Bayesian Change-Point Analysis of Tropical Cyclone Activity: The Central North Pacific Case*

PAO-SHIN CHU AND XIN ZHAO

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

(Manuscript received 19 August 2003, in final form 8 July 2004)

ABSTRACT

Bayesian analysis is applied to detect change points in the time series of annual tropical cyclone counts over the central North Pacific. Specifically, a hierarchical Bayesian approach involving three layers—data, parameter, and hypothesis—is formulated to demonstrate the posterior probability of the shifts throughout the time from 1966 to 2002. For the data layer, a Poisson process with gamma distributed intensity is presumed. For the hypothesis layer, a ‘‘no change in the intensity’’ hypothesis and a ‘‘single change in the intensity’’ hypothesis are considered. Results indicate that there is a great likelihood of a change point on tropical cyclone rates around 1982, which is consistent with earlier work based on a simple log-linear regression model. A Bayesian approach also provides a means for predicting decadal tropical cyclone variations. A higher number of tropical cyclones is predicted in the next decade when the possibility of the change point in the early 1980s is taken into account.

1. Introduction

Analyzing temporal changes in a climate time series is becoming increasingly important as we often need to know when a major shift in climate systems occurs. This information, if assessed appropriately, would aid researchers and planners in their strategy for more comprehensive analyses of complex climate systems and in sound decision-making processes. One such example is the well-known changing phase of the Pacific decadal oscillation (PDO) in the late 1970s. Studies have shown that the negative phase of the PDO is instrumental for the wintertime precipitation in the Pacific Northwest while the positive phase of the PDO does just the opposite (Mantua et al. 1997). Therefore, knowing the turning phase of a major climate system would be beneficial for many sectors such as agriculture and hydropower operations.

Bayesian analysis is an efficient way to provide a coherent and rational framework for distilling uncertainties by incorporating diverse information sources such as subjective beliefs, historical observations, model simulations, and new information. A comprehensive textbook introducing the Bayesian paradigm and its applications to atmospheric data is Epstein (1985). Solow (1988) applied a Bayesian method for inferences about climate change based on the two-phase regression model. Elsner and Bossak (2001) explicitly demonstrated the use of Bayesian analysis to the U.S. hurricanes by combining the less reliable historical accounts of hurricanes in the nineteenth century with the more reliable records from the twentieth century to yield a best estimate of the annual rates. Besides using a Bayesian technique for making inferences, Bayes' theorem can be applied in a predictive mode for the probability of future U.S. landfalling hurricanes (e.g., Epstein 1985; Elsner and Bossak 2001). This feature is applicable for disaster mitigation planning and insurance/reinsurance industries because landfalling hurricanes cause enormous property damage and their future occurrences are unknown in relation to climate variability.

Using a step function as an independent variable and taking a logarithmic transformation of the annual major hurricane rates over the North Atlantic as a dependent variable, Elsner et al. (2000) developed a model for detecting change points in the Atlantic hurricane time series. Chu (2002) used a similar log-linear regression method to model the shifts in annual tropical cyclone (TC) frequency over the central North Pacific (CNP) for the period 1966–2000. It was found that two change points are significant at $\alpha = 0.05$ level. A change-point time may be defined as the last year of an old epoch or the first year of a new epoch. Here the latter definition was adopted. The first change point occurred in 1982 with a t ratio of 2.45, and the second shift occurred in 1995 with a marginally significant t ratio of 2.04. As a

* School of Ocean and Earth Science and Technology Contribution Number 6490.

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 Road, H1G 331, Honolulu, HI 96822.
E-mail: chu@hawaii.edu

© 2004 American Meteorological Society

result, the entire 35-yr record is partitioned into three epochs, 1966–81, 1982–94, and 1995–2000.

Though the approach taken by Chu (2002) provided a simple and straightforward analysis as to when decadal variations in TC frequency occurred, it does not take into account the fact that the seasonal TC occurrence may be better described by a discrete Poisson process. Moreover, the log-linear model does not contain information related to the posterior probability for the change point, which is important for the prediction of future outcomes of the process. Elsner and Schmertmann (1993) and Wilson (1999) noted that the very low number of intense hurricanes over the North Atlantic and the lack of serial dependence in intense hurricanes from year to year justify the Poisson assumption. This may also be the case for the CNP because the annual TC events are rather few. For example, during the last 37 yr (1966–2002), 5 yr have no TC activity, 6 yr have only one TC occurrence in each year, and 8 yr have two TC occurrences in each year.

In view of the shortcoming of representing TC variations by a linear model, this study attempts to model the temporal changes of TC activity by a Poisson process with the Poisson parameter being treated as a gamma distribution. The resulting Bayesian analysis is then used to forecast the future TC activity over CNP. The essential issue of this study is that, rather than assuming the statistical distribution of the TC rate is time invariant throughout the observation period, we introduce a ''single change'' hypothesis under which there is a major shift on the TC activity rate. A few remarks justify the application of Bayesian methods in this study. As will be seen in section 4, the Bayesian approach is able to include other sources of less reliable data as prior information in the analysis, a certain advantage over the non-Bayesian method which often depends on more reliable but shorter portions of the historical records. Moreover, inferences about temporal shifts are couched in terms of probabilities, another desirable feature of the Bayesian paradigm. In contrast, the classical non-Bayesian method provides a deterministic estimate of the change-point location, but not probability information about the uncertainty of change points.

The structure of this paper is as follows. Section 2 describes the data source. The basic mathematical model of TC activity is reviewed in section 3. Section 4 introduces the three-level Bayesian analysis framework pertinent to our specific problem. Main results are described in section 5. A summary and discussion are found in section 6.

