

Shallow seismicity in the area of Greece: its character as seen by means of a stochastic model

L. C. Polimenakos

PANGAEA Consulting Engineers, Kifissias Avenue 131, GR 11524 Athens, Greece

Received 20 September 1994 - Accepted 20 June 1995 - Communicated by D. Sornette

Abstract. Occurrence of successive earthquake events in space is analysed by means of semi-stochastic processes. The analysis employs earthquake events with $M \geq 5.2$ from the area of Greece and its surroundings ($18-31^\circ$ E, $34-43^\circ$ N) for the time interval 1911-1985. The sequence of earthquake occurrences can be only marginally described by a first order Markov chain model. Substitutability analysis incorporates the results of Markov Chains, revealing, though, detailed interrelations of parts (subareas) of the study area, not appreciated in Markov Chain analysis. Reactivation of particular subareas provides an insight into the level of interaction between neighbouring seismogenic sources within a subarea. The earthquake occurrence pattern provides evidence for the effect of a significant stress diffusion through time in the sense of a stress front. Taking into account the limitations of the methodologies applied, results indicate the importance of large-scale monitoring of seismicity, which assist in the identification of particular characteristics of the earthquake occurrence in space and time.

earthquake epicentres are presented in Figure 1. The Hellenic Arc (HA) and Kefalonia Fault (KF) are in a compression regime while extension characterises the back-arc region (BA). The North Anatolian fault (NAF) and North Aegean Trough (NAT) exhibit a complex character: instead of a restraining bend they show a transtensional character (Pavlidis and Caputo, 1994). The epicentre distribution follows, in general, the major geotectonic features. However, this correspondence is not straightforward due to the great seismotectonic complexity of the area. Shallow as well as deep earthquakes occur in the area. The latter are associated with a Wadati/Benioff subduction zone in the South of the area (southern part of HA).

The spatial distribution of earthquakes in Greece and its surroundings has been discussed in several studies. The existence of two main centres of activity for $M \geq 6.8$ events and a E-W-E migration of large ($M \geq 7.0$), shallow and intermediate depth events, was proposed by Galanopoulos (1963 and 1971, respectively). An E-W migration of $M \geq 6.6$ earthquake events is reported by Papazachos and Papadimitriou (1984) for the North Aegean area. Papadimitriou *et al.* (1985) supported the idea that a migration of large, shallow and intermediate depth events ($M \geq 7.0$) occurs in the N-S direction approximately every 30 years. A cyclic eastward migration throughout the Hellenic Arc was proposed by Purcaru and Berckhemer (1982) and Papadopoulos (1988), the latter concerning also historical (i.e. pre-1900) intermediate depth events.

The above observations rely on qualitative interpretation and applies to generally large ($M \geq 6.6$ or 7.0) events. They also include intermediate depth events, which occur almost exclusively in the southern part of the area, thus introducing a bias in the interpretation of the results. The purpose of the present work is to examine, on a quantitative basis, possible interactions between areas of occurrence of successive shallow (i.e., crustal) earthquake events in Greece and its surroundings.

1 Introduction

The distribution of earthquake events in space is of considerable interest for it may reveal occurrence patterns with important seismotectonic and geodynamic implications and of potential predictive value. Moreover, the existence of potential spatiotemporal patterns of earthquakes is of primary importance for the understanding of the long- and intermediate-term seismic process.

In recent years, studies provided evidence that the tectonic framework of Greece and its surroundings ($18-31^\circ$ E, $34-43^\circ$ N) resembles to a large extent a converging plate boundary (see, for example, Papazachos, 1990 and references therein). The major geotectonic features along with the distribution of

2 Data

The present work utilises moderate to strong ($M \geq 5.2$ and $M \geq 6.0$), main, shallow earthquake events (foreshocks and aftershocks excluded) which occurred in the area bounded by the parallels 34° and 43° N and the meridians 18° and 31° E (see Fig.1). The time interval considered is 1911-1985. The magnitude threshold for the data used is selected upon the basis of their frequency of occurrence and significance with macroseismic effects. The selection of the depth range is made according to the genetical homogeneity of the

events, taking also into account that intermediate depth events occur almost exclusively in the southern part of the area. The catalogue of Comninakis and Papazachos (1986) is the principal source for the data. The catalogue of Makropoulos and Burton (1981) and another catalogue of $M \geq 6.0$ events, published in Papazachos and Papazachos (1989) were also taken into consideration. According to Comninakis and Papazachos (1986) data with $M \geq 5.2$ are complete for the time interval investigated in the present work, while the error in epicentre determination is of the order of 30 km (0.25°).

