VOLUME 14 JOURNAL OF CLIMATE 1 DECEMBER 2001

Bayesian Analysis of U.S. Hurricane Climate

JAMES B. ELSNER AND BRIAN H. BOSSAK

Department of Geography, The Florida State University, Tallahassee, Florida

(Manuscript received 12 March 2001, in final form 2 July 2001)

ABSTRACT

Predictive climate distributions of U.S. landfalling hurricanes are estimated from observational records over the period 1851–2000. The approach is Bayesian, combining the reliable records of hurricane activity during the twentieth century with the less precise accounts of activity during the nineteenth century to produce a best estimate of the posterior distribution on the annual rates. The methodology provides a predictive distribution of future activity that serves as a climatological benchmark. Results are presented for the entire coast as well as for the Gulf Coast, Florida, and the East Coast. Statistics on the observed annual counts of U.S. hurricanes, both for the entire coast and by region, are similar within each of the three consecutive 50-yr periods beginning in 1851. However, evidence indicates that the records during the nineteenth century are less precise. Bayesian theory provides a rational approach for defining hurricane climate that uses all available information and that makes no assumption about whether the 150-yr record of hurricanes has been adequately or uniformly monitored. The analysis shows that the number of major hurricanes expected to reach the U.S. coast over the next 30 yr is 18 and the number of hurricanes expected to hit Florida is 20.

1. Introduction

Landfalling hurricanes are of important social and economic concern in the United States. Strong winds and storm surge accompanying landfalling hurricanes kill people and destroy property. Their potential for destruction and loss of life rivals the potential for damage and casualties from earthquakes. Hurricane Andrew's strike on Florida during August of 1992 caused in excess of $30 billion in direct economic losses, while Hurricane Floyd's evacuation in 1999 disrupted the lives of 2.5 million of its residents. Knowledge of their past occurrence, even if it is incomplete, provides clues about their future frequency and intensity that goes beyond the capabilities of present climate prediction models in terms of specificity and lead time. This understanding is important for land-use planning, emergency management, hazard mitigation, (re)insurance application, and potentially the long-term derivative market.

Empirical and statistical research (Elsner et al. 2000a; Elsner et al. 1999; Gray et al. 1992) have identified factors that contribute to conditions favorable for hurricanes over the North Atlantic basin, which includes the Caribbean Sea and the Gulf of Mexico. Research shows that these factors influence the hurricane frequency differently depending on the particular region of the North Atlantic. For instance, the effect of an El Niño event on the frequency of hurricanes over the entire basin is significant, but the effect on the frequency of hurricanes forming over subtropical latitudes is small. In fact, additional factors are needed to explain the climate variation of hurricane activity locally (Jagger et al. 2001; Murnane et al. 2000; Lehmiller et al. 1997).

Here we consider the occurrence of tropical cyclones that make landfall in the continental United States as hurricanes and major hurricanes. The purpose of the present work is to provide a comprehensive climatology of U.S. coastal hurricanes. The focus on U.S. hurricanes allows us to use reliable data extending back at least to the start of the twentieth century. Moreover, less reliable but still useful information is available back through at least the second half of the nineteenth century. We are motivated to describe U.S. hurricane activity spatially and temporally using all available information since 1851. Results from a careful analysis define the climate that not only tells a story about the past, but can be used to gauge future activity.

The first step is to examine the observed hurricane landfall record. This is accomplished by dividing the 150-yr record into three 50-yr periods and comparing statistics. Statistics are compiled for the entire coastline and for three separate regions including the Gulf Coast, Florida, and the East Coast. For the entire coast, data are divided into hurricanes and major hurricanes. Hurricane statistics together with population data indicate a greater uncertainty in the landfall records from the nineteenth century. Yet a sharp decision to reject these earlier observations is wasteful. Instead, following the

---

Corresponding author address: James B. Elsner, Department of Geography, The Florida State University, Tallahassee, FL 32306.
E-mail: jelsner@garnet.fsu.edu

theory presented in Epstein (1985) we employ a Bayesian approach to determine a predictive climate distribution of coastal hurricane activity that uses the entire period and that makes no assumption about whether the 150-yr record has been adequately or uniformly monitored. In the following sections, we describe the observational record using statistics, present the Bayesian approach and results obtained by that method, and discuss ways in which the method can be extended. The last section gives a summary and lists the important conclusions.

