

RESEARCH ARTICLES

The effects of zoning methods on forecasting of the next earthquake occurrences, according to semi-Markov models

Ramin Sadeghian^{1,*}

¹ Bu Ali Sina University, Hamedan, Iran

Article history

Received February 2, 2010; accepted October 4, 2010.

Subject classification:

Semi-Markov model, Transition probability matrix, Probabilistic forecasting, Deterministic forecasting, Forecasting error, K94 zoning.

ABSTRACT

Earthquakes are natural phenomena that can be viewed in three dimensions: time, space and magnitude. Earthquakes can be investigated not only physically, but also mathematically. In this study, semi-Markov models are applied, which can be considered as useful methods to analyze and forecast the occurrence of future earthquakes based on previous earthquake data. In the present study, the target region, Iran, is divided into zones, and each zone is examined as one of the semi-Markov model states. Several methods to determine the levels of forecasting error are then introduced and applied to the target area. The results of the application of these semi-Markov models to investigate and forecast the occurrence of future earthquakes are obtained and analyzed mathematically. A new zoning method is developed and compared with that of Karakaisis, through the proposed forecasting method. Moreover, the effects of the type of zoning and the number of zones on the forecasting error of the next earthquake occurrences are investigated using several algorithms.

Editor's note. The present paper revisits and completes the paper by Sadeghian and Jalali-Naini [2008b] cited by the author in the References.

Introduction

Earthquakes can be examined both mathematically and physically [Sadeghian 2007]. Stochastic processes represent one of the branches of mathematics that can be applied to probabilistic investigations of these phenomena, and one of the stochastic process models that has been frequently applied over recent years is the Markov model, and especially the semi-Markov model. Semi-Markov models belong to those that can model events that have a relationship with previous events [Jalali-Naini 1997, Minh 2001, Rice 2007, Sadeghian and Jalali-Naini 2008a]. For analyzing of the three temporal, spatial, magnitudinal dimensions of earthquakes, Markov models, including semi-Markov models, can be applied [Patwardhan et al. 1980]. Markov models have already been applied for earthquake occurrence analysis [Di Luccio et al. 1997, Console 2001, Console et al. 2002]; moreover, in this regard, a few studies have recommended the use of semi-Markov models [Patwardhan et al. 1980,

Altinok and Kolcak 1999, Sadeghian and Jalali-Naini 2008b, Jalali-Naini and Sadeghian 2009].

One of the advantages of the model applied in this study is that it can forecast this triad of dimensions simultaneously, which has not been provided in previous studies [Kelleher et al. 1973, Nanjo et al. 2006, Papazachos 1992, Rikitake 1976]. For every semi-Markov model, and the inspection of each event, several states have to be defined. In this study, the Iran area was chosen as the target for study. Seismologically, this area was divided into various parts by Karakaisis [1994] (K94), and also in the present study. Each part is called a zone, and each zone can be considered as a state of a semi-Markov model. A similar classification has to be used for the magnitudes of the earthquake occurrences too.

Using the proposed semi-Markov models, later earthquakes can be forecast probabilistically. Then, to demonstrate whether the forecasting results obtained are mathematically valid or not, several methods are introduced to calculate the differences between the probabilities of the occurrences of actual earthquakes and those forecast, which from now on is called the forecasting error. Nava et al. [2005] presented another method for validation of their model. The forecasting error of the algorithms presented is obtained, and the results are investigated and analyzed. Moreover, a deterministic forecasting algorithm, based on the provided probabilistic one, is presented, and results related to deterministic forecasting are also mentioned.

Modeling

Applying semi-Markov models in the forecasting of the dimensions of earthquakes was completely examined by Altinok and Kolcak [1999] and Jalali-Naini [1997]. One of the most important elements of the semi-Markov models is the interval transition probability matrix. The probability of a transition from state i to state j in the interval $[0, n]$ requires the process to make at least one transition during that

interval. The interval transition probability matrix can be determined in matrix form as follows:

$$\begin{aligned} F(n) &= W^G(n) + \sum_{m=0}^n G \otimes T(m)F(n-m) = \\ &= W^G(n) + \sum_{m=0}^n C(m)F(n-m); \quad n = 0, 1, 2, \dots \end{aligned} \quad (1)$$

where n is the number of time intervals, $W^G(n)$ is a diagonal matrix where the i -th element is equal to $w_i^G(n)$, and is the probability that the waiting time in state i is greater than n , and G is a transition matrix, so that G_{ij} is the probability that if the last step is in state i , the next step will be in state j . In other words, G_{ij} is the probability of transition from state i to state j . Also $T(m)$ is the holding time matrix, which is defined as follows:

$$T_{ij}(m) = \Pr\{t_{ij} = m\}; \quad m = 1, \dots, n.$$

The interval transition probability $F(n)$ is obtained by a recursive procedure. Since $T(0)$ is equal to zero, $F(n)$ is just obtained for the interval $1 \leq m \leq n$. In the case $n = 0$, $F(n)$ is equal to the Kronecker Delta or identity matrix, which is defined as follows:

$$E_{ij}(0) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Also, $C(m)$ is the core matrix, which is defined as follows:

$$C_{ij}(m) = G_{ij}T_{ij}(m); \quad i, j = 1, \dots, N, m = 1, \dots, n.$$

Where N is the total number of states in the system [Jalali-Naini and Sadeghian 2009].

For the earthquake phenomenon, $F(n)$ can be used to study earthquake hazards and to evaluate their risk.

The two interval transition probability matrices that consist of $FR(k) \forall k = 1, \dots, n$ (for region-to-region transitions) and $FM(k) \forall k = 1, \dots, n$ (for magnitude-to-magnitude transitions) can be determined through Equation (1). If the last earthquake occurred in region r_0 with a magnitude m_0 , the matrix of probabilistic forecasting after $k; \forall k = 1, \dots, n$ time periods (i.e. $\hat{FRM}(k) \forall k = 1, \dots, n$) is obtained by the following formula:

$$\begin{aligned} \hat{FRM}_{r_i m_j}(k) &= FR_{r_0 r_i}(k) \times FM_{m_0 m_j}(k) \quad \forall k = 1, \dots, n; \\ & \quad j = 1, \dots, m; \quad k = 1, \dots, n \end{aligned} \quad (2)$$

where r is the number of zones of the assumed area considered, m is the number of classes for all magnitudes considered, and k specifies the number of time periods. Therefore, by computing $\hat{FRM}_{r_i m_j}(k) \forall i = 1, \dots, r; j = 1, \dots, m; k = 1, \dots, n$, all of the three dimensions of any earthquake can be taken into account and specified. As in any region, any earthquake

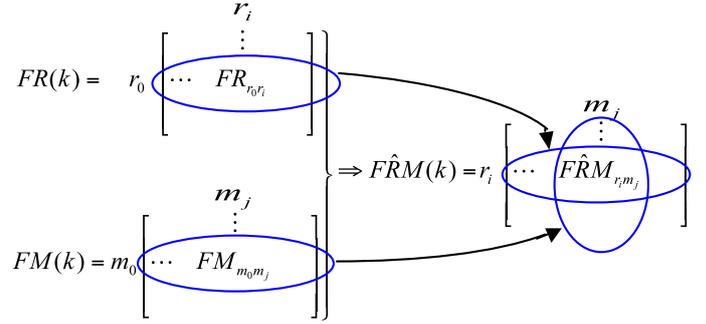


Figure 1. The ways that the \hat{FRM} matrix is determined [from Sadeghian and Jalali-Naini 2008b].

with any magnitude can occur, the region and magnitude can be considered as two independent random variables. Consequently, based on probability theory, if X and Y are two independent random variables, $f(x, y) = f(x) \cdot f(y)$, where f is a probability density function [Jalali-Naini 1997, Rice 2007]. The same relationship is also demonstrated by Equation (2).

Also $\hat{FRM}_{r_i m_j}(k) \forall i = 1, \dots, r; j = 1, \dots, m; k = 1, \dots, n$ is the probability that an earthquake will occur in the region r_i with the magnitude m_i after the time period k . Figure 1 illustrates this more clearly.

In this way, forecasting the dimensions of following earthquakes is possible, through determining the probability forecasting matrices (i.e. $\hat{FRM}(k) \forall k = 1, \dots, n$). In this study, the probabilistic forecasting matrix and deterministic forecasting matrix are demonstrated by the \hat{FRM} and FRM symbols respectively.