2. Data

The tropical cyclone records over the CNP at 6-h intervals come from the National Hurricane Center's (NHC) best tracks dataset, as described in Clark and Chu (2002). Here, tropical cyclones refer to tropical storms and hurricanes. In this study, tropical storms are defined as the maximum sustained surface wind speeds between 17.5 and 33 m s^{-1}, and hurricanes are defined as wind speeds at least 33 m s^{-1}. Mayfield and Rappaport (1992) suggest that reliable TC statistics in the CNP began in 1966, when satellite reconnaissance was initiated in the region. A second dataset used is the recent TC records compiled annually by the Central Pacific Hurricane Center, an entity of the National Weather Service Forecast Office in Honolulu, Hawaii. By combining these two datasets, reliable TC records extend from 1966 to 2002, which constitutes the main dataset for this study.

For Bayesian analysis, prior information of TCs before 1966 is needed. Data prior to 1966 can be found in Shaw (1981), who did the laudable task of compiling historical TC records for the CNP from various sources, including Mariners Weather Log, the Joint Typhoon Warning Center's annual typhoon reports, real-time cyclone tracks and advisories issued by the Central Pacific Hurricane Center, published and unpublished papers by emeritus Prof. James Sadler of the University of Hawaii, and others. Although historical TC accounts are thought to be less reliable, we are only concerned with the annual counts of TC in this study, not with the attributes of TC at 6-h intervals as detailed in the NHC's best track dataset.

3. Mathematical model of TC activity

The distribution of the annual counts of TCs in the vicinity of Hawaii is considered as a Poisson process (Chu and Wang 1998). The Poisson process is governed by a single parameter, λ, the intensity. Poisson events imply independence, meaning that the number of occurrences in a particular period of time is not affected by the outcomes in other time periods. A check of the sample annual TC counts over the CNP reveal very low autocorrelations (i.e., near 0.1) from lags of 1–5 yr during 1966–2002, suggesting small interannual correlations. Given the intensity parameter λ, the probability of h TCs occurring in T years is

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

where

$$h = 0, 1, 2, \ldots \quad \text{and} \quad \lambda > 0, \qquad T > 0.$$

The Poisson mean is simply the product of λT, so is its variance.

To examine whether the annual TC rates in the CNP follow a constant-rate Poisson process, we use a scheme developed by Keim and Cruise (1998). Denote R as the ratio of the sample variance and sample mean (i.e., $R = \sigma^2/\overline{x}$). This ratio is then compared against a critical R value (R_c), which is defined as $[\chi^2_{n-1,\alpha}/n - 1)]$, where n is the sample size for a one-sided test (Keim and Cruise 1998; Elsner et al. 2000). The α is the specified test level, chosen as 0.05 in this study. For the 1966–

2002 TC record, the R is 1.81 and R_c is 1.42, guiding us to reject a constant-rate Poisson process throughout the entire time period (Elsner et al. 2000). Thus, it is better not to treat λ as a determinant single-value parameter but a random variable. The resulting hierarchical feature fits well with the Bayesian inference because it is predicated on the assumption that the distribution parameter is treated as a random quantity and inferences are based on the probability distribution of the parameter given the data. A functional choice of λ is a gamma distribution (Epstein 1985) as expressed in the following form:

$$f(\lambda \mid h', T') = \frac{T'^{h'} \lambda^{h'-1}}{\Gamma(h')} \exp(-\lambda T'), \qquad \lambda > 0, \quad h' > 0, \quad T' > 0, \tag{2}$$

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

Given h TCs occurring in T yr, if the prior density for λ is gamma distributed with parameters h' and T', the posterior density for λ will also be gamma distributed with parameters $h + h'$ and $T + T'$. That is, the gamma density is the conjugate prior for λ. Referring to (2), the conditional expectation with respect to λ is $\mathrm{E}[\lambda \mid h', T'] = h'/T'$. In the later part of this paper, we will discuss how to find the prior information h' and T'.

Under the statistical model introduced earlier, the probability density function (PDF) of h TCs occurring in T yr when knowledge of the intensity is codified as a gamma density with prior parameters h' and T' is then the negative binomial distribution (Epstein 1985; Elsner and Bossak 2001; Gelman et al. 2004):

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

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

4. Bayesian approach for detection of shift in the TC series

a. Hypothesis model

In this study, we will mainly focus on the case in which the probability of more than one change point within the desired period is negligible. This simplifies the analysis to one of the two scenarios: a "no change-point" hypothesis versus a "single change-point" hypothesis. The following derivations are based on the mathematical model described in section 3. The annual tropical cyclone data, $h_1, h_2, \ldots, h_n$, are assumed to be described as a series of independent random variables. Mathematically, the two hypotheses models are postulated below.

1) Hypothesis H_0: "no change point of the rate" of the TC series:
$h_i \sim$ Poisson $(h_i \mid \lambda, T)$, $i = 1, 2, \ldots, n$, where T is the unit observation time
$\lambda \sim$ gamma (h', T'),
where the prior knowledge of the parameters h' and T' is given.
2) Hypothesis H_1: "a single change point of the rate" of the TC series:
$h_i \sim$ Poisson $(h_i \mid \lambda_1, T)$, when $i = 1, 2, \ldots, \tau - 1$
$h_i \sim$ Poisson $(h_i \mid \lambda_2, T)$, when $i = \tau, \ldots, n$
$\tau = 2, 3, \ldots, n$, T is as defined in the hypothesis H_0, and
$\lambda_1 \sim$ gamma (h'_1, T'_1)
$\lambda_2 \sim$ gamma (h'_2, T'_2),
where the prior knowledge of the parameters h'_1, T'_1, h'_2, T'_2 is given.

