.

5.6. Ambraseys et al. [33] (Aetal05)

This model presents equations for the estimation of PGA and pseudo-spectral acceleration caused by shallow crustal earthquakes by means of a set of 595 strong-motion records recorded in Europe and the Middle East.

5.7. Ozbey et al. [34] (Ozetal04)

The base is a subset of 195 records from 17 earthquakes used in the regression analyses. This model predicts PGA and acceleration response spectra with 5% damping for periods ranging from 0.1 to 4 seconds.

5.8. Kalkan and Gulkan [35] (KG04)

The corresponding authors used a dataset created from a set of 112 strong ground motion records from 57 earthquakes that occurred between 1976 and 2003 to develop horizontal GMPE relationships for Turkey.

5.9. Bindi et al. [36] (Bindi10)

The data set was composed of 561 three-component waveforms from 107 earthquake events occurred in Italy between 1972 and 2007 and recorded by 206 stations at distances up to 100 km. This model predicts PGA, PGV, and acceleration response spectra with 5% damping for periods ranging from 0.3 to 2 seconds.

5.10. Campbell and Bozorgnia [24] (CB08)

CB08 has been obtained based on a subset of the PEER NGA database (1661 records from 64 events) for predicting PGA, PGV, and acceleration response spectra with 5% damping according for periods ranging between 0.01 to 10 seconds. The CB08 includes the effects of magnitude saturation, magnitude-dependent GMPE, style of faulting, rupture depth, hanging-wall geometry, linear and nonlinear site response, 3-D basin response, and inter-event and intra-event variability.

5.11. Boore and Atkinson [23] (BA08)

This model is one of the NGA project models that were obtained based on 1574 records from 58 events for predicting PGA, PGV, and acceleration response spectra with 5% damping for periods ranging from 0 to 10 seconds. The main predictor parameters in BA08 are moment magnitude, closest horizontal distance to the surface projection of the fault plane (RJB), and the averaged shear-wave velocity from the surface to 30 m V_{S30}.

5.12. Chiou and Youngs [25] (CY08)

This model was driven by using the PEER NGA database of 1950 records from 125 events. CY08 predicts PGA, PGV, and acceleration response spectra with 5% damping for periods ranging between 0.01 and 10 seconds. The model incorporates the effect of seismic source scaling, path scaling, and site effects.

5.13. Abrahamson and Silva [22] (AS08)

AS08 has been obtained to predict PGA, PGV, and acceleration response spectra with 5% damping for periods ranging between 0.01 and 10 seconds. The corresponding authors used 2754 strong-motion records from 135 earthquake events of the PEER NGA database. This model obtained from site response model, hanging-wall model, depth-to-top of rupture model, large distance model, soil depth model, and constant displacement model.

The nominated GMPE models are summarized in Table (4) including the valid range of magnitude and distance. In this study, the result of Kaklamanos' technical note has been used in order to reduce uncertainties and convert all input variables of GMPE models into a unique definition [37]. In addition, all GMPE models use the moment magnitude scale except Getal07 model in which the transition equations for magnitude measures have been used [38].

6. Residuals Analysis

The residuals analysis is the main technique to choose an appropriate model among the numerous

Table 4. Nominated ground motion prediction equations.

No	GMPE	Abbreviation	Category	Dominant Region	Distance (km)	Magnitude
1	Saffari et al. [15]	Seta12	1	Iran	15-135	5.0-7.3
2	Zafarani et al. [29]	Zetal12	1	Iran	1-200	4.4-7.5
3	Ghodrati et al. [30]	Getal07	1	Iran	5-150	4.5-7.5
4	Akkar and Bommer [31]	AB10	2	Europe, Middle east	0-100	5.0-7.6
5	Akkar and Cagnan [32]	AC10	2	Turkey	0-200	3.5-7.6
6	Ambraseys et al. [33]	Aetal05	2	Europe, Middle east	0-100	5.0-7.5
7	Ozbey et al. [34]	Ozetal04	2	Turkey	5-300	5.0-7.4
8	Kalkan and Gulkan [35]	KG04	2	Turkey	1-250	4.0-7.5
9	Bindi et al. [36]	Bindi10	2	Italy	0-100	4.0-6.9
10	Campbell and Bozorgnia [24]	CB08	3	California	0-200	4.0-7.5
11	Boor and Atkinson [23]	BA08	3	California	0-200	5.0-8.0
12	Chiou and Youngs [25]	CY08	3	California	0-200	4.0-8.0
13	Abrahamson and Silva [22]	AS08	3	California	0-200	5.0-8.5

ground motion prediction equations. The residual is defined by Eq. (4) as the subtraction of the natural logarithm of the predicted value from the natural logarithm of the observed value in which each data point has one residual.

$$r_{ij} = LnY_{ij} - Ln\hat{Y}_{ij} = \eta_i + \varepsilon_{ij} \tag{4}$$

where LnY_{ij} is the observed value and LnY_{ij} is the predicted value of j^{th} record of i^{th} event. η_i and ε_{ij} are, respectively, the inter-event residual and the intra-event residual, Eqs. (5) and (6).

$$r_i^{[inter]} = \eta_i = \frac{1}{N_i}\sum_{j=1}^{N_i} r_{ij} \tag{5}$$

$$r_{ij}^{[intra]} = \varepsilon_{ij} = r_{ij} - \eta_i \tag{6}$$

6.1. Residuals Distribution

The perfect form of the obtained residuals in the previous section has a normal distribution with zero mean and unit variance ($\mu = 0$, $\sigma = 1$). The fitness of the resulted residuals to the ideal form indicates the level of compatibility of the applied ground motion model with the recorded data. There are various statistical tests in order to evaluate the goodness of fitness such as z-test and Lilliefors test [39]. The null hypothesis in the z-test is that the mean of the normalized residual set is zero when the residuals are assumed to have a normal distribution with unit variance [39]. The null hypothesis in the Lilliefors-test is that data come from a normal distribution when the mean and the variance of the distribution are unidentified [39].

The resulted p-values of these two hypothesis tests indicate acceptable or unacceptable null hypothesis with respect to a given data. A small p-value means significant difference between observed and predicted amounts of models and a large p-value referred to more acceptable model [40]. Table (5) shows the results of the mentioned tests in the current study. In this table, the logical value $H = 1$ belongs to the rejection of the null hypothesis at the 5% significance level in which $H = 0$ is reversed.

As seen in Table (5), all of the candidate models have normal distribution based on the Lilliefors test; however, the null hypothesis can be rejected for the majority of the candidate models based on the z-test. It worth emphasizing that the mentioned traditional hypothesis tests only check for one

Table 5. Results of the hypothesis tests.

Model	Lilliefors-Test		z-Test	
	P-value	H	P-value	H
*GP Model	0.4957	0	0.8892	0
Setal12	0.5774	0	0.9961	0
Zetal11	0.6977	0	0.0447	1
Getal07	0.8269	0	0.0547	0
AB10	0.3575	0	8.47E-7	1
AC10	0.2815	0	3.68E-50	1
Aetal05	0.8630	0	4.85E-5	1
Ozetal04	0.4968	0	2.6e-27	1
KG04	0.2574	0	0.8471	0
Bindi10	0.7171	0	7.62E-5	1
CB08	0.9553	0	0.3862	0
BA08	0.9153	0	2.94E-4	1
CY08	0.2619	0	3.27E-5	1
AS08	0.8010	0	0.0323	1

*GP Model: the obtained model by GP and optimized by GA hypothesis, i.e. normal distribution or zero mean; hence, they are not perfect tools for evaluating the nominated GMPEs. That is, additional techniques have been employed in order to assess the ground motion models as discussed in the following sections.