For model validation, assume that $\hat{FRM}(k)$ is a probability forecasting matrix, where its elements are the estimated probabilities of the next earthquake occurrences according to our proposed model, during the next k time periods, and $FRM(k)$ is a deterministic forecasting matrix defined as follows:

$$FRM_{ij}(k) \equiv FRM_{r_i m_j}(k) = \begin{cases} 1 & \text{If an earthquake occurs in region } r_i \\ & \text{with a magnitude } m_j \text{ in } k \text{ th time period} \\ 0 & \text{Otherwise} \end{cases}$$

There are three common ways for calculating the forecasting error: mean square error (MSE), mean absolute deviation (MAD) [Rice 2007], and mean absolute percent error (MAPE) [Haj Shirmohammadi 2003].

Let F_i be the forecasting value for period i , and D_i be the real data for period i , then the forecasting error (E_i) is calculated by $E_i = D_i - F_i$. These three attributes for comparing forecast precision can be used according to the following relationships:

$$MSE = \frac{\sum_i E_i^2}{N}$$

$$MAD = \frac{\sum_i |E_i|}{N}$$

$$MAPE = \frac{\sum_i \frac{|E_i|}{D_i}}{N} \times 100$$

where N is the number of historical data for the forecasting. Note that the MSE and MAD are not scale less, while the $MAPE$ is scale less. Therefore $MAPE$ is more suitable than MSE and MAD to compare several forecasting errors together.

Here, two fundamental questions are posed: The first question relates to the probable estimation of earthquake occurrences. In other words, can \hat{FRM} forecast the next earthquake occurrences successfully? The next question is to find out whether \hat{FRM} can be used for forecasting the next earthquake occurrences not only probabilistically, but also deterministically. The response to the first question can be found in methods 1, 2 and 3, and the response to the second question is explained in method 3, as:

Method 1: MSE and MAD [Rice 2007]

Method 2: $MAPE$ [Haj Shirmohammadi 2003]

Method 3: An innovative plan that can be used for determining probabilistic forecasting, deterministic forecasting and their forecasting errors (in this study, this is called the zero and one (0-1) method technically).

In methods 1 and 2, if the total number of the past data is equal to n , the first n_1 data will be used for forecasting the next n_2 data, which will be used to determine the forecasting error (i.e. $n = n_1 + n_2$). With all of the n data having already occurred and being available, the forecasting errors can be calculated. In method 3, the first n_1 data are used for forecasting the next n_2 data, which is applied for the making of a pattern, and the first $n_1 + n_2$ data are used for forecasting the next n_3 data, which is applied to the determination of the forecasting error (i.e. $n = n_1 + n_2 + n_3$). The whole n data have already occurred and are available; therefore, the forecasting errors can be calculated.

In these methods, calculation of the forecasting error can be performed in two forms: by using successive data or by using random data within the set of whole data. In the first form, the first n_1 data are used for forecasting the next n_2 data, (i.e. $n = n_1 + n_2$), while in the second form, the n_2 data within the main n data are eliminated randomly and then they will be forecast by their past data.

Method 1

In this method, the relationships between MSE and MAD are used. For this goal, an algorithm is presented for MSE , which can be used with some changes for both MAD and method 2.

Algorithm 1:

This algorithm can be used to determine the forecasting error by MSE as follows:

Step 0: Begin

Step 1: $n_2(0) = 0$

Step 2: $i = 0$

Step 3: Use the first $n_1 + n_2(i)$ data to determine \hat{FRM} , FRM (n_1 is the number of the first data that can be used for forecasting the next data, and $n_2(i)$ is the number of data that occurred during i time periods after the first n_1 data).

Step 4: Determine \hat{FRM} , FRM

$$\text{Step 5: } MSE(i+1) = \frac{\sum_{i=1}^r \sum_{j=1}^m (FRM_{ij}(1) - \hat{FRM}_{ij}(1))^2}{r \times m} \quad (3)$$

Step 6: $i = i + 1$

Step 7: If $n_1 + \sum_{j=1}^i n_2(j) < n$ then go to step 3; otherwise go to step 8 (n is the number of total data)

$$\text{Step 8: } MSE = \frac{\sum_{k=1}^i MSE(k)}{i}$$

Step 9: End.

For calculating the forecasting error by MAD , use algorithm (I); just be careful to substitute steps 5 and 8 in algorithm (I) by considering the following steps:

$$\text{Step 5: } MAD(i+1) = \frac{\sum_{i=1}^r \sum_{j=1}^m |FRM_{ij}(1) - \hat{FRM}_{ij}(1)|}{r \times m} \quad (4)$$

$$\text{Step 8: } MAD = \frac{\sum_{k=1}^i MAD(k)}{i}$$

For each forecast, if the past data are more complete, the forecast results are more accurate. Therefore applying $\hat{FRM}(k_2)$ after $\hat{FRM}(k_1)$ gives more accurate information than $\hat{FRM}(k_1 + k_2)$. Accordingly, only $\hat{FRM}(1)$ and $FRM(1)$ are used in Equations (3) and (4), and in the remaining equations.

One of the disadvantages of this method is that neither MSE nor MAD have any mathematical interpretation, nor is there any benchmark for their degree of goodness. In other words, if an amount is obtained for them, there is no common mathematical interpretation for every one. For example, by $MSE = 2$, regarding the forecasting error, it is not clear whether 2 is a suitable forecasting error or not.

Method 2

In this method, a similar equation to algorithm (I) is used, yet steps 5 and 8 in algorithm (I) are substituted by the following intended steps:

Step 5:

$$MAPE(i+1) = \frac{\sum_{i=1}^r \sum_{j=1}^m \frac{|FRM_{ij}(1) - \hat{FRM}_{ij}(1)|}{FRM_{ij}(1)}}{r.m} \times 100 \quad (5)$$

$$\text{Step 8: } MAPE = \frac{\sum_{k=1}^i MAPE(k)}{i}$$

where, as indicated above, $MAPE$ is the mean absolute percentage error. $MAPE$ is a useful equation, as it specifies the error percentage, which is both comprehensible and interpretable for every one. Therefore, this method will be concentrated upon in this study.

Method 3

Definition: the i -th maximum order

The i -th element in a descending sorted list, where none of the elements are equal to each other, is named the i -th maximum order.

This method is an innovative plan that can be used for two goals:

- a) Determining the forecasting error
- b) Deterministic forecasting of earthquake occurrences.

a) Determining the forecasting error

In this section, an algorithm to determine the forecasting error is presented. This algorithm has two parts. The first part of the algorithm is devoted to making a pattern, and the second part is assigned to determine the forecasting error. This algorithm consists of the following steps:

Algorithm II:

Step 0: Begin

Step 1: $n_2(0) = 0$

Step 2: $i = 0$

Step 3: Use first $n_1 + n_2(i)$ data for determining \hat{FRM} , FRM

Step 4: Determine \hat{FRM} , FRM

Step 5: $\hat{FRM1}(i+1) = \hat{FRM}(1)$, $\hat{FRM1}(i+1) = \hat{FRM}(1)$

Step 6: $i = i + 1$

Step 7: If $n_1 + \sum_{j=1}^i n_2(j) < n - n_3$ then go to step 3; otherwise $k_1 = i$ and go to step 8 (n_3 is the number of data used to determine the forecasting error)

Step 8: For $i = 1$ to k_1 do

$$\hat{FRM2}(i) = O \quad (O \text{ is a zero matrix})$$

$$MAPE2(i) = \frac{\sum_{i=1}^r \sum_{j=1}^m \frac{|FRM1_{ij}(i) - \hat{FRM2}_{ij}(i)|}{FRM1_{ij}(i)}}{r.m} \times 100$$

$$\text{Step 9: } MAPE(0) = \frac{\sum_{k=1}^{k_1} MAPE2(k)}{k_1}$$

Step 10: $i = 1$

Step 11: $M_{ij} = \{i\text{-th maximum order in } \hat{FRM1}(j), \forall j = 1, \dots, k_1\}$

Step 12: $\hat{FRM2}(j) = \left\lfloor \frac{\hat{FRM1}(j)}{M_{ij}} \right\rfloor, \forall j = 1, \dots, k_1$ ($\lfloor a \rfloor$ obtains the greatest integer number smaller than the real number a)

Step 13: All of the elements greater than or equal to 1 in $\hat{FRM2}(j) \forall j = 1, \dots, k_1$, are replaced by 1

Step 14: $MAPE2(j) =$

$$= \frac{\sum_{p=1}^r \sum_{q=1}^m \frac{|FRM1_{pq}(j) - \hat{FRM2}_{pq}(j)|}{FRM1_{pq}(j)}}{r.m} \times 100, \forall j = 1, \dots, k_1$$

$$\text{Step 15: } MAPE(i) = \frac{\sum_{j=1}^{k_1} MAPE2(j)}{k_1}$$

Step 16: $i = i + 1$

Step 17: If i -th maximum order is available in $\hat{FRM1}(j), \forall j = 1, \dots, k_1$ then go to step 11 otherwise go to step 18

Step 18: If $MAPE(t+1)$ is the first element greater than the $MAPE$ obtained in method 2, the i -th maximum order is the most correct one to be considered as a benchmark to calculate the errors.