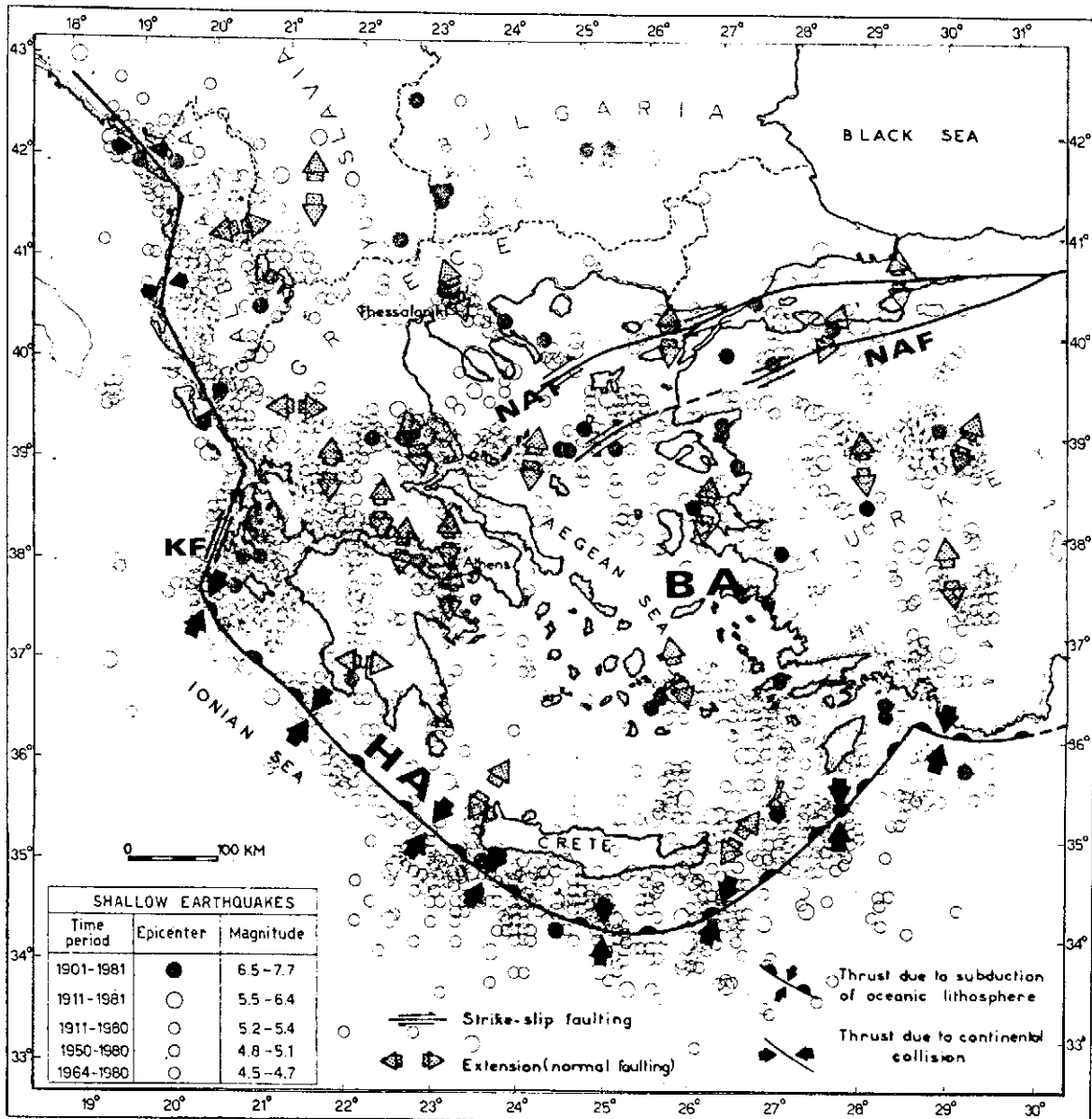


Fig. 1. Greece and its surroundings. Earthquake epicenters and tectonic features after Comninakis and Papazachos (1986) and Papazachos et al.(1986). KF: Kefalonia Fault; HA: Hellenic Arc; NAT: North Aegean Trough; NAF: North Anatolian Fault; BA: Back-Arc region (for explanation see text).

3 Methodology

For the purposes of quantitative analysis the data were grouped into spatial rectangular cells of equal area. These cells were chosen so as to contain enough events in order for the analysis to produce statistically meaningful results. Hence, the total area was divided into 6 and into 9 cells of equal area. The division into 6 cells (with dimension $4.5^\circ \times 4.3^\circ$) constitutes study Case A (hereafter referred to as Case A52 or A60, referring to $M \geq 5.2$ and $M \geq 6.0$ events, respectively). The division into 9 cells ($3^\circ \times 4.3^\circ$) constitutes study Case B (hereafter mentioned Case B52 and B60, respectively). Division into 9 cells made it possible to study the spatio-temporal distribution and potential migration of earthquakes in more detail.

The cell limits for the two cases are illustrated in Fig. 2. It should be pointed out that the above cells neither define nor represent any seismic 'zones', 'volumes' or 'sources', as for example those defined by Papazachos (1980, 1990). In fact, equivalence in area size is used to avoid separation into arbitrary defined seismicity 'zones', taking into account the uncertainty of epicentre locations. The edge effects were compensated by employing a guard region (see Davis, 1986; p.309) of about 20 km in every direction. The number of earthquake occurrence in every cell was counted first and is presented in Table 1. A χ^2 - test was carried out with the null hypothesis that the occurrence distribution is random. The test has $n-1$ degrees of freedom (hereafter denoted by dof), n being the number of cells. Results of the test (Table 1) imply that the distribution is not random, reflecting a natural selectivity in the distribution of earthquakes, as illustrated in Fig. 1.

To evaluate possible cell interactions, transition frequency matrices were formed, as Table 1 illustrates. The term 'transition' means the immediate passage from a state to another. In the present study, states are represented by cells. Occurrence of a state corresponds to activation of a cell, that is, occurrence of an earthquake in the cell. In Table 1, entries represent the total number of transitions between cells for the time interval 1911-1985. The matrix is read 'from rows to columns', meaning, for example, that a transition from a cell, i , to another cell, j , is counted as an entry in element a_{ji} of the matrix. That is, if we read in Case A52 (Table 1) from row N to column SW, we see that cell SW is activated after cell N (i.e. an earthquake occurred in cell SW after having occurred at cell N), 15 times in the earthquake sequence (element a_{15} of the matrix). When row totals of the transition probability matrix are divided by the total number of transitions, we obtain the relative proportions of the cells that are present in the sequence. This is called the *fixed* or *stationary probability vector* (denoted by FPV in Table 1). It allows us to predict what the transition probability should be if the activation of a cell was completely independent of that at the immediately preceding cell. Thus, the fixed probability vector is used for comparison of the observed transition counts to the expected ones.

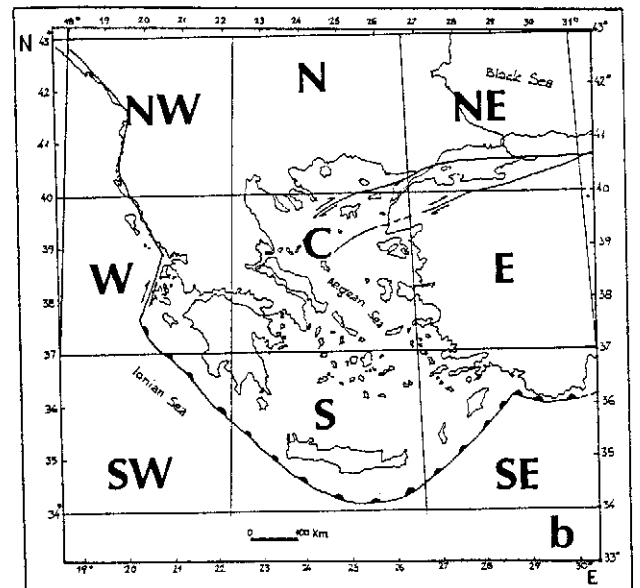
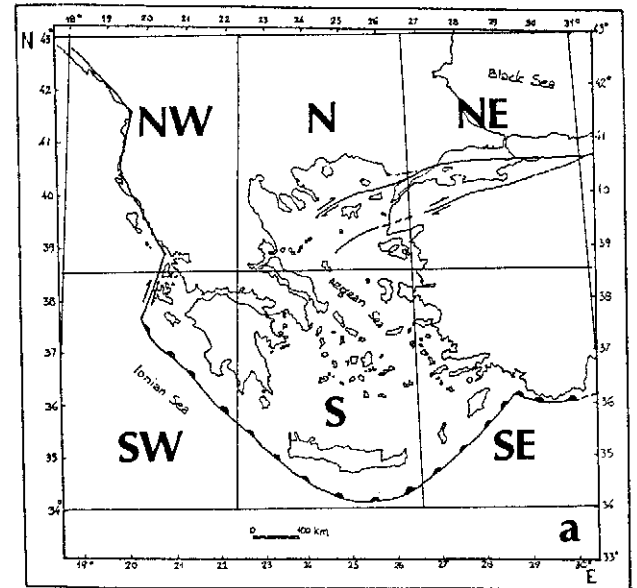


Fig. 2. Cell limits for Cases A and B. Labels refer to common geographic notation.

The tendency for one cell to succeed another is emphasised by normalising the transition counts, dividing each element in the i th row by the total of the i th row. In a probabilistic sense, these are estimates of the conditional probability, $P(j|i)$, that cell j will be the next to be activated, given that the present active cell is i . Since cells are mutually exclusive and exhaustive, one of the 6 or 9 must be activated and so the sum of conditional probabilities will be equal to 1.00. Conditional probabilities are also shown beneath the transition count in Table 1.