6.2. Error Terms, Coefficient of Determination, Information Theoretic Method and Coefficient of Efficiency

The error terms are the criteria for assessing the accuracy of the chosen GMPEs. In this study, two error criteria, i.e. Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) as written, respectively, in Eqs. (7) and (8) are used to quantify how accurate the models predict ground motion parameters.

$$RMSE = \sqrt{\frac{\sum_N (X_{obs} - X_{pre})^2}{N}} \tag{7}$$

$$MAE = \frac{\sum_N \left| X_{obs} - X_{pre} \right|}{N} \tag{8}$$

In statistics, the coefficient of determination (denoted by R^2) is a criterion to show how well a model predicts outcomes. This measure is most often seen as a number between zero and unity. A large value of the coefficient of determination indicates the model perfectly fits the data. Eq. (9) represents the mathematical form of the coefficient of determination which is used in the current study.

$$R^2 = \frac{\sum_N (X_{obs})^2 - \sum_N (X_{obs} - X_{pre})^2}{\sum_N (X_{obs})^2} \tag{9}$$

In Eqs. (7), to Eq. (9) X_{obs} and X_{pre} are, respectively, the observed and the predicted values and N is the total number of records in the data set.

The information theoretic method is a modern powerful technique for evaluating models [8]. The quantitative assessment between different candidate models requires a meaningful distance measure based on an information theoretic framework; this measure is given by the Kullback-Leibler distance [41]. The Kullback-Leibler distance between two probabilistic models g_1 and g_2 is presented as written in Eq. (10):

$$\begin{aligned} D(f,g_2) - D(f,g_1) = \\ & E_f[\log_2(f)] - E_f[\log_2(g_2)] - \\ & (E_f[\log_2(f)] - E_f[\log_2(g_1)]) = \\ & E_f[\log_2(f)] - E_f[\log_2(g_2)] - \\ & E_f[\log_2(f)] + E_f[\log_2(g_1)] = \\ & E_f[\log_2(g_1)] + E_f[\log_2(g_2)] \times \\ & \approx \left\langle \log_2(L(g_1|x)) - \log_2(L(g_2|x)) \right\rangle \end{aligned} \tag{10}$$

where E_f is the expected value taken with respect to f. Here, for a base 2 logarithm, its unit is bit. As a consequence, by means of the average sample log likelihood definition, the LLH criterion is defined, Eq. (11).

$$LLH := -\frac{1}{N}\sum_{i=1}^{N} \log_2(g(x_i)) \tag{11}$$

The low LLH value shows the good appropriateness of the candidate models.

The Nash-Sutcliffe efficiency coefficient (denoted by E) is employed to quantify the predicted values with the observed values [42]. This coefficient is determined in logarithmic space via Eq. (12):

$$E = [1 - \frac{\sum_{j=1}^{m}\sum_{i=1}^{n}(LnY_{ij} - Ln\hat{Y}_{ij})^2}{\sum_{j=1}^{m}\sum_{i=1}^{n}(LnY_{ij} - Ln\bar{Y})^2}] \times 100\% \tag{12}$$

where m is the number of periods under consideration (in this study, $m = 1$), n is the number of records in the database, LnY_{ij} are the observed values, $Ln\hat{Y}_{ij}$ are the predicted values, and $Ln\bar{Y}_{ij}$ is the mean of LnY_{ij}. This criterion can be varied between and 100%. The higher indicator represents better conformity between the predicted values and the observed values. An efficiency of 100% ($E = 1$) corresponds to a perfect match of predicted models to the observed data, whereas an efficiency of zero ($E = 0$) indicates that the model predictions are as accurate as the mean of the observed data, while the negative E values show that the arithmetic mean of the observed values has a greater prediction accuracy than the model itself [42]. As the Nash-Sutcliffe model is more sensitive to the additive and the multiplicative difference between the observations and the predictions than the other goodness-of-fit statics, the researchers find out this criterion as a better indicator [43]. Table (6) includes the result of the mentioned criteria.

It should be mentioned that the Nash-Sutcliffe efficiency can also be used to quantitatively describe the accuracy of predicting models and measure the dispersion about the one-to-one line. Figure (3) displays the comparison of the observed and predicted PGA by GP model and AC10 model within the selected database and also confirms the results of coefficient of efficiency shown in Table (6).

As it is illustrated in Table (6), GP model has the lowest LLH value among the other models; hence, the purpose of the defined fitness function for GP and GA is achieved according to this criterion. After GP model, Zetal12 model, which is based on the Iranian database, has lowest LLH value. According to the coefficient of efficiency criterion, GP model and Zetal12 have the best results among the other candidate models, whereas AC10 and Ozetal05 have efficiency less than zero. GP model and Zetal12 are the best models among the other models with the lowest error values and the highest coefficient of determination. In addition, AB10, AC10, Aetal05, Ozetal04, and Bindi10 models, which are corresponding to Europe and Middle East regions, have the lowest coefficient of determination and the highest error values. Therefore, the authors decided to exclude these two models in this stage of research. Furthermore, NGA models do not show enough reasonable results than the other models in which it seems they are not superior models for the study region.

Table 6. Result of Error terms (RMSE, MAE), LLH, E, and R^2.

Model	*LLH*	R^2	*E*	RMSE			MAE		
				r_{ij}	$r_i^{[inter]}$	$r_{ij}^{[intra]}$	r_{ij}	$r_i^{[inter]}$	$r_{ij}^{[intra]}$
GP Model	1.9368	0.9454	33.4252	0.9264	0.5643	0.7858	0.7470	0.4269	0.6272
Setal12	2.2913	0.9399	22.9985	0.9598	0.6383	0.7704	0.7787	0.4891	0.6085
Zetal12	1.9712	0.9460	34.1801	0.9212	0.5750	0.7651	0.7385	0.4135	0.6049
Getal07	2.9947	0.9393	24.4972	0.9742	0.6774	0.7771	0.7855	0.5424	0.6310
AB10	2.5466	0.9341	22.9090	1.0320	0.6920	0.7949	0.8524	0.5133	0.6388
AC10	3.2676	0.8651	-64.3621	1.4556	1.2127	0.7918	1.2117	1.1202	0.6251
Aetal05	2.4133	0.9364	25.7504	1.0115	0.6697	0.7907	0.8234	0.4865	0.6359
Ozetal05	3.6427	0.9032	-17.8296	1.2324	0.9556	0.7841	0.9839	0.8423	0.61152
KG04	2.2981	0.9444	32.3028	0.9342	0.6179	0.7716	0.7568	0.4760	0.6127
Bindi10	2.2288	0.9310	19.1787	1.0522	0.6815	0.8675	0.8593	0.4886	0.6787
CB08	2.7452	0.9433	30.9763	0.9433	0.6211	0.7710	0.7588	0.4607	0.6167
BA08	2.6496	0.9396	28.8792	0.9737	0.7443	0.6953	0.7660	0.5649	0.5382
CY08	2.5322	0.9378	24.5607	0.9889	0.6403	0.7816	0.8164	0.4568	0.6183
AS08	2.2605	0.9425	29.9721	0.9501	0.6120	0.7771	0.7708	0.4301	0.6199

