Earth Science Informatics (2022) 15:1635–1645
https://doi.org/10.1007/s12145-022-00810-x

RESEARCH ARTICLE

Poisson hidden markov model on earthquake occurrences in Metro Manila, Philippines

Edd Francis O. Felix[1] · Christian Alvin H. Buhat[1,2] · Jonathan B. Mamplata[1]

Received: 3 May 2021 / Accepted: 17 April 2022 / Published online: 20 May 2022

Abstract

The Philippines, as part of the Circum-Pacific belt, is considered as one of the most seismically active countries in the world. Earthquake occurrence is frequent and its effects vary depending on its size. Understanding how the occurrences happen is therefore important. Stochastic models of earthquake occurrence have been used to study seismic activities in various active earthquake zones globally. In this paper, we apply Poisson hidden Markov models (PHMM) using the January 1, 1960 to January 20, 2019 earthquake data of Metro Manila, Philippines. The parameters in the models are estimated using expectation-maximization (EM) algorithm. We determine using various statistical tests that the 5-state PHMM best represents the earthquake data and implement bootstrap algorithm to validate the acceptability of its parameter estimates. Moreover, we investigate the forecasting capability of the 5-state PHMM by comparing it to the ARIMA model. Using unscaled mean bounded relative absolute error (UMBRAE), we find that the 5-state PHMM gives closer one-step ahead forecasts and is a better forecasting model for the considered data.

Keywords Poisson hidden Markov model · EM algorithm · Earthquake prediction · Philippine earthquakes

Introduction

Earthquakes can be one of the most devastating natural phenomena. The randomness of its occurrence as well as its size make this event a potentially life-threatening disaster that could affect thousands of people and infrastructures (Kannan 2014). Analysis of earthquake data is therefore important to better understand earthquake behavior and create mechanism in countering its probable effects (Yip et al. 2017).

Stochastic modeling has been a traditional approach to study earthquake occurrences. Models such as Poisson (Utsu 1969; Lomnitz 1974; Dionysiou and Papadopoulos 1992) and negative binomial (Dionysiou and Papadopoulos 1992; Rao and Kaila 2010) models have been used to describe earthquake occurrences. These models assume that the previous occurrences do not affect the time for which the next one will occur. However, earthquakes of large magnitude could cause other large earthquakes to occur consecutively in a short span of time (Kannan 2014). This means that the occurrences are self-exciting, i.e, the previous occurrences make the future occurrences more likely to happen. Some researches on modelling the seismicity such as Bansal et al. (2012) and Spassiani and Sebastiani (2016) used the epidemic-type aftershock sequence (ETAS) model. ETAS model is a point process on modelling seismicity that is based on three assumptions: (i) the background seismicity follows the Poisson distribution, (ii) the number of aftershocks is proportional to $\exp(\alpha M)$, and (iii) the decrease in the number of aftershocks follows the modified Omori Law (Bansal et al. 2012). The ETAS model would be impossible to apply on the available

Communicated by: H. Babaie

✉ Jonathan B. Mamplata
jbmamplata@up.edu.ph

Edd Francis O. Felix
eofelix@up.edu.ph

Christian Alvin H. Buhat
chbuhat@uh.edu

[1] Institute of Mathematical Sciences and Physics, University of the Philippines Los Baños, Los Baños, Philippines

[2] Mathematics Department, University of Houston, Houston, TX, 77204, USA

local data since it is difficult to determine which among the earthquake occurrences are aftershocks or mainshocks. Hence, Markov model would be a suitable alternative since it uses information from the immediate past.

A hidden Markov model (HMM) is a stochastic process that involves random variables characterized as either an observed process or a hidden process. The hidden process is a sequence of unobservable events that directly affects the observed process. It is assumed to be a Markov process that governs the distribution of the observed process. One of the early works on the application of HMM was by Rabiner (1989) where they used the model in advanced speech recognition problems. Related studies also show how HMM can be used in molecular biology (Krogh et al. 2001), genetics (Pachter et al. 2002), engineering (Goh et al. 2012) and notably, seismology (Doganer and Calik 2013; Yip et al. 2017). The study of Yip et al. (2017) developed a novel HMM in modeling and predicting earthquakes. They used HMM recognizing that while earthquake occurrences are observable, the underlying underground dynamics, which involves the stress level around faults, are not. Through their model, they predicted the arrival time and magnitude of future earthquakes simultaneously using the data from the Southern California earthquake catalogues from 1981 to 2015. Meanwhile, Doganer and Calik (2013) focused on the use of HMM with forward algorithm in estimating the epicenter of many occurring earthquakes in East Anatolian Fault Zone. The use of HMM aided them in considering times of seismic inactivity. Their results showed a 0.73 chance of earthquake occurrence in Sincik- Lake Hazar. Different distributions can be used as the state-dependent probability distribution since HMM can be applied to both discrete and continuous data. This paper will employ Poisson distribution due to the discrete characteristic of the seismic data. Moreover, using Poisson hidden Markov model (PHMM) addresses the problem of over-dispersion of the data which is typically the case in earthquake data (Can et al. 2014).

Recent researches in PHMM include the application of this model to studies in video traffic (Rossi et al. 2015), infrastructure deterioration (Le Thanh et al. 2015), and insurance (Paroli et al. 2002). The study of Orfanogiannaki et al. (2011) introduced PHMM in modeling temporal seismicity changes. According to them, a PHMM can reveal unknown attributes of the earthquake mechanisms that produced the seismic data by providing a way to estimate the underlying hidden states of the system. Using PHMM, they were able to model the earthquake frequencies with local magnitude $M_L > 3.2$ in the seismogenic area of Killini, Ionian Sea, Greece, in the period 1990-2006. They allowed them to capture short-term precursory seismicity changes preceding strong mainshocks which the traditional analysis failed to recognize in the 1997 mainshock. Meanwhile, Can et al. (2014) applied PHMM to predict earthquake hazards in Bilecik, NW Turkey. They considered the annual frequencies of earthquakes occurring around the area from January 1900 to December 2012, with magnitude $M > 4$, and forecasted earthquake hazards for the years 2013-2047. In 2018, Orfanogiannaki et al. (2018) considered two main earthquakes that occurred between the Indo-Australian and the southeastern Eurasian plates, and used PHMM in identifying the temporal patterns in the time series of those two earthquakes. Their results showed the low seismicity in the region 400 days prior to the first earthquake, and a shift from low to high seismicity in between the two main earthquakes. The work of Orfanogiannaki and Karlis (2018) introduced the use of multivariate Poisson hidden Markov models in modeling earthquake occurrences. Each state of the multivariate model is associated with a different multivariate discrete distribution. They apply their model to the seismicity with magnitude $M > 5$ in three seismogenic subregions in the North Aegean Sea 1981 to 2008. Their results proved the migration of seismicity in adjacent subregions that share similar seismotectonic feature.

The Philippines is considered to have one of the most complex regions of plate interaction in the circum-pacific belt (Hopkins et al. 1991). This results to the high seismic activity in the country. Some of the recent memorable instances include the successive destructive earthquakes of magnitude higher than 6 that occurred in Mindanao as well as in other parts of the country in 2019 (Rappler.com 2019a, b; PHIVOLCS 2019a, b), the earthquake swarm that happened in Batangas in 2017 (PHIVOLCS 2017), and the 2013 Bohol earthquake that caused broken roads and damaged buildings (Rappler.com 2013). While there have been numerous studies on the use of PHMM in understanding seismic activities, none of these have used seismic data from the Philippines. This paper aims to apply PHMM in examining earthquake occurrences using seismic data of Metro Manila.

