Turkish Journal of Earth Sciences

http://journals.tubitak.gov.tr/earth/

Research Article

Turkish J Earth Sci
(2014) 23: 350-360

doi:10.3906/yer-1212-9

Seismic hazard assessment of Turkey by statistical approaches

Serpil ÜNAL[1,*], Salih ÇELEBİOĞLU[2], Bülent ÖZMEN[3]
[1]Department of Statistics, Faculty of Arts and Sciences, Uşak University, Uşak, Turkey
[2]Department of Statistics, Faculty of Sciences, Gazi University, Ankara, Turkey
[3]Earthquake Engineering Application and Research Center, Gazi University, Ankara, Turkey

Received: 31.12.2012 • **Accepted:** 11.02.2014 • **Published Online:** 21.03.2014 • **Printed:** 18.04.2014

Abstract: In this study, 2 probabilistic methods are presented for seismic hazard assessment in Turkey: Markov chains based on modeling the transition probabilities of states (related to the presence or absence of the earthquakes having magnitude $M \geq 4$ during the time interval $\Delta t = 0.07$ years in each region of Turkey located between 36°N and 42°N and 26°E and 45°E), and the Poisson model, used for computing occurrence probability and recurrence periods of earthquakes. In particular, it should be stated that in this study, our purpose is not to compare the results obtained from these 2 methods. The main purpose is to show that earthquakes occurring in Turkey can be modeled successfully by both a Markov chain, in which we have a different zoning, and the Poisson model, which can determine seismic hazard.

Key words: Markov chain, Poisson model, seismic hazard, entropy, earthquake in Turkey

1. Introduction

For a seismically active region, it is impossible to know for certain in advance the parameters of an earthquake regarding time, location, magnitude, severity, etc. However, statistical studies in the fields of geophysical, geological, and earthquake engineering show that the parameters of possible earthquakes can only be estimated probabilistically, called a seismic hazard.

Early attempts in constructing seismic hazard maps provided estimates of the severity of ground shaking or damage from known or likely earthquakes. Modern seismic hazard assessment began in the late 1960s, with the publication of a series of papers describing and applying the probabilistic seismic hazard assessment method (Lee et al., 2003). Since earthquakes demonstrate randomness according to parameters and there are various uncertainties (such as some deficiencies in the earthquake records), seismic hazard estimation with probabilistic methods is seen as the most appropriate method. The first models used for estimating seismic hazard were based on the assumption that earthquakes are independent from the times and places that they occur. The Poisson model (Yücemen and Akkaya, 1995; Kasap and Gürlen, 2003; Çobanoğlu et al., 2006) is the most common of such models. Afterwards, attempts to estimate seismic hazard were made using methods such as Gumbel extreme value distributions (Yücemen and Akkaya, 1995; Çobanoğlu et al., 2006), exponential distributions (Kasap and Gürlen, 2003; Çobanoğlu et al., 2006), and cluster analysis based on separating earthquakes into groups in time and Weibull distributions, a special form of the same model. In subsequent studies, the Markov model (Pınar et al., 1989; Yücemen and Akkaya, 1995; Ulutaş and Özer, 2000; Nava et al., 2005; Ünal and Çelebioğlu, 2011), based on the assumption that earthquakes indicate a dependence on the time dimension in connection with elastic rebound theory, and the semi-Markov model (Altınok, 1988; Altınok, 1991; Altınok and Kolçak, 1999), based on the assumption that earthquakes are dependent on the space dimension, have been used.

Turkey is located on the Alpine-Himalayan (Mediterranean) seismic belt, one of the most important seismic belts of the world. For this reason, the seismicity of this belt has been the subject of many studies and has attracted the attention of geologists. Nowadays, since increasing amounts of earthquake data are available in Turkey, the estimation of seismic hazard has gained greater importance (Bağcı, 1996).

Compared with a previous study (Ünal and Çelebioğlu, 2011), we have used a different approach for zoning to estimate seismic hazard in Turkey with Markov chains and the Poisson model. We think that the new approach

* Correspondence: serpil.unal@usak.edu.tr

will illuminate the interaction between the degrees of earthquake zonings regarding occurrence of earthquakes, whereas the previous study examined the interaction between the geographical zones from the same perspective.

2. Methods

In this section, we will summarize the methodology used for estimating the seismic hazard.

2.1. Markov chain

The Markov chain is an important class of stochastic processes concerning the sequence of random variables that correspond to the states of a certain system. In this system, the state at one time epoch depends only that in the previous time epoch (Çınlar, 1975), and hence the observed information in the last time epoch eliminates the need for the information in the preceding time epochs.

Let us consider a stochastic process $X = \{X_n{:}n = 0,1,2,...\}$ having a finite or countable infinite state space S. When $X_n = i$, we say that the process is in state i at time n. The probability that the process is in state j in the next time, provided that its present state is i, is denoted by P_{ij} and is called the (1-step) transition probability from state i to state j (Çınlar, 1975).

Definition 1. Suppose that there is a fixed probability P_{ij} being independent of time such that

$$P\{X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = i\} = P\{X_{n+1} = j \mid X_n = i\} = P_{ij}; n \geq 0$$

where $i_0, i_1, ..., i_{n-1}, i, j \in S$. Then $X = \{X_n{:}n = 0,1,2,...\}$ is called a Markov chain (Çınlar, 1975).

By this definition, a Markov chain is a sequence of random variables such that for any n, the next state of process X_{n+1} is independent of the past states $X_0, X_1, ..., X_{n-1}$, that is, the strong Markov property is to hold at randomly chosen times (Çınlar, 1975).

In the literature on Markov chains, it is customary to arrange the transition probabilities in a matrix form and call the resulting matrix the transition matrix. The elements of a transition matrix hold the following conditions:

a) for any 2 states $i, j \in, S, P_{ij} \geq 0$; and
b) for all $i \in S$, $\sum_j P_{ij} = 1$.

As seen in the next theorem and corollary, the joint distribution of $X_0, X_1, ... X_m$ can be completely specified for every m once the initial distribution and the transition matrix P are given (Çınlar, 1975).