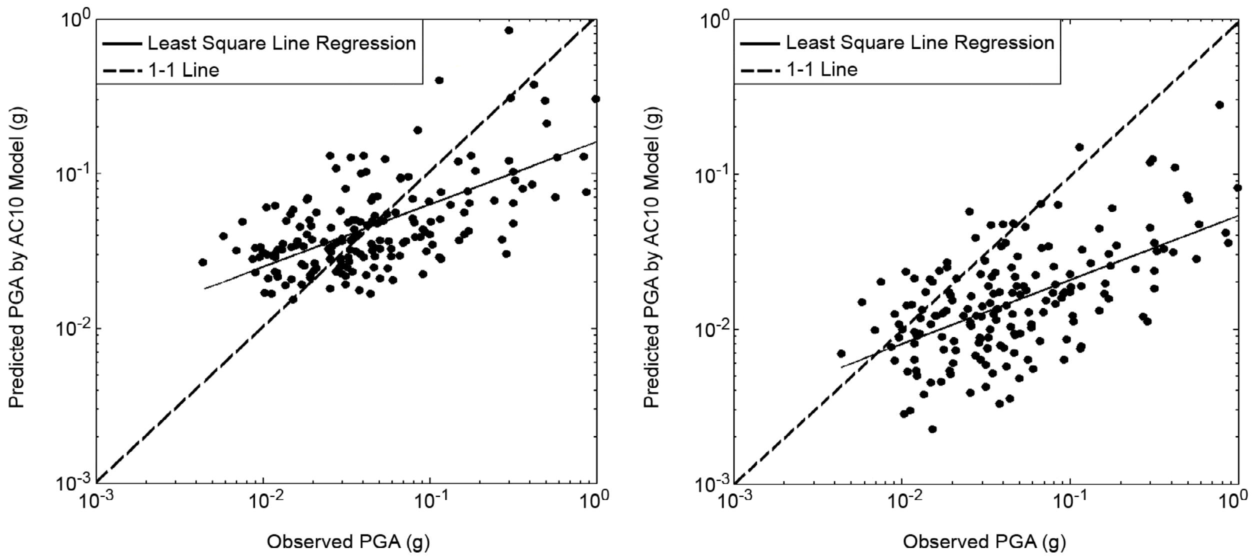

Figure 3. Comparison of the observed versus predicted PGA using the GP model and AC10 relationship, along with the least-squares regression line and the ideal one-to-one line.

6.3. Residuals Bias

One of the important requirements for choosing an appropriate GMPE model is unbiased residuals toward earthquake parameters [44]. In this study, a hypothetical test is applied by means of the generalized linear model regression in order to evaluate the inherent bias. The null hypothesis is that the slope and the Y-intercept of the fitted linear line to the residuals toward different seismic parameters is zero. Two P-values (P_a and P_b) are achieved by this hypothetical test in which the highest P-value indicates the lowest bias in the residuals. For the entire candidate models the residuals, Eq. (4), the inter-event residuals, Eq. (5), and the intra-event residuals, Eq. (6), are calculated based on the available Iranian strong-motion data, in order to examine the bias for the candidate model predictions. Tables (7a) and (7b) show the results of the hypothetical test to examine the bias.

As seen in Tables (7a) and (7b), the mean value of the residuals is calculated in order to show if the model prediction is over-estimated or under-estimated. As a consequence, AC10, Aetal05, Ozetal04, BA08, CY08, AB10, and Bindi10 are

Table 7a. Result of residuals biases.

Model	Mean of Residuals	P-values					
		M_w vs. r_{ij}		R vs. r_{ij}		V_{S30} vs. r_{ij}	
		P_b*	P_a**	P_b	P_a	P_b	P_a
GP model	-0.0104	0.7455	0.7344	0.5559	0.5594	0.9574	0.9878
Setal12	3.8E-04	0.0062	0.0065	0.0013	0.0036	0.8332	0.8477
Zetal12	0.1500	0.7744	0.9690	0.6143	0.1236	0.5498	0.1408
Getal07	-0.1452	0.0070	0.0038	0.1622	0.0304	0.0226	0.2263
AB10	0.3881	0.4316	0.8026	0.1744	0.3388	0.2801	0.0015
AC10	1.1131	0.5009	0.0155	0.6622	1.5E-12	0.2921	2.9E-08
Aetal05	0.3202	0.2048	0.4061	0.0646	0.8866	0.2647	0.0047
Ozetal04	0.8089	0.0309	6.3e-04	0.0588	2.1e-12	0.9052	2.9e-06
KG04	0.0144	0.0094	0.0105	0.7872	0.7365	0.7888	0.8769
Bindi10	0.3128	0.2950	0.1478	0.8660	0.1340	0.8627	0.1184
CB08	0.0648	0.0251	0.0337	0.6090	0.9829	0.6234	0.9569
BA08	0.2871	0.3799	0.6594	0.8932	0.0718	0.9298	0.1952
CY08	0.3114	0.0941	0.2403	0.0594	0.5370	0.7584	0.1223
AS08	0.0016	0.0647	0.1136	0.5566	0.5213	0.2666	0.972

*P_b: P-value for the slope of fitted line by regression.
**P_a: P-value for the Y-intercept of fitted line by regression.

Table 7b. Result of residuals biases.

Model	Mean of Residuals	P-values					
		M_w vs. $r_i^{[inter]}$		R vs. $r_{ij}^{[intra]}$		V_{S30} vs. $r_{ij}^{[intra]}$	
		P_b	P_a	P_b	P_a	P_b	P_a
GP Model	-0.0104	0.6613	0.6019	0.5735	0.6273	0.3669	0.4144
Setal12	3.8E-04	0.0267	0.0247	0.2324	0.2815	0.3007	0.3507
Zetal12	0.1500	0.6052	0.6866	0.5397	0.5968	0.6477	0.6794
Getal07	-0.1452	0.0357	0.0198	0.6396	0.6804	0.2960	0.3451
AB10	0.3881	0.2776	0.4895	0.6783	0.7075	0.9861	0.9875
AC10	1.1131	0.7841	0.1339	0.6371	0.6840	0.0559	0.0835
Aetal05	0.3202	0.2200	0.3409	0.4526	0.4974	0.8133	0.8322
Ozetal04	0.8089	0.1901	0.0301	0.4188	0.4855	0.4001	0.4465
KG04	0.0144	0.0428	0.0389	0.1488	0.2127	0.2089	0.2555
Bindi10	0.3128	0.6279	0.4485	0.7008	0.7279	0.2885	0.3406
CB08	0.0648	0.0535	0.0547	0.4862	0.5479	0.1797	0.2247
BA08	0.2871	0.3571	0.4770	0.6895	0.7291	0.4602	0.4903
CY08	0.3114	0.1312	0.2280	0.4407	0.5063	0.1394	0.1788
AS08	0.0016	0.1359	0.1711	0.8290	0.8522	0.0499	0.0758

remarkably under-estimated models. This table again implies that GP model gives more unbiased estimates of LnPGA for the considered database.