The paper is organized as follows: in "Poisson hidden Markov model", we discuss the PHMM and how the expectation-maximization (EM) algorithm is used to estimate the model's parameters. In "Numerical implementation", we present the results of the parameter estimations for PHMMs with different number of states. We use these results to determine which PHMM best describes the data. In "Benchmarking", we perform short-term forecasting using the PHMM and ARIMA model and compare the results. Lastly, in "Conclusion and recommendation", we present our conclusions and recommendations.

Poisson hidden Markov model

Let $(\Omega, \mathcal{F}, P)$ be the probability space such that $\mathbf{z}_k$ is a Markov chain in discrete time $k = 0, 1, 2, \ldots$. The Markov chain evolves according to the dynamics

$$\mathbf{z}_k = A\mathbf{z}_{k-1} + \mathbf{v}_k \tag{1}$$

where $\mathbf{v}_k$ is a martingale increment, that is,

$$E[\mathbf{v}_k|\mathcal{F}_k^{\mathbf{z}}] = 0 \tag{2}$$

where $\mathcal{F}_k^{\mathbf{z}}$ is the filtration generated by $\{\mathbf{z}_0, \mathbf{z}_1, ..., \mathbf{z}_k\}$. The Markov chain $\mathbf{z}_k$ represents the state, so if we are working with m states, then $\mathbf{z}_k \in \mathbb{R}^m$. In addition, $\mathbf{z}_k$ is a linear combination from the set $\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_m\}$ where $\mathbf{e}_i = (0, ..., 0, 1, 0, ..., 0)^\top \in \mathbb{R}^m$ is a zero vector with 1 on the i^{th} component, which is the canonical basis of $\mathbb{R}^m$. A is the transition probability matrix with entries $a_{ij}, i, j = 1, 2, ..., m$ which is the probability of transition from state i to state j from time $k-1$ to time k, that is,

$$a_{ij} = P(\mathbf{z}_k = \mathbf{e}_j|\mathbf{z}_{k-1} = \mathbf{e}_i) = P(\mathbf{z}_1 = \mathbf{e}_j|\mathbf{z}_0 = \mathbf{e}_i). \tag{3}$$

It is clear that $\sum_{j=1}^{m} a_{ij} = 1$ for $i = 1, ..., m$. Let $\xi = (\xi_1, ..., \xi_m)^\top$ be the marginal distribution of the initial Markov chain, $\mathbf{z}_0$, such that $\xi_i = P(\mathbf{z}_0 = \mathbf{e}_i)$. ξ represents the initial distribution of the Markov chain and $\sum_{i=1}^{m} \xi_i = 1$.

Furthermore, let Y_k, $k = 0, 1, 2, ...$ be the distribution of the observed process that depends only on $\mathbf{z}_k$. In other words, $\{Y_k\}$ is a sequence of conditionally independent random variables given the state process $\{\mathbf{z}_k\}$. In this study, we assume that Y_k given $\mathbf{z}_k$ is a Poisson random variable, thus we have the term Poisson hidden Markov model. The state space $\mathbf{z}_k$ determines the parameter of the Poisson process used to generate Y_k. See Paroli et al. (2002) for a comprehensive discussion on PHMM.

Let $\lambda = (\lambda_1, ..., \lambda_m)^\top$ be the parameter space for the observed process Y_k, that is λ_i is the parameter of the Poisson process at state i for $i = 1, ..., m$. Let η_{yi} be the conditional probability of Y_k given that the process is in state i, that is,

$$\eta_{yi} = P(Y_k = y|\mathbf{z}_k = \mathbf{e}_i) = e^{-\lambda_i}\frac{\lambda_i^y}{y!} \tag{4}$$

Note also here that $\sum_{y=0}^{\infty} \eta_{yi} = 1$ for $i = 1, ..., m$.

The entire process given by $\{(\mathbf{z}_k, Y_k)\}$ is what we are looking for. The Markov chain $\mathbf{z}_k$ is the latent process that has a semi-martingale representation given in (1). The process Y_k depends on the latent process $\mathbf{z}_k$ and this is depicted in Figure 1. The processes $\{\mathbf{z}_k\}$ and $\{Y_k\}$ are stationary processes, thus each Y_k has the same distribution for any value of k. Getting the probability mass function for Y_k, we have

$$\begin{aligned} P(Y_k = y) &= \sum_{i=1}^{m} P(Y_k, \mathbf{z}_k = \mathbf{e}_i) \\ &= \sum_{i=1}^{m} P(Y_k = y|\mathbf{z}_k = \mathbf{e}_i)P(\mathbf{z}_k = \mathbf{e}_i) \\ &= \sum_{i=1}^{m} \xi_i \eta_{yi} \end{aligned}$$

We can also observe that $E[Y_k] = \langle \xi, \lambda \rangle$ where $\langle \cdot, \cdot \rangle$ is the usual inner product.

Let Θ be the parameter space containing the set of plausible parameters for the process $\{(\mathbf{z}_k, Y_k)\}$. If $\theta \in \Theta$ is the maximum likelihood estimate for the process $\{(\mathbf{z}_k, Y_k)\}$, then

$$\theta = (a_{11}, a_{12}, ..., a_{mm}, \lambda_1, \lambda_2, ..., \lambda_m)^\top. \tag{5}$$

θ contains the entries of the transition probability matrix and the elements of the parameter space of the Poisson process. The transition probability matrix contains m elements but we know that $\sum_{j=1}^{m} a_{ij} = 1$ for $i = 1, ..., m$, so we can solve only for $m^2 - m$ entries of A. Setting the entries of the diagonal of A dependent on other entries of A, that is, $a_{ii} = 1 - \sum_{j=1, j\neq i}^{m} a_{ij}$. So, we can reduce θ by

$$\theta = (a_{12}, a_{13}, ..., a_{m,m-1}, \lambda_1, \lambda_2, ..., \lambda_m)^\top \tag{6}$$

which contains m^2 elements.

Suppose we have the observed process $\{y_0, y_1, ..., y_K\}$ up to time K and $\{\mathbf{z}_0, \mathbf{z}_1, ..., \mathbf{z}_K\}$ be the Markov chain process of the state up to time K. Thus, the set $\{\mathbf{z}_0, Y_0, \mathbf{z}_1, Y_1, ..., \mathbf{z}_K, Y_K\}$ is the set of complete data. The

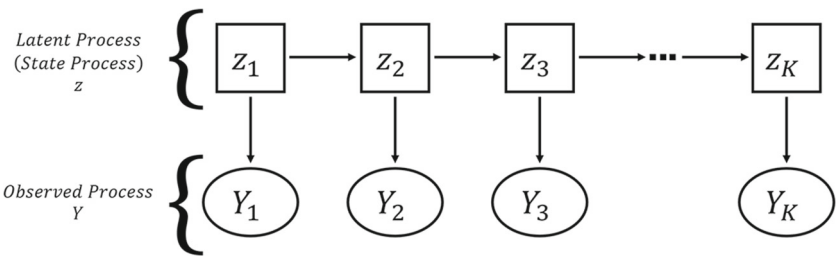

Fig. 1 Hidden Markov process

likelihood function for this set, denoted by $L^c(Y;\Theta)$ is given by

$$\begin{aligned} L^c(Y;\Theta) &= P(Y_0 = y_0, Y_1 = y_1, ..., Y_K = y_K, \mathbf{z}_0 = \mathbf{e}_{i_0}, ..., \mathbf{z}_K = \mathbf{e}_{i_K}) \\ &= \xi_{i_0}\eta_{y_0,i_0}\prod_{k=1}^{K} a_{i_{k-1},i_k}\eta_{y_k,i_k} \end{aligned} \tag{7}$$

