Iranian Journal of Science & Technology, Transaction A, Vol. 29, No. A1
Printed in The Islamic Republic of Iran, 2005

ARMA MODELLING OF ARTIFICIAL ACCELEROGRAMS FOR WESTERN IRAN*

S. M. FATEMI AGHDA[1]**, A. J. GANDOMI[1] AND A. BEITOLLAHI[2]

[1]Department of Geology, Tarbiat Moallem Univ., Tehran, I. R. of Iran
[2]Department of Eng. Seismology, BHRC, Tehran, I. R. of Iran

Abstract – The Artificial Accelerograms have been developed for assessing the dynamic response of structures. Considering seismological properties of the site are necessary for the best simulation of accelerograms. The real recorded accelergrams for simulating earthquake phenomenon are used in the Arma model. This is due to the fact that the Arma model can be considered more advantageous than the others.

In this paper, 25 recorded accelerograms are used for simulating accelerograms of a dam site in Hamedan province in western Iran. The parameters which are used in the stationary transformation are related to the physical parameters (eg. magnitude, epicentral distance and duration) via a regression analysis. To generate Artificial Accelerogram, the physical parameters enter the Arma model, therefore the output would be a stationary time history adjusted with recorded accelerograms. The generated stationary time history is changed into a non-stationary time history from the amplitude and frequency point of view.

Keywords – Artificial accelerogram, ARMA model, simulation of earthquake

1. INTRODUCTION

Western Iran is one of the most hazardous areas from a seismological and geotechnical point of view [1]. The investigated area is covered with thick alluvium sediments which show a similarity of seismic wave propagation behavior. There has recently been a considerable increase in the precise designing of structures in the urban and industrial development sectors, especially against the above mentioned problems such as land subsidence [1]. Also, the recent destructive earthquake in the northern part of Hamedan province (2001 Changureh earthquake, M=6.3) has turned attention to this region.

The ultimate goal of this research is to obtain simulation of earthquake ground motions of Hamedan Province, western Iran, that are statistically similar to the recorded motions of the area. For this purpose, it is useful to be able to describe the content of a recorded earthquake trace with a few parameters, which can be related to the recorded strong motion characteristics [2, 3, 4]. There are some methods that consider seismological characteristics of a region and use small earthquakes to model source faults such as empirical derived function and simulation based on attenuation characteristics via seismological modelling [5, 6]. These are called deterministic methods. A

*Received by the editor November 8, 2003 and in final revised form November 21, 2004
**Corresponding author

stochastic method which simulates strong ground motion based on recorded accelerograms is named ARMA modelling [3, 6-9].

Findell et al. [2]. used a method for describing a time history by transforming the record into a related stationary series using transformation function and fitting a single time-invariant ARMA model to this new series. They used this method to simulate accelerograms in the Istanbul region [2]. Also Chang et al. [3], in order to analyse and simulate El-Centro and San Fernando earthquakes, used the ARMA method [3].

With regard to the availability of uniformly digitized earthquake acceleration data for analysis, and the wide spread interest in generating artificial digitized accelerograms for structural response studies,it is worthwhile to consider the use of models that can be formulated explicitly in discrete time.

One of the most important discrete models is the Auto Regressive Moving Average (ARMA) model, which can be represented as a stochastically linear difference equation of finite order.

In this study, 25 recorded accelerograms are used to simulate the earthquake in the target region. The accelerogram data used in this study were obtained from the data bank of the Building and Housing Research Center (BHRC) of Iran [10].

The characteristics of the accelerograms used are shown in Table 1. Also the geographical distribution of selected accelerograms with a 300 Km distance around a specific target site are given in Fig. 1.

