Geophys. J. Int. (2008) **173,** 942–957 doi: 10.1111/j.1365-246X.2008.03773.x

Statistical identification of seismicity phases

A. Pievatolo and R. Rotondi

C.N.R. - Institute of Applied Mathematics and Information Technology, Via Bassini 15, 20133 *Milano, Italy. E-mail: reni@mi.imati.cnr.it*

Accepted 2008 February 26. Received 2008 January 23; in original form 2006 December 18

SUMMARY

We propose a statistical methodology to detect and quantify different seismicity phases on the basis of variations in certain characteristic features of seismic phenomenon, taking as examples two of the most studied aspects of seismicity: the occurrence rate and interevent time. Our objective is to provide an overall, compact picture of the activity, given a sufficiently long sequence of events, by identifying the flow of patterns that are singly described by well-known physical models in limited time intervals.

We assume that the sequence of events recorded in a seismically active region represents the set of an unknown number of consecutive phases, and that in each phase, the observations are the realization of a different version of the same stochastic process, in our case the Poisson or the Dirichlet process. The quantities to be estimated are the number of phases, the instants in which the system passes from one phase to another and the model parameters. We have formulated this problem in the framework of a multiple change-point problem. The inference is made possible by exploiting recently developed computer-intensive methods based on the stochastic simulation of Markov chains.

The data analysed are from the region of Kresna in southwestern Bulgaria. This area was hit by two strong earthquakes in the last century, one of $M_S = 7.8$ magnitude in 1904, and another of $M_S = 6.7$ magnitude in 1931.

Key words: Probability distributions; Seismic cycle; Earthquake dynamics; Statistical seismology.

1 INTRODUCTION

Despite many technological advances, we still have only a filtered image of the earthquake process. In fact, some key features cannot be directly observed; we can only record the effects, such as the frequency of events as well as their size and length. Analysis of the temporal distribution of seismicity is one of the most promising approaches in efforts to understand the underlying physical process. Here we cite just a few examples of studies in this direction. Evison & Rhoades (1998, 1999) propose, on the basis of quantitative empirical evidence, a qualitative physical process for the long-term seismogenesis of major earthquakes in subduction zones such as Japan, New Zealand and Greece. This process is constituted by three separate stages: swarms, main shocks and aftershocks, which are related by predictability. In other tectonic regimes, swarms are more appropriately replaced by more protracted seismicity anomalies. In Evison & Rhoades (2001), the causal relations among precursory seismicity, main shocks and aftershocks are explained by a three-stage faulting model comprised of crack formation, fracture and healing.

Zhuang (2000) has studied the earthquake sequence that occurred off Cape Palliser, New Zealand, in the period of 1978–1996; a preliminary exploratory analysis of the spatial-temporal distribution of the data leads to subdivision of the entire time interval into four periods: P_1, early background seismicity; P_2, a relatively quiescent period; P_3, main shocks and P_4, the aftershock sequence. To examine the apparent changes in the seismicity thus identified, some statistical models (the simple Omori law, the double Omori law with equal and different decay rates, the ETAS model, and their combinations) have been fitted to the global data set, or to its subsets, corresponding to the $P_\cdot$ periods. The performance of these models is measured by the residual analysis for point processes or Akaike's model selection criterion. Both the results of these procedures and the estimated values of the model parameters agree with the proposed partition of the sequence and suggest the hypothesis of four seismicity phases: interseismic, pre-seismic, coseismic and post-seismic. The four $P_\cdot$ subperiods initially identified correspond to these four phases in a seismic cycle.

A large number of studies in the literature combine the epidemic-type models and residual analysis to detect anomalies in seismic activity for the purpose of short-term earthquake prediction. Matsu'ura (1986) and Ogata (1989, 2001) find relative quiescences in aftershock sequences before major aftershocks, while Ogata *et al.* (1995, 1996) explore a statistical discrimination of foreshocks from earthquakes of other cluster types such as swarms or main shock–aftershocks sequences. Ogata (1992), analysing several data sets from a variety of local and global seismic regions, distinguishes three stages: ordinary seismicity; relative quiescence and recovery.

Wyss & Wiemer (2000) have examined spatial-temporal variations in the occurrence rate using the earthquake catalogue for southern California for the period 1981.0–1999.7, which includes

the Landers M 7.3 and Hector Mine M 7.1 earthquakes. They observe that the areas hit by these strong shocks had been in a state of precursory quiescence for a few years before earthquake occurrence, suggesting that a highly unusual decrease of seismicity rate without a nearby main shock, can be a precursor to a main shock and may indicate that the affected volume is primed for rupture (p. 1336, Wyss & Wiemer 2000). The same authors also point out that the redistribution of stress due to the Landers earthquake activated volumes in a temporary stress shadow and, conversely, turned off seismically active volumes, approximately displaying a pattern predicted by the Coulomb model. Goltz (2001) exploits a method based on eigenstructure (principal components) analysis to separate the seismicity rate patterns into a background component and components of change.

In their review, Zöller *et al.* (2007) seek an understanding of the underlying mechanisms of the seismicity patterns observed by adjusting numerical fault models to simulate the seismicity of real fault regions. They focus in particular on the detection of phenomena prior to larger earthquakes, which can be interpreted in terms of the advance towards a critical point in which an earthquake is seen as a phase transition.

We consider the observational data of the seismicity of a region as a realization of an underlying unknown (hidden) sequence of shifts among a finite number of possible states of the system states that can be characterized by variations in some features. Once the number and nature of these states have been identified together with their timing in a specific period, we must find a dynamic model that describes the fluctuation of the patterns. We can then predict which state is going to become active in the near future and, conditioned on this state, forecast the next earthquake. This perspective can be formulated in the framework of hidden Markov models—a class of models consisting of observations from different unknown states. The distribution of a future observation depends on the state of the system at that time, and the system moves from state to state according to the transition matrix of a Markov process.

This paper addresses the first part of this research project, aimed at identification and interpretation of phases based on variation of some features of seismic phenomenon. We consider the seismicity of a region as registered in the catalogue and propose a general statistical procedure that can answer these questions: does the examination of a significant feature indicate changes present in the seismicity pattern? If so, how many phases can be identified, and when did such changes occur? We formulate the issue in the general framework of a multiple change-point problem and solve the inferential aspects by resorting to advanced stochastic simulation methods, such as Markov chain Monte Carlo methods and the reversible jump method in particular (both methods are outlined in Appendix A). The procedure has been applied to the earthquake catalogue from the Kresna region of Bulgaria.

The paper is organised as follows. In Section 2, the tectonic structure of the region and the data used are described. In Section 3, we propose two indicators for the seismic pattern of a region: the occurrence rate and the distribution of interevent time. We model the variations of the former with a mixture of Poisson processes and those of the latter with a mixture of Dirichlet processes. We then express these models as a change-point problem, where every change-point marks the beginning of a different seismic pattern. The approach based on the distribution of interevent times includes the case of conditionally independent and identically distributed exponential distributions, corresponding to the approach based on the constant occurrence rate. In Section 4, we report the results obtained for each of the two indicators, and associate stages of a seismic cycle with the phases identified by the estimated change-points. In Section 5, we apply some tools of the free software package ZMAP to estimate seismicity rate changes to compare the potential of our methodology with that of a standard one.

2 DATA ANALYSIS

The choice of the data set for analysis is guided by the characteristics of the seismic, tectonic and geological processes. The Balkan region is characterized by destructive tectonic processes associated primarily with the vertical movement of tectonic blocks. Over the centuries, it has experienced strong earthquakes, although these are well documented only since 1900. In the last century, some of Europe's strongest events occurred in this area, in Pehcevo-Kresna (1904, $M_S = 7.8$) and Valandovo-Dojran (1931, $M_S = 6.7$). The earthquakes anticipated in these zones will be catastrophic for the entire Balkan region, including southwestern Bulgaria, Greece, Macedonia and the former Yugoslavia, in part because of the structural weakness of the prevailing traditional urban and rural architecture (Milutinovic 1998).

On the basis of these considerations, we examine the area located between latitudes 40.8°N–42.4°N and longitudes 22.0°E–24.4°E. It includes the Pehcevo–Kresna and Valandovo–Dojran seismogenic zones, which share seismotectonic properties (Simeonova *et al.* 2006; Tranos *et al.* 2006), and is comprised of southwestern Bulgaria together with the bordering regions of Greece and the Republic of Macedonia. The structure of this region is the result of differential vertical motion in extensional environments, characterized mainly by the Struma and Vardar lineaments tending NNW–SSE, and by transversal fault zones striking SW–NE (van Eck & Stoyanov 1996; Dineva *et al.* 1998). The most active seismic area is the Brezani fault zone (Krupnik fault) crossing the Struma lineament; in 1904, it suffered the strongest earthquakes recorded in this region. On the morning of 1904 April 4, a shock of magnitude $M_S = 7.1$ shook the Kresna region at 10:02 (GMT), followed 23 min later by another, stronger shock of magnitude $M_S = 7.8$ (Dineva *et al.* 1998). The size of these events is controversial (Shanov *et al.* 1999; Ambraseys 2001):a recent re-examination of the original seismograph records and bulletin data has re-determined their parameters; in particular, the magnitude estimates, $M_W = 6.8$ and $M_S = 7.2$, respectively, are both lower than those provided previously, whereas the occurrence times, $T_0 = 10{:}02$ and $T_0 = 10{:}26$, remain basically unchanged (Dineva *et al.* 2002). Since these magnitudes all exceed the threshold we indicate below, this is not a critical point for us. Moreover, Tsapanos *et al.* (2002) have estimated the seismic potential of the Kresna source as $m_{\max} 7.89 \pm 0.21$.

Few catalogues of historical earthquakes exist for the region under investigation; the data used in this study are taken from the recent catalogue by Dineva *et al.* (1999) dating from 1952 to 1995 and comprising 472 events of magnitudes ranging from 3.0 to 7.8, only seven of which occurred before 1800. Observing the curve of the cumulative number of events, one can note a change in slope and estimate the time of greatest change, by eye, as occurring around 1866–1895. To get a more precise answer, we have dealt with the completeness problem for different magnitude thresholds within the framework of the change-point problem (Rotondi & Garavaglia 2002, summarized in Appendix B). The longest reliable part of the catalogue is formed by $n = 130$ earthquakes of magnitude $M_S \geq 4.5$ that have occurred after 1890. They are listed in Table 1.

Besides the pair of strong earthquakes in 1904, the catalogue includes two other large shocks that hit the Valandovo area (SW of

Table 1. Index, date, epicentre and size of earthquakes of $M_S \geq 4.5$ that occurred after 1890 in the Kresna region.

