

Identification of temporal patterns in the seismicity of Sumatra using Poisson Hidden Markov models

Katerina Orfanogiannaki,^{1,2}
Dimitris Karlis,²
Gerassimos A. Papadopoulos¹

¹Institute of Geodynamics, National Observatory of Athens, Thissio;

²Athens University of Economics and Business, Department of Statistics, Athens, Greece

Abstract

On 26 December 2004 and 28 March 2005 two large earthquakes occurred between the Indo-Australian and the southeastern Eurasian plates with moment magnitudes $M_w=9.1$ and $M_w=8.6$, respectively. Complete data ($m_b \geq 4.2$) of the post-1993 time interval have been used to apply Poisson Hidden Markov models (PHMMs) for identifying temporal patterns in the time series of the two earthquake sequences. Each time series consists of earthquake counts, in given and constant time units, in the regions determined by the after-shock zones of the two mainshocks. In PHMMs each count is generated by one of m different Poisson processes that are called states. The series of states is unobserved and is in fact a Markov chain. The model incorporates a varying seismicity rate, it assigns a different rate to each state and it detects the changes on the rate over time. In PHMMs unobserved factors, related to the local properties of the region are considered affecting the earthquake occurrence rate. Estimation and interpretation of the unobserved sequence of states that underlie the data contribute to better understanding of the geophysical processes that take place in the region. We applied PHMMs to the time series of the two mainshocks and we estimated the unobserved sequences of states that underlie the data. The results obtained showed that the region of the 26 December 2004 earthquake was in state of low seismicity during almost the entire observation period. On the contrary, in the region of the 28 March 2005 earthquake the seismic activity is attributed to triggered seismicity, due to stress transfer from the region of the 2004 mainshock.

Introduction

A common approach when working with seismicity data is to count the number of

events in a given time period, *e.g.* one month, and then to examine the resulting series. The Poisson distribution is widely used as a probability model for such kind of data. One of the most known properties of the Poisson distribution is that the mean of the counts equals the variance. This can be used as a diagnostic for the appropriateness of the Poisson distribution. In some cases though the mean is greater than the variance and the data are overdispersed. The literature contains several competitors for this case. It is known that Poisson mixture models (PMMs) is an important candidate for modeling overdispersed heterogeneous data.^{1,2} However, data collected from the same area in successive time intervals tend to be dependent and, therefore, appropriate models for statistical modeling must accommodate this time-dependent structure. A class of models that allow for temporal dependence between the data in addition to overdispersion is Poisson Hidden Markov models (PHMMs). PHMMs are extensions of the well-known PMMs and they decay to PMMs in the case of independent observations.

In PHMMs each observation is generated by one of m Poisson distributions that are called states. The states are unobserved (hidden), hence the name PHMMs. Each state corresponds to a different rate, while the series of states is in fact a Markov chain. Through the transition probability matrix of the Markov chain, the state that generates the next observation depends on the state that generated the current observation. PHMMs allow us to estimate the unobserved sequence of states that underlie the observation sequence. In this way we may reveal unknown properties of the mechanism that generated the data and classify the observations with precision and objectivity.

PHMMs do not assume a constant rate for a long period of time. They incorporate a varying seismicity rate, which is more realistic than long-term constant rate. In fact, when long-term constant rate is assumed, short-term variations in seismicity are disregarded. However, short-term variations in seismicity are important in the evaluation of the seismic activity in a region. PHMMs assume a particular rate for each state. By estimating the state sequence observations are classified according to the rate that corresponds to them and changes to seismicity rate can be detected.

In recent years, applications of HMMs to geophysical and seismological data have gained increasing interest. Granat and Donnellan³ applied HMMs to GPS and seismicity data from the southern California region in order to distinguish different classes of observed seismic events. Temporal dependencies present in earthquake sequences cannot be accurately modeled by single distributions or finite mixture distributions. Li and Anderson-Sprecher⁴ used HMMs to model of

Correspondence: Katerina Orfanogiannaki, Institute of Geodynamics, National Observatory of Athens, Lofos Nymfon, 11810, Thissio, Athens, Greece.
Tel. +302103490123 - Fax: +302103490165.
E-mail: kath_orf@noa.gr

Key words: seismicity rate, Markov chains, Hidden states.

Acknowledgments: we would like to thank two anonymous referees for their thorough and insightful comments that helped us improve an initial version of this paper.

Contributions: the authors contributed equally.

Conflict of interests: the authors declare no potential conflict of interests.