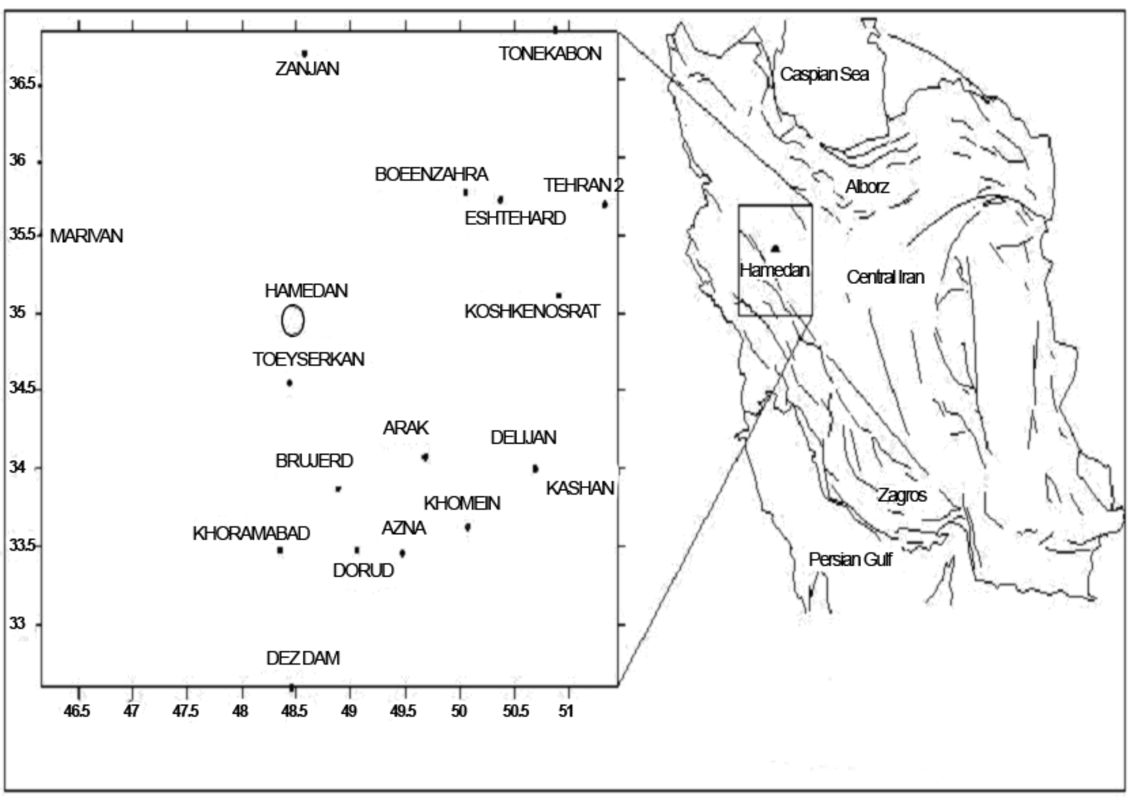

Fig. 1. Geographical distribution of used accelerograms

Table 1. Characteristics of used accelerograms

	Station Name	Lat	Lon	ED	Mb	Ms
1	BOOEINZAHRA	35.77	50.05	30	4.7	
2	MARIVAN	35.51	46.17	3	4.2	
3	KHORAMABAD	33.48	48.35	9	4.6	
4	KOSHKENOSRAT	35.11	50.9	71	5.5	5.8
5	KOSHKENOSRAT	35.11	50.9	72	5.4	5.2
6	KASHAN	33.98	51.44	91	5.5	5.8
7	DELIJAN	34	50.69	56	5.5	5.8
8	DELIJAN	34	50.69	55	5.4	5.2
9	ARAK	34.08	49.68	102	5.5	5.8
10	KHOMEIN	33.63	50.07	111	5.5	5.8
11	KHOMEIN	33.63	50.07	110	5.4	5.2
12	BRUKERD	33.88	48.88	9	4.6	3.8
13	DORUD	33.48	49.06	23	4.8	3.6
14	DEZDAM	32.6	48.46	142	4.5	
15	AZNA	33.46	49.47	15	4.5	
16	AZNA	33.46	49.47	8	4.4	
17	MARIVAN	35.51	46.17	35	4	
18	MARIVAN	35.51	46.17	34	4.6	3.5
19	DORUD	33.48	49.06	16	4.7	4
20	TOEYSERKAN	34.55	48.44	61	5	4.6
21	KHORAMABAD	33.48	48.35	61	5	4.6
22	TEHRAN2	35.69	51.33	209	6.4	7.7
23	TONEKABON	36.81	50.87	131	6.4	7.7
24	ZANJAN	36.66	48.57	75	6.4	7.7
25	ESHTEHARD	35.72	50.37	144	6.4	7.7