Theorem 1. Let $X = \{X_n : n \in N\}$ be a Markov chain. For any $m, n \in N$; $m \geq 1$ and $i_0, i_1, ..., i_m \in S$,

$$P\{X_{n+1} = i_1, X_{n+2} = i_2, \ldots X_{n+m} = i_m \mid X_n = i_0\} = P_{i_0 i_1} P_{i_1 i_2} \ldots P_{i_{m-1} i_m}$$

Corollary 1. For the Markov chain, let the initial probability distribution π_0 be given on the state space S; i.e. let $P\{X_0 = i\} = \pi_0(i)$ for all $i \in S$. Then for $m \in N$ and $i_0, i_1, ..., i_m \in S$,

$$P\{X_0 = i_0, X_1 = i_1, \ldots, X_m = i_m\} = \pi_0(i_0) P_{i_0 i_1} P_{i_1 i_2} \ldots P_{i_{m-1} i_m}$$

In some cases, it may be necessary to calculate the probabilities of the transitions between distant times for the Markov chain. The following definition can then be given.

Definition 2. For any $m \in N$, n-step transition probability from state i to state j is given by

$$P\{X_{m+n} = j \mid X_m = i\} = P_{ij}^{(n)}; i, j \in S, n \in N;$$

where $P_{ij}^{(n)}$ is the (i, j) th element of the nth power of transition probability matrix P.

The following theorem reflects how to calculate the steady state probabilities for the process (Ching and Ng, 2006).

Theorem 2. A vector $\pi' = (\pi_0, \pi_1, ..., \pi_{k-1})$ is said to be a stationary (limit) distribution of Markov chain if

a) $\pi_i \geq 0$ and $\sum_{i=1}^{k-1} \pi_i = 1$; and

b) $\pi' P = \pi'$, i.e. $\sum_{j=0}^{k-1} P_{ij} \pi_j = \pi_i$

are satisfied, where π is a column vector.

2.2. Entropy

2.2.1. Brief information on entropy

Entropy measures the uncertainty of a collection of events, while probability measures uncertainty about the occurrence of a single event (Karlin and Taylor, 1975). In other words, entropy is a measure of the uncertainty level for a system. According to studies on entropy, entropy is only useful in cases that include an uncertainty. Accordingly, the occurrence of events having a higher probability does not provide further information, whereas the occurrence of events having a lower probability does provide more information (Özkul, 2001; Karmeshu and Pal, 2003).

For the discrete random variables, entropy is defined as follows:

Definition 3. Let X be a random variable having the values $\{x_1, x_2, ..., x_n\}$ with the corresponding probabilities:

$p(x_i) = p(X = x_i) = p_i$; $i = 1,2,...,n$.

The entropy of X is given by:

$$H(X) = H(p) = -c \sum_{i=1}^{n} p_i \log p_i$$

where c is an arbitrary positive constant and is taken as $c = 1$ when the logarithm base is 2. In addition, it is assumed that $\log 0 = 0$ in calculations (Karmeshu and Pal, 2003).

2.2.2. Maximum entropy principle of Jaynes

According to Jaynes, if a distribution is chosen such that its entropy is less than maximum entropy, this reduction in entropy might have come from some additional information used consciously or unconsciously. However, in the case in which such information is not given, it would not be right to use the distribution having less entropy. Thus, only the distribution having the maximum entropy should be used (Jaynes, 1957; Karmeshu and Pal, 2003).

2.2.3. Entropy and Markov chains

Let $i, j \in S$ be the states of Markov chain, p_i be the probability of i, and $p_i(j) = p_{ij}$ be the conditional probability of j given i. For the Markov chains, the entropy $H(S)$ is defined as

$$H(S) = -\sum_i p_i \sum_j p_i(j) \log_2 p_i(j)$$

2.3. Poisson model

2.3.1. Magnitude–frequency relation

The basic magnitude–frequency relation suggested by Gutenberg and Richter (1954) is of great importance, since it is directly related to earthquake occurrence. The relationship between magnitude and frequency is expressed as:

$LogN = a - bM$,

where N is the cumulative earthquake number, M is magnitude, and a and b are the coefficients.

In the Gutenberg–Richter function, a large value of coefficient a points to numerous small earthquakes, whereas a small value of coefficient b indicates the predominance of big earthquakes (Çobanoğlu et al., 2006).

2.3.2. Determination of seismic hazard by Poisson model

One of the commonly used models in estimating earthquake occurrence is the Poisson model. According to this model, the distribution of waiting time for another earthquake is not affected by the time after the occurrence of the previous earthquake (Öztemir et al., 2000; Çobanoğlu et al., 2006). The statistical data also show that the Poisson model is valid, especially for big earthquakes (Çobanoğlu et al., 2006).

In the equation $LogN = a - bM$, the coefficients a and b can be computed by the least squares method. According to this method, normal equations are as follows (Sayıl and Osmanşahin, 2005):

$$\sum_{i=1}^{n} LogN_i = an - b\sum_{i=1}^{n} M_i$$

$$\sum_{i=1}^{n} M_i LogN_i = a\sum_{i=1}^{n} M_i - b\sum_{i=1}^{n} M_i^2.$$

Annual average earthquake occurrence number $N(M)$ calculated according to the coefficients and earthquake magnitude is expressed as:

$$N(M) = 10^{a - \log(b \ln 10) - \log T - bM}$$

where T stands for the investigated time periods and, accordingly, seismic hazard values $R(M)$ can be determined by the following equation:

$$R(M) = 1 - e^{-N(M)T^*}$$

where T^* shows the future time portion to be used in calculating seismic hazard. The recurrence period is determined as years by using the equation:

$$Q(M) = \frac{1}{N(M)}$$ (Çobanoğlu et al., 2006).

3. Analysis of the earthquake data

In this section, our aim is to estimate the seismic hazard of Turkey by using Markov chains and the Poisson model. Similar studies were done for Japan by Nava et al. (2005) and for Turkey by Ünal and Çelebioğlu (2011). In this study, with a new approach to the zoning and an additional analysis to the previous one, we obtain some new results.

For this study, we have used the earthquake data of Turkey having magnitude $M \geq 4$ between 36°N and 42°N and 26°E and 45°E. The data belonging to the years 1901–2006 were obtained from Boğaziçi University's Kandilli Observatory Earthquake Research Institute, National Earthquake Monitoring Center. We considered the seismic zones map of Turkey in Figure 1 published by the Ministry of Public Works and Settlement in 1996, which is also approved by the Council of Ministers, and used geographic information system analysis to divide Turkey into 4 regions as follows.