Note that there are two epochs in this model and τ is known as the first year of the new epoch, or the change point.

b. Hypothesis analysis

With the formula of Bayesian inference under hypothesis H_1 as described in the appendix, we will derive the Bayesian method to analyze the posterior probability of the hypothesis model H_0 and H_1 based on the given observation data and statistical assumption described previously. Basically, we need to determine the prior distribution for the hypothesis model H_0 and H_1, which can be of any discrete probability distribution function. A proper noninformative choice is the uniform distribution, that is, $P(H_0) = P(H_1) = 1/2$ since there is no prior information regarding which one of the hypotheses is preferable.

Provided the prior model distribution, using the Bayes' rule, we obtain its posterior distribution:

$$P(H_i \mid \mathbf{h}) = \frac{P(\mathbf{h} \mid H_i) P(H_i)}{\sum_{i=0}^{1} P(\mathbf{h} \mid H_i) P(H_i)}, \qquad i = 0, 1, \tag{4}$$

where $P(\mathbf{h} \mid H_1) = \sum_{\tau=2}^{n} P(\mathbf{h} \mid \tau, H_1) P(\tau \mid H_1)$ and $P(\mathbf{h} \mid \tau, H_1)$ is given in (A2). $P(\mathbf{h} \mid H_0)$ is expressed as

$$P(\mathbf{h} \mid H_0) = \prod_{i=1}^{n} \frac{\Gamma(h_i + h')}{\Gamma(h') h_i!} \left(\frac{T'}{1 + T'}\right)^{h'} \left(\frac{1}{1 + T'}\right)^{h_i}. \tag{5}$$

When the models H_0 and H_1 are compared, we use the ratio of posterior probability of the model as defined in the following:

$$\frac{P(H_1 \mid \mathbf{h})}{P(H_0 \mid \mathbf{h})} = \frac{P(\mathbf{h} \mid H_1) P(H_1)}{P(\mathbf{h} \mid H_0) P(H_0)}. \tag{6}$$

TABLE 1. Raftery's scale for interpreting the Bayes factors.

$2\ln B$	Evidence for Bayesian model
0–2	Not worth more than a bare mention
2–6	Positive
6–10	Strong

Rearranging terms in (6) leads to the definition of the term "Bayes factor":

$$B = \left[\frac{P(H_1|\mathbf{h})}{P(H_0|\mathbf{h})}\right] \bigg/ \left[\frac{P(H_1)}{P(H_0)}\right] = \frac{P(\mathbf{h}|H_1)}{P(\mathbf{h}|H_0)}. \quad (7)$$

The Bayes factor B is a measure of whether the observation data $\mathbf{h}$ have increased or decreased the odds on the hypothesis H_1 relative to the hypothesis H_0. It is only the ratio of the likelihood for both hypotheses and is independent of the hypothesis prior choice. In the uniform prior case, if $B > 1$, it is more plausible to select the hypothesis H_1 rather than H_0. However this criterion may be too loose for many problems and Raftery (1996) suggested a more conservative guideline to interpret this factor (Table 1), which favorably mitigates the effects of hypothesis prior bias.

c. *Predictive distributions*

Following the model selection procedure established in the previous section, in the case that the hypothesis H_0 is more likely, the predictive distribution for the number of TCs occurring in the future $\hat{T}$ yr, say $\hat{h}$, with the given prior knowledge h', T', and the data $\mathbf{h}$ observed in a period of T yr (here T is equal to n), is a straightforward negative binomial distribution:

$$P(\hat{h}|h', T', \hat{T}, H_0, \mathbf{h}, T) = \frac{\Gamma[\hat{h} + (h' + h_{\text{sum}})]}{\Gamma(h' + h_{\text{sum}})\hat{h}!} \times \left[\frac{T' + T}{\hat{T} + (T' + T)}\right]^{h'+h_{\text{sum}}} \left[\frac{\hat{T}}{\hat{T} + (T' + T)}\right]^{\hat{h}}, \quad (8)$$

where $h_{\text{sum}} = \sum_{i=1}^{T} h_i$, $\hat{h} = 0, 1, \ldots$ and $h' > 0$, $T' > 0$, $T > 0$, $\hat{T} > 0$.

In the case that we are under the hypothesis H_1, the predictive PDF for $\hat{h}$ TCs in future $\hat{T}$ yr conditional on the prior knowledge h'_1, T'_1, h'_2, T'_2 as well as the observation data $\mathbf{h}$ and observation length T will be

$$P(\hat{h}|h'_1, T'_1, h'_2, T'_2, \hat{T}, H_1, \mathbf{h}, T) = \sum_{\tau=2}^{T} P_{\text{nb}}\left(\hat{h}|\tilde{h}_2, \frac{\tilde{T}_2}{\hat{T} + \tilde{T}_2}\right) P(\tau|\mathbf{h}, H_1), \quad (9)$$

where $P_{\text{nb}}(\cdot)$ stands for the negative binomial distribution, $\tilde{h}_2$ and $\tilde{T}_2$ are as defined in the formula (A5), and $P(\tau|\mathbf{h}, H_1)$ can be calculated by (A3).

Finally, with the assumption of only two hypotheses, H_0 and H_1, a comprehensive formula of the predictive distribution for the TC frequency in a given period $\hat{T}$ yr, while suppressing the notation of the prior parameters and observation period length T, will be

$$P(\hat{h}|\mathbf{h}, \hat{T}) = P(\hat{h}|\mathbf{h}, \hat{T}, H_0)P(H_0|\mathbf{h}) + P(\hat{h}|\mathbf{h}, \hat{T}, H_1)P(H_1|\mathbf{h}), \quad (10)$$

where $P(\hat{h}|\mathbf{h}, \hat{T}, H_0)$ and $P(\hat{h}|\mathbf{h}, \hat{T}, H_1)$ come from (8) and (9), respectively, and $P(H_0|\mathbf{h})$ and $P(H_1|\mathbf{h})$ are obtained from (4).

