Turkish Journal of Science & Technology
Volume 8(1), 55-61, 2013

Estimates of Earthquake With Markov Models in the East Anatolian Fault Zone

Adem DOĞANER and Sinan ÇALIK
Firat University, Department of Statistics, Elazig, Turkey
adoganer@firat.edu.tr

(Geliş/Received: 03.08.2012; Kabul/Accepted: 26.09.200*)

Abstract

Stochastic methods are often used for estimates of future for environmental events realized. Markov chains are one of these methods. Markov chains can applied to estimates of the earthquake but it provides only information about a single parameter. In this article, epicenter of may occurring earthquake were estimated on East Anatolian Fault Zone. Epicenter parameter and times of seismic inactivity were associated for estimates of earthquake. Probabilities were provided by Hidden Markov Model with forwards algorithm and these probabilities were converted to state sequence. Transition probabilities of this sequence were examined with Discrete Time Markov Chains. As a result new earthquake were expected in 4th state (Sincik- Lake Hazar) with 0.73 probabilities. Results and method were discussed.

Key words: Hidden Markov model, Epicenter, East Anatolian Fault Zone

Markov Modelleri ile Doğu Anadolu Fay Zonu'ndaki Deprem Tahminleri

Özet

Stokastik yöntemler sıklıkla geleceğe yönelik çevresel olayların gerçekleşme tahminleri için kullanılabilmektedir. Markov zincirleri bu yöntemlerden biridir. Markov zincirleri deprem tahminlerinde uygulanabilmektedir, fakat bu yöntem sadece bir parametreye ilişkin bilgi sağlamaktadır. Bu çalışmada, Doğu Anadolu Fay Zonu üzerinde gerçekleşecek depremin dış merkezi tahmin edilmiştir. Olasılıklar, saklı Markov modeli ve ileri algoritması ile elde edilmiştir. Bu olasılıklar durum zincirine dönüştürülmüştür. Geçiş olasılıkları, kesikli parametreli Markov zincirleri ile incelenmiştir. Sonuç olarak yeni deprem 0.73 olasılıkla 4.durum olarak belirtilen Sincik- Hazar Gölü segmenti üzerinde gerçekleşebileceği tahmin edilmiştir. Sonuçlar ve metot tartışılmıştır.

Anahtar kelimeler: Saklı Markov modeli, Dış merkez, Doğu Anadolu Fay Zonu

1.Introduction

Earthquakes are seismic activities that occurred in the region caused great destruction and loss. Estimates of earthquake are a great importance for to minimize losses, but estimates of earthquakes precisely is not possible. Some statistical methods can be used for estimate of earthquake. Stochastic processes are one of these methods and were studied on these methods. Poisson processes and point processes were used by Ogata [1] and by Ferraes [2] for estimates of earthquake occurrences. Markov Chain can be used such as Poisson process for estimate of earthquake. Application of Markov Chains to geologic processes were discussed by Harbaugh and Bonham-Carter [3]. Markov Chains were applied by Nava et al. [4] to Japan region for the evolution of seismic risk. Besides Heng Tsai [5] used Markov Chains for estimates of earthquake recurrence. But Markov Chains only gives transition probabilities of a single parameter. Hidden Markov model have two processes known as sequence of unobserved state and sequence of observation. Ebel et al.[6] used Hidden Markov Model for earthquake forecasting in California. East Anatolian Fault System, which produce earthquake is an active fault system. Risk assessment for this fault is important.

In this study, probabilities of earthquake occurrence in epicenters of East Anatolian Fault

Zone were estimated by using hidden Markov model and forward algorithm. Transition probabilities of epicenters were estimated by using discrete parameter Markov chains. Relationship was obtained between epicenter and times of seismic inactivity.