Step 19: $n_3(0) = 0$

Step 20: $i = 0$

Step 21: Use the first $n_1 + n_2 + n_3(i)$ data to determine \hat{FRM} , FRM (n_1 is the number of the first data which can be used to forecast the next n_2 data, which can be used in making a pattern, and $n_3(i)$ is the number of data that occurred during i time periods after the first $n_1 + n_2$ data)

Step 22: Determine \hat{FRM} , FRM

Step 23: $\hat{FRM1}(i+1) = \hat{FRM}(1)$, $FRM1(i+1) = FRM(1)$

Step 24: $i = i + 1$

Step 25: If $n_1 + n_2 + \sum_{j=1}^i n_3(j) < n$ then go to step 21; otherwise $k_2 = i$ and go to step 26

Step 26: $M_j = \{t\text{-th maximum order in } \hat{FRM1}(j), \forall j = 1, \dots, k_2\}$

Step 27: $\hat{FRM2}(j) = \left\lfloor \frac{\hat{FRM1}(j)}{M_j} \right\rfloor, \forall j = 1, \dots, k_2$ ($\lfloor a \rfloor$ obtains the greatest integer number smaller than the real number a)

Step 28: All elements greater than or equal to 1 in $\hat{FRM2}(j), \forall j = 1, \dots, k_2$ are replaced by 1

Step 29: $MAPE2(j) =$

$$= \frac{\sum_{p=1}^r \sum_{q=1}^m \frac{|FRM1_{pq}(j) - \hat{FRM2}_{pq}(j)|}{FRM1_{pq}(j)}}{r.m} \times 100, \forall j = 1, \dots, k_2$$

$$\text{Step 30: } MAPE = \frac{\sum_{j=1}^{k_2} MAPE2(j)}{k_2}$$

Step 31: End.

Note that in algorithm (II), *MAPE* can be substituted by either *MSE* or *MAD*, but because of the aforementioned reasons, *MAPE* is preferred.

b) Deterministic forecasting of earthquake occurrences

Section (a) of method 3, along with the probabilistic forecasting matrixes, such as \hat{FRM} , can be used for deterministic forecasting of earthquake occurrences. Considering algorithm (II), and using the following steps, this makes it possible to forecast the earthquakes occurring in the next several time periods, deterministically (i.e. one and zero; so that one specifies the occurrence and zero specifies the lack of occurrence of any earthquake).

Algorithm III:

Step 0: Begin

Step 1: Use the past n data to determine \hat{FRM} (n is the number of total data)

Step 2: Determine \hat{FRM}

Step 3: $\hat{FRM1}(i) = \hat{FRM}(i) \forall i = 1, \dots, k$ (k is the number of predictable time periods in the future)

Step 4: $M_j = \{t\text{-th maximum order in } \hat{FRM1}(j), \forall j = 1, \dots, k\}$

Step 5: $\hat{FRM2}(j) = \left\lfloor \frac{\hat{FRM1}(j)}{M_j} \right\rfloor \forall j = 1, \dots, k$ ($\lfloor a \rfloor$ obtains the greatest integer number smaller than the real number a)

Step 6: Replace any element(s) greater than or equal to 1 in $\hat{FRM2}(j), \forall j = 1, \dots, k$ by 1 and name the resulting matrices $\hat{FRMD}(j), \forall j = 1, \dots, k$

Step 7: If $\hat{FRMD}_{rm}(j) = 1$ then an earthquake in region r with a magnitude m in j -th time period will occur, otherwise any earthquake in region r with a magnitude m in j -th time period will not occur

Step 8: End.

Application

In this section, the earthquake occurrences are forecast probabilistically and deterministically using the actual data collected from the Iran area. Then, two zoning methods of the Iran area are introduced and considered, consisting of zoning by K94 and a proposed zoning in the present study, and the forecasting error for these zoning methods is determined.

The Iran area was selected as the area of investigation. This area is bounded by longitudes 44.23° E , 63.33° E and latitudes 25.05° N , 39.78° N . The data were collected from the United States Geological Survey website

Magnitude (MB)	Space Distance (Km)	Time Interval (Days)
2.5	19.5	6
3	22.5	11.5
3.5	26	22
4	30	42
4.5	35	83
5	40	155
5.5	47	290
6	54	510
6.5	61	790
7	70	915
7.5	81	960
8	94	985

Table 1. A reference for distinguishing foreshocks and aftershocks from main-shocks [from Sadeghian and Jalali-Naini 2008b].

(<http://neic.usgs.gov/neis/epic/epic.html>). After filtering and removing the unsuitable data according to Table 1 [Gardner and Knopoff 1974], 3,179 data related to earthquakes that occurred during 1973-2007 are used. The maximum time interval between the times of the earthquake occurrences is 45 days, so by considering each 10 days as one time unit, forecasting during the next $\left\lfloor \frac{45}{10} \right\rfloor = 5$ time units in each time will be possible. Also, each zone in each zoning method is considered as a state in the region-to-region transitions of a semi-Markov model, and the magnitude of total occurrences is classified into 5 classes for each of the two zoning methods, and each class of magnitudes is considered as a state in the magnitude-to-magnitude transitions of a semi-Markov model.

Zoning by K94 was carried out according to occurrences or epicenters, while the zoning in the present study was carried out by fault lines. In other words, K94 considered the historical data of earthquake occurrences according to their longitudes and latitudes, and classified them into 22 classes; while the present study considered the historical data of earthquake occurrences according to the longitudes and latitudes of their fault lines, and obtained their linear regressions, and then classify these linear regressions into 10 classes. Since the epicenters are positioned on the fault lines, zoning by fault lines can be more accurate than by epicenters.

a) Zoning by K94 (based on epicenters)

Karakaisi [1994] divided the Iran area into 21 zones. In this section, his zoning method is applied. K94 does not considered the center of Iran as a zone; hence the zoning is used to divide Iran into 22 zones, i.e. 21 zones by K94, plus one central part, as shown in Figure 2.

Also, the magnitudes of the past occurrences were classified into 5 classes, as follows:

$$\begin{aligned} M_1: mb &\leq 3.6 \\ M_2: 3.6 &< mb < 4.8 \\ M_3: 4.8 &< mb < 5.4 \\ M_4: 5.4 &< mb < 6.3 \\ M_5: 6.3 &< mb \end{aligned} \quad (6)$$

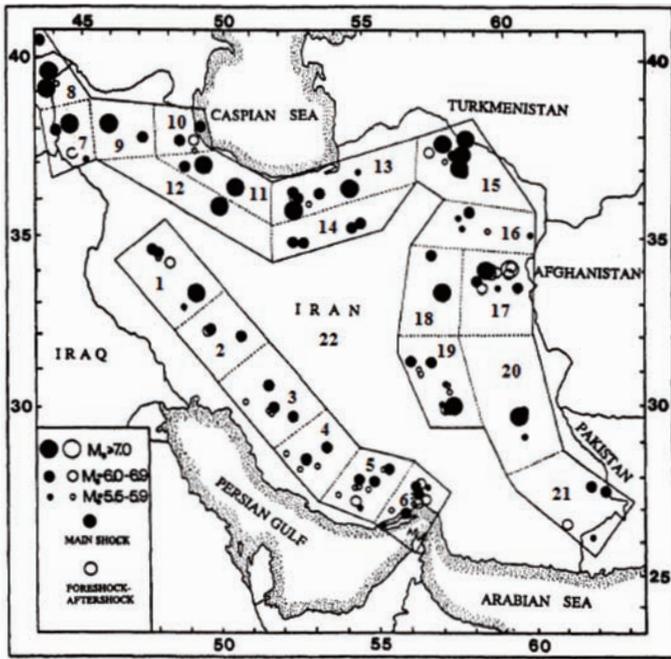


Figure 2. The seismicogenic source zones of Iran, as proposed by K94 [from Karakaisis 1994; re-edited by Sadeghian and Jalali-Naini 2008b].

