International Journal of Pure and Applied Mathematics
Volume 62 No. 1 2010, 1-15

A BAYESIAN ESTIMATION OF THE PROBABILITY OF ERUPTION PATTERNS OF THE MEXICAN VOLCANO *VOLCÁN DE COLIMA*

Luis Javier Álvarez[1], Eliane R. Rodrigues[2 §]

[1]Instituto de Matemáticas – Unidad Cuernavaca
Universidad Nacional Autónoma de México – UNAM
México, D.F., 04510, MEXICO
e-mail: lja@matcuer.unam.mx

[2]Instituto de Matemáticas
Universidad Nacional Autónoma de México – UNAM
Area de la Investigación Científica
Circuito Exterior, Ciudad Universitaria
México, D.F., 04510, MEXICO
e-mails: eliane@math.unam.mx

Abstract: In this paper we use a Markov chain model and real data of volcanic activity from *Volcán de Colima*–Mexico to estimate the probability of having a volcanic activity of interest with a given intensity. The data used here correspond to the activity pattern during the years 1893-1905. We assume that successive activities presented by the volcano follow a Markov chain of order K. This order is considered a random variable assuming values on a suitable set. Using a Bayesian approach we estimate the most likely order of the chain and the respective transition matrix. Based on these results the probability of having an activity with an intensity of interest is estimated.

AMS Subject Classification: 60J20, 62F15, 62M99, 92F05
Key Words: Bayesian inference, Markov chain models, volcanic activity

1. Introduction

According to documents kept in the *Archivo General de Indias* in Seville –

Received: December 21, 2009

§Correspondence author

Spain (see [14]) the historic documentation of eruptions of the Mexican volcano *Volcán de Colima* goes back to the mid 16-th century. As the scientific interest on volcanology in Mexico has grown, the quality of the recorded information has increased throughout the centuries. There have been several attempts to systematically analyse the existing information about the *Volcán de Colima*. In most cases the motivation has been to find some statistical pattern which may allow predictions of future activities of the volcano. Among the publications on this subject it is possible to quote [8] in which Poisson patterns of the activity of *Volcán de Colima* between 1560 and 1989 are identified.

The data used in this work is the one prepared by priest Severo Díaz for the X International Geological Congress, held in Mexico City in 1906. That work presents the results produced by observations carried out by Díaz and another priest, José María Arreola, from Zapotlán Meteorological and Volcanological Observatory, between 1893 and 1905. The original idea to carefully observe and register the volcano's activity is due to priest Arreola. Observations were carried out and compiled every day, and often every hour. The volcano's emissions were classified mainly into two main categories: steam emissions and eruptive activity.

In [9], priest Díaz gives extraordinarily detailed information on the activity of the *Volcán de Colima* over a 13-year period. Priest Díaz and his co-workers did not perform a theoretical interpretation of the information they were compiling. They only reported the observed activities. Those observations were initially recorded using a numerical scale, but this was substituted by a short description of the different types of activity. The information, when transformed into numerical values allows for the realisation of different types of analysis. In this work the approach taken to analyse the data is the Bayesian point of view.

According to the preliminary notes of [9] there was an attempt to use a numerical scale from 1 to 11 grouped into several sets. However, in the final presentation of the work this scale was not used. Instead, a series of descriptors were presented. In the present work, to those descriptors we assign consecutive numerical values according to the intensity of the eruptive activity.

Since volcanic activity can change during a given day, the number associated with a particular day is the maximum activity observed on that given day. Table 1 shows the arbitrary scale associated with the descriptors used by [9] and his co-workers. These are the basis of the data used.

Figure 1 shows the data used here, where the correspondence between descriptors and numerical values associated to them has already been made.

Some of the observations corresponding to "*small clouds of little importance*

Numerical Value	**Type of Activity**
1	*no activity*
2	*small clouds of little importance over the volcano*
3	*steamy small clouds*
4	*dense cloud*
5	*rapid eruption of dense smoke*
6	*small and slow eruption*
7	*small, dense and rapid eruption*
8	*medium size eruption*
9	*medium size, dense eruption*
10	*large and dense eruption*
11	*large and violent eruption*

Table 1: Numerical scale associated to the descriptors originally used by [9]

over the volcano", associated to number 2 in the numerical scale, are probably of volcanic origin, but may be also due to atmospheric condensation around the volcanic cone. Since the aim of this analysis is to detect the existence of a general trend in the volcano's behaviour the latter type was not taken into consideration when the data was compiled. Therefore, the final number of days for which data is available is 4748.

This paper is organised as follows. In Section 2 we describe the stochastic model considered here. Section 3 gives a Bayesian formulation of the model. In Section 4 results are applied to the data of the volcano Volcán de Colima. Finally, in Section 5 we present some remarks.

2. Description of the Stochastic Model

The model considered here is the one proposed by [3] to study the behaviour of the daily ozone peaks in Mexico City. (We refer the reader to that work if a more detailed description about the formulation described here is needed.)

