Markov Chain Monte Carlo Change-Point Analysis: An Application in Hurricane Climatology

James B. Elsner
Department of Geography, The Florida State University, Tallahassee, FL

Corresponding author address:
Dept. of Geography, The Florida State University, Tallahassee, FL 32306
Tel: 850-644-8374, Fax: 850-644-5913
E-mail: jelsner@garnet.fsu.edu

Xufeng Niu
Department of Statistics, The Florida State University, Tallahassee, Florida

Thomas H. Jagger
Insightful Corporation Seattle, Washington

May 16, 2002

Prepared for the *Journal of Applied Meteorology*

Abstract

Annual hurricane counts are analyzed using a Markov chain Monte Carlo change-point model. The approach simulates marginal posterior distributions of the Poisson rate parameter using the Gibbs sampler algorithm. The paper is expository. Procedures are first employed on a recently debated time series of major North Atlantic hurricanes. Results confirm our earlier findings that show significant shifts in major North Atlantic hurricane rates during the middle 1940's, the middle 1960's, and 1995. The methodology is then applied to annual counts of U.S. hurricanes. Results are consistent with a homogeneous Poisson process showing no abrupt changes to overall coastal hurricane rates during the 20th century. In contrast, the Florida hurricane record indicates consecutive downward shifts first during the early 1950's and again during the late 1960's. No significant rate shifts are noted for Gulf or East coast hurricane activity. With a slight modification, the Gibbs sampler is then used to examine climate influences on coastal hurricane activity. Results show a significantly lower U.S. hurricane rate during El Niño events and during the negative phase of the North Atlantic oscillation. The ENSO effects are most pronounced over Florida while the NAO effects are most pronounced along the Gulf coast.

1. Introduction

Applied climate research relies heavily on singular value decomposition (SVD) methods for data analysis and prediction. Techniques include empirical orthogonal functions (EOFs), singular spectrum analysis (SSA), canonical correlation analysis (CCA), factor analysis (FA), among others. These tools identify the largest portion of the temporal and/or spatial variability (dominant modes) in the observations or model output. Typically the objective is to reduce the dimension of a data set by considering a projection of the original values onto the dominant modes of variability. The reduced data set is then further analyzed using one of a variety of tools, including regression analysis. SVD techniques are often employed to describe trends and oscillatory behavior in climate. In these cases the underlying physical model of the climate assumes stationarity, and the oscillations are considered regularly varying against a background of correlated noise.

Contemporary understanding of the nature of global climate processes, including the El Niño-Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO) among others, suggest that climate may operate in two or more quasi-stationary states (Lockwood 2001; Tsonis, et al. 1998; Tsonis and Elsner 1990; Berger and Labeyrie 1987). Transitions between different climate states may occur rather abruptly rather than slowly varying as a consequence of the dissipative, non-linear, and non-equilibrium properties of the climate system (Vannitsem and Nicolis 1991). Under this scenario successive climate shifts may result in an observable that appears to have a low frequency oscillation. A more precise description might be randomly occurring change-points. In this case SVD-type data analytic tools are less appropriate as they assume stationarity and regularly varying changes. In short, climate research methods need to be consistent with the underlying physical model of the climate process.

Change-point models are used to quantitatively identify and describe shifts to climate variables. This is important in the context of studying climate variability and change (Solow 1988), but it also has relevance in pinpointing potential inhomogeneities in climate records arising from improved observational technologies and changes in the station location. This has utility for the weather derivative market which relies on homogeneous records for estimating call and put options. Here we show that a statistical change-point model provides an alternative tool for data analytical climate studies; a tool which is consistent with a physical climate model supporting abrupt rather than slowly varying transitions. Change-point models are used to study climate variations, but their lack of widespread appeal might be related to the often ad hoc decisions necessary for their application. This limitation is less severe with a Markov chain Monte Carlo (MCMC) approach. For instance, Elsner et al. (2000a) use a log-linear regression approach to detect change points in the time series of annual counts of North Atlantic major hurricanes during the period 1900–99. A similar approach is employed by Chu (2002) in examining hurricanes that visit the central North Pacific. The assumption is the annual count of hurricanes or major hurricanes, after a logarithm transformation, is approximately normally distributed.

This distribution limitation disappears with a MCMC approach which can be applied directly to non-normal distributions.

MCMC change-point analysis has received considerable attention from statisticians and engineers. A thorough treatment of its utility for hydrological data is presented in Perreault et al. (2000a, b). Beyond this the broader climatological community has yet to make use of these approaches. Our purpose here is twofold: (1) To introduce to a wider meteorological audience some of the essential ideas behind this data analytic approach; and (2) To shed additional light on the problem of hurricane climate changes. It is important to stress that this work is not an argument that change-point models are superior to SVD techniques. Both are important in climatic data analysis and both play a role depending on the purpose at hand. The essential message here is that change-point modeling can provide new insights into climate variability not accessible with SVD methods, and with the MCMC approach, some of the subjective decisions typically associated with change-point models can be dispensed with in favor of easier interpretation. Our purpose is to illustrate the utility of the approach by applying a particular MCMC algorithm to the problem of detecting and quantifying shifts in the rates of coastal hurricane activity. In section 2 we outline the basic philosophy of the Bayesian approach to change-point modeling. In section 3 we apply the algorithm to the annual counts of major North Atlantic hurricanes. In section 4 the approach is applied to time sequences of coastal hurricane activity. In section 5 we show how the method can be used to examine covariate relationships. In particular we examine the influence of the El Niño-Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO) on coastal hurricane activity. Section 6 provides a summary and final comments.

2. The MCMC Approach

a. Statistical inference

The MCMC approach is rooted in a Bayesian perspective. To help focus on this perspective it is useful to consider first the general problem of statistical inference, which plays an important role in climate science. Given a sample of data (from the climate or its simulation), what conclusions can be made about the entire 'population?' Inference about the statistical model can be formalized as follows: Let θ be a population parameter, then statistical inference amounts to a supposition about θ on the basis of observing the data. We contend that values of θ which give high probabilities to our specific data y are more likely than those which assign y low probability (maximum likelihood principle). In essence the inferences are made by specifying a probability distribution of y, $f(y|\theta)$, for a given value of θ.

If we treat θ as a constant we are in the domain of classical statistical theory. The difference for Bayesian inference is that θ is treated as a random quantity, and our inference is based on $p(\theta|y)$, a probability distribution of θ for a given data y, rather than on $f(y|\theta)$. This seems quite natural as we are interested in the probability distribution of the parameter given the data, rather than the data given the parameter. The cost of this, more natural approach, is that it is necessary to specify a prior probability distribution, $\pi(\theta)$, which represents our beliefs about the distribution of θ before we have any information from the data. Thus, in the Bayesian approach we combine the likelihood distribution of the data given the parameter with the prior distribution to obtain:

$$p(\theta|y) = \frac{f(y|\theta)\pi(\theta)}{\int f(y|u)\pi(u)du},$$

which is called Bayes' theorem. Having observed y, Bayes' theorem is used to determine the distribution of θ conditional on y. This is called the *posterior* distribution of θ, and is the subject of all Bayesian inference.

Any feature of the posterior distribution is legitimate for inference, including moments, quantiles, p-values, etc. These quantities can be expressed in terms of the posterior expectations of functions of θ. The posterior expectation of a function $g(\theta)$ is

$$E[g(\theta)|y] = \frac{\int g(\theta)\pi(\theta)f(y|\theta)d\theta}{\int \pi(\theta)f(y|\theta)d\theta}.$$

The integrals in the above expression are a source of practical difficulties in Bayesian inference, especially for more complex problems. Moreover, in most applications, analytic evaluation of the expected value of the posterior density is impossible. Since numerical approximation methods are difficult to employ in a general way, Monte Carlo integration methods have become popular.

Monte Carlo integration evaluates $E[g(y)]$ by drawing samples $\{y_s,\ s = 1, \ldots, N\}$ from a probability density $p(.)$. An asymptotic approximation is given by

$$E[g(y)] \approx \frac{1}{N}\sum_{s=1}^{n} g(y_s).$$

Thus the population mean of $g(y)$ is estimated by a sample mean. An invariant posterior distribution is guaranteed because the Monte Carlo sampling produces an irreducible Markov chain. The key here is that N is controlled by the analyst; it is not the size of a fixed data sample (Gilks et al. 1996).

b. Gibbs sampler algorithm

A common MCMC procedure is the so-called "Gibbs sampler." Let $\vec{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)'$ be a p-dimensional vector of parameters and let $p(\vec{\theta}|y)$ be its posterior distribution given the data y. Then the Gibbs sampler is given as:

1. Choose an arbitrary starting point $\vec{\theta}^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \ldots, \theta_p^{(0)})'$, and set $i = 0$.

2. Generate $\vec{\theta}^{(i+1)} = (\theta_1^{(i+1)}, \theta_2^{(i+1)}, \ldots, \theta_p^{(i+1)})'$ as follows:

- Generate $\theta_1^{(i+1)} \sim p(\theta_1|\theta_2^{(i)}, \ldots, \theta_p^{(i)}, y)$;
- Generate $\theta_2^{(i+1)} \sim p(\theta_2|\theta_1^{(i+1)}, \theta_3^{(i)}, \ldots, \theta_p^{(i)}, y)$;

  ... ...