N	Year	Month	Day	Lat	Long	M_S	N	Year	Month	Day	Lat	Long	M_S	N	Year	Month	Day	Lat	Long	M_S
1	1890	5	10	42.20	24.30	4.5	45	1904	8	1	42.00	23.20	5.0	89	1929	8	10	41.45	22.30	4.5
2	1890	6	7	42.40	23.70	4.6	46	1904	8	25	42.20	23.20	4.5	90	1931	3	7	41.35	22.35	6.0
3	1893	8	15	42.20	23.30	5.4	47	1904	10	29	41.90	23.20	5.0	91	1931	3	7	41.30	22.50	4.9
4	1894	11	25	42.20	23.00	4.9	48	1904	12	2	42.30	23.60	4.5	92	1931	3	8	41.30	22.50	6.7
5	1894	11	25	42.20	23.00	4.9	49	1904	12	11	42.00	23.40	4.5	93	1931	3	8	41.30	22.50	4.6
6	1894	11	26	42.20	23.20	5.3	50	1905	2	2	42.20	23.00	4.7	94	1931	7	23	42.00	23.40	4.9
7	1894	12	8	42.20	23.20	4.9	51	1905	7	9	42.00	24.20	5.1	95	1931	8	18	40.80	23.50	5.2
8	1894	12	19	42.40	23.30	5.2	52	1905	9	5	42.10	23.30	4.8	96	1932	4	23	41.30	22.70	5.0
9	1895	1	7	42.10	23.05	4.6	53	1905	10	8	41.80	23.10	6.4	97	1932	9	29	40.83	23.46	6.4
10	1895	1	7	42.10	23.30	5.1	54	1905	10	23	41.40	24.00	5.8	98	1932	11	1	41.70	23.60	4.5
11	1896	10	19	42.20	23.60	5.0	55	1905	11	18	41.00	23.00	5.6	99	1933	1	2	41.70	24.20	4.6
12	1899	9	21	42.30	23.60	4.5	56	1906	1	3	42.00	23.20	4.9	100	1933	5	8	41.50	24.20	4.9
13	1902	7	5	40.80	23.10	6.6	57	1906	4	23	41.80	23.20	4.5	101	1936	9	12	41.80	23.50	5.0
14	1903	11	25	42.10	23.20	5.5	58	1906	6	23	42.10	23.40	4.6	102	1937	9	8	41.60	24.10	4.8
15	1903	11	30	42.10	23.20	4.9	59	1906	10	8	41.80	23.10	6.4	103	1939	2	17	42.10	23.40	5.1
16	1903	12	1	42.10	23.40	4.6	60	1907	1	5	42.10	23.35	4.5	104	1944	3	14	41.70	23.90	5.1
17	1904	4	4	41.78	22.98	7.1	61	1907	8	17	41.30	22.50	4.9	105	1944	3	18	41.80	23.80	4.7
18	1904	4	4	41.80	23.10	5.1	62	1908	6	14	42.00	23.00	4.5	106	1952	6	27	40.80	23.50	4.7
19	1904	4	4	41.80	23.20	4.7	63	1909	3	10	42.00	23.00	4.6	107	1952	12	2	41.60	23.80	4.6
20	1904	4	4	42.10	23.40	5.2	64	1909	3	10	41.80	24.30	4.8	108	1952	12	17	41.80	23.50	4.5
21	1904	4	4	41.80	23.20	5.1	65	1909	4	1	42.00	24.00	4.5	109	1953	1	30	41.90	23.10	4.5
22	1904	4	4	41.80	23.10	7.8	66	1910	2	23	41.80	23.50	5.2	110	1953	9	4	41.90	23.10	4.7
23	1904	4	4	42.10	23.40	5.4	67	1910	3	22	41.20	22.00	5.0	111	1955	7	9	40.90	22.10	5.1
24	1904	4	4	42.10	23.30	4.8	68	1910	4	6	42.10	23.20	4.5	112	1955	7	10	40.90	22.10	4.5
25	1904	4	4	42.10	23.40	5.3	69	1911	3	6	41.00	22.00	4.5	113	1962	10	8	41.90	24.30	4.5
26	1904	4	4	42.20	23.10	4.5	70	1911	3	11	41.60	22.40	5.0	114	1962	10	8	42.00	24.20	4.9
27	1904	4	4	41.80	23.10	4.7	71	1911	3	11	42.00	23.00	5.5	115	1964	7	4	41.90	23.40	4.8
28	1904	4	4	42.10	23.40	4.8	72	1911	3	16	41.00	22.00	4.5	116	1965	3	19	41.40	22.90	4.5
29	1904	4	4	42.00	23.00	5.5	73	1912	11	7	41.90	24.00	4.8	117	1972	5	8	41.60	23.60	4.9
30	1904	4	4	42.00	23.20	4.5	74	1914	3	3	41.70	23.00	4.7	118	1972	7	8	41.60	23.60	4.7
31	1904	4	4	41.80	23.10	4.7	75	1914	3	22	41.70	23.00	4.7	119	1972	12	13	41.50	24.00	4.6
32	1904	4	4	42.20	24.40	4.5	76	1916	2	23	41.90	22.80	4.5	120	1974	6	21	41.28	22.70	4.7
33	1904	4	7	41.80	23.20	5.0	77	1916	4	10	41.70	22.00	4.8	121	1977	11	3	42.10	24.10	5.4
34	1904	4	9	42.10	23.40	4.5	78	1916	11	18	41.30	22.50	4.5	122	1978	6	21	40.81	23.06	5.0
35	1904	4	10	42.00	23.50	5.2	79	1917	4	4	42.00	23.00	5.0	123	1978	7	4	40.82	23.19	4.8
36	1904	4	13	42.40	22.80	5.5	80	1920	12	6	41.45	22.40	4.8	124	1978	12	31	41.99	23.22	4.6
37	1904	4	19	42.00	23.10	5.9	81	1921	5	4	42.00	22.80	5.0	125	1983	8	26	41.01	22.39	4.5
38	1904	4	21	42.10	23.20	4.5	82	1921	9	12	41.70	22.80	4.7	126	1985	9	28	41.50	22.31	4.9
39	1904	4	25	42.00	23.00	5.5	83	1925	1	7	42.00	22.40	4.7	127	1985	11	9	41.19	24.05	5.1
40	1904	5	6	42.30	23.60	4.8	84	1926	9	3	41.75	24.40	4.9	128	1985	11	9	41.23	24.02	5.4
41	1904	5	10	41.80	23.20	4.6	85	1927	7	23	41.70	22.70	4.9	129	1985	11	21	40.85	22.94	5.1
42	1904	5	12	41.80	23.20	5.1	86	1928	4	18	42.30	24.20	4.6	130	1990	12	21	40.94	22.42	5.9
43	1904	6	21	42.10	23.40	5.3	87	1928	4	27	42.40	24.40	4.9							
44	1904	7	24	42.20	23.00	4.7	88	1929	7	3	41.50	22.00	4.5							

the Krupnik earthquake) on 1931 March 7 and 8 with magnitudes of 6.0 and 6.7, respectively.

Fig. 1 shows the different components of the data set, with magnitudes and occurrence times in the top graph and the locations of the epicentres in the bottom graph. The low quality of event location in the catalogue makes many epicentres appear to coincide.

First of all, we perform an exploratory analysis of the data on the basis of an agglomerative hierarchical clustering procedure for partitioning the n objects (earthquakes) into all possible values k of the clusters (Kaufman & Rousseeuw 1990). The information is drawn from four attributes: magnitude, latitude, longitude and occurrence time. In this way, we arrange the measurements in a $4 \times n$ matrix. The algorithm starts with each object in a separate cluster and, in the first step, joins the least dissimilar pair of objects. In each of the following steps, the dissimilarity between pairs of clusters is evaluated, and the least dissimilar pair is merged; the algorithm proceeds until all the objects are in a single group. The hierarchy provided by the algorithm can be represented by a clustering tree. According to the 'complete linkage' method, the dissimilarity between two clusters, R and Q, is defined as the 'greatest' dissimilarity between an object i of R and an object j of Q, $d(R, Q) = \max_{i \in R, j \in Q} d(i, j)$, and, in our case, d is the Manhattan distance $d(i, j) = \sum_{k=1}^{4} |x_{i,k} - x_{j,k}|$. Between-cluster dissimilarity can be also defined in other ways, but we have preferred the more cautious complete linkage method where relatively similar objects will often remain in different clusters for a long time, and only strongly similar objects are placed in the same cluster.

In the initial clustering, we exploit all information (magnitude, latitude, longitude and occurrence time). Fig. 2 reproduces the part of the tree generated by the algorithm, where the clusters are well separated. The two nodes of the final branches of the tree are identified, respectively, by the dates of the oldest and the most recent earthquake belonging to the group. The same analysis is then repeated using only the temporal information (Fig. 3). The structure

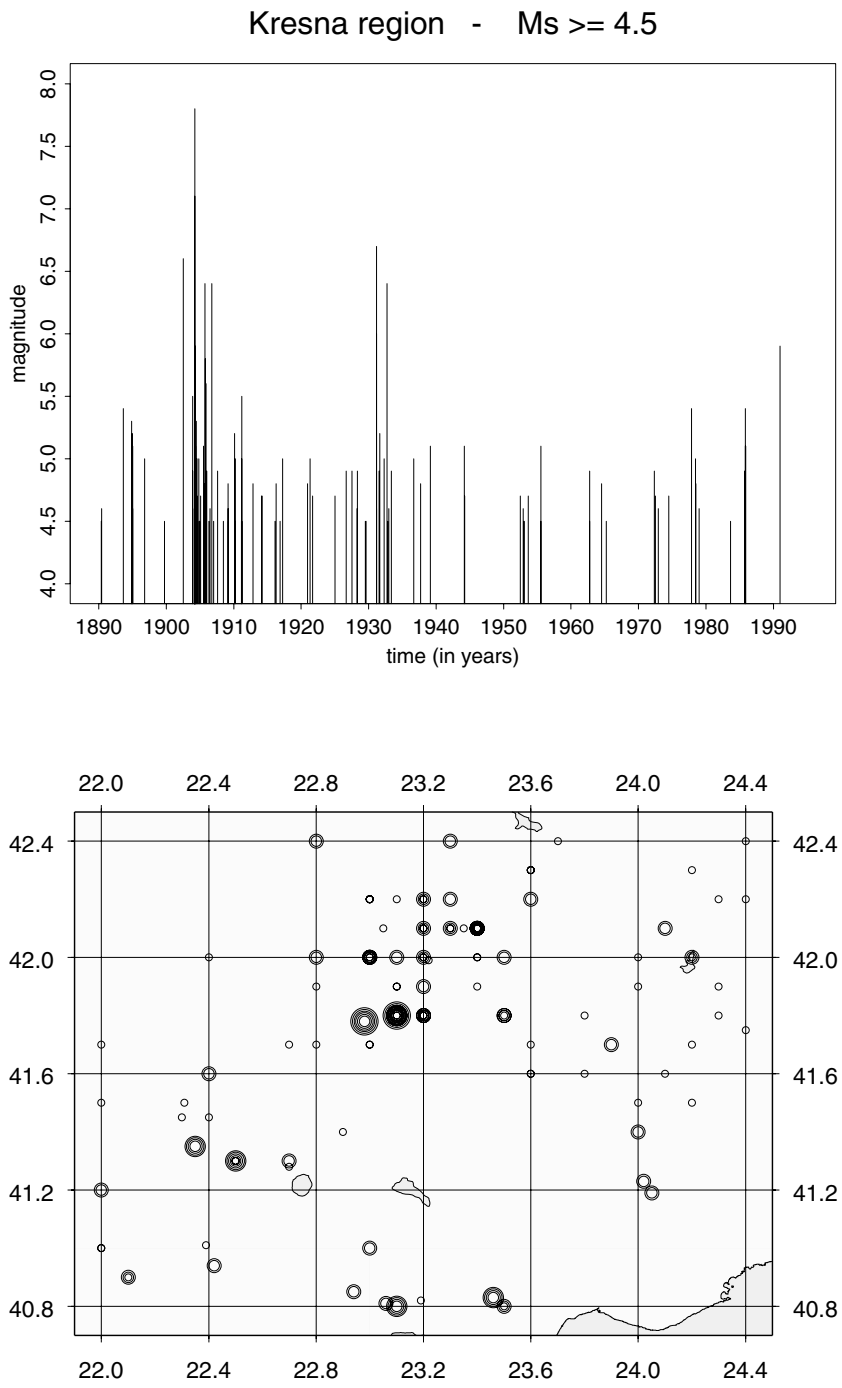

Figure 1. Occurrence times, magnitude (top) and locations (bottom) of $M_S \geq 4.5$ earthquakes in the Kresna region. Circle size indicates magnitude.

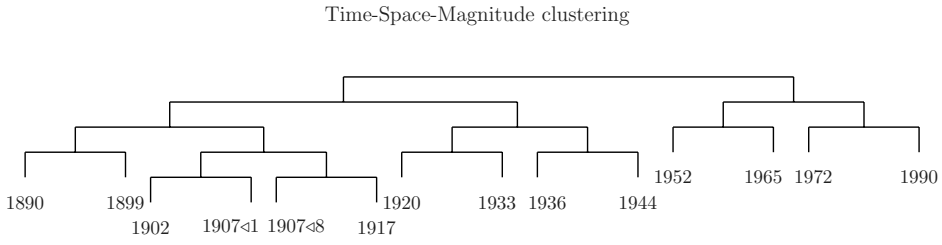

Figure 2. Clustering tree produced by agglomerative hierarchical clustering applied to the entire data set. The dates at the final nodes identify the first and the last earthquake in the cluster.

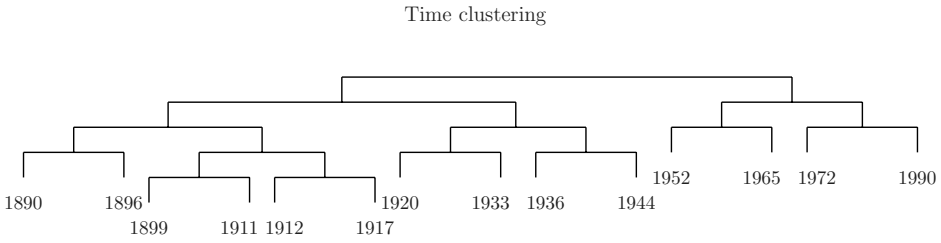

Figure 3. Clustering tree produced by agglomerative hierarchical clustering applied to the occurrence times. The dates at the final nodes identify the first and the last earthquake in the cluster.

of the two trees is the same; only the nodes of the cluster including the strongest earthquake change so that the cluster is smaller when all the data are analysed than when only the temporal information is considered. This may be due to the addition of temporal and spatial clustering effects immediately after a strong event, but it does not, at any rate, substantially affect the initial conclusions we can draw regarding the evolution of the phenomenon. Once the region and the magnitude threshold are appropriately chosen, the time appears to be the most expressive variable with regard to the objective of the present work. Consequently, we focus our efforts on examination of the temporal data, neglecting the spatial information and using the size of the events to choose the region and set the magnitude threshold.