Since the Markov process is a latent process, getting the sum over all possible values of the state process, we obtain the likelihood function of the incomplete data, given by

$$L(Y;\Theta) = \sum_{i_0=1}^{m}\sum_{i_1=1}^{m}\cdots\sum_{i_K=1}^{m}\xi_{i_0}\eta_{y_0,i_0}\prod_{k=1}^{K} a_{i_{k-1},i_k}\eta_{y_k,i_k} \tag{8}$$

where η_{y_k,i_k} is the state-dependent probability of y_k conditioned on state $\mathbf{z}_k$ given by

$$\eta_{y_k,i_k} = e^{-\lambda_{i_k}}\frac{\lambda_{i_k}^{y_k}}{y_k!}. \tag{9}$$

Solving for maximum likelihood estimates of the parameters is finding θ such that

$$\theta = \text{argmax } L(Y;\Theta) \tag{10}$$

Optimizing (8) is analytically intractable so we implement numerical method for estimating the parameters. We perform expectation-maximization (EM) algorithm to numerically estimate the parameters, see Ryden (1996) for a comprehensive discussion on this. EM algorithm is an iterative process that involve two main steps: E-step, the expectation step and M-step, the maximization step. Let

$$Q(\theta;\hat{\theta}) = E_{\hat{\theta}}[\log L_K^c(Y;\Theta)|\mathbf{y}] \tag{11}$$

where $\hat{\theta}\in\Theta$ and $\mathbf{y}$ is a vector of realized process for Y. The EM algorithm is described as follows:

1. Choose $\hat{\theta}_0\in\Theta$ such that $\hat{\theta}_0$ is a good approximate for the parameters.
2. E-step is calculating for (11), that is, finding $Q(\theta;\hat{\theta}_0)$ defined in (11).
3. M-step is finding $\hat{\theta}_{n+1}$ such that $Q(\hat{\theta}_{n+1};\hat{\theta}_n) \geq Q(\theta;\hat{\theta}_n)$ where
$$\hat{\theta}_{n+1} = \text{argmax}_{\theta\in\Theta}\ Q(\theta;\hat{\theta}_n). \tag{12}$$
4. The E and M steps are repeated in an alternating way until an optimal estimate is achieved, that is,
$$\left|\log L_K(\hat{\theta}_{n+1}) - \log L_K(\hat{\theta}_n)\right| < \epsilon \tag{13}$$
for some tolerance error ϵ.

Thus, we conclude that

$$\hat{\theta}_{n+1} = \left(a_{12}^{(n+1)}, a_{13}^{(n+1)}, ..., a_{m,m-1}^{(n+1)}, \lambda_1^{(n+1)}, \lambda_2^{(n+1)}, ..., \lambda_m^{(n+1)}\right)^\top \tag{14}$$

is an optimal estimate for the process $\{\mathbf{z}_k, Y_k\}$.

EM algorithm may be simplified by using forward and backward probabilities, commonly known as Baum-Welch algorithm (Baum et al. 1970; Baum and Petrie 1966). We denote $\alpha_k(i)$ be the forward probabilities of the past observations up to the present with the current state, then it is given by

$$\alpha_k(i) = P(Y_0 = y_0, Y_1 = y_1, ..., Y_k = y_k, \mathbf{z}_k = \mathbf{e}_i) \tag{15}$$

and the backward probabilities, $\beta_k(i)$, which is the probabilities of the future observations conditioned on the current state, that is,

$$\beta_k(i) = P(Y_{k+1} = y_{k+1}, ..., Y_K = y_K|\mathbf{z}_k = \mathbf{e}_i). \tag{16}$$

It can be shown that forward probabilities may be derived recursively as

$$\begin{aligned} \alpha_0(i) &= \xi_i\eta_{y_0,i} && \text{for } i = 1, ..., m \\ \alpha_k(j) &= \sum_{i=1}^{m}\alpha_{k-1}(i)a_{ij}\eta_{y_k,j} && \text{for } k = 1, ..., K \text{ and } j = 1, ..., m \end{aligned} \tag{17}$$

and backward probabilities as

$$\begin{aligned} \beta_K(i) &= 1 && \text{for } i = 1, ..., m \\ \beta_k(i) &= \sum_{j=1}^{m}\eta_{y_{k+1},j}\beta_{k+1}(j)a_{ij} && \text{for } k = K-1, ..., 0 \text{ and } j = 1, ..., m \end{aligned} \tag{18}$$

Evaluating (11) at the nth iteration of the parameters, $\hat{\theta}_n$, we get

$$\begin{aligned} Q(\theta,\hat{\theta}_n) = &\sum_{i=1}^{m}\frac{\alpha_0^{(n)}(i)\beta_0^{(n)}(i)}{\sum_{l=1}^{m}\alpha_k^{(n)}(l)\beta_k^{(n)}(l)}\log\xi_i \\ &+\sum_{i=1}^{m}\sum_{j=1}^{m}\frac{\sum_{k=0}^{K}\alpha_k^{(n)}(i)a_{ij}^{(n)}\eta_{y_{k+1},j}^{(n)}\beta_{k+1}^{(n)}(j)}{\sum_{l=1}^{m}\alpha_k^{(n)}(l)\beta_k^{(n)}(l)}\log a_{ij} \\ &+\sum_{i=1}^{m}\frac{\sum_{k=0}^{K}\alpha_k^{(n)}(i)\beta_k^{(n)}(i)}{\sum_{l=1}^{m}\alpha_k^{(n)}(l)\beta_k^{(n)}(l)}\log\eta_{y_k,i} \end{aligned}$$

where $\eta_{y_k,i}^{(n)}$, $\alpha_k^{(n)}(i)$ and $\beta_k^{(n)}(i)$ are derived based on $\hat{\theta}_n$ obtained at the nth iteration following EM algorithm. Thus, the maximum likelihood estimate of the entries of transition probability matrix derived at the $(n+1)^{\text{th}}$ iteration via EM algorithm is given by

$$a_{ij}^{(n+1)} = \frac{\sum_{k=0}^{K-1}\alpha_k^{(n)}(i)a_{ij}^{(n)}\eta_{y_{k+1},j}^{(n)}\beta_{k+1}^{(n)}(j)}{\sum_{k=0}^{K-1}\alpha_k^{(n)}(i)\beta_k^{(n)}(i)} \text{ for } i, j = 1, ..., m \tag{19}$$

and the parameter for the Poisson process is given by

$$\lambda_i^{(n)} = \frac{\sum_{k=0}^{K} \alpha_k^{(n)}(i)\beta_k^{(n)}(i)\, y_k}{\sum_{k=0}^{K} \alpha_k^{(n)}(i)\beta_k^{(n)}(i)} \text{ for } i = 1, ..., m. \quad (20)$$

You may refer to Elliott et al. (1995) for a comprehensive discussion on parameter estimation.