Duration of selected data in this study is between 6 to 45s and PGA (Peak Ground Acceleration) between 6 to 130 cm/ss.The investigated area is covered with thick alluvium in most parts, hence the accelerograms used in this study are selected from alluvial sites. As the site, geotechnical properties affect the seismic response of structures, strongly similarly selected accelerograms based on the characteristics of the geological site are very helpful for a better simulation.

2. ARMA MODEL AND ITS PARAMETERS

The ARMA model considered in this research is based on modelling the data sequence *x (n)* as the output of a linear system [11] which is characterized by a rational system function as follows:

$$H(Z) = \frac{B(z)}{A(z)} = \frac{\sum_{k=0}^{q} b_k z^{-k}}{H\sum_{k=1}^{p} a_k\, z^{-k}} \quad (1)$$

and the corresponding difference equation is

$$x(n) = -\sum_{k=1}^{p} a_k\, x(n-k) + \sum_{k=0}^{q} b_k\, w(n-k) \quad (2)$$

where *w(n)* is the input sequence to the system and the observed data, and *x(n)* represents the output sequence. In the ARMA modelling, if the observed data are characterized as a stationary random process, the input sequence will also be assumed to be a stationary random process.

The random process *x(n)* generated by the pole zero model in (l) or (2) is called an Auto Regressive Moving Average (ARMA) process of order (p,q) and is usually denoted as ARMA (p,q) if q=0 and b0 =1. The resulting system model has a system function H(z)=1/A(z) and its output is called an Auto Regressive (AR) process of order p. This is denoted as AR(p). The third possible model is obtained by setting A(Z) =1, so that H(z)=B(z). Its output *x(n)* is called a Moving Average (MA) process of order q and denoted as MA(q).

For example, the ARMA (2, 1) model for a stationary correlated process $\{g_k\}$ is defined by the second- order auto regressive/ first-order moving average difference equation as follows:

$$g_k - a_1\ g_{k-1} - a_2\, g_{k-2} = e_k - b_1\, e_{k-1} \tag{3}$$

in which it is assumed that *ek* is independently and identically distributed. That is to say, the input is stationary discrete white noise [3].

From the three mentioned linear models, the AR model is widely used, especially in signal processing studies [8, 11]. The reasons are twofold. The first is that the AR model is suitable for representing spectra with narrow peaks (resonance) [8, 12]. Second, the AR model can be shown in very simple linear equations for the AR parameters. On the other hand, the MA model, as a general rule, requires many more coefficients to represent a narrow spectrum. Consequently, it is rarely used alone by combining the above two models (AR and MA). From the point of view the number of mode parameters, the ARMA model provides a more efficient representation to simulate the earthquake phenomenon.

The ARMA model is particularly appropriate when our data have been corrupted by noise [8].

In evaluating ARMA mode parameters, the auto correlation of seismic trace is very important. The relationship between auto correlation and AR parameters can be considered as follows (4):

$$\begin{bmatrix} \gamma_{Xx}(q) & \gamma_{(xx)}(q-1) \ \dots & \gamma_{xx}(q-p+1) \\ \gamma_{zx}(q+1) & \gamma_{xx}(q) & \gamma_{xx}(q+p+2) \\ . & . & . \\ . & . & . \\ . & . & . \\ \gamma_{xx}(q+p+1) & \gamma_{xx}(q+p-2)\dots & \gamma_{xx}(q) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ . \\ . \\ . \\ a_p \end{bmatrix} = - \begin{bmatrix} \gamma_{xx}(q+1) \\ \gamma_{xx}(q+2) \\ . \\ . \\ . \\ \gamma_{xx}(q+p) \end{bmatrix} \tag{4}$$

In terms of the ARMA model, this indicates that the AR part of the model is of primary importance. Once this part is identified the MA part can be constructed [13].