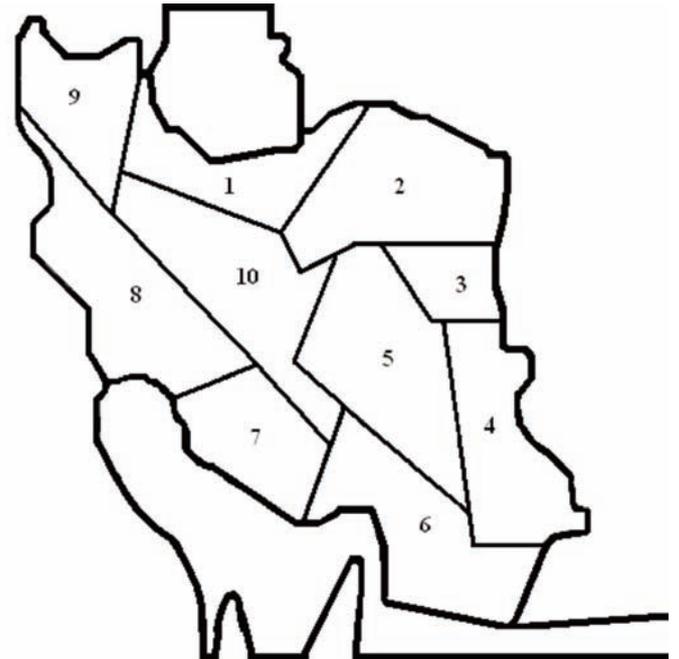


Figure 3. The seismicogenic source zones of Iran proposed by Sadeghian (present study).

These classes were obtained through the agglomerative nesting (AGNES) method, which is a technique for data clustering [Ghazanfari and Rezaei 2006, Jalali-Naini and Sadeghian 2009]. The minimum of the considered magnitudes is 3.1 MB (body wave magnitude) and their maximum is 7.1 MB.

With respect to the K94 zoning in Figure 2 and Equations (1), (2) and (6), the transition probability matrix for both the magnitude-to-magnitude and the region-to-region transitions are obtained as in Tables 7, 8a and 8b, which are included in Appendix 1.

By applying the data in Tables 7, 8a and 8b and Equation (1), interval transition probability matrices in both magnitude-to-magnitude and region-to-region transitions are determined. By using these and Equation (2), the probabilistic forecasting matrix for the next 5 time periods (i.e. the next 50 days) after normalizing are obtained (see also Tables 9a and 9b in Appendix 1, and more comments in the modeling section mentioned above).

Considering $n = 3179$, $n_1 = 3000$, $n_2 = 179$, hence, in zoning by K94, the values of the forecasting error of the algorithms mentioned are calculated as follows:

$$\begin{aligned} MSE &= 0.022 \\ MAD &= 0.05 \\ MAPE &= 5.00\% \end{aligned}$$

In method 3, according to algorithm (II) in the validation section, all of the data are divided into three classes. The data ranging from 1 to 3,000 are used for forecasting the next 104 data, which are used for making a pattern. The data ranging from 3,001 to 3,104, which is equal

to 28 time periods (each time unit is equal to 10 days), are used for deterministic forecasting of the data ranging from 3,105 to 3,179, which are the data later used for determining the forecasting error in the case of deterministic forecasting. In this way, according to algorithm (II) in the validation section, t is equal to 5. This value means that the element(s) greater than 5th maximum order in forecasting matrixes are replaced by 1 and the other elements are replaced by 0, then the deterministic forecasting is the nearest forecasting to the real occurrences and its forecasting error is the least. However by considering $t = 5$ in this zoning method, its MAPE is 2.398%.

b) The proposed zoning (based on fault lines)

In this zoning method, the estimated linear equations of fault lines (according to their longitudes and latitudes) were obtained. Next, the fault lines and their linear equations were classified into 10 zones according to specific attributes, including the Provinces of Iran, their importance, the length of the fault lines, etc. Therefore, Figure 3 was drawn approximately. The above steps were completely run on the MATLAB software.

In this way, the fault lines in the Iran area (shown in Figure 4) [Sadeghian 2007] are divided into 10 zones, as shown in Figure 3.

The classes of the magnitudes are considered similar to the K94 zoning method. The matrices G_R (i.e. probability matrix of region-to-region transitions) and $FRM(i) \forall i = 1, \dots, 5$ are determined through a method similar to the K94 zoning method, as with Table 10, in Appendix 2.

By using Tables 7 and 10, the interval transition probability matrices in both region-to-region and magnitude-

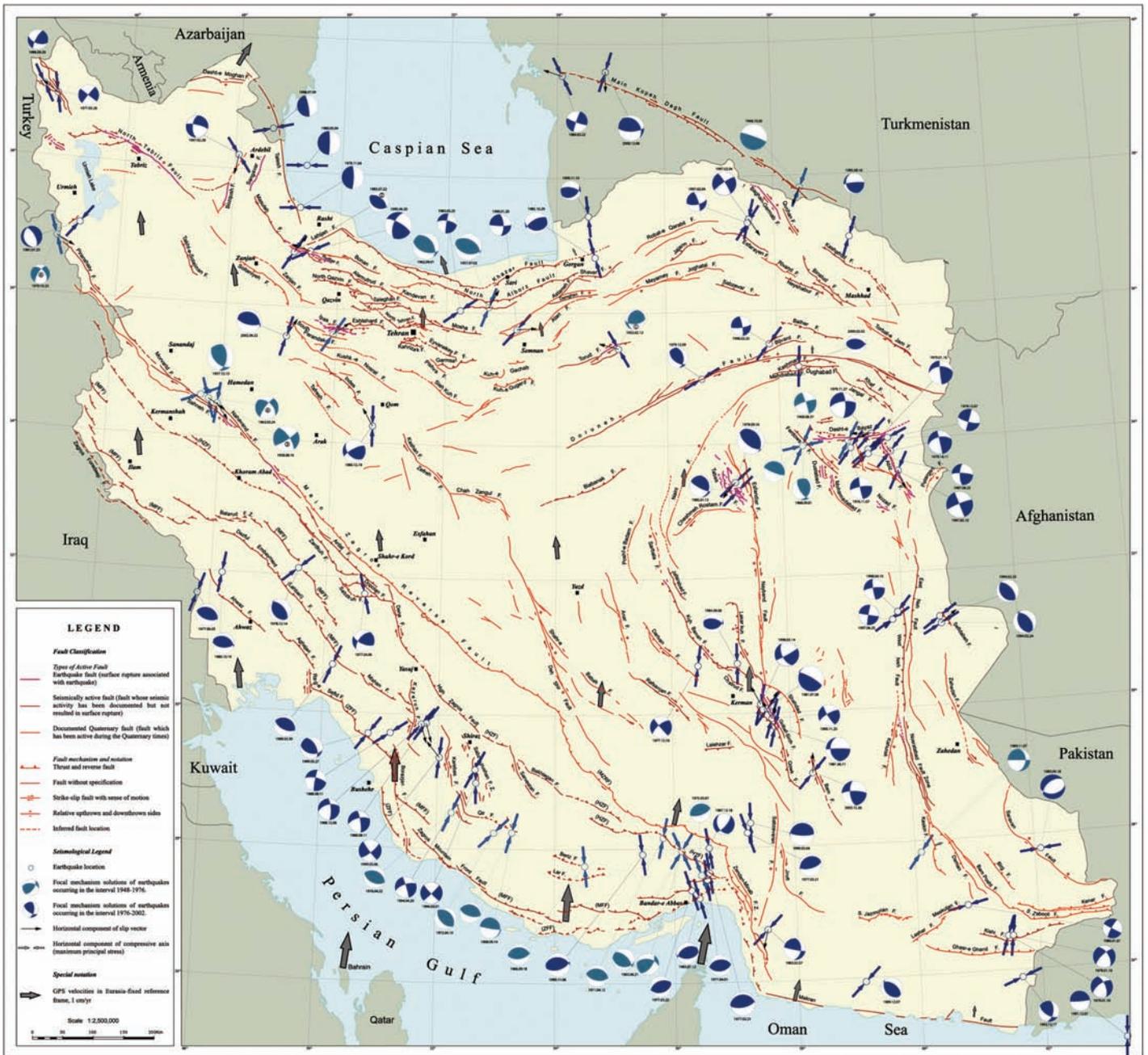


Figure 4. The fault lines in the Iran area [from the website of the International Institute of Earthquake Engineering and Seismology (Tehran, Iran), http://www.iiees.ac.ir/iiees/English/Seismology/eng_seis_faults.html].

to-magnitude transitions are determined; in addition, using these and Equation (2), the probabilistic forecasting matrices for the next 5 time periods (i.e. the next 50 days) after normalizing are determined. (See Tables 11a and 11b in Appendix 2).

