Selection of Ground Motion Prediction Models for Seismic Hazard Analysis in the Zagros Region, Iran

M. Noorizadeh, M.Mousavi
Arak University, Iran

SUMMARY:
The main objective of this paper is to assess a set of candidate ground motion models in the Zagros region of Iran. The candidate models were chosen from three categories: local models that have been developed based on the local data, regional models corresponding to Europe and Middle East data sets and finally NGA (Next Generation Attenuation) models. Two different statistical approaches were applied for evaluation of these models, the first the LH method and the second the LLH. One of the most significant results of this study is that local ground motion models show more consistency with the recorded data than do NGA models.

Keywords: epsilon Ground motion models, Evaluation of fitness, Ranking, Zagros, Iran.

1. INTRODUCTION

The selection of ground motion prediction models, and the determination of the contribution weight to assign to each of them, is a fundamental component of any seismic hazard analysis. It has been demonstrated that the uncertainty corresponding to the selection of the attenuation model influences the hazard results more than other aspects of seismicity modelling (Toro 2006). This epistemic uncertainty is often treated within the expert opinion approach through a logic tree framework (Budnitz et al 1997). The branch weights in a logic tree framework correspond to the degree of belief of experts in different prediction models. Although seemingly straightforward, the logic tree approach is a challenging tool to capture this uncertainty. Some professionals (e.g. Krinitzsky 1995) believe that any attempt to assign numbers to degrees of belief, which are by nature personal and indefinable. From another point of view, it is indicated that due to the informal selection of the branch models and weights, the potential pitfalls regarding the construction and the use of logic trees is a rational expection (Bommer and Scherbaum et al 2008). In addition to these general considerations, the absence of domestic experienced domestic experts is additional impediment with the use of logic trees in regions such as Iran. Because of these problems with "expert opinion" approaches we apply a recently developed statistically based scheme to assign the logic tree weights. The results can be used for seismic hazard studies in the Iran Zagros region, within a logic tree framework.

2. DATA DRIVEN GROUND MOTION MODEL SELECTION

Given a set of data recorded in the real conditions of a specified region, how can one quantitatively judge different candidate ground motion models? This is the key question of any data-driven model selection. The statistical analysis of the residuals is the prime technique to distinguish the validity of these models. Because ground-motion models are commonly expressed in terms of logarithmic quantities, the residual is defined as the subtraction of the logarithmic-model predictions from the logarithms of the observed values, divided by the corresponding standard deviations of the logarithmic model:

$$\mathrm{r} = \frac{\log(SA_{obs}) - \log(SA_{pre})}{s_{SA}} \tag{2.1}$$

Where, SA_{obs} corresponds to the observed acceleration response spectra in a specified period, and SA_{pre} and σ_{SA} are the mean and the standard deviation of the predicted response spectra, respectively, using a given ground motion model. Ideally, the residual so defined is normally distributed with zero mean and unit variance. The fitness degree of the resulting residuals to this distribution defines the compatibility of the applied ground motion model with the recorded data. Statistical tests to measure the goodness of the issued fitness can be invoke, for example, the z-test can be used to test the hypothesis that the mean of residuals is zero (Montgomery et al 2003). Also, the variance test may be used to test the residuals for unit variance (Montgomery et al 2003). In order to test the shape of the residual distribution, it is convenient to perform a Lilliefors test of the default null hypothesis that the residuals sample comes from a normal distribution (Montgomery et al 2003). It is important to emphasize that most of the traditional tests only checks for one hypothesis, i.e. normal distribution, zero mean or unit standard deviation. As a consequence, they are not perfect tools for evaluation and ranking of the considered ground motion models. Due to this limitation, the likelihood-based measure (LH) has been recently emerged as another goodness test which is suitable for measuring not only the model fit, but also the underlying statistical assumptions (Scherbaum et al 2004). One of the deficiencies of the above mentioned LH method is that it still requires a few subjective decisions e.g. thresholds for acceptability. The dependency of the results on the sample size is another drawback of these methods. In order to overcome to these problems, a modern information-theoretic approach has been proposed recently (Scherbaum et al 2009). This method is more general than LH method and, in addition, does not depend on ad hoc assumptions e.g. size of samples and significant thresholds (Scherbaum et al 2009). In this study, the information-theoretic approach in combination with the LH method and other goodness-of-fit measures are used to judge about the compatibility of the candidate ground motion models with the ground motion data recorded in the Zagros region, Iran.