Numerical implementation

We examine the earthquake data of Metro Manila (12^o - 17^oN Latitude, 119^o-123^oE Longitude) obtained from the DOST Philippine Institute of Volcanology and Seismology (DOST-PHIVOLCS). In particular, we consider the earthquake occurrences of magnitude greater than or equal to 4. It should be noted, however, that the magnitude entries in the data are either in local magnitude scale (M_l), body wave magnitude scale (M_b), or surface wave magnitude scale (M_s). We convert all the M_l and M_b entries to M_s using the formulas given by $M_s = \frac{M_b - 2.5}{0.63}$ and $M_s = -3.2 + 1.45M_l$ (Tobyás and Mittag 1991).

We determine the frequencies of earthquake occurrences over 30-day intervals from January 1, 1960 to January 20, 2019, and record 719 observed values whose summary is shown in Figure 2. We infer from the data that there is an over-dispersion of earthquake occurrences over 30-day intervals since the value of the standard deviation, 5.57309, is greater than the average number of earthquake occurrences, 1.64534.

The parameters of the model were estimated using MLE as described in the previous section. We implemented EM algorithm using a fixed set of states. Shown in Table 1 are the estimated parameters with the corresponding transition probability matrix from Poisson process to 6-state PHMM.

Based from the table, if we consider the Poisson process, the mean number of earthquakes in every 30-day interval is 1.645341. In the two-state model, the estimated parameters are 1.064042 and 21.212017, which shows two average

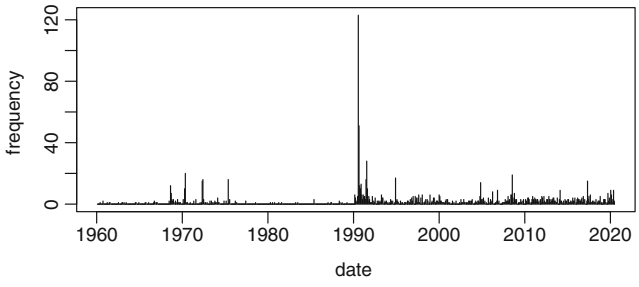

Fig. 2 30-day frequencies of earthquake occurrences from January 1, 1960 to January 20, 2019

number of occurrences that are relatively far from each other. Clearly, the estimated λ_2 accounts for the 30-day periods with high number of earthquake occurrences. Shown in Figure 3 is the distribution of states under the two-state regime. Based from the transition probability matrix, there is a minute chance that the number of occurrences on the next 30-day period is high. From 4-state to 6-state model, notice that one state has an average of 87 earthquake occurrences. From these three models, the probability that the next 30-day period stays in the same state is 0.5 while and the probability that it will shift to a lower state is also 0.5. The summary of state distributions under the two-state up to the 6-state regimes is shown in the Appendix.

In order to choose the best model, we use two metrics: Akaike information criterion (AIC) and Bayesian information criterion (BIC). The quality of the model relative to the other proposed model can be estimated by AIC (Akaike 1974). BIC, on the other hand, is similar to AIC that uses likelihood function to choose the most suitable model for the given data sets (Schwarz 1978). The formula for AIC and BIC are given by

$$AIC = 2p - 2\log L$$

$$BIC = p\log n - 2\log L$$

where p is the number of parameters estimated in the model, n is the number of data points considered and L is the likelihood value.

Shown in Table 2 are the AIC and BIC for the various models we implemented in this paper. AIC suggests that 5-state PHMM is the most suitable model while BIC suggests that 4-state suits the data sets.

Since these two metrics do not agree with model selection, we use likelihood ratio test to determine the better model. Let L_0 and L_1 be the maximum loglikelihood of the 4-state and 5-state PHMMs, respectively. We obtain the values $L_0 = -1029.098$ and $L_1 = -1002.291$. To perform the likelihood ratio test, we consider the following hypotheses:

H_0: The 4-state PHMM is better than the 5-state PHMM.
H_a: The 4-state PHMM is not better than the 5-state PHMM.

We calculate the χ^2 test statistic as follows:

$$\chi^2 = -2\,(L_0 - L_1) = 53.614$$

We choose $\alpha = 0.05$ and set the degrees of freedom (df) to 9, which is the difference in the number of parameters of the two models. With $\alpha = 0.05$ and $df = 9$, the corresponding value from the chi square distribution table is 16.919. Since $\chi^2 = 53.614 > 16.919$, we reject H_0. Thus, at 95% level of significance, we conclude that the 5-state PHMM is better than the 4-state PHMM.

Table 1 Estimated parameters and associated transition probability matrices

Model	$\hat{\lambda}$	Transition probability matrix
Poisson process	$\hat{\lambda} = 1.645341$	
2-state PHMM	$\hat{\lambda}_1 = 1.064042$ $\hat{\lambda}_2 = 21.212017$	$\begin{bmatrix} 0.9774 & 0.0226 \\ 0.7028 & 0.2972 \end{bmatrix}$
3-state PHMM	$\hat{\lambda}_1 = 0.2325329$ $\hat{\lambda}_2 = 2.0027568$ $\hat{\lambda}_3 = 23.2048761$	$\begin{bmatrix} 0.9731 & 0.0099 & 0.0170 \\ 0.0478 & 0.4395 & 0.5128 \\ 0.0242 & 0.0191 & 0.9566 \end{bmatrix}$
4-state PHMM	$\hat{\lambda}_1 = 0.2195573$ $\hat{\lambda}_2 = 1.8471073$ $\hat{\lambda}_3 = 13.2795612$ $\hat{\lambda}_4 = 87.000000$	$\begin{bmatrix} 0.9797 & 0.0094 & 0.0110 & 0.0000 \\ 0.0145 & 0.9585 & 0.0241 & 0.0028 \\ 0.0525 & 0.5725 & 0.3750 & 0.0000 \\ 0.0000 & 0.0000 & 0.5000 & 0.5000 \end{bmatrix}$
5-state PHMM	$\hat{\lambda}_1 = 0.1744851$ $\hat{\lambda}_2 = 0.851076600$ $\hat{\lambda}_3 = 2.463019200$ $\hat{\lambda}_4 = 13.948613400$ $\hat{\lambda}_5 = 87.000000000$	$\begin{bmatrix} 0.9927 & 0.0000 & 0.0032 & 0.0041 & 0.0000 \\ 0.0000 & 0.9156 & 0.0530 & 0.0314 & 0.0000 \\ 0.0055 & 0.0746 & 0.8973 & 0.0178 & 0.0048 \\ 0.0000 & 0.1354 & 0.5075 & 0.3571 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.5000 & 0.5000 \end{bmatrix}$
6-state PHMM	$\hat{\lambda}_1 = 0.074634240$ $\hat{\lambda}_2 = 0.690675720$ $\hat{\lambda}_3 = 0.988455300$ $\hat{\lambda}_4 = 3.051946270$ $\hat{\lambda}_5 = 14.808549700$ $\hat{\lambda}_6 = 87.000000180$	$\begin{bmatrix} 0.8650 & 0.1310 & 0.0000 & 0.0040 & 0.0000 & 0.0000 \\ 0.4814 & 0.4874 & 0.0000 & 0.0000 & 0.0312 & 0.0000 \\ 0.0000 & 0.0000 & 0.8519 & 0.1319 & 0.0162 & 0.0000 \\ 0.0000 & 0.0000 & 0.2494 & 0.7201 & 0.0238 & 0.0067 \\ 0.0000 & 0.1202 & 0.0000 & 0.5064 & 0.3734 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.5000 & 0.5000 \end{bmatrix}$