In this zoning method, the forecasting errors are determined through a method similar to the K94 zoning method, as follows:

$$\begin{aligned} MSE &= 0.0112 \\ MAD &= 0.0445 \\ MAPE &= 4.45\% \end{aligned}$$

In method 3, similar to the zoning by K94, all of data were divided into three sections. The value of t obtained was

4 and its MAPE in deterministic forecasting error is 3.552%.

Discussion

With respect to the two zoning methods investigated, the percentages of earthquake occurrences in each of the zones are shown in Figures 5 and 6. Moreover, the percentages of earthquake occurrences in each class of magnitudes, which are similar in both methods, are presented in Figure 7.

The transition probability matrices in the magnitude-to-magnitude transitions are the same with both of the zoning methods. Table 7 shows that the maximum and minimum probabilities in the magnitude-to-magnitude transitions are related to the transitions from M_3 to M_2 (1.00) and from M_2 to

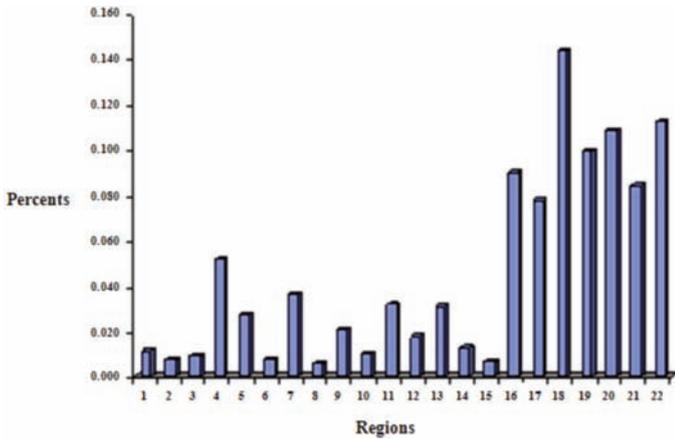


Figure 5. The percentage of occurrences in each region, with the zoning by K94 [from Sadeghian and Jalali-Naini 2008b].

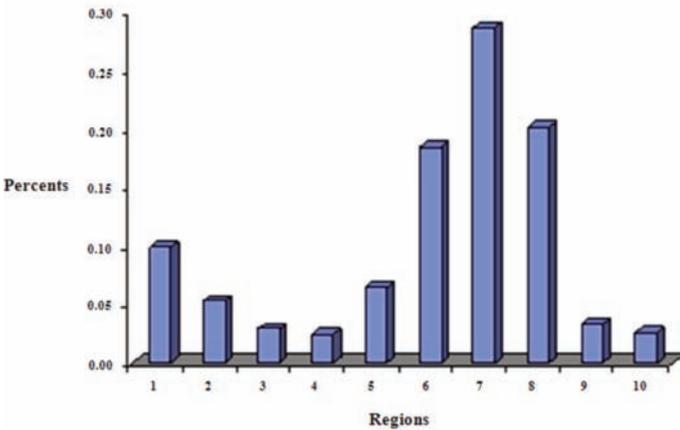


Figure 6. The percentage of occurrences in each region, with the zoning by Sadeghian (present study).

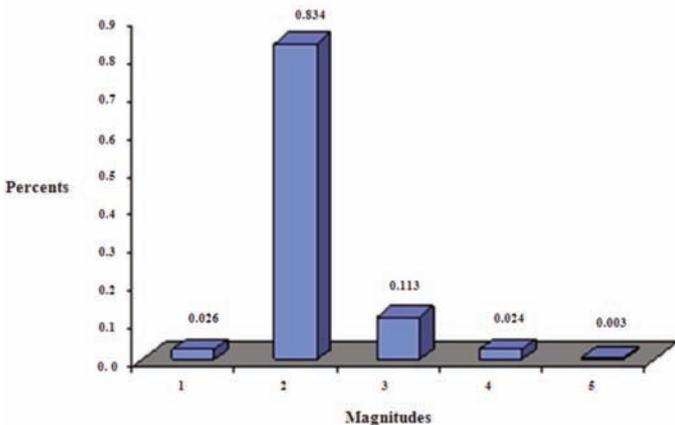


Figure 7. The percentage of occurrences in each class of magnitudes [from Sadeghian and Jalali-Naini 2008b].

M_5 (0.0026), respectively (except for none-zero elements). These results are illustrated in Figure 7.

With respect to the zoning by K94, presented in Tables 8a and 8b, it is clear that the maximum and the minimum probabilities in these tables are related to the $R_8 \rightarrow R_{18}$ (0.333) and $R_{22} \rightarrow R_6$, $R_{22} \rightarrow R_8$ (0.0028) transitions, respectively. Moreover, in the proposed zoning (Table 10), the maximum and minimum probabilities are related to $R_2 \rightarrow R_7$ (0.355) and $R_2 \rightarrow R_{10}$ (0.0059) transitions, respectively.

A summary of the minimum and maximum transition probabilities is shown in Table 2.

The results forecast for the probabilities of earthquake occurrence in the future 5 time periods with the zoning by K94 and the proposed zoning are shown in Tables 9a and 9b, and Tables 11a and 11b, respectively. These tables show that in zoning by K94, the maximum probability of earthquake occurrences pertains to $R_{22}M_2$ in the first time period, and $R_{18}M_2$ in the second, third, fourth and fifth time periods. Similarly, in the proposed zoning, the maximum probability of earthquake occurrences pertains to R_5M_2 in the first time period, and R_7M_2 in the second, third, fourth and fifth time periods. Ultimately, the forecasting errors in both of the zoning methods show that both the number of zones and the type of zoning affect the forecasting probabilities and the forecasting errors. MAPE in the proposed zoning is 4.45%, while in zoning by K94, it is 5.0%. Therefore, it can be concluded that the proposed zoning will be more successful and precise, although both of these methods have rather low forecasting errors.

A summary of the maximum probabilities in the two types of zoning is shown in Table 3.

In the deterministic forecast method, for different t 's, the minimum errors of the forecasting are calculated. So, by considering the forecasting errors obtained through method 3, it can be easily determined that the minimum error of forecasting is related to $t = 5$ in the zoning by K94, and to $t = 4$ in the proposed zoning. The MAPEs in the deterministic forecasting error of zoning by K94 and by the proposed zoning can be obtained through the aforementioned method (i.e. algorithm III). These values are 2.398% in the zoning by K94, and 3.552% in the proposed zoning. These values show that in comparison to the proposed zoning, the K94 zoning method has fewer errors. Of course, this priority has a rational reason, and that is because in method 3, although the sum of the absolute percentage error (SAPE) is the same for both methods, to find the MAPE in the K94 zoning method, the SAPE has to be divided by 22×5 , while, in the proposed zoning, it has to be divided by 10×5 (since the number of the classes of magnitudes is five, the number of the K94 zones is 22, while in the proposed zoning this is just 10); therefore, the MAPE obtained through the K94 zoning method is less than that obtained through the proposed method in the present study. Of course, this priority exists in

Transition	Minimum probability	Maximum probability
Magnitude-to-magnitude transition probability	$M_2 \rightarrow M_5$ (0.0026)	$M_5 \rightarrow M_2$ (1.00)
Region-to-region transition probability for the K94 zoning	$R_{22} \rightarrow R_6, R_{22} \rightarrow R_8$ (0.0028)	$R_8 \rightarrow R_{18}$ (0.333)
Region-to-region transition probability for the proposed zoning	$R_2 \rightarrow R_{10}$ (0.0059)	$R_2 \rightarrow R_7$ (0.355)

Table 2. Summary of minimum and maximum transition probabilities.

Zoning type	Period1	Period2	Period3	Period4	Period5
Maximum probabilities in K94 zoning	$R_{22} M_2$	$R_{18} M_2$	$R_{18} M_2$	$R_{18} M_2$	$R_{18} M_2$
Maximum probabilities in the proposed zoning	$R_5 M_2$	$R_7 M_2$	$R_7 M_2$	$R_7 M_2$	$R_7 M_2$

Table 3. Summary of maximum probabilities in the two types of zoning.

Zoning type	Future time periods 1 to 5
K94 zoning method	$R_{16} M_2, R_{18} M_2, R_{19} M_2, R_{20} M_2, R_{22} M_2$
Proposed zoning method	$R_5 M_2, R_6 M_2, R_7 M_2, R_8 M_2$

Table 4. Deterministic forecasting matrix during the first to fifth time periods for zoning by K94 and by the proposed zoning.