2. Data

A hurricane is a tropical cyclone with maximum sustained (1 min) 10-m winds of 65 kt or greater. Hurricane landfall occurs when all or part of the storm's eyewall passes directly over the coast or adjacent barrier islands. Since the eyewall extends outward a radial distance of 50 km or more from the hurricane center, landfall may occur even in the case where the exact center of lowest pressure remains offshore. A hurricane can make more than one landfall as Hurricane Andrew did when it struck southeast Florida and Louisiana in 1992. For U.S. hurricanes we only consider whether or not the observations indicate that the cyclone struck the continental United States at least once at hurricane intensity. The approximate coastal length affected by hurricanes is 6000 km. Major U.S. hurricanes of category 3 or higher on the Saffir–Simpson hurricane damage potential scale (Simpson 1974) have winds of 100 kt or greater at landfall. For regional frequencies we consider reports of multiple landfalls if they occur in different regions.

The North Atlantic Hurricane Data Base (HURDAT or best track) is the most complete and reliable source of North Atlantic hurricanes (Jarvinen et al. 1984). The dataset consists of the 6-hourly position and intensity estimates of tropical cyclones back to 1886 (Neumann et al. 1999). Important additional contributions to the knowledge of past hurricanes were made by interpreting written accounts of tropical cyclones from ship logs, newspapers, and other nontraditional archives (Ludlum 1963; Fernández-Partagás and Diaz 1996). These studies update and add information about hurricane landfalls during the period 1851–1900. For instance, the *New York Times*'s reports of damage and casualties often contain sufficient details to reconstruct the location and intensity of a hurricane at landfall. Arguably these additional sources of U.S. hurricane information provide justification to extend the U.S. hurricane record back to the preindustrial era (Elsner and Kara 1999; Elsner et al. 2000b). Recently, the National Oceanic and Atmospheric Administration (NOAA) embarked on a 3-yr hurricane reanalysis project. The motivation was, in part, to reduce the level of uncertainty surrounding the historical reports of hurricanes during the last half of the nineteenth century. The hurricane-landfall phase of the project was completed in July 2000. A concatenated dataset consisting of landfalling hurricane accounts from historical archives and modern direct measurements is carefully analyzed here.

TABLE 1. Annual U.S. hurricane and major hurricane statistics. Values include the mean, variance, maximum, minimum, and quantiles of the mean from a bootstrap resampling (number of bootstrap samples is 1000).

Years	Total No.	Mean	Variance	Max	Min	Quantiles of the mean 5%	95%
All U.S. hurricanes							
1851–1900	88	1.76	2.349	7	0	1.42	2.14
1901–50	92	1.84	1.770	6	0	1.54	2.16
1951–2000	72	1.44	1.558	6	0	1.16	1.72
Major U.S. hurricanes							
1851–1900	26	0.52	0.500	3	0	0.36	0.68
1901–50	36	0.72	0.696	3	0	0.54	0.92
1951–2000	27	0.54	0.458	3	0	0.38	0.70

3. Multidecadal comparisons

The principal objective of this paper is a climatology of continental coastal hurricane activity based on the available observations during the past 150 years. This is important as it provides a benchmark against which future activity as well as forecasts of future activity can be gauged. Because we are utilizing a long record of storms it is worthwhile to compare samples of storms in three consecutive 50-yr intervals: 1851–1900, 1901–50, and 1951–2000. To appreciate the length of record used in this analysis we note that 1851 marked Oersted's discovery of the magnetism of an electric current, Lord Kelvin's introduction of the absolute temperature scale, and Foucault's invention of the pendulum for demonstrating the earth's rotation (Heilbron and Bynum 2001). We begin the comparison by considering hurricane activity along the entire coastline of the United States.

a. Entire coast

Table 1 provides descriptive statistics of U.S. hurricane and major hurricane activity during the three consecutive time periods. The average annual number of U.S. hurricanes ranges from a high of 1.8 during the period 1901–50 to a low of 1.4 during the period 1951–2000. Each subperiod has at least 1 yr of 6 or more hurricanes, with the most in a single year occurring in 1886. The 90% confidence (credible) intervals generated from a bias-corrected bootstrap procedure (Efron and Tibshirani 1993), indicate large overlap suggesting minor differences in the observed rates of U.S. hurricanes between consecutive periods. The bootstrap procedure considers each year as independent and resamples, with replacement, the annual counts. The number of bootstrap samples is 1000. The independence assumption is valid since the lag-1 temporal autocorrelation value is a negligible -0.002.