Formula (10) is at the heart of the Bayesian prediction scheme; however, it also leads to considerable computational difficulty. For the sake of convenience, sometimes we may be able to just fix the hypothesis and the change-point position (when the H_1 hypothesis is significantly more likely) to its maximum a posterior (MAP) estimation, which is defined as the value with the largest posterior probability for a random variable when calculating the predictive PDF of the TC counts in a desired period. If the H_0 hypothesis is dominant, we only need to use (8) and if the H_1 hypothesis prevails, with a fixed change point, formula (10) is reduced to

$$P(\hat{h}|h'_1, T'_1, h'_2, T'_2, \hat{T}, H_1, \mathbf{h}, T, \hat{\tau}) = P_{\text{nb}}\left[\hat{h}|h' + \sum_{i=\hat{\tau}}^{T} h_i, \frac{T' + T - \hat{\tau} + 1}{\hat{T} + (T' + T - \hat{\tau} + 1)}\right] \quad (11)$$

where

$$\hat{\tau} = \underset{\tau}{\operatorname{argmax}}\,[P(\tau|\mathbf{h}, H_1), \tau = 2, 3, \ldots, T].$$

This simplified equation (11) is obviously biased relative to (10). However, if the posterior probability of the H_1 hypothesis is much larger than that of the H_0 hypothesis and the MAP estimation of the change point is comparatively much higher than the probability of other years, this bias should be within tolerance. As will be shown later, this simplification does not much impact the final prediction result.

d. *Calculation of the prior parameters*

So far, we have constructed the theoretical framework for a three-level hierarchical Bayesian analysis, but we have not mentioned how to obtain the prior knowledge of the distribution, that is, the estimation of prior parameters h' and T' for H_0 hypothesis and parameters h'_1, T'_1, h'_2, T'_2 for H_1 hypothesis. The prior parameters represent the degree of belief in the outcomes prior to a particular data sample being observed. For estimating prior parameters, data prior to 1966 are needed. We used the records from Shaw's report for the period of 1957–65 as our prior knowledge to estimate the parameters for the "before change" period. The mean of TCs during this short time period is 2.2 TCs yr^{-1}. Data prior to 1957 are also available but tend to be rather unreliable. For example, there is not a single TC reported for six consecutive years from 1951 to 1956.

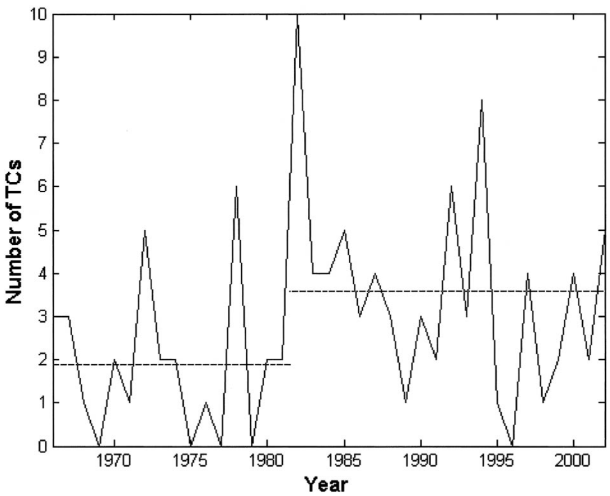

FIG. 1. Time series of annual tropical cyclone counts over the central North Pacific from 1966 to 2002. Broken lines denote the means for the period 1966–81 and 1982–2002, respectively.

It is also necessary to have prior parameters for the post change point (e.g., Tapsoba et al. 2004). In Chu (2002), a significant shift in 1982 and a marginally significant shift in 1995 were suggested. To make inferences about the shift around 1982, it is reasonable to choose the data before 1995 as the target period for the ‘‘after change’’ period under the H_1 hypothesis. To maximize the available sample size for the modeling period, we thus select the annual TC observations from 1990–94 as the prior. As for the H_0 hypothesis, we just combine this set of data with the short, pre-1966 dataset to form the prior knowledge.

For the two prior periods aforementioned, one needs to estimate parameters h' and T'. A straightforward method to estimate the distribution parameters of the negative binomial is to use the moment statistics (Carlin and Louis 2000). It is sufficient to just use the first two moments of the data to estimate h' and T', which can be formulated as in the following:

$$\hat{T}' = \frac{m_h}{s_h^2 - m_h},$$
$$\hat{h}' = m_h\hat{T}', \qquad (12)$$

where $m_h = 1/n \sum_{i=1}^{n} h_i$, $S_h^2 = 1/(n-1) \sum_{i=1}^{n} (h_i - m_h)^2$ are the sample mean and sample variance for the given data respectively. In most cases, the moment method can give reasonable estimation of model parameters when the sample size is large. However, when the sample mean is very close to the sample variance or when the sample mean is larger than the sample variance, Eq. (12) breaks down.

In this study, under H_1, for the ‘‘before change-point’’ period, sample mean and sample variance are equal to 2.22 and 5.44, respectively. From (12), this leads to $h_1' = 1.53$ and $T_1' = 0.69$. For the ‘‘after change-point’’ period, sample mean and sample variance are equal to 4.40 and 6.30, respectively, yielding $h_2' = 10.19$ and $T_2' = 2.32$. In contrast, under H_0, sample mean and sample variance are equal to 3.00 and 6.46; therefore, the resulting prior h' and T' are equal to 2.60 and 0.87, respectively.

TABLE 2. Results of the Bayesian analysis on change-point of annual TC counts over the CNP. Here, τ stands for the change-point year, B is the Bayes factor, λ_1 and λ_2 represent the TC rate before and after the change point under H_1 hypothesis, respectively, and $P(H_1|\mathbf{h})$ is the posterior probability of hypothesis H_1. The analysis period is 1966–89.