3. THE TESTING GROUND MOTION DATASET

To test the effectiveness of candidate ground motion models 21 significant earthquakes have been chosen from the Zagros region of Iran, each with moment magnitude between 5.0 to 6.2. Five significant events from East-Central Iran region have been added to the testing dataset due to the inadequate magnitude range of recorded ground motions in the target region. Table 1 shows the information about each of the events with the corresponding reference .A total of 114 records have been extracted from the website of the Building and Housing Research Center (BHRC) in Iran. Fig. 1 shows the magnitude-distance distribution of the employed ground motion records. The different stations are categorized into three different soil classes: Rock for V_{S30}>750 m/s, Stiff soil for 375 m/s<V_{S30}<750 m/s and soft soil for V_{S30}<275 m/s as shown in Fig. 1.

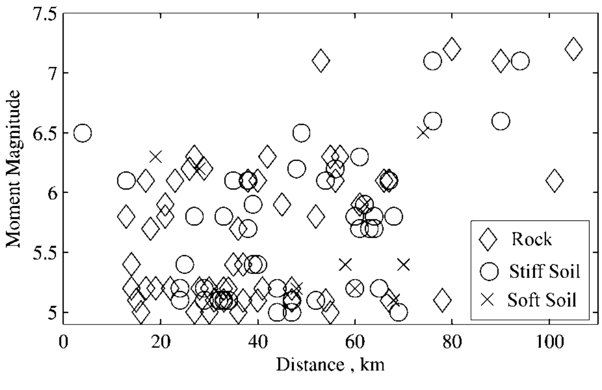

Figure 1. The magnitude – distance distribution of the employed records used in this study

Table 1. Information about the Zagros earthquakes used in this study

No.	Event Date	Time	Mw	Depth (km)	§N	Reference of Mw
*1	1979/11/27	15:36	7.1	10	4	HRVD
*2	1997/05/10	07:57	7.2	13	3	HRVD
*3	1998/03/14	19:40	6.6	5	2	HRVD
4	1999/08/21	05:31	5.0	25	3	HRVD
5	1999/05/06	23:00	6.2	7	5	HRVD
6	1999/05/06	23:13	5.7	10	3	HRVD
7	1999/10/31	15:09	5.2	15	4	HRVD
8	2002/04/24	19:48	5.4	25	6	HRVD
9	2002/12/24	17:03	5.2	20	6	HRVD
10	2003/07/10	17:06	5.8	10	4	HRVD
11	2003/07/10	17:40	5.7	15	4	HRVD
12	2003/11/28	23:19	5.0	25	3	HRVD
*13	2003/12/26	01:56	6.5	3	3	HRVD
*14	2005/02/22	02:25	6.3	10	6	HRVD
15	2005/11/27	10:22	5.9	12	6	HRVD
16	2006/03/30	19:36	5.1	20	8	HRVD
17	2006/03/31	01:17	6.1	12	9	HRVD
18	2006/03/31	11:54	5.1	26	6	HRVD
19	2006/06/28	21:02	5.8	12	4	HRVD
20	2008/05/05	21:57	5.2	12	3	HRVD
21	2008/09/10	11:00	6.1	12	5	HRVD
22	2008/09/11	02:16	5.2	7	3	HRVD
23	2008/09/17	17:43	5.2	12	3	HRVD
24	2008/12/07	13:36	5.4	12	4	HRVD
25	2008/12/08	14:41	5.1	12	3	HRVD
26	2008/12/09	15:09	5.0	14	3	HRVD

* Selected from East-Central Iran;
§N, Number of used records;
HRVD: Harvard seismology

4. CANDIDATE GROUND MOTION ATTENUATION MODELS

Based on different studies on the seismotectonic characteristics of Iran, it has been shown that all of the Iranian plateau earthquakes are shallow, intra-plate events (Berberian 1976). According to these criteria, candidate ground motion models were selected from three categories:

- Ground motion models developed specially for the region of Iran (Category 1)
- Ground motion models developed for the Mideast-Europe region (Category 2)
- Global ground motion models developed by the "Next Generation of Ground-Motion Attenuation Models" (NGA) project (Category 3)

The Next Generation Attenuation (NGA) project has developed a series of ground motion models intended for application to geographically diverse regions; the only constraint is that the region be tectonically active with earthquakes occurring in the shallow crust. Ground motion models have been selected according to the criteria proposed recently by Bommer (Bommer et al 2010). Two significant points were particularly considered;

- Models that have been superseded by a more recent publication were avoided.
- Models that lack either non-linear magnitude dependence or magnitude-dependent decay with

distance were excluded.