The range of activity intensities of the volcano is partitioned into $I+1$ parts. The extremes of this partition are indicated by $L_1 < L_2 < \cdots < L_I$. We assume that successive intervals to which the states of the volcano belong to are governed by a time-homogeneous Markov chain. Let K be the order of this chain.

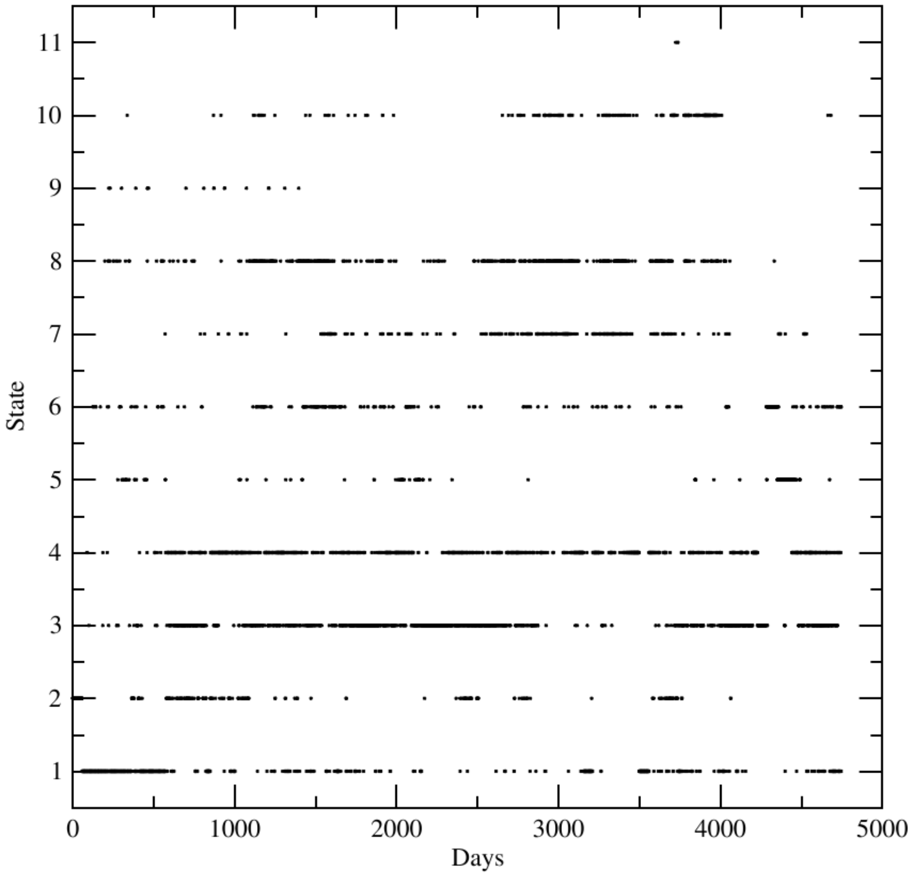

Figure 1: Numerical data associated to the descriptors provided by [9] that was used in the analysis presented here

We assume that K is a random variable with state space $S = \{0, 1, \ldots, M\}$. Let N be a fixed natural number such that $N \geq K$ with probability one (the number N accounts for the amount of observed data).

Denote by $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_N)$ the sequence recording the volcano's daily activity. Let $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_N)$ be the sequence of random variables such that, for $i = 1, 2, \ldots, N$,

$$Y_i = \begin{cases} 1, & \text{if } Z_i \leq L_1 \\ k, & \text{if } L_{k-1} < Z_i \leq L_k, \qquad k = 2, 3, \ldots, I \\ I + 1, & \text{if } Z_i > L_I. \end{cases}$$

(Note that, Y_i, $i = 1, 2, \dots, N$ indicates to which interval the volcano's activity on the i-th observed day belongs.)

Let the sequence $\mathbf{Y}$ be governed by a time-homogeneous Markov chain of order K indicated by $X^{(K)} = \{X_n^{(K)} : n = 1, 2, \dots\}$. The sequence $X^{(K)}$ has as its state space the set

$$\chi_1^{(K)} = \{(x_1, x_2, \dots, x_K) : x_i \in \{1, 2, \dots I + 1\}, i = 1, 2, \dots, K\}, \tag{1}$$

$K = 1, 2, \dots, M$. When $K = 0$ we have that $\chi_1^{(0)} = \chi_1^{(1)}$.

Given the nature of the state space (1) it is worthwhile to consider an alternative way of representing it. One solution is to associate each $(z_1, z_2, \dots, z_K) \in \chi_1^{(K)}$ with a number in

$$\chi_2^{(K)} = \{1, 2, \dots, (I + 1)^K\}, \tag{2}$$

using the function $f : \chi_1^{(K)} \to \chi_2^{(K)}$ given by

$$f((z_1, z_2, \dots, z_K)) = \sum_{l=0}^{K-1} (z_{l+1} - 1)(I + 1)^l + 1.$$

Unless otherwise stated, from now on we use the set given by (2) to represent the state space of $X^{(K)}$ and use the notation $(x_1, \dots, x_K) \leftrightarrow \overline{m}$ to indicate that $(x_1, \dots, x_K) \in \chi_1^{(K)}$ corresponds to $\overline{m} \in \chi_2^{(K)}$.