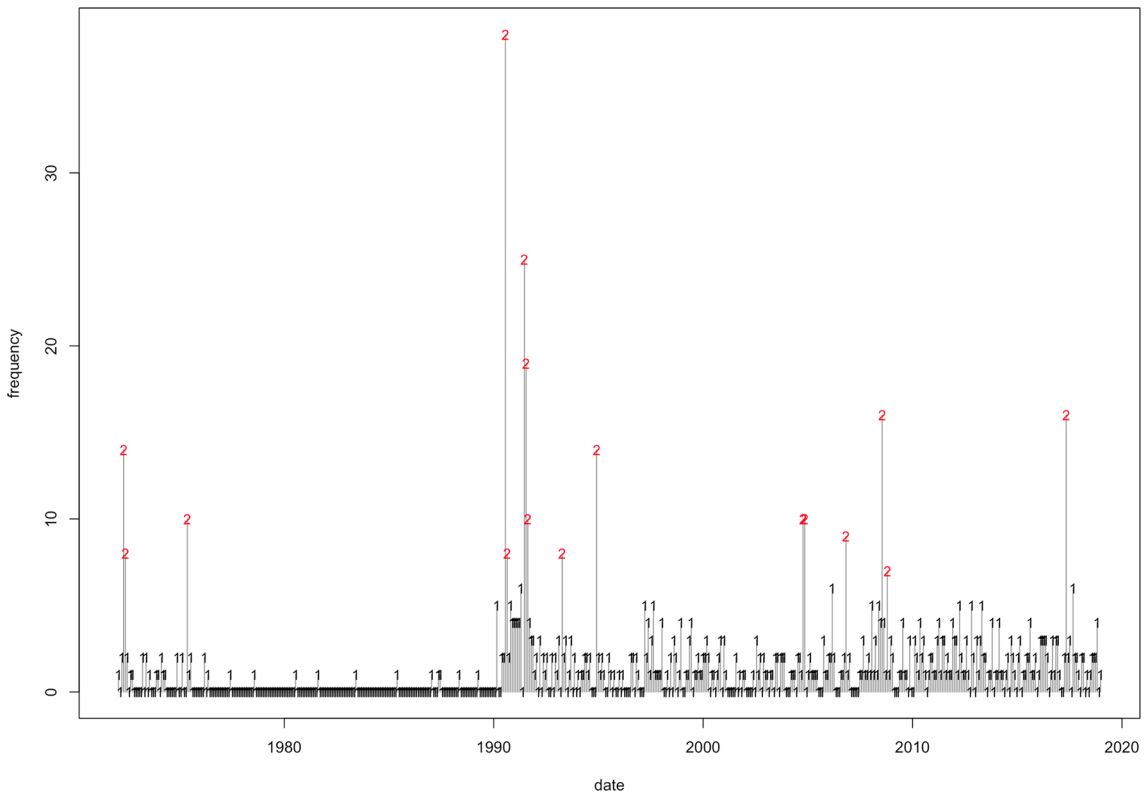

Fig. 3 State distribution under 2-state PHMM

Table 2 Values of AIC and BIC

Model	p	AIC	BIC
Poisson process	1	4249.852	4254.430
2-state PHMM	4	2801.154	2824.043
3-state PHMM	9	2341.429	2391.785
4-state PHMM	16	2096.197	2183.176
5-state PHMM	25	2062.582	2195.340
6-state PHMM	36	2071.253	2258.945

Acceptability of the 5-state PHMM parameter estimates To investigate the acceptability of the parameters of 5-state PHMM, we perform bootstrapping. We partition the original data set into subintervals of 30 points and resample it with replacement from each subinterval. From the resampled data points, we perform estimation of parameters using the algorithm presented. We select 10,000 set of bootstrap samples; then, we calculate for the average value of each of the parameter estimates (Xiong and Mamon 2017). Table 3 summarizes the values of each parameters obtained from bootstrapping, and their corresponding confidence interval. The mean, std dev, 95% lower cl and 95% upper cl represent the average, standard deviation, 2.5% quantile and 97.5% quantile of the 10,000 bootstrap samples, respectively.

Notice that as the estimated parameter increases, the corresponding standard deviation from the bootstrap sample increases as well. This indicates that the state with a higher parameter would have larger fluctuations on the number of earthquake occurrences over the 30-day period. Also, all the parameters have values that lie inside their corresponding confidence interval. This means that their values are acceptable estimates of the average 30-day earthquake occurrences.

5-state PHMM on the Metro Manila earthquake data Shown in Figure 4 are different portions of the graph of the 30-day earthquake occurrences in Metro Manila. The graphs also show the major earthquake occurrences and in which 30-day interval they are a part of. Take note that $\hat{\lambda}_1 = 0.174485100$, $\hat{\lambda}_2 = 0.851076600$, $\hat{\lambda}_3 = 2.463019200$, $\hat{\lambda}_4 = 13.948613400$, and $\hat{\lambda}_5 = 87$.

We observe that most of the major earthquakes are part of a 30-day interval that is currently in state 4, i.e., the interval has an average of 13.95 earthquake occurrences. It may be due to the occurrence of foreshocks and aftershocks. A large earthquake is causally preceded by foreshock or multiple-shock activities (Fukao and Furumoto 1975) and the stress increase caused by the major earthquake could result to the occurrence of subsequent minor earthquakes (King et al. 1994). We also notice that the July 22 - August 21, 1990 interval has the most number of earthquake occurrences with 123, and the next interval has the second most occurrences with 51. These are the only intervals that are in state 5, i.e, the interval has an average of 87 earthquake occurrences. The major Luzon earthquake with magnitude 7.8 that happened in July 16, 1990 could have triggered many aftershocks. This major earthquake is also considered to have a possible relation to the eruption of Mt. Pinatubo in July 1991 (Bautista et al. 1996). Lastly, we observe that whenever the 30-day interval is in state 4, whether one-time or consecutive times, the next interval will be in state 2 or 3 which means that the average number of earthquake occurrences in the interval is 0.85 or 2.46, respectively. This is consistent with the explanation in (Yip et al. 2017) that the underground stress goes back to normal after it builds up, reaches a certain threshold, and gets released in the form of an earthquake.

Benchmarking

In forecasting time series, a common model is the Autoregressive integrated moving average (ARIMA) (Fattah et al. 2018) model. Several ARIMA models of varying time steps have been used in forecasting large-scale earthquake occurrences through fitting the model and smoothing the data with a sequence of empirical recurrence rates (ERR) time series (Amei et al. 2012; Ho and Bhaduri 2015). In the study of Ho and Bhaduri (2015), historical earthquake data of Parkfield was modeled through applying ARIMA techniques to an ERR time series. Such applications were also done to similar time series data sets from hurricanes and volcanoes (Ho and Bhaduri 2017; Bhaduri and Ho 2019). For our earthquake time series, ARIMA was applied to forecast future occurrences.

