

# Bayesian Multiple Changepoint Analysis of Hurricane Activity in the Eastern North Pacific: A Markov Chain Monte Carlo Approach

XIN ZHAO

*Department of Information and Computer Sciences, University of Hawaii at Manoa, Honolulu, Hawaii*

PAO-SHIN CHU

*Department of Meteorology, School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, Honolulu, Hawaii*

(Manuscript received 12 November 2004, in final form 27 July 2005)

## ABSTRACT

A Bayesian framework is developed to detect multiple abrupt shifts in a time series of the annual major hurricanes counts. The hurricane counts are modeled by a Poisson process where the Poisson intensity (i.e., hurricane rate) is codified by a gamma distribution. Here, a triple hypothesis space concerning the annual hurricane rate is considered: “a no change in the rate,” “a single change in the rate,” and “a double change in the rate.” A hierarchical Bayesian approach involving three layers—data, parameter, and hypothesis—is formulated to demonstrate the posterior probability of each possible hypothesis and its relevant model parameters through a Markov chain Monte Carlo (MCMC) method.

Based on sampling from an estimated informative prior for the Poisson rate parameters and the posterior distribution of hypotheses, two simulated examples are illustrated to show the effectiveness of the proposed method. Subsequently, the methodology is applied to the time series of major hurricane counts over the eastern North Pacific (ENP). Results indicate that the hurricane activity over ENP has very likely undergone a decadal variation with two changepoints occurring around 1982 and 1999 with three epochs characterized by the inactive 1972–81 epoch, the active 1982–98 epoch, and the inactive 1999–2003 epoch. The Bayesian method also provides a means for predicting decadal major hurricane variations. A lower number of major hurricanes are predicted for the next decade given the recent inactive period of hurricane activity.

## 1. Introduction

It is well known that climate states are often characterized by different regimes lasting for a decade or longer. Examples of this can be found in the Pacific decadal oscillation index series and the major Atlantic hurricane series during the last century (e.g., Mantua et al. 1997; Goldenberg et al. 2001). Because the change of a climate basic state affects many facets of society (e.g., agriculture, water resources, environment), it is desirable to know when the change is likely to have occurred and possibly develop a means for predicting its future variation once a shift can be identified from a sound statistical analysis. Furthermore, many diagnostic stud-

ies are rooted on the basis of comparing active and inactive (or positive and negative) phases of climate states when the knowledge of a regime shift in the system is given (e.g., Deser et al. 2004).

Based on a log-linear regression model, Chu (2002) identified the shifts in the annual rates of tropical cyclone (TC) frequency over the central North Pacific for the period 1966–2000. More recently, Chu and Zhao (2004) applied an hierarchical Bayesian changepoint analysis to this set of data with two hypothesis models. The first hypothesis assumes no changepoint while the second hypothesis considers a single shift in the TC rates. Under a single changepoint hypothesis, the posterior probability of a shift around 1982 is high. Because the TC activity over the central North Pacific is related, to a large extent, to variations in TCs over the eastern North Pacific (Clark and Chu 2002), a question arises as to whether the major hurricane series over the eastern North Pacific (ENP) has also undergone a similar decadal variation. The ENP hurricanes sometimes

---

*Corresponding author address:* Dr. Pao-Shin Chu, Department of Meteorology, School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, 2525 Correa Rd., Honolulu, HI 96822-2291.  
E-mail: chu@hawaii.edu

wreak havoc on Central America and/or the southwest United States, bringing heavy rainfall, strong winds, and coastal storm surges.

When dealing with the changepoint analysis in Bayesian context, it is often difficult to analytically evaluate complex integral quantities of posterior distributions. One efficient way to overcome difficult posterior quantities is through the use of the Markov chain Monte Carlo (MCMC) approach (e.g., Lavielle and Labarbarier 2001; Gelman et al. 2004). Applications of MCMC to climate research are emerging. For instance, Berliner et al. (2000) used a MCMC approach to update distribution parameters of their physically based statistical model for predicting the Pacific sea surface temperatures. Elsner et al. (2004) successfully applied a MCMC approach to detect shifts in the Atlantic major hurricane series.

Strictly speaking, the method employed in Chu and Zhao (2004) is only applicable to detecting a single change in the tropical cyclone time series. For detecting more than one shift, a more elaborate algorithm is needed. In fact, when analyzing a climate time series even as short as 30 years, it is not unusual to find more than one phase shift in the records. To address this issue, the major objective of this study is to develop a method for detecting and quantifying the finite number of shifts in a series of hurricane activity within the framework of hierarchical Bayesian models.

Specifically, the annual major hurricane counts over the ENP are modeled by a Poisson process where the Poisson intensity is codified by its conjugate, gamma distribution. As in Chu and Zhao (2004), a three-layer hierarchical structure involving data, parameter, and hypothesis space is espoused in this study. But, different from Chu and Zhao (2004), a new scheme with three possible hypotheses is entertained. These include “a no change in the Poisson intensity,” “a single change in the Poisson intensity,” and “a double change in the Poisson intensity” hypothesis. Also new to the current study is the use of the Gibbs sampler. After developing a practical approach to approximately sampling from an informative prior, we deliberately design an algorithm to allow for calculation of the posterior probability of each possible hypothesis and its relevant model parameters, including the Poisson intensity and the time of the abrupt shifts, through a MCMC approach.

Section 2 discusses the dataset, and section 3 outlines the basic model for describing hurricane activity. In section 4, we present a complete Bayesian analysis for detecting up to two abrupt shifts in the hurricane series through a MCMC approach. Two simulated examples and the analysis results for annual major hurricane counts in the eastern North Pacific are given in section

5. Section 6 provides the summary and conclusions, where we also discuss a general concept about how to detect any finite number of multiple shifts in a time series under a given model.

## 2. Data

The annual major hurricane counts over the ENP come from the National Hurricane Center in Miami, Florida. Major hurricanes refer to Category 3 or higher on the Saffir–Simpson destruction potential scale. This corresponds to the maximum sustained wind speeds exceeding  $50 \text{ m s}^{-1}$ . Whitney and Hobgood (1997) suggested that reliable hurricane statistics over the ENP began in the early 1970s, when the Dvorak scheme for the estimation of the intensity of tropical cyclones was operationally implemented. The year 1972 was chosen as the starting year by Collins and Mason (2000) for investigating interannual variation of local environmental conditions and TC activity over the ENP; thus, the data from 1972 to 2003 are used in this study.

## 3. Mathematical model of hurricane activity

A Poisson process is a proper probability model for describing independent rare event counts. Given the intensity parameter  $\lambda$  (i.e., annual rate of major hurricanes), the probability mass function (PMF) of  $h$  major hurricanes occurring in  $T$  years is (Epstein 1985)

$$P(h|\lambda, T) = \exp(-\lambda T) \frac{(\lambda T)^h}{h!}, \quad (1)$$

where  $h = 0, 1, 2, \dots$  and  $\lambda > 0, T > 0$ . The Poisson mean is simply the product of  $\lambda$  and  $T$ , so is its variance.

In many cases, hurricane time series cannot be simply described by a constant rate Poisson process (e.g., Elsner and Jagger 2004). Thus the Poisson intensity,  $\lambda$ , should not be treated as a determinant single-value parameter but as a random variable. This resulting hierarchical feature also fits well with the Bayesian inference. A functional choice of  $\lambda$  is a gamma distribution (Epstein 1985) as expressed in the following form:

$$f(\lambda|h', T') = \frac{T'^{h'} \lambda^{h'-1}}{\Gamma(h')} \exp(-\lambda T'),$$

$$\lambda > 0, \quad h' > 0, \quad T' > 0, \quad (2)$$

where the gamma function is defined as  $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ .

Given  $h$  TCs occurring in  $T$  years, if the prior density for  $\lambda$  is gamma distributed with parameters  $h'$  and  $T'$ , the posterior density for  $\lambda$  will also be gamma distrib-

uted with parameters  $h + h'$  and  $T + T'$ ; that is, the gamma density is the conjugate prior for  $\lambda$ . Referring to (2), the conditional expectation with respect to  $\lambda$  is  $E[\lambda|h', T'] = h'/T'$ . Obviously, when  $h'$  and  $T'$  approach infinity, the model will converge to a constant rate Poisson distribution. In the later part of this paper, we will discuss how to properly estimate the prior information  $h'$  and  $T'$  given the observation.

Under the statistical model introduced above, the marginal PMF of  $h$  TCs occurring in  $T$  years when the intensity is gamma distributed with prior parameters  $h'$  and  $T'$  is a negative binomial distribution (Epstein 1985; Elsner and Bossak 2001; Chu and Zhao 2004):

$$\begin{aligned} P(h|h', T', T) &= \int_0^\infty P(h|\lambda, T)f(\lambda|h', T') d\lambda \\ &= \frac{\Gamma(h + h')}{\Gamma(h')h!} \left(\frac{T'}{T + T'}\right)^{h'} \left(\frac{T}{T + T'}\right)^h \\ &= P_{nb}\left(h/h', \frac{T}{T + T'}\right), \end{aligned} \quad (3)$$

where  $h = 0, 1, \dots, h' > 0, T' > 0, T > 0$ , and  $P_{nb}(\cdot)$  stands for the negative binomial distribution.

#### 4. MCMC approach for detection of multiple changepoints in the hurricane series

##### a. MCMC approach to Bayesian analysis and Gibbs sampling

In general, let us assume  $\theta$  to be the set of model parameters and  $\mathbf{h}$  the data for the analysis. The fundamental Bayesian model can be described by the mathematical statement:

$$P(\mathbf{h}|\theta) = \frac{P(\mathbf{h}|\theta)P(\theta)}{\int P(\mathbf{h}|\theta)P(\theta) d\theta} \propto P(\mathbf{h}|\theta)P(\theta),$$

where “ $\propto$ ” means “proportional.” Here  $P(\mathbf{h}|\theta)$  is the conditional distribution of data  $\mathbf{h}$  given the model parameters  $\theta$  (or the likelihood with given model) and  $P(\theta)$  is a prior probability. The Bayes formula provides the posterior probability  $P(\theta|\mathbf{h})$ , the probability of  $\theta$  after the data  $\mathbf{h}$  are observed. It is also clear that data affects the posterior inference only through the likelihood function  $P(\mathbf{h}|\theta)$ . To make predictive inference, we rely on the posterior predictive distribution  $P(\hat{\mathbf{h}}|\mathbf{h}) = \int P(\hat{\mathbf{h}}|\theta)P(\theta|\mathbf{h}) d\theta$ , where  $\hat{\mathbf{h}}$  denotes the prediction (Gelman et al. 2004). Here  $P(\hat{\mathbf{h}}|\mathbf{h})$  is the posterior predictive distribution since it is conditional on the observed data  $\mathbf{h}$  and provides a prediction for the unknown observable  $\hat{\mathbf{h}}$ . This formula is at the heart of Bayesian analysis.