3 MULTIPLE CHANGE-POINT PROBLEM

Let us assume that it is possible to draw information on meaningful facets of the seismic pattern in a region from a seismic catalogue. Changes in these features signal changes in seismic activity. If we were able to identify some cyclic trend in the sequence of these changes, their identification could have considerable value in forecasting future events. Assuming that the occurrence rate and the time between subsequent events are among the most significant aspects of the seismic phenomenon, and given the sequence of earthquakes recorded in a region with sizes larger than a set threshold, our aim is to identify the number, location, and size of variations of the above-mentioned features. Sections 3.1 and 3.2 present the theoretical modelling of this problem for the occurrence rate and for the interevent time respectively.

3.1 Changes in the occurrence rate: A mixture of Poisson processes

Let $\mathbf{t} = (t_1, \ldots, t_n)$ be the occurrence times of the n earthquakes recorded in the time interval $\mathcal{T} = [T_1, T_2)$ in the region under study, and let $\bigcup_{j=0}^{k}[s_j, s_{j+1})$ be a partition of the interval $\mathcal{T}$ with $s_0 = T_1$, $s_{k+1} = T_2$ and $T_1 < s_1 < s_2 < \cdots < s_k < T_2$. We assume that the observed times are realizations of a Poisson process with different rates h_j in the $(k+1)$ disjoint subintervals $[s_j, s_{j+1})$, $j = 0, \ldots, k$, which is a mixture of Poisson processes characterized by the intensity step function

$$\lambda(t) = \sum_{j=0}^{k} h_j \, \delta_{[s_j, s_{j+1})}(t) \tag{1}$$

where $\delta_{(a,b)}(t)$ is equal to 1 if and only if $t \in (a, b)$. This means that the seismic activity has changed in the interval $\mathcal{T}$; the number k, the positions $\mathbf{s} = \{s_j\}_{j=1}^{k}$ of the changes and the different rates $\mathbf{h} = \{h_j\}_{j=0}^{k}$ must be estimated. The log-likelihood has the form:

$$\log \mathcal{L}(\mathbf{t} \mid \mathbf{h}, \mathbf{s}, k) = \sum_{j=0}^{k} \left[\sum_{i=1}^{n} \delta_{[s_j, s_{j+1})}(t_i) \, \log h_j \right] - \sum_{j=0}^{k} h_j \, (s_{j+1} - s_j). \tag{2}$$

In the Bayesian framework, to complete the definition of the model, we must assign *a priori* distributions to the parameters, as an expression of our beliefs regarding the physical process. Little is known about k, apart from the fact that, at the scale in which we are studying the phenomenon, a seismicity phase is at least a few years long. Hence, it is possible to assign an upper bound $k_{\max}$ to the number of changes and to assume k is uniformly distributed on $\{0, k_{\max}\}$. Having k, one has the model M_k with parameters $\mathbf{s} = (s_1, s_2, \ldots, s_k)$ and $\mathbf{h} = (h_0, h_1, \ldots, h_k)$; the locations of the change-points are drawn independently from the uniform distribution on $\mathcal{T}$, and then ordered. As for the rates $h_0, h_1, \ldots, h_k$, we can obtain information on their global features, on the variability range of their average value, for instance, from examination of the seismic activity of tectonically similar regions. We assume, therefore, that they follow the $Gamma(a_0, b)$ distribution and, to take into account that uncertainty, we add a further stage to the model by assuming that b follows the $InvGamma(c_0, d_0)$ distribution. For a more detailed presentation and an application of this model to some Italian regions, see Rotondi (1999).

To estimate the parameters, we have to compute the posterior joint distribution $p(k, \boldsymbol{\theta}_k \mid \mathbf{t})$, with $\boldsymbol{\theta}_k = (\mathbf{h}, \mathbf{s}, b)$, and its marginals. This, however, requires the computational burden of evaluating multidimensional integrals. A feasible alternative solution is offered by a class of iterative methods, the so-called Markov chain Monte Carlo (MCMC) methods, based on the simulation of a Markov chain of which $p(k, \boldsymbol{\theta}_k \mid \mathbf{t})$ is the equilibrium distribution. The generation of such a chain presents two problems. First, generating different k means jumping, inside the same procedure, between models M_k of different dimensions. The reversible jump Markov chain sampler proposed by Green (1995) enables us to solve this. Then, having set k, that is, having assigned the model M_k, the parameter vector $\boldsymbol{\theta}_k$ can be estimated through Gibbs sampling, one of the best known MCMC algorithms. The reader is referred to Green (1995) for details on the basic reversible jump methodology and to Rotondi (2002) for a more efficient version of that method in the context of the multiple change-point problem.

3.2 Changes in the probability distribution of interevent time: a mixture of Dirichlet processes

Let us denote by $(t_1, t_2, \ldots, t_{n+1})$ the vector of the occurrence times, and focus on the sequence of interevent times $(y_1, \ldots, y_n)$, where $y_i = t_{i+1} - t_i$, $i = 1, \ldots, n$, which we suppose to be stochastically

independent. Hence, the interevent times follow a renewal process, and we assume that the probability distribution of these variables changes when the seismicity phase changes. In particular, we denote by r_j the index of the first interevent time y_{r_j} of the jth phase, with $j = 1, \ldots, k$, implying that t_{r_j} is the date of the last event of the $(j-1)$th phase, and t_{r_j+1} is that of the first of the jth phase. Summarizing, we denote by $\mathbf{r} = (r_1, \ldots, r_k)$ the locations of the k change-points, where $r_j \in \{2, \ldots, n\}$, whereas the interevent times $y_{r_j}, \ldots, y_{r_{j+1}-1}$ follow the same probability distribution F_j, $j = 0, \ldots, k$, with $r_0 = 1$ and $r_{k+1} = n + 1$. Formally, this is expressed by the joint distribution

$$\Pr(Y_1 \le y_1, \ldots, Y_n \le y_n \mid \mathbf{r}, k) = \prod_{j=0}^{k} \prod_{i=r_j}^{r_{j+1}-1} F_j(y_i). \tag{3}$$

Many parametric distributions have been proposed in the literature to model the time between consecutive earthquakes (Utsu 1984; Wang & Kuo 1998), but the problem has not been altogether solved; therefore, it is clear that any probability distribution we were to choose for the interevent times would leave out some aspects of the phenomenon. Consequently, we have decided to consider F_j, for each $j = 0, \ldots, k$, as a random distribution that follows a Dirichlet process with parameter $MG(\cdot\,; \theta_j)$, where M is a constant, and G is a distribution function over $(0, \infty)$ with unknown parameter θ_j. A Dirichlet process over $(0, \infty)$ is a multivariate distribution whose domain is a set of probability distributions F on $(0, \infty)$ (Ferguson 1973); that is, it is a powerful tool for assigning the probability to the unit vectors $[F(B_1), \ldots, F(B_m)]$, given any measurable partition $\{B_1, \ldots, B_m\}$ of the support $(0, \infty)$ of F. The density function of a Dirichlet distribution is

$$f(x_1, \ldots, x_m) = \frac{\Gamma(\nu_1 + \ldots + \nu_m)}{\Gamma(\nu_1) \ldots \Gamma(\nu_m)} x_1^{\nu_1 - 1} \ldots x_m^{\nu_m - 1},$$

where $x_i \ge 0$, $\sum_{i=1}^{m} x_i = 1$ and the parameters $\nu_i, i = 1, \ldots, m$, are positive; in our case, $x_i = F_j(B_i)$ and $\nu_i = MG(B_i; \theta_j)$, $i = 1, \ldots, m$, with $G(B_i; \theta_j)$ being the probability assigned by the distribution $G(\cdot\,; \theta_j)$ to the set B_i.

As the distribution $G(\cdot)$ is the *a priori* expectation of F, we have set it equal to a generalized gamma distribution with density function:

$$g(y; \eta, \xi, \rho) = \frac{\eta\, \xi^{\rho} y^{\rho\eta - 1} e^{\xi y^{\eta}}}{\Gamma(\rho)}, \tag{4}$$

where $\boldsymbol{\theta}_j = (\eta_j, \xi_j, \rho_j)$, $j = 0, \ldots, k$. This distribution was chosen because it includes the most used distributions for the recurrence time:exponential ($\eta = 1$, $\rho = 1$), gamma ($\eta = 1$), Weibull ($\rho = 1$), and asympotically lognormal (as $\rho \to \infty$). The literature tells us that MLE numerical methods often fail to converge unless the sample size is very large; MCMC methods for Bayesian estimation, on the contrary, do not seem to suffer from this problem.

A fundamental property of the Dirichlet process is that the conditional distribution of, let us say, F_j, given the sample $\mathbf{y}$, is still a Dirichlet process with parameters that are easily updated:

$$(F_j \mid \boldsymbol{\theta}_j, r_j, r_{j+1}, \mathbf{y}) \sim \mathcal{D}\left\{ MG(\cdot\,; \theta_j) + \sum_{i=r_j}^{r_{j+1}-1} \delta_{y_i}(\cdot) \right\}, \tag{5}$$

where $\delta_y(\cdot)$ is the measure that concentrates mass 1 at point y (Ferguson 1973).

We now complete the model by assigning the prior distributions of the remaining parameters. Given k, the vector $\boldsymbol{r}$ is chosen at random in the set of all the possible combinations of k elements taken into $\{2, \ldots, n\}$; as in the previous case, k is uniformly distributed on $\{0, k_{\max}\}$ with $k_{\max} = n - 1$, and the parameters of the generalized gamma distribution are independently distributed:$\eta_j \sim Gamma(g, \tau)$, $\xi_j \sim Gamma(a, \beta)$ and $\rho_j \sim \mathrm{Exp}(b)$, $\forall\ j = 0, \ldots, k$, with $\boldsymbol{\omega} = (\beta, \tau)$ as random variables. At the last level of the hierarchical model, we have $\beta \sim Gamma(c, d)$ and $\tau \sim Gamma(h, f)$. The letters g, a, b, c, d, h and f denote fixed parameters of the priors, called hyperparameters. Considering the parameters $\boldsymbol{\theta}, \boldsymbol{\omega}$ as random variables, each distribution F is distributed as a mixture of Dirichlet processes. The model is described in greater detail in Pievatolo & Rotondi (2000), where a method is proposed to assign the hyperparameters of the priors on the parameters of the generalized gamma distribution.

A basic theorem of Antoniak (1974) characterizes the distribution of a sample from a mixture of Dirichlet processes, enabling us to express the likelihood of sample $\mathbf{y}$ as

$$\mathcal{L}(k, \mathbf{r}, \boldsymbol{\theta}; \mathbf{y}) = \prod_{j=0}^{k} \frac{M^{r_{j+1} - r_j}}{M^{[r_{j+1} - r_j]}} \prod_{i=r_j}^{r_{j+1}-1} g(y_i; \eta_j, \xi_j, \rho_j), \tag{6}$$

where $M^{[m]} = M\,(M+1)\ldots(M+m-1)$.

Summarizing, the posterior joint distribution is given by:

$$\pi(k, \boldsymbol{r}, \boldsymbol{\theta}, \boldsymbol{\omega}) \propto L(k, \boldsymbol{r}, \boldsymbol{\theta}; \mathbf{y})\, \pi(k)\, \pi(\mathbf{r} \mid k)\, \pi(\boldsymbol{\theta} \mid k, \boldsymbol{\omega})\, \pi(\boldsymbol{\omega}) \tag{7}$$

with $\pi(\boldsymbol{\omega}) = \pi(\beta)\ \pi(\tau)$. Like the mixture of Poisson processes, the MCMC algorithm for sampling from this posterior joint distribution is the combination of the reversible jump MCMC method and Gibbs sampling (see Appendix A). For a full explanation of its implementation, see Pievatolo & Rotondi (2000).