4 Markov Chain analysis

Following Krumbein (1968), Doveton and Skipper (1974) and Davis (1986), partial dependency, in a

Table 1. Occurrence and transition counts. Elements in *italics* refer to reactivation cases.

$M \geq 5.2$								$M \geq 6.0$													
Case A								Case B													
Occurrence counts								Occurrence counts													
N	NE	SE	S	SW	NW	Sum		N	NE	E	SE	S	SW	W	NW	C	Sum	$\chi^2 = 62.2$		$\chi^2 = 158.3$	
66	43	91	97	76	141	514		29	16	63	35	82	17	134	66	52	514	$\chi_{0.05,5}^2 = 11.7$		$\chi_{0.05,8}^2 = 15.5$	
Transition counts								Transition counts													
	N	NE	SE	S	SW	NW	FPV		N	NE	E	SE	S	SW	W	NW	C	FPV			
N	8	2	12	6	15	23	(0.13)	N	1	0	5	1	2	0	11	6	3	(0.06)			
	<i>0.12</i>	0.03	0.18	0.09	0.23	0.35			<i>0.03</i>	0.00	0.11	0.03	0.07	0.00	0.38	0.21	0.10				
NE	5	1	4	10	7	16	(0.08)	NE	1	0	2	1	1	0	10	1	0	(0.03)			
	0.12	<i>0.02</i>	0.09	0.23	0.16	0.37			0.06	<i>0.00</i>	0.13	0.06	0.06	0.00	0.63	0.06	0.00				
SE	12	10	11	17	16	25	(0.18)	E	3	0	5	6	15	2	16	9	7	(0.12)			
	<i>0.13</i>	0.11	<i>0.12</i>	0.19	0.18	0.27			0.05	0.00	<i>0.08</i>	0.10	0.24	0.03	0.25	0.14	0.11				
S	9	17	17	16	10	28	(0.19)	SE	4	4	4	4	5	4	11	11	8	(0.11)			
	0.09	0.18	0.18	<i>0.16</i>	0.10	0.29			0.07	0.07	0.07	<i>0.07</i>	0.09	0.07	0.20	0.20	0.15				
SW	16	3	17	19	9	12	(0.15)	S	5	7	14	7	14	0	19	11	5	(0.16)			
	0.21	<i>0.04</i>	0.22	<i>0.21</i>	<i>0.12</i>	0.16			0.06	0.09	0.17	0.09	<i>0.17</i>	0.00	0.23	0.13	0.06				
NW	16	10	30	29	19	36	(0.27)	SW	0	0	2	3	2	0	0	3	7	(0.03)			
	0.11	0.07	0.21	0.21	0.14	0.26			0.00	0.00	0.12	0.18	0.12	<i>0.00</i>	0.00	0.18	0.41				
								W	9	3	19	20	26	5	28	12	12	(0.26)			
									0.07	0.02	0.14	0.15	0.19	0.04	<i>0.21</i>	0.09	0.09				
								NW	4	2	5	6	12	3	19	6	8	(0.13)			
									0.06	0.03	0.08	0.09	0.18	0.05	0.29	<i>0.09</i>	0.12				
								C	2	0	7	7	5	3	20	6	2	(0.10)			
									0.04	0.00	0.13	0.13	0.10	0.06	0.38	0.12	<i>0.04</i>				
Occurrence counts								Occurrence counts													
N	NE	SE	S	SW	NW	Sum		N	NE	E	SE	S	SW	W	NW	C	Sum	$\chi^2 = 17.3$		$\chi^2 = 50.7$	
21	15	13	17	12	32	110		7	6	18	4	13	1	28	15	18	110	$\chi_{0.05,5}^2 = 11.7$		$\chi_{0.05,8}^2 = 15.5$	
Transition counts								Transition counts													
	N	NE	SE	S	SW	NW	FPV		N	NE	E	SE	S	SW	W	NW	C	FPV			
N	2	0	6	6	2	5	(0.19)	N	0	0	2	0	2	0	1	1	1	(0.06)			
	<i>0.10</i>	0.00	0.29	0.29	0.10	0.24			<i>0.00</i>	0.00	0.29	0.00	0.29	0.00	0.14	0.14	0.14				
NE	6	2	0	3	1	3	(0.14)	NE	1	1	0	0	1	0	1	0	2	(0.05)			
	0.40	<i>0.13</i>	0.00	0.20	0.07	0.20			0.17	<i>0.17</i>	0.00	0.00	0.17	0.00	0.17	0.00	0.33				
SE	1	3	2	3	2	2	(0.12)	E	2	1	2	2	4	0	3	1	3	(0.16)			
	0.08	0.23	<i>0.15</i>	0.23	0.15	0.15			0.11	0.06	<i>0.11</i>	0.11	0.22	0.00	0.17	0.06	0.17				
S	4	4	0	0	2	7	(0.15)	SE	0	0	1	0	0	1	1	1	0	(0.04)			
	0.24	0.24	0.00	<i>0.00</i>	0.12	0.41			0.00	0.00	0.25	<i>0.00</i>	0.00	0.25	0.25	0.25	0.00				
SW	0	1	2	3	1	4	(0.10)	S	2	1	1	0	0	0	6	2	1	(0.12)			
	0.00	0.09	0.18	0.27	<i>0.09</i>	0.36			0.15	0.08	0.08	0.00	<i>0.00</i>	0.00	0.46	0.15	0.08				
NW	8	5	3	2	4	10	(0.29)	SW	0	0	0	0	0	0	1	0	0	(0.01)			
	0.25	0.16	0.09	0.06	0.13	0.31			0.00	0.00	0.00	0.00	0.00	<i>0.00</i>	1.00	0.00	0.00				
								W	0	1	5	2	3	0	8	4	4	(0.25)			
									0.00	0.04	0.19	0.07	0.11	0.00	<i>0.30</i>	0.15	0.15				
								NW	2	2	1	0	1	0	3	3	3	(0.14)			
									0.13	0.13	0.07	0.00	0.07	0.00	0.20	<i>0.20</i>	0.20				
								C	0	0	6	0	2	0	5	1	4	(0.16)			
									0.00	0.00	0.33	0.00	0.11	0.00	0.28	0.06	<i>0.22</i>				

2 : transitions count FPV : fixed probability vector $\chi_{0.05,5}^2$: 0.05= confidence interval; 5= degrees of freedom
0.10 : conditional probabilities = transition frequencies

probabilistic sense, of a state on a preceding state is called a *Markov Chain*. When a statistical dependency exists between a state and its immediate predecessor, this constitutes the simplest or first-order Markov chain model.

To avoid arbitrary selection of sampling intervals (e.g. number of earthquake occurrences per cell) transitions from a state to itself were not permitted (c.f. Doveton, 1971 and Doveton and Skipper, 1974). This is called an *embedded* Markov Chain. Table 2 illustrates the matrix of embedded transitions along with the expected transitions, calculated by an independent events model. The resulting information loss from omission of the diagonal elements of the original matrices is compensated by a separate analysis in a later section.