The MCMC approach is one of the efficient algorithms for Bayesian inference. The general Bayesian analysis method described above essentially involves calculating the posterior expectation

$$E[a|\mathbf{h}] = \int_{\theta} a(\theta)P(\theta|\mathbf{h}) d\theta,$$

where  $a(\theta)$  can be any function conditional on model  $\theta$ . This expectation, however, is very difficult to integrate in most practical models. Alternatively, a numerical way to calculate such an expectation is to use Monte Carlo integration by

$$E[a|\mathbf{h}] \approx \frac{1}{N} \sum_{i=1}^N a(\theta^{[i]}),$$

where  $\theta^{[1]}, \theta^{[2]}, \dots, \theta^{[N]}$  are independently sampled from  $P(\theta|\mathbf{h})$ . When  $N$  goes to infinity, this approximation will converge to its analytical integral under very general conditions.

This method is straightforward, but it is often infeasible to generate such an independent series  $\theta^{[1]}, \theta^{[2]}, \dots, \theta^{[N]}$  when  $P(\theta|\mathbf{h})$  is complicated. Nonetheless, in most applications, it may be possible to generate a series of dependent values by using a Markov chain (MC) that has  $P(\theta|\mathbf{h})$  as its stationary distribution. The MC is defined by giving an initial distribution for the first state of the chain  $\theta^{[1]}$  and a set of transition probabilities for a new state  $\theta^{[i+1]}$  given the current state  $\theta^{[i]}$ . Under very general conditions (the Markov chain is ergodic), the distribution for the state will converge to a unique stationary distribution. As long as this stationary distribution is  $P(\theta|\mathbf{h})$ , the Monte Carlo integration described above still gives an unbiased estimate for  $E[a|\mathbf{h}]$  (Ripley 1987).

One of the most widely used MCMC algorithms is known as Gibbs sampler. Suppose there are  $p$  parameters in the model, thus we define  $\theta = [\theta_1, \theta_2, \dots, \theta_p]'$  as a  $p$ -dimensional vector of parameters. Presumably, directly sampling from the posterior distribution  $P(\theta|\mathbf{h})$  is unlikely; we assume we can generate a value from the conditional distribution for any component of  $\theta$  given values for the rest of the other components of  $\theta$ . In detail, Gibbs sampling involves successive sampling from the complete conditional posterior densities  $P(\theta_k|\mathbf{h}, \theta_1, \dots, \theta_{k-1}, \theta_{k+1}, \dots, \theta_p)$ , where  $k$  is from 1 to  $p$ .

The Gibbs sampling algorithm is briefly described below.

- 1) Choose arbitrary starting values:  $\theta^{[0]} = [\theta_1^{[0]}, \theta_2^{[0]}, \dots, \theta_p^{[0]}]$ .

2) Start at  $j = 1$  and complete the single cycle by drawing values from the  $p$  distributions given by

$$\begin{aligned} \theta_1^{[j]} &\sim P(\theta_1|\mathbf{h}, \theta_2^{[j-1]}, \theta_3^{[j-1]}, \dots, \theta_{p-1}^{[j-1]}, \theta_p^{[j-1]}) \\ \theta_2^{[j]} &\sim P(\theta_2|\mathbf{h}, \theta_1^{[j]}, \theta_3^{[j-1]}, \dots, \theta_{p-1}^{[j-1]}, \theta_p^{[j-1]}) \\ \theta_3^{[j]} &\sim P(\theta_3|\mathbf{h}, \theta_1^{[j]}, \theta_2^{[j]}, \dots, \theta_{p-1}^{[j-1]}, \theta_p^{[j-1]}) \\ \theta_{p-1}^{[j]} &\sim P(\theta_{p-1}|\mathbf{h}, \theta_1^{[j]}, \theta_2^{[j]}, \dots, \theta_{p-2}^{[j]}, \theta_p^{[j-1]}) \\ \theta_p^{[j]} &\sim P(\theta_p|\mathbf{h}, \theta_1^{[j]}, \theta_2^{[j]}, \dots, \theta_{p-2}^{[j]}, \theta_{p-1}^{[j]}). \end{aligned}$$

3) Set  $j = j + 1$  and repeat the second step until convergence.

Once convergence is reached, the conditional distributions from simulations contain sufficient information to reach the true posterior distribution of interest. Thus, we can approximately calculate  $E[a|\mathbf{h}]$  by

$$\frac{1}{N} \sum_{i=1}^N a(\boldsymbol{\theta}^{[i]})$$

with sufficiently large  $N$ , where  $\boldsymbol{\theta}^{[i]}$  is the  $i$ th sample drawn from Gibbs sampling algorithm within each iteration after convergence.

*b. Hypothesis model*

In this study, it is assumed that the probability of more than two changepoints within the desired period is negligible. In principle, the method espoused here can be easily extended to more than two changepoints, although the complexity of the problem is correspondingly increased. We will discuss this in the summary section.

Consider the following three hypotheses: (i) "a no-changepoint" hypothesis  $H_0$ , (ii) "a single changepoint" hypothesis  $H_1$ , and (iii) "a double changepoint" hypothesis  $H_2$ . The following derivations are all based on the mathematical model described in section 3. Note that the annual major hurricane data,  $\mathbf{h} = [h_1, h_2, \dots, h_n]'$ , are assumed to be described as a series of independent random variables. The three hypotheses models are postulated below.

1) Hypothesis  $H_0$ : "A no change in the rate" of the hurricane series:

$$\begin{aligned} h_i &\sim \text{Poisson}(h_i|\lambda_0, T), \\ i &= 1, 2, \dots, n, \end{aligned}$$

where  $T$  is the unit observation time, which is a year in this study.

$$\lambda_0 \sim \text{gamma}(h'_0, T'_0),$$

where the prior knowledge of the parameters  $h'_0$  and  $T'_0$  is given.

2) Hypothesis  $H_1$ : "A single change in the rate" of the hurricane series:

$$h_i \sim \text{Poisson}(h_i|\lambda_{11}, T),$$

when  $i = 1, 2, \dots, \tau - 1$

$$h_i \sim \text{Poisson}(h_i|\lambda_{12}, T),$$

when  $i = \tau, \dots, n$

$$\tau = 2, 3, \dots, n,$$

where  $T$  is as defined in the hypothesis  $H_0$ , and

$$\lambda_{11} \sim \text{gamma}(h'_{11}, T'_{11})$$

$$\lambda_{12} \sim \text{gamma}(h'_{12}, T'_{12}),$$

where the prior knowledge of the parameters  $h'_{11}$ ,  $T'_{11}$ ,  $h'_{12}$ ,  $T'_{12}$  is given.

Note that there are two epochs in this model and  $\tau$  is defined as the first year of the second epoch, or the changepoint.

3) Hypothesis  $H_2$ : "A double change in the rate" of the hurricane series:

$$h_i \sim \text{Poisson}(h_i|\lambda_{21}, T),$$

when  $i = 1, 2, \dots, \tau_1 - 1$

$$h_i \sim \text{Poisson}(h_i|\lambda_{22}, T),$$

when  $i = \tau_1, \tau_1 + 1, \dots, \tau_2 - 1$

$$h_i \sim \text{Poisson}(h_i|\lambda_{23}, T),$$

when  $i = \tau_2, \tau_2 + 1, \dots, n$ .

Note that  $\tau_1\tau_2 = 2, 3, \dots, \tau_2 - 1$ ,  $\tau_2|\tau_1 = \tau_1 + 1, \dots, n$ ,  $T$  is as defined in the hypothesis  $H_0$ , and

$$\lambda_{21} \sim \text{gamma}(h'_{21}, T'_{21})$$

$$\lambda_{22} \sim \text{gamma}(h'_{22}, T'_{22})$$

$$\lambda_{23} \sim \text{gamma}(h'_{23}, T'_{23}),$$

where the prior knowledge of the parameters  $h'_{21}$ ,  $T'_{21}$ ,  $h'_{22}$ ,  $T'_{22}$ , and  $h'_{23}$ ,  $T'_{23}$  is given.

Note that there are three epochs in this model:  $\tau_1$  is defined as the first year of the second epoch, or the first changepoint, and  $\tau_2$  is defined as the first year of the third epoch, or the second changepoint.

*c. Bayesian inference under each hypothesis*

As described in the model, throughout this study, the unit period for the observation data  $\mathbf{h}$  is always 1.

### 1) BAYESIAN INFERENCE UNDER $H_0$ HYPOTHESIS

There is only one parameter  $\lambda_0$  under this hypothesis. Since gamma is the conjugate prior for Poisson, the conditional posterior density function for  $\lambda_0$  is straightforward:

$$\lambda_0 | \mathbf{h}, H_0 \sim \text{gamma} \left( h'_0 + \sum_{i=1}^n h_i, T'_0 + n \right).$$

### 2) BAYESIAN INFERENCE UNDER $H_1$ HYPOTHESIS

Under this hypothesis, there are three parameters,  $\lambda_{11}$ ,  $\lambda_{12}$ , and  $\tau$ . Following the conjugate prior property again and with a given  $\tau$ , the conditional posterior density function for  $\lambda_{11}$  and  $\lambda_{12}$  is

$$\lambda_{11} | \mathbf{h}, \tau, H_1 \sim \text{gamma} \left( h'_{11} + \sum_{i=1}^{\tau-1} h_i, T'_{11} + \tau - 1 \right)$$

$$\lambda_{12} | \mathbf{h}, \tau, H_1 \sim \text{gamma} \left( h'_{12} + \sum_{i=\tau}^n h_i, T'_{12} + n - \tau + 1 \right).$$

A detailed derivation for the conditional posterior density of  $\tau$  given the parameters  $\lambda_{11}$  and  $\lambda_{12}$  is described in appendix A (and can also be found in Elsner et al. 2004). With a noninformative prior assumption (uniform distribution) for changepoint  $\tau$ , the conditional posterior density of  $\tau$  is formulated as [Eq. (A3)]

$$P(\tau | \mathbf{h}, H_1, \lambda_{11}, \lambda_{12}) \propto e^{-(\tau-1)(\lambda_{11} - \lambda_{12})} \left( \frac{\lambda_{11}}{\lambda_{12}} \right)^{\sum_{i=1}^{\tau-1} h_i}$$

$$\tau = 2, 3, \dots, n.$$

Thus, we have completed the Markov chain for the  $H_1$  hypothesis.