The problem of order selection for the ARMA(p,q) model has been investigated by Bruzzone and Kaveh [14]. For this purpose, the minimum of the *AIK* index:

$$AIK(p,q) = In\,\hat{\sigma}^2 + \frac{2(p+q)}{N} \tag{5}$$

can be used where $\hat{\sigma}^2$ is an estimation of variance of the input error.

The ARMA model that is selected in this research is an ARMA (4, 1) model which is best fitted to input accelerograms used in this study which can be represented as follows:

$$Z(t) = a_1 Z_{t-1} + a_2 Z_{t-2} + a_3 Z_{t-3} - a_4 Z_{t-4} - e_t - e_{t-1} b_1 \tag{6}$$

in which τ_t is the stationary output process, *a* and *b* are, respectively, AR and MA parameters and e_t is input white noise. The calculation of this modelling has been done in MATLAB software environment.

3. MODELLING PROCEDURE

The ARMA approach for analysing timeseries is only applicable to stationary time series [2, 8]. However, the acceleration of ground motion during an earthquake is not a stationary process [3, 15]. This means that before an ARMA model is applied to an earthquake trace, the trace must be altered to a stationary one.

Most obvious trends in any earthquake time series are included in amplitude content. So if the variations are removed, a stationary series, which can then be modeled by an ARMA process, will remain.

The first step in the modelling procedure is to isolate the period of significant shaking. This generally entails removing firstly 1%, and finally 2% [2, 4]. This is due to the presence of noise in the two end sides of a trace, especially in the final end of the trace. This usually shortens the time series significantly, but isolates the significant portion of the earthquake shaking from the noise that is only slightly above the background level. The duration of this shortened time series is denoted.

With this shortened series, the next step is to remove the amplitude variation .The trends in this variation can be described by the standard deviation of the amplitude at every point within the series. Due to the impossibility of calculation of standard deviation for a single point (single time value), it is determined for a time window of 50 points around each point. For records with equal time intervals an envelope is given below

$$\sigma^2(i) = \frac{1}{nwind + 1} \sum_{j=i-nwind/2}^{i+nwind/2} a(j)^2 \tag{7}$$

nwind value is 50 point, and for traces with a 0.005 discrete time interval, it is equal to 0.25.

For most earthquakes, an exponentially decaying function of the form equation(8) can be fitted to the experimental standard deviation envelope.

$$\sigma(t) = c_1(a - k_1)\left(\frac{t}{\tau}\right)^p \exp\left[-\left(\frac{c_2}{T}\right)t\right] + k_1$$
$$c_1 = \left[\frac{c_2 \exp(1)}{P}\right]^p, c_2 = 2\sqrt{3}, P = \left(\frac{c_2}{\tau}\right)_{t_{\max}} \tag{8}$$

Equation (8) is proposed by Findell, et al. [2]. Each of the parameters in the above function has some physical meaning related to the real earthquake trace. α is a measure of strong shaking, which is calculated maximum value of standard deviation. k_1 is a measure of weak shaking and determined as an average of the first one-third of the trace. τmeasures the duration of strong shaking which is the duration elapsed between 2% and 95% of cumulative energy, t_{max} is the rise time to peak amplitude.

Figure 2 shows accelerograms and standard deviation envelope, which are calculated for some of the records used in this study. Calculated parameters for each accelerogram were used to transform the recorded accelerograms into stationary ones (Table 2).