The hypothesis that the activation of a cell is independent of the activation of the immediately preceding cell was tested with means of a χ^2 -test against the null hypothesis in which cells are independent of their immediate predecessors. The test is in the form of a contingency table. Each element of the observed transition matrix is compared to its expected value to calculate the test statistic. The test has $(m-1)^2 - m$ degrees of freedom (dof), where m is the number of cells involved (a dof is lost from each row because the probabilities in the row sum to 1.00). In case the test prerequisite of a minimum expected entry of 5 for each cell is not fulfilled by the entire matrix, only cells with predicted entries equal to or greater than 5 are employed in the statistic computation, leading to a conservative test of significance. If the null hypothesis is rejected, then a dependency between successive cells is established.

Higher order Markov properties can also exist. However, a decision has been made not to go into complex, multidependent models at this stage, especially since no temporal variations of the transition behaviour have not been considered either. Hence, the results reflect an 'average transition behaviour'.

4.1. Results of Markov Chain analysis

Both Cases A52 and B52 do exhibit a Markovian property. However, the values of the χ^2 -test at the 95% level of significance is not exceeded by a large amount (see Table 2). This implies a tendency of interaction of some cells, but not a particularly strong one. This is apparent from the very few transitions which result in a positive modified χ^2 -test result, illustrated in form of a flow diagram (Fig. 3, the arrow meaning "from...to..."). Moreover, as seen in the expected transition counts (illustrated beneath observed counts in Table 2), although some observed counts are high, the expected values are high also. It is interesting to notice that the above high counts are mainly transitions from and to the cells representing the western parts of the area of Greece and its surroundings, which eventually are the most active parts of the area (see Fig.1 and Table 1). Hence the low general as well as individual χ^2 values.

Nevertheless, the diagram illustrates that, regarding Case A52, activity in the southern part of HA (cells S and SW; see Introduction and Fig.2) shows a probabilistic tendency to be succeeded by activity in the NAF region (cell NE) and the northern part of BA (cell N). Low probability of interaction is observed

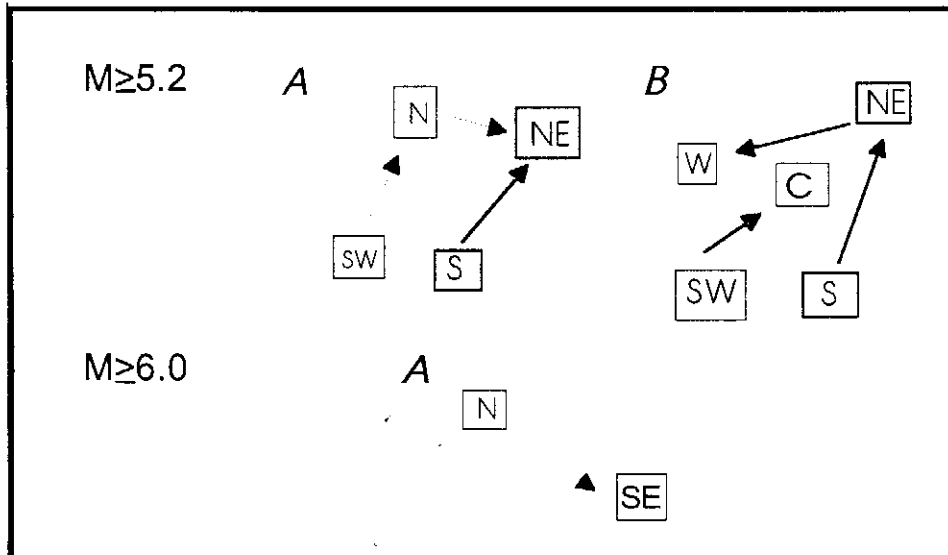


Fig. 3. Flow diagram of Markov Chain transitions with positive modified χ^2 -test result. Heavy arrows: primary transitions (χ^2 -value > critical acceptance value at $\alpha=0.05$); dotted arrows: minor transitions (critical acceptance value at $\alpha=0.10 < \chi^2$ -value < critical acceptance value at $\alpha=0.05$).

The structure of this dependency may be further investigated. Each transition count can be independently tested against its independent events alternative by a modified χ^2 procedure, applying Yates' correction (Davis, 1986; p.237). Transitions with a positive modified test result are linked in the form of a flowchart as an optimum transition scheme.

between cell N and cell NE. There appears no interaction of parts of HA. An interaction of parts of HA and of the back-arc region (BA) in favour of an interaction within the arc is also observed in Case B52, indicated by transitions S->NE, SW->C and NE->W. Notice that other parts of the area do not exhibit any significant relation with each other or with the ones

Table 2. Observed and Expected embedded transitions. Expected transitions calculated by an independent events model.

$M \geq 5.2$	Case A						Case B															
	N	NE	SE	S	SW	NW		N	NE	E	SE	S	SW	W	NW	C						
N	-	2	12	6	15	23		N	-	0	5	1	2	0	11	6	3					
		5.9	12.1	12.2	9.8	16.7				0.98	3.7	3.3	4.5	1.04	7.7	3.9	3.2					
NE	5	-	4	10	7	16		NE	1	-	2	1	1	0	10	1	0					
			8.6	8.7	7.0	11.8					0.98	2.0	1.9	2.5	0.61	4.3	2.2	1.8				
SE	12	10	-	17	16	25		E	3	0	-	6	15	2	16	9	7					
				17.9	14.4	24.4						3.7	2.0	7.1	9.8	2.3	16.6	8.4	6.9			
S	9	17	17	-	10	28		SE	4	4	4	-	5	4	11	11	8					
					14.5	24.7							3.3	1.8	7.1	8.5	1.97	14.6	7.3	6.0		
SW	16	3	17	19	-	12		S	5	7	14	7	-	0	19	11	5					
						19.8								4.5	2.4	9.9	8.5	2.7	19.9	10.0	8.2	
NW	16	10	30	29	19	-		SW	0	0	2	3	2	-	0	3	7					
															0.04	0.53	2.2	1.9	2.6	4.6	2.3	1.9
								W	9	3	19	20	26	5	-	12	12					
															7.7	4.2	16.9	14.6	20.3	4.6	17.1	14.0
								NW	4	2	5	6	12	3	19	-	8					
															3.9	2.1	8.4	7.2	10.1	2.2	17.2	7.1
								C	2	0	7	7	5	3	20	6	-					
															3.2	1.8	6.9	6.0	8.2	1.9	14.0	7.1

$\chi^2 = 39.8 \quad \chi_{0.05}^2 = 30.1 \quad (19)$

$\chi^2 = 87.4 \quad \chi_{0.05}^2 = 72.0 \quad (55)$

$M \geq 6.0$	Case A						Case B															
	N	NE	SE	S	SW	NW		N	NE	E	SE	S	SW	W	NW	C						
N	-	0	6	6	2	5		N	-	0	2	0	2	0	1	1	1					
			3.2	2.5	4.2	2.3	5.9				0.4	1.4	0.3	1.1	0.10	1.8	1.1	1.2				
NE	6	-	0	3	1	3		NE	1	-	0	0	1	0	1	0	2					
				1.7	2.8	1.5	3.8					0.4	0.9	0.2	0.7	0.10	1.1	0.7	0.8			
SE	1	3	-	3	2	2		E	2	1	-	2	4	0	3	1	3					
					2.2	1.2	3.1					1.4	0.9	0.7	2.7	0.2	4.2	2.5	2.9			
S	4	4	0	-	2	7		SE	0	0	1	-	0	1	1	1	0					
						5.1							0.3	0.2	0.7	0.6	0.04	0.9	0.5	0.6		
SW	0	1	2	3	-	4		S	2	1	1	0	-	0	6	2	1					
						2.8								1.1	0.7	2.7	0.6	0.1	3.3	2.0	2.3	
NW	8	5	3	2	4	-		SW	0	0	0	0	0	-	1	0	0					
															0.1	0.1	0.2	0.04	0.1	0.2	0.1	0.2
								W	0	1	5	2	3	0	-	4	4					
															1.8	1.1	4.2	0.9	3.3	0.2	3.1	3.5
								NW	2	2	1	0	1	0	3	-	3					
															1.1	0.7	2.5	0.5	2.0	0.1	3.1	2.1
								C	0	0	6	0	2	0	5	1	-					
															1.2	0.8	2.9	0.6	2.3	0.2	3.5	2.1