Through the ARIMA R package (Hyndman and Khandakar 2008), a best fit model was used for a one-step-ahead

Table 3 Results from bootstrap sampling

	Estimated parameter	Mean	Std dev	95% lower cl	95% upper cl
$\hat{\lambda}_1$	0.174485100	0.115344	0.0794548	1.24×10^{-15}	0.2519691
$\hat{\lambda}_2$	0.851076600	0.764185	0.3867902	0.1647585	1.7630368
$\hat{\lambda}_3$	2.463019200	1.809700	0.7982738	0.6349577	3.5338221
$\hat{\lambda}_4$	13.948613400	8.598575	5.281904	1.921417	16.299988
$\hat{\lambda}_5$	87.000000000	60.49747	41.77191	12.95755	123.000000

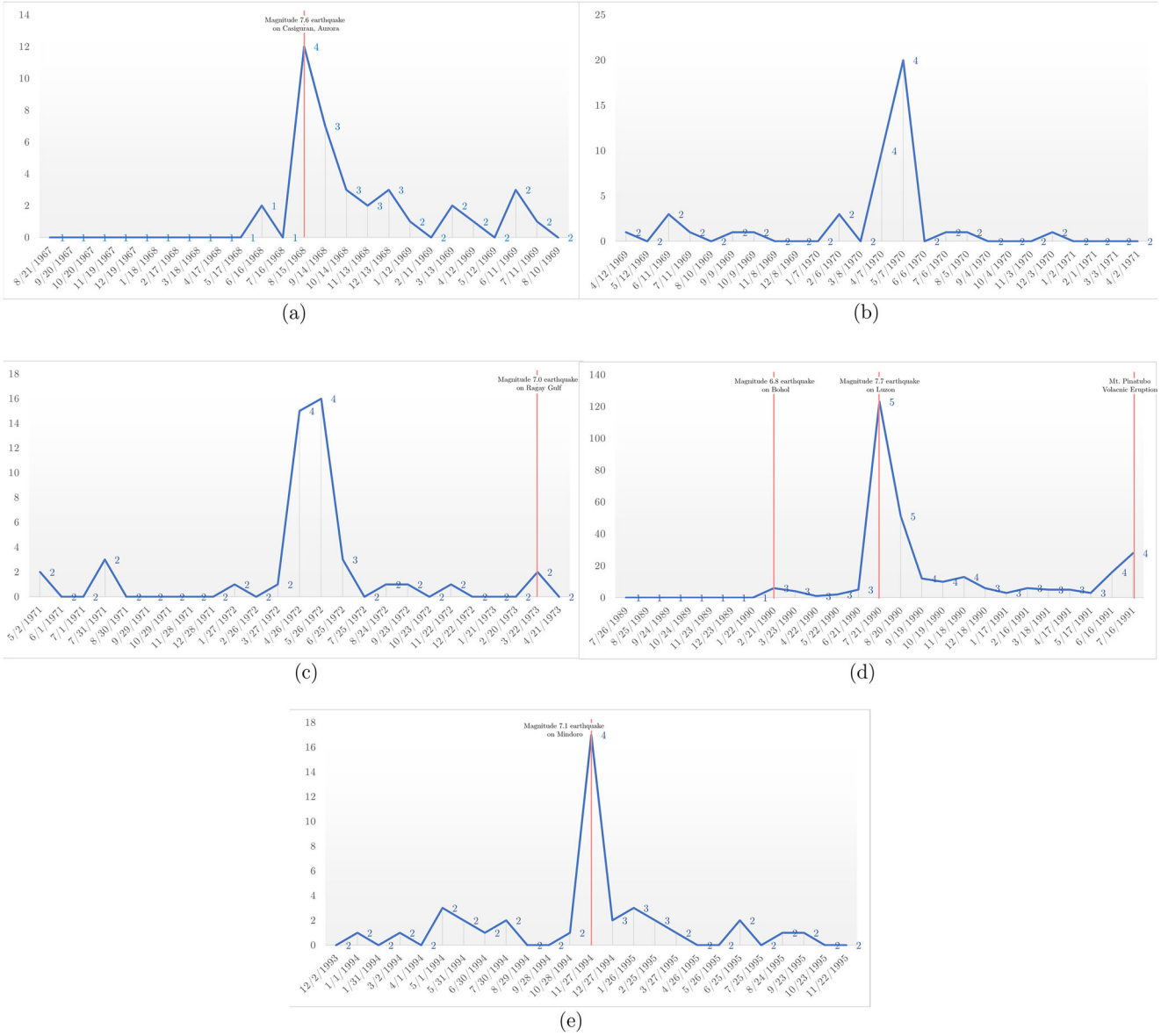

Fig. 4 State distribution under 5-state PHMM of some portions of the Metro Manila data

forecast. We do 18 one-step-ahead forecasts, with each training data set increasing by 1 as we include the real data from a previously forecasted time interval. We then compare these forecasts with the values of the one-step-ahead forecasts from the PHMM and the actual January 2019- June 2020 earthquake data as shown in Figure 5. The order of the best ARIMA model that fits the data is (2, 3, 1).

Notice that both the 5-state PHMM and ARIMA(2, 3, 1) forecast values are relatively close to each other, with the 5-state PHMM having a better estimate in most forecasts in terms of closeness to the actual value. There are three peaks in the actual data (9/17/2019, 1/15/2020, and 5/14/2020) that both models underestimate and fail to capture. In the first two peaks, the ARIMA(2, 3, 1) forecasts has a slight edge over those of the 5-state PHMM, while for the last peak (and most of the other data points), the forecasts for the 5-state PHMM are closer to the actual data than those of ARIMA(2, 3, 1).

We also perform an analysis of the deviations of the forecast values from both models to the actual using the unscaled mean bounded relative absolute error (UMBRAE) developed in (Chen et al. 2017). We find that the 5-state PHMM performs roughly 15.35% better than the ARIMA(2, 3, 1) model.