Conference presentation: part of this paper was presented at the 11th International Congress of the Geological Society of Greece, 2007 May 24-26, Athens, Greece.

Received for publication: 29 June 2013

Revision received: 6 November 2013.

Accepted for publication: 22 November 2013.

This work is licensed under a Creative Commons Attribution NonCommercial 3.0 License (CC BY-NC 3.0).

©Copyright K.Orfanogiannaki et al., 2013
Licensee PAGEPress, Italy
Research in Geophysics 2014; 4:4969
doi:10.4081/rg.2014.4969

waiting times in the 1985 Yellowstone earthquake swarm. They assumed exponential distribution associated with each state and showed that in the case of the 1985 Yellowstone swarm, the HMMs improved the modeling of the waiting time distribution compared to single distributions (Exponential, Weibull and Log-normal) or finite mixture distributions (Exponential - mixture) that are traditionally used to model earthquake waiting times. Ebel *et al.*⁵ applied an HMM using not only the intervened times between earthquakes but also the spatial quadrant in which they occur. The time-dependent interaction of seismic events that is extracted by the transition probability matrix of HMMs identifies relationships between earthquakes due to stress changes within a fault system. The outline of this methodology was used by Chambers *et al.*⁶ to produce 1, 2 and 10 days forecasts in the southern California and western Nevada regions. Orfanogiannaki *et al.*⁷ used HMMs as a tool to identify through the estimated sequence of hidden states the seismic cycle of strong earthquakes in the area of Killini, Western Greece. It was an application of HMMs to model the number of earthquakes occurring in daily and monthly time intervals.

Recently, Votsi *et al.*⁸ applied HMMs to groups of earthquake magnitudes and estimated the sequence of hidden states, where each state corresponds to different stress field levels. In this approach we follow the methodology introduced by Orfanogiannaki *et al.*⁷ and examine the seismicity in two adjacent areas of Sumatra where two large earthquakes occur in 2004 and 2005. The objective of this paper is to reveal the characteristics of the regions examined and the geophysical processes that led to the occurrence of the two mainshocks.

Data

The data sources are the USGS and ISC earthquake catalogues for the region denoted by the rectangle E with coordinates 1.00N - 15.00N and 91.00E - 100.00E (Figure 1). At first, the entire region E, is divided into two sub-regions, N (north) and S (south), based on the rupture zones of the two big earthquakes of 26.12.04 and 28.03.05, respectively.⁹ The solid line in the map of Figure 1, shows the boundary between these two regions, while the red and blue stars correspond to the epicenters of the 2004 and 2005 mainshocks, respectively. According to geophysical evidence the rupture in the sub-region N was not uniform.¹⁰ The rupture started at the south part of the region and then propagated further to the north. Based on the progress of the rupture we divided sub-region N into two smaller regions N1 and N2. Data completeness analysis based on the magnitude-frequency relationship showed that the data in all regions are complete for $mb \geq 4.2$ for the time interval from 1994 onwards. This common completeness magnitude is adopted for all sub-regions and only events with magnitude above this cut-off are used in the analysis that follows. The time distribution of earthquake magnitudes is plotted *versus* time in Figure 2. All data sets are actually discrete valued time series, since they count the number of events in twenty-three daytime periods. The time interval between the two big earthquakes of 26.12.04 and 28.03.05 is 92 days. The entire time interval can be divided into 4 equal time intervals of 23 days in such a way that the two mainshocks do not fall in the middle of any time interval but only at the edges. We are interested to examine if there are any patterns before the two mainshocks and in the time interval between them.

Hidden Markov models: definition and notation

HMMs are stochastic processes that consist

of two parts. The first part is an unobserved (hidden) finite state Markov chain $\{C_i: i \in \mathbb{N}\} \mathbb{R}$ with m states. The second part is a sequence of random variables $\{Y_i: i \in \mathbb{N}\}$ that, conditionally on C_i , are mutually independent. Each state is associated with a probability distribution function f from the same parametric family. When the state of the model at time i (C_i), is known and equal to c_i , Y_i takes the value y_i with probability $f(y_i | c_i)$. f may be either continuous or discrete. In the discrete case f may be selected among a variety of different distribution families like Poisson, Negative Binomial etc. We focus on discrete valued hidden Markov models and we assume that each observation is generated from a Poisson distribution, and thus we derive Poisson hidden Markov models (PHMMs). In PHMMs f takes the form:

$$f(y_i | c_i = j) = \frac{e^{-\lambda_j} \lambda_j^{y_i}}{y_i!} \quad (1)$$

where $\lambda_j \geq 0$, $j = 1, \dots, m$ is the parameter of the Poisson distribution that corresponds to state j and $y_i = 0, 1, \dots$, for all $i = 1, \dots, n$ is the observation that corresponds to the i -th point in time. In our case the sequence of random variables $\{Y_i: i \in \mathbb{N}\}$ is restricted to take only nonnegative integer values. The transition among the different states is determined by the transition probabilities of the transition probability matrix of the Markov chain:

$$\Gamma = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1m} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{m1} & \gamma_{m2} & \cdots & \gamma_{mm} \end{pmatrix} \quad (2)$$

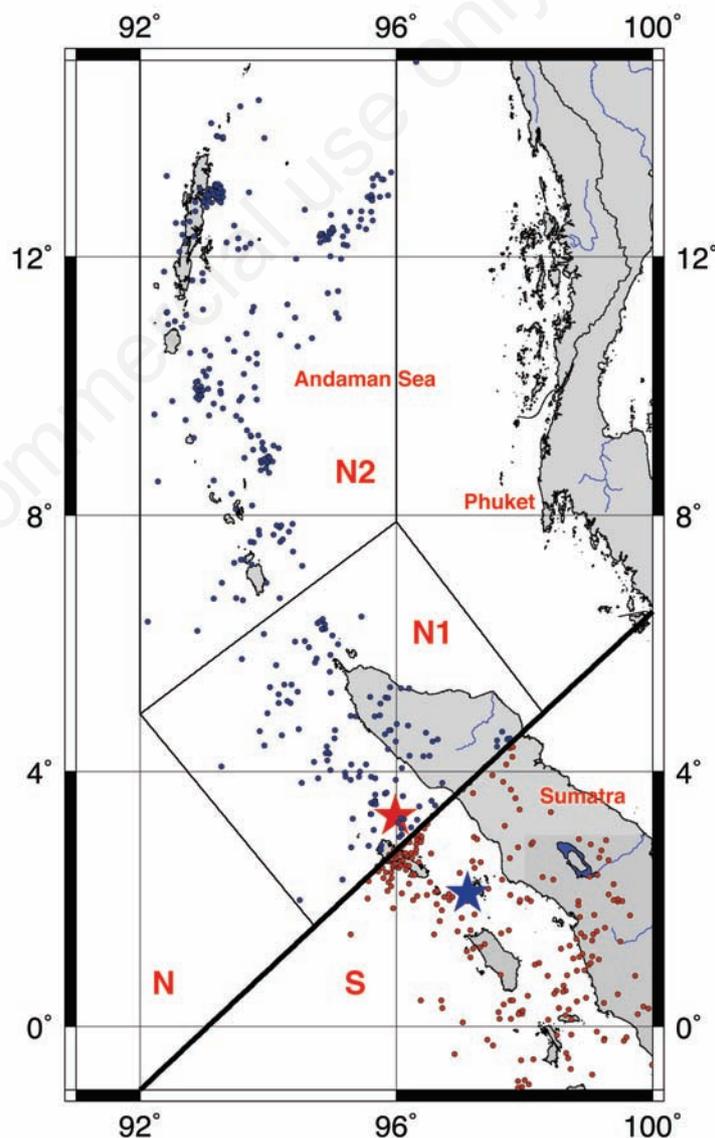


Figure 1. Map of the area. N1, N2, and S are the two major regions examined (see details in the text). Stars represent earthquake epicentres.

The transition probabilities γ_{lj} are defined as: $\gamma_{lj} = P(C_i = j | C_{i-1} = l)$ where $l, j = 1, \dots, m$. This is the probability that given the hidden process was in state l at the previous time point, it will be in state j at the current. The parameters of the model are the transition probabilities of the Markov chain and the parameters of the Poisson distributions that are associated with the states. If we denote with Ψ the vector of model parameters to be estimated then the likelihood of an HMM is:

$$L(\Psi | y_1, \dots, y_m) = \sum_{c_1=1}^m \dots \sum_{c_m=1}^m P(C_1 = c_1) f(y_1 | c_1) \prod_{i=2}^m \gamma_{c_{i-1}c_i} f(y_i | c_i) \quad (3)$$

where n is the length of the observation sequence used for the model parameter estimation. For the likelihood of a HMM to be calculated, the backward, $\beta_j(i)$, and forward, $\alpha_j(i)$, probabilities were introduced by Baum *et al.*¹¹ $\beta_j(i) = P(y_{i+1}, \dots, y_n | C_i = j)$ and $\alpha_j(i) = P(y_1, \dots, y_i, C_i = j)$. The likelihood can then be calculated in terms of the forward probabilities as:

$$L = \sum_{i=1}^m a_i(n) \quad (4)$$

For a very thorough review of the theory of HMMs one may refer to Ephraim and Merhav.¹²