$\chi^2 = 27.6 \quad \chi_{0.05}^2 = 30.1 \quad (19)$

$\chi^2 = 67.5 \quad \chi_{0.05}^2 = 72.0 \quad (55)$

2 : observed embedded transitions count
 2.1 : expected embedded transitions count

$\chi_{0.05}^2$: 0.05= confidence level
 (19)= degrees of freedom

aforementioned. However, the above relations should be regarded with caution, since they may not reflect a global characteristic of the area but only a 'local' one, confined to the particular earthquake set employed in the present work.

No Markovian property is found for $M \geq 6.0$ events. This is to be mainly attributed to the sensitivity of Markov chain analysis to the amount of available data (Davis, 1986), rather than to an independence of successive cells since, also, the values of χ^2 -test are not much lower from the critical values of acceptance (see Table 2). The only positive individual modified χ^2 -test result implies a tendency for the SE part of HA to be preceded by the N part of BA (Case A60). However, no assessment of the validity of this dependence can be made.

5 Substitutability analysis

An alternative treatment of the observed transition frequencies is provided by substitutability analysis. Let us regard a process as being capable of producing any of the states a_1, a_2, \dots, a_n . Occurrence of a state is considered to be conditional upon the occurrence of all other states. The tendency of two or more states to occur in a *common context* is called *substitutability*. That is, the states are somehow related and one can substitute for the other state in a preferred succession (c.f. Davis, 1986; p.276).

A direct application of this concept requires knowledge of the infinite set of conditional probabilities between states in the sequence. However, with a finite sequence, higher order conditional probabilities can be assumed zero, thus permitting employment of only first-order conditional probabilities (see Rosenfeld *et al.*, 1968).

From the transition probability matrix we can develop two quantities, the first order *left substitutability* and the first order *right substitutability*. High substitutability between two states indicates that there is a strong tendency for the two states to be followed (or preceded) by the same states. That is, the two states have a similar transitional behaviour. If two states are *succeeded* by a third common state, they have a high *left substitutability*; a high *right substitutability*, if *preceded* by a third common state; and a high *mutual substitutability*, if preceded *and* followed by equivalent states.

Assessment of substitutability has been applied to analysis of parts of speech and processing of satellite photographs (Rosenfeld *et al.*, 1968) as well as to stratigraphic columns (Davis and Cocke, 1972; Doveton and Skipper, 1974). Calculation of substitutability proceeds as follows: suppose we have an $m \times m$ transition probability matrix where P_{ij} is the conditional probability that state (cell) a_j follows a_i . Left substitutability of states a_r and a_s is defined by the cross-product ratio between rows of the transition probability matrix as:

$$S_{rs} = \frac{\sum_{j=1}^m P_{rj} P_{sj}}{\sqrt{\sum_{j=1}^m P_{rj}^2 \sum_{j=1}^m P_{sj}^2}} \quad (1)$$

where r, s denote rows of the transition probability matrix. $S_{rs} = 1.0$ if rows r and s are identical. Hence, the substitutability measure is constrained to the range $0.0 \leq S_{rs} \leq 1.0$. Since the P_{ij} 's are calculated by dividing each entry in a row of the transition matrix by the row totals, the same result can also be achieved by directly computing the cross-product ratio between rows of the transition matrix. The result of computations is a symmetrical $m \times m$ matrix of cross-product ratios. Its diagonal elements will be 1.0, the other elements indicating the degree of similarity between one state and another, based on the percentage of times they are followed by a common third state. Right substitutabilities are calculated in an analogous manner, by computing the cross-product between columns of the transition probability matrix.

The symmetry of the substitutability matrices about the main diagonal enables interpretation in a hierarchical manner, using conventional clustering procedures. In the present work, a *weighted average group* method is applied (see Davis, 1986; p.504). The method proceeds by iteratively averaging initially clustered substitutabilities with others until finally two clusters remain. The results are presented in a dendrogram. However, the process of averaging introduces an error that becomes increasingly apparent as successive levels of clusters are averaged. The significance of this error can be evaluated by comparing the substitutability matrix with a matrix consisting of the results of the averaging. This is called the matrix of *cophenetic values* (see Davis, 1986; p.507). By calculating the correlation coefficient between the two matrices we obtain a measure of their similarity, the *cophenetic correlation coefficient* (hereafter referred to as CCC). Since there is no formal statistical procedure available for testing the results of hierarchical clustering, calculation of CCC is a way of evaluating the reliability and significance of clustered substitutabilities. Correlation coefficients below about 0.8 may indicate errors at lower linkages in the dendrogram so large that the diagram is misleading.

Substitutability matrices and derived dendrograms retain the full information contained in transition matrices. Hence, they may show significant collective patterns of spatiotemporal behaviour of earthquake activity that may not be easy to trace in the piecemeal statistical testing of individual transition pairs employed in Markov chain analysis. However, the transition probabilities which substitutability matrices are based on are estimates of their population parameters, which depend on the differing sizes of available observation data.

5.1. Results of substitutability analysis

The results of clustering appear in dendrogram form in Fig. 4.

Left substitutability dendrograms illustrate the various levels at which the component cells and their grouping show similarity with regard to the cells that immediately follow them. Right substitutability dendrograms illustrate collective properties in transition behaviour pattern with regard to immediately preceding cells or cell groups.

Regarding $M \geq 5.2$ events, high substitutability values are observed for cells representing the central and eastern part of Greece and its surroundings. That is, cells NE, SE, N, S and partly cell C (Case B5). These cells show a tendency to be followed (left substitutability) or preceded (right substitutability) by particular cells which, eventually, have low substitutability values, that is, cells NW and SW (Case A) and W (Case B).

The above high similarity of central and eastern cells implies that these behave in a similar way regarding the character of the earthquake generation process, as expressed through earthquake events. This reflects a similarity (a) in stress buildup and relaxation pattern and (b) the seismotectonic character of the cells in question. The degree of behavioural similarity in the dendrogram is understood to be a function of the role (significance) of the particular cell group in the regional geodynamic process, and it is not solely dependent on the seismotectonic structure of every cell.

Also noticeable is a high similarity of cells E and NW (Case B), reflecting a similar seismogenic behaviour of the two edges of HA (see Fig.1 for location). From the above, a transition scheme between the BA and HA is inferred, involving activation of either northern or southern cells in a general E-W direction. Alternatively, the sequence may initiate at either ends of HA and resume into the SSW part of HA. Figure 5a illustrates the case.