2.Material and Methods

2.1. Data

East Anatolian Fault Zone, known as Türkoğlu-Antakya segment, Gölbaşı-Türkoğlu Segment, Çelikhan-Erkenek Segment, Lake Hazar-Sincik Segment, Palu- Lake Hazar Segment and Karlıova-Bingöl Segment, is composed six segments. Occurred earthquakes in area radius of 20 km the East Anatolian Fault Zone were included in this study. Earthquake data were supplied from Boğaziçi University Kandilli Observatory and Earthquake Research Institute National Earthquake Monitoring Centre. Data were included $M_L \geq 4.0$ earthquakes occurred between 1975 and 2011.

In this study, data were included about known as one of earthquake parameters epicenter and times of seismic inactivity. The parameter is epicenter. Epicenters were categorized according to segments of East Anatolian Fault Systems. Epicenters were represented as $S_1 = Türkoğlu - Antakya\ segment,$ $S_2 = Gölbaşı - Türkoğlu\ segment,$ $S_3 = Çelikhan - Erkenek\ segment,$ $S_4 = Lake\ Hazar - Sincik\ segment,$ $S_5 = Palu - Lake\ Hazar\ segment$ and $S_6 = Karlıova - Bingöl\ segment.$ In this study, times of inactivity were also considered. Times of inactivity were categorized as $O_1(Time = 0 - 90\ days,$
$O_2(Time = 91 - 180\ days),$
$O_3(Time = 181 - 270\ days),$
$O_4(Time = 271 - 360\ days),$
$O_5(Time \geq 361\ days),$ respectively.

229 Seismic activity with $M_L \geq 4$ were occurred in fault zone between 1975 and 2011. Aftershocks occurs that depend on main shocks. Aftershocks were excluded in this study, therefore each random variables must be independent the other according to Markov property. 126 earthquakes were included in this study.

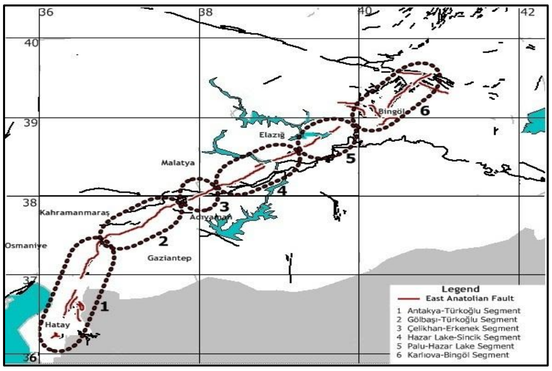


Figure 1. Study Area(East Anatolian Fault Zone and segments)

2.2. Parameter Relationship

In this study, between the parameters of earthquake statistical relationship was examined. Transitions of earthquake epicenter in East Anatolian Fault Zone were wanted to estimate. Outputs of earthquake were considered, since transitions of earthquake epicenter cannot directly observed. Time of seismic inactivity is an output of earthquake. The relationship between epicenter parameter and times of seismic inactivity were analyzed with Chi-Squared test for independence. Results of analysis were represented table 1. The relationship between epicenter parameter and times of seismic inactivity statistically were significant.

Table 1. Parameter Relationship

	X^2	p
Epicenter*Times of seismic inactivity	42.052	0.03

2.3. Markov Chains and Hidden Markov Model

2.3.1. *Theoretical Background*

Markov Chains base on, when state i at the time t were known, state j at the time $t+1$ depend on only state i at the time t. In other words, a stochastic process with the Markov property is expressed by;

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 i_0) =$$

$$P(X_{n+1} = j | X_n = i) \qquad (1)$$

This equation often is called as memorylessness [7].

Hidden Markov model is Markov Chains having additional features. Theory of Hidden Markov Model was developed by Baum et al.[8,9,10,11,12]. Hidden Markov model has a two stochastic processes. Each Hidden Markov Model is defined by states, state probabilities, transition probabilities, emission probabilities and initial probabilities [13,15].

Hidden Markov model were characterized by the following [14],

1.The N states of the Model, defined by $S = \{S_1, S_2, \dots, S_N\}$.

2. M observation symbols per state $V = \{v_1, v_2, \dots, v_m\}$.

3. The State transition probability distribution $A = \{a_{ij}\}$, where a_{ij} is the probability that the state at the time $t+1$ is S_j is given when the state at time t is S_i.