Estimation of the unknown parameters

Due to the underlying structure of HMMs that allow for a missing data representation of the model Estimation Maximization algorithm, known as EM algorithm,¹³ is adopted for Maximum Likelihood estimation of the parameters of interest. The EM-algorithm consists of two steps: the expectation step (E-step) and the maximization step (M-step). At the E-step, the conditional expectations of the missing data are computed. The algorithm augments the observed data to a set of complete data using the values obtained from the E-step. The complete data log-likelihood is maximized with respect to the model parameters at the M-step of the algorithm. In a HMM the hidden states and the transitions from one state to another are treated as missing data in the EM algorithm. We define the indicator random variables $U_j(i)$ and $V_{jk}(i)$ respectively. The random variable $U_j(i)$ is equal to 1 if the state of the model at time i is j and 0 otherwise, *i.e.* $U_j(i) = 1$, if $C_i = j$ and 0 otherwise. The random variable $V_{jk}(i)$ is equal to 1 if a transit from state j to state k occurs at i and 0 otherwise, *i.e.* $V_{jk}(i) = 1$, if $C_{i-1} = j$ and $C_i = k$ and 0 otherwise. If we denote with Ψ the parameter vector to be estimated, then the complete data log-likelihood in term of the indicator variables is:

$$L(\Psi | y_1, \dots, y_m) = \log P(C_1 = c_1) + \sum_{i=2}^m \sum_{j=1}^m \sum_{l=1}^m v_{lj}(i) \log \gamma_{lj} + \sum_{i=2}^m \sum_{j=1}^m u_j(i) \log \text{Pois}(y_i | \lambda_j) \quad (5)$$

At the E-step of the EM algorithm we estimate U and V through their conditional expectations:

$$\hat{u}_j(i) = P(C_i = j | y_1, \dots, y_n) \quad (6)$$

and

$$\hat{v}_{jk}(i) = P(C_i = k, C_{i-1} = j | y_1, \dots, y_n) \quad (7)$$

$\hat{u}_j(i)$ is the probability that the state of the process at time i is j , given the observation sequence. $\hat{u}_j(i)$, for $i = 1, \dots, n$ and $j = 1, \dots, m$ are also called state probabilities, and they are used in the estimation of the unobserved sequence of states that underlie the data. $\hat{v}_{jk}(i)$ expresses the probability that given the hidden process is in state j today, it will be in state k tomorrow.

At the M-step of the EM algorithm we maximize the complete data log-likelihood and obtain parameter estimates in terms of $\hat{u}_j(i)$ and $\hat{v}_{jk}(i)$. Closed form equations are available for the parameter estimates.¹⁴ The sequence of hidden states that underlie the data is

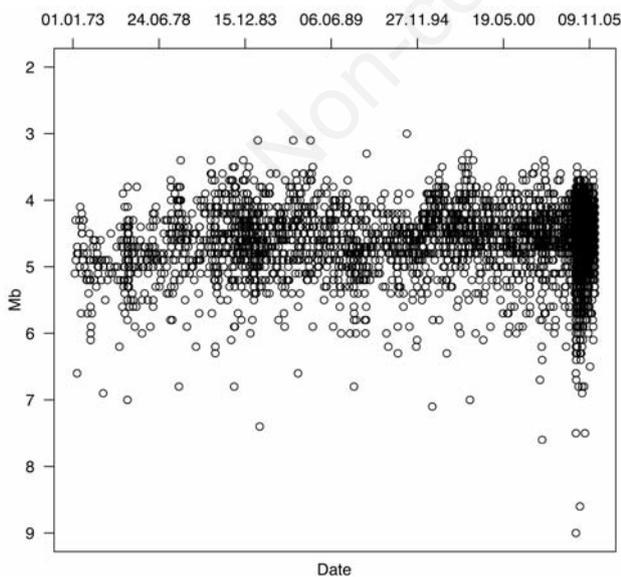


Figure 2. Time-magnitude plot for events occurring in the entire region examined from 01.01.1973 to 14.03.2006. The cut-off magnitude for completeness has been selected equal to $mb=4.2$ for the post-1993 time interval.