- Generate $\theta_p^{(i+1)} \sim p(\theta_p|\theta_1^{(i+1)}, \theta_2^{(i+1)}, \ldots, \theta_{p-1}^{(i+1)}, y)$.

3. Set $i = i + 1$, and go to Step 2.

In this way each component of θ is visited in order and a cycle through the scheme results in a sequence of p random numbers (Chen et al. 2000). Under general conditions the sequence of θs forms a Markov chain, and the stationary distribution of the chain is the posterior distribution. Typically, the chain is run for a large number of generations until the sample output is stable. A large number of additional generations are run, the output of which is analyzed as if it were a sample from the posterior distribution (Coles 1999).

3. North Atlantic Major Hurricane Activity

A hurricane is a tropical cyclone with maximum sustained (one-minute) 10 m winds of 33 ms^{-1} (65 kt) or greater. A major hurricane is one in which winds exceed 50 ms^{-1} (category 3 or higher on the Saffir/Simpson hurricane destruction potential scale). The long-term average number of major hurricane over the North Atlantic is close to two per year. Landsea et al. (1996) note a downward trend in the occurrence of these powerful hurricanes. In contrast, Wilson (1999) suggests a possible increase in activity beginning with 1995. The classical change-point model employed by Elsner et al. (2000a) shows indeed that 1995 is the start of the most recent epoch of greater major hurricane activity.

We begin our applications of the MCMC change-point model by revisiting this case. As in Elsner et al. (2000a) we consider 1900 as the first year of the record. It is understood that annual counts are likely biased prior to 1943 before the advent of aircraft reconnaissance, but the intention here is to identify shifts in the time series of annual counts regardless of their origin (natural or artificial). In fact one of the points made by Elsner et al. (2000a) is that, if the model is worthwhile it should detect a shift in activity during the middle 1940s. Landsea (1993) argue that an overestimation of hurricane intensity might have occurred even after 1943 during the period spanning the 1940s through the 1960s. Since there is still debate on this issue, and since corrections have yet to be made in the best-track data set, we do not consider the effect of this potential bias in the present study. In any event this decision does not influence the work presented here.

a. MCMC change-point algorithm

Given a time series of annual major hurricane counts, a change point occurs in the series if at some point t the values come from a distribution with a common rate up to that time and come from the same distribution but with a different rate afterward. The change-point separates the series into two parts, and we define the change-point year as the first year of the new epoch. Thus in a series of n annual counts, if a change is detected between year k and $k+1$, we say that $k+1$ is the change-point year.

An algorithm for detecting change points using a Markov chain Monte Carlo Gibbs sampler consists of two steps. Step one uses the entire record to determine *candidate* change points based on the expected value of the transition kernel of the Markov chain as a function of year. Higher mean values indicate candidate change-point years. A plot of the expected values as a function of year along with a 95th percentile line identifies the candidate years. Step two determines the posterior distributions of the relevant statistics before and after the candidate change point. The fraction of the posterior density of $\lambda_a - \lambda_b > 0$ (or $\lambda_a - \lambda_b < 0$) provides evidence against the hypothesis of no rate difference. From a frequentist perspective this amounts to a p-value against the null hypothesis of no change. As with all output associated with the Markov chain, the p-value is a random variable so additional runs are used to obtain ensemble averaged values.

For the present problem we are interested in two parameters; the hurricane rates before and after some change point. Let $\tilde{\lambda} = (\lambda_a, \lambda_b)$ be a vector of two parameters, where λ_a is the mean hurricane rate after the change and λ_b is the mean hurricane rate before the change and we wish to simulate from the posterior $f(\tilde{\lambda}|y)$ as described previously. More specifically, the data are counts (Y_i) of the annual number of major hurricanes observed over the North Atlantic each year during the period 1900–2001 ($n = 102$). Thus the model describes a Poisson number of hurricanes per year with a mean rate λ_b during years $i = 1, \ldots, k$ and a different mean rate λ_a during years $i = k+1, \ldots, n$. Formally the Poisson/Gamma model is given as:

$$Y_i \sim \text{Poisson}(\lambda_b); \quad i = 1, \ldots, k;$$

$$Y_i \sim \text{Poisson}(\lambda_a); \quad i = k+1, \ldots, n;$$

where $\lambda_b \sim \text{Gamma}(\alpha_1, \beta_1)$, $\lambda_a \sim \text{Gamma}(\alpha_2, \beta_2)$, k is discrete uniform over $\{1, \ldots, 101\}$, each independent, and $\beta_1 \sim \text{Gamma}(\gamma_1, \epsilon_1)$ and $\beta_2 \sim \text{Gamma}(\gamma_2, \epsilon_2)$ (Coles 1999). This specification leads to the following conditions:

$$\lambda_b|Y, \lambda_a, \beta_1, \beta_2, k \sim \text{Gamma}\left(\alpha_1 + \sum_{i=1}^{k} Y_i, k + \beta_1\right)$$

$$\lambda_a|Y, \lambda_b, \beta_1, \beta_2, k \sim \text{Gamma}\left(\alpha_2 + \sum_{i=k+1}^{n} Y_i, n - k + \beta_2\right)$$

$$\beta_1|Y, \lambda_b, \lambda_a, \beta_2, k \sim \text{Gamma}(\alpha_1 + \gamma_1, \lambda_b + \epsilon_1)$$

$$\beta_2|Y, \lambda_b, \lambda_a, \beta_1, k \sim \text{Gamma}(\alpha_2 + \gamma_2, \lambda_a + \epsilon_2)$$

and

$$p(k|Y, \lambda_b, \lambda_a, \beta_1, \beta_2) = \frac{L(Y; k, \lambda_b, \lambda_a)}{\sum_{j=1}^{n} L(Y; j, \lambda_b, \lambda_a)},$$

where the likelihood function is

$$L(Y; k, \lambda_b, \lambda_a) = \exp\{k(\lambda_a - \lambda_b)\}(\lambda_b/\lambda_a)^{\sum_{i=1}^{k} Y_i}.$$

Starting with some initial (prior) values for α_1, α_2, γ_1, γ_2, ϵ_1, and ϵ_2 the Gibbs sampler generates sequences of λ_a and λ_b which form Markov chains. The stationary distributions of the chains are the posterior distributions for each of the parameters.

b. Practical considerations

There are several practical issues that need to be addressed. First is the definition of years relative to the suspected change point. In the above model specification year k is the last year of the old epoch with $k+1$ the first year of the new epoch. To be consistent our earlier results (Elsner et al. 2000a) we plot the change-point as the first year of the new epoch and refer to this year as the change-point year.

Second is the choice of starting (or initial) values. In theory if the chain is irreducible meaning that it can reach any non-empty set with positive probability, then the choice of initial values will not influence the final stationary (invariant) posterior distribution. Since the Poisson rate parameter is Gamma(α, β) with mean α/β and variance α/β^2, we choose $\alpha_1 = \alpha_2 = 0.3$ and $\gamma_1 = \gamma_2 = 0.1$, and $\epsilon_1 = \epsilon_2 = 1$ as our starting values. Thus the mean values for β_1 and β_2 are $0.1/1 = 0.1$ and the mean values for λ_a and λ_b are $0.3/0.1 = 3$, which is close to the average annual number of major hurricanes per year. In practice it is useful to perform a number of simulations with different starting values to check if the posterior distribution is sensitive to the choice of initial values. Results from these simulations are given in the next section.

Third is the issue of chain length and burn-in. If the chain is irreducible, aperiodic, and positive recurrent then it will converge to a stationary posterior distribution (Roberts 1996). In practice the chain is run for a large number of iterations until the sample output is stable with the first hundred or so iterations discarded as "burn-in" and the remain values considered samples from the stationary distribution. The length of burn-in depends on the initial values and the rate of convergence, which is related to how fast the chain mixes. Developing rigorous criteria for deciding chain length and burn-in requires a detailed study of the convergence properties of the chain (Jones and Hobert 2001) that is beyond the scope of the present work. Trial-and-error using visual inspection of the chain's output is a commonly used method for determining length of burn-in, and it is the one adopted here. Using the above prescribed initial values, the Gibbs sampler change-point algorithm is run on the annual counts of major North Atlantic hurricanes (1900–2001) with values of β_1 and β_2 plotted for each iteration (Figure 1). Convergence is quick.

The distribution of values for both β_1 and β_2 do not appear to change as the chain is run for a greater number of iterations. This is typical. However, it is still good practice to remove the early iterates to allow the chain to 'forget' its starting position. Throughout the present work we choose a burn-in of 50 iterations and estimate the posterior distribution from the next 1000 iterations.