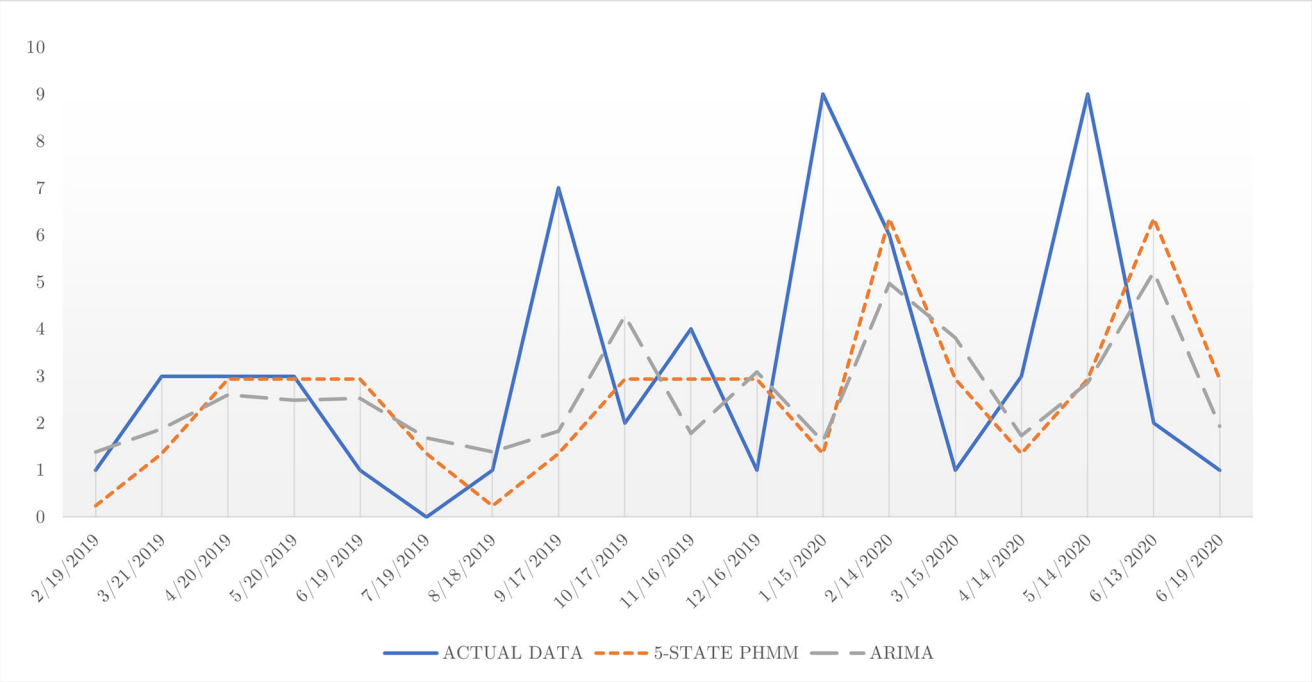

Fig. 5 Plot of the 18 one-step-ahead forecasts of magnitude greater than or equal to 4 earthquakes in Metro Manila from January 2019 - June 2020 of the 5-state PHMM model and the ARIMA(2, 3, 1) model vs the actual data

Conclusion and recommendation

In this study, we modeled the earthquake activity in Metro Manila using Poisson hidden Markov model. We considered the 30-day earthquake occurrences (with magnitude 4 or greater) data of Metro Manila from January 1960 to January 2019. We identified the 5-state Poisson hidden Markov model, with parameters $\lambda_1 = 0.174485100$, $\lambda_2 = 0.851076600$, $\lambda_3 = 2.463019200$, $\lambda_4 = 13.948613400$, and $\lambda_5 = 87$, as the best fit for the earthquake data. In addition, we investigated the forecasting capability of this model by comparing its 18 one-step-ahead forecasts to those of the ARIMA. Using various error metrics, the 5-state PHMM gave closer forecast values.

Our study has shown that the number of earthquake occurrences in Metro Manila can be modeled using PHMM. The model can help researchers to further understand the seismic behavior in the area. In particular, it can provide insights on how earthquakes of magnitude 4 or greater behave by observing the patterns of the states for which the 30-day intervals are in. We recommend the use of PHMM on the earthquake occurrences in other areas of the country. Also, the model can be used to develop an early warning signal by identifying trigger points that suggest an upcoming period with high number of occurrences or an occurrence of a major earthquake. It can be used to alert disaster risk management agencies so they can be prepared for a possible calamity. We also recommend the use of some modifications to PHMM such as through empirical recurrence rates relations similar to Bhaduri (2020). If possible, use an earthquake data with moment magnitude (M_w) since it is a more accurate measurement of the earthquake size and try to benchmark against other methods depending on the data set. This data, however, is not available in the earthquake catalogue that the DOST-PHIVOLCS can provide.

Appendix: State distributions

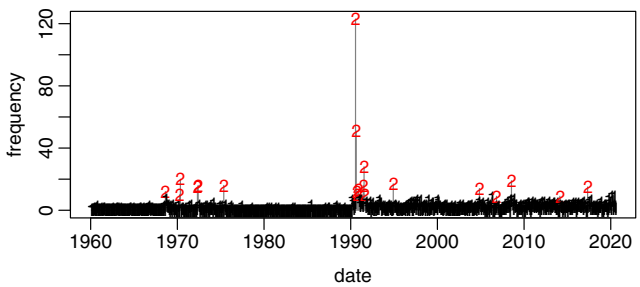

Fig. 6 State distributions under 2-state PHMM

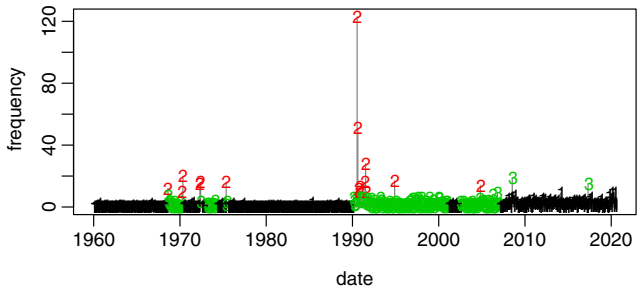

Fig. 7 State distributions under 3-state PHMM

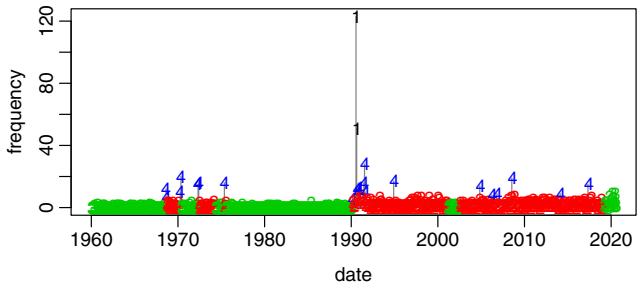

Fig. 8 State distributions under 4-state PHMM

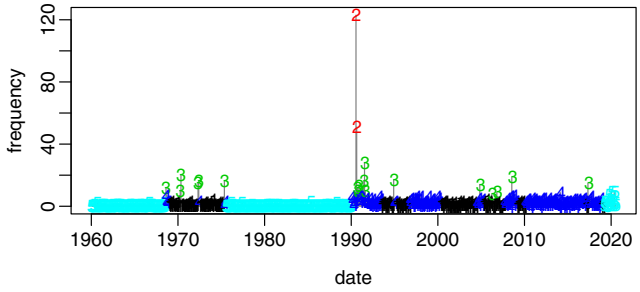

Fig. 9 State distributions under 5-state PHMM

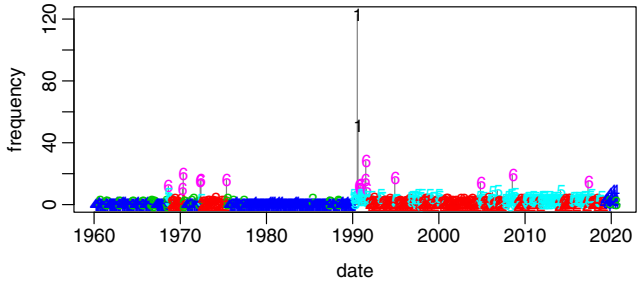

Fig. 10 State distributions under 6-state PHMM

Supplementary Files The R Code for PHMM model and ARIMA model, as well as the data, can be found online at: Codes and DOST-PHIVOLCS Data.

Acknowledgments We extend our gratitude to the Department of Science and Technology - Philippine Institute of Volcanology and Seismology (DOST-PHIVOLCS) for sharing the necessary earthquake data for the completion of the study.

Declarations

Conflict of interests The authors declare that they have no conflict of interest.