Variable	Value
$\hat{\tau}$	1982
$\underline{P}(\hat{\tau}\|\mathbf{h}, H_1)$	0.31
$\underline{\lambda}_1$	1.88
λ_2	3.57
$2 \ln(B)$	2.22
$P(H_1\|\mathbf{h})$	0.75

5. Results

a. Results of shift in intensity of TCs

Figure 1 shows the time series of annual TC counts over the CNP since 1966. The average rate prior to 1982 is about 1.9 TCs yr^{-1} but it increases to almost 3.6 TCs yr^{-1} thereafter. The result of the Bayesian analysis on the shift year of the annual TC counts in CNP is listed in Table 2. From this table, we can see that the measure of Bayes factor [$2 \ln(B)$] for the annual TC counts during the 1966–89 period is 2.22, which favors H_1 over the H_0 hypothesis with a uniform prior for the hypothesis layer. The posterior probability that a change has occurred is rather high, reaching 0.75. Figure 2 shows the posterior probability of the change point of TC activity, plotted as a function of time. Larger probabilities on year *i* imply a more likely change occurring with *i* being the first year of a new epoch. The maximum probability

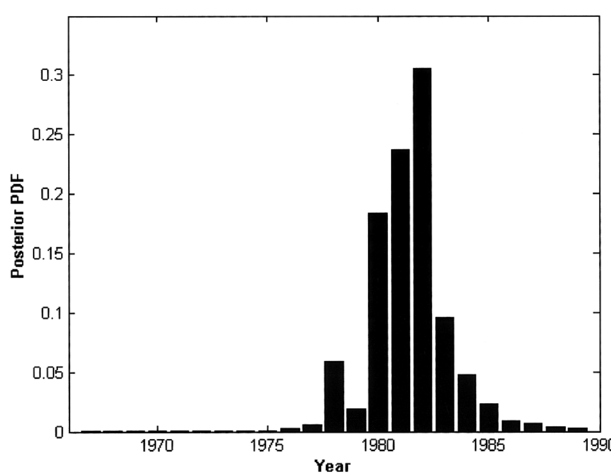

FIG. 2. Posterior probability distribution of the change point, $P(\tau\,|\,\mathbf{h}, H_1)$ of TC series over the CNP.

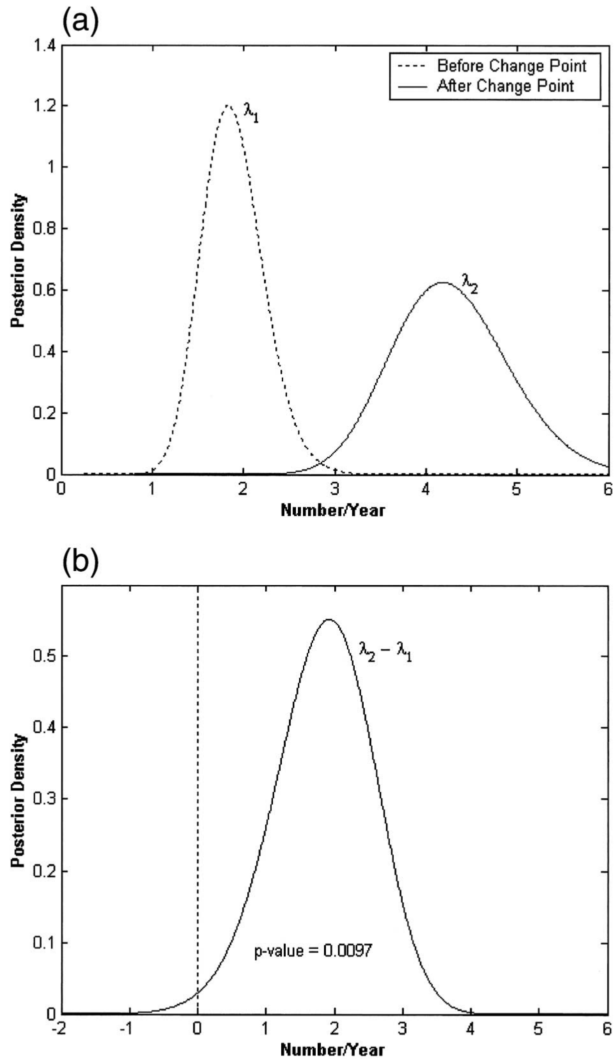

FIG. 3. (a) Posterior density function of annual TC intensity before the shift, $P(\lambda_1|\mathbf{h}, H_1)$, and after the shift, $P(\lambda_2|\mathbf{h}, H_1)$, with the change-point year being set in 1982. (b) Posterior density of $(\lambda_2 - \lambda_1)$.

of 0.31 occurs in 1982. This suggests that the most likely year of the new epoch is 1982 although other change-point years such as 1981 and 1980 are plausible candidates.

The posterior PDFs of TC intensity before and after the change point, λ_1 and λ_2, are plotted in Fig. 3a. The posterior distribution represents a combination of the prior distribution and the likelihood function. In this plot, the change-point year is fixed in 1982. Recalling from Table 2, the sample rate before 1982 is 1.88 and after 1982 is 3.57. Figure 3a shows very little overlapping in the tail areas between these two posterior distributions, implying a rate increasing for the ''after the shift'' distribution beginning with 1982. Figure 3b displays the posterior density of $(\lambda_2 - \lambda_1)$. The p value of this difference $[P(\lambda_2 - \lambda_1 < 0|H_1, \mathbf{h})]$ is very small (<0.01), strongly supporting the contention of a shift toward a higher rate of annual TC intensity since 1982.