Fourth is the issue of statistical significance. Even if the time series of annual counts comes from a homogeneous Poisson process implying no rate shifts, a sequence of random values generated from such a process might result in a series that has one or more change points. Thus we need to compare the results of the analysis on our observed time series against the null hypothesis of no change point. To do this, we generate a set of 1000 surrogate time series from a homogeneous Poisson process with the rate equal to the average number hurricanes observed over the 102-year period. We then run the Gibbs sampler as before (1050 iterations discarding the first 50 as burn-in) on each of the surrogate series. The confidence level associated with each of the individual years being a change-point must be adjusted upward to produce a simultaneous confidence level across the entire period. It makes little sense if each year is the start of a new epoch as then each epoch would last only a single year. Realistically we assume that an epoch lasts a decade so that the confidence level of $m = n/10$ individual epochs are adjusted upward to the 99.5th percentile. This procedure provides a simultaneous 95% confidence level as an application of the Bonferroni inequality in probability theory. The simultaneous confidence interval is a baseline for identifying candidate change-point years. Note we could sample the assumed homogeneous Poisson rate and then use the rate to generate random Poisson samples. However, the annual probability of a change point depends more on record length then on variation in the rates. In fact, for this study we use a single Bonferroni confidence level irrespective of the rates. The confidence level is re-estimated when we consider a shorter record.

c. Results

Figure 2 shows the results from the MCMC change-point algorithm applied to the annual counts of major North Atlantic hurricanes during the period 1900–2001. The expected value of the transition kernel of the Markov chain p is plotted as a function of year. The expected value represents the average posterior probability of the year being the first year of a new epoch. Large probabilities indicate a change occurred with year t. The dashed line shows the 95% simultaneous confidence line against the null hypothesis of no change point generated from 10^3 simulations of a homogeneous Poisson process with rate parameter equal to the mean number of major hurricanes over the period 1900–2001. The line is smoothed using a 5-year normal kernel smoother. Years with probabilities above the 95% simultaneous confidence line include 1906, 1943, and 1995. Note that several years around 1943 are also candidate change-point years. This indicates that although the algorithm chooses 1943 as the most likely year of the new epoch there is statistical uncertainty associated with whether the new epoch begins with 1943 or 1944. This is not the case with 1995 or 1906 where no other years appear to be in contention. Also note that years near the beginning and end of the record require substantially larger posterior probabilities to surpass the nominal significance level. The U-shaped confidence level indicates that there is a greater uncertainty (larger variance on the posterior probabilities) about candidate change-points close to the record end points. Caution is warranted when interpreting large probabilities on these years as the chance of a false detection is greater. Ensemble runs as discussed below can help in this regard.

As previously mentioned it is important to examine the influence the choice of initial values has on the average posterior probability and thus the selection of candidate years. This is done here by running the Gibbs sampler algorithm 30 times for each value of the α priors equal to 0.1, 0.3, 0.5, and 0.7. The algorithm is run using 1050 iterations with the first 50 discarded as burn-in. Figure 3 shows the distribution of probabilities for the candidate years using box and whisker plots. Overall the results demonstrate that the choice of prior values is not a critical factor in identifying candidate years as the chain quickly finds a stationary distribution regardless of where it is started. Variability in the average posterior probabilities is largest for candidate years 1906 and 1995. As noted above, this results from the fact that these years are near the beginning and end of the time series. Posterior probability distributions based on relatively few data points will have a greater spread.

The MCMC change-point algorithm continues by checking the significance of the candidate change-point years. We look first at 1943 since the jump in annual major hurricane counts at this time is most likely due to the use of aircraft reconnaissance investigations (Neumann et al. 1999; Jarvinen et al. 1984). Posterior density estimates of the relevant statistics (β_1, β_2, λ_a, λ_b, and $\lambda_a - \lambda_b$) from the Gibbs sampler are shown in Fig. 4. We focus on the probability densities of the annual hurricane rate parameters before and after (including) 1943. Densities are smoothed versions of the histograms and are based here on a normal kernel with bandwidth equal to 4 times the standard deviation of the values (Venables and Ripley 1999).

The Gibbs sampler is run 30 times to get an ensemble average of the mean Poisson rate before $\langle\bar{\lambda}_b\rangle$ and after $\langle\bar{\lambda}_a\rangle$ the change point. The ensemble average of the mean rate parameter is 1.51 before 1943 and 2.51 thereafter. The posterior densities of the rate parameters indicate little overlap implying a significant rate increase beginning with the 1943 season. This is examined directly by considering the posterior density of the rate differences. Only a negligible fraction of the posterior distribution of $\lambda_a - \lambda_b$ is less than zero. The number of differences less than zero provides a one-sided p-value as evidence against the hypothesis of equal rates before and after 1943. The ensemble average p-value is less than 0.001. There is convincing evidence of a rate difference. As stated above the increase in observed activity starting in the middle 1940s is likely due in part to the start of aircraft reconnaissance. Therefore we continue the investigation of major hurricane activity by examining the record only from 1943 onward.

Figure 5 shows the results from the MCMC change-point algorithm applied to the annual counts of major North Atlantic hurricanes during the period 1943–2001. Both 1965 and 1995 are years with high probabilities. Additional high probability years clustering around 1965 include 1962, 1966, and 1967. Estimates of the posterior densities relative to 1965 are shown in Fig. 6. As was done before the Gibbs sampler is run 30 times to get an ensemble average of the mean Poisson rate before $\langle\bar{\lambda}_b\rangle$ and after $\langle\bar{\lambda}_a\rangle$ the change point. The ensemble averaged mean rate is 3.41 before 1965 and 1.97 thereafter. The posterior densities of the rates indicate little overlap implying a significant decrease in activity beginning with the 1965 season. The ensemble p-value against the null hypothesis of no rate change is less than 0.001. Similar results are obtained for 1962, 1966, and 1967 indicating that the decline in abundance of major North Atlantic hurricanes might have begun as early as 1962 or as late as 1967 with the most likely year being 1965.

Next we consider 1995. Estimates of the posterior densities are shown in Fig. 7. The ensemble averaged mean hurricane rate before 1995 is 1.98 and 3.57 thereafter. The density for the hurricane rate since 1995 (λ_a) is considerably flatter owing to the relatively few years of data in the record following this year (7). The greater uncertainty about the annual rate at the end of the hurricane record creates more overlap on the rate distributions and thus a larger p-value on the rate difference. Even still, evidence is convincing that 1995 represents an upward shift in hurricane activity.

Thus a picture emerges of significant quantifiable shifts in the frequency of major North Atlantic hurricanes during the 20th century. The results for the middle 1940's, 1965, and 1995 are consistent with results obtained using a non-probabilistic change-point model (Elsner et al. 2000a). Beginning with the era of aircraft surveillance, we see that major hurricanes occurred at an average annual rate of nearly 3.5 per year. The rate dropped significantly to about 2 major hurricanes per year beginning sometime during the middle 1960's with the new epoch most likely starting with the 1965 season. This modern era of fewer major hurricanes ends abruptly with the 1995 season. For the next seven seasons through 2001 the mean rate is more than 3.5 hurricanes per year. The advantage of the probabilistic model is that it provides natural uncertainty estimates on the rates before and after the change-point. It also provides ensemble estimates of rate differences and p-values. We now turn our attention to a systematic examination of change points in U.S. hurricane activity during the 20th century.

4. U.S. Hurricane Activity

In the previous section we demonstrate the change-point algorithm by applying it to major hurricanes activity over the entire North Atlantic basin. Here we apply the methodology to records of U.S. hurricane activity. We are unaware of change-point studies on these records. Hurricane landfall occurs when all or part of the eye wall (the central ring of deep atmospheric convection, heavy rainfall, and strong winds) passes directly over the coast or adjacent barrier island. A U.S. hurricane is a hurricane that makes at least one landfall. A reliable list of the annual counts of U.S. hurricanes back to 1900 is available from the U.S. National Oceanic and Atmospheric Administration (Neumann et al. 1999). These data represent a blend of historical archives and modern direct measurements. An updated climatology of annual coastal hurricane activity is given in Elsner and Kara (1999) and Elsner and Bossak (2001).

a. Overall activity

We consider first overall U.S. hurricane activity. The annual time series of U.S. hurricane counts appears to be stationary over the period (Elsner and Kara 1999). The lag-one autocorrelation is a negligible -0.02. Figure 8 shows the probability of each year being a change-point in the series. In contrast to the posterior probabilities computed above from the series of annual major hurricane counts, the probabilities computed based on counts of U.S. hurricanes are considerably lower and all below the simultaneous confidence limit (95%) estimated from a homogeneous Poisson process. Notice that no hurricanes reached the U.S. coast during 2000 and 2001, so the algorithm hints at a possible change-point following the 1999 season. The evidence however is not strong as there are other two-year periods without hurricanes (1930–31 and the more recent 1981–82). Thus, somewhat surprisingly, the shifts in overall major hurricane activity noted in the previous section are not reflected in overall landfall rates in the United States.