Period	K94 zoning method	Proposed zoning method
1	$R_{16} M_2, R_{20} M_2, R_{21} M_2$	$R_6 M_2, R_8 M_2,$
2	$R_2 M_2, R_{13} M_2, R_{14} M_2$	$R_4 M_2, R_5 M_2, R_9 M_2,$
3	$R_7 M_2, R_{17} M_2, R_{18} M_2, R_{19} M_2, R_{22} M_2$	$R_2 M_2, R_4 M_2, R_7 M_2, R_8 M_2,$
4	$R_{12} M_2, R_{18} M_2, R_{18} M_2, R_{19} M_2, R_{20} M_3$	$R_5 M_2, R_6 M_3, R_7 M_1, R_7 M_2,$
5	$R_{20} M_2, R_{21} M_2, R_{22} M_3$	$R_6 M_2, R_6 M_3, R_7 M_2,$

Table 5. Actual earthquake occurrences during the following 5 time periods after the last earthquake occurrence.

Zoning type	Forecast (%)			
	Completely correct	Region correct, Magnitude wrong	Adjacent region correct, Magnitude correct	Not forecast
K94 zoning	42	16	16	26
Proposed zoning	56	19	19	6

Table 6. Summary of the results of the forecasting.

deterministic forecasting just due to the greater number of zones in the K94 zoning method, so it does not exist for probabilistic forecasting.

Conclusion

In this study, first, the method of applying the semi-Markov model in forecasting these triad dimensions of earthquakes was mentioned. Then, the method of calculating the probabilistic and deterministic forecasting was elaborated on. Moreover, several algorithms for validating the proposed model and obtaining the forecasting errors are introduced. Finally, while introducing two zoning methods, as those of K94 and of the proposed zoning, a case study on the Iran area is carried out and the forecasting errors of the proposed models are obtained for both of these zoning methods. In conclusion, it can be noted that both the number of zones and the type of zoning affect the forecast probabilities and the forecasting errors. The forecasting errors show that the *MAPE* for the K94 zoning is 5.0%, while it is just 4.45% in the proposed zoning. These levels show that the forecasting error can be reduced through manipulating the numbers of zones or the types of zones.

With respect to method 3, in calculating the deterministic forecasting errors, this method can also forecast the next earthquake occurrences deterministically. With respect to the collected data in the case study, the last earthquake considered in this study occurred on March 26, 2007, with a magnitude of 4.9 M_B , M_s , and in the regions of R_{22} and R_5 (in the K94 and the proposed zonings, respectively). Accordingly, the deterministic forecasting during the future 5 time periods (the future 50 days) are determined in the K94 and the proposed zoning by algorithm III, as shown in Table 4.

During the 50 days after carrying out the proposed model in the Iran area, some earthquakes occurred for both of the zoning methods. These are presented in Table 5.

These results obtained show that in the zoning by K94, 42% of the earthquakes were correctly forecast. The areas of 16% of the earthquakes were correctly forecast, but not their magnitudes. Then 16% of the earthquakes occurred in areas adjacent to the forecast earthquakes, with the correct magnitudes, and 26% of the earthquakes were not forecast at all.

In the proposed zoning, 56% of the earthquakes were correctly forecast in all of the three dimensions. The areas of 19% of the earthquakes were correctly forecast, but not their magnitudes. Then 19% of the earthquakes occurred in areas adjacent to the forecast earthquakes, with the correct magnitudes, and 6% of the earthquakes were not forecast at all. The summary of these results is given in Table 6.

In addition to the errors obtained and explained above, the statistics mentioned here show that in comparison to the K94 zoning, the proposed zoning provides more accurate

forecasting. One of the reasons for the priority of the proposed zoning over that of K94 is that the K94 zoning was not performed correctly (i.e. it did not take into account the center of Iran, which is a very extended area, while this was part of the zones considered in the present study).

References

- Altinok, Y. and D. Kolcak (1999). An application of the semi-Markov model for earthquake occurrences in north Anatolia, Turkey, *J. Balkan Geophys. Soc.*, 2 (4), 90-99.
- Console, R. (2001). Testing earthquake forecast hypotheses, *Tectonophysics*, 338, 261-268.
- Console, R., D. Pantosti and G. D'Addezio (2002). Probabilistic approach to earthquake prediction, *Annals of Geophysics*, 45 (6), 723-731.
- Di Luccio, F., R. Console, M. Imoto and M. Murru (1997). Analysis of short time-space range seismicity pattern in Italy, *Annals of Geophysics*, 40 (4), 783-798.
- Gardner, J.K. and L. Knopoff (1974). Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian?, *B. Seismol. Soc. Am.*, 64, 1363-1367.
- Ghazanfari, M. and M. Rezaei (2006). Introduction on Fuzzy Theory and its Applications, IUST Press, Tehran.
- Haj Shirmohammadi, A. (2003). Principle of production and inventory planning and control, Arkan Press, Isfahan.
- Jalali-Naini, G.R. (1997). Stochastic processes and its applications in biology, agriculture, management and industries, Statistics and Information Office Press, Tehran.
- Jalali-Naini, G.R. and R. Sadeghian (2009). Forecasting time and place of earthquakes using a semi-Markov model, *Mathematical Modelling and Applied Computing*, 2, 1.
- Karakaisis, G.F. (1994). Long-term earthquake prediction in Iran based on the time- and magnitude-predictable model, *Phys. Earth Planet. In.*, 83, 129-145.
- Kelleher, J., L. Sykes and O. Jack (1973). Possible Criteria for Predicting Earthquake Locations and Their Application to Major Plate Boundaries of the Pacific and the Caribbean, *J. Geophys. Res.*, 78 (14), 2547-2585.
- Minh, D.L. (2001). Applied Probability Models, Duxbury, Thomson Learning Press, California.
- Nanjo, Z.K., R.J. Holliday, C. Chen, B.J. Rundle and L.D. Turcotte (2006). Application of a modified pattern informatics method to forecasting the locations of future Large earthquakes in central Japan, *Tectonophysics*, 424, 351-366.
- Nava, F.A., C. Herrera, J. Frez and E. Glowacka (2005). Seismic hazard evaluation using Markov chains: application to the Japan area, *Pure Appl. Geophys.*, 162, 1347-1366.
- Papazachos, B.C. (1992). A time- and magnitude-predictable model for generation of shallow earthquakes in the Aegean area, *Pure Appl. Geophys.*, 138 (2), 287-308.
- Patwardhan, S.A., B.R. Kulkarni and D. Tocher (1980). A semi-Markov model for characterizing recurrence of great earthquakes, *B. Seismol. Soc. Am.*, 70 (1), 323-347.

- Rice, J.A. (2007). *Mathematical Statistics and Data Analysis*, Thomson Press, California.
- Rikitake, T. (1976). Recurrence of great earthquakes at subduction zones, *Tectonophysics*, 35, 335-362.
- Sadeghian, R. and G.R. Jalali-Naini (2008a). A new probability density function in earthquake occurrences, *J. Ind. Engineering Int.*, 4 (6), 1-6.
- Sadeghian, R. and G.R. Jalali-Naini (2008b). Applying semi-Markov models for forecasting the triple dimensions of next earthquake occurrences, *Dynamics of Continuous, Discrete and Impulsive Systems, Series B: Applications & Algorithms*, 15, 395-416.
- Sadeghian, R. (2007). *Whatever You Don't Know about Earthquakes*, Shooka Press, Tehran.

*Corresponding author: Ramin Sadeghian,
Bu Ali Sina University, Hamedan, Iran; e-mail: rsadeghian@basu.ac.ir.