### 3) BAYESIAN INFERENCE UNDER $H_2$ HYPOTHESIS

Under this hypothesis, there are five parameters,  $\lambda_{21}$ ,  $\lambda_{22}$ ,  $\lambda_{23}$ , and  $\tau_1, \tau_2$ . With similar conjugate prior logic, we have the conditional posterior density function for  $\lambda_{21}$ ,  $\lambda_{22}$ , and  $\lambda_{23}$  with the given changepoints  $\tau_1$  and  $\tau_2$ :

$$\lambda_{21} | \mathbf{h}, \tau_1, H_2 \sim \text{gamma} \left( h'_{21} + \sum_{i=1}^{\tau_1-1} h_i, T'_{21} + \tau_1 - 1 \right)$$

$$\lambda_{22} | \mathbf{h}, \tau_1, \tau_2, H_2 \sim \text{gamma} \left( h'_{22} + \sum_{i=\tau_1}^{\tau_2-1} h_i, T'_{22} + \tau_2 - \tau_1 \right)$$

$$\lambda_{23} | \mathbf{h}, \tau_2, H_2 \sim \text{gamma} \left( h'_{23} + \sum_{i=\tau_2}^n h_i, T'_{23} + n - \tau_2 + 1 \right).$$

In appendix B, we derive the conditional posterior density of  $\tau_1$  and  $\tau_2$  given the parameters  $\lambda_{21}$ ,  $\lambda_{22}$ , and  $\lambda_{23}$ . With a noninformative prior (uniform) assumption

for changepoint  $\tau_1 | \tau_2$  and  $\tau_2 | \tau_1$ , the conditional posterior densities of  $\tau_1$  and  $\tau_2$  are [Eqs. (B1) and (B2)]:

$$P(\tau_1 | \mathbf{h}, H_2, \lambda_{21}, \lambda_{22}, \lambda_{23}, \tau_2) \propto e^{-(\tau_1-1)(\lambda_{21} - \lambda_{22})} \left( \frac{\lambda_{21}}{\lambda_{22}} \right)^{\sum_{i=1}^{\tau_1-1} h_i},$$

$$\tau_1 = 2, 3, \dots, \tau_2 - 1$$

$$P(\tau_2 | \mathbf{h}, H_2, \lambda_{21}, \lambda_{22}, \lambda_{23}, \tau_1) \propto e^{-(\tau_2 - \tau_1)(\lambda_{22} - \lambda_{23})} \left( \frac{\lambda_{22}}{\lambda_{23}} \right)^{\sum_{i=\tau_1}^{\tau_2-1} h_i},$$

$$\tau_2 = \tau_1 + 1, \dots, n.$$

The Markov chain under  $H_2$  hypothesis is also completed.

With the formula introduced in this section, plus the prior knowledge, we can apply the Gibbs sampler described in section 4a to draw the samples from the posterior distribution of the model parameters under each respective hypothesis.

### d. Hypothesis analysis

Based on the Bayes formula, the posterior PMF for hypothesis is

$$P(H | \mathbf{h}) = \frac{P(\mathbf{h} | H) P(H)}{\sum_H P(\mathbf{h} | H) P(H)},$$

where  $P(H)$  can be any discrete probability distribution function. Generally a proper noninformative choice for hypothesis space is a uniform distribution; thus we can simplify this formula to  $P(H | \mathbf{h}) \propto P(\mathbf{h} | H)$ .

Let  $\boldsymbol{\theta}$  be the vector of model parameters. In this study,  $\boldsymbol{\theta} = [\boldsymbol{\theta}'_0, \boldsymbol{\theta}'_1, \boldsymbol{\theta}'_2]'$  = [ $\lambda_0, \lambda_{11}, \lambda_{12}, \lambda_{21}, \lambda_{22}, \lambda_{23}, \tau, \tau_1, \tau_2$ ]', where  $\boldsymbol{\theta}_0 = \lambda_0$  represents the parameter under the  $H_0$  hypothesis;  $\boldsymbol{\theta}_1 = [\lambda_{11}, \lambda_{12}, \tau]'$  represents the parameters under the  $H_1$  hypothesis; and  $\boldsymbol{\theta}_2 = [\lambda_{21}, \lambda_{22}, \lambda_{23}, \tau_1, \tau_2]'$  represents the parameters under the  $H_2$  hypothesis. Obviously,  $\boldsymbol{\theta}_0, \boldsymbol{\theta}_1$ , and  $\boldsymbol{\theta}_2$  are nonoverlapping.

In this section, we propose a simple approach to estimating the marginal likelihood under the hypothesis  $H$ ,  $P(\mathbf{h} | H)$ . This is motivated by the formula,

$$P(\mathbf{h} | H) = \int_{\boldsymbol{\theta}} P(\mathbf{h} | \boldsymbol{\theta}, H) P(\boldsymbol{\theta} | H) d\boldsymbol{\theta},$$

which could be approximated using the Monte Carlo integration method:

$$P(\mathbf{h} | H) = \frac{1}{N} \sum_{i=1}^N P(\mathbf{h} | \boldsymbol{\theta}, H)^{[i]},$$

$$N \rightarrow \infty, \quad (4)$$

where  $\boldsymbol{\theta}$  is drawn from the prior  $P(\boldsymbol{\theta} | H)$  and the superscript  $[i]$  denotes the  $i$ th sampling, recalling that we

already have the likelihood function for each hypothesis [Eq. (C1)] as given in appendix C.

Specifically, under the  $H_0$  hypothesis, we generate a value for  $\lambda_0$  from the prior distribution; under the  $H_1$  hypothesis, we generate values for parameters  $\lambda_{11}$ ,  $\lambda_{12}$ , and  $\tau$  from the prior distribution; and under the  $H_2$  hypothesis, we generate values for  $\lambda_{21}$ ,  $\lambda_{22}$ ,  $\lambda_{23}$ , and  $\tau_1$ ,  $\tau_2$  from the prior distribution. The parameters of the changepoint  $\tau$  (under the  $H_1$  hypothesis) or  $\tau_1|\tau_2$  and  $\tau_2|\tau_1$  (under the  $H_2$  hypothesis) are drawn from the uniform distribution, which does not need any prior knowledge. As for the rate parameters under each hypothesis,  $[\lambda_0, \lambda_{11}, \lambda_{12}, \lambda_{21}, \lambda_{22}, \lambda_{23}]$ , they are generated from the prior distributions as described in section 4b with given prior parameters, namely,  $h'_0$  and  $T'_0$  for  $H_0$ ;  $h'_{11}$ ,  $T'_{11}$ ,  $h'_{12}$ ,  $T'_{12}$  for  $H_1$ ; and  $h'_{21}$ ,  $T'_{21}$ ,  $h'_{22}$ ,  $T'_{22}$ ,  $h'_{23}$ ,  $T'_{23}$  for  $H_2$ .

Generally speaking, a flat, noninformative prior is preferred in the Bayesian inference. A proper noninformative prior for the rate parameters is a gamma( $c, d$ ) where  $c = d = 0.001$ . A noninformative prior may be suitable in the Gibbs sampler proposed in section 4c under each hypothesis since the likelihood is very narrow. However, a noninformative prior for the rate parameters is inappropriate for Eq. (4) because it only works well when priors are relatively close to the region of the highest likelihood. A noninformative prior such as gamma( $c, d$ ) is apparently too flat, thus Eq. (4) may never converge within a large number of iterations. In this regard, a proper informative prior is chosen. Accordingly, we will propose a practical approach to sampling prior from a distribution, which is conforming to the peak region of the likelihood. This approach is hereafter referred to as the informative prior estimation (IPE) method. We will first look on the simplest hypothesis,  $H_0$  and then extend it to the  $H_1$  and  $H_2$  hypotheses in a similar manner.

As discussed in section 3, for the hurricane series, if the two prior parameters are given, the marginal distribution for the observed data is a negative binomial. With this model, under the  $H_0$  hypothesis, and guided by Eq. (3), we have

$$P(h_i|h'_0, T'_0, T = 1, H_0) \sim P_{\text{nb}}\left(h_i|h'_0, \frac{T'_0}{1 + T'_0}\right),$$

$$i = 1, 2, \dots, n.$$

For the sake of convenience, we would choose a moment method to estimate the parameters  $h'_0$  and  $T'_0$ . Nonetheless, when  $T'_0$  is relatively large, the moment estimation could be significantly biased. In this case, the maximum likelihood estimation (MLE) method appears to be preferable. The formula for moment estimation is as below (Carlin and Louis 2000; Chu and Zhao 2004)

$$\hat{T}'_0|H_0 = \frac{q_0}{1 - q_0},$$

$$\hat{h}'_0|H_0 = m_{h_0}\hat{T}'_0, \tag{5}$$

where  $q_0 = m_{h_0}/s_{h_0}$  and  $m_{h_0} = (1/n)\sum_{i=1}^n h_i$ ,  $s_{h_0} = (1/n - 1)\sum_{i=1}^n (h_i - m_{h_0})^2$  are the sample mean and sample variance for the given data, respectively.

When  $q_0 \geq 1$ , the moment estimation breaks down. Thus, in this case, we may regard this series as a constant rate Poisson process and set  $\hat{T}'_0$  as a large enough value and let  $\hat{h}'_0 = m_{h_0} * \hat{T}'_0$ . To draw rate  $\lambda_0$  from its prior or posterior density, one just sets it equal to the sample mean  $m_{h_0}$ .