Recent studies have shown inter-annual to decadal changes to the spatial patterns of U.S. hurricane activity related to large-scale climate factors (Elsner et al. 2000b). For instance, in La Niña years during which the North Atlantic oscillation is weak, the probability of a hurricane strike to the central Gulf coast increases significantly (Jagger et al. 2001; Saunders et al. 2000). It is therefore instructive to consider regional coastal hurricane activity. We divide the coast into three zones; Gulf coast, Florida, and East coast and consider the possibility of rate changes over each. Florida, with its 2171 km of coastline, leads the United States in frequency of hurricanes. The Gulf coast is defined as the region from Texas to Alabama, while the East coast is defined as the region from Georgia to Maine. Clearly, other divisions are possible.

b. Regional activity

Figure 9 shows the average posterior probability of each year being a change-point for the records of Gulf coast, Florida, and East coast hurricanes. The hurricane rates along the Gulf and East coasts appear to be rather stable over the 102-year period. No probabilities extend above the simultaneous confidence limit estimated from a constant-rate Poisson process. Thus, as with overall coastal hurricane activity, we find no significant shifts in the rates of Gulf or East coast hurricanes. The situation is different in Florida where there is evidence of rate shifts during the early 1950's and again during the late 1960's.

The Gibbs sampler is run 30 times to get an ensemble average of the mean Poisson rate before and after 1952. The ensemble average of the mean rate parameter is 0.83 before 1952 and

0.48 thereafter indicating a decrease in Florida hurricanes beginning in the early 1950's. The difference is significantly offset from zero with an ensemble p-value of 0.015. Surprisingly, the shift at 1969 is also downward with a mean rate of 0.79 before and 0.40 thereafter. Here the difference is significantly offset from zero with an ensemble p-value of 0.008. Posterior density estimates of the rate parameters and their differences are shown for both candidate change points in Fig. 10. The densities indicate two consecutive downward shifts in Florida hurricane activity during the later half of the 20th century.

5. ENSO and the NAO

Importantly, the Gibbs sampler can also be used to examine the effect of covariates on hurricane activity. In particular the influence of the El Niño/Southern Oscillation (ENSO) and the North Atlantic oscillation (NAO) on annual coastal hurricane numbers is of interest and can be examined with a slight modification to the Gibbs sampler algorithm. Here it is assumed that each year is independent which is reasonable for annual hurricane counts. The statistical relationship between ENSO and U.S. hurricanes is well-known (Bove et al. 1998; Elsner et al. 1999; Elsner and Kara 1999; Jagger et al. 2001), but the relationship between NAO and U.S. hurricanes is less well recognized (Elsner et al. 2000b; Elsner et al. 2001).

A reliable time record of the Pacific ENSO is obtained by using basin-scale equatorial fluctuations of sea surface temperatures (SST). Average SST anomalies over the region bounded by 6°N to 6°S latitude and 90°W to 180°W longitude are called the "cold tongue index" (CTI) (Deser and Wallace 1990). Values of CTI are obtained from the *Joint Institute for the Study of the Atmosphere and the Oceans* as monthly anomalies (base period: 1950–79) in hundredths of a degree Celsius. Monthly values of the CTI are strongly correlated with values from other ENSO SST indices. Since the Atlantic hurricane season runs principally from August through October, a 3-month averaged (Aug–Oct) CTI from the data set is used. Values of an index for the NAO are calculated from sea level pressures at Gibraltar and at a station over southwest Iceland (Jones et al. 1997), and are obtained from the *Climatic Research Unit.* The values are first averaged over the pre- and early-hurricane season months of May and June. This is a compromise between signal strength and timing relative to the hurricane season. The signal-to-noise ratio in the NAO is largest during the boreal winter and spring, whereas the U.S. hurricane season begins in June (see Elsner et al. 2001).

For both the ENSO and NAO we divide the range of values occurring over the 102-year period into equal interval terciles describing below, normal, and above normal years. The upper and lower tercile values of the Aug–Oct average CTI are 0.90 and −0.23°C, respectively. The upper and lower tercile values of the May-Jun average NAO are 1.05 and −0.85 s.d., respectively. Years of above (below) normal CTI correspond to El Niño (La Niña) events. We remove the group of normal years and compare the hurricane rates for years of above and below normal climate conditions. In this way there are 14 (39) above (below) normal ENSO years and 11 (34) above (below) normal NAO years during the 20th century.

Figure 11 shows the posterior densities of the rate differences (above normal years minus below normal years) using the Gibbs sampler. As anticipated we see that during El Niño years (above normal) the annual rate is significantly less than the rate during La Niña years (below normal). A 30-member ensemble gives an average rate of 0.72 hurricanes/yr during El Niño years compared with 2.18 hurricanes/yr during La Niña years. This difference results in a p-value that is less than 0.001. The influence of El Niño appears all along the coast, but is strongest over Florida which has a mean rate of 0.37 hurricanes/yr during El Niño compared with 0.93 hurricanes/yr during La Niña. This difference corresponds to a p-value of 0.011. Figure 11 also shows the effect of NAO on U.S. hurricanes. During its strong phase (above normal), the ensemble average rate is 1.02 hurricanes/yr compared with 2.21 hurricanes/yr during its weak (or negative) phase (below normal). This provides a p-value of 0.003. Unlike the influence of ENSO which is felt all along the coast, the influence of the NAO is only significant along the Gulf coast. Here the annual rate is 0.38 hurricanes/yr during the NAO strong phase and 0.86 during the NAO weak phase. This is consistent with the hypothesis that the NAO is linked to hurricane steering mechanisms (Elsner et al. 2000b, Elsner et al. 2001).

6. Summary and Comments

This paper demonstrates an application of a general statistical framework for determining sudden changes at unknown times in climatological records involving counts. The presentation is expository. The approach is rooted in Bayesian theory. The iterative Monte Carlo method, known as the Gibbs sampler, produces a Markov chain, the output of which corresponds to a (correlated) sample from the joint posterior distribution. Our purpose is to illustrate the utility of the approach by applying it to the problem of detecting and quantifying shifts in the rates of coastal hurricane activity.

The procedure is first applied to annual counts of major North Atlantic hurricanes. Results are consistent with those generated from a classical change-point model (Elsner et al. 2000a) including an ominous rate increase starting in 1995. When the algorithm is applied to annual counts of overall U.S. hurricane activity there is little evidence for significant rate changes during the 20th century. Grouping counts by region including the Gulf coast, Florida, and the East coast and applying the algorithm to each region separately indicates significant decreases in the number of Florida hurricanes during the early 1950's and again during the late 1960's. The statistically significant and consecutive decreases in Florida hurricane activity occur over a period of substantial growth in the state's population.

The algorithm is also used to study the influence of ENSO and the NAO on coastal hurricane activity. Climate data representing these two modes of variability are divided into terciles representing above normal, normal, and below normal conditions. As expected from previous studies, we find a statistically significant linkage to the ENSO. During El Niño years coastal hurricane rates are reduced from Texas to Maine. The most pronounced effect occurs over Florida. The NAO also plays a role. During years in which the NAO index is below normal more than twice as many hurricanes reach the coast on average. However, unlike the ENSO's influence which is felt all along the coast, NAO's influence is significant only for the Gulf coast from Texas to Alabama.

Given its utility and ease of application, MCMC change-point algorithms should become a standard research tool in climate-related sciences.

Acknowledgements. Partial support for this study was provided by the National Science Foundation, the Risk Prediction Initiative, and by the Florida State University's Cornerstone Program. The views expressed within are those of the author's and do not reflect those of the funding agencies.

References

- Berger, W.H., and L.D. Labeyrie, 1987: Abrupt climatic change, evidence and implications. NATO ASI Series C, Mathematical and Physical Sciences; 216.

- Bove, M.C., J.B. Elsner, C.W. Landsea, X. Niu, and J.J. O'Brien, 1998: Effect of El Niño on U.S. landfalling hurricanes, revisited. *Bull. Amer. Meteor. Soc.*, **79**, 2477–2482.

- Chen, M.-H., Q.-M. Shao, and J.G. Ibrahim, 2000: *Monte Carlo Methods in Bayesian Computation.* Springer, New York, 386pp.

- Chu, P.-S., 2002: Large-scale circulation features associated with decadal variations of tropical cyclone activity over the central North Pacific. *J. Climate*, in press.

- Coles, S., 1999: Course Notes 351–Statistical Inference: (Bayesian Inference). June 10, 1999 [Available from http://www.stats.bris.ac.uk/∼masgc/].

- Deser, C., and J.M. Wallace, 1990: Large-scale atmospheric circulation features of warm and cold episodes in the tropical Pacific. *J. Climate*, **3**, 1254–1281.

- Elsner, J.B., and A.B. Kara, 1999: *Hurricanes of the North Atlantic: Climate and Society.* Oxford University Press, New York, 488 pp.

- Elsner, J.B., A.B. Kara, and M.A. Owens, 1999: Fluctuations in North Atlantic hurricanes. *J. Climate*, **12**, 427–437.

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

- Elsner J.B., K.-b. Liu, and B. Kocher, 2000b: Spatial variations in major U.S. hurricane activity: Statistics and a physical mechanism. *J. Climate*, **13**, 2293–2305.

- Elsner, J.B., B.H. Bossak, and X.-F. Niu, 2001: Secular changes to the ENSO-U.S. hurricane relationship. *Geophys. Res. Lett.*, **28**, 4123–4126.