On the seismic zones map of Turkey,
First-degree seismic zone is taken as region 1,
Second-degree seismic zone is taken as region 2,
Third-degree seismic zone is taken as region 3,
Fourth-degree and fifth-degree seismic zones are taken as region 4.

Here, the reason for combining the fourth-degree and fifth-degree seismic zones is to avoid a shortage of data and hence to reduce the number of zero probabilities in the transition matrix for the analysis.

3.1. Markov chain approach

Given a seismic catalog, the state of each rth region s_r can have one of 2 values during each time interval: 0 or 1, corresponding to the absence or presence of the earthquakes with magnitude $M \geq 4$ in each region,

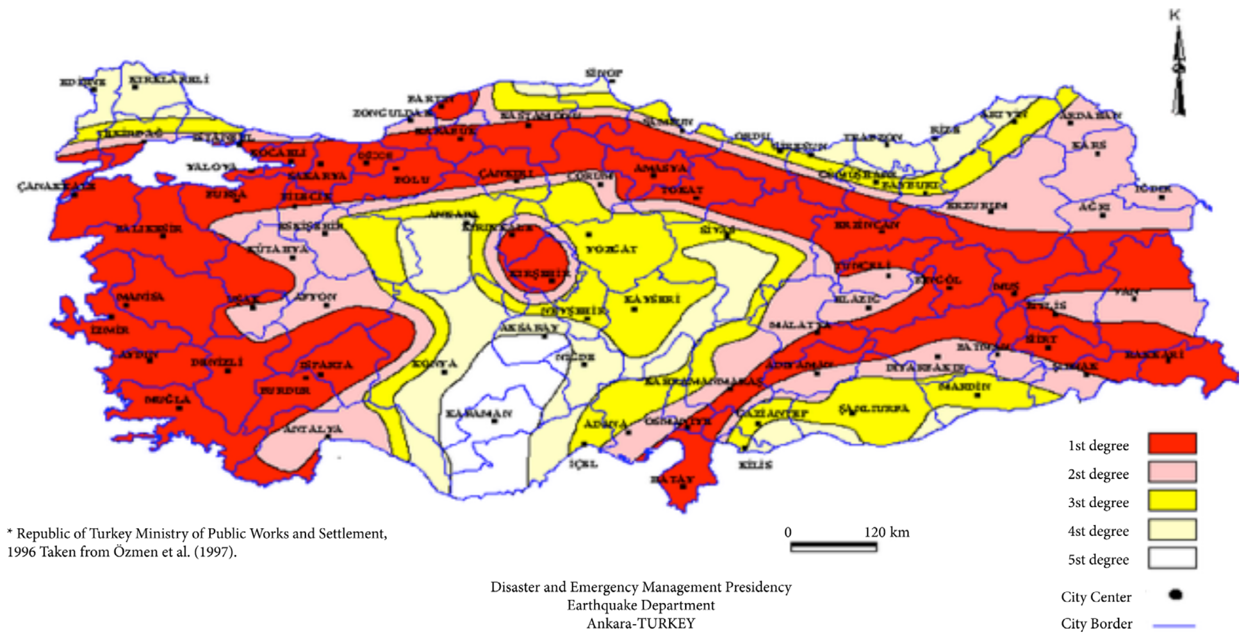

Figure 1. Seismic zones map of Turkey.

respectively. In this study, there are 4 regions and hence $2^4 = 16$ states. Therefore, the set of all possible states is S = {0,1,2...,15} .

For a given interval Δt, if there are no earthquakes in any region, we write 0000 for the state 0; if there is earthquake(s) only in region 1, we write 1000 for state 1; if there is earthquake(s) only in region 2, we write 0100 for state 2, and so on; and if there is earthquake(s) in all regions, we write 1111 for state 15. The regions and corresponding states are shown in Table 1.

Table 1. The regions and corresponding states.

State		Region			
		1	2	3	4
0	=	0	0	0	0
1	=	1	0	0	0
2	=	0	1	0	0
3	=	1	1	0	0
4	=	0	0	1	0
5	=	1	0	1	0
6	=	0	1	1	0
7	=	1	1	1	0
8	=	0	0	0	1
9	=	1	0	0	1
10	=	0	1	0	1
11	=	1	1	0	1
12	=	0	0	1	1
13	=	1	0	1	1
14	=	0	1	1	1
15	=	1	1	1	1

For the different values of the time interval Δt, transition matrices were obtained, and according to the maximum entropy principle and some conditions (Nava et al., 2005, pp. 1349–1350) it was found that the most suitable transition matrix corresponds to Δt = 0.07 years. From the data, the matrix of transition frequencies and the transition matrix obtained from this matrix are estimated as follows.

The matrix of transition frequencies:

State(i)	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	407	146	42	28	12	5	1	7	9	1	1	0	0	0	0	1
1	131	133	30	49	10	23	2	7	2	11	1	2	0	1	1	2
2	46	26	8	10	2	3	0	1	1	0	0	1	0	0	0	2
3	27	48	8	37	5	7	0	2	2	3	1	8	0	1	1	1
4	13	7	2	3	0	3	1	0	0	0	0	0	0	0	0	0
5	9	20	3	9	0	4	2	3	0	1	1	1	1	0	0	0
6	4	1	0	1	0	1	0	1	0	0	1	0	0	0	0	0
7	5	7	2	5	0	3	0	3	0	0	0	0	1	0	0	0
8	9	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0
9	4	8	1	4	0	1	0	1	0	1	0	0	0	0	0	0
10	2	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0
11	1	3	1	2	0	1	2	0	0	2	0	0	0	0	1	0
12	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
13	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
14	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
15	0	1	0	1	0	2	0	1	0	0	0	1	0	0	0	0

The transition matrix:

State (i)	0	1	2	3	4	5	6	7
0	0.6167	0.2212	0.0636	0.0424	0.0182	0.0076	0.0015	0.0106
1	0.3235	0.3284	0.0741	0.1210	0.0247	0.0568	0.0049	0.0173
2	0.4600	0.2600	0.0800	0.1000	0.0200	0.0300	0.0000	0.0100
3	0.1788	0.3179	0.0530	0.2450	0.0331	0.0464	0.0000	0.0132
4	0.4483	0.2414	0.0690	0.1034	0.0000	0.1034	0.0345	0.0000
5	0.1667	0.3704	0.0556	0.1667	0.0000	0.0741	0.0370	0.0556
6	0.4444	0.1111	0.0000	0.1111	0.0000	0.1111	0.0000	0.1111
7	0.1923	0.2692	0.0769	0.1923	0.0000	0.1154	0.0000	0.1154
8	0.6429	0.2857	0.0000	0.0714	0.0000	0.0000	0.0000	0.0000
9	0.2000	0.4000	0.0500	0.2000	0.0000	0.0500	0.0000	0.0500
10	0.4000	0.0000	0.2000	0.2000	0.0000	0.0000	0.0000	0.0000
11	0.0769	0.2308	0.0769	0.1538	0.0000	0.0769	0.1538	0.0000
12	0.0000	0.0000	0.5000	0.0000	0.0000	0.5000	0.0000	0.0000
13	0.0000	0.0000	0.5000	0.0000	0.0000	0.0000	0.5000	0.0000
14	0.3333	0.3333	0.0000	0.3333	0.0000	0.0000	0.0000	0.0000
15	0.0000	0.1667	0.0000	0.1667	0.0000	0.3333	0.0000	0.1667

State (i)	8	9	10	11	12	13	14	15
0	0.0136	0.0015	0.0015	0.0000	0.0000	0.0000	0.0000	0.0015
1	0.0049	0.0272	0.0025	0.0049	0.0000	0.0025	0.0025	0.0049
2	0.0100	0.0000	0.0000	0.0100	0.0000	0.0000	0.0000	0.0200
3	0.0132	0.0199	0.0066	0.0530	0.0000	0.0066	0.0066	0.0066
4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5	0.0000	0.0185	0.0185	0.0185	0.0185	0.0000	0.0000	0.0000
6	0.0000	0.0000	0.1111	0.0000	0.0000	0.0000	0.0000	0.0000
7	0.0000	0.0000	0.0000	0.0000	0.0385	0.0000	0.0000	0.0000
8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9	0.0000	0.0500	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
10	0.0000	0.2000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
11	0.0000	0.1538	0.0000	0.0000	0.0000	0.0000	0.0769	0.0000
12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
15	0.0000	0.0000	0.0000	0.1667	0.0000	0.0000	0.0000	0.0000

Each value in the transition matrix indicates the transition probabilities between the states. For example, using 1901 as the beginning year, it can be seen that in the case in which earthquakes only in region 1 having magnitude $M \geq 4$ occur(s) in any period (with length $\Delta t = 0.07$ years), the probability of earthquake occurrences only in region 2 having magnitude $M \geq 4$ in the next period (with length $\Delta t = 0.07$ years) is about 7.4%. From states 0 through 11, earthquakes either occur only in region 1, or they do not occur with the highest probability. Especially in region 1, a period having earthquake(s) follows a second period having earthquake(s) with the highest probability.

3.1.1. Chi-square analysis

Chi-square analysis is conducted to test whether or not the estimated transition matrix fits the data by simulation study. In the simulation study we have the same total frequency with observed frequencies, and the expected frequencies are obtained as follows:

Expected frequencies

State (i)	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	401	149	39	30	11	9	1	6	10	3	0	0	0	0	0	1
1	127	136	31	51	8	23	5	6	2	7	4	1	0	1	0	3
2	45	22	8	14	2	3	0	1	3	0	0	0	0	0	0	2
3	30	40	7	41	6	10	0	4	1	1	1	8	0	1	1	0
4	12	8	3	3	0	2	1	0	0	0	0	0	0	0	0	0
5	6	19	1	13	0	1	2	6	0	0	1	2	3	0	0	0
6	4	0	0	2	1	1	0	0	0	1	0	0	0	0	0	0
7	2	7	4	6	0	4	0	2	0	0	0	0	1	0	0	0
8	11	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0
9	5	9	1	4	0	0	0	1	0	0	0	0	0	0	0	0
10	2	0	0	1	0	0	0	0	2	0	0	0	0	0	0	0
11	2	3	0	4	0	1	2	0	0	1	0	0	0	0	0	0
12	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
13	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
14	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0
15	0	2	0	0	0	2	0	0	0	0	0	2	0	0	0	0

Observed frequencies

State (i)	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	407	146	42	28	12	5	1	7	9	1	1	0	0	0	0	1
1	131	133	30	49	10	23	2	7	2	11	1	2	0	1	1	2
2	46	26	8	10	2	3	0	1	1	0	0	1	0	0	0	2
3	27	48	8	37	5	7	0	2	2	3	1	8	0	1	1	1
4	13	7	2	3	0	3	1	0	0	0	0	0	0	0	0	0
5	9	20	3	9	0	4	2	3	0	1	1	1	1	0	0	0
6	4	1	0	1	0	1	0	1	0	0	1	0	0	0	0	0
7	5	7	2	5	0	3	0	3	0	0	0	0	1	0	0	0
8	9	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0
9	4	8	1	4	0	1	0	1	0	1	0	0	0	0	0	0
10	2	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0
11	1	3	1	2	0	1	2	0	0	2	0	0	0	0	1	0
12	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
13	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
14	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
15	0	1	0	1	0	2	0	1	0	0	0	1	0	0	0	0

In the goodness-of-fit test for the transition matrix, we will be interested in testing:

H_0: Estimated transition matrix fits the data, versus

H_1: Estimated transition matrix does not fit the data.

We have the results for the test as follows.

The calculated value of the test statistics is $\chi^2_{cal} = 17.93$ and under the significance level of 0.05, the critical value is $\chi^2_{44,0.05} \cong 55.76$. Since $\chi^2_{cal} < \chi^2_{44,0.05}$, the null hypothesis cannot be rejected.

Moreover, by comparison of the observed and expected frequencies, we conclude that we have an 85.21% aftcast (forecast of data already used to evaluate the hazard) success rate in the average for the entire catalog. In Ünal and Çelebioğlu's study (2011), the aftcast success rate was 81.1%.