This estimation method could be extended to multiple changepoint hypothesis with given changepoints. The similar argument will apply to the  $q_{11}$ ,  $q_{12}$ ,  $q_{21}$ ,  $q_{22}$  or  $q_{23} \geq 1$  scenario as  $q_0 \geq 1$  under  $H_0$ . The moment estimation formulas for prior parameters under the  $H_1$  and  $H_2$  hypotheses are given by Eqs. (6) and (7), respectively:

$$\hat{T}'_{11}|H_1, \tau = \frac{q_{11}}{1 - q_{11}},$$

$$\hat{h}'_{11}|H_1, \tau = m_{h_{11}}\hat{T}'_{11}$$

$$\hat{T}'_{12}|H_1, \tau = \frac{q_{12}}{1 - q_{12}},$$

$$\hat{h}'_{12}|H_1, \tau = m_{h_{12}}\hat{T}'_{12}, \tag{6}$$

where

$$q_{11} = m_{h_{11}}/s_{h_{11}},$$

$$q_{12} = m_{h_{12}}/s_{h_{12}},$$

and

$$m_{h_{11}} = \frac{1}{\tau - 1} \sum_{i=1}^{\tau-1} h_i,$$

$$s_{h_{11}} = \frac{1}{\tau - 2} \sum_{i=1}^{\tau-1} (h_i - m_{h_{11}})^2,$$

$$m_{h_{12}} = \frac{1}{n - \tau + 1} \sum_{i=\tau}^n h_i,$$

$$s_{h_{12}} = \frac{1}{n - \tau} \sum_{i=\tau}^n (h_i - m_{h_{12}})^2,$$

$$\hat{T}'_{21}|H_2, \tau_1, \tau_2 = \frac{q_{21}}{1 - q_{21}},$$

$$\hat{h}'_{21}|H_2, \tau_1, \tau_2 = m_{h_{21}}\hat{T}'_{21}$$

$$\hat{T}'_{22}|H_2, \tau_1, \tau_2 = \frac{q_{22}}{1 - q_{22}},$$

$$\begin{aligned}
\hat{h}'_{22}|H_2, \tau_1, \tau_2 &= m_{h22}\hat{T}'_{22} \\
\hat{T}'_{23}|H_2, \tau_1, \tau_2 &= \frac{q_{23}}{1 - q_{23}}, \\
\hat{h}'_{23}|H_2, \tau_1, \tau_2 &= m_{h23}\hat{T}'_{23},
\end{aligned} \tag{7}$$

where

$$\begin{aligned}
q_{21} &= m_{h21}/S_{h21}, \\
q_{22} &= m_{h22}/S_{h22}, \\
q_{23} &= m_{h23}/S_{h23},
\end{aligned}$$

and

$$\begin{aligned}
m_{h21} &= \frac{1}{\tau_1 - 1} \sum_{i=1}^{\tau_1-1} h_i, \\
S_{h21} &= \frac{1}{\tau_1 - 2} \sum_{i=1}^{\tau_1-1} (h_i - m_{h21})^2, \\
m_{h22} &= \frac{1}{\tau_2 - \tau_1} \sum_{i=\tau_1}^{\tau_2-1} h_i, \\
S_{h22} &= \frac{1}{\tau_2 - \tau_1 - 1} \sum_{i=\tau_1}^{\tau_2-1} (h_i - m_{h22})^2, \\
m_{h23} &= \frac{1}{n - \tau_2 + 1} \sum_{i=\tau_2}^n h_i, \quad \text{and} \\
S_{h23} &= \frac{1}{n - \tau_2} \sum_{i=\tau_2}^n (h_i - m_{h23})^2.
\end{aligned}$$

Apparently, we could insert Eqs. (5), (6), and (7) into the Markov chain derived in section 4c to complete the Gibbs sampler within each iteration. In real applications, we would choose a noninformative prior during the burn-in phase because before convergence, the estimation of prior parameters for rate based on (5), (6), and (7) could be biased. After reaching convergence, we then insert the estimation of the prior parameters in each iteration of the Gibbs sampler. That is, after the burn-in period, in each iteration of Gibbs sampling, calculate the estimation of prior parameters with given changepoints, then generate the Poisson rates with their relative estimated priors, and then calculate the likelihood for each hypothesis. With guidance of Eqs. (4) and (C1), under the uniform prior assumption in the hypothesis space and after normalization, we obtain  $P(H|\mathbf{h})$ . In the simulated examples of this study, we find this estimation method converges very fast. Since there are two parameters to estimate for each epoch, the minimum sample size is two. This constraint is imposed in a real Gibbs sampler design.

Although the method described here does work well if the model can adequately represent the data, it still has a limitation. For example, if the ratio between the

mean and variance of the data is larger than one in an epoch with given changepoint(s), the prior has to be set as a gamma distribution with an infinite rate. As a result, one is forced to conclude that the data do not update prior beliefs. This is against the fundamental rule of Bayesian analysis. To avoid this problem, one may use some other estimation approaches or build some other models (e.g., Elsner et al. 2004).

One of the simplest marginal likelihood estimation methods is called ‘‘harmonic mean estimator’’ (Congdon 2003). Inspired by the identity,

$$\int_{\boldsymbol{\theta}} \frac{P(\boldsymbol{\theta}|\mathbf{h}, H)}{P(\mathbf{h}|\boldsymbol{\theta}, H)} d\boldsymbol{\theta} = \frac{1}{P(\mathbf{h}|H)},$$

an alternative approach to estimating  $P(\mathbf{h}|H)$  by using Monte Carlo integration is formulated by

$$P(\mathbf{h}|H) = 1 / \left( \frac{1}{N} \sum_{i=1}^N \frac{1}{P(\mathbf{h}|\boldsymbol{\theta}, H)^{[i]}} \right), \tag{8}$$

$N \rightarrow \infty,$

where  $\boldsymbol{\theta}$  is drawn from its posterior probability  $P(\boldsymbol{\theta}|\mathbf{h}, H)$  and the superscript  $[i]$  denotes the  $i$ th sampling.

Unlike the approach given by Eq. (4), Eq. (8) does not depend on sampling from the prior distribution. Instead, it only needs to draw samples from the posterior distribution, which in this study is equivalent to the output of the proposed Gibbs sampler after the burn-in period. Thus, a noninformative prior for the rate parameters is proper for Eq. (8). A harmonic mean estimator (HME) has been widely used, but it may be unstable if by chance a few low likelihood values are present in the sampling output. In the simulated examples of this study, this approach converges much slower than the proposed IPE approach. However, since the MHE approach is only based on a noninformative prior and needs outputs from a Gibbs sampler for estimation, it may serve as a useful comparison with the IPE approach. A simple extension to this method is given by Gelfand and Dey (1994).

#### e. Predictive distribution

After running the Gibbs sampling algorithm with the Markov chain described in sections 4c and 4d, the stationary output will be  $P(\boldsymbol{\theta}, H|\mathbf{h})$ . Therefore the predictive distribution for future  $\hat{T}$  years will be

$$P(\hat{h}|\mathbf{h}, \hat{T}) = \frac{1}{N} \sum_{i=1}^N P(\hat{h}|\hat{T}, \boldsymbol{\theta}, H)^{[i]}, \tag{9}$$

$N \rightarrow \infty,$

where  $\boldsymbol{\theta}$  and  $H$  are drawn from  $P(\boldsymbol{\theta}, H|\mathbf{h})$  and superscript  $[i]$  denotes the  $i$ th independent sampling. Alter-

nately, since both hypothesis choice and the position of changepoints are discrete, another simple method for prediction is to find the optimum [maximum a posterior (MAP)] estimation of changepoints and hypothesis and apply them to the likelihood model. This yields

$$P(\hat{h}|\mathbf{h}, \hat{T}) = P_{\text{nb}}\left(\hat{h} \mid \sum_{i=1}^n h_i, \frac{n}{\hat{T} + n}\right),$$

when  $\hat{H} = H_0$

$$P(\hat{h}|\mathbf{h}, \hat{T}, \hat{\tau}) = P_{\text{nb}}\left(\hat{h} \mid \sum_{i=\hat{\tau}}^n h_i, \frac{n - \hat{\tau} + 1}{\hat{T} + (n - \hat{\tau} + 1)}\right),$$

when  $\hat{H} = H_1$

$$P(\hat{h}|\mathbf{h}, \hat{T}, \hat{\tau}_1, \hat{\tau}_2) = P_{\text{nb}}\left(\hat{h} \mid \sum_{i=\hat{\tau}_2}^n h_i, \frac{n - \hat{\tau}_2 + 1}{\hat{T} + (n - \hat{\tau}_2 + 1)}\right),$$

when  $\hat{H} = H_2$ . (10)

### 5. Simulation and results

For each following example, the length of the burn-in period is 500 and the number of iterations of the Gibbs sampler after the burn-in period is chosen as 10 000 for the MCMC–IPE approach and 50 000 for the HME approach. In all simulations, the noninformative prior for the rate parameters is a gamma(0.001, 0.001).

#### a. Simulated examples

To test the validity of the MCMC–IPE method, two simulated examples are shown subsequently. We apply the model building approach outlined in section 4b to the following two simulated time series; the first one is characterized by one single changepoint and the second one by two changepoints.

##### 1) EXAMPLE ONE

In the first example, a Poisson time series with a sample size of 300 is generated. For the first 160 points, the rate for each point is generated from a gamma(2, 1) while for the remaining 140 points, the rate for each point is drawn from a gamma(4, 1). Figure 1a illustrates the simulated series. After running the proposed method described in the preceding section, the output for the posterior probability of three hypotheses is  $P(H_0|\mathbf{h}) = 0$ ,  $P(H_1|\mathbf{h}) = 0.981$ , and  $P(H_2|\mathbf{h}) = 0.019$ , suggesting that the single changepoint hypothesis is prevailing. The posterior PMF for changepoint  $\tau$  is illustrated as a solid line in Fig. 1b, from which we see that the model can precisely catch the prescribed changepoint, which is shown as a broken line. We also plotted the estimated prior PDF and the true model

prior PDF of the Poisson rate for the epoch before and after the changepoint (namely  $\lambda_1$  and  $\lambda_2$ ) in Figs. 1c and 1d, respectively. The estimated prior is very close to the true model under  $H_1$  hypothesis. We also apply the harmonic mean estimator approach to this series and obtain almost exactly same PMF plot for the changepoint. The estimation of the posterior probability for hypothesis  $H_0, H_1$ , and  $H_2$  is 0, 0.645, and 0.355, respectively.