Note that for observations $(y_1, y_2, \dots, y_N)$, if the present state of $X^{(K)}$ is $X_n^{(K)} = (y_{n+1}, y_{n+2}, \dots, y_{n+K}) \leftrightarrow \overline{m}$, $0 \leq n \leq N - K - 1$, then the transition probability of $X^{(K)}$ is different from zero if the next state $X_{n+1}^{(K)}$ is $(y_{n+2}, \dots, y_{n+K}, y_{n+K+1}) \leftrightarrow \overline{m}'$. Additionally, $\overline{m}'$ occurs if and only if the observation following $y_{n+1}, y_{n+2}, \dots, y_{n+K}$ is y_{n+K+1}. This allows us to use a reduced transition matrix for $X^{(K)}$. This matrix is denoted by $P^{(K)} = \left(P_{\overline{m}j}^{(K)}\right)$, $\overline{m} \in \chi_2^{(K)}$, $j \in \{1, 2, \dots, I + 1\}$ and is defined by

$$P_{\overline{m}j}^{(K)} = P(Y_{n+K+1} = j \,|\, X_n^{(K)} = (y_{n+1}, \dots, y_{n+K}) \leftrightarrow \overline{m}),$$

for $\overline{m} \in \chi_2^{(K)}$, $j \in \{1, 2, \dots, I + 1\}$, and $0 \leq n \leq N - K - 1$.

The interest of this work is to make predictions about the probability of having a given volcanic behaviour in a given day into the future. In order to do so, the order K of the Markov chain $X^{(K)}$ and its transition probabilities $P_{\overline{m}j}^{(K)}$, $\overline{m} \in \chi_2^{(K)}$, $j \in \{1, 2, \dots, I + 1\}$ must be estimated.

Therefore, the parameter that ought to be estimated is $\theta = \{K, P^{(K)}\}$. This parameter belongs to the space

$$\Theta = \bigcup_{K=0}^{M} \left(\{K\} \times (\Delta_{I+1})^{(I+1)^K} \right),$$

where Δ_l indicates the simplex $\{(x_1,\, x_2,\, \ldots, x_l) \in \mathbb{R}^l \,:\, x_i \geq 0,\ i = 1, 2, \ldots, l;\ \sum_{i=1}^{l} x_i = 1\}$.

3. A Bayesian Estimation of the Parameters

Bayesian models have been applied to analyse data in areas such as image recovery, molecular biology and air pollution (see for example [1], [2], [3], [4], [5], [6], [7], [12], [13], and references therein). In this work we use the marginal posterior distribution of the parameters to obtain information about them. Therefore, if $L(\mathbf{Y} \,|\, K)$ and $P(K \,|\, \mathbf{Y})$ indicate the marginal likelihood function of the model and the marginal posterior distribution of K, then we have that

$$P(K \,|\mathbf{Y}) \propto L(\mathbf{Y} \,|\, K)\, P(K), \tag{3}$$

where $P(K)$ represent the prior distribution of K. Furthermore, for $P(P^{(K)} \,|\, K, \mathbf{Y})$ representing the marginal posterior distribution of $P^{(K)}$ given K, we have that

$$P(P^{(K)} \,|\, K, \mathbf{Y}) \propto L(\mathbf{Y} \,|\, K, P^{(K)})\, P(P^{(K)} \,|\, K), \tag{4}$$

where $L(\mathbf{Y} \,|\, K, P^{(K)})$ and $P(P^{(K)} \,|\, K)$ are the likelihood function of the model and the prior distribution of $P^{(K)}$ given K, respectively. The different components of (3) and (4) are given as follows.

The *prior distribution of the order* K will be a truncated Poisson distribution on the set $S = \{0, 1, 2, \ldots, M\}$ with parameter $\lambda > 0$ (other distributions may also be considered). The rows $P_{\overline{m}}^{(K)}$, $\overline{m} \in \chi_2^{(K)}$ of the matrix $P^{(K)}$ are assumed to be independent and have as *prior distributions* Dirichlet distributions with parameters $\alpha_{\overline{m}i} > 0$, $i = 1, 2, \ldots, I+1$. The *likelihood function* of the model is given by (see for example [11] among others)

$$L(\mathbf{Y} \,|\, K, P^{(K)}) = \prod_{\overline{m} \in \chi_2^{(K)}} \left[\left(\prod_{i=1}^{I} \left(P_{\overline{m}i}^{(K)} \right)^{n_{\overline{m}i}^{(K)}} \right) \left(1 - \sum_{i=1}^{I} P_{\overline{m}i}^{(K)} \right)^{n_{\overline{m}I+1}^{(K)}} \right], \tag{5}$$

and the *marginal likelihood function* is

$$L(\mathbf{Y} \,|\, K) \propto \prod_{\overline{m} \in \chi_2^{(K)}} \left\{ \frac{\Gamma\left(\sum_{i=1}^{I+1} \alpha_{\overline{m}i} \right)}{\Gamma\left(\sum_{i=1}^{I+1} [n_{\overline{m}i}^{(K)} + \alpha_{\overline{m}i}] \right)} \prod_{i=1}^{I+1} \left[\frac{\Gamma(n_{\overline{m}i}^{(K)} + \alpha_{\overline{m}i})}{\Gamma(\alpha_{\overline{m}i})} \right] \right\}, \tag{6}$$

where $n_{\overline{m}j}^{(K)}$ records the number of transitions such that the state of the Markov chain corresponding to $\overline{m} \in \chi_2^{(K)}$ is followed by the observation $j \in \{1, 2, \ldots, I+1\}$.