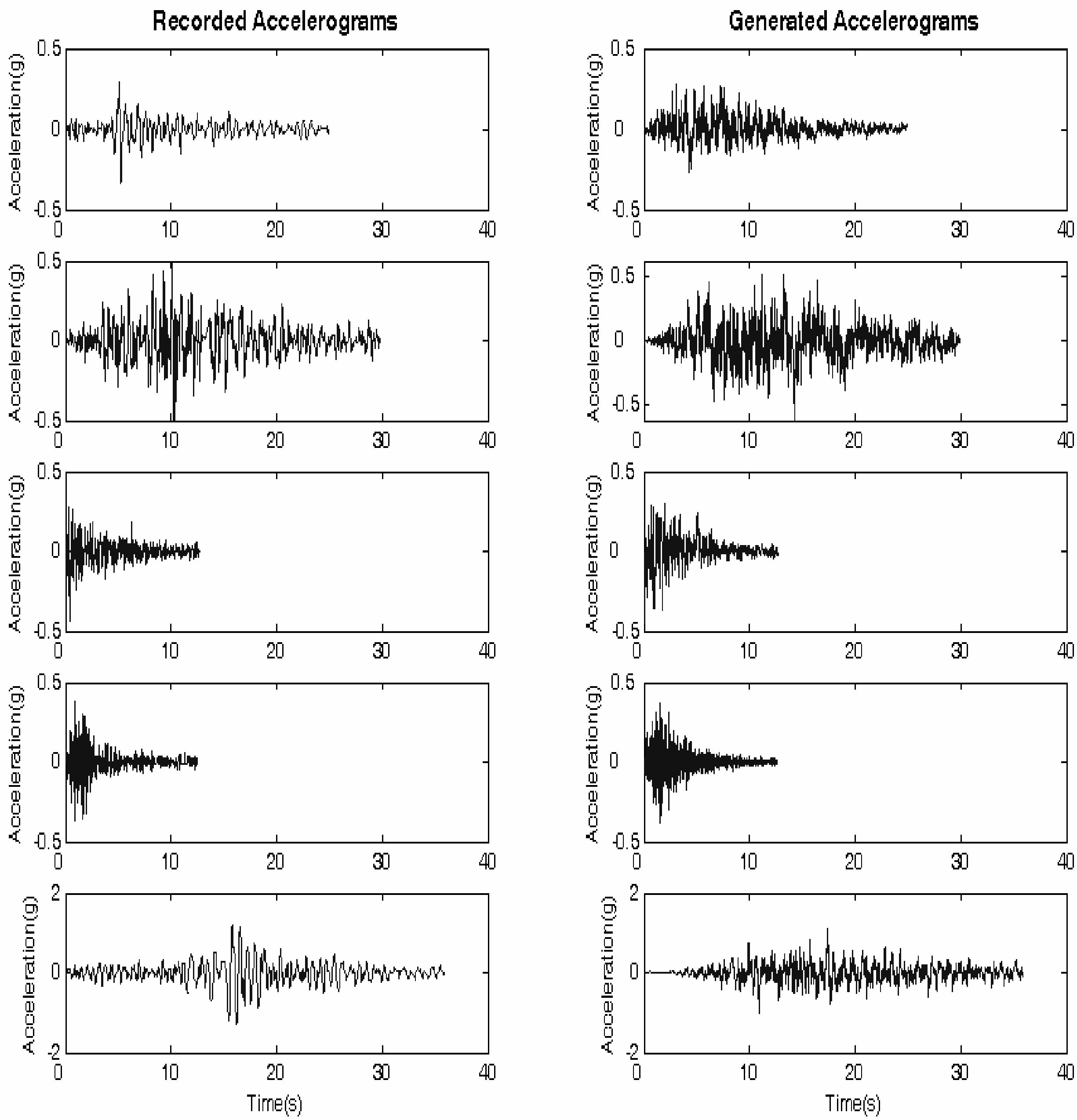

Fig. 2. Five recorded traces in western Iran and simulated ones

Table 2. Calculated nonstationary parameters for each recorded accelerogram

	Station Name	c2	b2	tav	tt	k1	alpha
1	BOOEINZAHRA	4	0.11	10.6	12	4.8	0.18
2	MARIVAN	3.5	0.19	6	8.6	0.16	0.12
3	KHORAMABAD	4	0.21	3.9	6.5	0.14	0.24
4	KOSHKENOSRAT	4	0.16	15	15.6	5	0.13
5	KOSHKENOSRAT	2.5	0.15	8.3	9.4	1.1	0.15
6	KASHAN	2	0.06	15.7	18.6	5.2	0.25
7	DELIJAN	4	0.10	18.4	23	10.3	0.36
8	DELIJAN	4	0.16	12.1	15.8	1.8	0.21
9	ARAK	6	0.30	8.7	11.4	0.5	0.2
10	KHOMEIN	4.5	0.13	15.3	16.3	13	0.18
11	KHOMEIN	3	0.23	8	9	1.7	0.13
12	BRUKERD	4.5	0.31	1.8	3.5	3.7	0.49
13	DORUD	4.2	0.18	4	8.2	0.5	0.18
14	DEZDAM	4.5	0.38	3.2	3.6	0.23	0.15
15	AZNA	6	0.43	5.9	6.9	3.4	0.1
16	AZNA	3.5	0.25	7	7.5	1.9	0.06
17	MARIVAN	8	0.64	4.5	5.7	0.12	0.04
18	MARIVAN	7	0.55	8.2	8.5	0.3	0.04
19	DORUD	4	0.18	7	7.9	1.34	0.06
20	TOEYSERKAN	4	0.12	6.1	8.7	1	0.23
21	KHORAMABAD	4.5	0.20	4	9.5	0.2	0.1
22	TEHRAN2	4	0.24	7.7	7.9	3.5	0.04
23	TONEKABON	2	0.05	16.7	20	16.2	0.22
24	ZANJAN	6	0.10	25.6	26.2	14.6	0.78
25	ESHTEHARD	8	0.15	20	22.7	19.7	0.4

The time series is stabilized by dividing it through the fitted standard deviation envelop at discrete time steps.

$$\hat{a}(t) = \frac{a(t)}{\hat{\sigma}(t)} \tag{9}$$

In order to stabilize the records due to the frequency content, the following step is performed. One means of characterizing the frequency content of a time series with zero-mean is to measure the rate of zero-axis crossing [2]. Again, a time window of 50 points around each point is used to determine the rate of zero-crossing $F_c(i)$ within that window

$$F_c(1) = \frac{Cors\sin g\ between\, i \pm nwind/2}{(nwind)(\Delta t)}\,. \tag{10}$$

Then the variance-stabilized series is altered by essentially stretching the series where the crossing rate is high, and compressing it where the crossing rate is low, so that the frequency is made approximately constant throughout the duration of the series.

Now the time series is stationary and ready for an auto regressive moving average (ARMA) characterization.

4. GENERATION OF ARTIFICIAL ACCELEROGRAMS

In order to develop earthquake simulations that are dependent on both the seismological characteristics of a particular region (Lam et al.) [6] and the physical variables describing an event, a regression analysis is performed between the variables, specifically magnitude, epicentral distance, strong shaking parameter α and its duration τ and rise time to peak amplitude $t_{\max}$.

A combined regression between α and the two physical parameters of earthquake M (magnitude) and R (epicentral distance) is as follows:

$$\log(\alpha) = 0.2 + 0.12 \log\left(10^m / R^2\right) \qquad (11)$$

This is in the form that best fits the data that have an epicentral distance of more than 30 km. Note that this equation is derived for the target region (show in Fig.1) and in each region for generating artificial accelerograms, this relationship and other ones must be derived.