##### 2) EXAMPLE TWO

In the second example, the same model as in the first one is used to generate a different Poisson time series. This time, the time series has two changepoints. The length of this time series is also set to be 300. For the first 50 points, the rate for each point is generated from a gamma(4, 2); for the middle 170 points, the rate for each point is drawn from a gamma(5, 1) and for the final 80 points, the rate for each point is drawn from a gamma(9, 3). The simulated series are plotted in Fig. 2a and after running the MCMC–IPE approach, the output for the posterior probability of three hypotheses is  $P(H_0|\mathbf{h}) = 0$ ,  $P(H_1|\mathbf{h}) = 0.011$ , and  $P(H_2|\mathbf{h}) = 0.989$ ; this result indicates that the double changepoint hypothesis is dominant, a feature consistent with our preset model. The posterior PMF for these two changepoints, namely  $\tau_1$  and  $\tau_2$ , are drawn in Figs. 2b and 2c, respectively; again, the method demonstrated can catch both changepoints precisely under the  $H_2$  hypothesis. Figures 2d,e,f plot the estimated prior for the rate of each epoch, namely,  $\lambda_1, \lambda_2$ , and  $\lambda_3$ , superimposed on their relative model prior PDFs. It is clear that the estimated PDF approximates the true PDF for each epoch without much bias under the  $H_2$  hypothesis. For the HME approach, we obtain almost exactly the same PMF plots for the changepoints and the estimation of the posterior probability for hypotheses  $H_0, H_1$ , and  $H_2$  is 0, 0.003, and 0.997, respectively.

#### b. Changepoint analysis of the hurricane rates in the ENP basin

Now let us turn our attention to modeling the actual records. We apply the proposed method to the series of annual major hurricane (MH) counts in ENP. Figure 3a shows the time series of annual MH counts over the ENP from 1972 to 2003. Under the uniform prior assumption in the hypothesis space, with the MCMC–IPE approach, the calculated posterior probability for each hypothesis is  $P(H_0|\mathbf{h}) = 0.021$ ,  $P(H_1|\mathbf{h}) = 0.195$ , and  $P(H_2|\mathbf{h}) = 0.784$ ; for the HME approach, the calculated posterior probability for each hypothesis is  $P(H_0|\mathbf{h}) =$

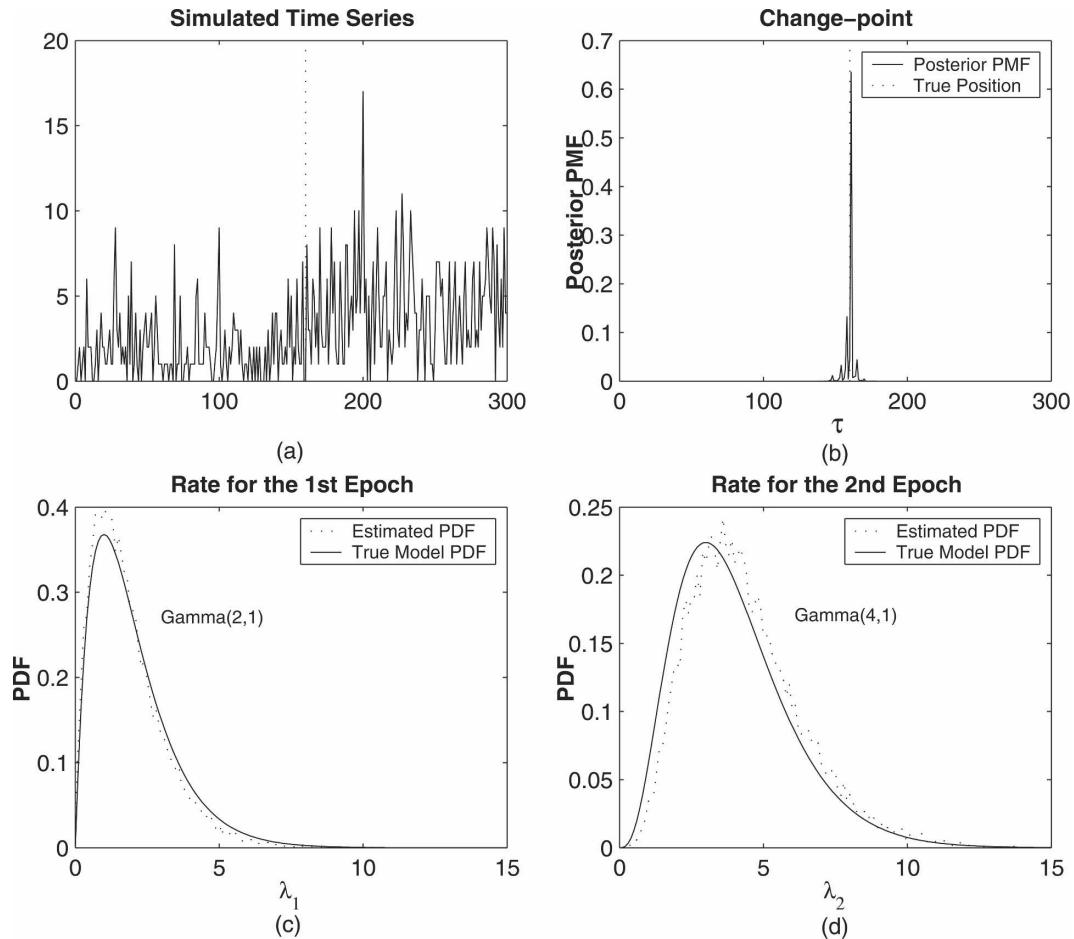


FIG. 1. Result for the first simulated example. A Poisson time series with sample size 300 is generated. For the first 160 points, the rate for each point is generated from a gamma(2, 1) while for the remaining 140 points, the rate for each point is drawn from a gamma(4, 1). (a) The simulated time series. (b) The posterior PMF for changepoint  $\tau$  (solid line). The broken line in (a) and (b) denotes the prescribed changepoint position. The estimated prior PDF (broken line) and the true model prior PDF (solid line) of the Poisson rate for the epoch (c) before and (d) after the changepoint (i.e.,  $\lambda_1$  and  $\lambda_2$ ).

0.089,  $P(H_1|\mathbf{h}) = 0.237$ , and  $P(H_2|\mathbf{h}) = 0.674$ . Regardless of which approach is taken, the  $H_2$  hypothesis is by far the most likely choice. Figure 4 displays the successive sample values for each parameter under the  $H_2$  hypothesis and their relative autocorrelation plots with lags up to 40. Under the  $H_2$  hypothesis, the posterior PMF for both changepoints are shown in Figs. 3b and 3c, through which one can see that the best choice for the first (second) changepoint appears to be 1982 (1999) in terms of MAP estimation.

The posterior PDF for the rate of each epoch is plotted in Figs. 3d,e,f, respectively. The average rate prior to 1982 is about 2.64  $\text{MH yr}^{-1}$ , and increases to almost 4.41  $\text{MH yr}^{-1}$  from 1982 to 1998, and drops back to 1.4  $\text{MH yr}^{-1}$  thereafter (Table 1). Figures 5a,b show the posterior PDF for the rate shift from the first epoch to

the second epoch ( $\lambda_2 - \lambda_1$ ), and the rate shift from the second to the third epoch ( $\lambda_2 - \lambda_3$ ), respectively. The  $p$  values for the posterior PDF of both shifts are all very small (0.006 for the first and 0.004 for the second), which strongly implies the existence of two changepoints in the hurricane time series.

### c. Decadal tropical cyclone prediction

After having identified two changepoint years in the major hurricane series, we will use Eqs. (9) and (10) to predict major hurricane counts in the ENP for the next decade. Figures 6a and 6b display the predictive distributions of the annual counts in the next decade. A bimodal distribution of future annual major hurricane counts is noted if a weighted-average formula such as (9) is used (triangles in Fig. 6a). The two peaks and the

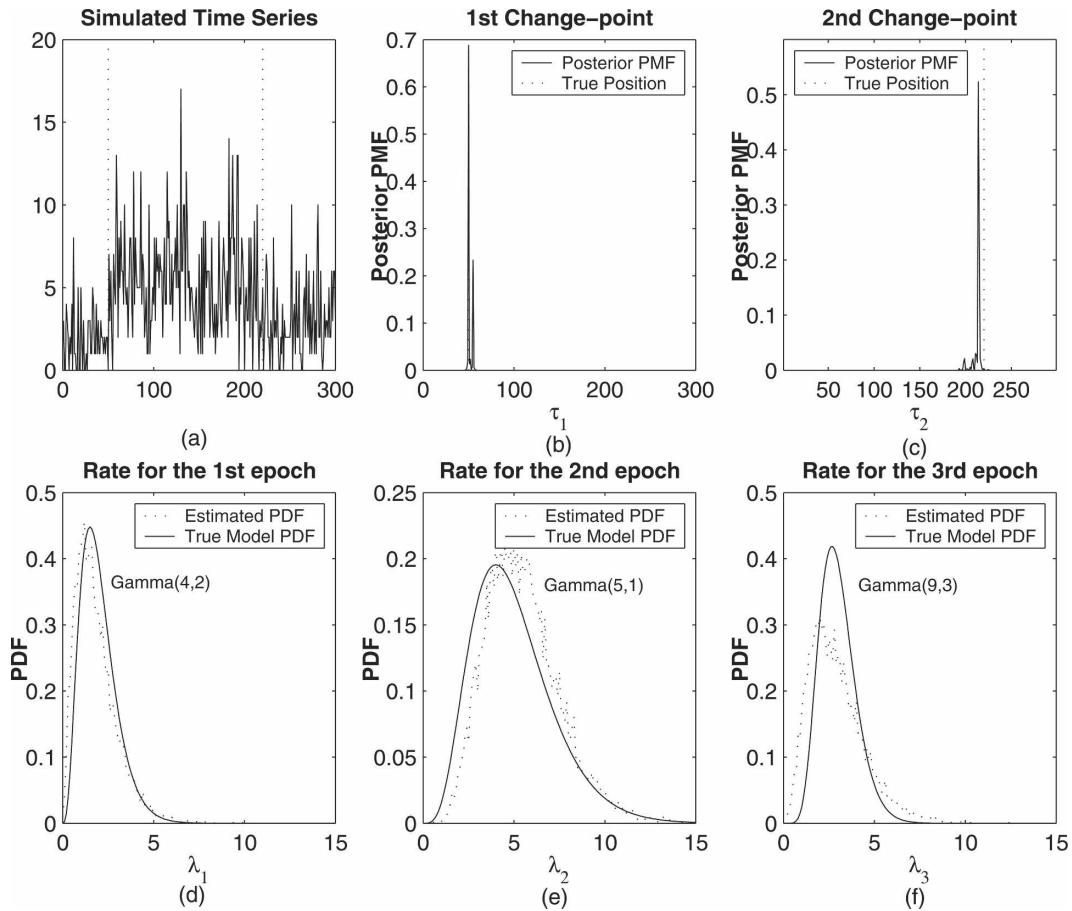


FIG. 2. Result for the second simulated example. A Poisson time series with sample size 300 is generated. The time series has two prescribed changepoints. For the first 50 points, the rate for each point is generated from a gamma(4, 2); for the middle 170 points, the rate for each point is drawn from a gamma(5, 1); and for the final 80 points, the rate for each point is drawn from a gamma(9, 3). (a) The simulated series are plotted. The posterior PMF for these two changepoints, namely (b)  $\tau_1$  and (c)  $\tau_2$  are shown. Broken lines in (a)–(c) denote the prescribed changepoint positions. The estimated prior (broken line) for the rate of each epoch, (d)  $\lambda_1$ , (e)  $\lambda_2$ , and (f)  $\lambda_3$ , superimposed on their theoretical prior (solid line) PDFs.

distribution are similar to that shown in Fig. 3f, where a bimodal distribution is also exhibited in the posterior distribution of the hurricane rate for the third epoch. Since there are two changepoints identified and the rate is sharply decreased after the second changepoint (Table 1), a straight application of Eq. (10) results in a distribution that tends to yield only one peak near 12 major hurricanes (asterisks in Fig. 6a).