From a log-linear regression analysis, Chu (2002) noted that after a major shift in 1982, a second shift, albeit weak, appears to occur in 1995. In order to test whether the latter shift also can be identified from the Bayesian framework, a similar analysis is performed for the period 1985–2002. For the H_1 hypothesis, two rather short prior periods, 1984–88 and 1998–2002, are chosen. For the H_0 hypothesis, these two periods are combined as the prior. Results indicate that $2\ln(B) = -0.38$, which means the odd is in favor of H_0 hypothesis (i.e., no change point) over H_1 hypothesis for the post-1982 period.

b. Decadal tropical cyclone prediction

After having identified a change-point year in the TC series, our next goal is to predict TC activity over the CNP on a climate time scale. One way to calculate this predictive distribution is to use formula (10); however, this form may be computationally complicated. Based on the Bayesian analysis results presented in Table 2 and Fig. 2, it is reasonable to choose the H_1 hypothesis. Thus, we opt to use the simplified formula (11) with the fixed change point at 1982.

The final decadal predictive PDF and cumulative distribution function (CDF) of TC counts calculated from both (10) and (11) over the CNP are plotted in Fig. 4a and Fig. 4b, respectively. As a comparison, we also plot the predictive PDF and CDF that do not involve the hypothesis layer. In other words, only the traditional approach involving the H_0 hypothesis is assumed and (8) is applied. In Fig. 4a, the PDF is narrower when only H_0 is assumed and becomes broader after considering H_1 hypothesis. Thus, one may expect larger variability in TC rates for the next decade when the H_1 is assumed. Moreover, there is an overlap between the two predictive PDFs under H_0 and H_1, but a significant shift toward the right is clearly seen when H_1 is considered. Figure 4b displays the CDFs of predicting no more than a particular number of cyclones over the next decade. For example, the probability of predicting no more than 40 TCs in the next 10 yr when we only consider the H_1 hypothesis is 0.74 while predicting the same TC numbers under the H_0 hypothesis is 0.98, an almost guaranteed probability of occurrence. Moreover, the difference between the predictive distribution calculated from (10) and the simplified form from (11) is negligible, implying the simplified formula (11) works well for our problem.

6. Summary and discussion

In this study, a hierarchical Bayesian change-point analysis of tropical cyclone counts is developed. Specifically, the annual tropical cyclone counts over the central North Pacific are described by a Poisson process

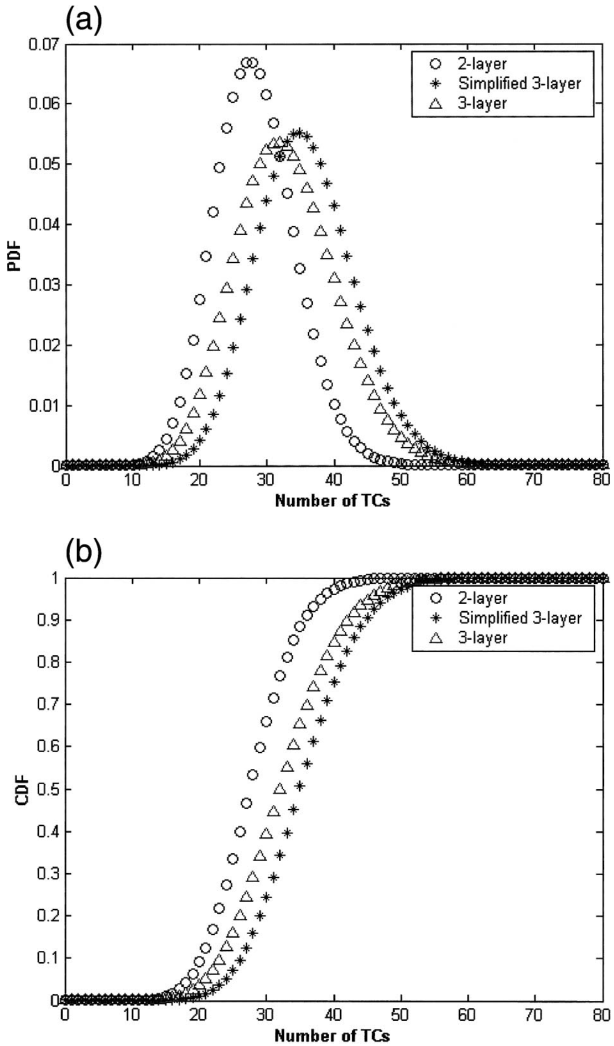

FIG. 4. (a) Predictive PDF and (b) predictive CDF of decadal TC counts over the CNP, where the circle refers to the two-layer Bayesian analysis, the triangle refers to the complete three-layer Bayesian analysis, and the asterisk refers to the simplified three-layer Bayesian analysis by using Eq. (11).

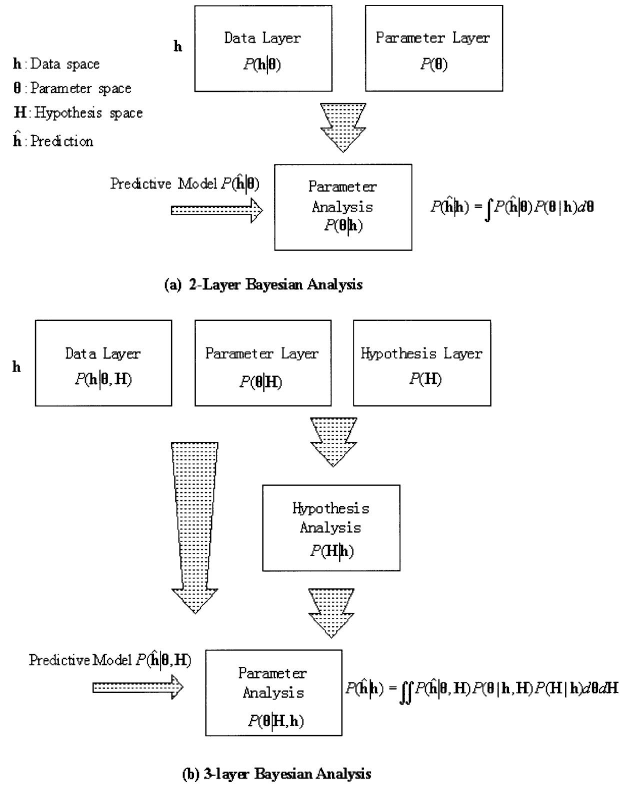

FIG. 5. Hierarchical structure of the Bayesian analysis methodology.