3.1.2. Regional transition probabilities

From the transition matrix, it can be found that the conditional probabilities of an earthquake occurrence in region L (L = 1,2,3,4), given that the system is in state i (i = 0,1,…,15), (Nava et al., 2005, pp. 1355–1359) are as follows:

State(i)	1	2	3	4
0	0.284848	0.121212	0.039394	0.018182
1	0.562963	0.232099	0.113580	0.049383
2	0.430000	0.220000	0.080000	0.040000
3	0.708609	0.384106	0.112583	0.112583
4	0.448276	0.206897	0.137931	0.000000
5	0.703704	0.351852	0.185185	0.074074
6	0.444444	0.333333	0.222222	0.111111
7	0.692308	0.384615	0.269231	0.038462
8	0.357143	0.071429	0.000000	0.000000
9	0.750000	0.300000	0.100000	0.050000
10	0.400000	0.400000	0.000000	0.200000
11	0.615385	0.461538	0.307692	0.230769
12	0.500000	0.500000	0.500000	0.000000
13	0.000000	1.000000	0.500000	0.000000
14	0.666667	0.333333	0.000000	0.000000
15	1.000000	0.500000	0.500000	0.166667

From the above matrix, for example, looking at line 1, it can be seen that in the case in which no earthquakes having magnitude $M \geq 4$ occur in any period (with length Δt = 0.07 years), the probabilities of earthquake occurrences in each region for the next period are low. Furthermore, the aftcasts of regional activity have a 92.61% success rate on average; the highest aftcast has about a 95.72% success rate.

3.1.3. Limit distribution

The limit distribution of the Markov chain is found as follows:

$$\pi' = (0.4391\ 0.2703\ 0.0667\ 0.1016\ 0.0194\ 0.0361\ 0.0060\ 0.0174\ 0.0093\ 0.0134\ 0.0033\ 0.0087\ 0.0013\ 0.0013\ 0.0020\ 0.0040)$$

The WinQSB software package states that the system can reach its steady state after 16 periods (a period in excess of approximately 1 year, 16×0.07 = 1.12 years) on the average. This limit distribution can be interpreted as follows: in the long run, there will be no earthquakes in all the regions 43.9% of the times, there will be earthquake(s) only in region 1 at 27.0% of the time, only in region 2 at 6.7% of the time, only in region 3 at 1.9% of the time, only in region 4 at 0.9% of the time, only in regions 1 and 2 at 10.2% of the times…, and there will be earthquake(s) in all the regions 0.4% of the time.

The ratio $\pi(k)/\pi(j)$obtained from the limit distribution can be interpreted as the expected number of transitions to k between 2 transitions to j for the Markov chains (Çınlar, 1975). Under this interpretation, we can give the following matrix:

State(i)	0	1	2	3	4	5	6	7
0	1.0000	1.6247	6.5831	4.3217	22.6860	12.1720	72.9940	25.3060
1	0.6155	1.0000	4.0519	2.6600	13.9630	7.4921	44.9280	15.576
2	0.1519	0.2468	1.0000	0.6565	3.4461	1.8490	11.0880	3.8441
3	0.2314	0.3759	1.5233	1.0000	5.2493	2.8166	16.8900	5.8555
4	0.0441	0.0716	0.2902	0.1905	1.0000	0.5366	3.2176	1.1155
5	0.0822	0.1335	0.5408	0.3550	1.8637	1.0000	5.9967	2.0789
6	0.0137	0.0223	0.0902	0.0592	0.3108	0.1668	1.0000	0.3467
7	0.0395	0.0642	0.2601	0.1708	0.8965	0.4810	2.8845	1.0000
8	0.0213	0.0345	0.1400	0.0919	0.4823	0.2588	1.5519	0.5380
9	0.0305	0.0495	0.2005	0.1316	0.6908	0.3707	2.2227	0.7706
10	0.0076	0.0124	0.0501	0.0329	0.1727	0.0926	0.5555	0.1926
11	0.0199	0.0323	0.1307	0.0858	0.4505	0.2417	1.4496	0.5026
12	0.0030	0.0049	0.0200	0.0131	0.0690	0.0370	0.2221	0.0770
13	0.0031	0.0050	0.0201	0.0132	0.0692	0.0371	0.2227	0.0772
14	0.0046	0.0074	0.0301	0.0198	0.1039	0.0557	0.3343	0.1159
15	0.0091	0.0148	0.0601	0.0394	0.2070	0.1111	0.6661	0.2309

8	9	10	11	12	13	14	15
47.0360	32.8400	131.4000	50.3530	328.6900	327.7100	218.3600	109.5900
28.9510	20.2130	80.8760	30.9930	202.3100	201.7100	134.4000	67.4530
7.1450	4.9885	19.9600	7.6489	49.9300	49.7810	33.1710	16.6470
10.8840	7.5987	30.4040	11.6510	76.0550	75.8280	50.5270	25.3580
2.0734	1.4476	5.7920	2.2196	14.4890	14.4460	9.6256	4.8308
3.8642	2.6979	10.7950	4.1367	27.0030	26.9220	17.9390	9.0032
0.6444	0.4499	1.8001	0.6898	0.0005	4.4896	2.9915	1.5014
1.8587	1.2977	5.1924	1.9898	12.9890	12.9500	8.6290	4.3307
1.0000	0.6982	2.7935	1.0705	6.9880	6.9672	4.6425	2.3299
1.4323	1.0000	4.0012	1.5333	10.0090	9.9791	6.6494	3.3372
0.3580	0.2499	1.0000	0.3832	2.5015	2.4940	1.6619	0.8340
0.9341	0.6522	2.6095	1.0000	6.5277	6.5082	4.3366	2.1764
0.1431	0.0999	0.3998	0.1532	1.0000	0.9970	0.6643	0.3334
0.1435	0.1002	0.4010	0.1537	1.0030	1.0000	0.6663	0.3344
0.2154	0.1504	0.6017	0.2306	1.5052	1.5007	1.0000	0.5019
0.4292	0.2997	1.1990	0.4595	2.9993	2.9903	1.9925	1.0000

For example, the element in the fourth row and ninth column of this matrix can be interpreted as follows: between the 2 earthquakes that occurred only in region 4, it is expected that approximately 11 earthquakes occurred only in regions 1 and 2, simultaneously.