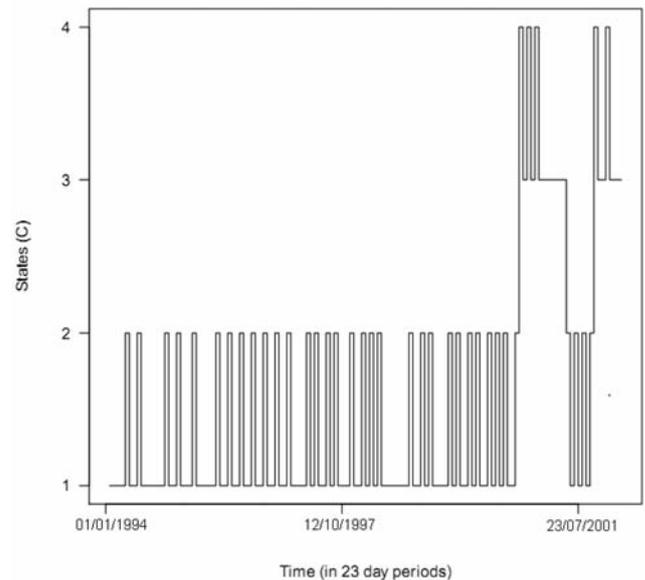


Figure 3. Estimated states, C , that underlie the data against time (in 23-day periods) for the entire region (E) examined. The zero point of time is 01.01.1994. The dates on the horizontal axis represent the beginning of the 23-day interval.

obtained by maximizing the joint probability distribution of the hidden states given the observation sequence, also known as the Viterbi algorithm¹⁵ (see also Forney¹⁶).

Analysis

We applied PHMMs to regions E, S, N, N1 and N2 for different number of hidden states. The optimum number of states for each region is selected based on the Akaike information criterion (AIC).¹⁷ Consider the AIC criterion defined as: $AIC(m) = -2L(m) + 2m^2$, where $L(m)$ is the maximized log-likelihood for a model with m states. For each region we choose the number of states m to be the number that minimizes $AIC(m)$.

In region E, the model with 4 states was selected as the best model to describe the data. The application of PHMM in the complete data set for the entire region E showed (Figure 3) that the state of seismicity ranges only from state 1 to 2 in the interval 1994-2002 that is the seismicity was relatively low. From 2002 onwards a transition to higher states of seismicity is observed; that is in states 3 and 4, with rates 9.61 and 28.20 (events/23days), respectively (Figure 3). To focus on the period of the increased seismicity, we narrow the time window examined. Seismicity state was investigated for the time interval 01.01.2000-25.12.2004 inclusive for sub-regions N, N1 and N2 as well as for the time interval 1.1.2000-27.3.2005 inclusive for sub-region S. In sub-regions S, N and N2 the model with 3 components was selected while sub-region N1 has

one state less. The parameter estimates of PHMMs for all 4 sub-regions are summarized in Table 1.

The two sub-regions N and S that correspond to the rupture zones of the 26.12.04 and 28.03.05 respectively, are both characterized by intense seismic activity in the time period examined. In sub-region N, 7 earthquakes with magnitudes $mb \geq 5.9$ occurred within almost 5 years while in sub-region S the number of earthquakes of the same magnitude class for the same time period is 6. The time distribution of magnitudes for the two sub-regions are illustrated in Figures 4A and 5A. We observe that strong events are distributed uniformly across time in both sub-regions. The number of earthquakes counted in 23-day periods is plotted versus time for the two sub-regions in Figures 4B and 5B. The mean seismicity rate for sub-region N is 3.14 events/23-day periods while for sub-region S is 2.78. That is almost 3 earthquakes of magnitude $mb \geq 4.2$ occurred in both sub-regions every 23 days.