4 RESULTS AND INTERPRETATION

4.1 Results for variations in rate

The algorithm implementing the MCMC method for estimation was run for $\nu = 400\,000$ iterations, obtaining the sequence of sampled vectors $(\mathbf{s}^{(1)}, \mathbf{h}^{(1)}, \beta^{(1)}, k^{(1)})$, $(\mathbf{s}^{(2)}, \mathbf{h}^{(2)}, \beta^{(2)}, k^{(2)})$, $\ldots$, $(\mathbf{s}^{(\nu)}, \mathbf{h}^{(\nu)}, \beta^{(\nu)}, k^{(\nu)})$. One of the conditions used to ensure convergence of the chain to the equilibrium distribution requires that its initial vector $\mathbf{x}^{(0)} = (\mathbf{s}^{(0)}, \mathbf{h}^{(0)}, \beta^{(0)}, k^{(0)})$ be drawn from that distribution, or as in our case, from the posterior joint distribution $\pi\ (\mathbf{x} \mid \mathbf{t})$. Of course, this is not easily done, but the problem is usually dealt with by discarding the first b iterations, generally 10 per cent of ν (*burn-in* period), so that the next term in the chain can be assumed to have a distribution close to $\pi\ (\mathbf{x} \mid \mathbf{t})$. From the whole chain, we obtain the pointwise estimate of the conditional intensity function of the generalized Poisson process described in Section 3.1, or in other words, the posterior mean rate of occurrence:

$$\tilde{\lambda}(t) = 1/\nu \sum_{l=1}^{\nu} \sum_{j=0}^{k_{\max}} h_j^{(l)}\, \delta_{[s_j^{(l)}, s_{j+1}^{(l)})}(t). \tag{8}$$

This rate is shown in Fig. 4, together with the 5 and 95 per cent sampled quantiles at any time t; alternately higher and lower levels of seismic activity are clearly recognizable in the curve.

Using the sequence of values $\{k^{(l)}\}_{l=b+1}^{\nu}$, we estimate the probability function $\pi(k \mid \mathbf{t})$. The result is quite sharp, with 59 per cent of the probability concentrated on the values $k = 4, 5$: precisely $\pi(k = 4 \mid \mathbf{t}) = 0.333$ and $\pi(k = 5 \mid \mathbf{t}) = 0.257$. Hence, the posterior mode is $\hat{k} = 4$, while the posterior mean is $\tilde{k} = 4.79$. According to these results, M_4 and M_5 are the models that best describe the phenomenon. This means that the seismic activity could present four or five change-points and, consequently, five or six different phases. We examine both possibilities, estimating the posterior density functions $\pi(s_j \mid k)$, $j = 1, \ldots, k$, of the positions of the change-points and

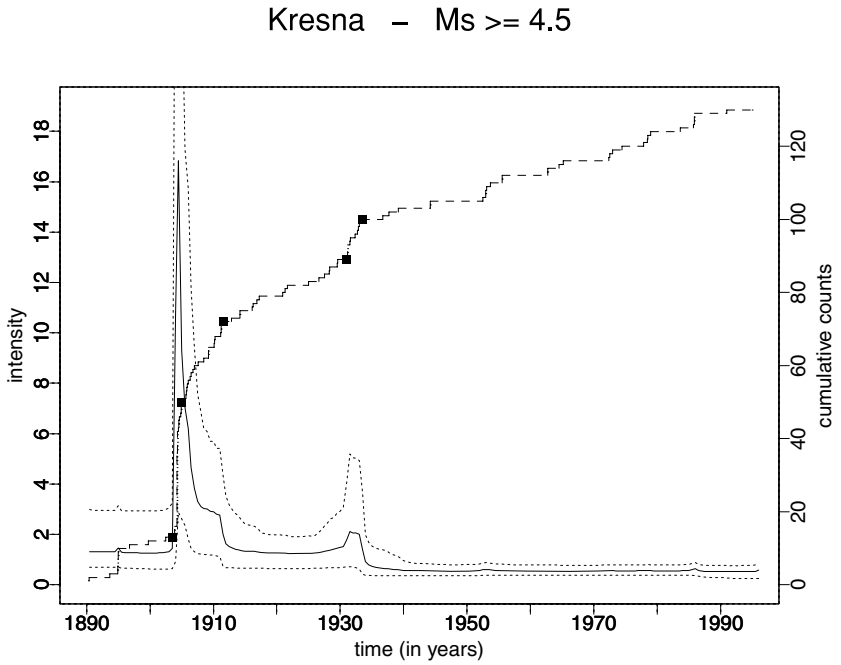

Figure 4. Ergodic average of the conditional intensity function (solid line), sampled 5 and 95 per cent quantiles (dotted lines) and cumulative counting process (dashed line). The solid squares denote the estimated positions of the change-points $s_1, \ldots, s_5$ on the cumulative curve.

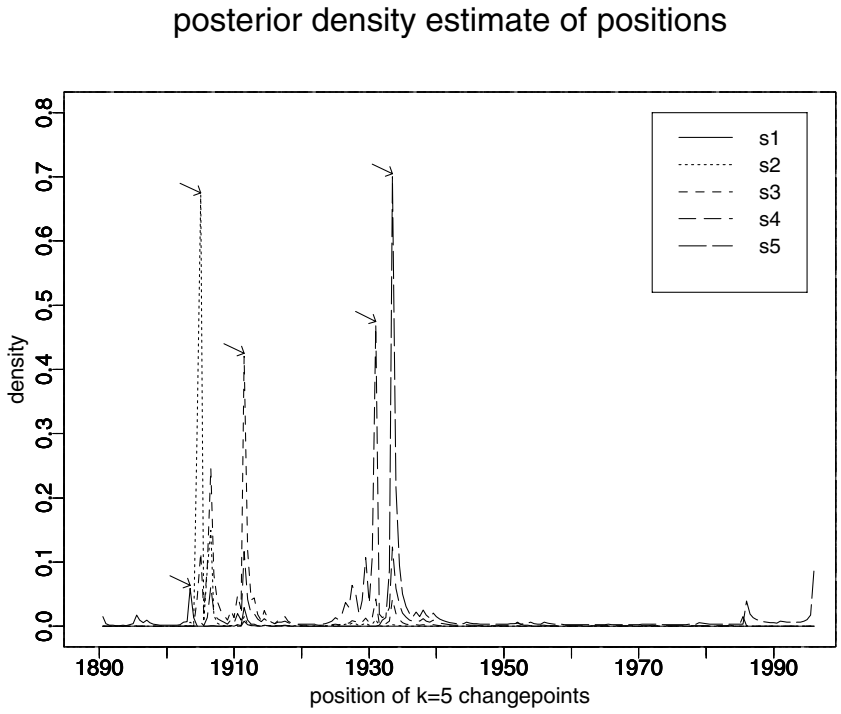

Figure 5. Posterior estimates of the density functions $\pi(s_j \mid k)$, $j = 1, \ldots, k$, of the positions of the change-points conditional on $k = 5$; the arrows indicate the value at the mode of each density function, that is at the estimated position of a change-point.

the posterior density functions $\pi(h_j \mid k)$, $j = 0, \ldots, k$, of the rates in the different phases, given both $k = 4$ and $k = 5$. The posterior densities of the location of the change-points when $k = 5$ are shown in Fig. 5, and in Fig. 6 when $k = 4$. Their piecewise exponential shape suggests that it is generally appropriate to adopt the posterior mode as an estimator of $s_1, \ldots, s_k$, since the posterior mean may be located at a point of very low probability.

In Figs 5 and 6, the arrows indicate the value of the mode of every density, that is, of the estimated position of each change-point. The posterior densities of the rates of segments of the step function (1) are shown in Fig. 7 for $k = 5$ and in Fig. 8 for $k = 4$.

The strong multimodality of the posterior densities $\pi(h_1 \mid k = 5)$ and $\pi(h_4 \mid k = 5)$ matches that of the densities $\pi(s_1 \mid k = 5)$ and $\pi(s_4 \mid k = 5)$; this issue, which arises when the variables are highly

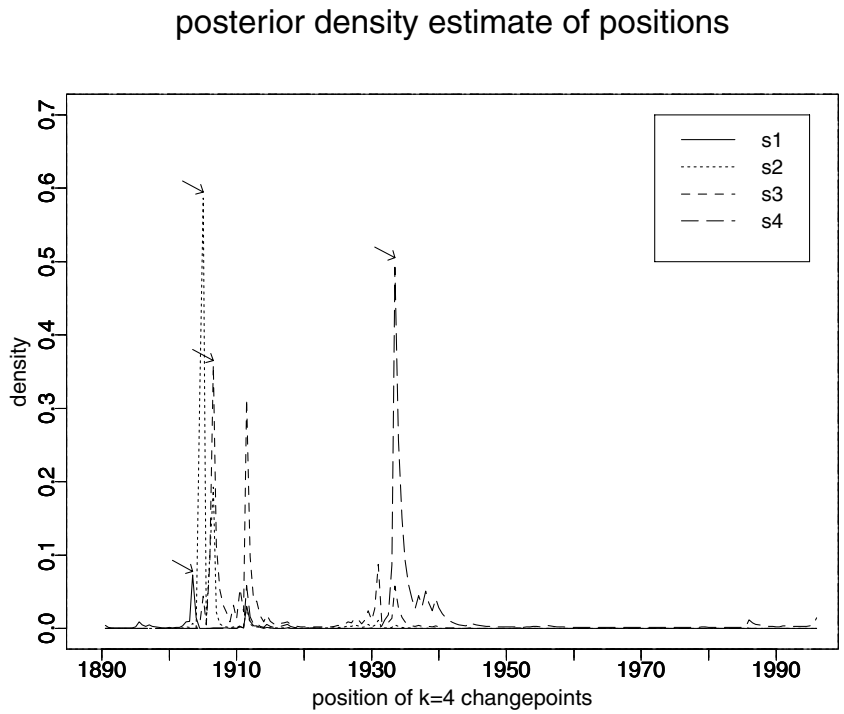

Figure 6. Posterior estimates of the density functions $\pi(s_j \mid k)$, $j = 1, \ldots, k$, of the positions of the change-points conditional on $k = 4$; the arrows indicate the value at the mode of each density function, that is at the estimated position of a change-point.

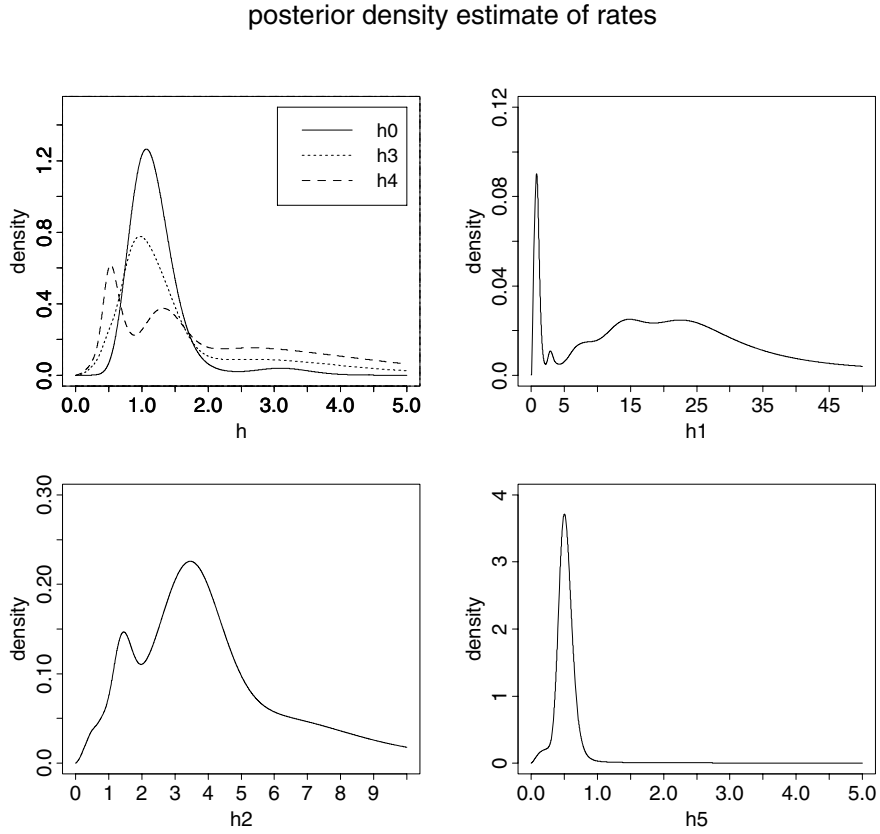

Figure 7. Posterior estimates of the density functions $\pi(h_j \mid k)$, $j = 0, \ldots, k$, of the rates conditional on $k = 5$ change-points.