3.1.4. Estimated distributions of earthquakes in Turkey for the future

In this section, using 2006 as the beginning year, we have tried to predict the earthquakes in Turkey in the future years.

Let n be the number of periods after the year 2006, such that 2006 + 0.07 *year* × n The distribution of earthquakes in period n, π'_n, is thus given by:

$$\pi'_n = \pi'_0 P^n;\ n = 1,2,\dots,$$

where π'_0 is the initial distribution, and it can be taken as follows:

$$\pi'_0 = \begin{pmatrix} 0.0625 & 0.0625 & 0.0625 & 0.0625 & 0.0625 & 0.0625 & 0.0625 & 0.0625 \\ 0.0625 & 0.0625 & 0.0625 & 0.0625 & 0.0625 & 0.0625 & 0.0625 & 0.0625 \end{pmatrix}$$

The estimated distributions of earthquakes in Turkey for the next 5 periods from the beginning of 2006 are given in Table 2.

According to the estimation of the first period, between 1 January 2006 and 26 January 2006 ($\Delta t = 0.07$ years, about 26 days), the probability that there are no earthquakes having magnitude $M \geq 4$ in any region is 28%, earthquake(s) only in region 1 is 22.1%…and earthquake(s) in all regions is 0.2%. As the number of periods increases, the single region-specific probabilities, except for region 2, increase; the probabilities for double and triple combinations of regions decrease and the probability for all regions increases for the first period and then decreases; and all of these probabilities become rapidly stationary; i.e. approximate to the steady-state distribution.

3.2. Poisson model

3.2.1. Evaluation of the magnitude–frequency relation

The magnitude distributions of earthquakes that occurred in each region are given in Tables 3–6. The scatterplots and mathematical models of the magnitude–frequency relations obtained from the distributions for each region are given in Figures 2–5.

From the figures it is easily seen that the highest determination coefficient emerges in region 1 (98.8%) and the lowest emerges in region 3 (95.2%). In all of the regions, the changes in frequencies of earthquakes are explained by the changes in their magnitudes and only a portion of less than 5% remain unexplained. The orders of determination coefficients of the regions are consistent with the complexity of their seismic structures; that is, region 1 is connected to almost every side of the country and regions 3 and 2 are within the interior of the country. Statements similar to those for region 1 can be used to describe region 4. At the same time, in the 3 models (other than that for the first region), the frequencies of extreme magnitudes, i.e. the small and large magnitudes, are underestimated a bit and the frequencies in the middle are overestimated a bit, while the model for the first region estimates the frequencies almost on the regression line.

3.2.2. Determination of seismic hazard and recurrence periods by Poisson models

Calculated seismic hazard values and recurrence periods for each region are shown in Tables 7–10.

For example, to interpret Table 7, in the first region, the occurrence probability of an earthquake having a magnitude of 5.0 in 10 years is 100.00%. Moreover, the recurrence period for an earthquake having a magnitude of 5.0 was found to be 0.58 years. This can be interpreted such that there are approximately 2 earthquakes of magnitude 5.0 per year in the first region. As expected, in all regions as the magnitude increases, the recurrence period also increases. The recurrence period attains its maximum in the region 3 for the magnitude 7.5, for example.

When the tables are compared to each other regarding recurrence periods, the recurrence periods in region 2 are 4 to 5.4 times longer than those in region 1. Similarly, the recurrence periods in region 3 are 13.5 to 17.4 times longer, and the recurrence periods in region 4 are 6.2 to 26.2 times longer than those in region 1. The recurrence periods in region 3 are 3.4 to 5.4 times longer than those in region 2. For magnitudes 4.0–5.0, the recurrence periods in region 3 are at most 1.9 times shorter, and for magnitudes 5.5–7.5, are at most 2.8 times longer than those in region 4.

4. Conclusions

In the Erzincan earthquake (26 December 1939), one of the largest earthquakes of the 20th century, and in the Marmara earthquake (17 August 1999), thousands of people died and tens of thousands were wounded; additionally, hundreds of thousands of buildings collapsed. These experiences indicate that we may face these types of destructive earthquakes in the future. For this reason, with some statistical analysis and predictions done in this study, we tried to show that the casualties and damage that occurred in the results of those earthquakes in Turkey, an earthquake zone, could have been prevented to some extent. Furthermore, in the Van earthquake (23 October 2013), great destruction and casualties demonstrated once again the importance of such studies.

Compared with the previous study (Ünal and Çelebioğlu, 2011), in this study a different approach to zoning made it easier to see that the earthquakes occurring in Turkey can be modeled more successfully by a Markov

Table 2. Estimated distributions for the first 5 periods from 1 January 2006.

2006 + 0.07 years	0.280	0.221	0.112	0.138	0.006	0.094	0.046	0.034	0.003	0.029	0.009	0.016	0.004	0.001	0.005	0.002
2006 + 0.14 years	0.380	0.277	0.065	0.124	0.017	0.046	0.008	0.026	0.008	0.017	0.009	0.012	0.003	0.001	0.003	0.005
2006 + 0.21 years	0.414	0.274	0.068	0.111	0.019	0.040	0.007	0.019	0.009	0.016	0.004	0.010	0.002	0.0015	0.002	0.004
2006 + 0.28 years	0.429	0.272	0.067	0.105	0.019	0.038	0.006	0.018	0.009	0.014	0.004	0.009	0.001	0.001	0.002	0.004
2006 + 0.35 years	0.435	0.271	0.067	0.103	0.019	0.037	0.006	0.018	0.009	0.014	0.003	0.009	0.001	0.001	0.002	0.004

Table 3. Distribution of magnitudes of earthquakes that occurred in the first region.

Magnitude	Frequency (N)	Cumulative frequency (Ni)
4.0–4.4	874	2059
4.5–4.9	711	1185
5.0–5.4	298	474
5.5–5.9	127	176
6.0–6.4	35	49
6.5–6.9	11	14
7.0–7.4	1	3
7.5+	2	2

Table 4. Distribution of magnitudes of earthquakes that occurred in the second region.