Therefore, from (3), (6) and the expression for $P(K)$, we have that $P(K \mid \mathbf{Y})$ is the discrete distribution given by

$$P(K \mid \mathbf{Y}) = \frac{1}{c}\left(\prod_{\overline{m}\in\chi_2^{(K)}}\left\{\frac{\Gamma\left(\sum_{i=1}^{I+1}\alpha_{\overline{m}i}\right)}{\Gamma\left(\sum_{i=1}^{I+1}[n_{\overline{m}i}^{(K)}+\alpha_{\overline{m}i}]\right)}\prod_{i=1}^{I+1}\left[\frac{\Gamma(n_{\overline{m}i}^{(K)}+\alpha_{\overline{m}i})}{\Gamma(\alpha_{\overline{m}i})}\right]\right\}\right)\frac{\lambda^K}{K!}, \quad (7)$$

where

$$c = \sum_{k\in S}\left(\prod_{\overline{m}\in\chi_2^{(k)}}\left\{\frac{\Gamma\left(\sum_{i=1}^{I+1}\alpha_{\overline{m}i}\right)}{\Gamma\left(\sum_{i=1}^{I+1}[n_{\overline{m}i}^{(k)}+\alpha_{\overline{m}i}]\right)}\prod_{i=1}^{I+1}\left[\frac{\Gamma(n_{\overline{m}i}^{(k)}+\alpha_{\overline{m}i})}{\Gamma(\alpha_{\overline{m}i})}\right]\right\}\right)\frac{\lambda^k}{k!}$$

is the normalizing constant. Furthermore, from (4), (5) and the assumption made for $P(P^{(K)}|K)$, the *marginal posterior distribution of* $P^{(K)}$ *given* K is

$$P(P^{(K)} \mid K, \mathbf{Y}) = \prod_{\overline{m}\in\chi_2^{(K)}}\left\{\frac{\Gamma\left(\sum_{i=1}^{I+1}[n_{\overline{m}i}^{(K)}+\alpha_{\overline{m}i}]\right)}{\prod_{i=1}^{I+1}\Gamma\left(n_{\overline{m}i}^{(K)}+\alpha_{\overline{m}i}\right)}\left(\prod_{i=1}^{I}\left(P_{\overline{m}i}^{(K)}\right)^{n_{\overline{m}i}^{(K)}+\alpha_{\overline{m}i}-1}\right)\right.$$
$$\left.\times\left(1-\sum_{i=1}^{I}P_{\overline{m}i}^{(K)}\right)^{n_{\overline{m}I+1}^{(K)}+\alpha_{\overline{m}I+1}-1}\right\}, \quad (8)$$

i.e., is a product of Dirichlet distributions with parameters $n_{\overline{m}i}^{(K)} + \alpha_{\overline{m}i}$, $\overline{m} \in \chi_2^{(K)}$, $i \in \{1, 2, \ldots, I+1\}$.

There are several ways of estimating K and $P^{(K)}$. The approach considered here is the following. We take the value of K that maximizes (7) and given that value, we use as an estimate of the transition matrix $P^{(K)}$ the value that maximizes (8).

Note that, given $K = k$, the mode of $P(P_{\overline{m}}^{(k)} \mid K = k, \mathbf{Y})$, $\overline{m} \in \chi_2^{(k)}$ is

$$P_{\overline{m}i}^{(k)} = \frac{n_{\overline{m}i}^{(k)} + \alpha_{\overline{m}i} - 1}{\sum_{j=1}^{I+1}(n_{\overline{m}j}^{(k)} + \alpha_{\overline{m}j} - 1)}, \qquad i = 1, 2, \ldots, I+1 \quad (9)$$

(see for example [10]).

4. Predicting the Behaviour of the Mexican Volcano *Volcán de Colima*

In this section we apply the theoretical description presented in earlier sections to the *Volcán de Colima* data reported by [9]. The data consists, as it was mentioned above, of the daily observation of the volcano's activity from 1893 until 1905 presenting a total number of $N = 4748$ observations. The methodology is applied to two cases. One of them is such that the intervals provide the exact classification given by [9]. In the other one we consider what we call condensed classification. This classification gathers together states with similar behaviour in one class. These two cases are given in Sections 4.1 and 4.2, respectively.