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

- Gilks, W.R., S. Richardson, and D.J. Spiegelhalter, 1996: *Markov Chain Monte Carlo in Practice*, Chapman & Hall/CRC, Boca Raton, 486pp.

- Jagger, T.H., J. B. Elsner, and X. F. Niu, 2001: A dynamic probability model of hurricane winds in coastal counties of the United States. *J. Appl. Meteor.*, **40**, 853–863.

- Jarvinen, B.R., C.J. Neumann, and M.A.S. Davis, 1984: A tropical cyclone data tape for the North Atlantic Basin, 1886–1883: Contents, limitations, and uses. *NOAA Tech. Memo.*, NWS NHC-22, Coral Gables, FL, 21pp.

- Jones, G.L., and J.P. Hobert, 2001: Honest exploration of intractable probability distributions via Markov chain Monte Carlo. *Statistical Science*, **4**, 312–324.

- Jones, P.D., T. Jónsson, and D. Wheeler, 1997: Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and South-West Iceland. *Int. J. Climatol.* **17**, 1433–1450.

- Landsea, C.W., 1993: A climatology of intense (or major) Atlantic hurricanes. *Mon. Wea. Rev.*, **121**, 1703–1713.

- Landsea, C.W., N. Nicholls, W.M. Gray, and L.A. Avila, 1996: Downward trends in the frequency of intense Atlantic hurricanes during the past five decades. *Geophys. Res. Lett.*, **23**, 1697–1700.

- Lockwood, J.G., 2001: Abrupt and sudden climatic transitions and fluctuations: A review. *Int. J. Climatol.*, **21**, 1153–1179.

- Neumann, C.J., B.R. Jarvinen, C.J. McAdie, and G.R. Hammer, 1999: *Tropical Cyclones of the North Atlantic Ocean, 1871–1998.* Historical Climatological Series 6-2, Prepared by the National Climatic Data Center, Asheville, NC, in cooperation with the Tropical Prediction Center/NHC, Miami, FL, 206 pp.

- Perreault, L., J. Bernier, B. Bobée, and E. Parent, 2000a: Bayesian change-point analysis in hydrometeorological time series. Part 1. The normal model revisited. *J. Hydrology*, **235**, 221–241.

- Perreault, L., J. Bernier, B. Bobée, and E. Parent, 2000b: Bayesian change-point analysis in hydrometeorological time series. Part 2. Comparison of change-point models and forecasting. *J. Hydrology*, **235**, 242–263.

- Roberts, G.O., 1996: Markov chain concepts related to sampling algorithms. *Markov Chain Monte Carlo in Practice*, W.R. Gilks, S. Richardson, and D.J. Spiegelhalter, Eds., Chapman & Hall/CRC, London, 45–58.

- Saunders, M.A., R.E. Chandler, C.J. Merchant, and F.P. Roberts, 2000: Atlantic hurricanes and NW Pacific typhoons: ENSO spatial impacts on occurrence and landfall. *Geophys. Res. Lett.*, **27**, 1147–1150.

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

- Tsonis, A.A., P.J. Roebber, and J.B. Elsner, 1998: A characteristic time scale in the global temperature record, *Geophys. Res. Lett.*, **25**, 2821–2823.
- Tsonis, A.A., and J.B. Elsner, 1990: Multiple attractors, fractal basins and longterm climate dynamics, *Atmos. Physics*, **63**, 171–176.
- Vannitsem, S., and C. Nicolis, 1991: Detecting climatic transitions: Statistical and dynamical aspects. *Atmos. Physics*, **64**, 245–254.
- Venables, W.N., and B.D. Ripley, 1999: *Modern Applied Statistics with S-PLUS.* Springer, New York, 501 pp.
- 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.

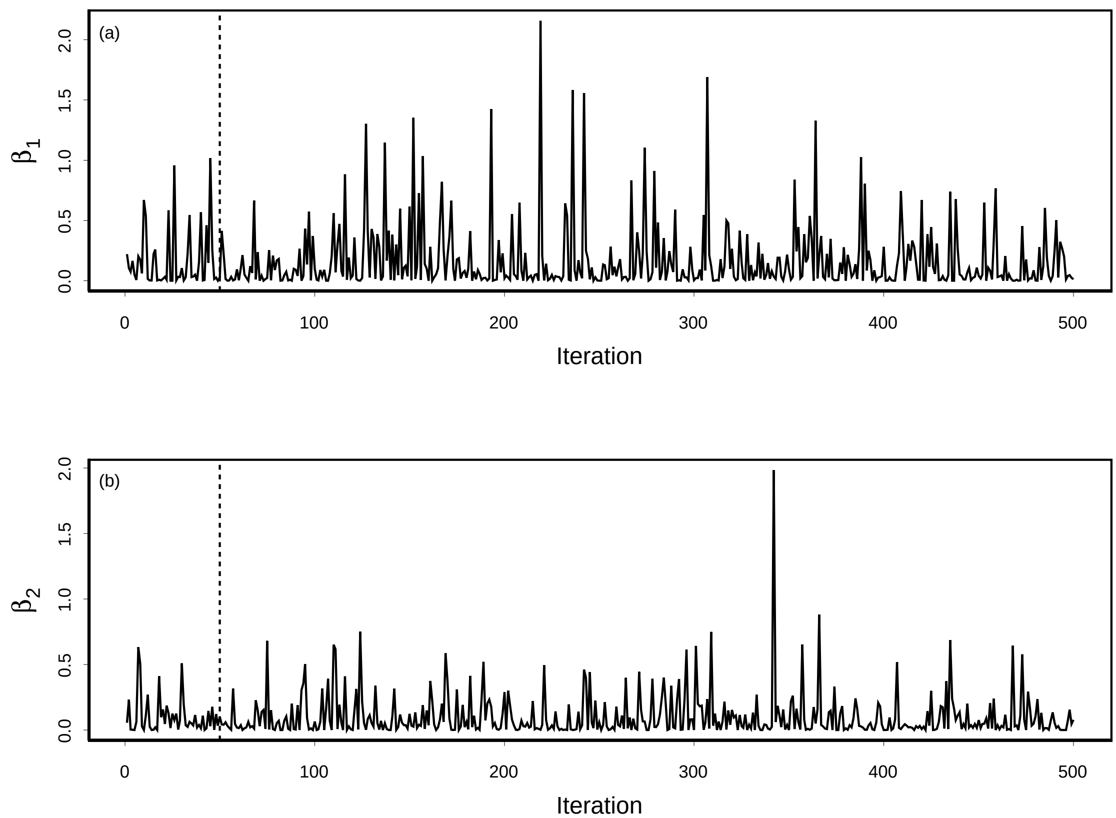

Figure 1: The first 500 iterations of the Gibbs sampler change-point algorithm applied to the annual counts of North Atlantic hurricanes over the period 1900–2001. Initial values are $\alpha_1 = \alpha_2 = 0.3$ and $\gamma_1 = \gamma_2 = 0.1$, and $\epsilon_1 = \epsilon_2 = 1$. The first 50 iterations (left of the dashed vertical line) are discarded as burn-in.

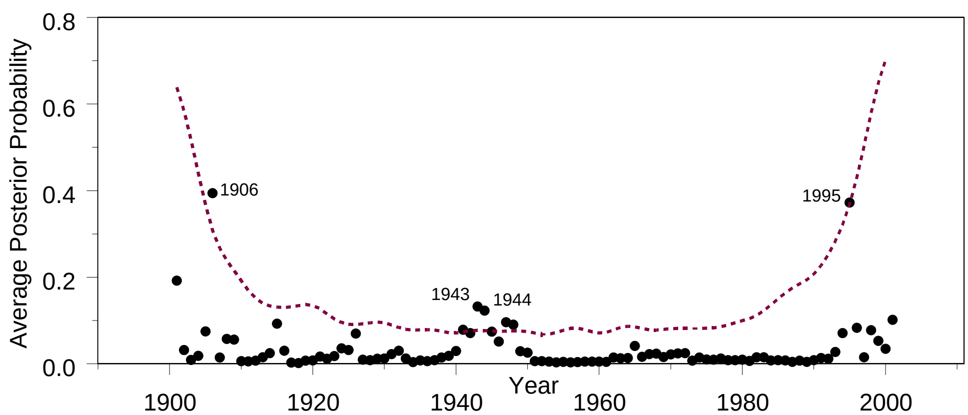

Figure 2: Average posterior probabilities of each year being the first year of a new epoch. Large probabilities on year t indicate a change likely occurred with t as the first year of the new epoch. The dashed line represents the 95% simultaneous confidence level based on 1000 simulations of a constant rate (homogeneous) Poisson process. In general years with points lying above this line are considered candidate change-point years.