With $M \geq 6.0$ events, high substitutability is observed for cells representing the NNE and WSW parts of Greece and its surroundings that is, the northeastern part of BA and the WSW part of HA (see Fig.1 for location). These tend to be followed or preceded by the SSE part of HA. Alternatively, the pattern initiates in the NNW part of the back-arc region, resuming again in the south-southeastern edge of HA. The transition pattern has thus a rotational sense, as illustrated in Figure 5b.

The overall shape of dendrograms for both cases A and B indicates that the main clusters are formed at relatively similar levels, as for example do clusters {NE,SE} and {N,S} in Case A52. Hence, for Case A, clusters are well defined while a transitional tendency is not very strongly expressed. On the other hand, dendrograms of Case B indicate that a complex relation evolves around a reasonably well defined

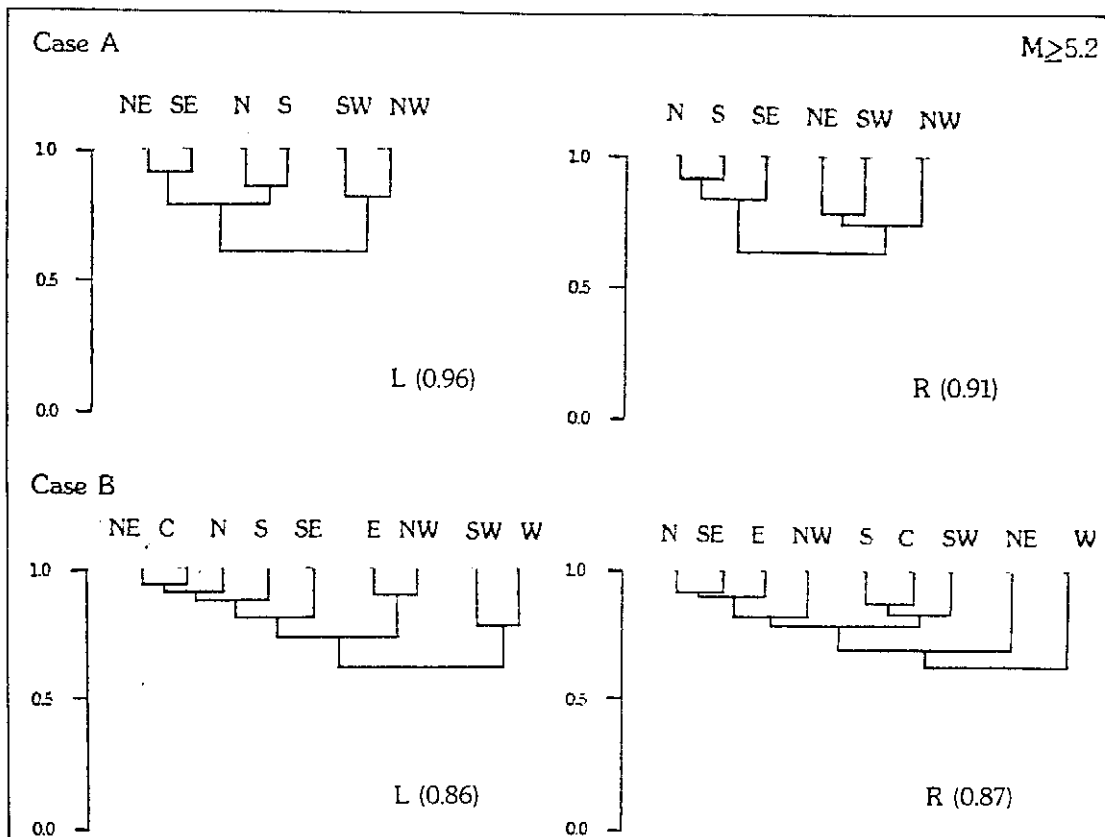


Fig. 4. Clustered left (L) and right (R) substitutabilities. Numbers in parentheses correspond to cophenetic correlation coefficients (CCC's). For notation see text and Fig. 2.

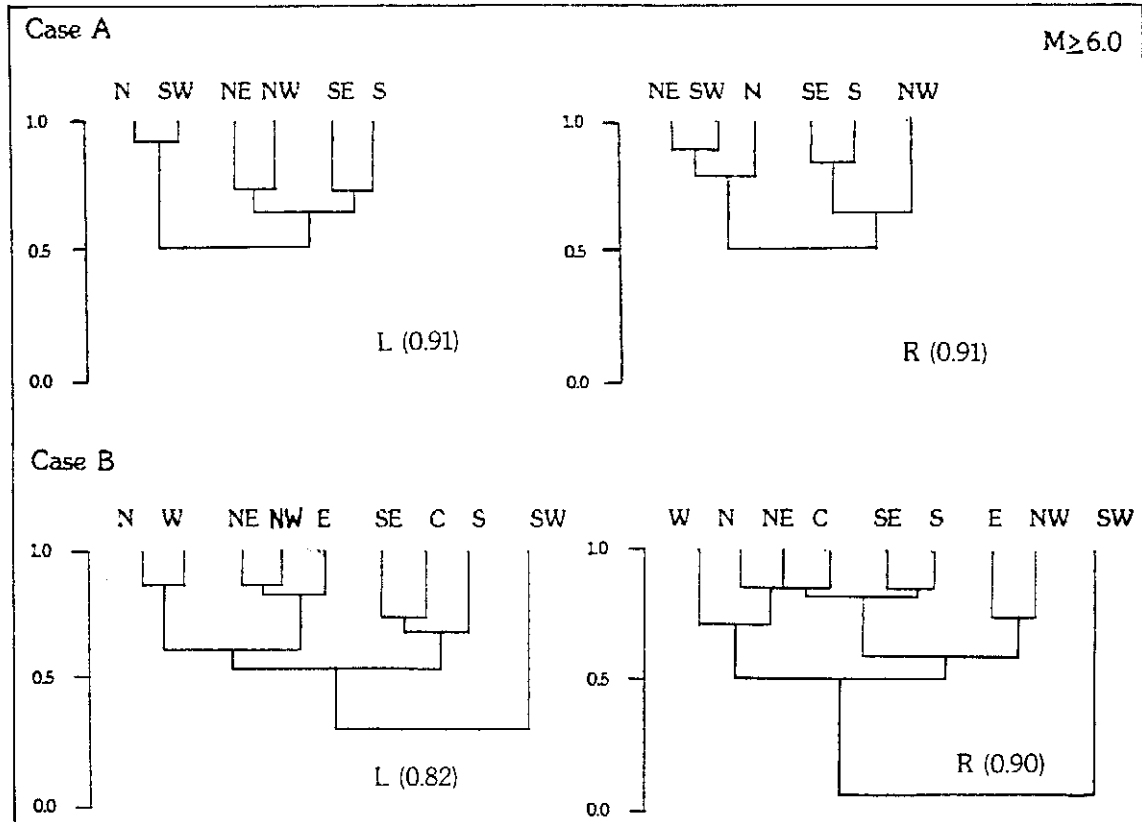


Fig. 4. Continued.

cluster 'core'. See, for example, cluster cores {NE,C} and {N,SE} in the left and right substitutability dendrograms, respectively, of Case B52. This complexity in cell formation implies that, considering the number of cells employed, more data were presumably needed to obtain a more clear image. This is particularly evident with $M \geq 6.0$ events, where the cell number is increased with respect to $M \geq 5.2$ events. Finally, the obtained CCC's in the range of 0.82-0.91, indicate that lower level linkages are of critical significance, implying a rather good discrimination of clusters (Case A) or cluster 'cores' (Case B) but not of other cells associated with clusters (Case B). The implication for these cells is that they enter the transition pattern in a more-or-less random manner.