Finally, it should be noted that epistemic uncertainties may be influenced by different measures of distance and magnitude. Different attenuation relationships use different forms of distance measures (such as R_{epi}, R_{hypo}, R_{rup}, R_{jb}, etc.) and magnitude scales (M_s, M_b, M_w, etc.) for prediction of ground motion parameters. Here, the selected ground motion models are summarized in Table 2.

Fig. 2 compares the selected ground motion models for the scenario M_w=5.8 and R=45km. This scenario corresponds to the average magnitude and distance of the used dataset, strike slip faulting mechanism, and shear wave velocity 750 m/sec.

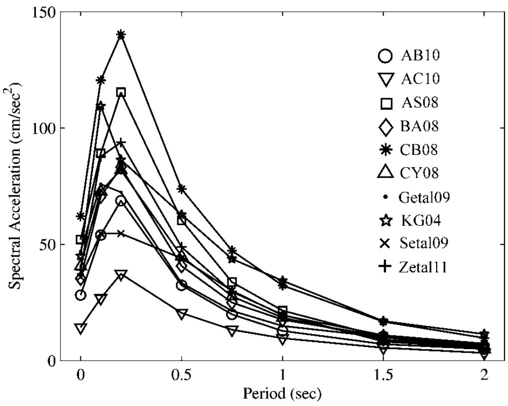

Figure 2. Comparison of spectral acceleration of different ground motion models for scenario M_w5.8 and R=45km, strike slip faulting mechanism, and shear wave velocity 750 m/sec.

Table 2. Candidate ground motion models

No	Model	Abb.	Dominant Region	Category	Mw	Distance
1	Zafarani et al. (2011)	Zetal11	Iran	1	4.4 - 7.5	2 - 200 km
2	Ghasemi et al. (2009)	Getal09	Iran	1	5.0 - 7.4	5 - 500 km
3	Sharma et al. (2009)	Setal09	India, Iran	2	5.0 - 7.0	0 - 200 km
4	Akkarand Cagnan(2010)	AC10	Turkey	2	3.5 - 7.6	0 - 200 km
5	Akkar & Bommer(2010)	AB10	Europe, Middle east	2	5.0 - 7.6	0 - 100 km
6	Kalkan & Gulkan (2004)	KG04	Turkey	3	4.0 - 7.4	1 - 200 km
7	Abrahamson & Silva (2008)	AS08	California	3	5.0- 8.5	0 - 200 km
8	Boore & Atkinson (2008)	BA08	California	3	5.0-8.0	0 - 200 km
9	Campbell & Bozorgn (2008)	CB08	California	3	4.0- 7.5	0 - 200 km
10	Chiouand & Youngs(2008)	CY08	California	3	4.2- 7.9	0 - ~100 km

5. GROUND MOTION MODELS RANKING

For each of the ground motion records synthetic acceleration response spectra, S_a (T), have been generated using the 10 candidate ground motion prediction models presented in Table 2 over seven periods including (0.1 sec, 0.2 sec, 0.5 sec, 0.75 sec, 1.0 sec, 1.5 sec, and 2 sec) and the peak ground acceleration (PGA). By using Eqn. 2.1, the residual set associated by each model is achievable for any arbitrary period. As an example, the residual distribution of S_a (T=1.0 sec) is shown in Fig. 3 for all of ground motion models. The standard normal distribution, as the ideal distribution of the residuals, is also plotted for each case in Fig. 3.