4. The observation symbol probability distribution in each state, $B = \{b_j(k)\}$, where $b_j(k)$ is the probability that symbol v_k is emitted in state S_j.

$$b_j(k) = p\{o_t = v_k | q_t = j\},$$
$$1 \leq j \leq N, \qquad 1 \leq k \leq \qquad (2)$$

Where v_k denotes the k^{th} observation symbol in the alphabet, and o_t the current parameter vector.

5. The Hidden Markov Model is the initial state distribution $\pi = p\{\pi_i\}$, where π_i is the probability that the model is in state S_i at the time $t = 0$ with

$$\pi_i = p\{q_1 = i\}\ and\ 1 \leq i \leq N$$

Discrete parameter hidden Markov model generally is expressed as $\lambda = (A, B, \pi)$ [13,14,15].

2.4. Application

In this study, 126 earthquakes and epicenter data and times of inactivity about this earthquake were examined in order to epicenter data and times of inactivity adjust to Hidden Markov Model.

First step, transition probabilities matrix of state, emission probabilities matrix and initial probabilities vector were provided about Hidden Markov model. Transition probabilities matrix of state, emission probabilities matrix and initial probabilities vector were showed in Table 2.

Transition probabilities from each state to others state were given in transition probabilities matrix. Example transition probability from State S_1 known as Türkoğlu-Antakya segment to State S_4 known as Hazar Lake-Sincik segment were determined as 0.80.

The observation symbols probabilities for each state were given in emission probabilities matrix. Example State 1 known as Türkoğlu-Antakya segment produced 0.60 probability observation symbol O_1 and 0.40 probability observation symbol O_3.

Initial state probabilities equally for each state were distributed since initial state probabilities distribution were not known. Each initial state have 0.16666667 probability.

Table 2. Hidden Markov Model Parameters.

a. Transition Probabilities Matrix

	S_1	S_2	S_3	S_4	S_5	S_6
S_1	0	0	0.20	0.80	0	0
S_2	0	0.20	0	0.40	0	0.40
S_3	0.07	0.07	0.33	0.20	0.13	0.20
S_4	0.02	0	0.10	0.54	0.15	0.19
S_5	0.08	0.04	0.08	0.16	0.52	0.12
S_6	0.04	0.07	0.07	0.33	0.16	0.33

b. Emission Probabilities Matrix

	O_1	O_2	O_3	O_4	O_5
S_1	0.60	0	0.40	0	0
S_2	0.60	0.20	0	0.20	0
S_3	0.54	0.20	0.13	0.13	0
S_4	0.77	0.13	0.02	0.06	0.02
S_5	0.54	0.42	0.04	0	0
S_6	0.92	0.08	0	0	0

c. Initial Probabilities Vector

	S_1	S_2	S_3	S_4	S_5	S_6
$\pi =$	[0.166	0.166	0.166	0.166	0.166	0.167]

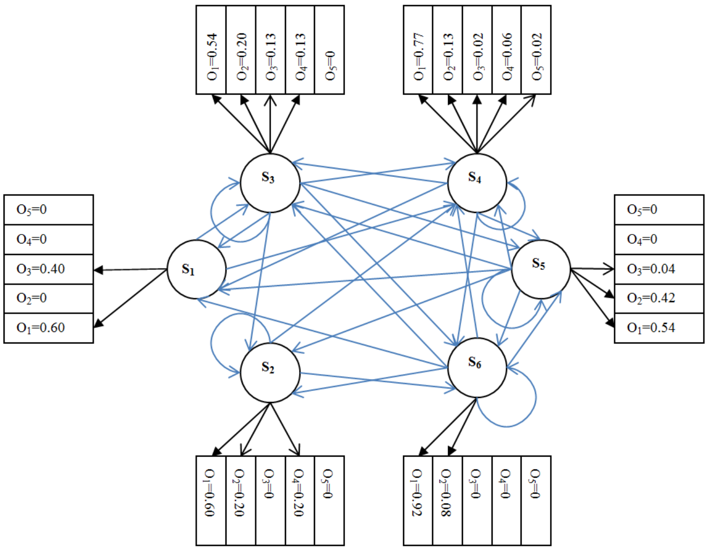


Figure 2. Emission probabilities and transitions of between states.