Table 2. Test of differences in mean hurricane rates. Numbers are *p* values from a Wilcoxon signed rank test under the null hypothesis of no difference in mean rates between periods.

Difference in periods	All	Major	Gulf Coast	Florida	East Coast
1851–1900 minus 1901–50	0.563	0.230	0.298	0.198	0.552
1901–50 minus 1951–2000	0.082	0.333	0.338	0.011	0.465
1851–1900 minus 1951–2000	0.301	0.781	0.887	0.281	0.897

We test for differences in mean rates between the three periods using a Wilcoxon signed rank test. Table 2 shows the *p* values based on the large-sample approximation for the null hypothesis of no difference against the two-sided alternative. The values indicate little evidence against the null hypothesis of equal hurricane rates, so we make the assumption of stationarity for these data and time periods.

In contrast, the variance of interannual activity decreases from 2.3 during the last half of the nineteenth century to 1.6 during the last half of the twentieth century, indicating a potential bias during the earliest 50-yr period. This is because information is measured in terms of precision, which is the inverse of the variance. Larger interannual variance (lower precision) during the nineteenth century might result from an incomplete record. A hurricane striking southeastern Florida or southern Texas during the 1850s could have gone undetected as these areas were undeveloped at that time. Or a tropical storm at landfall might have been misclassified as a hurricane due to insufficient or inaccurate historical accounts near the storm center.

Similar statistics on the annual occurrence of major U.S. hurricanes are provided in Table 1. The numbers indicate an average of approximately two major U.S. hurricanes every four years during the periods 1851–1900 and 1951–2000, and an average of two major hurricanes every five years during the period 1901–50. Each subperiod has at least one year in which three major hurricanes reached the coast. The lag-1 autocorrelation for annual major hurricane counts is $+0.030$ and the Wilcoxon tests provide no evidence against the null hypothesis of constant rates (Table 2).

Overall the data on U.S. hurricanes and U.S. major hurricanes provide little or no evidence of statistically significant differences in the level (rate) of activity between the three 50-yr periods. Figure 1 provides histograms of U.S. hurricane and major hurricane activity over the three 50-yr periods. For U.S. hurricanes the distributions are relatively flat in each of the three periods. During the first half of the twentieth century, a greater number of years have two or more hurricanes, while during the second half of the century a greater number of years have exactly one hurricane. Similarly, only small differences are noted in the distributions of U.S. major hurricanes.

Intraseasonal variations in coastal hurricane activity also are examined for each of the three 50-yr periods (Fig. 2). The average date of landfall occurs within the first week of September, although the most active of these 10-day intervals extends from 9 to 18 September. The standard deviations range from 32 days during the periods 1851–1900 and 1951–2000 to 35 days during the period 1901–50. This display is important in highlighting the apparent steadiness in U.S. hurricane statistics since early industrial times.

b. Gulf Coast, Florida, and the East Coast

While we find no discernible differences in the rates of overall U.S. hurricane activity between the 50-yr intervals, there exists the possibility of regional shifts in

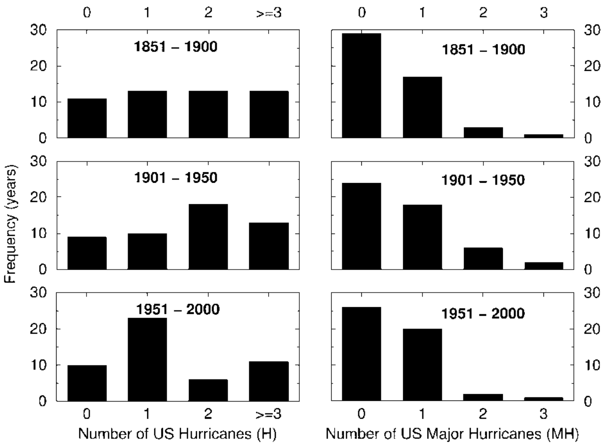

Fig. 1. Histograms of the observed annual occurrence of U.S. hurricanes and major hurricanes during each of the 50-yr periods.