6.4. Stability of Ground Motion Prediction Equations [45]

Each GMPE is obviously obtained based on a specific ground motion database. A small change in the chosen ground motion database should not affect significantly on the GMPE outputs. In other words, if a GMPE is strongly sensitive to a small change of ground motion database, then, the predicted values may not be so reliable. To quantify this phenomenon, a sensitivity of the GMPE models to their own databases is evaluated in this section as the following steps:

1. For each GMPE, a reduced number of ground motion records, say N, is selected based on uniformly random selection.
2. The P-values based on Magnitude, Distance and Site condition as well as LLH, R-squared and RMSE are calculated based on the reduced database defined in Step 1.
3. Steps 1 and 2 are repeated for optimized K times to avoid any bias from the random selection process.
4. Step 1, 2 and 3 are repeated for N = 70 to N = 'maximum number of records within data base, with the increment of 10.
5. The obtained indicators that were calculated in Step 2 can be shown versus N e.g. Figure (4) in the case of GP model.

The process of choosing the optimized K factor is summarized as the following steps and the result is shown in Figure (5):

- Select an initial assumption for the number of subsets (GMRs or events), say K = 50, with a constant number of GMRs, say N = 1000, in this study.
- The p-values corresponding to the residuals, versus different types of seismic input parameters, are calculated based on the chosen subset, which was defined in Step 1 (e.g. intra-event residuals versus RRUP).
- The median p-value is calculated and stored.
- Steps 1, 2 and 3 are repeated for T times to avoid any bias from the random selection process, say T = 50, in this study.
- The interval between the maximum and minimum of the stored median p-values in step 3 is

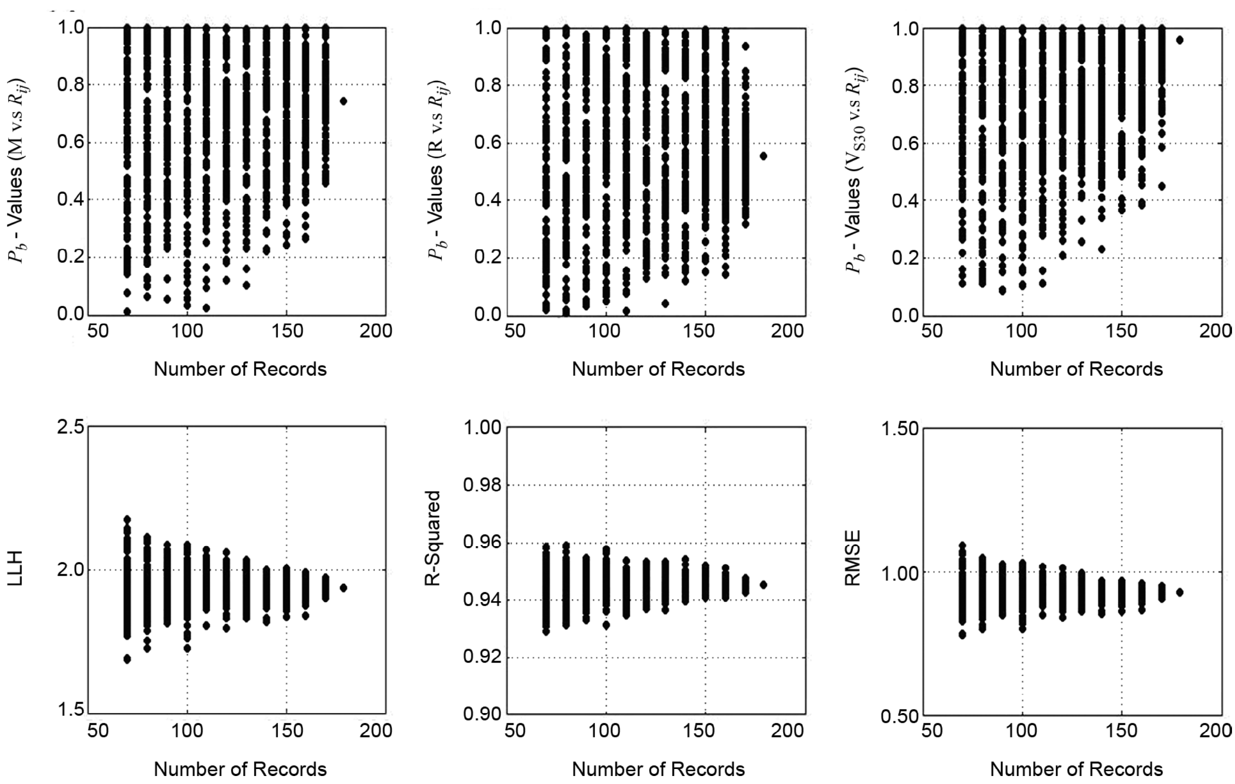

Figure 4. Re-sampling of the GP model for 400 uniformly random selected databases, towards moment magnitude, distance measure, shear-wave velocity, LLH, R^2, and RMSE.

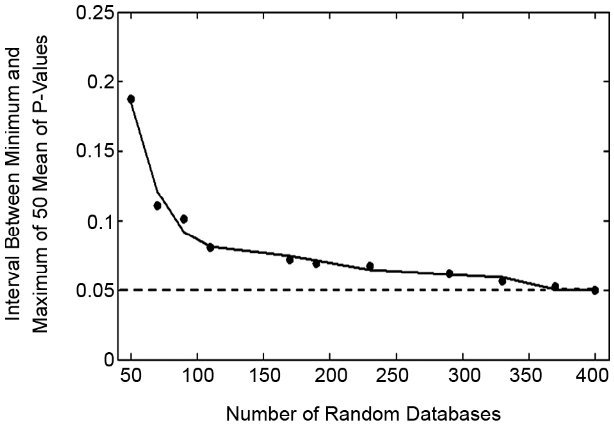

Figure 5. The presses of choosing the number of random data bases for re-sampling analysis.

calculated.

- Steps 1 to 5 are repeated by a constant increment (e.g. 50) in K parameter until the interval in step 5 is less than 5%.
- The obtained intervals calculated in Step 5, are shown versus K factor.

6. The average of each indicator, in Step 2 for a specific N samples, can be calculated as a final indicator. It makes possible to show the final indicator in one plot for all GMPEs as seen in Figure (6) in the case of Moment magnitude, Distance, Site condition, LLH, R-squared and RMSE.