4.1. A Partition with Eleven Possible States

In this section we consider the partition $L_j = j$, $j = 1, 2, \ldots, 11$. Table 1 gives the interpretation of these values in terms of the possible activity intensities that the volcano may present. This classification follows the different activities according to the description provided by [9].

In order to apply the methodology considered here we need to locate the time-homogeneous segments of the sequence $\mathbf{Y}$. That is made by considering a window of length 1500 and moving it through the sequence $\mathbf{Y}$ and calculating the proportion of each state in $S = \{1, 2, \ldots, 11\}$. It was possible to identify three time-homogeneous segments. One of them is obtained by considering the first 1000 observations. The second time homogeneous segment goes from the 1001-st observation to the 2000-th. The remaining observations form the third time-homogeneous segment. We refer to these cases as (i), (ii) and (iii), respectively.

In each case the maximum value of the marginal posterior distribution $P(K \mid \mathbf{Y})$ is achieved at $K = 1$. The difference from one case to another is reflected in the values of $P^{(K)}$ that maximise the marginal posterior distribution $P(P^{(K)} \mid K, \mathbf{Y})$. The value of the three transition matrices are given in Table 2. We write (i), (ii) and (iii) next to the states in each row to indicate which time-homogeneous segment we are considering. For instance, the first two rows of Table 2 indicate the transition probabilities from state 1 to all states when the first time-homogeneous segment is considered. The second row of Table 2 gives the transition probabilities from state 2 to all states when the second time-homogeneous segment is considered, and so on.

Note that, if the volcano's activity is "*no activity*", i.e., the chain has state

states	1	2	3	4	5	6	7	8	9	10	11
1(i)	0.77626	0.01663	0.03119	0.03742	0.03326	0.02911	0.00624	0.02703	0.01247	0.00624	0.00416
1(ii)	0.35514	0.06542	0.12150	0.18692	0.01869	0.04673	0.02804	0.11215	0.01869	0.02804	0.01869
1(iii)	0.63736	0.02930	0.10989	0.09524	0.00733	0.03297	0.02198	0.02564	0.00733	0.02564	0.00733
2(i)	0.03315	0.63536	0.09392	0.13260	0.01105	0.01657	0.01105	0.03315	0.01105	0.01105	0.01105
2(ii)	0.07937	0.26984	0.15873	0.25397	0.04762	0.03175	0.03175	0.03175	0.03175	0.03175	0.03175
2(iii)	0.07692	0.27885	0.19231	0.12500	0.01923	0.04808	0.05769	0.09615	0.01923	0.06731	0.01923
3(i)	0.10204	0.14286	0.36735	0.21769	0.01361	0.02721	0.02041	0.04082	0.04082	0.01361	0.01361
3(ii)	0.04207	0.01618	0.48867	0.19094	0.01294	0.05825	0.03883	0.10356	0.01294	0.02913	0.00647
3(iii)	0.04222	0.01847	0.62005	0.13984	0.01715	0.04090	0.03430	0.05409	0.00264	0.02639	0.00396
4(i)	0.08571	0.09524	0.14286	0.56190	0.00952	0.02381	0.02381	0.01429	0.01429	0.01905	0.00952
4(ii)	0.05161	0.03548	0.17742	0.50645	0.01935	0.04516	0.02258	0.10323	0.01613	0.01613	0.00645
4(iii)	0.02493	0.02632	0.14543	0.55679	0.02355	0.02770	0.06233	0.08864	0.00277	0.03878	0.00277
5(i)	0.47826	0.04348	0.04348	0.06522	0.08696	0.04348	0.04348	0.06522	0.04348	0.04348	0.04348
5(ii)	0.05882	0.05882	0.11765	0.14706	0.08824	0.08824	0.11765	0.14706	0.05882	0.05882	0.05882
5(iii)	0.01935	0.01290	0.07742	0.10968	0.63871	0.03871	0.04516	0.01290	0.01290	0.01935	0.01290
6(i)	0.37255	0.03922	0.15686	0.03922	0.03922	0.07843	0.03922	0.11765	0.03922	0.03922	0.03922
6(ii)	0.05785	0.01653	0.17355	0.10744	0.02479	0.25620	0.06612	0.23140	0.01653	0.03306	0.01653
6(iii)	0.05028	0.01676	0.11732	0.12849	0.03352	0.45810	0.09497	0.05028	0.01117	0.02235	0.01676
7(i)	0.07143	0.10714	0.10714	0.14286	0.10714	0.07143	0.10714	0.07143	0.07143	0.07143	0.07143
7(ii)	0.06173	0.02469	0.14815	0.17284	0.03704	0.06173	0.20988	0.19753	0.02469	0.03704	0.02469
7(iii)	0.02308	0.02692	0.10769	0.19615	0.03077	0.04231	0.26538	0.20769	0.00769	0.08462	0.00769
8(i)	0.35185	0.07407	0.11111	0.09259	0.05556	0.09259	0.03704	0.05556	0.05556	0.03704	0.03704
8(ii)	0.04639	0.01546	0.22165	0.15464	0.02062	0.13918	0.04124	0.29897	0.01031	0.04124	0.01031
8(iii)	0.01754	0.02924	0.15205	0.16959	0.00585	0.02339	0.14620	0.31287	0.00585	0.13158	0.00585
9(i)	0.29545	0.04545	0.15909	0.06818	0.04545	0.04545	0.04545	0.04545	0.15909	0.04545	0.04545
9(ii)	0.11111	0.07407	0.07407	0.18519	0.07407	0.07407	0.07407	0.11111	0.07407	0.07407	0.07407
9(iii)	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091
10(i)	0.08000	0.08000	0.12000	0.08000	0.12000	0.08000	0.08000	0.12000	0.08000	0.08000	0.08000
10(ii)	0.07018	0.05263	0.15789	0.12281	0.03509	0.07018	0.07018	0.22807	0.03509	0.12281	0.03509
10(iii)	0.06512	0.01860	0.08837	0.13023	0.00930	0.01395	0.10698	0.20930	0.00930	0.33953	0.00930
11(i)	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091
11(ii)	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091	0.09091
11(iii)	0.08333	0.08333	0.16667	0.08333	0.08333	0.08333	0.08333	0.08333	0.08333	0.08333	0.08333