		M_1	M_2	M_3	M_4	M_5	
$G_M =$		M_1	0.0482	0.9157	0.0241	0.0120	0.0000
		M_2	0.0275	0.8368	0.1108	0.0222	0.0026
		M_3	0.0112	0.8123	0.1317	0.0448	0.0000
		M_4	0.0260	0.7532	0.1948	0.0130	0.0130
		M_5	0.0000	1.0000	0.0000	0.0000	0.0000

	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	R_{11}	R_{12}	R_{13}	R_{14}	R_{15}	R_{16}	R_{17}	R_{18}	R_{19}	R_{20}	R_{21}	R_{22}
R_1	0.000	0.028	0.000	0.083	0.000	0.000	0.000	0.028	0.000	0.000	0.000	0.000	0.111	0.000	0.000	0.083	0.111	0.222	0.083	0.083	0.056	0.111
R_2	0.000	0.000	0.045	0.091	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.045	0.091	0.091	0.091	0.227	0.136	0.182
R_3	0.069	0.000	0.034	0.034	0.000	0.000	0.034	0.000	0.000	0.138	0.000	0.000	0.103	0.034	0.000	0.034	0.069	0.103	0.069	0.103	0.034	0.138
R_4	0.012	0.006	0.018	0.079	0.067	0.012	0.024	0.006	0.018	0.000	0.012	0.018	0.024	0.012	0.012	0.091	0.061	0.140	0.110	0.073	0.061	0.140
R_5	0.000	0.000	0.000	0.126	0.264	0.000	0.011	0.000	0.000	0.011	0.011	0.000	0.023	0.023	0.000	0.080	0.080	0.069	0.092	0.057	0.080	0.069
R_6	0.000	0.000	0.000	0.000	0.043	0.000	0.000	0.000	0.000	0.043	0.000	0.000	0.000	0.000	0.000	0.043	0.043	0.174	0.000	0.217	0.217	0.217
R_7	0.009	0.009	0.000	0.018	0.053	0.000	0.035	0.000	0.035	0.018	0.026	0.035	0.061	0.000	0.009	0.044	0.070	0.167	0.088	0.114	0.088	0.123
R_8	0.000	0.000	0.056	0.056	0.000	0.000	0.000	0.000	0.056	0.000	0.000	0.000	0.000	0.000	0.000	0.167	0.056	0.333	0.167	0.111	0.000	0.000
R_9	0.015	0.000	0.000	0.090	0.000	0.015	0.030	0.015	0.045	0.000	0.030	0.015	0.060	0.000	0.000	0.075	0.045	0.134	0.134	0.119	0.104	0.075
R_{10}	0.000	0.000	0.000	0.065	0.032	0.000	0.065	0.000	0.000	0.032	0.000	0.032	0.000	0.032	0.000	0.065	0.097	0.065	0.161	0.194	0.000	0.161
$G_R =$	R_{11}	0.000	0.000	0.010	0.040	0.010	0.040	0.000	0.030	0.010	0.280	0.010	0.010	0.010	0.010	0.110	0.030	0.080	0.050	0.080	0.050	0.120
	R_{12}	0.000	0.018	0.018	0.018	0.018	0.054	0.018	0.000	0.036	0.071	0.179	0.036	0.000	0.000	0.018	0.125	0.125	0.000	0.089	0.018	0.161
	R_{13}	0.051	0.010	0.010	0.051	0.031	0.010	0.051	0.010	0.010	0.020	0.031	0.163	0.031	0.010	0.102	0.061	0.082	0.020	0.092	0.071	0.082
	R_{14}	0.000	0.000	0.000	0.073	0.000	0.024	0.000	0.000	0.024	0.000	0.000	0.000	0.098	0.000	0.122	0.049	0.171	0.049	0.073	0.146	0.171
	R_{15}	0.000	0.000	0.000	0.048	0.048	0.000	0.000	0.000	0.048	0.000	0.000	0.095	0.000	0.000	0.190	0.048	0.048	0.190	0.095	0.095	0.095
	R_{16}	0.007	0.021	0.010	0.031	0.031	0.010	0.028	0.010	0.007	0.028	0.010	0.017	0.014	0.007	0.140	0.101	0.143	0.119	0.084	0.056	0.091
	R_{17}	0.016	0.000	0.012	0.053	0.016	0.041	0.016	0.016	0.016	0.012	0.024	0.020	0.012	0.008	0.093	0.093	0.167	0.077	0.114	0.085	0.106
	R_{18}	0.015	0.007	0.013	0.046	0.018	0.037	0.004	0.020	0.000	0.024	0.015	0.020	0.018	0.007	0.094	0.074	0.182	0.098	0.098	0.092	0.107
	R_{19}	0.013	0.003	0.006	0.054	0.006	0.035	0.003	0.028	0.013	0.016	0.009	0.044	0.013	0.006	0.070	0.089	0.130	0.120	0.073	0.108	0.155
	R_{20}	0.009	0.003	0.003	0.035	0.009	0.047	0.009	0.026	0.017	0.032	0.015	0.017	0.009	0.003	0.076	0.084	0.160	0.102	0.172	0.087	0.073
	R_{21}	0.004	0.007	0.004	0.067	0.026	0.049	0.000	0.011	0.007	0.022	0.015	0.015	0.011	0.004	0.086	0.086	0.109	0.101	0.120	0.116	0.131
	R_{22}	0.011	0.011	0.008	0.054	0.017	0.003	0.034	0.003	0.025	0.023	0.011	0.028	0.006	0.014	0.099	0.056	0.152	0.127	0.124	0.076	0.107

Appendix 1: Transition probability matrices in the zoning according to K94
 Table 7 (top). Transition probability matrix of the magnitude-to-magnitude transitions in the zoning by K94. Table 8a/b (bottom, left/right). Probability matrix of the region-to-region transitions in the zoning by K94.

Appendix 1: Transition probability matrices in the zoning according to K94

Table 9a (left to right). Probabilistic forecasting matrixes during the first, second, third and fourth time periods in the zoning by K94.

	M_1	M_2	M_3	M_4	M_5		M_1	M_2	M_3	M_4	M_5		M_1	M_2	M_3	M_4	M_5
R_1	0.0007	0.0536	0.0190	0.0034	0.0000	R_1	0.0018	0.0668	0.0136	0.0018	0.0000	R_1	0.0025	0.0712	0.0109	0.0017	0.0000
R_2	0.0007	0.0536	0.0190	0.0034	0.0000	R_2	0.0018	0.0533	0.0108	0.0018	0.0000	R_2	0.0017	0.0469	0.0067	0.0017	0.0000
R_3	0.0007	0.0400	0.0142	0.0027	0.0000	R_3	0.0018	0.0515	0.0108	0.0018	0.0000	R_3	0.0017	0.0528	0.0084	0.0017	0.0000
R_4	0.0027	0.2266	0.0807	0.0142	0.0000	R_4	0.0108	0.3631	0.0732	0.0117	0.0009	R_4	0.0109	0.3554	0.0536	0.0101	0.0008
R_5	0.0007	0.0801	0.0285	0.0054	0.0000	R_5	0.0045	0.1536	0.0307	0.0045	0.0000	R_5	0.0050	0.1660	0.0251	0.0050	0.0008
R_6	0.0000	0.0136	0.0047	0.0007	0.0000	R_6	0.0009	0.0434	0.0090	0.0018	0.0000	R_6	0.0017	0.0453	0.0067	0.0017	0.0000
R_7	0.0014	0.1201	0.0427	0.0075	0.0000	R_7	0.0081	0.2773	0.0560	0.0090	0.0009	R_7	0.0075	0.2490	0.0377	0.0075	0.0008
R_8	0.0000	0.0136	0.0047	0.0007	0.0000	R_8	0.0009	0.0361	0.0072	0.0009	0.0000	R_8	0.0008	0.0402	0.0059	0.0008	0.0000
R_9	0.0014	0.1065	0.0380	0.0068	0.0000	R_9	0.0045	0.1617	0.0325	0.0054	0.0000	R_9	0.0042	0.1475	0.0226	0.0042	0.0000
R_{10}	0.0007	0.0536	0.0190	0.0034	0.0000	R_{10}	0.0018	0.0650	0.0126	0.0018	0.0000	R_{10}	0.0017	0.0645	0.0101	0.0017	0.0000
R_{11}	0.0014	0.1065	0.0380	0.0068	0.0000	R_{11}	0.0054	0.1798	0.0361	0.0054	0.0000	R_{11}	0.0059	0.1953	0.0293	0.0059	0.0008
R_{12}	0.0007	0.0400	0.0142	0.0027	0.0000	R_{12}	0.0036	0.1174	0.0235	0.0036	0.0000	R_{12}	0.0034	0.1157	0.0176	0.0034	0.0000
R_{13}	0.0014	0.1201	0.0427	0.0075	0.0000	R_{13}	0.0072	0.2240	0.0452	0.0072	0.0009	R_{13}	0.0067	0.2205	0.0335	0.0067	0.0008
R_{14}	0.0000	0.0265	0.0095	0.0020	0.0000	R_{14}	0.0018	0.0696	0.0145	0.0018	0.0000	R_{14}	0.0025	0.0788	0.0117	0.0025	0.0000
R_{15}	0.0007	0.0665	0.0237	0.0041	0.0000	R_{15}	0.0018	0.0524	0.0108	0.0018	0.0000	R_{15}	0.0017	0.0436	0.0067	0.0017	0.0000
R_{16}	0.0047	0.4132	0.1479	0.0265	0.0000	R_{16}	0.0199	0.6631	0.1328	0.0208	0.0018	R_{16}	0.0193	0.6320	0.0964	0.0184	0.0017
R_{17}	0.0027	0.2266	0.0807	0.0142	0.0000	R_{17}	0.0154	0.5068	0.1021	0.0163	0.0009	R_{17}	0.0168	0.5608	0.0855	0.0168	0.0017
R_{18}	0.0075	0.6398	0.2286	0.0414	0.0000	R_{18}	0.0307	1.0000	0.2014	0.0316	0.0027	R_{18}	0.0302	1.0000	0.1517	0.0293	0.0025
R_{19}	0.0061	0.5197	0.1859	0.0332	0.0000	R_{19}	0.0226	0.7525	0.1518	0.0235	0.0018	R_{19}	0.0201	0.6731	0.1023	0.0193	0.0017
R_{20}	0.0061	0.4796	0.1716	0.0305	0.0000	R_{20}	0.0253	0.8347	0.1680	0.0262	0.0018	R_{20}	0.0226	0.7653	0.1165	0.0226	0.0017
R_{21}	0.0041	0.3331	0.1194	0.0217	0.0000	R_{21}	0.0181	0.5827	0.1174	0.0181	0.0018	R_{21}	0.0176	0.5809	0.0880	0.0168	0.0017
R_{22}	0.0122	1.0000	0.3575	0.0645	0.0000	R_{22}	0.0271	0.8862	0.1780	0.0280	0.0018	R_{22}	0.0243	0.8013	0.1215	0.0235	0.0017