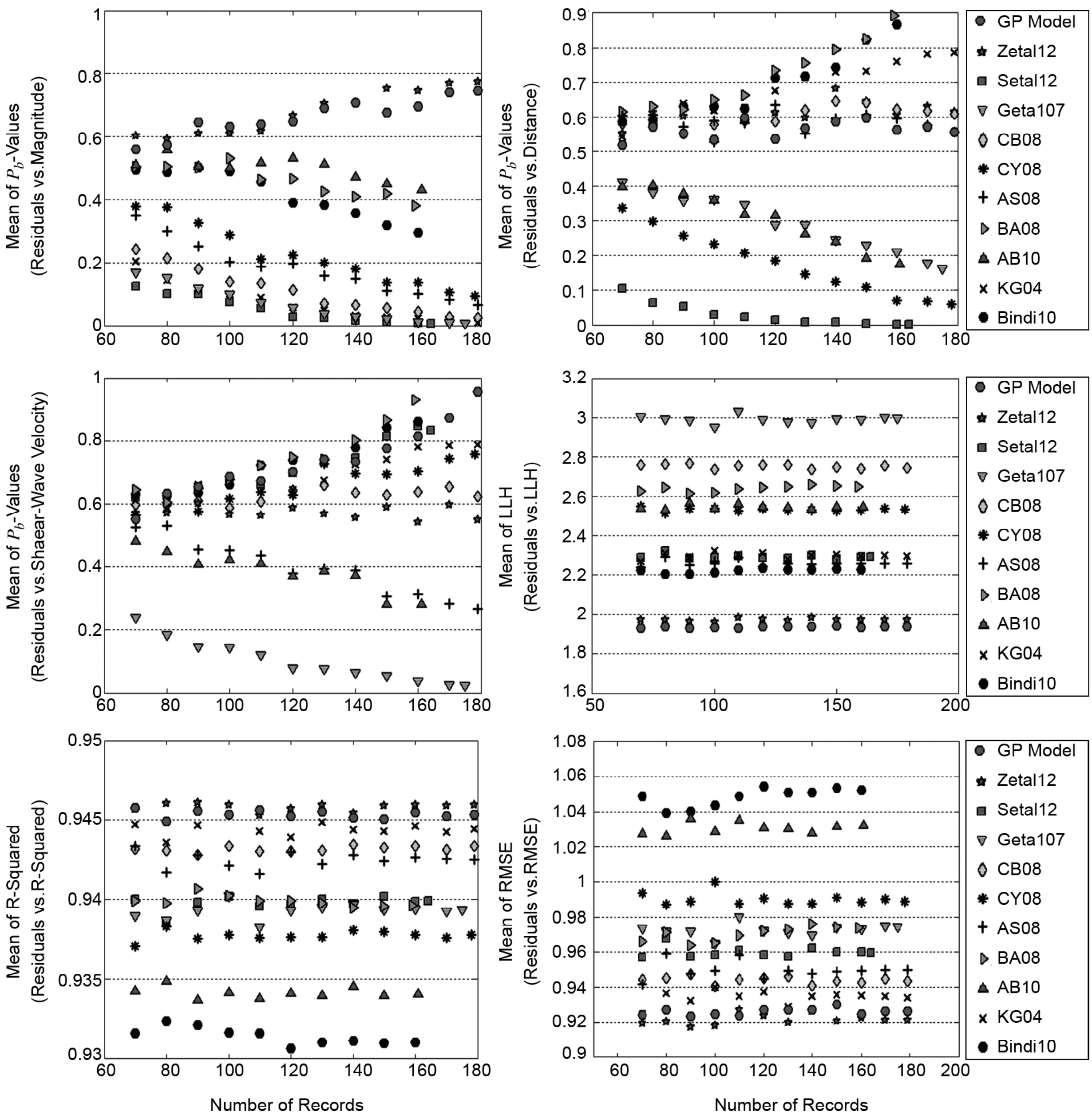

Figure 6. Re-sampling of the candidate models for 400 uniformly random selected databases, towards moment magnitude, distance measure, and shear-wave velocity, LLH, R^2, and RMSE.

It is worth mentioning that an unbiased model should move ascending, as seen in Figure (6), while the sample size is increasing. Except GP model and Zetal12 model, other models have descending trend towards moment magnitude as shown in Figure (6). Additionally, as seen in Figure (6), except Getal07, AB10, CY08, and Setal12 in the case of Pb-values versus distance measure and CB08, AB10, and Getal07 models in the case of Pb-values versus shear-wave velocity measure, other models are more stable. GP model has the lowest LLH value. GP model and Zetal12 have the lowest RMSE values and the highest R-squared values. As a consequence, the re-sampling results confirm that only the achieved model in this study, GP model is superior and it is more stable with ascending behaviour versus all the seismic variables when compared with the other models.

7. Conclusion

In this study, the new predictive PGA model has been obtained by means of the new fitness function based on LLH criteria and re-sampling analysis, for Iranian seismic plateau database. Furthermore, in order to assess the obtained model, the traditional and modern approaches have been employed to evaluate the ground motion prediction equations, which are nominated to be used for Iranian database. Based on the traditional hypothesis test, such as Z-test and Lilliefors test, all the candidate models have normal distribution; however, the majority of the models do not have zero mean residuals. Other results from different statistical and mathematical methods such as error terms (RMSE & MAE), coefficient of determination (R-squared), the information-theoretical method (LLH) and coefficient of efficiency (E) indicate that GP model has the best performance in comparison to the other models. Another important test, which is applied in this paper, is checking for no bias on residuals. GP model is not significantly biased between the other models in this case. Finally, the results based on the new proposed approach in this study for evaluating stability of models, by re-sampling analysis, indicate that GP model is more stable in comparison to the other selected models.

Acknowledgement

The authors are very grateful to the Building and Housing Research Center (BHRC) for providing us with the earthquake catalogue data.

References

1. McGuire, R.K. (1995) Probabilistic seismic hazard analysis and design earthquakes: closing the loop. *Bulletin of the Seismological Society of America*, **85**(5), 1275-1284.

2. Kermani, E., Jafarian, Y., and Baziar, M.H. (2009) New predictive models for the ratio v_{max}/a_{max} of strong ground motions using genetic programming. *International Journal of Civil Engineering*, **7**(4), 246-239.

3. Cabalar, A.F. and Cevik, A. (2009) Genetic programming-based attenuation relationship: an application of recent earthquakes in Turkey. *Computers and Geosciences*, **35**(9), 1884-1896.

4. Sobhaninejad, Gh., Noorzad, A., and Ansari, A. (2007) Combination of generalized approximation method (anfis) and global optimization techniques (genetic algorithm) in estimation strong-ground-motion attenuation laws. *Proceedings of the Fifth International Conf. on Seismology and Earthquake Engineering*, Tehran, Iran.

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

6. Koza, J.R. (1992) Genetic programming: on the programming of computers by means of natural selection. MIT Press, Cambridge, Massachusetts, 840.

7. Goldberg, D.E. (1989) Genetic algorithms in search, optimization, and machine learning. first edition, Addison-Wesley publication, Inc., Reading, MA, 432.

8. Scherbaum, F., Delavaud, E., and Riggelsen, C. (2009) Model selection in seismic hazard analysis: an information-theoretic perspective. *Bulletin of the Seismological Society of America*, **99**(6), 3234-3247.

9. Holland, J.H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press,

Ann Arbor, 228.

10. Chaiyaratana, N. and Zalzala, A.M.S. (1997) Recent developments in evolutionary and genetic algorithms: theory and applications. Proceeding of Genetic Algorithms in Engineering Systems: Innovations and Applications. GALESIA 97, *Second International Conference on (Conf. Publ. No. 446). IET*.