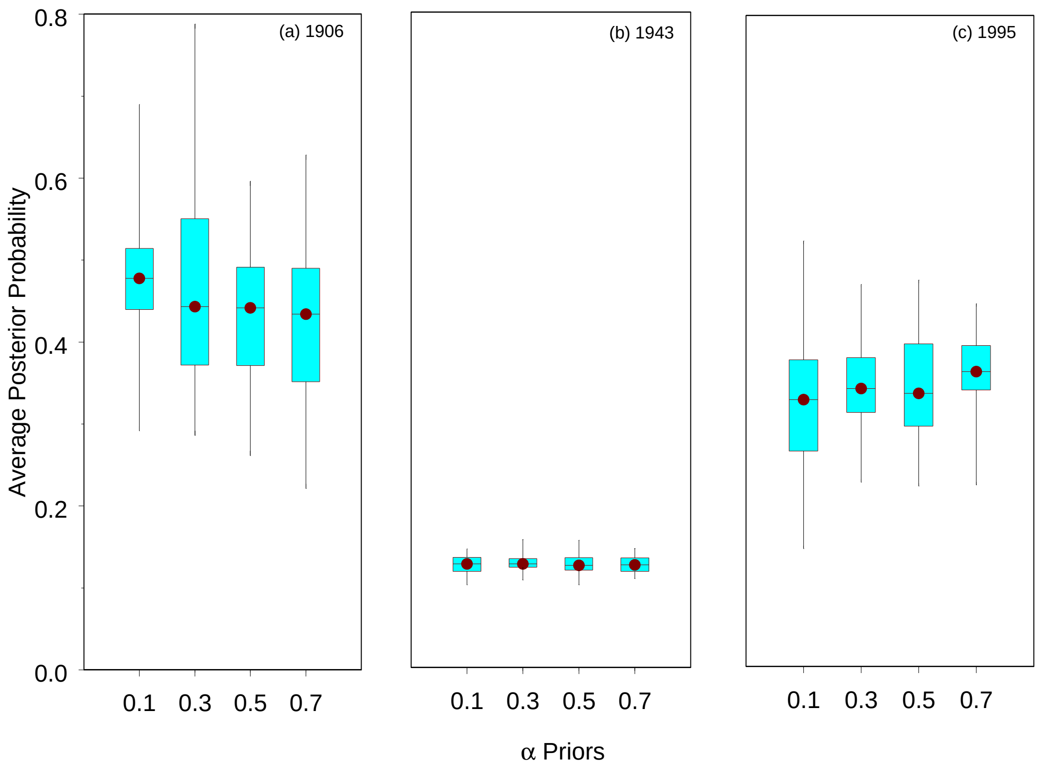

Figure 3: Distributions of the average posterior probabilities that (a) 1906, (b) 1943, and (c) 1995 are change-point years. Distributions based on 30 simulations are shown with a box and whisker plot where the quartiles are represented by the top and bottom of the box and the median is shown with a dot and horizontal line inside the box. Whiskers extend from the box to the largest and smallest value in the distribution. The distributions are based on 30 independent simulations.

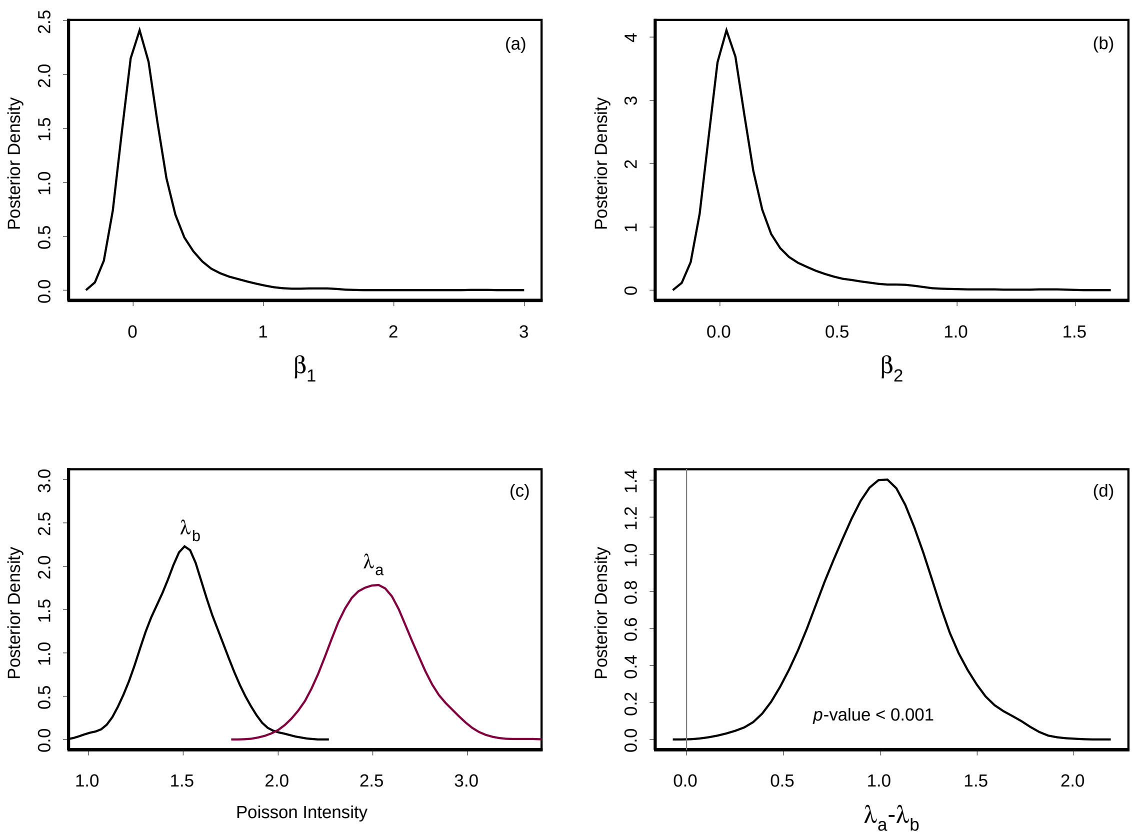

Figure 4: Estimates of the posterior densities from the Gibbs sampler applied to the time series of major hurricane counts. Density of (a) β_1, (b) β_2, (c) λ_b (1900–1942) and λ_a (1943–2001), and (d) $\lambda_a - \lambda_b$. The ensemble average Poisson rate before 1943 $\langle\bar{\lambda}_b\rangle = 1.51$ and the ensemble average rate after (and including) 1943 $\langle\bar{\lambda}_a\rangle = 2.51$. This provides an ensemble average p-value on the rate difference that is less than 0.001, where the p-value is based on the fraction of values from the posterior density of $\lambda_a - \lambda_b$ that are less than zero.

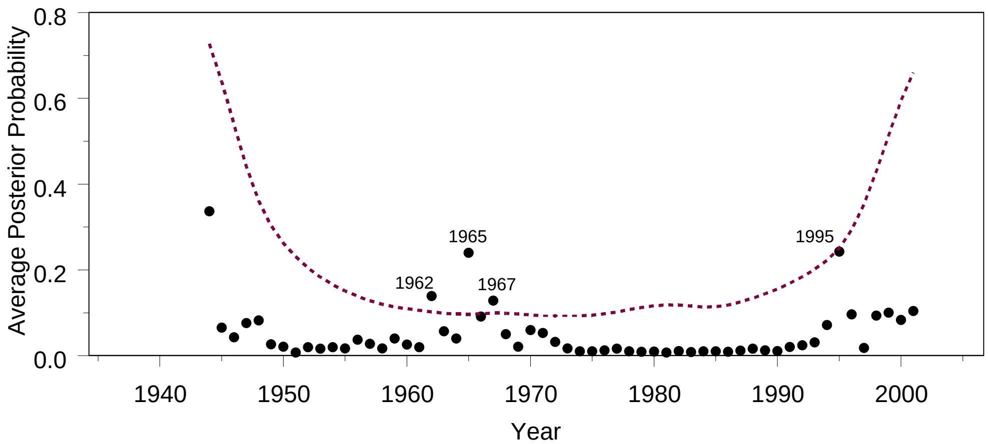

Figure 5: Same as Fig. 2 expect the years prior to 1943 are excluded from the analysis.