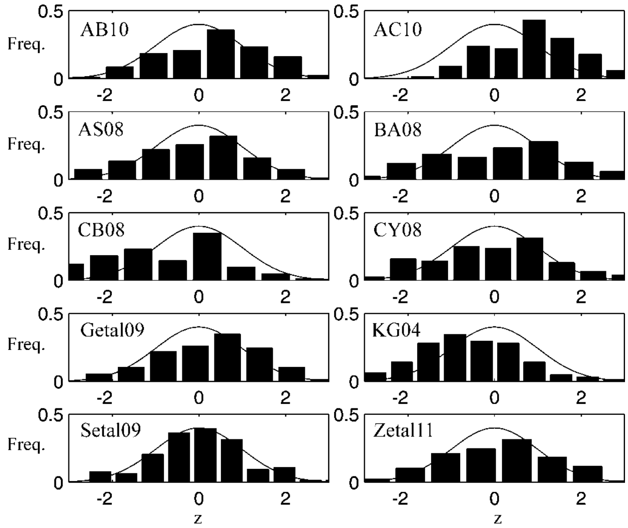

Figure 3. Residual distribution of Sa(T=1.0sec) with respect to different ground-motion models. Solid line shows the expected distribution function for a standard normal distribution.

Here, three statistical analyzes are applied to gain an insight into the goodness of a standard normal distribution to the residuals. Due to the space limitations, the statistical analysis results for S_a(T=1.0sec) are explained briefly, here.

5.1. The z-test

The null hypothesis is that the mean of the normalized residual set is zero. The residuals are assumed to be gained from a normal distribution of known variance (unit). The p-value indicates the smallest level of significance that would lead to rejection of the null hypothesis with the given data. Table 3 includes the z-test p-values for different ground motion models for residual distribution of Sa(T=1.0sec). According to this table, the null hypothesis can be rejected for the majority of the given models.

Table 3. Traditional tests p-values; T=1.0 sec

Model Name	z-test	lil-test
AB10	0.00	0.500
AC10	0.00	0.500
AS08	0.40	0.346
BA08	0.150	0.298
CB08	0.00	0.063
CY08	0.81	0.26
Getal09	0.00	0.500
KG04	0.00	0.500
Setal09	0.81	0.500
Zetal11	0.10	0.500

5.2. The Lilliefors Test

The Lilliefors test was used to test the null hypothesis that data come from a normally distributed population, when the null hypothesis does not specify the mean and variance of the distribution. Table 3 includes the Lilliefors test p-values for different ground motion models for residual distribution of Sa (T=1.0sec). According to this table, there is not enough evidence to reject the null hypothesis for any of models. Therefore, as also mentioned earlier, the traditional tests are not perfect tools for evaluation and ranking of the models (Scherbaum et al 2004).

5.3. The LH Test

The distribution of LH values for S_a(T=1.0sec) are shown in Fig. 4.The statistical measurements of LH values are shown in Table 4, as well as some other measurements of the residuals for acceleration response spectra in period T=1.0 sec, S_a(T=1.0 sec). The goodness-of-fit-measures in this method are: the median LH values (MEDLH) and the median, mean, and the standard deviation of the normalized residuals (MEDNR, MEANNR, and STDNR, respectively). The corresponding standard deviations of these measures (σ) are calculated using the bootstrap technique through data re-sampling (Efron and Tibshirani 1993). By using these measures and based on the scheme presented in the former sections, the ground motion models are ranked in the categories A, B, C, or D in the last column. The earlier table may be repeated for the different periods. The relative similarity of ranking results for different periods may be interpreted as a sign of method stability. This hypothesis is studied. Since the ranking results are more or less stable for the different periods, it has been decided to merge all residuals into a unit set and then repeat the ranking procedure.

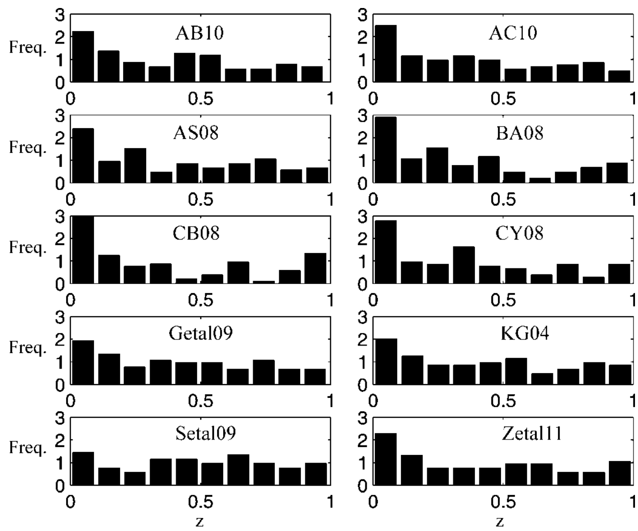

Figure 4. Distribution of LH values for Sa(T=1.0sec) with respect to different ground-motion models