Magnitude	Frequency (N)	Cumulative frequency (Ni)
4.0–4.4	200	492
4.5–4.9	169	292
5.0–5.4	71	123
5.5–5.9	39	52
6.0–6.4	10	13
6.5–6.9	3	3

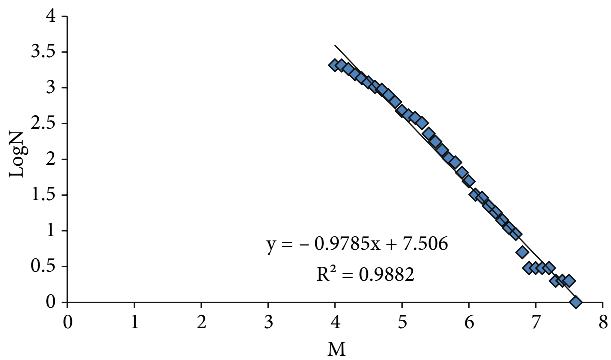

Figure 2. Magnitude–frequency relationship for the first region.

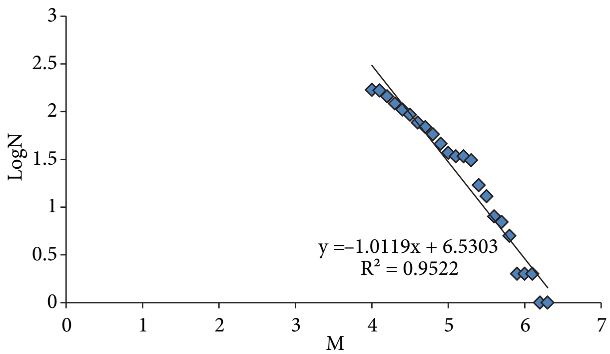

Figure 4. Magnitude–frequency relationship for the third region.

Table 5. Distribution of magnitudes of earthquakes that occurred in the third region.

Magnitude	Frequency (N)	Cumulative frequency (Ni)
4.0–4.4	76	169
4.5–4.9	56	93
5.0–5.4	24	37
5.5–5.9	11	13
6.0–6.4	2	2

Table 6. Distribution of magnitudes of earthquakes that occurred in the fourth region.

Magnitude	Frequency (N)	Cumulative frequency (Ni)
4.0–4.4	49	95
4.5–4.9	26	46
5.0–5.4	8	20
5.5–5.9	10	12
6.0–6.4	2	2

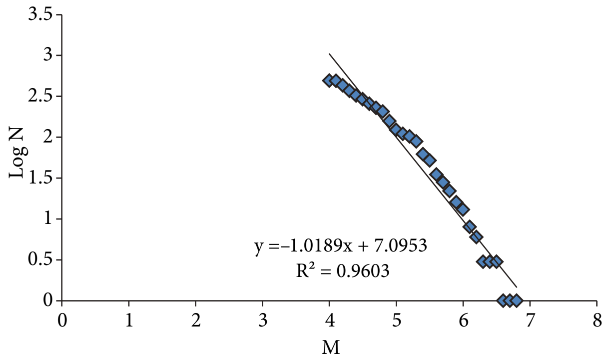

Figure 3. Magnitude–frequency relationship for the second region.

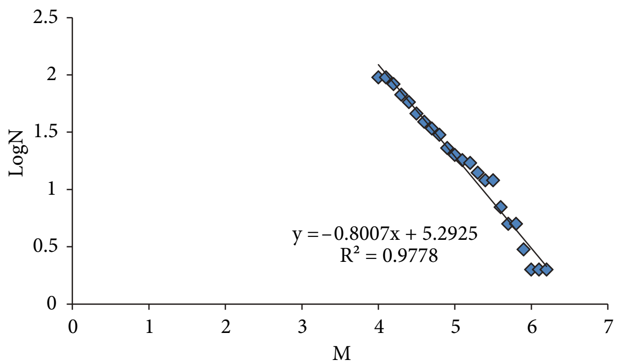

Figure 5. Magnitude–frequency relationship for the fourth region.

Table 7. Obtained seismic hazard and recurrence periods for the first region.

M	Seismic hazard (%) (for the first region)										Recurrence period Q (years)
	Years										
	10	20	30	40	50	60	70	80	90	100	
4.0	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	0.06
4.5	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	0.19
5.0	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	0.58
5.5	99.64	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	1.78
6.0	83.86	97.40	99.58	99.93	99.99	100.00	100.00	100.00	100.00	100.00	5.48
6.5	44.64	69.35	83.03	90.61	94.80	97.12	98.41	99.12	99.51	99.73	16.9
7.0	17.44	31.84	43.73	53.54	61.65	68.34	73.86	78.42	82.18	85.29	52.2
7.5	6.02	11.68	17.00	22.00	26.70	31.12	35.27	39.17	42.83	46.27	161

Table 8. Obtained seismic hazard and recurrence periods for the second region.

M	Seismic hazard (%) (for the second region)										Recurrence period Q (years)
	Years										
	10	20	30	40	50	60	70	80	90	100	
4.0	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	0.24
4.5	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	0.76
5.0	98.29	99.97	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	2.46
5.5	71.59	91.93	97.71	99.35	99.81	99.95	99.99	100.00	100.00	100.00	7.95
6.0	32.25	54.10	68.90	78.93	85.73	90.33	93.45	95.56	96.99	97.96	25.7
6.5	11.35	21.41	30.33	38.24	45.25	51.46	56.97	61.86	66.19	70.02	83
7.0	3.66	7.18	10.58	13.85	17.01	20.04	22.97	25.79	28.50	31.12	268
7.5	1.15	2.28	3.40	4.51	5.60	6.69	7.76	8.81	9.86	10.89	867

Table 9. Obtained seismic hazard and recurrence periods for the third region.

M	Seismic hazard (%) (for the third region)										Recurrence period Q (years)
	Years										
	10	20	30	40	50	60	70	80	90	100	
4.0	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	0.81
4.5	97.92	99.96	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	2.58
5.0	70.14	91.08	97.34	99.20	99.76	99.93	99.98	99.99	100.00	100.00	8.27
5.5	31.41	52.95	67.73	77.86	84.81	89.58	92.86	95.10	96.64	97.69	26.5
6.0	11.09	20.96	29.73	37.52	44.45	50.61	56.09	60.96	65.29	69.14	85
6.5	3.60	7.07	10.42	13.65	16.76	19.75	22.64	25.43	28.12	30.70	273
7.0	1.14	2.26	3.37	4.47	5.56	6.63	7.70	8.75	9.78	10.81	874
7.5	0.36	0.71	1.06	1.42	1.77	2.12	2.47	2.81	3.16	3.51	2802

Table 10. Obtained seismic hazard and recurrence periods for the fourth region.