Table 2: Transition matrices for the Markov chains of order one for the three time-homogeneous segments. We indicate by (i), (ii) and (iii) the transitions from states corresponding to the first, second and third time-homogeneous segments, respectively.

$j = 1$ and we consider any of the three time-homogeneous segments, we have that the most likely state that the volcano will present in the next day is "*no activity*". When considering the first time homogeneous segment, we have that the chance of going to another state is very small when compared to the transition "*no activity*" to "*no activity*". However, for the second and third time-homogeneous segments we have that there is also a not so small probability of going from "*no activity*" to "*rapid eruption of dense clouds*" and "*dense clouds*", respectively.

In general, when considering the probabilities of going from one state to another, the behaviour of the volcano's activity can be described as follows. In all three time-homogeneous segments, the probability of going from state j to state j, $j = 1, 2, 3, 4$ are larger than the transition from state j to a state l, $l \neq j$. In later days, observation 2001 to the end, we have that this behaviour persists for states $j = 5, 6, 7, 8, 10$ and is valid for $j = 6, 7, 8$ when considering the second time-homogeneous segment.

If we focus on the first one thousand observations and states $j = 5, 6, 8, 9$, then the most likely transition is to state $l = 1$, i.e., if the activity is either "*rapid eruption of dense smoke*", "*small and slow eruption*", "*medium size eruption*" or "*medium size, dense eruption*", then it is more likely that in the next day there will be "*no activity*". Still considering the first time-homogeneous segment, we have that for states $j = 7, 10, 11$ the behaviour is not as homogeneous. If we have that today there is a "*small, dense and rapid eruption*", then the most likely state tomorrow is "*dense cloud*". If today there is a "*large and dense eruption*", then the largest probability is of going to either "*steamy small clouds*", "*rapid eruption of dense smoke*" or "*medium size eruption*" with the same value. Since the occurrence of "*large and violent eruption*" is rare and the fact that there was none during the first one thousand days observed, we have that the transition from state $j = 11$ to any state is equally likely.

When considering the second time-homogeneous segment, we have that if today there is a "*rapid eruption of dense smoke*", then the most likely transitions are to either "*dense cloud*" or "*medium size eruption*" with the same probability. If today there is a "*medium size eruption*", then tomorrow the most likely transition is to state "*dense cloud*". When we have that today a "*large and dense eruption*" occurs, then it is more likely that tomorrow we will have "*steamy small clouds*". The state "*large and violent eruption*" produces the same conclusion as in the case of the first time-homogeneous segment.

During the 2001-st and 4847-th days observed there was only one occurrence of a "*large and violent eruption*" and therefore, the transition to different states produce different results. The most likely transition is to go to state "*steamy small clouds*". The state that produces equally likely transitions to any state in the third time-homogeneous segment is "*medium size, dense eruption*".

4.2. Condensed States

In this subsection we consider the condensed states obtained by grouping the eleven states corresponding to the classification given by [9] and presented in Table 1 into five states. These states are formed by grouping together states in Table 1 that can be considered very similar. Hence, we have the following partition. Take those values in Table 1 that are smaller or equal to $L_1 = 1$ (name them state 1); those values that are larger than L_1 and smaller or equal to $L_2 = 4$ (name them state 2); those values that are larger than L_2 and smaller or equal to $L_3 = 7$ (name them state 3); those values that are larger than L_3 and smaller or equal to $L_4 = 9$ (name them state 4) and those that are larger than L_4 (named state 5). These groups are classified as "*no activity*", "*small eruptions*", "*regular eruptions*", "*medium size eruptions*" and "*large eruptions*", respectively.

The identification of the time-homogeneous segments for the sequence $\mathbf{Y}$ in the present case is made in the same manner as when we had eleven states but now with $S = \{1, 2, 3, 4, 5\}$. In this case we also have three time-homogeneous segments. One is formed by the first three hundred observations. The second time-homogeneous segment ranges from the 301-st observation to the 2000-th. The last segment goes from observation 2001-st to the end of the data set.