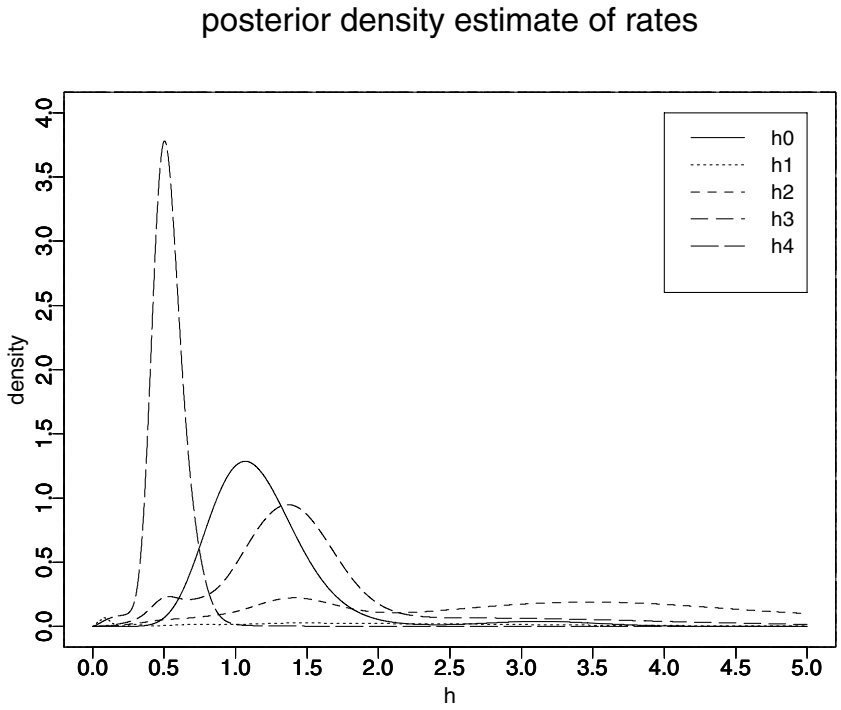

Figure 8. Posterior estimates of the density functions $\pi(h_j \mid k)$, $j = 0, \ldots, k$, of the rates conditional on $k = 4$ change-points.

Table 2. Estimates of the rates and of the locations of the change-points for the models M_4 and M_5 in the analysis of variations in the occurrence rate.

$\mathbf{k} = 5$	$\tilde{h}_0$	$\tilde{h}_1$	$\tilde{h}_2$	$\tilde{h}_3$	$\tilde{h}_4$	$\tilde{h}_5$
	1.26	37.67	6.46	2.04	2.32	0.59
	$\hat{s}_1$	$\hat{s}_2$	$\hat{s}_3$	$\hat{s}_4$	$\hat{s}_5$	
	1903.5	1905.	1911.5	1931.	1933.5	
$\mathbf{k} = 4$	$\tilde{h}_0$	$\tilde{h}_1$	$\tilde{h}_2$	$\tilde{h}_3$	$\tilde{h}_4$	
	1.23	38.16	4.63	1.66	0.54	
	$\hat{s}_1$	$\hat{s}_2$	$\hat{s}_3$	$\hat{s}_4$		
	1903.5	1905.	1906.5	1933.5		

correlated, requires careful use of the MCMC output in computing expectations. Table 2 reports the estimates of the posterior means $\tilde{h}_j$, $j = 0, \ldots, k$, and of the modes $\hat{s}_j$, $j = 1, \ldots, k$, for $k = 4$ and $k = 5$. As we have adopted a step equal to 0.5 yr for the pointwise estimate of the densities $\pi(s_\cdot)$ and $\pi(h_\cdot)$, the positions of the change-points are approximated with an error of less than half a year.

Comparing the results for M_5 and M_4, we see that the first two and the last change-points coincide, and the rates of the first two and the last phases are also very similar. In both models, the second phase (1903.5–1905) covers the pair of strong earthquakes in 1904; this interval shows an exceptional peak of occurrences, which one would expect to be followed by a period of frequent secondary shocks. In fact, in the period from 1904 to 1906, in addition to the two shocks of magnitude 7.1 and 7.8, there was one of $M_S = 5.9$ on 1904 April 19, and two of $M_S = 6.4$ on 1905 October 8 and on 1906 October 8. According to the formula—max (100, $10^{0.5\,M-1}$) in days—for the time span after a main shock of magnitude M given in Ogata *et al.* (1996), the combination of the triggering effect due to these large events should be felt longer than 1907.2. This agrees more closely with the position of the third change-point, 1911.5, in M_5. The two remaining change-points of this model, 1931 and 1933.5, bound the fifth phase, which is characterized by a higher rate than the adjacent phases, including, as it does, two strong earthquakes in 1931, the shock of magnitude $M_S = 6.4$ of 1932 September 29, and the events triggered by these. Moreover, observing Fig. 4, we see that the estimated rates for model M_5 change consistently with the jumps of the estimated conditional intensity function of the generalized Poisson process. In light of these considerations, we conclude that the model M_5 better describes the seismic pattern of the Kresna region.

In this study, we have assigned the following values to the hyperparameters: $a_0 = 2$; $c_0 = 3$ and, $d_0 = 0.5$. For the criterion of choice, see Rotondi (2002).

4.2 Results for variations in the interevent time distribution

The MCMC algorithm provides samples from the posterior distribution of $(k, r_1, \ldots, r_k, \theta_0, \ldots, \theta_k \mid \mathbf{t})$, which we can use to estimate the probability that a given interevent time, let us say for example, the ith y_i is the first of a new phase, by taking the observed relative frequency of the event $\{r_j = i, \forall\, j\}$ in the simulated sequence. Assuming a fixed k, we calculate this frequency by $\sum_{j=1}^{k} I(r_j = i)$ to obtain the estimate of the probability $\pi(i \mid k, \mathbf{t})$ that one of the k changes occurred at t_i. Marginalizing this distribution with respect to k, that is, varying k, we estimate $\pi(i \mid \mathbf{t})$ through $\sum_{k=1}^{k_{\max}} \sum_{j=1}^{k} I(r_j = i)$.

As in the rate-variation approach, we may first consider $\pi(k \mid \mathbf{t})$, whose mode is $\hat{k} = 8$ with $\pi(k = 8 \mid \mathbf{t}) = 0.130$, and then examine $\pi(i \mid \mathbf{t}, k = 8)$ as i varies in $\{2, \ldots, n\}$ (Fig. 9), in search of candidate positions for change-points. The eight highest values of $\pi(i \mid \mathbf{t}, k = 8)$ correspond to $i = 10$, 14, 17, 32, 63, 72, 90 and 93; the first interevent time of each phase, $y_i = t_{i+1}$-t_i, starts, therefore, with the earthquake that occurred at the date: $t_{10} = 1895/01/07$, $t_{14} = 1903/11/25$, $t_{17} = 1904/04/04$, $t_{32} = 1904/04/04$,

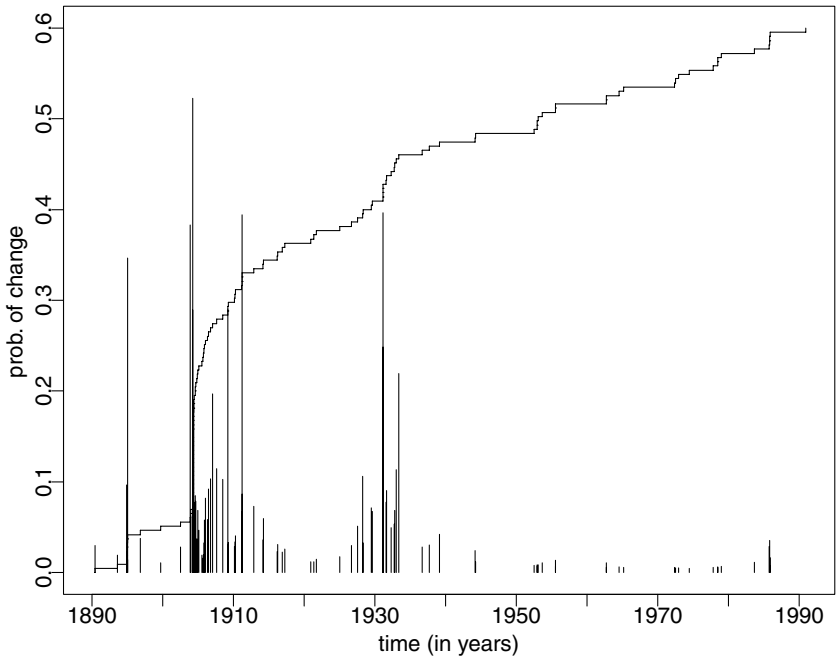

Figure 9. Cumulative number of earthquakes. Bars show the posterior probabilities that a change has occurred at t_i, $i = 2, \ldots, n$, for number of change-points $k = 8$.

$t_{63} = 1909/03/10$, $t_{72} = 1911/03/16$, $t_{90} = 1931/03/07$ and $t_{93} = 1931/03/08$. We note that less than 4 hr separate the shock of $M_S = 7.1$ at t_{17} from that at t_{32}, and the quakes in 1931 March were little over a day apart.

In this study, we have assigned to the parameters of the prior distributions the following values: $a_0 = 3$, $b_0 = 1$, $c_0 = 18$, $d_0 = 11$, $g_0 = 1.25$, $h_0 = 11$, $f_0 = 8$ and $M = 260$; for the criterion of choice, we refer to Pievatolo & Rotondi (2000).

We observe the following regarding these initial results.

(1) Intervals between change-points are sometimes so short that they cannot be plausibly held to possess the physical significance of seismicity phases as defined in the introduction. This happens, as we shall also see, in correspondence to the peak of activity immediately after a strong earthquake. In fact, the times between consecutive aftershocks could be better modelled as self-exciting rather than as renewal processes because, in general, they are very short at the beginning of the sequence and lengthen as the time after the main shock increases. However, we hold that even so, the same stochastic process can appropriately describe all the phases when, as in our case, the objective is a compact summary of the seismic activity. Therefore, although we maintain the subdivision produced by the algorithm, we do not interpret these time intervals as seismicity phases.

(2) It is possible that the positions of the k change-points with the highest marginal posterior probability $\pi(i \mid \mathbf{t})$ are never generated in the same iteration by the simulation algorithm. This can be checked by memorizing the vector of indices r_j, $j = 1, \ldots, k$, which identifies the positions of the change-points in each iteration. Given k, the total number of these vectors is $\binom{n-1}{k}$, the number of combinations of $n - 1$ elements of class k, which, for $n = 130$ and $k = 8$, is of the order of 10^{12}. Since the number of possible cases is so large, the frequencies of observed vectors of indices will probably be small and will not differ significantly from one another. Consequently, a clear-cut choice of the best positions for the change-points cannot be made on the basis of these frequencies.

An alternative strategy is to examine the plot in Fig. 10, indicating, at each t_i, the probability $\pi(i \mid \mathbf{t})$ that it is a change-point, and estimating as positions of the change-points the t_i's associated with the values $\pi(i \mid \mathbf{t})$ exceeding a given threshold. We must choose between two opposite random events: t_i is either a change-point, or it is not. The Bayes factor can be a criterion for deciding which of these two alternatives is more supported by the data. It is given by the ratio of the posterior probabilities of the two options divided by the ratio of their prior probabilities:

$$BF = \frac{\hat{\pi}(i \mid \mathbf{t})}{1 - \hat{\pi}(i \mid \mathbf{t})} \Big/ \frac{n}{n-2}.$$

Jeffreys (1961) has proposed a scale to assess the significance of the Bayes factor according to which $BF > 0.316$ indicates slight evidence against the first alternative, t_i location of a change-point. As this inequality corresponds to $\hat{\pi}(i \mid \mathbf{t}) > 0.243$, we adopt this value as the threshold and obtain the following dates: $t_{10} = 1895/01/07$, $t_{14} = 1903/11/25$, $t_{17} = 1904/04/04$, $t_{32} = 1904/04/04$, $t_{63} = 1909/03/10$, $t_{72} = 1911/03/16$, $t_{90} = 1931/03/07$, $t_{93} = 1931/03/08$ and $t_{100} = 1933/05/08$, corresponding to the vertical lines in Fig. 10 that top the other ones.

We have $k = 9$ t_i, which identify ten intervals. By roughly estimating the stress release as a function of magnitude $S \propto 10^{0.75 M_S}$ (Vere-Jones 1978), and the seismicity rate of each interval $[t_{r_j}, t_{r_{j+1}})$, $j = 0, \ldots, 9$, we observe the following about the seismicity in these intervals.

$[\mathbf{t_{17}}, \mathbf{t_{32}})$ and $[\mathbf{t_{90}}, \mathbf{t_{93}})$: They cover about 1 d and are characterized by an extremely high stress release, four to five orders higher than that in the other intervals; they are, therefore, 'peaks of activity', globally related to the two strongest crises in the sequence, which, for their length, are not considered phases.