Table 4. Ranking of models based on LH method with respect to S_a(T=1.0sec)

T=1.0 sec

Model Name	Rank	MEDLH	σ	MEDNR	σ	MEANNR	σ	STDNR	σ
Setal09	A	0.48	0.04	0.00	0.13	0.01	0.10	1.07	0.08
Getal09	B	0.36	0.04	0.38	0.16	0.32	0.11	1.23	0.09
KG04	C	0.40	0.05	-0.64	0.15	-0.56	0.10	1.10	0.08
AB10	C	0.37	0.06	0.56	0.16	0.46	0.12	1.27	0.09
Zetal11	C	0.34	0.05	0.21	0.14	0.15	0.12	1.31	0.10
CY08	C	0.32	0.04	0.02	0.25	0.04	0.13	1.41	0.08
AS08	C	0.30	0.05	0.04	0.20	-0.15	0.13	1.35	0.09
AC10	D	0.32	0.04	0.96	0.12	0.91	0.10	1.09	0.07
BA08	D	0.24	0.03	0.17	0.25	0.21	0.14	1.50	0.09
CB08	D	0.22	0.05	-0.88	0.27	-0.84	0.14	1.56	0.10

Table 5 shows the ranking of models based on this united residuals set. Table 5 can be accounted as the final ranking of models based on LH method. According to this ranking, two models CB08, and AC10 should be excluded from the acceptable models. An interesting result of this table is that all models developed specially for Iran region (Category 1) are ranked B, and the NGA models (Category 3) are ranked C, and D. On the other hand, models that were categorized as Europe and Middle East models (Category 2) show a wide range of performance, from B to D. We next exclude CB08 and AC10 from the usable models, and study the ranking of the remaining models using the information-theory method. The agreement of the two thus-obtained rankings provides an estimate of the reliability of the results.

Table 5. Final ranking of models based on LH method for united residuals, all events

All periods, all events

Model Name	Rank	MEDLH	σ	MEDNR	σ	MEANNR	σ	STDNR	σ
Setal09	B	0.44	0.02	-0.01	0.05	0.06	0.04	1.17	0.03
Zetal11	B	0.42	0.02	0.03	0.04	0.04	0.04	1.17	0.03
Getal09	B	0.40	0.02	0.16	0.05	0.13	0.04	1.21	0.03
AB10	C	0.37	0.02	0.39	0.05	0.40	0.04	1.26	0.03
KG04	C	0.34	0.02	-0.52	0.07	-0.42	0.04	1.26	0.03
AS08	C	0.33	0.02	-0.22	0.07	-0.24	0.04	1.34	0.03
CY08	C	0.31	0.02	0.03	0.06	0.05	0.05	1.37	0.03
BA08	C	0.28	0.02	0.05	0.07	0.17	0.05	1.48	0.03
AC10	D	0.32	0.02	0.96	0.04	0.98	0.04	1.06	0.02
CB08	D	0.21	0.02	-0.86	0.08	-0.82	0.05	1.55	0.03

5.4. Ranking of Ground Motion Models by Using The Information-theory Method

The average sample log likelihood (LLH) has been calculated for each of the considered periods. Fig. 5 compares the LLH value for the candidate ground motion models in different periods. This Fig. has been prepared for all records, records with M_w<6.25, and records with M_w>6.25, separately. As shown in Fig. 5, it seems that some of ground motion models are more compatible with the observational data for nearly all periods. We emphasize here, that the two models, Getal09 and Setal09, are developed just for the response spectra values, excluding the PGA value. Therefore, these two models are excluded from comparisons. A final period- independent ranking can be created by averaging on LLH values of all periods, as shown in Table 6. By comparing the Table 5 with Table 6, the agreement of LH and the information-theory method in ranking of the models is confirmed. The two models, Getal09, and Zetal11, which are located in top levels of the ranking belong to category 1. In contrast, NGA models are consistently in the lower half of table. The rational judgment gained from this result is that using the application of NGA models for Iran, which is a common practice in many hazard analysis projects, may be questionable (particularly, note that CB08 was fully rejected according to LH method).

Now, the main objective is to find corresponding weights to be used in seismic hazard analysis. The LLH values can be transformed into compatible weights. However, in order to combine the two methods LH and information-theory method, this procedure is undertaken in two steps:

(1) We assign a general weight to three top models as well as for the five bottom models. The criterion for this weighting is the arithmetic average value of LLH values. The arithmetic average value of LLH index for the three top models is equal to 2.05 and for the other five models is 2.21. The overall resulted weights are 0.53, and 0.47, respectively.