The sub-region N was further divided into sub-regions N1 and N2 and PHMMs were applied to all 4 sub-regions in order to reveal the pattern of hidden states that underlie the data. In Figure 6 appears the estimated sequence of hidden states for all 4 sub-regions. The sequence of states for sub-region N and sub-region N2 have the same pattern and the Poisson rates that correspond to the 3 states of the two sub-regions have similar values (Table 1). The 3 states correspond to periods of low (~ 1.4 events/23-day period), medium (~ 5 events/23-day period) and high (~ 18.5 events/23-day period) seismic activity. State 3 of high seismicity is attributed to aftershock

activity associated with strong earthquakes. The seismicity in region N took place in region N2 and region N1 was always much less active. In fact, no state of high seismic activity appears in sub-region N1 for the entire period examined. State 2 is associated with some clusters of earthquakes of similar magnitude concentrated in time. During the observation sequence sub-region N1 was in state of long-term failure propagation. Evidence of instability comes from the 4 earthquake clusters with mb ranging from 4.7 to 5.9 that occurred in N1. However, it appears that the system had not reached the maturity stage that would result to rupture.

One of the most interesting aspects of these data sets is how the seismic activity in sub-region S seems to be affected by the seismic activity in sub-region N. In particular, both areas experienced moderately strong earthquakes in 2002 (N area first followed almost 2 months later by area S). A similar pattern repeated at the end of 2004 and in early 2005 and PHMMs pick up these seismicity increases as jumps to higher states. In particular at the end of 2004 the state of seismicity in sub-region S changes from 1 to 2 and 3 immediately after this change is observed in sub-region N. When sub-region N transits to state 3 for the second time sub-region S transits again to state 3 immediately after the occurrence of the 26.12.2004 mainshock. However this time sub-region S does not jump back to state 2 like before but it remains to state 3 for 3 subsequent time intervals until the occurrence of the 28.3.05 mainshock. The change of the system from a state of medium seismicity to a state of increased seismicity is not associated

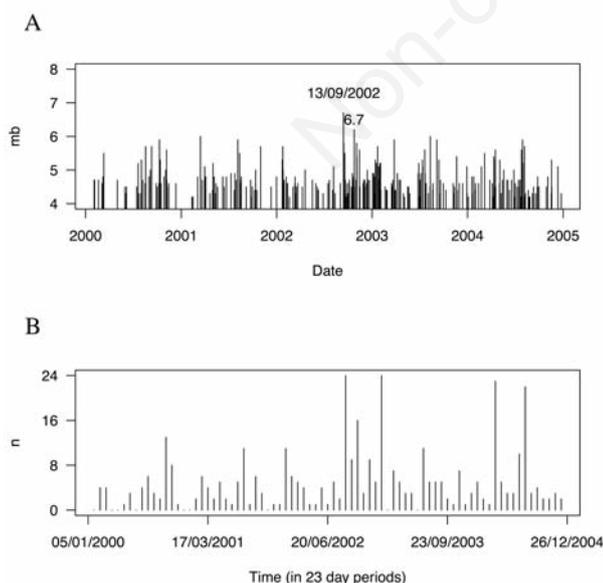


Figure 4. Time distribution of: A) magnitudes; and B) the event counts (in 23-day periods) for sub-region N. The dates on the horizontal axis represent the beginning of the 23-day interval.

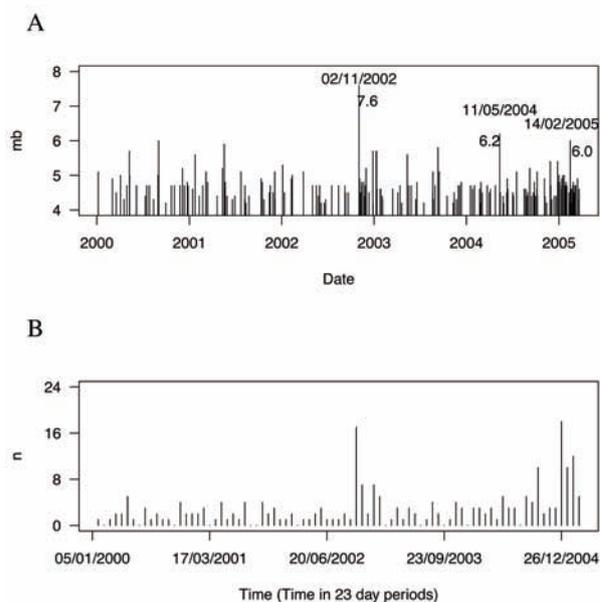


Figure 5. Time distribution of: A) magnitudes; and B) the event counts (in 23-day periods) for sub-region S. The dates on the horizontal axis represent the beginning of the 23-day interval.

with the occurrence of another strong earthquake in the sub-region S or neighboring areas. Sub-region S represents rather a triggered seismicity, due to stress increase from sub-region N to sub-region S.¹⁸