$[\mathbf{t_{10}}, \mathbf{t_{14}})$ and $[\mathbf{t_{14}}, \mathbf{t_{17}})$: They precede the shock of $M_S = 7.8$ and are characterized by a general increase in activity: in the former interval, the average stress release per event increases whereas, in the latter, the seismicity rate and the stress release rate increase. In brief, we could say that there are more events in the latter interval,

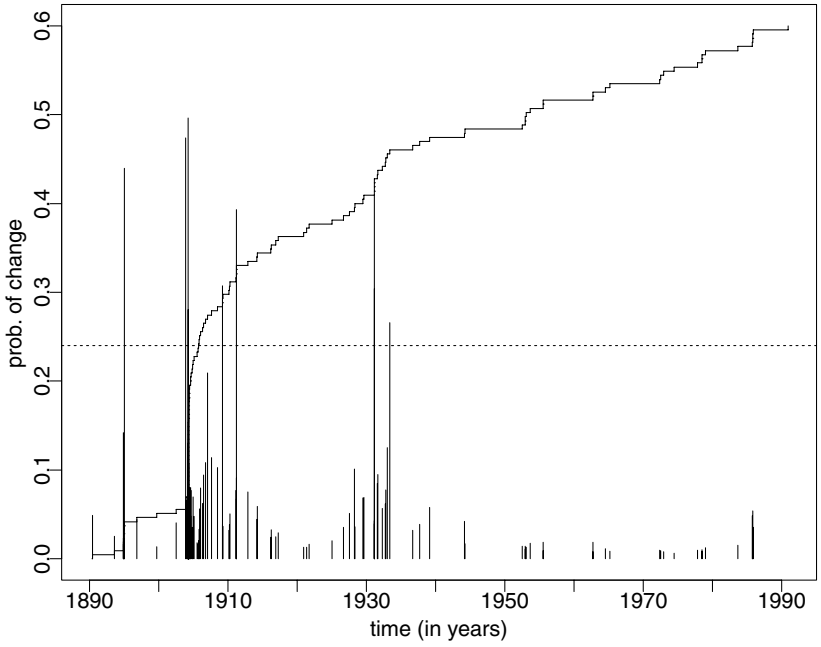

Figure 10. Cumulative number of earthquakes. Bars show the posterior probabilities that a change has occurred at t_i, $i = 2, \ldots, n$, not conditional on the number k of change-points.

but they are stronger in the former. We consider these intervals components of a single 'foreshock phase' with different features.

$[\mathbf{t_{32}}, \mathbf{t_{63}})$ and $[\mathbf{t_{63}}, \mathbf{t_{72}})$: They follow the shock of $M_S = 7.8$ in 1904 and, contrary to the above-mentioned intervals, are characterized by a growing decrease in the activity, which agrees with the lengthening of the interevent times that characterize self-exciting processes, more suitable to fit the aftershock activity. Therefore, we associate these two intervals with the monotonic trend of the same 'aftershock phase'.

$[\mathbf{t_{93}}, \mathbf{t_{100}})$: It follows the earthquake of $M_S = 6.7$ in 1931. In agreement with the Ogata *et al.* formula (1996) for the time span after a main shock, the combination of the triggering effect due to the quakes of 1931/03/07 ($M_S = 6.0$), 1931/03/08 ($M_S = 6.7$) and 1932/09/29 ($M_S = 6.4$) is felt up to the first months of 1933. We consider this interval another 'aftershock phase'.

$[\mathbf{t_1}, \mathbf{t_{10}})$, $[\mathbf{t_{72}}, \mathbf{t_{90}})$, $[\mathbf{t_{100}}, \mathbf{t_{130}})$: Moving t_1 back to 1874/05/05, the date of the last event of the incomplete part of the catalogue, we find very similar values, the lowest in the sequence, for the stress release rates and the seismicity rates in these three intervals; hence, we consider them 'background phases'.

The final subdivision into seismicity phases could then be as follows: $t_{10} = 1895/01/07$, $t_{14} = 1903/11/25$, $t_{17} = 1904/04/04$ ($t_{32} = 1904/04/04$), $t_{63} = 1909/03/10$, $t_{72} = 1911/03/16$, $t_{90} = 1931/03/07$ ($t_{93} = 1931/03/08$) and $t_{100} = 1933/05/08$. The times in parentheses are considered the beginning of the next phases.

These results do not contradict those given in Section 4.1. On the contrary, they provide a clearer and more detailed description of the activity. For instance, the increase in seismicity preceding the very strong 1904 earthquake is signalled first by the increase in the stress release and then, by the increase in the number of events, whereas the aftershock sequence, from 1904 to 1911, is divided into two segments with a decreasing rate.

Having fixed the number and the position of the change-points, we estimate pointwise the distribution functions of the interevent time in each of the eight phases we have identified. By exploiting the simulated Markov chains for the parameters $\theta_i = (\eta_i, \xi_i, \rho_i)$, we evaluate the *a posteriori* expected value of F_i through the following relationship deduced from (5):

$$E\{F_i(\cdot) \mid \theta_i, r_i, r_{i+1}, \mathbf{y}\} = \frac{M\, G(\cdot;\, \theta_i) + \sum_{j=r_i}^{r_{i+1}-1} \delta_{y_j}(\cdot)}{M\, +\, r_{i+1} - r_i}, \qquad (9)$$

where G is the generalized gamma cumulative distribution and $\delta_y(t) = 1$ when $t = y$. Fig. 11 shows these estimates (solid line) with the box plots of the cumulative probability at a suitable grid of points as measures of the estimate uncertainty. The dashed lines represent the plug-in estimates of the generalized gamma distribution obtained by substituting for the parameters η, ξ and ρ, their posterior means given by the ergodic mean of the corresponding Markov chains. In the same way, Fig. 12 shows the plug-in estimates of the generalized gamma density functions (solid line) and of the corresponding hazard functions (dashed line) in the eight phases illustrated in this section.

The curves in Fig. 12 are further results of our analysis. We remind the reader that we are interested in their trend on a finite timescale of the same order as the corresponding phase. For instance, Phase 2 lasts only 4 months whereas the curves for density and hazard cover 10 yr. From this prospective, we may observe that: Phase 1, 8 yr long, has a clearly increasing hazard and unimodal density function that differs from all the others. Phase 5, 20 yr long, shows almost constant hazard as expected in a Poisson model and the increasing hazard function of Phase 2 (2 yr long) differs in behaviour from that expected in typical aftershock periods such as those indicated in Phases 3 and 4.

5 COMPARISON WITH THE ZAMP SOFTWARE PACKAGE

We have also tried to solve the issues arising here by using one of the available software packages. Tens of publications that made use of the ZMAP package (http://www.seismo.ifg.ethz/staff/stefan) to identify and evaluate spatial and temporal variations in seismicity led us to choose that software. ZMAP (Wiemer 2001) is a set of

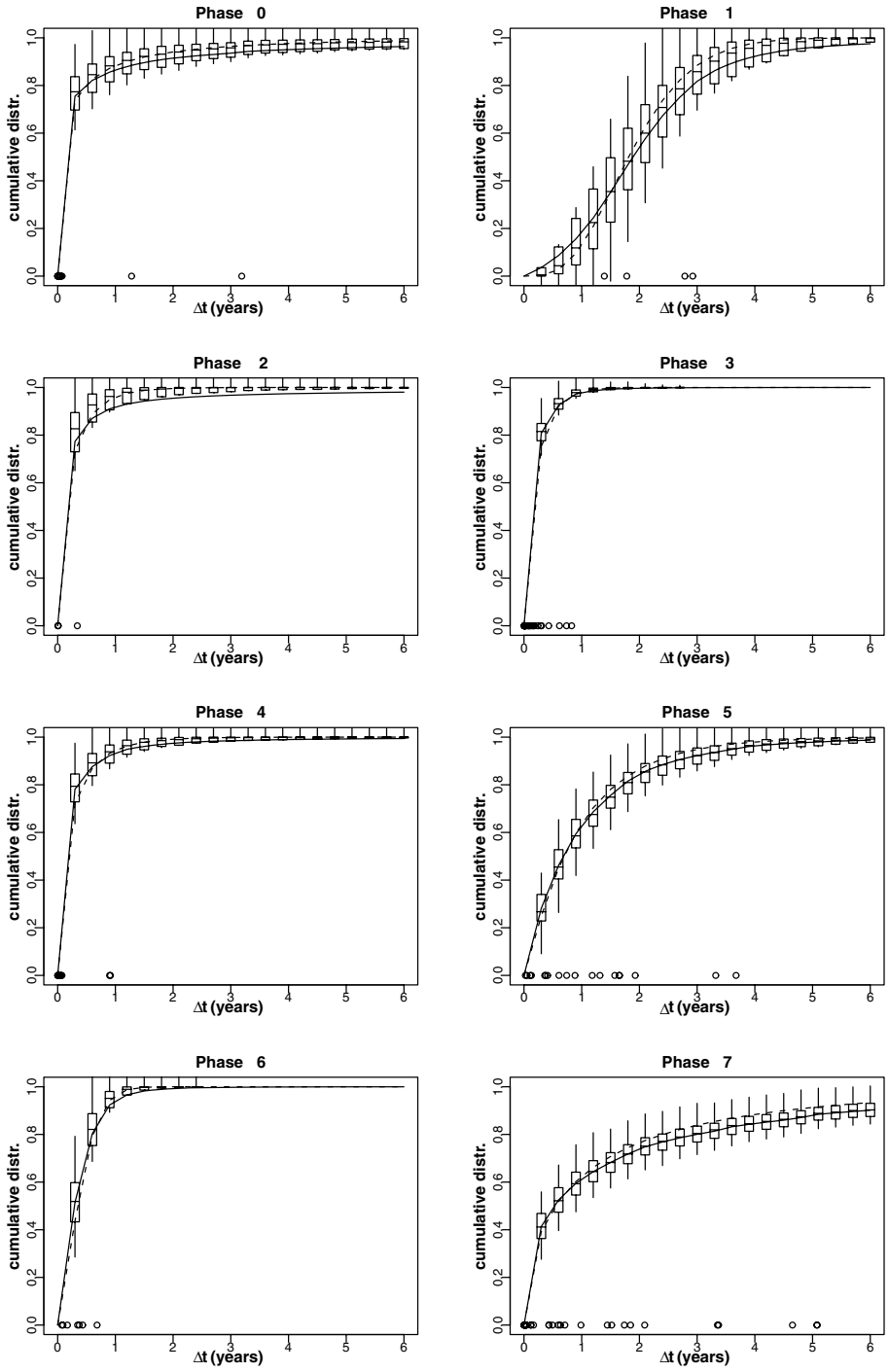

Figure 11. MCMC estimates of the cumulative distribution functions in the eight detected regimes: box plots; averaged conditional expected values of DP (solid line) and plug-in estimates of the generalized gamma distribution with the posterior means of η, ξ and ρ (dashed line).

tools, driven by a graphical user interface, that are designed to help seismologists analyse catalogue data. It was first published in 1994 and has continued to grow over the past years, up to the most recent version, ZMAP v. 6. Rate changes are analysed using the function termed *AS(t)* (Habermann 1988); it is the *z*-value

$$z = \frac{M_1 - M_2}{\left(\left(S_1^2/N_1 + S_2^2/N_2\right)\right)^{1/2}}, \tag{10}$$

resulting from comparison of the mean rates M_1 and M_2 during the two periods (t_0, t) and (t, t_f), where t moves from the extremes $(t_0 + t_w)$ and $(t_f - t_w)$, S_1 and S_2 are the standard deviations of those rates, N_1 and N_2 are the number of samples in the two periods, and t_w is the window length. The z statistic (10) is approximately normally N(0,1) distributed (i.e. $z = 1.96$ is 95 per cent significant, and $z = 2.57$ is 99 per cent significant). The maximum of the function *AS(t)* is usually employed to detect a possible anomalous period of quiescence before a strong earthquake, indicated by a high *z*-value, which denotes rate decrease. As we are interested in both rate increases and decreases, we have iteratively analysed the data set storing, at each sweep, the time t_a corresponding to the