M	Seismic hazard (%) (for the fourth region)										Recurrence period Q (years)
	Years										
	10	20	30	40	50	60	70	80	90	100	
4.0	99.83	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	1.57
4.5	92.00	99.36	99.95	100.00	100.00	100.00	100.00	100.00	100.00	100.00	3.96
5.0	63.39	86.60	95.09	98.20	99.34	99.76	99.91	99.97	99.99	100.00	9.95
5.5	32.95	55.04	69.86	79.79	86.45	90.91	93.91	95.92	97.26	98.16	25
6.0	14.70	27.24	37.94	47.06	54.84	61.48	67.15	71.98	76.10	79.61	62.9
6.5	6.08	11.80	17.16	22.20	26.93	31.38	35.55	39.47	43.15	46.61	159
7.0	2.48	4.91	7.27	9.57	11.82	14.01	16.15	18.23	20.26	22.25	397
7.5	1.00	1.98	2.96	3.92	4.88	5.83	6.77	7.69	8.61	9.52	999

chain model (an 85.21% aftcast success rate on average for the new model against a 81.1% aftcast success rate for the entire catalog for the former model). In addition to this result, the recurrence periods of earthquakes of different magnitudes in each region were estimated by the Poisson model, which is particularly valid for big earthquakes. Thus, the recurrence periods of earthquakes, especially the big ones, can be known with great accuracy and necessary precautions can be taken. As was also emphasized by Özmen (2012), in Turkey, there are overexpected losses of earthquakes whose most important role are played by the seismic zoning maps, the determining rules of construction for each zone, and nonconforming constructions in accordance with regulations for ground surveys. Therefore, we think that seismic hazard assessments can be much improved by the information based on newly developed seismic zoning maps. We look forward to these maps.

Acknowledgment

We want to thank the anonymous referee for helpful comments that improved this paper.

References

Altınok Y (1988). Seismic risk estimation of the North Anatolian Faulth Zone using semi-Markov model. Jeofizik 2: 44–58 (article in Turkish with English abstract).

Altınok Y (1991). Evaluation of earthquake risk in West Anatolia by semi-Markov model. Jeofizik 5: 135–140 (article in Turkish with English abstract).

Altınok Y, Kolçak D (1999). An application of the semi-Markov model for earthquake occurrences in North Anatolia, Turkey. J Balkan Geophys Soc 2: 90–99.

Bağcı G (1996). Earthquake occurrences in Western Anatolia by Markov model. Jeofizik 10: 67–75.

Ching W, Ng MK (2006). Markov Chains: Models, Algorithms and Applications. New York, NY, USA: Springer, pp. 1–15.

Çınlar E (1975). Introduction to Stochastic Processes. Englewood Cliffs, NJ, USA: Prentice Hall, pp. 106–277.

Çobanoğlu İ, Bozdağ Ş, Dinçer İ, Erol H (2006). Statistical approaches to estimating the recurrence of earthquakes in the Eastern Mediterranean Region. İstanbul Üniv Müh Fak Yerbilimleri Dergisi 19: 91–100.

Gutenberg B, Richter CF (1954). Earthquake magnitude, intensity, energy and acceleration. Bull Seism Soc Am 32: 163–191.

Jaynes ET (1957). Information theory and statistical mechanics. Phys Rev 106: 620–630.

Karlin S, Taylor HE (1975). A First Course in Stochastic Processes. 2nd ed. San Diego, CA, USA: Academic Press, pp. 495–502.

Karmeshu, Pal NR (2003). Uncertainty, Entropy and Maximum Entropy Principle – An Overview in Entropy Measures, Maximum Entropy Principle and Emerging Applications. Berlin, Germany: Springer-Verlag.

Kasap R, Gürlen Ü (2003). Obtaining the return period of earthquake magnitudes: as an example Marmara Region. Doğuş Üniversitesi Dergisi 4: 157–166 (article in Turkish with English abstract).

Lee WHK, Kanomori H, Jennings PC, Kisslinger C (2003). International Handbook of Earthquake and Engineering Seismology, Part B. San Diego, CA, USA: Academic Press, pp. 1234–1237.

Nava FA, Herrera C, Frez J, Glowacka E (2005). Seismic hazard evaluation using Markov Chains: application to the Japan Area. Pure Appl Geophys 162: 1347–1366.

Özkul S (2001). Entropy-based assessment of water quality monitoring networks. Turk J Engin Environ Sci 25: 435–452.

Özmen B (2012). The historical development of seismic zoning maps of Turkey. Geol Bull Turk 55: 43–55.

Özmen B, Nurlu M, Güler H (1997). Analysis of Earthquake Zones with Geographical Information System. Ankara, Turkey: Republic of Turkey Ministry of Public Works and Settlement, Disaster and Emergency Management Presidency, Earthquake Department.

Öztemir F, Necioğlu A, Bağcı G (2000). Focal mechanism solutions and seismicity for Antakya Region and its surrounding. Jeofizik 14: 87–102 (article in Turkish with English abstract).

Pınar R, Akçığ Z, Demirel F (1989). The investigation of Western Anatolia seismicity by the Markov method. Jeofizik 3: 56–66 (article in Turkish with English abstract).

Sayıl N, Osmanşahin İ (2005). Investigation of seismicity of the Marmara Region. In: Proceedings of the Earthquake Symposium, 23–25 March 2005, Kocaeli, Turkey, pp. 1417–1426.

Ulutaş E, Özer FM (2000). Seismic hazard estimation of Çukurova Region by using Markov model. Jeofizik 14: 103–112 (article in Turkish with English abstract).

Ünal S, Çelebioğlu S (2011). A Markov chain modelling of the earthquakes occurring in Turkey. GU J Sci 24: 263–274.

Yücemen MS, Akkaya A (1995). Stochastic models for the estimation of seismic hazard and their comparison. In: Proceedings of the 3rd Earthquake Engineering Conference, İstanbul, Turkey, pp. 466–477.