Discussion and Conclusions

PHMMs can provide a diagnostic tool for identifying changes in seismicity states. The model incorporates a varying seismicity rate; it detects the changes of the rate over time and corresponds a particular rate to each one state. The transition probabilities of the Markov chain determine the transitions among the different states. Estimation of the sequence of unobserved states that underlie the data is attained with relative ease. The entire region was divided in sub-regions based in the progress of the rupture of the 2004 earthquake and the rupture zone of the 2005 earthquake. PHMMs were applied to all four sub-regions and the sequence of hidden states was estimated. The estimated sequence of states revealed that the regions of the two Sumatra mainshocks have different seismotectonic characteristics that need further more detailed analysis to be revealed. The analysis showed that we shouldn't consider the North region uniformly and revealed exclusive characteristics of the sub-regions. The rupture of the 2004 earthquake was not uniform it started from the South (sub-region N1) and propagated further

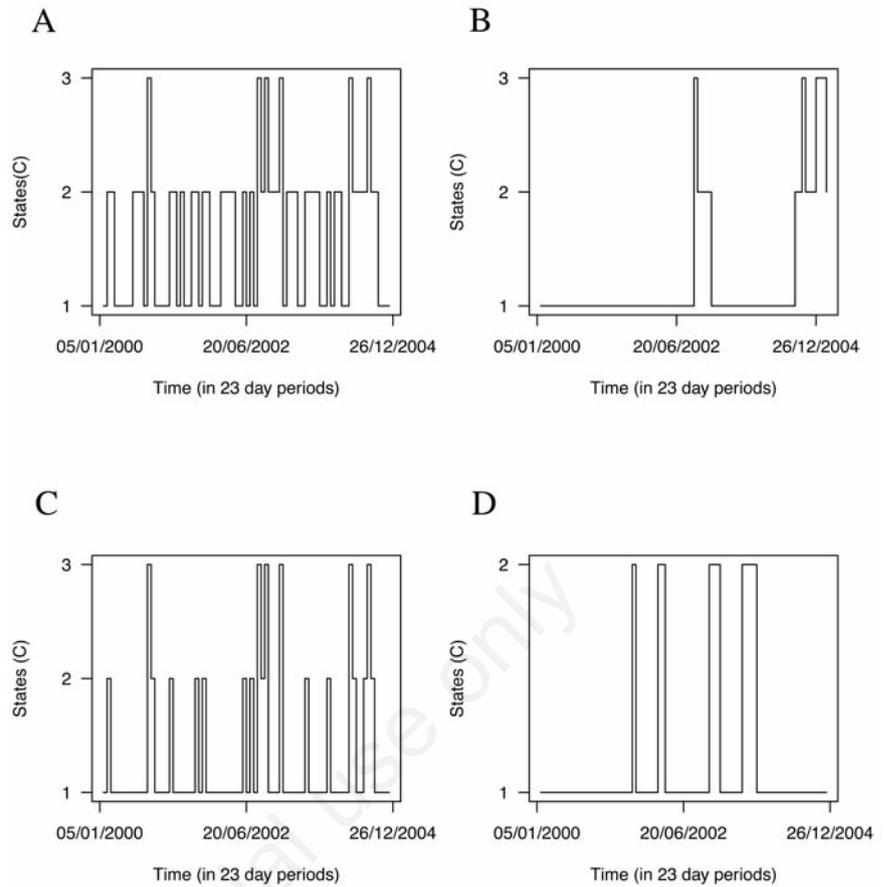


Figure 6. Estimated states, C, that underlie the data against time (in 23-day periods): A) sub-region N; B) sub-region S; C) sub-region N2; and D) sub-region N1. The zero point of time is 01.01.2000 for all the sub-regions examined. The dates on the horizontal axis represent the beginning of the 23-day interval.

Table 1. Model parameter estimates.