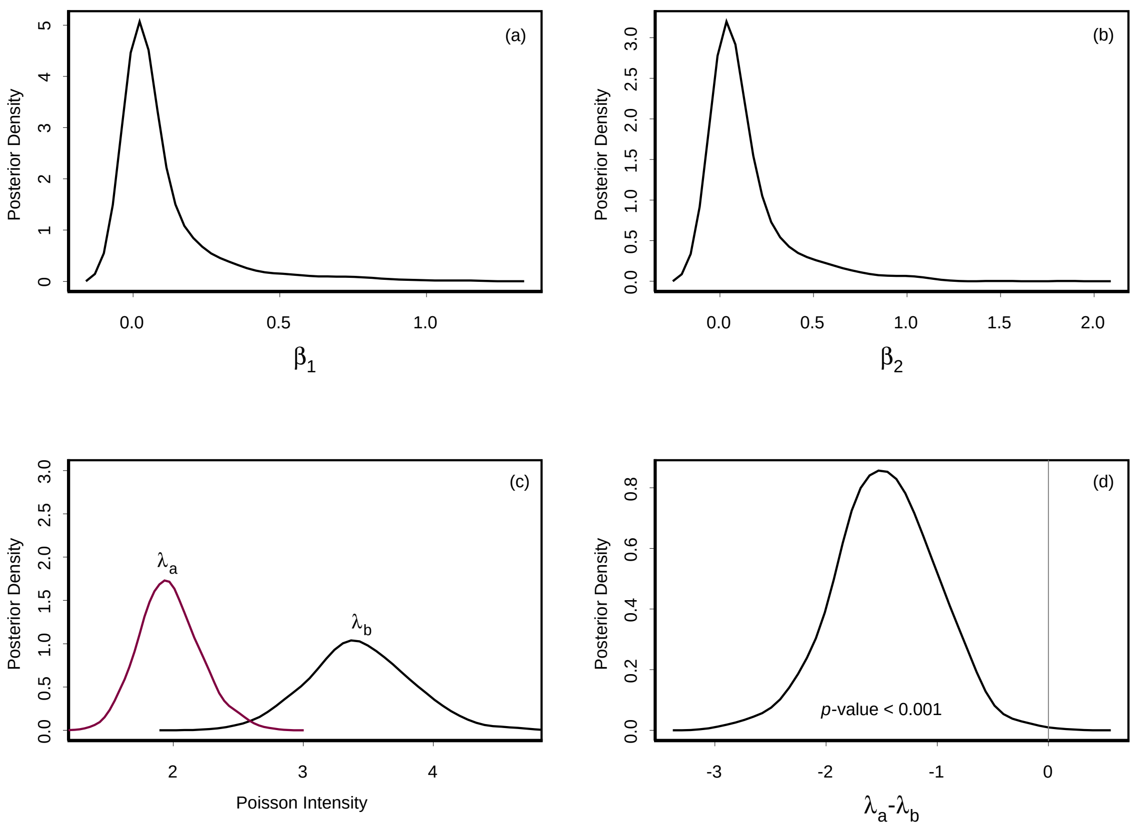

Figure 6: Estimates of the posterior densities from the Gibbs sampler applied to the time series of major hurricane counts. Density of (a) β_1, (b) β_2, (c) λ_b (1943–1964) and λ_a (1965–2001), and (d) $\lambda_a - \lambda_b$. The ensemble average Poisson rate before 1965 $\langle \bar{\lambda}_b \rangle = 3.41$ and the ensemble average rate after (and including) 1965 $\langle \bar{\lambda}_a \rangle = 1.97$. This provides an ensemble average p-value on the rate decrease of less than 0.001.

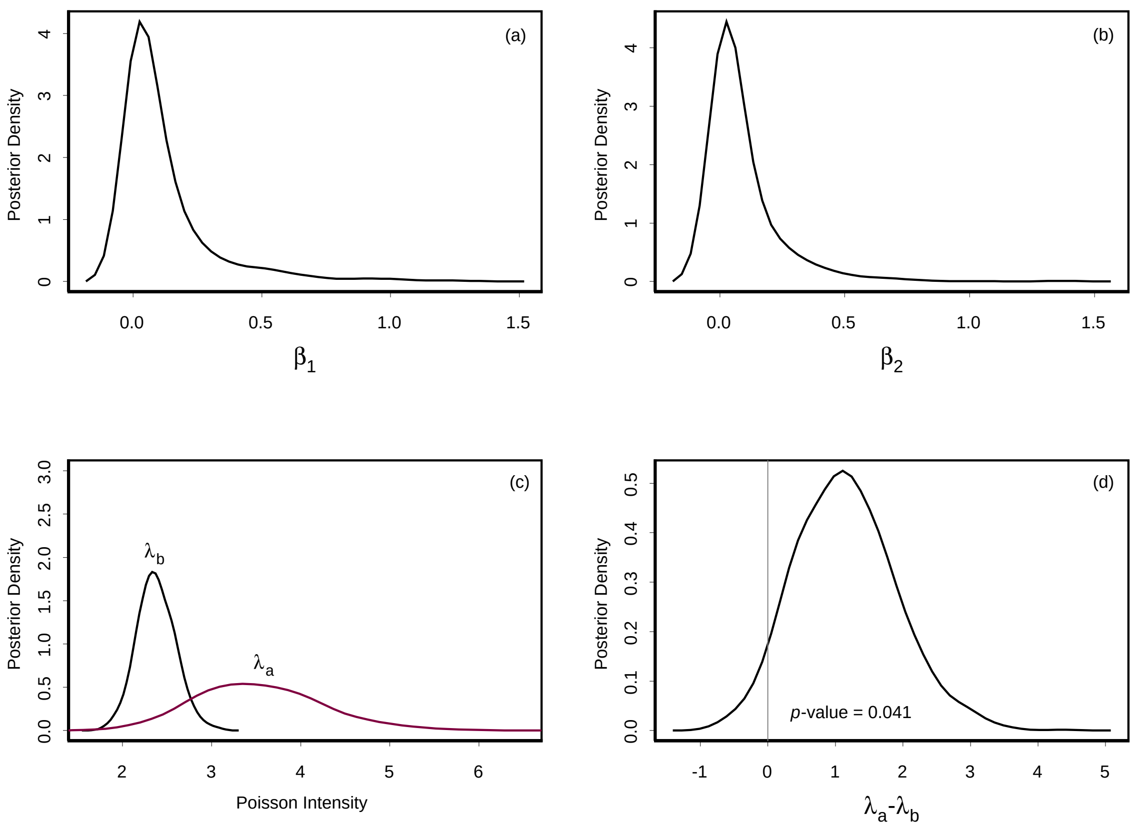

Figure 7: Estimates of the posterior densities from the Gibbs sampler applied to the time series of major hurricane counts. Density of (a) β_1, (b) β_2, (c) λ_b (1900–1994) and λ_a (1995–2001), and (d) $\lambda_a - \lambda_b$. The ensemble average Poisson rate before 1995 $\langle\bar{\lambda}_b\rangle = 2.37$ and the ensemble average rate after (and including) 1995 $\langle\bar{\lambda}_a\rangle = 3.57$. This provides an ensemble average p-value on the rate increase equal to 0.041.

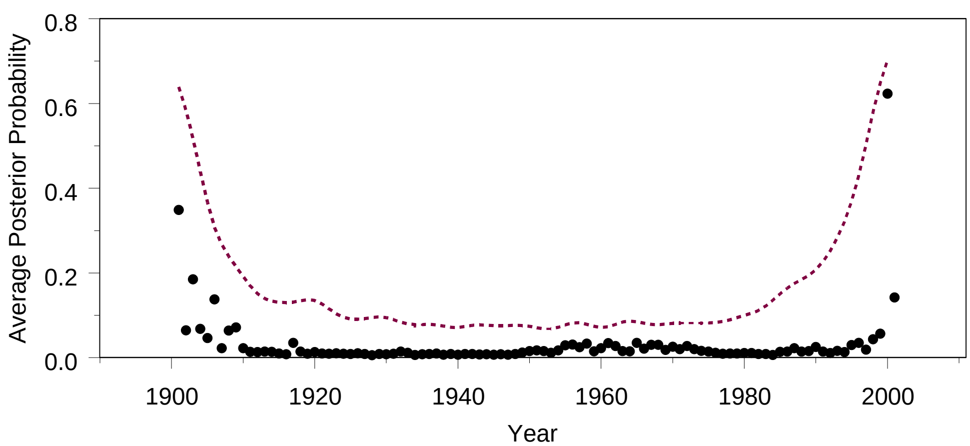

Figure 8: Same as Fig. 2 except for U.S. hurricanes. Note in this case there are no years with posterior probabilities that exceed the 95$ confidence level from the constant rate Poisson model.