Also shown in Figs. 6a and 6b is the predictive distribution when no changepoint is presumed (open circles), a common practice in the traditional Bayesian analysis. This refers to  $H = H_0$  in Eq. (10). Clearly, a shift toward a higher number of major hurricanes is evident when one compares the predictive distribution of the two changepoints to that of a no changepoint scenario (asterisks to open circles). Thus, without taking into account the different hurricane rates through-

out the time, it would be naive to project active hurricane activity in the next decade. The difference in the cumulative predictive distributions between a no changepoint and a two changepoint model is also very clear in Fig. 6b. The curve based on the weighted average (triangles) resembles a compromise between these two distributions (Fig. 6b).

## 6. Summary and discussion

Traditionally, major hurricane rates over an ocean basin have been modeled in a data parameter two-layer hierarchical Bayesian framework (Elsner and Bossak 2001). In this view, hurricane rates are assumed to be invariant throughout time, and the annual hurricane counts are described by a stationary Poisson process whose rate is conditional on a gamma distribution. To

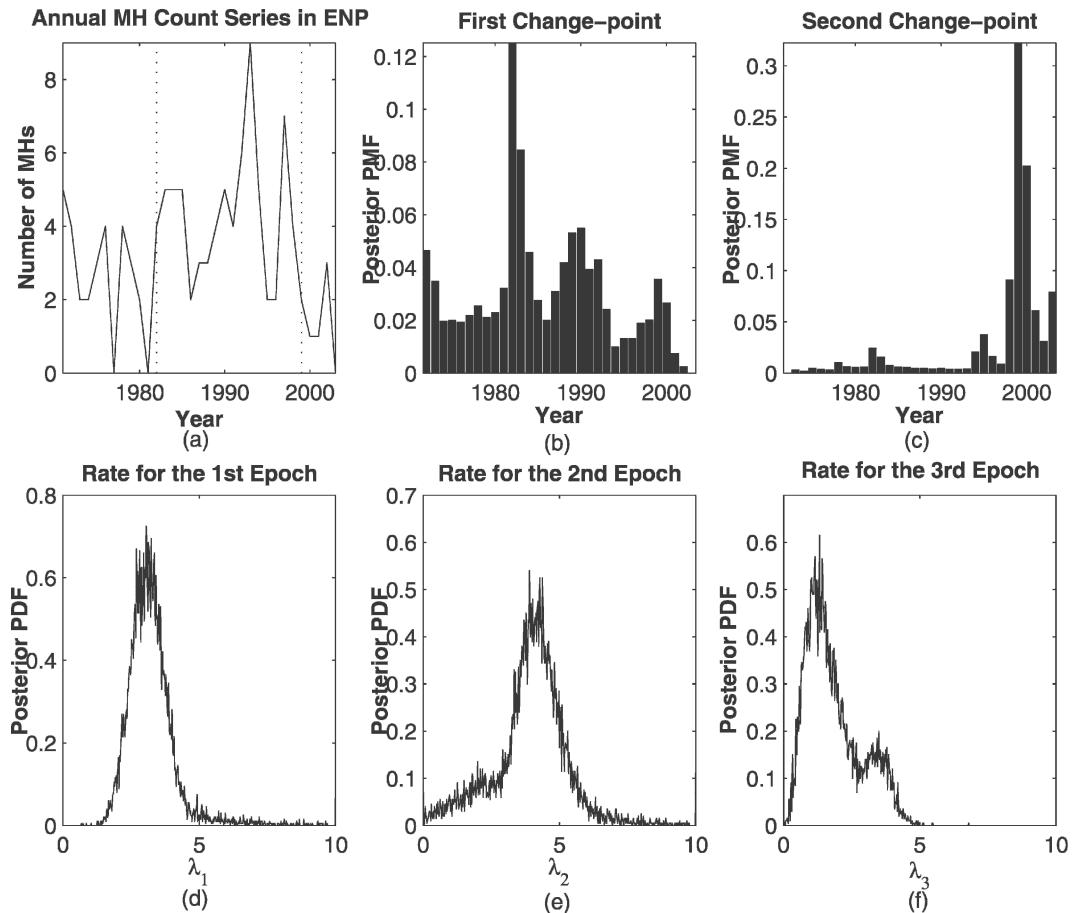


FIG. 3. Bayesian changepoint analysis result for the annual MH counts in ENP: (a) The time series of annual MH counts over the ENP from 1972 to 2003; (b), (c) the posterior PMF for both changepoints; and (d), (e), (f) the posterior PDF for the rate of each epoch.

detect potential abrupt shifts in tropical cyclone series, Chu and Zhao (2004) introduced a data parameter hypothesis, three-layer Bayesian paradigm so that the hypothesis layer allows the possibility of both a single change and no change of the annual hurricane rate. Elsner et al. (2004) recently modeled the hurricane count time series based on a nonconstant Poisson rate process with an hierarchically constructed prior. Building and extending from the three-layer, two-hypothesis paradigm for the central North Pacific tropical cyclone series, in this study the hypothesis space allows for three scenarios: “a no change in the rate,” “a single change in the rate,” and “a double change in the rate” for the annual major hurricane counts over the ENP. This extension is necessary because in the real world it is not unusual to have encountered more than one abrupt shift in a climate time series.

With a noninformative assumption on both hypothesis space and the position of changepoints, after constructing the complete Markov chains of a Gibbs sam-

pler under each hypothesis, we propose an approach to estimating the posterior distribution of the hypotheses and the parameters under each hypothesis. It is based on sampling from an estimated informative prior. As a test of the informative prior estimation (IPE) approach, two examples are simulated and their results are compared favorably to the harmonic mean estimator approach. The simulation results demonstrate that the IPE approach can catch the proper hypothesis and the associated changepoints quickly. In addition, the samples drawn from the estimated prior are very close to that of the true prior distribution of the underlying model.

The new Bayesian algorithm is subsequently applied to the annual major hurricane counts over the ENP during 1972–2003 and a double change hypothesis is most likely over the last 32 years with abrupt shifts being identified in 1982 and 1999. In terms of epoch definitions, the period 1972–81 is regarded as inactive while the epoch of 1982–98 is active. Beginning with

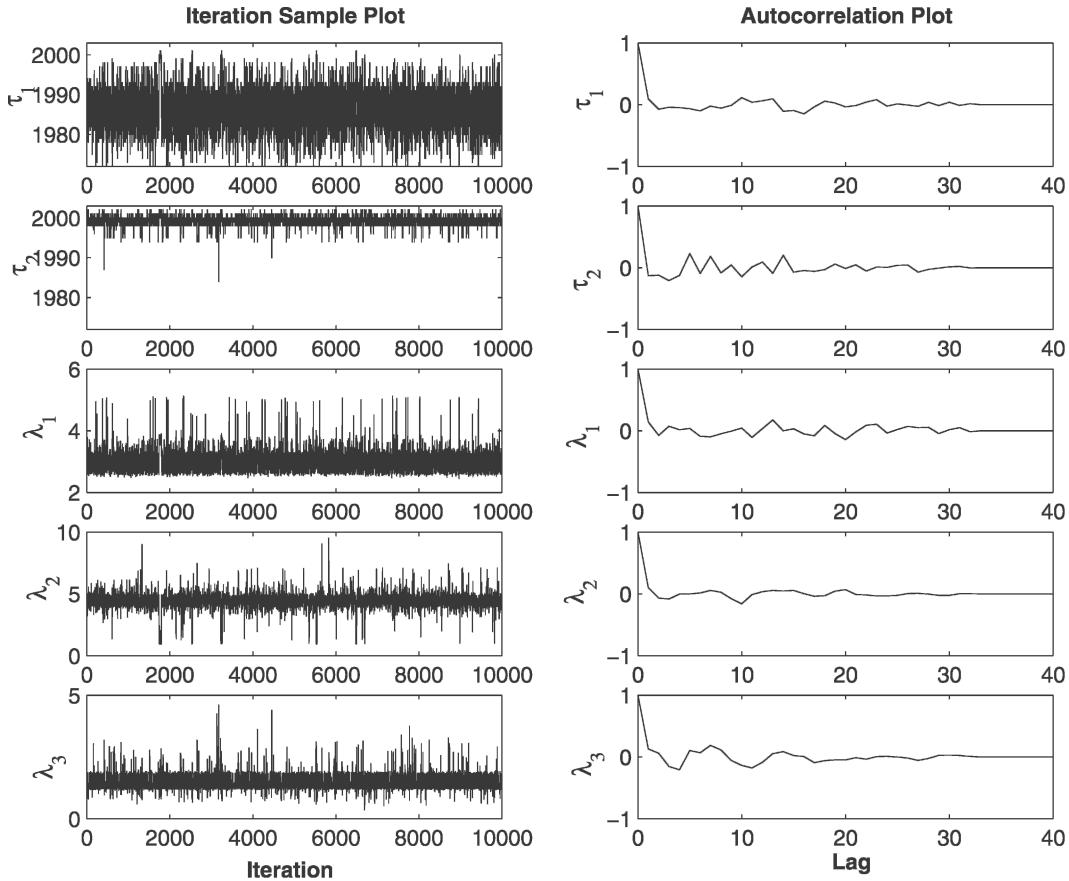


FIG. 4. Successive sample values for each parameter under (left) the  $H_2$  hypothesis and (right) their autocorrelation plots when running the MCMC/IPE approach to the annual MH count series in the eastern North Pacific.