| Segment | Number of components | Component number i | Parameters estimates | | |
|---------|----------------------|--------------------|------------------------------|--|-------------|
| | | | Poisson rates λ _i | Transition matrix | Probability |
| E | 4 | 1 | 4.85 | $\begin{pmatrix} 0 & 0.963 & 0 & 0.037 \\ 0.465 & 0.526 & 0 & 0.009 \\ 0.104 & 0 & 0.658 & 0.237 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ | |
| | | 2 | 8.86 | | |
| | | 3 | 9.61 | | |
| | | 4 | 28.20 | | |
| S | 3 | 1 | 1.78 | $\begin{pmatrix} 0.966 & 0.016 & 0.018 \\ 0.133 & 0.689 & 0.178 \\ 0 & 0.573 & 0.427 \end{pmatrix}$ | |
| | | 2 | 4.13 | | |
| | | 3 | 13.15 | | |
| N | 3 | 1 | 1.46 | $\begin{pmatrix} 0.490 & 0.424 & 0.086 \\ 0.406 & 0.495 & 0.099 \\ 0.213 & 0.787 & 0 \end{pmatrix}$ | |
| | | 2 | 5.21 | | |
| | | 3 | 19.42 | | |
| N1 | 2 | 1 | 1.38 | $\begin{pmatrix} 0.911 & 0.089 \\ 0.460 & 0.540 \end{pmatrix}$ | |
| | | 2 | 5.17 | | |
| N2 | 3 | 1 | 1.29 | $\begin{pmatrix} 0.742 & 0.206 & 0.052 \\ 0.803 & 0 & 0.197 \\ 0.334 & 0.666 & 0 \end{pmatrix}$ | |
| | | 2 | 4.15 | | |
| | | 3 | 18.15 | | |

North (sub-region N2). Sub-region N2 is characterized by increased seismic activity that dominates the entire region. On the contrary, the sub-region N1 was in state of low seismicity during almost the entire observation period. It appears that the system was in state of long-term failure propagation during the observation period and had not reach the maturity stage that would result to rupture. The seismic activity in sub-region S is affected by the seismic activity in sub-region N. Both areas experienced moderately strong earthquakes in 2002 and at the end of 2004 and in early 2005. Sub-region N was activated first, followed by sub-region S. PHMMs pick up these seismicity increases as jumps to higher states. This may imply a triggered seismicity, due to stress transfer from the region of the 2004 mainshock to the region of the 2005 mainshock and a resulting acceleration of the stress loading that prepared the 2005 rupture.

References

1. Titterton M, Makov U, Smith AFM. Statistical analysis of finite mixtures. New York: Wiley; 1985.
2. McLachlan G, Peel D. Finite mixture models. New York: Wiley; 2000.
3. Granat R, Donnellan A. A hidden Markov model based tool for geophysical data exploration. *Pure Appl Geophys* 2002;159: 2271-83.
4. Li Y, Anderson-Sprecher R. Hidden Markov modeling of waiting times in the 1985 Yellowstone Earthquake Swarm. *Pure Appl Geophys* 2013;170:785-95.
5. Ebel JE, Chambers DW, Kafka AL, Baglivo JA. Non-poissonian earthquake clustering and the hidden Markov model as bases for earthquake forecasting in California. *Seismol Res Lett* 2007;78:57-65.
6. Chambers DW, Baglivo JA, Ebel JE, Kafka AL. Earthquake forecasting using hidden Markov models. *Pure Appl Geophys* 2012; 169:625-39.
7. Orfanogiannaki K, Karlis D, Papadopoulos GA. Identifying seismicity levels via Poisson hidden Markov Models. *Pure Appl Geophys* 2010;167:919-31.
8. Votsi I, Limnios N, Tsaklidis G, Papadimitriou E. Hidden Markov models revealing the stress field underlying the earthquake generation. *Physica A* 2013; 392:2868-85.
9. Lay T, Kanamori H, Ammon CJ, et al. The great Sumatra-Andaman earthquake of 26 December 2004. *Science* 2005;308:1127-33.
10. Ammon CJ, Ji C, Thio HK, et al. Rupture Process of the 2004 Sumatra-Andaman earthquake. *Science* 2005;308:1133-39.
11. Baum LE, Petrie T, Soules G, Weiss N. A maximization technique in the statistical analysis of probabilistic functions of Markov chains. *Ann Math Stat* 1970;41: 164-71.
12. Ephraim Y, Merhav N. Hidden Markov processes. *IEEE Trans Inform Theory* 2002;48:1518-69.
13. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J Roy Stat Soc B* 1977;39:1-38.
14. MacDonald IL, Zucchini W. Hidden Markov models and other models for discrete-valued time series. London: Chapman and Hall; 1997.
15. Viterbi AJ. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE T Inform Theory* 1967;IT-13.
16. Forney GD. The Viterbi algorithm. *Proc IEEE* 1973;61.
17. Akaike H. New look at the statistical model identification. *IEEE T Automat Contr* 1974;AC-19.
18. McCloskey J, Nalbant SS, Steacy S. Earthquake risk from co-seismic stress. *Nature* 2005;434:291.