11. Seyed-Nabavi, M. (1979) Seismic activity in Iran 1963-1976. *Journal of the Earth and Space Physics*, **8**, 13-64.

12. Shoja-Taheri, J. and Niazi, M. (1981) Seismicity of the Iranian plateau and bordering regions. *Bulletin of the Seismological Society of America*, **71**(2), 477-489.

13. Ambraseys, N.N. and Melville, C.P. (1982) *A History of Persian Earthquakes*. Cambridge University Press, New York.

14. Berberian, M. and Mohajer-Ashjai, A. (1977) Seismic Risk Map of Iran, A Proposal, 1:5,000,000. In: *Contribution to the Seismotectonics of Iran, Part III* (ed. M. Berberian). Geol. Min. Surv. Iran.

15. Saffari, H., Kuwata, Y., Takada, S., and Mahdavian, A. (2012) Updated PGA, PGV, and spectral acceleration attenuation relations for Iran. *Earthquake Spectra*, **28**(1), 257-276.

16. Building and Housing Research Center of Iran (2007) *Iranian Code of Practice for Seismic Resistant Design of Buildings*. Standard No. 2800, 3rd edition, Tehran, Iran.

17. Silva, S. (2007) *GPLAB_A Genetic Programming Toolbox for MATLAB*, Version 3, http://gplab.sourceforg.net.

18. Johari, A., Habibagahi, G., and Ghahramani, A. (2006) Prediction of soil-water by characteristic curve using genetic programming. *Journal of Geotechnical and Geoenvironmental Engineering*, **132**(5), 661-665.

19 Berberian, M. (1976) *Contribution to the Seismotectonics of Iran (Part II)*. Geological Survey of Iran, Report 39, 518p.

20. Chen, S.Z. and Atkinson, G.M. (2002) Global comparisons of earthquake source spectra. *Bulletin of the Seismological Society of America*, **92**(3), 885-895.

21. PEER, Pacific Earthquake Engineering Research Center, *NGA Database*, University of California, Berkeley. http://peer.berkeley.edu/ngawest/nga.

22. Abrahamson, N. and Silva, W. (2008) Summary of the abrahamson & Silva NGA ground-motion relations. *Earthquake Spectra*, **24**(1), 67-97.

23. Boore, D.M. and Atkinson, G.M. (2008) Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s. *Earthquake Spectra*, **24**(1), 99-138.

24. Campbell, K.W. and Bozorgnia, Y. (2008) NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s. *Earthquake Spectra*, **24**(1), 139-171.

25. Chiou, B.S. and Youngs, R.R. (2008) An NGA model for the average horizontal component of peak ground motion and response spectra. *Earthquake Spectra*, **24**(1), 173-215.

26. Idriss, I.M. (2008). An NGA empirical model for estimating the horizontal spectral values generated by shallow crustal earthquakes. *Earthquake Spectra*, **24**(1), 217-242.

27. Bommer, J.J., Douglas, J., Scherbaum, F., Cotton, F., Bungum, H., and Fah, D. (2010) On the selection of ground-motion prediction equations for seismic hazard analysis. *Seismological Research Letter*, **81**(5), 783-793.

28. Cotton, F., Pousse, G., , Bonilla, F., and Scherbaum, F. (2008) On the discrepancy of recent european ground-motion observations and predictions from empirical models: analysis of KiK-net accelerometric data and point-sources stochastic simulations. *Bulletin of the Seismological Society of America*, **98**(5), 2244-2261.

29. Zafarani, H. and Soghrat, M. (2012) Simulation of ground motion in the Zagros region of Iran

using the specific barrier model and the stochastic method. *Bulletin of the Seismological Society of America*, **102**(5), 2031-2045.

30. Ghodrati Amiri, G., Mahdavian, A., and Manouchehri Dana, F. (2007) Attenuation Relationships for Iran. *Journal of Earthquake Engineering*, **11**(4), 469-492.

31. Akkar, S. and Bommer, J.J. (2010) Empirical equations for the prediction of PGA, PGV, and spectral accelerations in europe, the mediterranean region, and the middle east. *Seismological Research Letter*s, **81**(2), 195-206.

32. Akkar, S. and Cagnan, Z. (2010) A local ground-motion predictive model for Turkey, and its comparison with other regional and global ground-motion models. *Bulletin of the Seismological Society of America*, **100**(6), 2978-2995.

33. Ambraseys, N.N., Douglas, J., Sarma, S.K., and Smit, P.M. (2005) Equations for the estimation of strong ground motions from shallow crustal earthquakes using data from europe and the middle east: horizontal peak ground acceleration and spectral acceleration. *Bulletin of Earthquake Engineering*, **3**(1), 1-53.

34. Ozbey, C., Sari, A., Manuel, L., Erdik, M., and Fahjan, Y. (2004) An empirical attenuation relationship for northwestern Turkey ground motion using a random effects approach. *Soil Dynamics and Earthquake Engineering*, **24**(2), 115-125.

35. Kalkan, E. and Gülkan, P. (2007) Empirical attenuation equations for vertical ground motion in Turkey. *Earthquake Spectra*, **20**(3), 853-882.

36. Bindi, D., Luzi, L., Massa, M., and Pacor, F. (2010) Horizontal and vertical ground motion prediction equations derived from the Italian accelerometric archive (ITACA). *Bulletin of Earthquake Engineering*. **8**(5), 1209-1230.

37. Kaklamanos, J., Baise, L.G., and Boore, D.M. (2011) Estimating unknown input parameters when implementing the NGA ground-motion prediction equations in engineering practice. *Earthquake Spectra*, **27**(4), 1219-1235.

38. Elnashai, A.S. and Di Sarno, L. (2008) *Fundamentals of Earthquake Engineering*. Wiley, 347p.

39. Montgomery, C.D. and Runger C.G. (2003) *Applied Statistics and Probability for Engineers*. Wiley, 706p.

40. Scherbaum, F., Cotton, F., and Smit, P. (2004). On the use of response spectral reference data for the selection of ground-motion models for seismic hazard analysis: the case of rock motion. *Bulletin of the Seismological Society of America*, **94**(6), 341-348.

41. Delavaud, E., Scherbaum, F., Kuehn, N., and Riggelsen, C. (2009) Information-theoretic selection of ground-motion prediction equations for seismic hazard analysis: an applicability study using Californian data. *Bulletin of the Seismological Society of America*, **99**(6), 3248-3263.

42. Nash, J.E. and Sutcliffe, J.V. (1970) River flow forecasting through conceptual models: part I, a discussion of principles. *Journal Hydrology*, **10**(3), 282-290.

43. Legates, D.R. and McCabe, G.J. (1999) Evaluating the use of "Goodness-of-Fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, **35**(1), 233-241.

44. Campbell, K.W. and Bozorgnia, Y. (2007) *Campbell-Bozorgnia NGA Ground Motion Relations for the Geometric Mean Horizontal Component of Peak and Spectral Ground Motion Parameters*. PEER 2007/02, Pacific Earthquake Engineering Research Center.

45. Azarbakht, A., Rahpeyma, S., and Mousavi, M. (2014) A new methodology for assessment the stability of ground motion prediction equations. Accepted in *Bulletin of the Seismological Society of America*.

Appendix Table A.