(2) The resulting weights are now shared between the corresponding models.

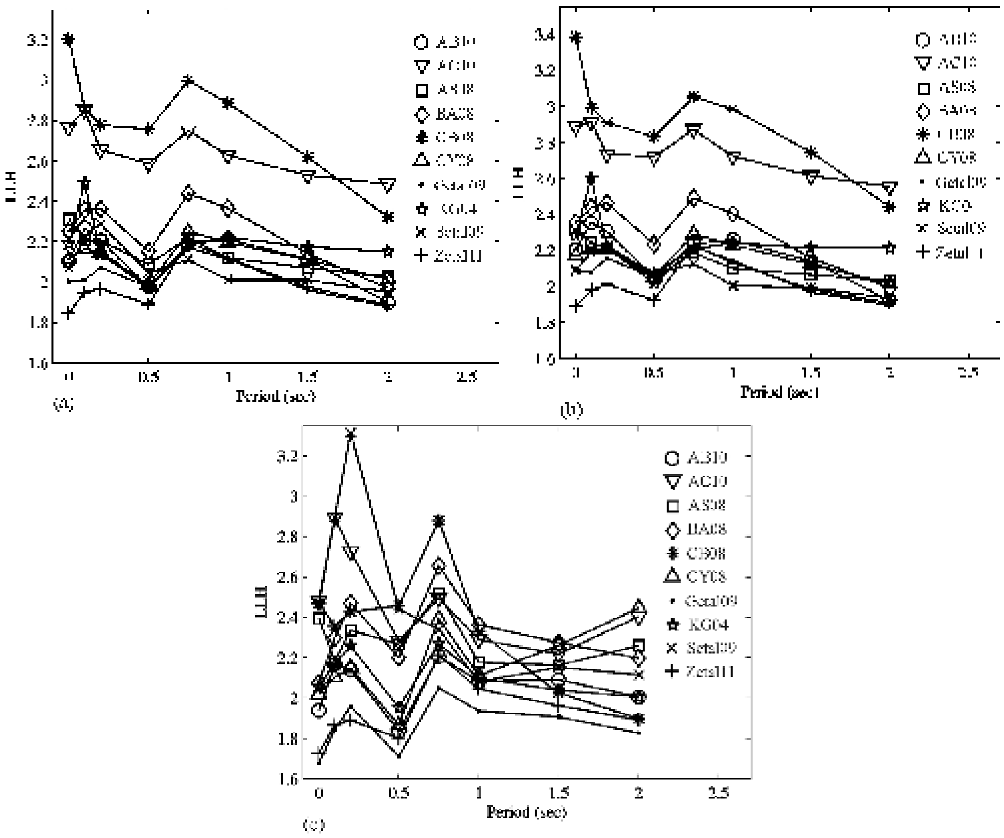

Figure 5. Comparison of LLH values for different ground motion models
(a) all records, (b) record with Mw<6.25, and (c) records with Mw>6.25

Table 6. Final ranking of models based on information-theoretic method for all periods, all events

All periods, all Events		
Rank	LLH	Model
1	1.98	Zetal11
2	2.03	Getal09
3	2.09	Setal09
4	2.12	AB10
5	2.12	CY08
6	2.16	AS08
7	2.19	KG04
8	2.25	BA08
9	2.66	AC10
10	2.80	CB08

The final result of weighting calculations is shown in Fig. 6. It is worth emphasizing that the proposed logic tree provides just an offer and never takes the place of the expert judgment. In other words, the quantitative values of each branch obtained through the above procedure could be considered as a numerical guide for subjective weighting by experts. Using this approach, the uncertainty arising from differences in expert opinions can be decreased. It should be noted that since the major source of uncertainty in the probabilistic seismic hazard analysis originates from ground motion models, any quantitative framework of obtaining weights of logic tree is of great value.