2.5. Forward Algorithm

2.5.1.*Theoretical Background*

The variable $\alpha_t(i)$ was defined as probability of state S_i is situated as a result of partial observation until at the time t [14].

Equation is defined as;

$$\alpha_t(i) = P(O_1, O_2, \ldots, O_t, q_t = S_i|\lambda) \qquad (3)$$

Forward algorithm contain three parts [14].

Initialization:

$$\alpha_1(i) = \pi_i b_i(O_1) \quad 1 \leq i \leq N \qquad (4)$$

Induction:

$$\alpha_{t+1}(j) = \left[\textstyle\sum_{i=1}^{N} \alpha_t(t)\alpha_{ij}\right] b_j(O_{t+1}) \qquad (5)$$

$$1 \leq t \leq T-1$$
$$1 \leq j \leq N.$$

Termination:

$$P(O|\lambda) = \textstyle\sum_{i=1}^{N} \alpha_t(i) \qquad (6)$$

the forward algorithm aim calculated of probability of realization $P(O|\lambda)$ observation sequence with λ model for $O = O_1 O_2 \ldots O_t$ observation sequence and $\lambda = \{A, B, \pi\}$ model [14].

The purpose of use of forward algorithm in this study, realization probabilities of state sequence in accordance with hidden Markov model were wanted to estimate by observation sequence. Result of forward algorithm showed in Table 3.

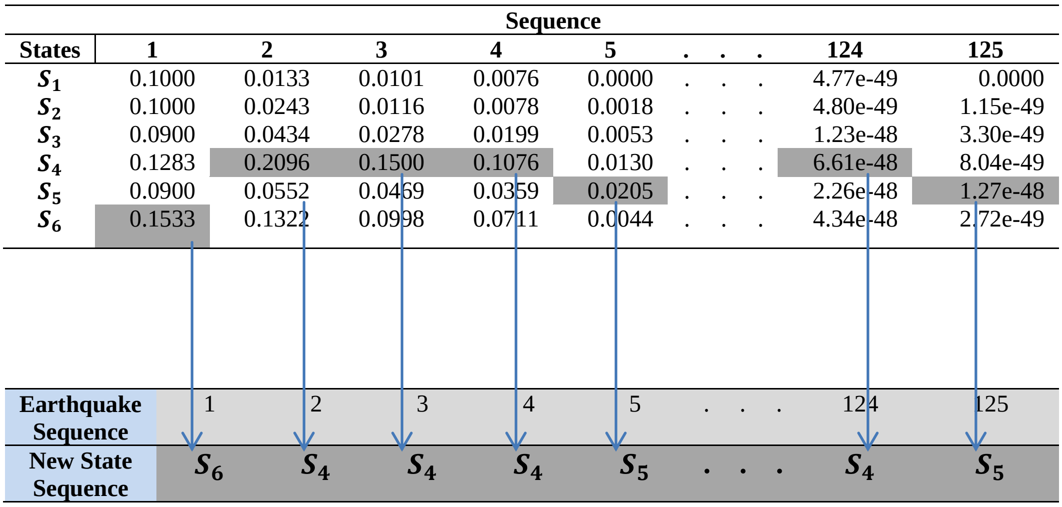
Table 3. State Probability for earthquake sequence and result of Forward algorithm