Regression analysis of horizontal records led to the parametric equations that are shown in Table 3. To address the accuracy of this approach of earthquake modelling, real traces were compared in both the frequency and time domain with simulated traces that are generated according to magnitude, duration and epicentral distance of the given real event. Figures 2 and 3 compare the time series and response spectrum for damping ratio=0.05 of real events with simulated ones.

Table 3. Parametric relations obtained from strong motion data

$\log(\alpha) = 0.2 + 0.12\log\left(10^m / R^2\right)$ $\tau = 0.9363.tt - 1.2255$ $\alpha = 0.0156.tt + 0.0189$ $t_{\max} = 0.7894.\tau - 3.2791$	$\tau_{\min} = 4s$ $\tau_{\max} = 26.2s$ $\alpha_{\min} = 0.1g$ $\alpha_{\max} = 0.78g$

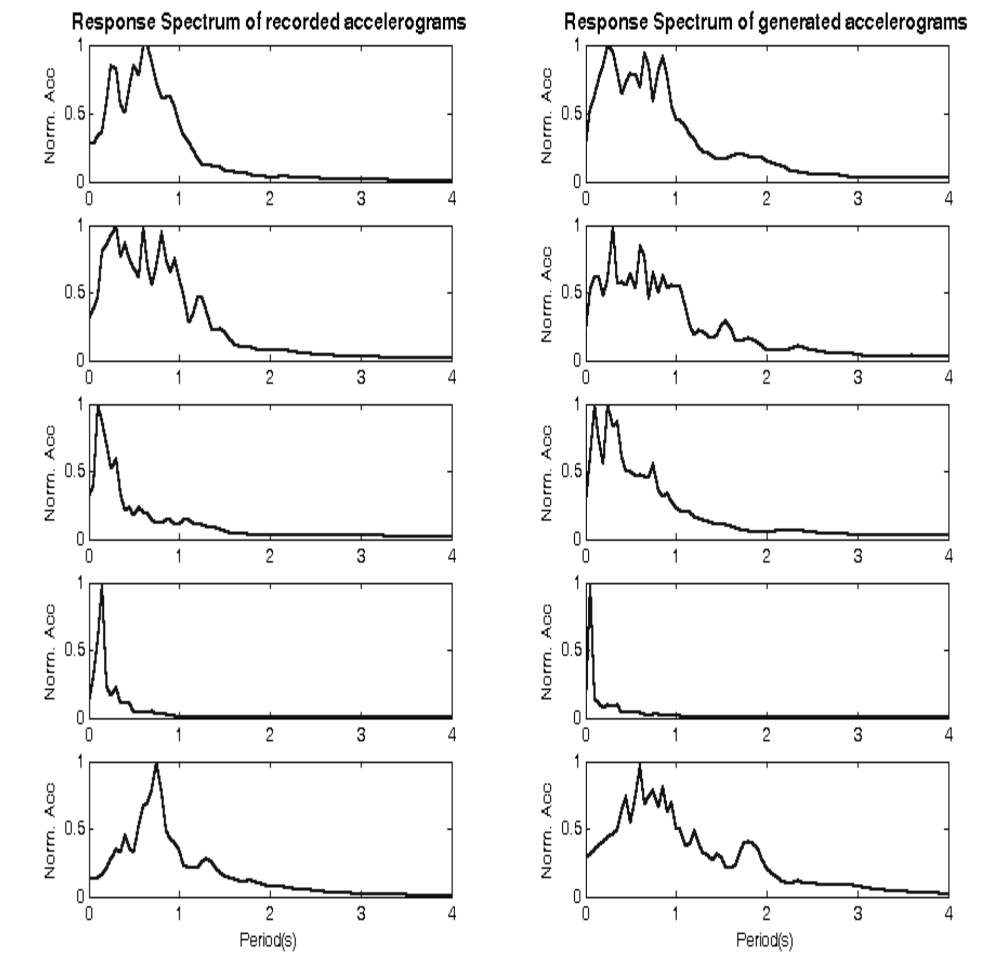

Fig. 3. Response spectrum of the time series in Fig. 3 for damping ratio=0.05

5. CONCLUSIONS

Due to the lack of real accelerograms and also the need for a detailed and precise dynamic design of structures, artificial accelerogams are generated. This paper focuses on one of the stochastic methods to generate artificial accelerograms named ARMA modelling (Auto Regressive Moving Average).