that is conditional on gamma distributions. The method focuses on the scenario in which the probability of more than one shift is negligible. Considering two equiprobable hypotheses, H_0 and H_1, we perform a hierarchical Bayesian analysis of making inferences about shifts in the tropical cyclone series. Inferences are based on the posterior probabilities of the possible shifts. Results suggest that there is a great likelihood of a change point in TC intensity in 1982 over the CNP, which is consistent with our earlier analysis based on a simple log-linear regression method (Chu 2002). Bayesian analysis is also used for predicting decadal tropical cyclone variations, and higher TC frequency is predicted in the next decade when the change point is taken into account. The predicted TC frequency may serve as a benchmark to gauge the future observed TC activity over the central North Pacific.

In the fundamental Bayesian framework, only two layers—a data layer and a parameter layer—are considered for deriving the posterior distribution $P(\boldsymbol{\theta}|\mathbf{h})$ and obtaining the optimum predictive distribution $P(\hat{h}|\mathbf{h}) = \int P(\hat{h}|\boldsymbol{\theta})P(\boldsymbol{\theta}|\mathbf{h})\, d\boldsymbol{\theta}$. As illustrated in Fig. 5, the data layer is embodied by a likelihood distribution $P(\mathbf{h}|\boldsymbol{\theta})$ and a parameter layer is embodied by prior information $P(\boldsymbol{\theta})$. In this framework, no change points are assumed. Expanding from this two-layer thinking, we introduce a new layer, called the hypothesis layer, which is embodied by prior information $P(\mathbf{H})$, where $\mathbf{H}$ represents hypothesis. In this three-layer paradigm (Fig. 5), both the data layer and parameter layer are conditional on hypothesis selection so they are described by $P(\mathbf{h}|\boldsymbol{\theta}, \mathbf{H})$ and $P(\boldsymbol{\theta}|\mathbf{H})$, respectively.

Following the same Bayes' rule, we obtain the posterior distribution for both hypotheses and parameters, $P(\mathbf{H}|\mathbf{h})$ and $P(\boldsymbol{\theta}|\mathbf{H}, \mathbf{h})$. The predictive distribution is thus $P(\hat{h}|\mathbf{h}) = \iint P(\hat{h}|\boldsymbol{\theta}, \mathbf{H})P(\boldsymbol{\theta}|\mathbf{H}, \mathbf{h})P(\mathbf{H}|\mathbf{h})\, d\boldsymbol{\theta}\, d\mathbf{H}$. For the sake of computational simplicity, we also used the simplified formula $P(\hat{h}|\hat{\boldsymbol{\theta}}, \hat{\mathbf{H}})$ instead of the entire integration in the aforementioned predictive distribution, where $\hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{H}}$ are the MAP estimation of $\boldsymbol{\theta}$ and $\mathbf{H}$, respectively. Apparently, the traditional two-layer Bayesian thinking (e.g., Elsner and Bossak 2001) can be viewed as a special case of a three-layer framework. Recently, Tapsoba et al. (2004) used a similar three-layer Bayesian analysis to detect change points in the western Africa rainfall time series. However, normal and inverted gamma distributions were used in their analysis.

Also recently, Elsner et al. (2004) applied a Markov chain Monte Carlo (MCMC) approach based on Gibbs sampling algorithm to detect change points in the Atlantic hurricane series. Gibbs sampling assumes that a value for one element of a multidimensional parameter can be generated when values for all other elements of this parameter are given. With some initial prior values of distribution parameters being prescribed, Gibbs sampling produces sequences of the parameters such as the hurricane rates before and after a change point. This approach provides an alternative to the classical Bayesian change-point analysis involving the prior, likelihood function, and the posterior distribution as presented in this study. While our study and Elsner et al. (2004) focus on a single change-point scenario, more elaborate multiple hypothesis choices such as the ''double change points'' hypothesis have been proposed by Lavielle and Labarbier (2001). It is yet to be demonstrated how such complicated modeling processes can be applied to detecting more than one change point in the hurricane time series.

Acknowledgments. We thank two anonymous reviewers and Francis Zwiers for suggestions that led to improvements in the manuscript. Partial support for this study has been provided by NOAA Grant NA17RJ1230.

APPENDIX

Bayesian Inference for the Hypothesis H_1

A Bayesian approach is considered as making inferences under a single change in a time series. That is, the hypothesis H_1 is being considered here. Since the annual TC counts are used, the unit observation period is 1 yr; therefore, $T = 1$. Following the formula shown in (A3), the likelihood of the annual count of TCs under the H_1 hypothesis is as shown next (with the given change-point τ):

$$P(h_i \mid h_1', T_1', h_2', T_2', \tau, H_1) = \begin{cases} \dfrac{\Gamma(h_i + h_1')}{\Gamma(h_1')h_i!}\left(\dfrac{T_1'}{1 + T_1'}\right)^{h_1'}\left(\dfrac{1}{1 + T_1'}\right)^{h_i}, & i = 1, 2, \ldots, \tau - 1 \\ \dfrac{\Gamma(h_i + h_2')}{\Gamma(h_2')h_i!}\left(\dfrac{T_2'}{1 + T_2'}\right)^{h_2'}\left(\dfrac{1}{1 + T_2'}\right)^{h_i}, & i = \tau, \ldots, n, \end{cases} \quad \text{(A1)}$$

where $\tau = 2, 3, \ldots, n$.