When the first and the second time-homogeneous segments are considered the value of K that maximises the posterior distribution $P(K \mid \mathbf{Y})$ is 2. The third time-homogeneous segment is more likely to be governed by a Markov chain of order 3. Due to the size of the transition matrix of the chain of order 3 (125 lines and 5 columns) we omit it here. However, in Table 3 we present the transition matrices for the first two time-homogeneous segments. We have indicated by (i) and (ii) the lines corresponding to the first and second time-homogeneous segments, respectively. The rows correspond to the transformed states $\overline{m} \in \chi_2^{(K)}$. Hence the first row corresponds to state $(1, 1)$, i.e., two consecutive days with no activity at all.

As an example of how to use the information produced by the results presented here, assume that we consider the second time-homogeneous segment and that today and yesterday we have observed a "*small eruption*", and we want to know what the change is of having a "*regular eruption*" tomorrow, i.e.,

states	1	2	3	4	5
1(i)	0.865	0.04	0.04	0.045	0.01
1(ii)	0.75424	0.11441	0.06780	0.04661	0.01695
2(i)	0.56522	0.17391	0.08696	0.08696	0.08696
2(ii)	0.41892	0.41892	0.04054	0.09459	0.02703
3(i)	0.63636	0.09091	0.09091	0.09091	0.09091
3(ii)	0.57143	0.11905	0.14286	0.11905	0.04762
4(i)	0.54167	0.08333	0.16667	0.125	0.08333
4(ii)	0.35897	0.33333	0.20513	0.05128	0.05128
5(i)	0.2	0.2	0.2	0.2	0.2
5(ii)	0.25	0.25	0.16667	0.16667	0.16667
6(i)	0.56522	0.13043	0.13043	0.08696	0.08696
6(ii)	0.28947	0.51316	0.09211	0.07895	0.02632
7(i)	0.05405	0.86486	0.02703	0.02703	0.02703
7(ii)	0.04051	0.82658	0.04937	0.06835	0.01519
8(i)	0.18182	0.18182	0.18182	0.27273	0.18182
8(ii)	0.02857	0.71429	0.11429	0.1	0.04286
9(i)	0.2	0.2	0.2	0.2	0.2
9(ii)	0.07292	0.59375	0.11458	0.1875	0.03125
10(i)	0.2	0.2	0.2	0.2	0.2
10(ii)	0.125	0.5	0.16667	0.125	0.08333
11(i)	0.52174	0.13043	0.13043	0.13043	0.08696
11(ii)	0.55263	0.07895	0.18421	0.13158	0.05263
12(i)	0.27273	0.18182	0.18182	0.18182	0.18182
12(ii)	0.08861	0.39241	0.32911	0.16456	0.02532
13(i)	0.27273	0.18182	0.18182	0.18182	0.18182
13(ii)	0.10256	0.30769	0.30769	0.24359	0.03846
14(i)	0.27273	0.18182	0.18182	0.18182	0.18182
14(ii)	0.05085	0.27119	0.32203	0.30508	0.05085
15(i)	0.2	0.2	0.2	0.2	0.2
15(ii)	0.2	0.33333	0.13333	0.13333	0.2
16(i)	0.65217	0.08696	0.08696	0.08696	0.08696
16(ii)	0.325	0.1	0.225	0.3	0.05
17(i)	0.27273	0.18182	0.18182	0.18182	0.18182
17(ii)	0.05263	0.50526	0.10526	0.29474	0.04211
18(i)	0.18182	0.18182	0.27273	0.18182	0.18182
18(ii)	0.10169	0.32203	0.18644	0.32203	0.0678
19(i)	0.2	0.2	0.2	0.2	0.2
19(ii)	0.07895	0.39474	0.23684	0.23684	0.05263
20(i)	0.2	0.2	0.2	0.2	0.2
20(ii)	0.11765	0.29412	0.23529	0.23529	0.11765

Table 3: Transition matrices for the Markov chains of order two for the three time-homogeneous segments. We indicate by (i) and (ii) the transitions from states corresponding to the first and second time homogeneous segments, respectively.

we want to know the probability of the chain assuming value three in the next step, given that the states at the present and the previous one are 2. Then, we should look at the line $7(ii)$ which corresponds to the state $(2, 2)$ of $X^{(2)}$.

states	1	2	3	4	5
21(i)	0.2	0.2	0.2	0.2	0.2
21(ii)	0.25	0.16667	0.25	0.16667	0.16667
22(i)	0.2	0.2	0.2	0.2	0.2
22(ii)	0.1	0.43333	0.1	0.3	0.06667
23(i)	0.2	0.2	0.2	0.2	0.2
23(ii)	0.15385	0.23077	0.23077	0.15385	0.23077
24(i)	0.2	0.2	0.2	0.2	0.2
24(ii)	0.125	0.25	0.125	0.25	0.25
25(i)	0.2	0.2	0.2	0.2	0.2
25(ii)	0.15385	0.15385	0.30769	0.23077	0.15385

Table 3: Continuation

Therefore, the probability sought is 0.027. However, given that today and yesterday there were "*small eruptions*", tomorrow we will have a probability of 0.86 of having "*small eruption*".