Since the ARMA model can only be used on the stationary data, accelerograms used must be stationary. For this reason a standard deviation envelope is used to change the accelerograms into a stationary one. Some regressive relationships are assessed between accelerogram parameters such as total duration, significant duration (the duration between 5-95% of arias intensity) and strong shaking parameter (depends on PGA). Regressive analysis shows that the significant duration parameter increases with distance.

A program that is developed in MATLAB language evaluates ARMA parameters. Note that the output of the ARMA filter is a stationary time series, so again it must be changed to be non-stationary.

To test the capability of the model, some of the recorded accelerograms are simulated and their response spectrum for 5 percent damping is calculated. This shows good agreement between recorded and simulated ones.

REFERENCES

1. Aliari, A. R. (1993). Investigation cause of ground failures in the vicinity of Hamedan (Shahid Mofatteh) power plant. MS.c Thesis, Tarbiat Moallem University, Tehran, Iran, 2002.
2. Findell, K. L., Koyluglu, H. U. & Cakamak, A. S. (1993). Modelling and simulating earthquake accelerograms using strong motion data from the Istanbul, Turkey region. *Soil Dynamic and Earthquake Engineering*, *12*, 51-59.
3. Chang, M. K., Kwiatokowski, J. W., Nau, R. F., Oliver, R. M. & pister, K. K. (1982). ARMA model for earthquake ground motions. *Earthquake Engineering and Structural Dynamics*, *10*, 651-662.
4. Gandomi, A. J. (2003). Generation of artificial accelerograms using artificial neural network. MS.c Thesis, Tarbiat Moallem University, Tehran, Iran.
5. Hadon, R. A. W. (1996). Use of empirical green's functions, spectral ratios and kinematic source models for simulating strong ground motion. *BSSA*, *86*, 597-615.
6. Lam, N., Wilson, J. & Hutchiston, G. (2000). Generation of synthetic earthquake accelerograms using seismological modelling: A Review. *Journal of earthquake engineering*, *4*(3), 321-354.
7. Jurkevics, A. & Ulrych, T. J. (1979). Auto regressive parameters for a suite of strong motion accelerograms. *BSSA*, *69*(6).
8. Proakis, J. & Manolakis, D. G. (1989). Introduction to digital signal processing. *Maxwell & Macmillan*, 944.
9. Suzuki, S. Hada, K. & Asano, K. (1998). Simulation of strong ground motion based on recorded accelerograms and the stochastic method. *Soil Dynamic and Earthquake Engineering*, *17*, 551-556.
10. Building and Housing Research Center. (1993). A Collection of the accelerogram network data of Islamic Republic of Iran, Pub. No. 179.
11. Box, G. E. P. & Jenkins, G. M. (1976). *Time series analysis: Forecasting and Control*. San Francisco, Holden-Day.
12. Oliver, R. M. & Dister, K. S. (1982). ARMA model for earthquake ground motions. *Earthquake Engineering and Structural Dynamics*, *10*, 651-662.
13. Dargahi-Nobari, G. R. (1992). Models for seismic records, why uniformly modulated ARMA. *Soil Dynamic and Earthquake Engineering*, *11*, 381-385.
14. Bruzzone, S. P. & Kaveh, M. (1980). On some suboptimum ARMA spectral estimator. *IEEE, Trans. Acoustics, speech and signal processing*, *28*(6), 753-755.
15. Sabetta, F. & Dugliese, A. (1996). Estimation of Response Spectra and simulation of nonstationary earthquake ground motion. *BSSA*, *86*(2), 337-352.