$\hat{FRM}(4)$

$\hat{FRM}(3)$

$\hat{FRM}(2)$

$\hat{FRM}(1)$

	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}
R_1	0.2173	0.0319	0.0160	0.0319	0.0639	0.1534	0.2396	0.1981	0.0288	0.0192
R_2	0.0651	0.1006	0.0178	0.0000	0.0947	0.2012	0.3550	0.1183	0.0414	0.0059
R_3	0.0652	0.0761	0.2826	0.0326	0.0543	0.1630	0.1630	0.0978	0.0326	0.0326
R_4	0.0513	0.0641	0.0256	0.0513	0.0513	0.1795	0.3205	0.2308	0.0128	0.0128
R_5	0.0891	0.0693	0.0396	0.0297	0.1634	0.1436	0.1881	0.1881	0.0545	0.0347
R_6	0.1111	0.0496	0.0239	0.0256	0.0359	0.2376	0.2632	0.2000	0.0222	0.0308
R_7	0.0757	0.0516	0.0209	0.0231	0.0527	0.1866	0.3304	0.1998	0.0340	0.0252
R_8	0.1000	0.0484	0.0156	0.0234	0.0672	0.1594	0.3000	0.2219	0.0359	0.0281
R_9	0.0377	0.0377	0.0377	0.0189	0.0849	0.1981	0.2547	0.2453	0.0566	0.0283
R_{10}	0.0488	0.0610	0.0122	0.0244	0.0488	0.1585	0.2927	0.3171	0.0122	0.0244

$G_R =$

	M_1	M_2	M_3	M_4	M_5
R_1	0.0041	0.3198	0.1141	0.0203	0.0000
R_2	0.0029	0.2601	0.0927	0.0168	0.0000
R_3	0.0017	0.1599	0.0574	0.0104	0.0000
R_4	0.0012	0.1199	0.0429	0.0075	0.0000
R_5	0.0122	1.0000	0.3575	0.0643	0.0000
R_6	0.0058	0.5000	0.1790	0.0319	0.0000
R_7	0.0081	0.6599	0.2358	0.0423	0.0000
R_8	0.0081	0.6802	0.2433	0.0435	0.0000
R_9	0.0023	0.1999	0.0713	0.0127	0.0000
R_{10}	0.0017	0.1402	0.0498	0.0093	0.0000

$\hat{FRM}(1)$

$\hat{FRM}(2)$

Appendix 2: Transition probability matrices in the proposed zoning (present study)

Table 10 (top). Probability matrix of the region-to-region transitions in the proposed zoning.

Table 11a (bottom). Probabilistic forecasting matrices during the first and second time periods in the proposed zoning.

	M_1	M_2	M_3	M_4	M_5
R_1	0.0025	0.0712	0.0099	0.0025	0.0000
R_2	0.0017	0.0480	0.0066	0.0017	0.0000
R_3	0.0017	0.0613	0.0083	0.0017	0.0000
R_4	0.0108	0.3634	0.0522	0.0108	0.0008
R_5	0.0050	0.1722	0.0248	0.0050	0.0008
R_6	0.0017	0.0464	0.0066	0.0017	0.0000
R_7	0.0075	0.2500	0.0356	0.0075	0.0008
R_8	0.0008	0.0406	0.0058	0.0008	0.0000
R_9	0.0041	0.1440	0.0207	0.0041	0.0000
R_{10}	0.0017	0.0637	0.0091	0.0017	0.0000
R_{11}	0.0058	0.2028	0.0290	0.0058	0.0008
R_{12}	0.0033	0.1175	0.0166	0.0033	0.0000
R_{13}	0.0066	0.2185	0.0315	0.0066	0.0008
R_{14}	0.0025	0.0836	0.0116	0.0025	0.0000
R_{15}	0.0017	0.0464	0.0066	0.0017	0.0000
R_{16}	0.0190	0.6283	0.0894	0.0182	0.0017
R_{17}	0.0166	0.5629	0.0803	0.0166	0.0017
R_{18}	0.0298	1.0000	0.1424	0.0298	0.0025
R_{19}	0.0199	0.6523	0.0927	0.0190	0.0017
R_{20}	0.0224	0.7434	0.1060	0.0215	0.0017
R_{21}	0.0174	0.5828	0.0828	0.0174	0.0017
R_{22}	0.0232	0.7740	0.1101	0.0232	0.0025

$\hat{FRM}(5)$

Appendix 1.

Table 9b. Probabilistic forecasting matrixes during the fifth time period in the zoning by K94.

Appendix 2: Transition probability matrices in the proposed zoning (present study)

Table 11b (left to right). Probabilistic forecasting matrices during the third, fourth and fifth time periods in the proposed zoning.

	M_1	M_2	M_3	M_4	M_5		M_1	M_2	M_3	M_4	M_5		M_1	M_2	M_3	M_4	M_5
R_1	0.0106	0.3527	0.0536	0.0101	0.0009	R_1	0.0103	0.3502	0.0511	0.0099	0.0009	R_1	0.0102	0.3451	0.0492	0.0102	0.0008
R_2	0.0057	0.1944	0.0295	0.0057	0.0004	R_2	0.0056	0.1897	0.0275	0.0056	0.0004	R_2	0.0055	0.1865	0.0267	0.0055	0.0004
R_3	0.0031	0.1038	0.0158	0.0031	0.0004	R_3	0.0030	0.0996	0.0146	0.0030	0.0004	R_3	0.0030	0.0971	0.0140	0.0030	0.0004
R_4	0.0026	0.0809	0.0123	0.0022	0.0000	R_4	0.0026	0.0828	0.0120	0.0026	0.0004	R_4	0.0025	0.0835	0.0119	0.0025	0.0004
R_5	0.0088	0.2863	0.0435	0.0084	0.0009	R_5	0.0073	0.2412	0.0352	0.0069	0.0009	R_5	0.0068	0.2289	0.0326	0.0068	0.0004
R_6	0.0202	0.6706	0.1020	0.0198	0.0018	R_6	0.0197	0.6579	0.0957	0.0189	0.0017	R_6	0.0195	0.6499	0.0924	0.0191	0.0017
R_7	0.0299	1.0000	0.1522	0.0290	0.0026	R_7	0.0300	1.0000	0.1459	0.0288	0.0026	R_7	0.0301	1.0000	0.1424	0.0292	0.0025
R_8	0.0215	0.7216	0.1099	0.0211	0.0018	R_8	0.0215	0.7236	0.1056	0.0206	0.0021	R_8	0.0216	0.7202	0.1026	0.0212	0.0021
R_9	0.0035	0.1165	0.0176	0.0035	0.0004	R_9	0.0034	0.1155	0.0167	0.0034	0.0004	R_9	0.0034	0.1132	0.0161	0.0034	0.0004
R_{10}	0.0031	0.0954	0.0145	0.0026	0.0004	R_{10}	0.0030	0.0957	0.0137	0.0026	0.0004	R_{10}	0.0030	0.0950	0.0136	0.0030	0.0004
	$\hat{FRM}(3)$						$\hat{FRM}(4)$						$\hat{FRM}(5)$				