	Sequence									
States	**1**	**2**	**3**	**4**	**5**	**.**	**.**	**.**	**124**	**125**
$\boldsymbol{S_1}$	0.1000	0.0133	0.0101	0.0076	0.0000	.	.	.	4.77e-49	0.0000
$\boldsymbol{S_2}$	0.1000	0.0243	0.0116	0.0078	0.0018	.	.	.	4.80e-49	1.15e-49
$\boldsymbol{S_3}$	0.0900	0.0434	0.0278	0.0199	0.0053	.	.	.	1.23e-48	3.30e-49
$\boldsymbol{S_4}$	0.1283	0.2096	0.1500	0.1076	0.0130	.	.	.	6.61e-48	8.04e-49
$\boldsymbol{S_5}$	0.0900	0.0552	0.0469	0.0359	0.0205	.	.	.	2.26e-48	1.27e-48
$\boldsymbol{S_6}$	0.1533	0.1322	0.0998	0.0711	0.0044	.	.	.	4.34e-48	2.72e-49

2.6. Model Evolution

Hidden Markov model and forward algorithm were applied to Earthquake data until this stage of study.

At the stage of model evolution, probabilities of state by forward algorithm were examined. In a result of forward algorithm, a state having highest probability for each earthquake in sequence were included as a new epicenter of earthquake to our model. Example, State S_6 has the highest probability for first earthquake in table 3, So a new state sequence were obtained. Obtaining this new state sequence were shown figure 3.



States	Sequence 1	2	3	4	5	. . .	124	125
S_1	0.1000	0.0133	0.0101	0.0076	0.0000	. . .	4.77e-49	0.0000
S_2	0.1000	0.0243	0.0116	0.0078	0.0018	. . .	4.80e-49	1.15e-49
S_3	0.0900	0.0434	0.0278	0.0199	0.0053	. . .	1.23e-48	3.30e-49
S_4	0.1283	0.2096	0.1500	0.1076	0.0130	. . .	6.61e-48	8.04e-49
S_5	0.0900	0.0552	0.0469	0.0359	0.0205	. . .	2.26e-48	1.27e-48
S_6	0.1533	0.1322	0.0998	0.0711	0.0044	. . .	4.34e-48	2.72e-49
Earthquake Sequence	1	2	3	4	5	. . .	124	125
New State Sequence	S_6	S_4	S_4	S_4	S_5	. . .	S_4	S_5

Figure 3. Development of new state sequence

Table 3. New State Sequence

S.	State	S.	State	S.	State	S.	State	S.	State	S.	State	S.	State
1	S_6	21	S_5	41	S_4	61	S_4	81	S_4	101	S_4	121	S_4
2	S_4	22	S_4	42	S_4	62	S_4	82	S_4	102	S_4	122	S_4
3	S_4	23	S_1	43	S_4	63	S_5	83	S_5	103	S_4	123	S_4
4	S_4	24	S_4	44	S_4	64	S_4	84	S_5	104	S_5	124	S_4
5	S_5	25	S_4	45	S_4	65	S_4	85	S_5	105	S_4	125	S_5
6	S_4	26	S_6	46	S_4	66	S_4	86	S_5	106	S_4		
7	S_5	27	S_6	47	S_4	67	S_4	87	S_4	107	S_5		
8	S_4	28	S_4	48	S_4	68	S_4	88	S_4	108	S_4		
9	S_5	29	S_4	49	S_4	69	S_5	89	S_4	109	S_4		
10	S_1	30	S_1	50	S_4	70	S_4	90	S_4	110	S_4		
11	S_4	31	S_4	51	S_4	71	S_4	91	S_4	111	S_4		
12	S_4	32	S_4	52	S_4	72	S_4	92	S_4	112	S_4		
13	S_5	33	S_4	53	S_4	73	S_4	93	S_4	113	S_4		
14	S_5	34	S_4	54	S_4	74	S_4	94	S_5	114	S_4		
15	S_5	35	S_4	55	S_5	75	S_4	95	S_4	115	S_4		
16	S_1	36	S_5	56	S_4	76	S_4	96	S_6	116	S_4		
17	S_4	37	S_4	57	S_4	77	S_5	97	S_4	117	S_4		
18	S_4	38	S_4	58	S_4	78	S_5	98	S_4	118	S_4		
19	S_4	39	S_4	59	S_4	79	S_5	99	S_4	119	S_4		
20	S_1	40	S_6	60	S_5	80	S_5	100	S_4	120	S_4		