1999, the ENP is dominated by a quiescent epoch. Interestingly, the Pacific decadal oscillation also seems to change its phase from a positive to a negative one in 1999 (more information available online at [http://](http://sealevel.jpl.nasa.gov/science/pdo.html)

TABLE 1. Bayesian analysis results on the changepoint of annual MH counts in the ENP. Here, under the uniform assumption in the hypothesis space,  $P(H_0|\mathbf{h})$ ,  $P(H_1|\mathbf{h})$ , and  $P(H_2|\mathbf{h})$  denote the posterior probability of the  $H_0$ ,  $H_1$ , and  $H_2$  hypotheses, respectively;  $\hat{\tau}_1$  and  $\hat{\tau}_2$  are the maximum likelihood estimates for the first and second changepoints under the  $H_2$  hypothesis, respectively;  $\bar{\lambda}_1|\hat{\tau}_1, \hat{\tau}_2$ ,  $\bar{\lambda}_2|\hat{\tau}_1, \hat{\tau}_2$ , and  $\bar{\lambda}_3|\hat{\tau}_1, \hat{\tau}_2$  denote the average rate, under the  $H_2$  hypothesis, in three consecutive epochs (i.e., 1972–81, 1982–98, and 1999–2003), respectively, given changepoints in 1982 and 1999.

Term	Value
$P(H_0 \mathbf{h})$	0.021
$P(H_1 \mathbf{h})$	0.195
$P(H_2 \mathbf{h})$	0.784
$\hat{\tau}_1$	1982
$\hat{\tau}_2$	1999
$\bar{\lambda}_1 \hat{\tau}_1, \hat{\tau}_2$	2.64
$\bar{\lambda}_2 \hat{\tau}_1, \hat{\tau}_2$	4.41
$\bar{\lambda}_3 \hat{\tau}_1, \hat{\tau}_2$	1.40

sealevel.jpl.nasa.gov/science/pdo.html). Whether the change in hurricane series is related to the PDO phase shift is beyond the scope of the current study and merits further investigation. As expected, the predicted major hurricane count in ENP for the next decade based on this three-layer Bayesian analysis, given the possibility of the changepoints, is much lower than that based on the traditional two-layer analysis.

For the sake of simplicity, we neglect the possibility of more than two changepoints in this study. However, it would also be possible to carry out an analysis with more than two changepoints. The general idea is outlined in the following. For example, if we assume that there are possibly at most up to  $K - 1$  changepoints in the time series, we then have  $K$  possible hypotheses in hypothesis space. In detail, under the  $k$  changepoint hypothesis,  $0 \leq k \leq K - 1$ , we can introduce  $k$  changepoint position random variables, namely  $\tau_1, \tau_2, \dots, \tau_k$  and  $k + 1$  random Poisson rates, namely  $\lambda_1, \lambda_2, \dots, \lambda_{k+1}$ , for each relative epoch. Thus there will be  $2k + 1$  parameters under the  $k$  changepoint hypothesis. For any single target parameter under this hypothesis, if all

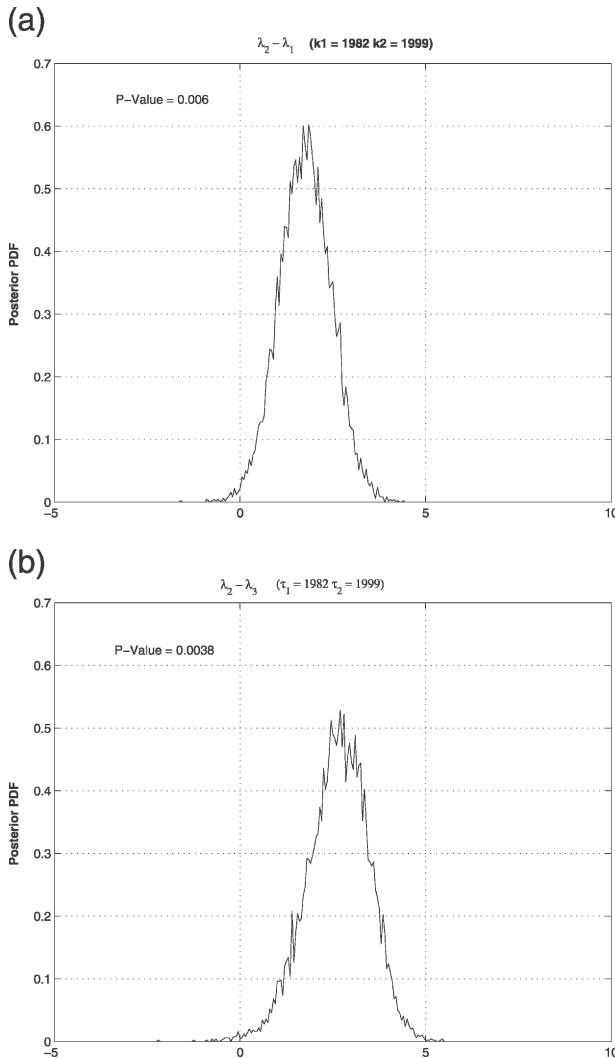


FIG. 5. Posterior density function of the shift for the major hurricane series over the ENP (a)  $P(\lambda_2 - \lambda_1 | \mathbf{h}, H_2, \tau_1 = 1982, \tau_2 = 1999)$  and (b)  $P(\lambda_2 - \lambda_3 | \mathbf{h}, H_2, \tau_1 = 1982, \tau_2 = 1999)$ .

other parameters are given, the derivation of its conditional posterior distribution is fairly similar to the extension from the  $H_1$  to the  $H_2$  hypothesis with very minor revisions (refer to the argument and derivation in appendix B). Furthermore, under the similar deduction in appendix C, it is also easy to formulate the likelihood  $P(\mathbf{h} | \boldsymbol{\theta}, H_k)$  under the given model. Once the Markov chain for Gibbs sampler is set, we can estimate both posterior and prior distributions in a way similar to that introduced in section 4d for each parameter under this hypothesis. In essence, it is quite straightforward to extend the whole framework described in section 4 to up to any finite number of multiple change-points in hypothesis space. However, the extended framework inevitably increases the computational complexity.

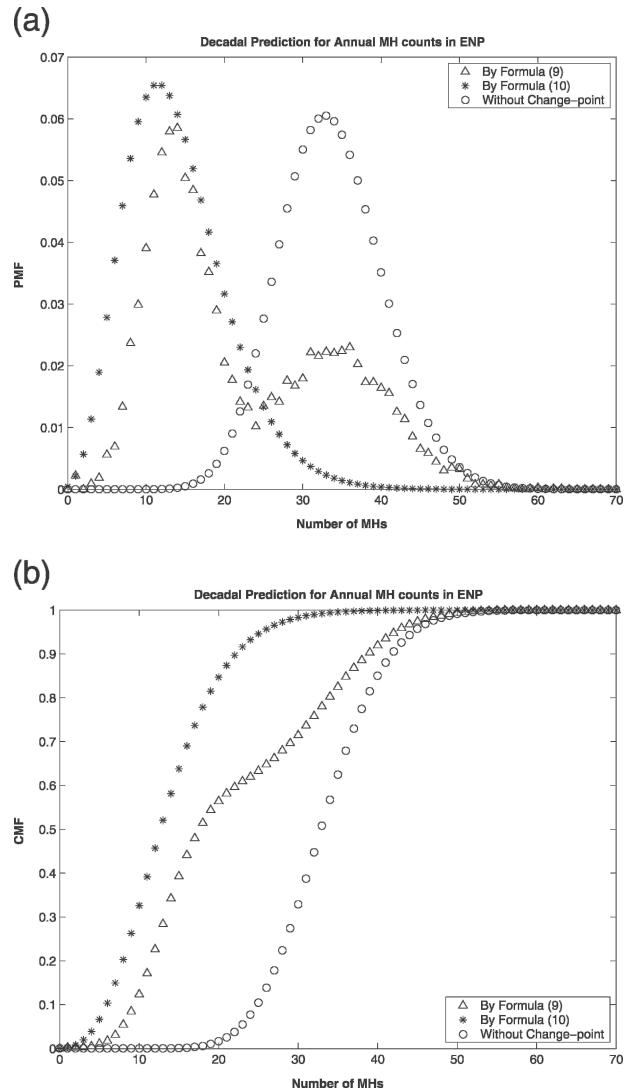


FIG. 6. Decadal predictive distribution of MH counts in the ENP: (a) probability and (b) cumulative mass function. Triangles indicate the prediction made by the weighted average; asterisks denote the prediction by the two changepoint hypotheses; open circles denote the prediction without considering the changepoint.

*Acknowledgments.* We are grateful to one anonymous reviewer for many constructive comments and criticisms. This paper is funded in part by a grant from NOAA. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any of its subagencies.

### APPENDIX A

#### Derivation of a Conditional Posterior PDF

$$P(\tau | \mathbf{h}, H_1, \lambda_{11}, \lambda_{12})$$

Given  $\mathbf{h} = [h_1, h_2, \dots, h_n]$  and  $h_i \approx \text{Poisson}(h_i | \lambda_{11}, T = 1)$  when  $i = 1, 2, \dots, \tau - 1$  and  $h_i \approx \text{Poisson}(h_i | \lambda_{12}, T = 1)$  when  $i = \tau, \dots, n$ , we have

$$\begin{aligned}
 P(\mathbf{h}|H_1, \lambda_{11}, \lambda_{12}, \tau) &= \left( \prod_{i=1}^{\tau-1} \frac{e^{-\lambda_{11}\lambda_{11}^{h_i}}}{h_i!} \right) \left( \prod_{i=\tau}^n \frac{e^{-\lambda_{12}\lambda_{12}^{h_i}}}{h_i!} \right) \\
 &= \left( \prod_{i=1}^n \frac{e^{-\lambda_{12}\lambda_{12}^{h_i}}}{h_i!} \right) \left( \frac{\prod_{i=1}^{\tau-1} e^{-\lambda_{11}\lambda_{11}^{h_i}}}{\prod_{i=1}^{\tau-1} e^{-\lambda_{12}\lambda_{12}^{h_i}}} \right) = \left( \prod_{i=1}^n \frac{e^{-\lambda_{12}\lambda_{12}^{h_i}}}{h_i!} \right) e^{-(\tau-1)(\lambda_{11} - \lambda_{12})} \left( \frac{\lambda_{11}}{\lambda_{12}} \right)^{\sum_{i=1}^{\tau-1} h_i}. \tag{A1}
 \end{aligned}$$

With guidance of the Bayesian formula, we have

$$\begin{aligned}
 P(\tau|\mathbf{h}, H_1, \lambda_{11}, \lambda_{12}, \tau) &= \frac{P(\mathbf{h}|H_1, \lambda_{11}, \lambda_{12}, \tau)P(\tau)}{\sum_{\tau=2}^n P(\mathbf{h}|H_1, \lambda_{11}, \lambda_{12}, \tau)P(\tau)}, \\
 \tau &= 2, 3, \dots, n \tag{A2}
 \end{aligned}$$

where  $P(\tau)$  is the prior density for the changepoint.