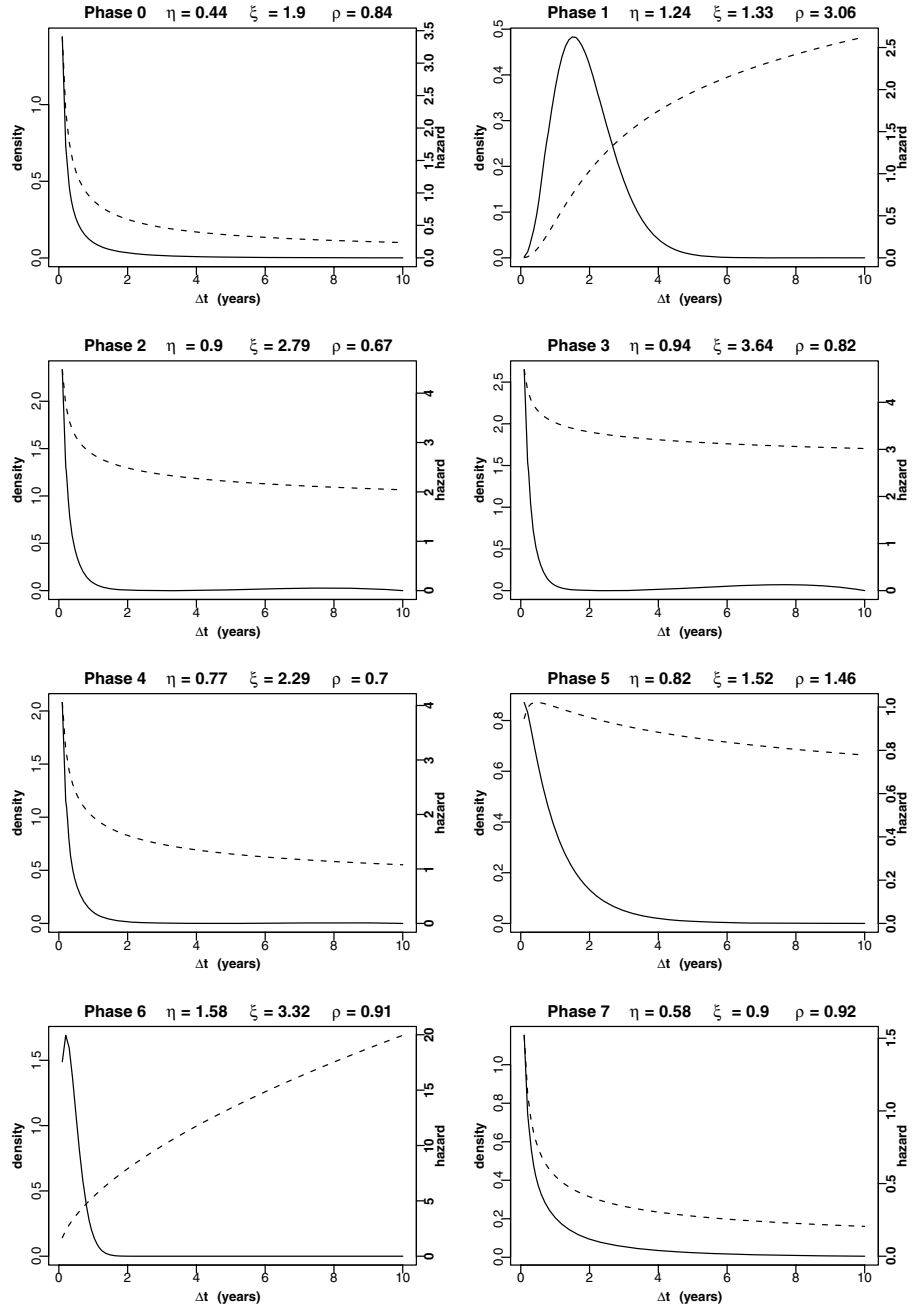

Figure 12. Plug-in estimates of the generalized gamma density functions (solid line) and of the corresponding hazard functions (dashed line) in the eight phases illustrated in Section 4.2.

maximum absolute z-value. At the first step, we set $t_a \rightarrow \mathrm{t}^{(1)}$ and we obtained the binary partition $(t_0, t^{(1)})$ and $(t^{(1)}, t_f)$ of the interval (t_0, t_f); in the subsequent step we evaluated the function *AS(t)* in the two subintervals and we memorized the two absolute maxima $t^{(2)}$ and $t^{(3)}$ obtaining 2^2 subintervals. We repeated this procedure as long as significant rate changes were found (i.e. $|z| \geq 1.96$). The results of this analysis are reported in Table 3; the parameters used are 1 yr for the window length and ten days for the bin length. Further trials performed with 0.5–1.5 yr and 15–28 d, respectively, produced similar results. To quantify each change, we chose the option 'compare two rates' from the Ztools menu in the cumulative number window of ZMAP and compared the b-values during the period before and after each change-point, in light of the probability estimated according to Utsu (1992) that the two samples come from the same indistinguishable population of magnitudes. This hypothesis is rejected with over 0.95 significance (that is there is also a change in the magnitude distribution) at $t = 1894.88$, 1903.89 and 1933.34; the remaining change-points fail the test. Only by adding the point 1927.54, with z-value -1.76, in the interval (1911.20, 1933.34), do both the change-points, 1911.20 and 1927.54, present a change in magnitude.

We have also applied another procedure to define the times of greatest change. First, we evaluated the *AS(t)* function in the entire interval and stored the times corresponding to both the maximum and minimum z-value independently of the significance of z (in fact, all the minima are negative and larger than -1.96), and then we

Table 3. Significant rate changes provided by ZMAP software:for each iteration of the binary partition of the interval (1890, 1995) we report the time of the rate change and the corresponding *z*-value (in brackets).

Iteration	*t* (min AS)	*t* (max AS)
1		1985.89 (3.81)
2		1933.34 (3.76)
3	1894.88 (−2.79)	
4		1911.20 (2.31)
5	1903.89 (−2.42)	
6		1907.02 (2.19)

Table 4. Rate changes corresponding to the minimum and maximum *z*-value (in brackets) provided by ZMAP software applied in telescopic intervals.

Step	*t* (min AS)	*t* (max AS)
1	1894.88 (−1.39)	1985.89 (3.81)
2	1903.89 (−0.65)	1933.34 (3.27)
3	1931.16 (−0.87)	1911.22 (2.81)

repeated this step in the interval having those times as extremes until the *AS(t)* function showed no significant rate change. The results are found in Table 4. In the second step, we neglected the global minimum 1984.82 [$AS(1984.82) = -0.76$], almost the same as the previously-found value for 1985.89. As for the *b*-values, the change-points $t = 1894.88$, 1903.89, 1911.22, 1931.16 and 1933.34 reject, with high significance, the hypothesis that the two sets of observed magnitudes come from the same population, indicating a change in the magnitude distribution.

In summary, the two procedures jointly identify the following dates as change-points: 1894.88, 1903.89, 1907.02, 1911.20, 1931.16 and 1933.34. Consequently, there is substantial agreement between the analysis performed by ZMAP and that based on variations in the distribution of interevent time. There are, however, some important differences: (a) ZMAP sets the foreshock phase a year ahead and fails to identify the interval (1903–1904) with a rate higher than that in (1895–1903); (b) it also does not detect the peaks of activity in 1904 and 1931 and (c) ZMAP marks the aftershock interval following the 1904 earthquake as ending even sooner than the date given by the Ogata *et al.* formula (1996).

6 CONCLUSION

Using statistical procedures, we have answered the questions asked in Section 1: Are there changes in the seismic activity? How many? When? We have produced an objective decomposition of the observed seismicity into phases on the basis of variations in these characteristic features: occurrence rate and probability distribution of interevent time. Other physical features, such as the distribution of the size of the events, could be similarly analysed. An inherent limitation of the proposed method is that interevent times are conditionally independent and identically distributed within every phase, while this is not always true, particularly during aftershock sequences. The disadvantage of this simplified modelling is mainly that more phases than those really existing may be estimated in aftershock intervals, since the progressive change in the expected length of interevent times, corresponding, for instance, to a power-law decay rate, is fitted through a sequence of different interevent time distributions. However, knowing this, a user should be able to ascribe to these artificially created phases the same physical significance and obtain a consistent sequence of patterns.

The analysis of rate variations provides preliminary results that are clarified by the use of interevent time distributions. Moreover, the estimates of the probability distributions, the densities, and the hazard functions (Figs 11 and 12) provide a physical interpretation of the different phases, which add information useful for assessing time-dependent seismic hazard. Furthermore, the box plots in Fig. 11 allow us to quantify the uncertainty of the estimates, as is done in Figs 5–8 for the variations in rate.

The two procedures applied using ZMAP software almost completely confirm our results regarding the detection of the change-points if we take into account points of lower significance level, such as 1931.16 in Table 4. Moreover, the ZMAP software is not equipped to deal with questions regarding the interevent time probability distribution. We also note that the two-sample test performed by ZMAP is based on the assumptions, not always found in real situations, that the two compared populations are normally distributed with known variances, or that the size of the samples is sufficiently large.

In conclusion, the estimates of the rates (Table 2, Fig. 4), interevent time distributions, densities and hazard functions (Figs 11 and 12) produce a global and objective description of the phases identified, as follows.

(i) (1890–1895)—The seismic activity is low, the rate is about 0.5; we classify this phase as the final part of a 'background' activity period.

(ii) (1895–1903), (1903–1904)—At first, only the stress release rate rises, but then the occurrence rate rises as well; an activation period starts with a 'foreshock' of $M_S = 6.6$ in 1902 July.

(iii) (1904–1909), (1909–1911)—The 'aftershock' sequence continues with a series of secondary shocks, of $M_S = 6.4$ in 1905 October, and again, 1 yr later, of $M_S = 6.4$. The global occurrence rate remains high, but the activity decreases progressively during this phase.

(iv) (1911–1931)—Although a large amount of accumulated stress was released in the previous phase, the rate is at an intermediate level; no event, however, exceeds a magnitude of 5.

(v) (1931–1933)—After the two main shocks of $M_S = 6$–6.7 in 1931 March and a secondary shock of $M_S = 6.4$ in 1932 September, this seismic crisis reproduces that of 1904–1911 on a reduced time-size scale.

(vi) (1933–1995)—The occurrence rate has returned to the low rate registered at the beginning of the period under examination.

The time period we have considered presents two seismic cycles which, even if of different strengths, are characterized by similar patterns: background activity (i, iv and vi), foreshocks (ii), main earthquakes (iii), and aftershocks sequence (iii and v), corresponding to the stages indicated by Zhuang (2000): interseismic, pre-seismic, coseismic and post-seismic. The pre-seismic phase may be absent, not clearly identifiable, or confused with the interseismic one. This may mean that the evolution in time is determined by hidden random factors, which constitute a higher level in the hierarchical structure of the stochastic process modelling the earthquake process.

Our future research will attempt to give a probabilistic structure for the assignment of the system to a specific state and to combine this with the probability that an event occurs when the system is in that state.

ACKNOWLEDGMENTS

This work is the continuation of research carried out within the framework of the International Cooperation-COPERNICUS project

'Assessment of seismic potential of European large earthquake areas' supported by the European Union Directorate General XII, contract ERBIC-15CT-97-0200. The authors wish to thank the editor and the two referees for their accurate and constructive remarks that have greatly contributed to improving the paper. The authors are also grateful to Dr Gospodinov for the helpful discussions on the seismotectonic properties of the Kresna region. The map in Fig. 1 was produced using the GMT software (Wessel & Smith 1998).

REFERENCES

Ambraseys, N.N., 2001. The Kresna earthquake of 1904 in Bulgaria, *Ann. Geofis.,* **44,** 95–117.

Antoniak, C.E., 1974. Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems, *Ann. Stat.*, **2,** 1152–1174.

Dineva, S., Sokerova, D. & Mihaylov, D., 1998. Seismicity of South-Western Bulgaria and border regions, *J. Geodyn.*, **26**(2–4), 309–325.

Dineva, S., Sokerova, D. & Mihaylov, D., 1999. *Catalog of earthquakes in SW Bulgaria and border regions (52 to 1995),* manuscript, Geophysical Institute of Bulgarian Academy of Sciences.

Dineva, S., Batllo, J., Mihaylov, D. & van Eck, T., 2002. Source parameters of four strong earthquakes in Bulgaria and Portugal at the beginning of the 20th century. *J. Seism.*, **6,** 99–123.

Evison, F.F. & Rhoades, D.A., 1998. Long-term seismogenic process for major earthquakes in subduction zones, *Phys. Earth planet. Inter.*, **108,** 185–199.

Evison, F.F. & Rhoades, D.A., 1999. The precursory earthquake swarm in Japan: hypothesis test, *Earth Planets Space*, **51,** 1267–1277.

Evison, F.F. & Rhoades, D.A., 2001. Model of long-term seismogenesis, *Ann. Geophys., 44*, **1,** 81–93.