References

Akaike H (1974) Information theory and an extension of the maximum likelihood principle. In: Petrov B (ed) 2nd International Symposium on Information Theory, pp 267–281. Akademiai

Amei A, Fu W, Ho C (2012) Time series analysis for predicting the occurrences of large scale earthquakes. International Journal of Applied Science and Technology 2(7):64–75

Bansal A, Dimri V, Babu K (2012) Epidemic type aftershock sequence (etas) modeling of northeastern himalayan seismicity. J Seismol 17:255–264. https://doi.org/10.1007/s10950-012-9314-7

Baum L, Petrie T (1966) Statistical inference for probabilistic functions of finite state markov chains. The Annals of Mathematical Statistics 37(6):1554–1563

Baum L, Petrie T, Soules G, Weiss N (1970) A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. The Annals of Mathematical Statistics 41(1):164–171

Bautista B, Bautista M, Stein R, Barcelona E, Punongbayan R, Laguerta E, Rasdas A, Ambubuyog G, Amin E (1996) Relationship of regional and local structures to mount pinatubo activity. Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines. https://pubs.usgs.gov/pinatubo/bbautist/

Bhaduri M (2020) On modifications to the poisson-triggered hidden markov paradigm through partitioned empirical recurrence rates ratios and its applications to natural hazards monitoring. Scientific Reports 10:15,889–15,907

Bhaduri M, Ho C (2019) On a temporal investigation of hurricane strength and frequency. Environ Model Assess 24:495–507

Can C, Ergun G, Gokceoglu C (2014) Prediction of earthquake hazard by hidden markov model (around bilecik, nw turkey). Central European Journal of Geosciences 6:403–414

Chen C, Twycross J, Garibaldi J (2017) A new accuracy measure based on bounded relative error for time series forecasting. PLOS ONE 12:1–23

Dionysiou D, Papadopoulos G (1992) Poissonian and negative binomial modelling of earthquake time series in the aegean area. Physics of The Earth and Planetary Interiors - PHYS EARTH PLANET INTERIORS 71:154–165

Doganer A, Calik S (2013) Estimates of earthquake with markov models in the east anatolian fault zone. Turkish Journal of Science and Technology 8(1):55–61

Elliott R, Aggoun L, Moore J (1995) Hidden Markov models: Estimation and control Springer-Verlag

Fattah J, Ezzine L, Aman Z, Moussami H, Lachhab A (2018) Forecasting of demand using arima model. International Journal of Engineering Business Management 10:1–9. 10.1177/1847979018808673

Fukao Y, Furumoto M (1975) Foreshocks and multiple shocks of large earthquakes. Physics of The Earth and Planetary Interiors - PHYS EARTH PLANET INTERIORS 10:355–368

Goh C, Dauwels J, Mitrovic N, Asif M, Oran A, Jaillet P (2012) Online map-matching based on hidden markov model for real-time traffic sensing applications. pp 776–781

Ho C, Bhaduri M (2015) On a novel approach to forecast sparse rare events: Applications to parkfield earthquake prediction. Nat Hazards 78(1):669–679

Ho C, Bhaduri M (2017) A quantitative insight into the dependence dynamics of the kilauea and mauna loa volcanoes, hawaii. Math Geosci 49:893–911

Hopkins D, Clark W, Matuschka T, Sinclair J (1991) The philippines earthquake of july 16, 1990. Bulletin of the New Zealand Society for Earthquake Engineering 24:3–95

Hyndman R, Khandakar Y (2008) Automatic time series forecasting: the forecast package for R. Journal of Statistical Software 26(3):1–22. https://www.jstatsoft.org/article/view/v027i03

Kannan S (2014) Improving innovative mathematical model for earthquake prediction. Eng Fail Anal 41:89–95

King G, Stein R, Lin J (1994) Static stress changes and the triggering of earthquakes. Bulletin - Seismological Society of America 84(3):935–953

Krogh A, Larsson B, Heijne G, Sonnhammer E, Bioinformatics S (2001) Predicting transmembrane protein topology with a hidden markov model: Application to complete genomes. J Mol Biol 305:567–580

Le Thanh N, Kaito K, Kobayashi K (2015) Infrastructure deterioration prediction with a poisson hidden markov model on time series data. Journal of Infrastructure Systems 21(3):04014,051

Lomnitz C (1974) Global tectonics and earthquake risk, Developments in Geotectonics, vol 46. Elsevier

Orfanogiannaki K, Karlis D (2018) Multivariate poisson hidden markov models with a case study of modelling seismicity. Australian and New Zealand Journal of Statistics 60:301–322

Orfanogiannaki K, Karlis D, Papadopoulos G (2011) Identifying seismicity levels via poisson hidden markov models. Pure Appl Geophys 167:65–77

Orfanogiannaki K, Karlis D, Papadopoulos G (2018) Identification of temporal patterns in the seismicity of sumatra using poisson hidden markov models. Bull Geol Soc Greece 40:1–7

Pachter L, Alexandersson M, Cawley S (2002) Applications of generalized pair hidden markov models to alignment and gene finding problems. Journal of Computational Biology: A Journal of Computational Molecular Cell Biology 9:389–399

Paroli R, Redaelli G, Spezia L (2002) Poisson hidden markov models for time series of overdispersed insurance counts. In: Proceedings of the XXXI International ASTIN Colloquium, XXXI International ASTIN Colloquium, pp 461–474

PHIVOLCS (2017) Primer on the april 2017 series of earthquakes in batangas province. https://www.phivolcs.dost.gov.ph/index.php/news/628-primer-on-the-april-2017-series-of-earthquakes-in-batangas-province

PHIVOLCS (2019) Primer on the 22 april 2019 magnitude 6.1 central luzon earthquake. https://www.phivolcs.dost.gov.ph/index.php/news/8233-primer-on-the-22-april-2019-magnitude-6-1-central-luzon-earthquake-23-april-2019

PHIVOLCS (2019) Primer on the 23 april 2019 magnitude 6.5 eastern samar earthquake. https://www.phivolcs.dost.gov.ph/index.php/news/8234-primer-on-the-23-april-2019-magnitude-6-5-eastern-samar-earthquake

Rabiner L (1989) A tutorial on hidden markov models and selected. Proc IEEE 77:257–286

Rao N, Kaila K (2010) Application of the negative binomial to earthquake occurrences in the alpide-himalayan belt. Geophys J R Astron Soc 85:283–290

Rappler.com (2013) Strong quake rocks visayas; 28 dead. https://www.rappler.com/nation/41372-central-bohol-earthquake

Rappler.com (2019) Magnitude 6.5 earthquake rocks parts of mindanao. https://www.rappler.com/nation/243810-strong-aftershock-minda nao-october-31-2019

Rappler.com (2019) Magnitude 6.9 earthquake strikes davao del sur. https://www.rappler.com/nation/247244-earthquake-davao-del-su r-updates-december-15-2019

Rossi L, Chakareski J, Frossard P, Colonnese S (2015) A poisson hidden markov model for multiview video traffic. IEEE/ACM Trans Networking 23:547–558

Ryden T (1996) An em algorithm for estimation in markov-modulated poisson processes. Computational Statistics and Data Analysis 21:431–447

Schwarz G (1978) Estimating the dimension of a model. Annals Statistics 6(2):461–464

Spassiani I, Sebastiani G (2016) Magnitude-dependent epidemic-type aftershock sequences model for earthquakes. Phys Rev E 93(042):134. https://doi.org/10.1103/PhysRevE.93.042134

Tobyás V, Mittag R (1991) Local magnitude, surface wave magnitude and seismic energy. Studia Geophysica Et Geodaetica - STUD GEOPHYS GEOD 35:354–357

Utsu T (1969) Aftershocks and earthquake statistics(1) : Some parameters which characterize an aftershock sequence and their interrelations. Journal of the Faculty of Science, Hokkaido University, Series 7, Geophysics 3:129–195

Xiong H, Mamon R (2017) Putting a price tag on temperature. CMS 15:259–296. https://doi.org/10.1007/s10287-017-0291-8

Yip T, Ng W, Yau C (2017) A hidden markov model for earthquake prediction. Stoch Env Res Risk A 32:44–53