The latest stage of study, transition probabilities were estimated for new state sequence. Transition probabilities matrix were included by discrete parameter Markov Chains method. Obtaining probabilities by discrete parameter Markov Chains give to transition probabilities of between the epicenter. According to probabilities in this matrix, Earthquake may occur with 0.80 probability in state S_4 known as Hazar Lake-Sincik segment, after earthquake occurred in state S_1 known as Antakya-Türkoglu segment. Initial probabilities were estimated for new state sequence. Initial probabilities gives distribution of epicenter in new state sequence.

The product of initial probabilities vector and transition probabilities matrix gives probability distribution vector of epicenter for the next earthquake. Transition probabilities matrix, initial probabilities vector and probability distribution vector of epicenter for the next earthquake were represented P_0, π_0 and P_1 respectively. P_0, π_0 and P_1 were shown in Table 4.

Table 4. Result Table

a. Transition Probabilities Matrix.

$$\begin{array}{c} \\ S_1 \\ S_4 \\ S_5 \\ S_6 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_4 & S_5 & S_6 \end{array} \\ \begin{bmatrix} 0.00 & 0.80 & 0.20 & 0.00 \\ 0.03 & 0.76 & 0.18 & 0.03 \\ 0.09 & 0.56 & 0.35 & 0.00 \\ 0.00 & 0.80 & 0.00 & 0.20 \end{bmatrix} \end{array}$$

b. Initial Probabilities Vector

$$\begin{array}{rcccc} & S_1 & S_4 & S_5 & S_6 \\ \pi_0 = & [0.04 & 0.73 & 0.19 & 0.04] \end{array}$$

c. Next step transition matrix.

$$\begin{array}{rcccc} & S_1 & S_4 & S_5 & S_6 \\ P_1 = & [0.04 & 0.73 & 0.20 & 0.03] \end{array}$$

According to results, next earthquake was estimated to occur with 0.73 probability in State 4 known as Hazar Lake-Sincik segment.

3. Results

In study, Hidden Markov Model and forward algorithm were applied earthquake data of East Anatolian Fault Zone between 1975 and 2011. Transition probabilities matrix, initial probabilities vector, emission probabilities matrix and new state sequence were obtained. Discrete parameter Markov chains method was applied to new state sequence. Transition probabilities from each epicenter to others epicenter were calculated. Epicenter of next earthquake was estimated as a probabilistic.

According to obtained result, Earthquake may occur with 0.80 probability in State S_4 after an earthquake occurred in State S_1, with 0.76 probability in State S_4 after a earthquake occurred in State S_4, with 0.57 probability in State S_4 after a earthquake occurred in State S_5, with 0.80 probability in State S_4 after a earthquake occurred in State S_6.

As a result next earthquake were estimated to occur with 0.73 probability in State S_4. Emission probabilities matrix show probabilities of times of inactivity for each epicenter. According to this matrix, earthquake will occur with 0.54 probabilities within between 1 and 90 days after an earthquake occurred in State S_5. The latest earthquake in East Anatolian Fault Zone occurred in State S_5.

In this study, Hidden Markov model methods were shown to be available for epicenter estimates.

4. Discussion

In seismic surveys, estimates of earthquake epicenter were been point of interest. Estimates of Earthquake can be used different methods. Some statistical methods such as Gumbell distribution can be used estimates of magnitude and determination of earthquake risk.

Discrete parameter Markov Chain is a using methods in estimates of earthquake epicenter as well as geological approaches. Discrete parameter Markov Chain is successful to examine as a probabilistic transition of earthquake epicenter. But this method do not base on a relationship transition

of earthquake epicenter. This method only accept as a sequence this transitions and it gives a probability distribution depending on process. Hidden Markov model makes more significant this model. Because Hidden Markov Model consider times of seismic inactivity as well as earthquake epicenter. Times of inactivity depend on earthquake epicenter.