5. Discussion

The results produced by the theoretical framework used here corroborate the general belief among volcanologists that there is a short term dependence of successive volcanic activities. This can be seen when using the detailed numerical classification associated to the original observed states given by [9]. The classification gives the several possible states and they are presented in an increasing level of activities without skipping any of them. The corroboration of the general belief follows when the order of the chain in the first case studied (eleven possible states) in one and the largest transition probabilities from one state are in general associated to the state itself and its immediate neighbours.

In addition to the results provided by the detailed classification, it is possible to see that when the condensed states are used we also have a small range of dependence (two or three). The increase in the order of dependence can be explained by the fact that to go from one condensed state to another, the volcano activity has in some cases to go through some of the intermediate states and therefore, more information is needed about the past behaviour of the volcanic activity to predict the probability of the next state. For instance, if we take the condensed state 2 and the state previous to it was condensed state 1, and look at the possible transitions that the activity may perform, then we have that the actual state can be any one of the different classification of steamy clouds. Depending on the intensity of clouds the activity could stay the same or go to a different state perhaps within the same category. Then we have also to know what the previous state was in order to infer if the activity is steadily increasing, decreasing or stable.

A refinement of the methodology presented here in order to make it useful as a predictive tool for varying activity in long term volcanic eruptions could be done if one had detailed information of currently ongoing eruptions all over the world. In that case, it would be suitable to analyse the volcanic observations from the Bayesian probabilistic point of view.

Acknowledgments

L.J. Álvarez thanks Don José Rogelio Álvarez for kindly providing an original copy of the *Efemérides del Volcán de Colima* for its study and also thanks Guadalupe Núñez for patiently digitising the data. E.R. Rodrigues was partially funded by CONACyT Project Number 45684-F. We thank David Flores for providing a subroutine that allows for infinite precision calculations. All calculations were performed on the supercomputer granted by Fundación Clínica Médica Sur, A. C. to Instituto de Matemáticas, Unidad Cuernavaca, UNAM.

References

[1] J.A. Achcar, A.A. Fernández-Bremauntz, E.R. Rodrigues, G. Tzintzun, Estimating the number of ozone peaks in Mexico City using a non-homogeneous Poisson model, *Environmetrics*, **19** (2008), 469-485.

[2] L.J. Álvarez, N.L. Garcia, E.R. Rodrigues, Comparing the performance of a reversible jump Markov chain Monte Carlo algorithm for DNA sequences alignment, *Journal of Statistical Computation and Simulation*, **76** (2006), 567-584.

[3] L.J. Álvarez, A.A. Fernández-Bremauntz, E.R. Rodrigues, G. Tzintzun, *Maximum a posteriori* estimation of the daily ozone peaks in Mexico City, *Journal of Agricultural, Biological and Environmental Statistics*, **10** (2005), 276-290.

[4] R.J. Boys, D.A. Henderson, On determining the order of Markov dependence of an observed process governed by a hidden Markov model, *Special Issue of Scientific Programming*, **10** (2002), 241-251.

[5] R.J. Boys, D.A. Henderson, Bayesian approach to DNA segmentation (with discussion), *Biometrics*, **60** (2004), 573-588.

[6] R.J. Boys, D.A. Henderson, D.J. Wilkinson, Detecting homogeneous segments in DNA sequences by using hidden Markov models, *Applied Statistics*, **49** (2000), 269-285.

[7] G.A. Churchill, B. Lazareva, Bayesian restoration of a hidden Markov chain with applications to DNA sequencing, *Journal of Computational Biology*, **6** (1999), 261-277.

[8] S. de la Cruz-Reyna, Random patterns of occurrence of explosive eruptions at Colima volcano, México, *Journal of Volcanology and Geothermal Res.*, **55** (1993), 51-68.

[9] S. Díaz, *Efemérides del Volcán de Colima*, Imprenta y Fototipia de la Secretaría de Fomento, Spain (1906).

[10] M. Evans, T. Swartz, *Approximating Integrals via Monte Carlo and Deterministic Methods*, Oxford Statistical Sciences Series **20**, Oxford University Press, UK (2000).

[11] T.-H. Fan, C.-A. Tsai, A Bayesian method in determining the order of a finite state Markov chain, *Communications in Statistics - Theory and Methods*, **28** (1999), 1711-1730.

[12] P.A. Ferrari, A. Frigessi, P.G. Gonzaga de Sá, Fast approximate maximum a posteriori restoration of multicolour images, *Journal of the Royal Statistical Society Series B*, **57** (1995), 485-500.

[13] D. Geman, Random fields and inverse problems in imaging, *Lecture Notes in Mathematics*, Springer-Verlag, USA, **1427** (1990), 113-193.

[14] F. Medina-Martínez, Analysis of the eruptive history of the Volcán de Colima, México (1560-1980), *Geophys. J. Int.*, **22** (1983), 127-178.