Substitutability analysis incorporates the results of Markov chains in that it can not provide strong evidence for characteristic transition patterns. It mainly provides significant information regarding possible genetic relations between the studied cells.

6 Reactivation - The effect of earthquake reoccurrence in a cell

The diagonal elements of the original transition matrices (Table 1) provide evidence that in certain cases successive earthquake events occur in the same cell; that is, the cell can be termed as being

'reactivated'. The observation deserves further attention since (a) it constitutes about 10-30% of the transitions in every cell and about 15-20% of the total transitions and (b) it is not observed for all cells.

Reactivation of $M \geq 5.2$ events is highest in the western part of the study area (cell NW; Case A and W; Case B, respectively). The southern part (cell S) has the second highest reactivation in both cases A and B. The remaining cells in Case A show reactivation values not much different from cell S, while in Case B, values of reactivation are very low. No reactivation is observed for distant cells SW and NE in Case B. For $M \geq 6.0$ events, western regions show also high reactivation values. Southern cells show very low (Case A) or no (Case B) reactivation. The rest of the cells show also very low or no reactivation at all.

The frequency distribution of the diagonal elements was tested by means of a χ^2 -test against the null hypothesis of being random. The test has $n-1$ degrees of freedom (dof), where n is the number of cells. Results are shown in Table 3 and indicate that the distribution is not random regarding $M \geq 5.2$ events, while it is critically random for $M \geq 6.0$ events.

Considering the strain development in particular cells, it may be suggested that, in cases of reactivation, the accumulated seismic strain is gradually relaxed through the activation of several seismogenic sources in the immediate vicinity of the initially activated source. This behaviour of interacting neighbouring sources shall depend on the seismotectonic structure of

the region in question (e.g. fault network and size, geodynamic state, etc.).

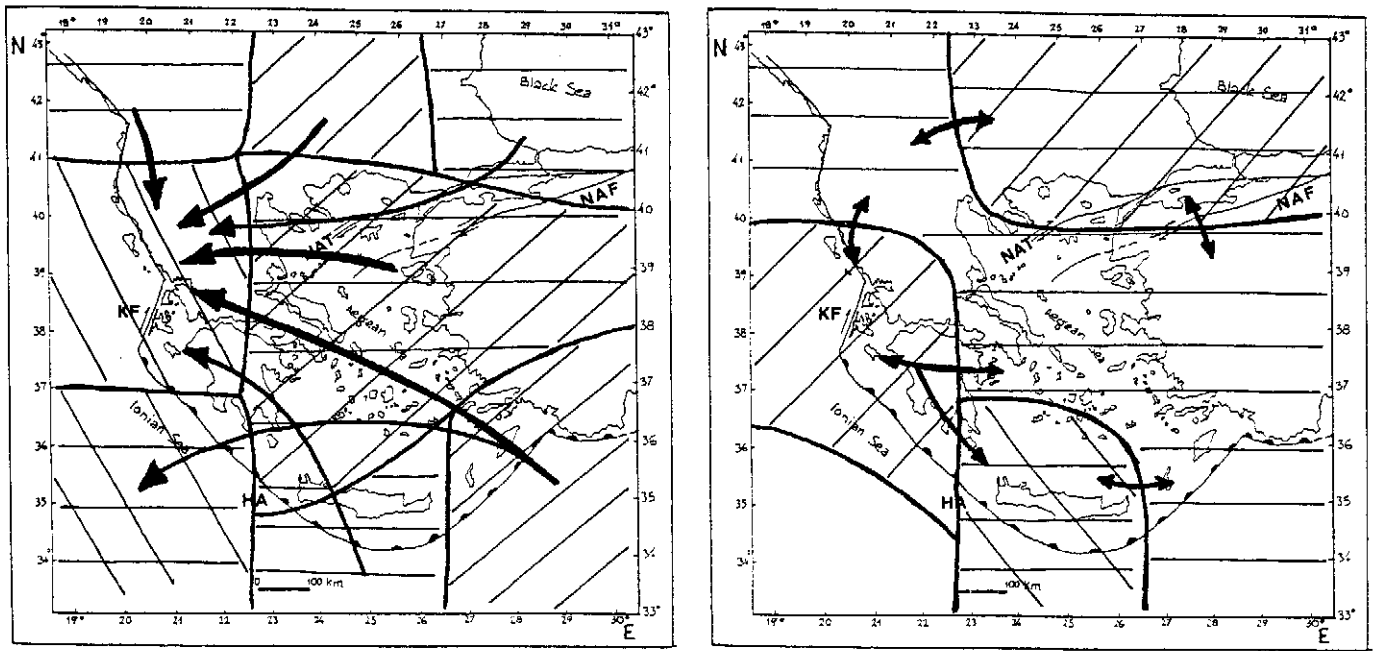
Table 3. Results of χ^2 - test for randomness of reactivation frequencies; (dof= degrees of freedom); 0.05 denotes the 5% level of significance.

$M \geq 5.2$ Case:	A	B
χ^2	15.3	23.4
$\chi^2_{0.05}(\text{dof})$	11.1(5)	15.5(8)

$M \geq 6.0$ Case:	A	B
χ^2	8.1	14.2
$\chi^2_{0.05}(\text{dof})$	11.1(5)	15.5(8)

7 Discussion and concluding remarks

Markov Chain and Substitutability analysis are simple procedures that can provide an image of potential association and interaction of different states of a physical system, in a probabilistic sense. A first-order Markov model implies that the system behaves stochastically and the generating process of future states is memoryless; that is, it depends only on the present state and not on past ones. Substitutability 'zooms in' properties of a sequence of states and reveals genetic relations. Although the earthquake generating process can be thought of being oversimplified when a Markov model is employed, results are still satisfactory in that they establish a simple 'one-dimensional' relation of particular regions.



$M \geq 5.2$

$M \geq 6.0$

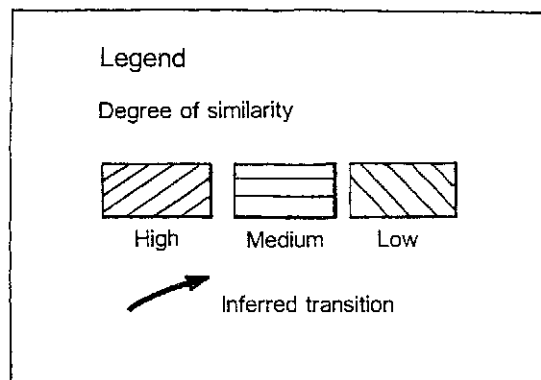


Fig. 5. Seismogenic relations (hatched areas) and migration paths (arrows) in Greece and its surroundings, as suggested by the results of the present work. Hatched patterns indicate degree of seismicogenic similarity; mixed patterns refer to mixed conditions. Arrow thickness indicates inferred relative importance of migration path. For the meaning of KF, HA, NAT, and NAF see Fig. 1.