Table A. The records that is used in this study

N.O	Event	Station Name	Record ID	Latitude (deg)	Longitude (deg)	Epi-Distance (km)	Vs30 (m/s)
1	1978/09/16	Bajestan	1086.00	34.52	58.18	166.00	710.00
2		Boshroyeh	1083-1	33.86	57.43	74.00	564.00
3		Deyhook	1082-1	33.29	57.50	19.00	843.00
4		Tabas	1084-1	33.58	56.92	58.00	645.00
5	1979/11/27	Gonabad	1142-1	34.35	58.68	94.00	529.00
6		Ghaen	1139.00	33.73	59.22	53.00	889.00
7		Khezri	1140.00	34.02	58.81	76.00	701.00
8		Bajestan	1134-2	34.52	58.18	143.00	710.00
9		Birjand	1137.00	32.87	59.21	137.00	787.00
10		Kakhk	1135.00	34.14	58.66	91.00	1961.00
11		Sadeh	1138-1	33.33	59.24	90.00	1180.00
12	1990/06/20	Abbar	1362-1	36.93	48.95	32.00	691.00
13		Ghazvin	1353-1	36.26	50.00	98.00	456.00
14	1994/06/20	Babandar	1498.00	28.98	53.22	55.00	885.00
15		Farashband	1497.00	28.87	52.07	63.00	630.00
16		Firuzabad	1493-2	28.84	52.57	29.00	894.00
17		Kavar	1491.00	29.20	52.69	18.00	753.00
18		Maymand	1490-2	28.87	52.75	24.00	881.00
19		Zanjiran	1502-9	29.07	52.62	10.00	936.00
20		Zarrat	1492-16	29.09	52.85	19.00	720.00
21	1997/02/04	Marvertappeh	1674.00	37.90	55.96	136.00	538.00
22	1997/02/28	Germi	1702.00	39.05	48.06	109.00	712.00
23		Kariq	1833-2	37.92	48.06	35.00	589.00
24		Namin	1724.00	38.42	48.48	71.00	1236.00
25	1997/05/10	Marak	1750-2	32.92	59.43	105.00	872.00
26		Mussaviyeh	1770.00	33.29	58.91	105.00	848.00
27		Sangan	1753.00	34.40	60.25	80.00	941.00
28		Khezri	1740.00	34.02	58.81	188.00	701.00
29		Feyzabad	1741.00	35.01	58.78	162.00	561.00
30		Gonabad	1742.00	34.37	58.68	120.00	683.00
31		Mud	1751.00	32.71	59.52	127.00	961.00
32	1998/03/14	Abaraq	1864-1	28.10	57.23	90.00	641.00
33		Baqein	1866.00	30.19	56.82	76.00	516.00
34	1999/08/21	Noorabad	2251.00	34.07	47.97	36.00	758.00
35		Boroujerd	2183-1	33.89	48.75	69.00	579.00
36		Aleshtar	2196-2	33.86	48.25	47.00	621.00
37	1999/05/06	Ghaemiyeh	2126-3	29.85	51.59	48.00	617.00
38		Kazeroon	2121-2	29.62	51.67	28.00	352.00
39		Balaadeh	2131-2	29.83	52.40	29.00	1380.00
40		Gooyom	2123-2	29.83	52.40	56.00	598.00
41		Khan Zeynioun	2130-1	29.67	52.15	26.00	773.00
42	1999/05/06	Balaadeh	2131-3	29.29	51.94	18.00	1380.00
43		Gooyom	2123-3	29.83	52.40	64.00	598.00
44		Khan Zeynioun	2130-2	29.67	52.15	36.00	773.00
45	1999/10/31	Kazeroon	2216-1	29.62	51.67	36.00	352.00
46		Romghan	2217.00	29.37	52.16	34.00	1362.00
47		Ghaemiyeh	2218-3	29.85	51.59	60.00	617.00
48		Balaadeh	2219-12	29.29	51.94	19.00	1380.00
49	2002/04/24	Armanijan	2706-2	34.61	47.35	25.00	390.00
50		Aran	2707-2	34.41	47.92	58.00	175.00
51		Bistoon	2708-2	34.38	47.43	35.00	750.00
52		Sahneh	2710-2	34.47	47.68	39.00	375.00
53		Sonqor	2711-2	34.78	47.60	37.00	1477.00
54		Lenj Ab	2747-2	34.87	47.28	40.00	375.00
55	2002/06/22	Abegarm	2748-1	35.76	49.28	18.00	199.00
56		Bahar	2750.00	34.89	48.44	108.00	913.00
57		Bakandi	2787-1	36.40	49.57	93.00	308.00
58		Buinzahra	2759.00	35.77	50.06	79.00	255.00
59		Darsejin	2769-2	36.02	49.24	46.00	636.00
60		Deh-Jalal	2768.00	36.32	48.70	88.00	748.00
61		Ghohrud	2778.00	35.47	48.06	105.00	414.00
62		Goltappeh	2777.00	35.22	48.20	102.00	1077.00
63		Kabodarahang	2754-1	35.21	48.72	65.00	613.00
64		Razan	2756-1	35.39	49.03	33.00	314.00
65		Saei-Ghale	2772.00	36.31	49.07	71.00	642.00
66		Shirinsu	2781.00	35.49	48.45	70.00	813.00
67	2002/12/24	Armanijan	2933-3	34.61	47.35	24.00	390.00
68		Aran	2934.00	34.41	47.92	48.00	175.00
69		Bistoon	2935.00	34.38	47.43	28.00	750.00
70		Sahneh	2936-1	34.47	47.68	29.00	375.00
71		Sonqor	2937-1	34.78	47.60	34.00	1477.00
72		Lenj Ab	2999-1	34.87	47.28	44.00	375.00
73	2003/07/10	Hajiabad-3	3040-1	28.35	54.42	27.00	561.00
74		Jouyom	3041-1	28.26	53.98	21.00	1244.00
75		Zahedshahr	3042-1	28.74	53.80	60.00	390.00
76		Jahrom	3045-1	28.50	53.55	64.00	375.00

Table A. Continue.