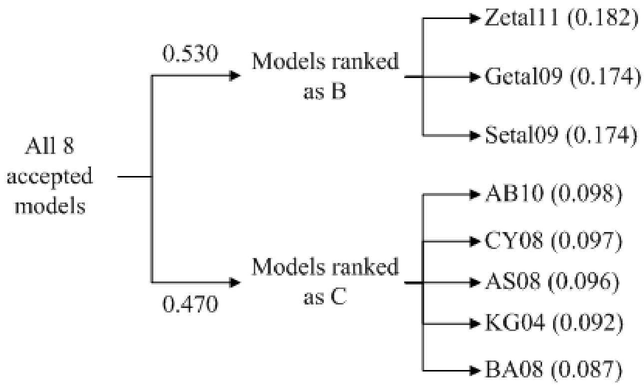

Figure 6. Final weighting results based on combination of LH and information-theoretic methods

6. CONCLUSIONS

Two different approaches have been used here to evaluate candidate ground motion models for the Zagros region of Iran. First, by using a set of recorded ground motion data, the computed residuals with respect to different ground motion models were analyzed by using the LH method. Based on this method, two models were unacceptable and the remaining models were ranked as B or C. Second, information theory was employed to rank the models, again. The good agreement of these two methods confirms the reliability of the final ranking. One of most significant results of this study was that the regional ground motion models show more consistency with observed data than do models developed using NGA models. Finally, from a combination of the two methods, coherent weights can be calculated that provide a quantitative alternative to expert opinions in seismic hazard projects. These weights can be used to complement expert opinions, where these may be available, or replace expert opinions when these are unavailable. Due to a paucity of data, the testing of the method developed here does not include data from earthquakes with Mw > 6.5 and R < 50 km.

REFERENCES

Abrahamson, N. and Silva, W. (2008). Summary of the Abrahamson & Silva NGA Ground Motion Relations. *Earthquake Spectra* **24:1**, 67–97.

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 Letters* **81:2**, 195-206.

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.

Berberian, M. (1976).Contribution to the seismotectonics of Iran (Part 2), Geological Survey of Iran, Report 39.

Bommer, J.J. and Scherbaum, F. (2008). The use and misuse of logic-trees in probabilistic seismic hazard analysis, *Earthquake Spectra* **24:4**, 997–1009.

Bommer, J.J., Douglas, J., Scherbaum, F., Cotton, F., Bungum, H., and Fäh, D. (2010). On the Selection of Ground-Motion Prediction Equations for Seismic Hazard Analysis, Seismological Research **81:5**, 783 – 793.

Boore, D. and Atkinson, G. (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.

Budnitz, R.J., Apostolakis, G., Boore, D.M., Cluff, L.S., Coppersmith, K., Cornell, C.A. and Morris, B.J. (1997). Recommendations for probabilistic seismic hazard analysis, in Guidance on Uncertainty and Use of Experts, Lawrence Livermore National Laboratory, **Vol I**.

Campbell, K. 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.

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.

Efron, B. and Tibshirani, R.J. (1993). An Introduction to the Bootstrap, Chapman & Hall/CRC, New York.

Ghasemi, H., Zare, M., Fukushima Y., Koketsu, K. (2009). An empirical spectral ground-motion model for Iran, J. Seism. doi 10.1007/s10950-008-9143-x.

Kalkan, E., and Gülkan, P. (2004). Site-dependent spectra derived from ground motion records in Turkey. *Earthquake Spectra* **20:4**, 1111–1138.

Krinitzsky, E. L. (1995). Problems with logic trees in earthquake hazard evaluation, *Eng. Geol* **39:1,** 1–3.

Montgomery, C.D., and Runger C.G. (2003). Applied statistics and probability for engineers, John Wiley & Sons, Inc.

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, *Bull. Seismol. Soc. Am*. **94:6**, 341–348.

Scherbaum, F., Delavaud, E. and Riggelsen, C. (2009). Informationtheoretic Selection of Ground-motion Prediction Equations for Seismic Hazard Analysis: An Applicability Study using California Data. *Bulletin of the Seismological Society of America, BSSA* **99:6**, 3234–3247.

Sharma, M. (2009). Ground-motion prediction equations based on data from the Himalayan and Zagros regions. *Journal of Earthquake Engineering* **13:8,** 1191-1210.

Toro, G. (2006). The effects of ground-motion uncertainty on seismic hazard results: Examples and approximate results, in *Annual Meeting of the Seismo Seismological Society of America*, San Francisco.

Zafarani et al. (2011) . Simulation of ground motion in the Zagros region, Iran using the specific barrier model and stochastic method. submitted to *Bulletin of the Seismological Society of America*. Under review