Here  $P(\tau)$  can be of any discrete distribution, however, the proper noninformative choice is uniform distribution. Under this assumption, we have

$$\begin{aligned}
 P(\tau|\mathbf{h}, H_1, \lambda_{11}, \lambda_{12}, \tau) &\propto P(\mathbf{h}|H_1, \lambda_{11}, \lambda_{12}, \tau) \\
 &\propto e^{-(\tau-1)(\lambda_{11} - \lambda_{12})} \left( \frac{\lambda_{11}}{\lambda_{12}} \right)^{\sum_{i=1}^{\tau-1} h_i}, \\
 \tau &= 2, 3, \dots, n. \tag{A3}
 \end{aligned}$$

APPENDIX B

**Derivation of Conditional Posterior PDF**

$$\begin{aligned}
 &P(\tau_1|\mathbf{h}, H_2, \lambda_{21}, \lambda_{22}, \lambda_{23}, \tau_2) \text{ and} \\
 &P(\tau_2|\mathbf{h}, H_2, \lambda_{21}, \lambda_{22}, \lambda_{23}, \tau_1)
 \end{aligned}$$

For  $P(\tau_1|\mathbf{h}, H_2, \lambda_{21}, \lambda_{22}, \lambda_{23}, \tau_2)$ , with given  $\tau_2$ , for the data in the first and second epochs,  $[h_1, h_2, \dots, h_{\tau_2-1}]$ , the case is equivalent to the single changepoint hypothesis as described in appendix A except that the data

range from 1 to  $\tau_2 - 1$  and the rates for the first and the second epochs are  $\lambda_{21}$  and  $\lambda_{22}$ , respectively. Thus, with noninformative prior for  $\tau_1|\tau_2$ , the conditional posterior for  $\tau_1$  is

$$\begin{aligned}
 P(\tau_1|\mathbf{h}, H_2, \lambda_{21}, \lambda_{22}, \lambda_{23}, \tau_2) &\propto e^{-(\tau_1-1)(\lambda_{21} - \lambda_{22})} \left( \frac{\lambda_{21}}{\lambda_{22}} \right)^{\sum_{i=1}^{\tau_1-1} h_i}, \\
 \tau_1 &= 2, 3, \dots, \tau_2 - 1. \tag{B1}
 \end{aligned}$$

By similar argument, with given  $\tau_1$  the data in the second and third epochs,  $[h_{\tau_1}, h_{\tau_1+1}, \dots, h_n]$ , is also with only a single changepoint. Thus, with given the rate for the epochs before and after  $\tau_2$ ,  $\lambda_{22}$ , and  $\lambda_{23}$ , respectively, also with noninformative assumption for the prior  $\tau_2|\tau_1$ , we have the conditional posterior for  $\tau_2$ :

$$\begin{aligned}
 P(\tau_2|\mathbf{h}, H_2, \lambda_{21}, \lambda_{22}, \lambda_{23}, \tau_1) &\propto e^{-(\tau_2 - \tau_1)(\lambda_{22} - \lambda_{23})} \left( \frac{\lambda_{22}}{\lambda_{23}} \right)^{\sum_{i=\tau_1}^{\tau_2-1} h_i}, \\
 \tau_2 &= \tau_1 + 1, \dots, n. \tag{B2}
 \end{aligned}$$

APPENDIX C

**Derivation of Likelihood**

$$P(\mathbf{h}|\boldsymbol{\theta}, H_i) \quad i = 0, 1, 2$$

Given

$$\begin{aligned}
 \ln(P(\mathbf{h}|\boldsymbol{\theta}, H_0)) &= \ln \left( \prod_{i=1}^n \frac{e^{-\lambda_0}\lambda_0^{h_i}}{h_i!} \right) = -n\lambda_0 + \ln(\lambda_0)^* \sum_{i=1}^n h_i - \sum_{i=1}^n \ln(h_i!) \\
 \ln(P(\mathbf{h}|\boldsymbol{\theta}, H_1)) &= \ln \left[ \left( \prod_{i=1}^{\tau-1} \frac{e^{-\lambda_{11}\lambda_{11}^{h_i}}}{h_i!} \right) \left( \prod_{i=\tau}^n \frac{e^{-\lambda_{12}\lambda_{12}^{h_i}}}{h_i!} \right) \right] \\
 &= -[\lambda_{11}(\tau - 1) + \lambda_{12}(n - \tau + 1)] + \ln(\lambda_{11})^* \left( \sum_{i=1}^{\tau-1} h_i \right) + \ln(\lambda_{12})^* \left( \sum_{i=\tau}^n h_i \right) - \sum_{i=1}^n \ln(h_i!) \\
 \ln(P(\mathbf{h}|\boldsymbol{\theta}, H_2)) &= \ln \left[ \left( \prod_{i=1}^{\tau_1-1} \frac{e^{-\lambda_{21}\lambda_{21}^{h_i}}}{h_i!} \right) \left( \prod_{i=\tau_1}^{\tau_2-1} \frac{e^{-\lambda_{22}\lambda_{22}^{h_i}}}{h_i!} \right) \left( \prod_{i=\tau_2}^n \frac{e^{-\lambda_{23}\lambda_{23}^{h_i}}}{h_i!} \right) \right] \\
 &= -[\lambda_{21}(\tau_1 - 1) + \lambda_{22}(\tau_2 - \tau_1) + \lambda_{23}(n - \tau_2 + 1)] + \ln(\lambda_{21})^* \left( \sum_{i=1}^{\tau_1-1} h_i \right) + \ln(\lambda_{22})^* \left( \sum_{i=\tau_1}^{\tau_2-1} h_i \right) \\
 &\quad + \ln(\lambda_{23})^* \left( \sum_{i=\tau_2}^n h_i \right) - \sum_{i=1}^n \ln(h_i!),
 \end{aligned}$$

obviously,  $\sum_{i=1}^n \ln(h_i!)$  does not contain any information about the parameters or the hypothesis choice. Thus,

after removing the constant part  $e^{-\sum_{i=1}^n \ln(h_i!)}$ , we get the following formula:

$$P(\mathbf{h}|\boldsymbol{\theta}, \mathbf{H}) \propto \begin{cases} e^{-n\lambda_0 * \sum_{i=1}^n h_i} & \text{if } H = H_0 \\ e^{-(\lambda_{11}(\tau-1) + \lambda_{12}(n-\tau+1)) * \sum_{i=1}^{\tau-1} h_i * \sum_{i=\tau}^n h_i} & \text{if } H = H_1 \\ e^{-(\lambda_{21}(\tau_1-1) + \lambda_{22}(\tau_2-\tau_1) + \lambda_{23}(n-\tau_2+1)) * \sum_{i=1}^{\tau_1-1} h_i * \sum_{i=\tau_1}^{\tau_2-1} h_i * \sum_{i=\tau_2}^n h_i} & \text{if } H = H_2 \end{cases} \quad (\text{C1})$$

#### REFERENCES

- Berliner, L. M., C. K. Wilke, and N. Cressie, 2000: Long-lead prediction of Pacific SSTs via Bayesian dynamic modeling. *J. Climate*, **13**, 3953–3968.
- Carlin, B. P., and T. A. Louis, 2000: *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman & Hall/CRC, 419 pp.
- Chu, P.-S., 2002: Large-scale circulation features associated with decadal variations of tropical cyclone activity over the central North Pacific. *J. Climate*, **15**, 2678–2689.
- , and X. Zhao, 2004: Bayesian change-point analysis of tropical cyclone activity: The central North Pacific case. *J. Climate*, **17**, 4893–4901.
- Clark, J. D., and P.-S. Chu, 2002: Interannual variation of tropical cyclone activity over the central North Pacific. *J. Meteor. Soc. Japan*, **80**, 403–418.
- Collins, J. M., and I. M. Mason, 2000: Local environmental conditions related to seasonal tropical cyclone activity in the northeast Pacific basin. *Geophys. Res. Lett.*, **27**, 3881–3884.
- Congdon, P., 2003: *Bayesian Statistical Modeling*. John Wiley & Sons, 529 pp.
- Deser, C., A. S. Phillips, and J. W. Hurrell, 2004: Pacific interdecadal climate variability: Linkages between the Tropics and the North Pacific during boreal winter since 1900. *J. Climate*, **17**, 3109–3124.
- Elsner, J. B., and B. H. Bossak, 2001: Bayesian analysis of U.S. hurricane climate. *J. Climate*, **14**, 4341–4350.
- , and T. H. Jagger, 2004: A hierarchical Bayesian approach to seasonal hurricane modeling. *J. Climate*, **17**, 2813–2827.
- , X. Niu, and T. H. Jagger, 2004: Detecting shifts in hurricane rates using a Markov Chain Monte Carlo approach. *J. Climate*, **17**, 2652–2666.
- Epstein, E. S., 1985: *Statistical Inference and Prediction in Climatology: A Bayesian Approach*. Meteor. Monogr., No. 42, Amer. Meteor. Soc., 199 pp.
- Gelfand, A. E., and D. Dey, 1994: Bayesian model choice: Asymptotics and exact calculations. *J. Roy. Stat. Soc. Ser. B*, **56**, 501–514.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin, 2004: *Bayesian Data Analysis*. 2d ed. Chapman & Hall/CRC, 668 pp.
- Goldenberg, S. B., C. W. Landsea, A. M. Mestas-Nunez, and W. M. Gray, 2001: The recent increase in Atlantic hurricane activity: Causes and implications. *Science*, **293**, 474–479.
- Lavielle, M., and M. Labarbier, 2001: An application of MCMC methods for multiple change-points problem. *Signal Process.*, **81**, 39–53.
- Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis, 1997: A Pacific interdecadal oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.*, **78**, 1069–1079.
- Ripley, B. D., 1987: *Stochastic Simulation*. John Wiley, 237 pp.
- Whitney, L. D., and J. S. Hobgood, 1997: The relationship between sea surface temperatures and maximum intensities of tropical cyclone intensities in the eastern North Pacific. *J. Climate*, **10**, 2921–2930.