Subject to the assumption that the number of occurrence of annual TCs over the CNP is independent from year to year, and dropping the notations for parameters h_1', T_1', h_2', T_2' in (A1) for the sake of simplicity, the vector form of the likelihood function with a given change point becomes

$$P(\mathbf{h} \mid \tau, H_1) = \prod_{i=1}^{n} P(h_i \mid \tau, H_1), \quad \text{(A2)}$$

where $\mathbf{h} = [h_1, h_2, \ldots, h_n]'$ is the vector form of the observation data.

Central to the Bayesian thinking is updating or revising knowledge about subjective probability assessments consistent with new information. This updating of knowledge involves both the prior and likelihood functions. When these two are combined, they yield a posterior distribution that represents the best information about the unknown parameter of interest conditional on the observed data. As a result, the posterior distribution function of the change-point τ under the hypothesis H_1 will be

$$P(\tau \mid \mathbf{h}, H_1) = \frac{P(\mathbf{h} \mid \tau, H_1)P(\tau \mid H_1)}{\sum_{\tau=2}^{n} P(\mathbf{h} \mid \tau, H_1)P(\tau \mid H_1)} \propto P(\mathbf{h} \mid \tau, H_1)P(\tau \mid H_1), \quad \text{(A3)}$$

where the prior knowledge of τ under H_1 hypothesis, $P(\tau \mid H_1)$, can be of any probability distribution function for discrete variables or probability density function for continuous variables. In the uninformative prior case, a proper choice for $P(\tau \mid H_1)$ could be the uniform distribution and (A3) is thus reduced to

$$P(\tau \mid \mathbf{h}, H_1) = \frac{P(\mathbf{h} \mid \tau, H_1)}{\sum_{\tau=2}^{n} P(\mathbf{h} \mid \tau, H_1)} \propto P(\mathbf{h} \mid \tau, H_1). \quad \text{(A4)}$$

Note that the denominator in both (A3) and (A4) is just a normalization factor.

Making use of (A2) and keeping in mind the conjugate property of gamma distribution, the conditional posterior distribution (with a given change-point τ) of the intensity before and after the change point, say λ_1 and λ_2, will also be a gamma distribution and the formula is

$$f(\lambda_1 | h'_1, T'_1, \tau, H_1, \mathbf{h}) = \frac{\tilde{T}_1^{\tilde{h}_1} \lambda_1^{\tilde{h}_1 - 1}}{\Gamma(\tilde{h}_1)} \exp(-\lambda_1 \tilde{T}_1),$$

$$f(\lambda_2 | h'_2, T'_2, \tau, H_1, \mathbf{h}) = \frac{\tilde{T}_2^{\tilde{h}_2} \lambda_2^{\tilde{h}_2 - 1}}{\Gamma(\tilde{h}_2)} \exp(-\lambda_2 \tilde{T}_2), \quad \text{(A5)}$$

where

$$\tilde{h}_1 = h'_1 + \sum_{i=1}^{\tau-1} h_i, \qquad \tilde{T}_1 = T'_1 + \tau - 1,$$

$$\tilde{h}_2 = h'_2 + \sum_{i=\tau}^{n} h_i, \qquad \tilde{T}_2 = T'_2 + (n - \tau + 1),$$

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

Finally, the posterior PDF for intensity λ_1 and λ_2 under H_1 hypothesis is

$$f(\lambda_i | h'_i, T'_i, H_1, \mathbf{h}) = \sum_{\tau=2}^{n} f(\lambda_i | h'_i, T'_i, \tau, H_1, \mathbf{h}) P(\tau | \mathbf{h}, H_1), \quad i = 1, 2. \quad \text{(A6)}$$

REFERENCES

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 J. Wang, 1998: Modeling return periods of tropical cyclone intensities in the vicinity of Hawaii. *J. Appl. Meteor.,* **37,** 951–960.

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.

Elsner, J. B., and C. P. Schmertmann, 1993: Improving extended-range seasonal predictions of intense Atlantic hurricane activity. *Wea. Forecasting,* **8,** 345–351.

——, and B. H. Bossak, 2001: Bayesian analysis of U.S. hurricane climate. *J. Climate,* **14,** 4341–4350.

——, T. Jagger, and X.-F. Niu, 2000: Changes in the rates of North Atlantic major hurricane activity during the 20th century. *Geophys. Res. Lett.,* **27,** 1743–1746.

——, 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.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin, 2004: *Bayesian Data Analysis.* 2d ed. Chapman & Hall/CRC, 668 pp.

Keim, B. D., and J. F. Cruise, 1998: A technique to measure trends in the frequency of discrete random events. *J. Climate,* **11,** 848–855.

Lavielle, M., and M. Labarbier, 2001: An application of MCMC methods for the 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.

Mayfield, M., and E. N. Rappaport, 1992: Eastern North Pacific hurricane season of 1991. *Mon. Wea. Rev.,* **120,** 2697–2708.

Raftery, A. E., 1996: Approximate Bayes factors and accounting for model uncertainty in generalized linear models. *Biometrika,* **83,** 251–266.

Shaw, S. L., 1981: A history of tropical cyclones in the central North Pacific and the Hawaiian Islands: 1832–1979. Central Pacific Hurricane Center Rep., NOAA/National Weather Service Forecast Office, Honolulu, HI, 137 pp.

Solow, A. R., 1988: A Bayesian approach to statistical inference about climate change. *J. Climate,* **1,** 512–521.

Tapsoba, D., M. Haché, L. Perreault, and B. Bobée, 2004: Bayesian rainfall variability analysis in West Africa along cross sections in space–time grid boxes. *J. Climate,* **17,** 1069–1082.

Wilson, R. M., 1999: Statistical aspects of major (intense) hurricanes in the Atlantic basin during the past 49 hurricane seasons (1950–1998): Implications for the current season. *Geophys. Res. Lett.,* **26,** 2957–2960.