Ferguson, T.S., 1973. A Bayesian analysis of some parametric problems, *Ann. Stat.*, **1,** 209–230.

Goltz, C., 2001. Decomposing spatio-temporal seismicity patterns, *Nat. Hazards Earth Sys. Sci.*, **1,** 83–92.

Green, P.J., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrika*, **82**(4), 711–732.

Habermann, R.E., 1988. Precursory seismic quiescence:past, present and future, *Pure appl. Geophys.*, **126**(2–4), 279–318.

Jeffreys, H., 1961. *The Theory of Probability,* 3rd edn, Clarendon Press, Oxford.

Kaufman, L. & Rousseeuw, P.J., 1990. *Finding Groups in Data: An Introduction to Cluster Analysis,* John Wiley & Sons, Inc., New York.

Matsu'ura, R.S., 1986. Precursory quiescence and recovery of aftershock activities before some large aftershocks, *Bull. Earthq. Res. Inst.* University of Tokyo, **61,** 1–65.

Milutinovic, Z., 1998. Seismic hazard and countermeasures in the city of Skopie, available at http://iisee.kenken.go.jp/net/macedonia/SeismicityReport.html

Ogata, Y., 1989. Statistical model for standard seismicity and detection of anomalies by residual analysis, *Tectonophysics*, **169,** 159–174.

Ogata, Y., 1992. Detection of precursory relative quiescence before great earthquakes through a statistical model, *J. geophys. Res.*, **97,** 19 845–19 871.

Ogata, Y., 2001. Increased probability of large earthquakes near aftershock regions with relative quiescence, *J. geophys. Res.*, **106**(B5), 8729–8744.

Ogata, Y., Utsu, T. & Katsura, K., 1995. Statistical features of foreshocks in comparison with other earthquake clusters, *Geophys. J. Int.*, **121,** 233–254.

Ogata, Y., Utsu, T. & Katsura, K., 1996. Statistical discrimination of foreshocks from other earthquakes clusters, *Geophys. J. Int.*, **127,** 17–30.

Pievatolo, A. & Rotondi, R., 2000. Analysing the interevent time distribution to identify seismicity phases: a Bayesian nonparametric approach to the multiple-changepoint problem, *Appl. Stat.*, **49**(4), 543–562.

Rotondi, R., 1999. Statistical analysis of temporal variations of seismicity level in some Italian regions, *Natural Hazards*, **19**(2–3), 139–150.

Rotondi, R., 2002. On the influence of the proposal distributions on a reversible jump MCMC algorithm applied to the detection of multiple change-points, *Comput. Stat. Data Anal.*, **40**(3), 633–653.

Rotondi, R. & Garavaglia, E., 2002. Statistical analysis of the completeness of a seismic catalogue, *Natural Hazards*, **25**(3), 245–258.

Shanov, S., Boykova, A. & Stoev, D., 1999. Geophysical model of the coseismic rupture from the april 4th 1904, Krupnik earthquake, M = 7.8, S.W. Bulgaria, in *Proceedings of the 2nd Balkan Geophys. Congress, Abstract,* 232–233, Istanbul.

Simeonova, S.D., Solakov, D.E., Leydecker, G., Busche, H., Schmitt, T. & Kaiser, D., 2006. Probabilistic seismic hazard map for Bulgaria as a basis for a new building code. *Nat. Hazards Earth Sys. Sci.*, **6,** 881–887.

Tranos, M.D., Karakostas, V.G., Papadimitriou, E.E., Kachev, V.N., Ranguelov, B.K. & Gospodinov, D.K., 2006. Major active faults of SW Bulgaria: implications of their geometry, kinematics and regional active stress regime, in *Tectonic Development of the Eastern Mediterranean Region,* Vol. 260, pp. 671–687, eds Robertson, A.H.F. & Mountrakis, D., Geological Society, London, Special Publications.

Tsapanos, T.M., Vlachos, S.D. & Georganta, H.Ch., 2002. The seismic potential assessment for some seismogenic sources in the central northern Greece and its adjacent areas based on the maximum credible magnitude, *J. Balkan geophys. Soc.*, **5**(4), 123–130.

Utsu, T., 1984. Estimation of parameters for recurrence models, *Bull.Earthq. Res. Inst.*, **59,** 53–66.

Utsu, T., 1992. On seismicity, in: Report of the Joint Research Institute for Statistical Mathematics, Institute for Statistical Mathematics, Tokyo, p. 139–157.

Van Eck, T. & Stoyanov, T., 1996. Seismotectonic and seismic hazard modelling for Southern Bulgaria, *Tectonophysics*, **262,** 77–100.

Vere-Jones, D., 1978. Earthquake prediction: a statistician's view. *J. Phys. Earth,* **26,** 129–146.

Wang, J.-H. & Kuo, C.-H., 1998. On the frequency distribution of interoccurrence times of earthquakes, *J. Seism.*, **2,** 351–358.

Wessel, P. & Smith, W.H.F., 1998. New, improved version of Generic Mapping Tools released. *EOS, Trans. Am. geophys. Un.*, **79** (47), pp. 579.

Wiemer, S., 2001. A software package to analyze seismicity: *ZMAP, Seismological Research Letters*, **72,** 2, 374–383.

Wyss, M. & Wiemer, S., 2000. Change in the probability for earthquakes in Southern California due to the Landers magnitude 7.3 earthquake, *Science*, **290,** 1334–1338.

Zhuang, J., 2000. Statistical modelling of seismicity patterns before and after the 1990 Oct 5 Cape Palliser earthquake, New Zealand, *New Zealand J.Geol. Geophys.*, **43,** 447–460.

Zöller, G., Hainzl, S., Ben-Zion, Y. & Holschneider, M., 2007. Critical states of seismicity: from models to practical seismic hazard estimates, in *Encyclopedia of Complexity and System Science*, Springer-Verlag, Heidelberg.

APPENDIX A: MCMC METHODS

We give only the main points of the MCMC methodology. Suppose that one wishes to sample from a distribution $\pi(dx)$ for $x \in X$, in general the posterior distribution $p(x \mid y)$ in Bayesian framework, and also suppose that it cannot be done directly. If it is possible to construct a Markov chain of state space X with equilibrium distribution $\pi(dx)$, then, under suitable regularity conditions, asymptotic results allow us to estimate features of $\pi(dx)$. In particular, if $X^{(1)}$, $X^{(2)}, \ldots, X^{(t)}, \ldots$ is a realization from such a chain, it can be proven that

$$X^{(t)} \to^{d}_{t\to\infty} X \approx \pi(x)\,, \quad \frac{1}{t}\sum_{i=1}^{t} f(X^{(i)}) \to_{t\to\infty} E_{\pi}\{f(X)\} \quad a.s.. \tag{A1}$$

The first result emphasizes that the chain must be run for a sufficiently long time for the sampled values $X^{(t)}$ to be considered drawn from π. The second result shows that, in this way, it is possible to compute any property of $\pi(dx)$ that can be represented in terms of expectation under π, hence the existing moments and even the probabilities of specified events when f is an indicator function. Constructing a Markov chain means assigning a transition kernel P

so that $P(x \to \mathrm{x}')$ is the probability associated with the move from the current state x to the new one $x\prime$. P must be chosen so as to guarantee the convergence of the chain to π, taking into account various issues: the speed of exploration of the support of π, the degree of correlation between the $X^{(i)}$ values and the accuracy of approximation by the empirical average (A1). The choice of P characterizes the MCMC algorithm; in Gibbs sampling, the kernel is given by the product of the full conditional distributions, whereas in Metropolis–Hastings type algorithms, the potential new state or 'proposal' x' is generated from an essentially arbitrary proposal distribution q (x, $dx\prime$), and $x\prime$ is then accepted with the probability

$$\alpha(x, x') = \min\left\{1, \frac{\pi(dx')q(x', dx)}{\pi(dx)q(x, dx')}\right\}, \tag{A2}$$

or else x is retained as the next state.

If, as in the case of the multiple change-point problem, the model adopted belongs to the class of variable dimension models, the MCMC algorithm must generate a chain that moves into a finite or denumerable collection of parameter subspaces of different dimensionality. This can be done through a more advanced algorithm; the reversible jump algorithm (Green 1995), referring to the formulation given for the mixture of Poisson processes in Section 3.1, can be described as follows. Assuming that the Markov chain is at the state (model) M_k, different types of moves m can be made in which, respectively, (1) the parameters of the model M_k are updated; (2) a further change-point is added or (3) one of the present change-points is removed. At each transition, one of these moves is chosen according to a suitable probability distribution. In the first type of move, the dimension of the parameter vector ($\mathbf{s}$, $\mathbf{h}$, β, k) is fixed; hence an ordinary MCMC method can be applied. In particular, we have used the Gibbs sampling as, once we have chosen the priors for β and h_j in conjugate families of distributions, it is easy enough to sample from the conditional distributions. Instead, other types of moves change the dimension of the parameter vector, and this gives rise to the issue of the existence of a common dominating symmetric measure, with respect to which the probabilities in the ratio in eq. (A2) have densities. This problem can be overcome by applying the reversible jump MCMC method as follows. Supposing that move m takes $x = (\mathbf{s}, \mathbf{h}, \beta, k)$ to $x' = (\mathbf{s}', \mathbf{h}', \beta', k + 1)$ in a higher-dimensional space, a continuous random vector u is drawn from a distribution with density $q(u)$ with respect to the Lebesgue measure and the new state is given by an invertible deterministic function $x\prime\,(x, u)$. Then, the acceptance probability (A2) becomes

$$\alpha(x, x') = \min\left\{1, \frac{p(x' \mid y)}{p(x \mid y)} \times \frac{p_m(x')}{p_m(x)q(u)} \times \left|\frac{\partial x'}{\partial(x, u)}\right|\right\},$$

or, informally $\alpha(x, x') = \min\{1, (\textit{posterior ratio}) \times (\textit{proposal ratio}) \times (\textit{Jacobian})\}$, where $p_m(\cdot)$ is the probability of move m, and the last term is a Jacobian ensuing from the change of variable. Obviously, the dimension of $x\prime$ must be equal to the sum of the dimensions of x and u.

APPENDIX B: STATISTICAL COMPLETENESS OF A SEISMIC CATALOGUE

The statistical method adopted is based on the assumption that the sequence of occurrence times $y_1, y_2, \ldots, y_n$ in the interval $[T_1, T_2]$ is a realization of a Poisson process with intensity function $\lambda(t) = h_1\, I_{[T_1,s)}(t) + h_2\, I_{[s,T_2)}(t)$; that is, the occurrence rate has the constant value h_1 up to time s, after which it takes the value of h_2. We assume the change-point s indicates the date from the time at which the catalogue can be considered complete over a fixed magnitude threshold. The likelihood function has the form:

$$p(\mathbf{y} \mid h_1, h_2, s) = h_1^{\sum_{i=1}^{n} \delta_{[T_1,s)}(y_i)}\, \mathrm{e}^{-h_1\,(s-T_1)}\, h_2^{\sum_{i=1}^{n} \delta_{[s,T_2)}(y_i)}\, \mathrm{e}^{-h_2\,(T_2-s)}.$$

In the Bayesian approach, both the seismicity rates h_1, h_2 and the change-point s are random variables, to which we assign probability distributions exploiting our *a priori* knowledge of the phenomenon. The inference is based on stochastic simulation methods—Markov chain Monte Carlo methods (MCMC)—through which we generate a sample of dependent values for the parameters to be estimated, $(s_{(j)}, h_1^{(j)}, h_2^{(j)})_{j=1}^{m}$. Estimates of the distribution of s, h_1 and h_2 and of their summaries can be obtained from this sample. Thus, for example, the estimate of the probability density $f(s)$ is given by:

$$\tilde{f}(s) = \sum_{j=1}^{m} f(s \mid h_1^{(j)}, h_2^{(j)}, \mathbf{y})/m$$

where $f(s \mid \cdot)$ denotes the conditional density of s, and the estimates of the mean and of the variance of s are given, respectively, by

$$\tilde{s} = \sum_{j=1}^{m} s_{(j)}/m, \qquad \tilde{\sigma}_s^2 = \sum_{j=1}^{m} (s_{(j)} - \tilde{s})^2/m.$$

The same holds for h_1 and h_2. We can also obtain an estimate of $\hat{\lambda}(t)$:

$$\hat{\lambda}(t) = \frac{1}{m} \sum_{j=1}^{m} \left\{h_1^{(j)}\, I_{[T_1,s_{(j)})}(t) + h_2^{(j)}\, I_{[s_{(j)},T_2)}(t)\right\},$$

which represents the posterior mean rate of the Poisson process.