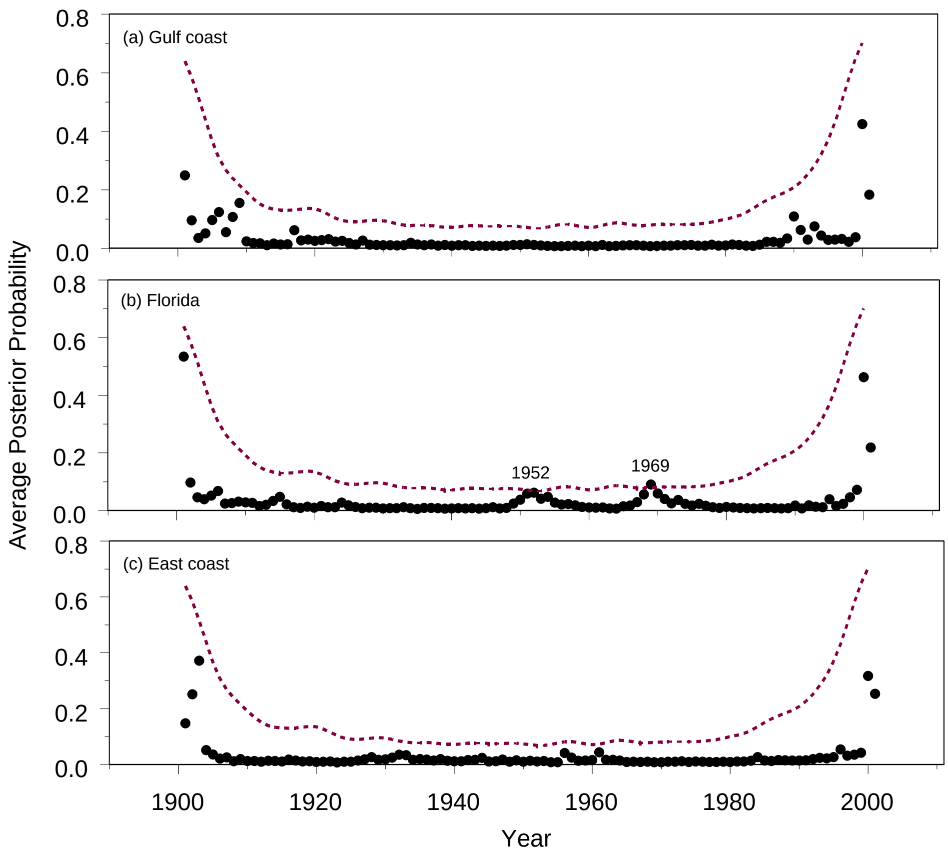

Figure 9: Same as Fig. 9 except for annual counts of hurricanes affecting the (a) Gulf coast, (b) Florida, and (c) East coast.

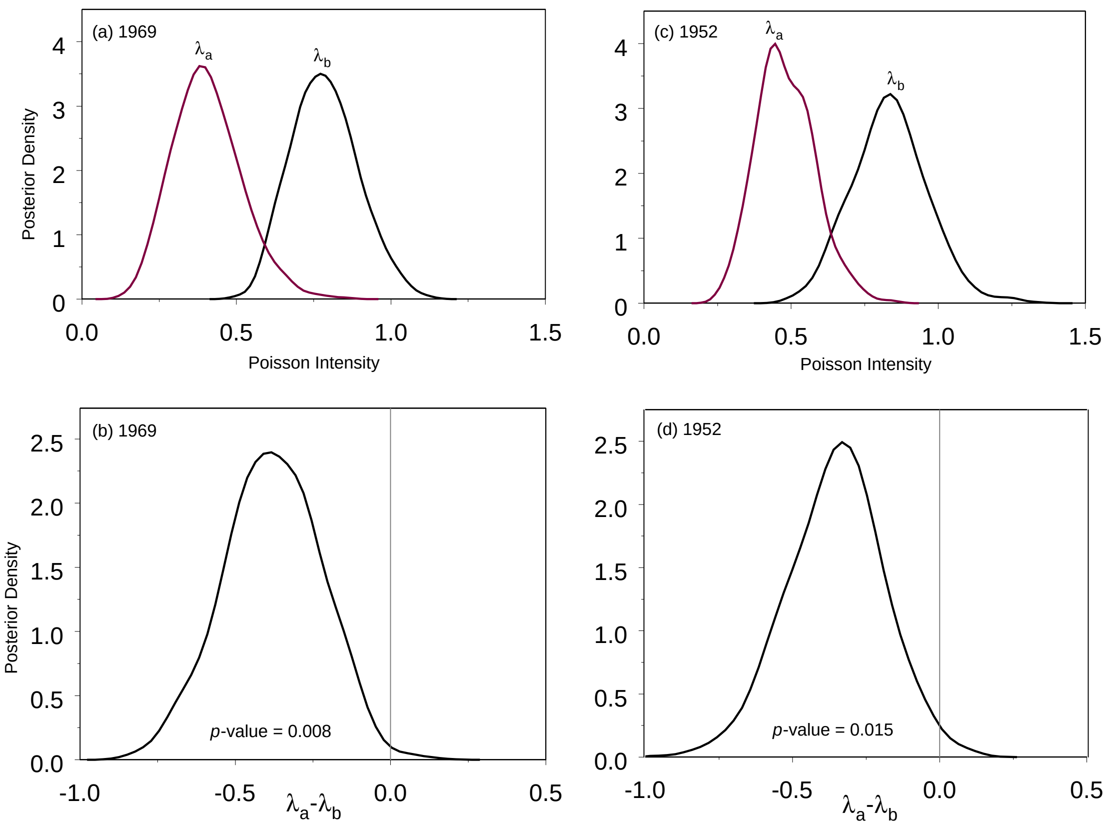

Figure 10: Estimates of the posterior densities from the Gibbs sampler applied to the time series of Florida hurricane counts. Density of (a) λ_b (1900–1968) and λ_a (1969–2001), and (b) $\lambda_a - \lambda_b$. The ensemble average Poisson rate before 1969 $\langle\bar{\lambda}_b\rangle = 0.79$ and the ensemble average rate after (and including) 1969 $\langle\bar{\lambda}_a\rangle = 0.40$. This provides an ensemble average p-value on the rate difference that equal to 0.008. Density of (c) λ_b (1900–1951) and λ_a (1952–2001), and (d) $\lambda_a - \lambda_b$. The ensemble average Poisson rate before 1952 $\langle\bar{\lambda}_b\rangle = 0.83$ and the ensemble average rate after (and including) 1952 $\langle\bar{\lambda}_a\rangle = 0.48$. This provides an ensemble average p-value on the rate difference equal to 0.015.

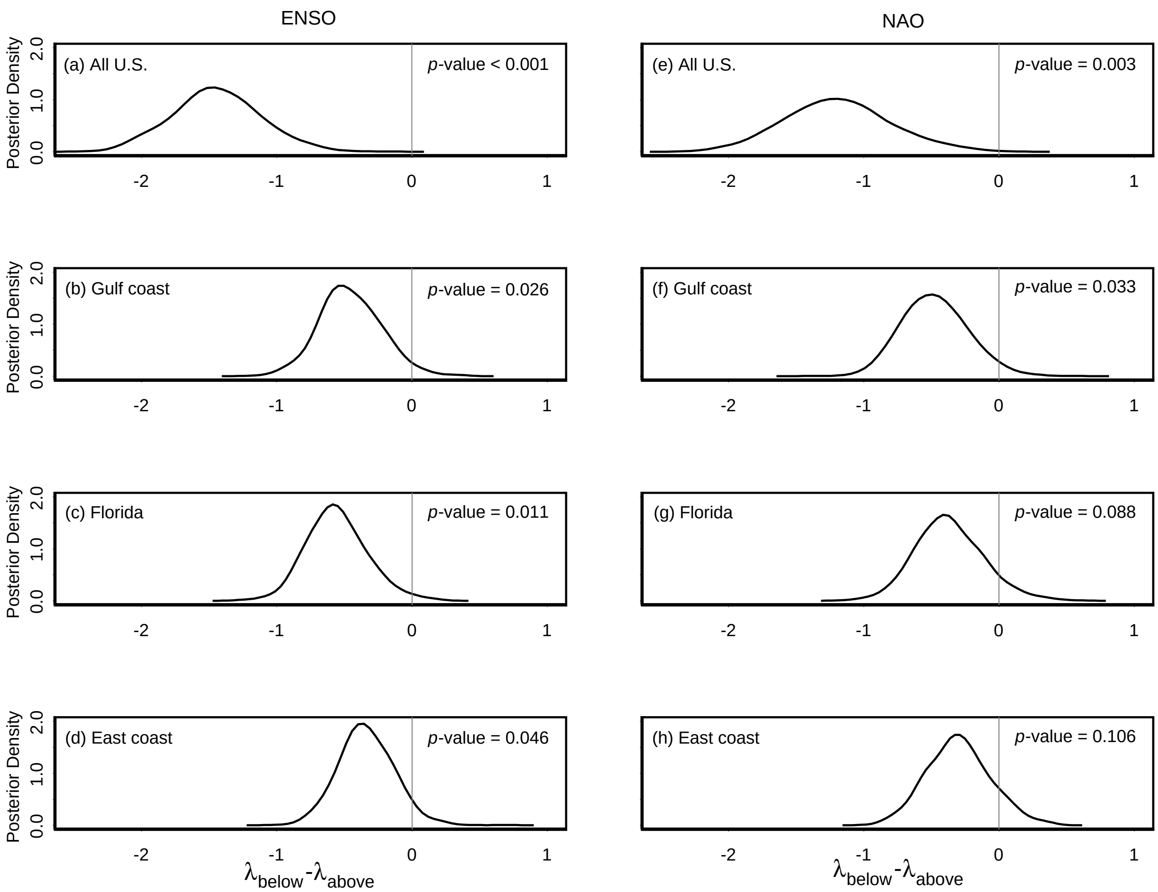

Figure 11: Posterior densities of the hurricane rate differences (below normal years minus above normal years) for (a) All U.S. hurricanes and ENSO, (b) Gulf coast hurricanes and ENSO, (c) Florida hurricanes and ENSO, (d) East coast hurricanes and ENSO, (e) All U.S. hurricanes and NAO, (f) Gulf coast hurricanes and NAO, (g) Florida hurricanes and NAO, and (h) East coast hurricanes and NAO.