Considering earthquake occurrence in Greece and its surroundings, the above techniques provide evidence for the tendency of some parts of the area (a) to, primarily, behave in a seismogenic similar manner and (b), secondarily, to be activated after or before certain other parts in the earthquake sequence.

The above information is to be seen in the light of the action of the principal agents in the geodynamic evolution of the area, namely: (a) the collision and subduction in the southern part of the Aegean in a principal SW-NE direction; (b) the eastward motion of the Adriatic microplate in the NW and (c) the westward motion of the Turkish microplate, expressed through the activity of the North Anatolian Fault in the ENE (Papazachos, 1988). Recently also, the role of the Kefalonia transform fault (KF) has been gained much appreciation (Reuther *et al.*, 1993). This structure seems to exercise a major control over the distribution of stresses throughout the whole area of Greece and its surroundings.

According to the above, the associative pattern and seismogenic similarities obtained through the aforementioned statistical techniques for events with $M \geq 6.0$ (illustrated in Fig. 5b) reflects the effect of a NE-SW acting stress field, controlled by the SW part of the HA and the NAF structure. Hence the observed high behavioural similarity expressed in the substitutability dendrograms. It is suggested that, during the stress relaxation in the areas located along the above trend, a stress buildup occurs in their neighbours, in a rotational sense. Events with $M \geq 5.2$ show a complex pattern due to the diffusion of earthquake activity into smaller scale seismotectonic features. Moreover, the pattern of $M \geq 5.2$ events is strongly influenced by the western part of the study area, which accommodate the highest earthquake activity (c.f. occurrence counts in Table 1).

Evidence of statistical association between parts of an active geotectonic region may be related to the models of geodynamic interaction proposed for example by Keilis-Borok (1990). These models consider the fragmented lithosphere as a hierarchy of volumes, which move to each other. The movement of these volumes is controlled by a wide variety of processes, concentrated in the boundary zones between the volumes. This movement is realised, to a large extent, through earthquakes. Modelling of these processes requires investigation at large scale. The present work hopes to have provided a preliminary empirical evidence at least for the validity of the above concept.

Acknowledgements. I am indebted to Dr.Ch.Kavouras at UCL for his help with the literature on Substitutability and Markov Chains. I wish to thank Ass.Prof. S. Pavlides and Prof. B. Papazachos for their ideas and encouragement. The comments and criticism of two referees have greatly assisted in improving the original manuscript. They are gratefully appreciated. An anonymous reviewer is also acknowledged for her help in improving the manuscript.

References

- Comninakis, P.E. and Papazachos, B.C., A catalogue of earthquakes in Greece and the surrounding area for the period 1901-1985, *Geophys. Lab. Univ. Thessaloniki Publ. 1*, 1-167, 1986.
- Davis, J.C., *Statistics and data analysis in Geology*, 2nd edition, Wiley, New York, 1986.
- Davis, J.C. and Cocke, J.M., Interpretation of complex lithologic succession by substitutability analysis, In: D.F.Merriam (Ed.), *Mathematical models of sedimentary processes*, Plenum, New York, 1972.
- Doveton, J.H., An application of Markov chain analysis to the Ayrshire Coal Measures succession, *Scott.J.Geol.*, 7, 11-27, 1971.
- Doveton, J.H. and Skipper, K., Markov Chain and substitutability analysis of a turbidite succession, Cloridorme Formation (Middle Ordovician), Gaspé, Quebec. *Can.J.Earth Sci.*, 11, 472-488, 1974.
- Galanopoulos, A.G., On mapping of seismic activity in Greece. *Ann. Geophys.*, 16, 10-37, 1963.
- Galanopoulos, A.G., Space-time seismicity of Greece. *Prakt.Akad. Athens*, 46, 216-224, 1971.
- Keilis-Borok, V.I., The lithosphere of the earth as a non-linear system with implications for earthquake prediction. *Rev. Geophys.*, 28, 1, 19-34, 1990.
- Krumbein, W.C., Statistical models in sedimentology. *Sedimentology*, 10, 7-26, 1968.
- Makropoulos, K.C. and Burton, P.W., A catalogue of seismicity in Greece and adjacent areas. *Geophys. J.R.astr.Soc.*, 65, 741-762, 1981.
- Papadimitriou, E.E., Karacostas, B.G., Karakaisis, G.F., and Papazachos, B.C., Space-time patterns of seismicity in the Aegean and surrounding area, *Procs. 3rd Intl. Sympos. Anal. Seismicity and Seismic Risk, Liblice, Czech., June 17-22*, 1985.
- Papadopoulos, G.A., Synchronized earthquake occurrence in the Hellenic Arc and implications for earthquake prediction in the Dodecanese Islands (Greece). *Tectonophysics*, 145, 343-3, 1988.
- Papazachos, B.C., Seismicity rates and long-term prediction in the Aegean area. *Quaterniones Geod.*, 3, 171-190, 1980.
- Papazachos, B.C., Active tectonics in the Aegean Sea and surrounding area. In: *Seismic Hazard in Mediterranean regions* (edited by J.Bonnin et al.), *Modern Appr.in Geophys.*, 7, 301-331, 1988.
- Papazachos, B.C., Seismicity of the Aegean and surrounding area. *Tectonophysics*, 178, 287-308, 1990.
- Papazachos, B.C., Kiratzi, A.A., Hatzidimitriou, P.M. and Karacostas, B.G., Seismotectonic properties of the Aegean area that restrict valid geodynamic models. *"Wegener/Medias Conference, Athens, Greece, 1986"*.
- Papazachos, B.C. and Papadimitriou, E.E., 1984. Seismicity regularities in Greece as premonitory patterns. *Procs. Intl. Sympos. Conti. Seismicity and Earthq.Prediction, Beijing, China, September 1982*, 191-197, 1982.
- Papazachos, B.C. and Papazachos, C., *The earthquakes of Greece*, Ziti, Thessaloniki, 1989.
- Pavlides, B. and Caputo, R., The North Aegean region: a tectonic paradox?, *Terra Nova*, 6, 37-44, 1994.
- Purcaru, G. and Berckhemer, H., Regularity patterns and zones of seismic potential for future large earthquakes in the Mediterranean region, *Tectonophysics*, 85, 1-30, 1982.
- Reuther, C.D., Ben-Avraham, Z., and Grasso, M., Origin and role of major strike-slip transfers during plate collision in the central Mediterranean, *Terra Nova*, 5, 249-257, 1993.
- Rosenfeld, A., Huang, H.K., and Schneider, V.H., An Application of Cluster Detection to Text and Picture Processing, *Univ. Maryland Computer Science Center, College Park, Md., Office Naval Res. Grant Nonr 5144(00), Tech.Rep. 68-68*, 1968.
- Sackin, M.J. and Merriam, D.F., Autoassociation, a new geological tool, *Journ.Intern. Assoc.Math.Geol.*, 1, 1-16, 1969.