N.O	Event	Station Name	Record ID	Latitude (deg)	Longitude (deg)	Epi-Distance (km)	Vs30 (m/s)
77	2003/07/10	Hajiabad-3	3040-3	28.35	54.42	38.00	561.00
78		Jouyom	3041-2	28.26	53.98	18.00	1244.00
79		Zahedshahr	3042-2	28.74	53.81	63.00	390.00
80		Jahrom	3045-2	28.50	53.55	61.00	375.00
81	2003/08/21	Fahraj	3067.00	28.95	58.89	87.00	280.00
82		Nosratabad	3069.00	29.86	59.98	97.00	1154.00
83		Ryqan	3070.00	28.65	59.01	85.00	437.00
84	2003/11/28	Hajiabad-3	3134-2	28.35	54.42	44.00	561.00
85		Jouyom	3135.00	28.26	53.98	27.00	1244.00
86		Doobaran	3136-5	28.41	54.18	30.00	1363.00
87	2003/12/26	Bam	3168-2	29.09	58.35	4.00	499.00
88		Jiroft	3170-2	28.67	57.74	74.00	343.00
89		Mohamadabad	3162-1	28.91	57.89	49.00	507.00
90		Anduhjerd	3164.00	30.23	57.75	143.00	566.00
91		Golbaf	3155-2	29.89	57.73	111.00	365.00
92		Joshan	3156.00	30.12	57.61	140.00	776.00
93	2004/05/28	Hasan Keyf	3333.00	36.50	51.15	45.00	339.00
94		Moalemkelayeh	3367.00	36.45	50.47	100.00	490.00
95		Noshahr	3368-1	36.65	51.49	43.00	165.00
96		Nur	3369-1	36.57	52.01	53.00	178.00
97		Taleqn	3318.00	36.18	50.76	72.00	462.00
98	2004/10/07	Aliabad	3542.00	36.90	54.85	52.00	562.00
99		Bandar-e-Gaz	3557-2	36.76	53.95	63.00	347.00
100		Dibaj	3590.00	36.43	54.23	87.00	526.00
101		Gomishan	3546.00	37.07	54.08	47.00	322.00
102		Gonbad-e-Kavoos	3544.00	37.24	55.16	69.00	402.00
103		Gorgan	3545.00	36.84	54.39	46.00	291.00
104		Inche Broun	3560-1	37.46	54.72	51.00	283.00
105		Minoodasht	3639-1	37.23	55.37	86.00	449.00
106		Ramyan	3551.00	37.02	55.14	68.00	827.00
107	2005/01/10	Aq Qala	3608.00	37.01	54.46	59.00	341.00
108		Aliabad	3612.00	36.90	54.85	75.00	562.00
109		Bandar-e-gaz	3609.00	36.76	53.95	98.00	347.00
110		Gomishan	3607.00	37.07	54.08	67.00	322.00
111		Gorgan	3623.00	36.84	54.39	77.00	291.00
112		Inche Broun	3618.00	37.46	54.72	36.00	283.00
113		Minoodasht	3639-5	37.23	55.37	85.00	449.00
114		Ramyan	3621-2	37.02	55.14	80.00	827.00
115	2005/02/22	Chatrud	3660-1	30.61	56.91	27.00	852.00
116		Davaran	3702.00	30.58	56.19	57.00	752.00
117		Deh-e-Loulo	3679.00	30.53	57.29	61.00	617.00
118		Horjand	3688.00	30.68	57.15	42.00	999.00
119		Ravar	3661.00	31.26	56.79	55.00	853.00
120		Zarand	3671-1	30.81	56.58	19.00	226.00
121	2005/11/27	Bandar-e-Abas1	3912.00	27.19	56.29	62.00	337.00
122		Bandar-e-Abas2	3917.00	27.19	56.30	62.00	375.00
123		Bandar-e-Khamir	3913.00	26.95	55.58	39.00	679.00
124		Kahoorestan	3910.00	27.22	55.56	61.00	807.00
125		Qeshm	3909.00	26.96	56.28	45.00	757.00
126		Fin	3916.00	27.63	55.90	96.00	681.00
127		Suza	3915-1	26.78	56.07	21.00	1334.00
128	2006/03/30	Chalan Choolan	4027-5	33.66	48.91	24.00	428.00
129		Boroujerd	4023-2	33.89	48.75	34.00	579.00
130		Dorood	4022-1	33.49	49.06	37.00	771.00
131		Khoramabad1	4019-1	33.49	48.36	47.00	375.00
132		Chaghalvandi	4018-2	33.66	48.55	29.00	616.00
133		Shool Abad	4055-2	33.18	49.19	67.00	1084.00
134		Tooshk-e-Ab-e-Sar	4035-2	33.77	48.57	31.00	891.00
135		Darreh-Asbar	4052-2	33.45	49.06	40.00	935.00
136	2006/03/31	Chaghalvandi	4018-3	33.66	48.55	35.00	616.00
137		Khoram Abad	4019-2	33.49	48.36	54.00	375.00
138		Dorood	4022-2	33.49	49.06	23.00	771.00
139		Aleshtar	4025.00	33.87	48.26	67.00	621.00
140		Chalan Choolan	4027-8	33.66	48.91	13.00	428.00
141		Darreh-Asbar	4052-3	33.45	49.06	67.00	935.00
142		Noor Abad	4024.00	34.07	47.97	101.00	758.00
143		Shool Abad	4055-3	33.18	49.19	56.00	1084.00
144		Tooshk-e-Ab-e-Sar	4035-3	33.77	48.57	38.00	891.00
145	2006/03/31	Khoram Abad	4136.00	33.49	48.36	52.00	375.00
146		Shool Abad	4055-4	33.18	49.19	78.00	1084.00
147		Tooshk-e-Ab-e-Sar	4035-6	33.77	48.57	31.00	891.00
148		Dorood	4032.00	33.49	49.06	47.00	771.00
149		Boroujerd	4034.00	33.89	48.75	32.00	579.00
150		Chaghalvandi	4044.00	33.66	48.55	33.00	616.00
151	2006/06/28	Tomban	4147-13	26.77	55.86	13.00	778.00
152		Bandar-e-Khamir	4152.00	26.95	55.58	33.00	679.00
153		Qeshm	4128.00	26.96	56.28	52.00	757.00
154		Bandar-e-Abas2	4144.00	27.19	56.30	68.00	375.00

Table A. Continue.

N.O	Event	Station Name	Record ID	Latitude (deg)	Longitude (deg)	Epi-Distance (km)	Vs30 (m/s)
155	2008/05/05	Doobaran	4573.00	28.41	54.18	33.00	1363.00
156		Jouyom	4574.00	28.26	53.98	14.00	1244.00
157		Zahedshahr	4575.00	28.74	53.81	65.00	390.00
158	2008/09/10	Bandar-e-Khamir	4672.00	26.95	55.58	38.00	679.00
159		Tomban	4686-3	26.77	55.86	23.00	778.00
160		Suza	4678-1	26.78	56.07	40.00	1334.00
161		Tabl	4675-1	26.76	55.73	17.00	931.00
162		Kahoorestan	4676.00	27.22	55.56	66.00	807.00
163	2008/09/11	Tabl	4675-2	26.76	55.73	22.00	931.00
164		Suza	4678-5	26.78	56.07	47.00	1334.00
165		Tomban	4686-19	26.77	55.86	30.00	778.00
166	2008/09/17	Bandar-e-Abas1	4687-1	27.19	56.29	60.00	337.00
167		Suza	4690-1	26.78	56.07	17.00	1334.00
168		Qeshm	4688-1	26.96	56.28	41.00	757.00
169	2008/12/07	Suza	4732-2	26.78	56.07	35.00	1334.00
170		Tabl	4735.00	26.76	55.73	14.00	931.00
171		Bandar-e-Abas1	4734.00	27.19	56.29	70.00	337.00
172		Bandar-e-Khamir	4736.00	26.95	55.58	25.00	679.00
173	2008/12/08	Bandar-e-Abas1	4742.00	27.19	56.29	68.00	337.00
174		Suza	4739-1	26.78	56.07	33.00	1334.00
175		Tabl	4741-1	26.76	55.73	15.00	931.00
176		Qeshm	4737-1	26.96	56.28	54.00	757.00
177	2008/12/09	Suza	4739-2	26.78	56.07	30.00	1334.00
178		Qeshm	4737-2	26.96	56.28	55.00	757.00
179		Tabl	4741-2	26.76	55.73	16.00	931.00