The new state sequence was obtained by forward algorithm and this sequence was converted to transition probabilities matrix by discrete parameter Markov Chains.

5. References

1. Ogata, Y., (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of American statistical association* **83**(401):9-27.

2. Ferraes, S.G., (1967). Test of Poisson for earthquakes in Mexico City. *J. Geophys. Res.* **72**:3741-3742.

3. Harbaugh, J.W., Bonham-Carter G (1970). Computer simulation in geology. Wiley, Newyork, 575pp.

4. Nava, F.A., Herrera, C., Frez, J., Glowacka, E., (2005). Seismic hazard evaluation using Markov chains: Application to the Japan area. *Pure applied geophysics,* **162**(2005):1347-1366.

5. Tsai, H., (2002). Estimates of earthquake recurrences in the Chiayi-Tainan area, Taiwan. *Engineering Geology,* **63**(1-2):157-168.

6. Ebel, J.E., Chambers, D.W., Kafka, A.L., Baglivo, J.A, (2007). Non-Poissonian Earthquake clustering and the hidden Markov model as bases for earthquake forecasting in California. *Seismological Research Letters***.** **78**(1):57-65.

7. Markov, A.A., (1908). Wahrscheinlichkeitsrechnung. B. G. Teubner, Leipzig, Berlin.

8. Baum, L. E., Petrie, T., (1966). Statistical Inference for Probabilistic Functions of Finite State Markov Chains. *The Annals of Mathematical Statistics,* **37** (6): 1554–1563.

9. Baum, L. E., Eagon, J. A., (1967). An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. *Bulletin of the American Mathematical Society,* **73** (3). 360-363.

10. Baum, L. E., Sell, G. R., (1968). Growth transformations for functions on manifolds. *Pacific Journal of Mathematics,* **27**(2):211–227.

11. Baum, L. E., Petrie, T., Soules, G., Weiss, N., (1970). A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains. *The Annals of Mathematical Statistics,* **41**(1): 164-171.

12. Baum, L.E., (1972). An Inequality and Associated Maximization Technique in Statistical Estimation of Probabilistic Functions of a Markov Process. *Inequalities* **3**: 1–8.

13. Kouemou, G.L., (2011). History and theoretical basics of hidden Markov models. Intech, Croatia.

14. Rabiner, L., (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceeding of IEEE.* **77**(2):257-286.

15. Dymarski, P. (*Edited by*) (2011). Hidden Markov models, theory and applications. Intech, Croatia.

16. Barkat, M., (2005). Signal detection and estimation, second edition. Artech house Boston.

17. Cappe, O., Moulines, E., Ryden, T., (2005). Inference in hidden Markov models, Springer, Newyork.

18. Durbin, R., Eddy, S., Krogh, A., Mitchison, G., (1998). Biological sequence analysis, probabilistic models of proteins and nucleic acids. Cambridge University Press, Cambridge.

19. Dynkin, E.B., (1961), Theory of Markov processes. Pergamon Press, London.

20. Elliott, R.J., Aggoun, L., Moore. J.B., (2008). Hidden Markov models estimation and control. Stochastic Modelling and applied probability, Springer.

21. Erdik, M., Doyuran, V., Akkas, N., Gulkan, P., (1985). A probabilistic assessment of the seismic hazard in Turkey. *Tectonophysics,* **117**(3-4):295-344.

22. Fraser, A.M., (2008). Hiddden Markov models and dynamical systems. Society for industrial and applied mathematics, Philadelphia.

23. Karakaisis, G.F., (1994). Long term earthquake prediction along the north and east Anatolian fault zones based on the time and magnitude predictable model. *Geophysical journal international***,** **116**(1):198-204.

24. Kemeny, J.G., Snell. J.L., (1976). Finite Markov chains. Springer-Verlag, Newyork.

25. Nalbant, S.S., McCloskey, J., Steacy, S., Barka, A.A., (2002). Stress accumulation and increased seismic risk in eastern Turkey. *Eart and Planetary science letters,* **195**(3-4):291-298.