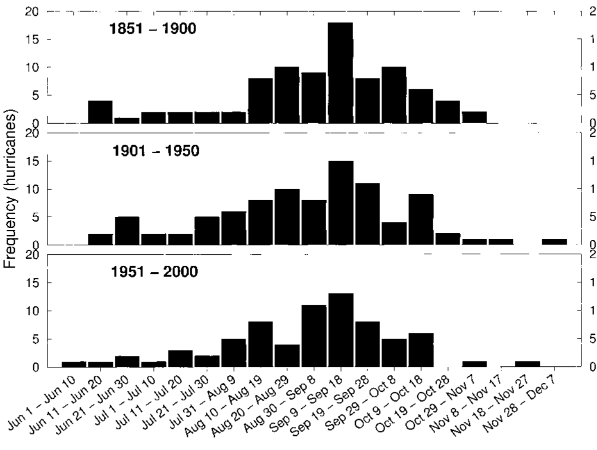

Fig. 2. Histograms of the observed intraseasonal occurrence of landfalling hurricanes during each of the 50-yr periods.

TABLE 3. Annual regional hurricane statistics. The three regions include the Gulf Coast, Florida, and the East Coast. Values include the mean, variance, maximum, minimum, and quantiles of the mean from a bootstrap resampling (number of bootstrap samples is 1000).

Years	Total No.	Mean	Variance	Max	Min	Quantiles of the mean 5%	Quantiles of the mean 95%
Gulf Coast hurricanes							
1851–1900	32	0.64	0.807	4	0	0.44	0.84
1901–50	38	0.76	0.676	3	0	0.58	0.96
1951–2000	30	0.60	0.531	3	0	0.42	0.74
Florida hurricanes							
1851–1900	35	0.70	0.867	3	0	0.48	0.90
1901–50	43	0.86	0.653	2	0	0.66	1.04
1951–2000	24	0.48	0.540	3	0	0.32	0.64
East Coast hurricanes							
1851–1900	25	0.50	0.582	3	0	0.34	0.68
1901–50	19	0.38	0.362	2	0	0.22	0.52
1951–2000	26	0.52	0.622	3	0	0.34	0.68

activity. Thus, we next divide the coast into three geographical regions including the Gulf Coast, Florida, and the East Coast, and separately examine the number of hurricanes affecting each region. The Gulf Coast includes the region from Brownsville, Texas, to the Alabama–Florida line, and the East Coast extends from the Florida–Georgia line to Eastport, Maine. The approximate precentages of total coastline are 22.3, 34.8, and 42.9 for the Gulf Coast, Florida, and East Coast, respectively.[1] Note that a single U.S. hurricane may count as a Florida and a Gulf Coast hurricane if it makes separate landfalls in both regions. However, a hurricane making more than one landfall in the same region is counted only once. The lag-1 autocorrelations on the annual hurricane counts are 0.009, 0.008, and 0.014 for the Gulf Coast, Florida, and the East Coast, respectively.

Table 3 provides descriptive statistics for landfalling hurricanes along the Gulf Coast, across Florida, and along the East Coast during the three periods. Overall the East Coast has somewhat fewer hurricanes compared with either Florida or the Gulf Coast. The greatest concentration of hurricanes over Florida occurred during the first half of the twentieth century. Yet during the second half of the twentieth century Florida experienced fewer hurricanes than the East Coast. There is a positive correlation between activity over Florida and the Gulf Coast, and an anticorrelation between activity along the Gulf and East Coasts (Elsner and Kara 1999). In these three subperiods when hurricane activity is above (below) the long-term average over the Gulf Coast and Florida, it is below (above) the average along the East Coast. Spatial variations in major U.S. hurricanes are examined in Elsner et al. (2000b), Jagger et al. (2001), and Elsner and Bossak (2001).

Quantile intervals around the mean show large overlap between successive periods providing scant observational evidence of secular changes in the rates of regional hurricane activity. Figure 3 provides histograms of landfalling hurricanes in the three regions for the three 50-yr periods. With the exception of the Gulf Coast and Florida during the first half of the twentieth century, the majority of years are without a landfalling hurricane. Also, of the three regions, the East Coast is most likely to have a year without a landfall. Results of the Wilcoxon tests on differences in average rates are shown in Table 2. Again we see no statistically significant differences in rates over time. One exception is Florida between the first and second halves of the twentieth century. Here the p value is 0.011. However, since we are performing 15 independent tests, a Bonferonni adjustment requires the individual p value to be less than 0.0033 (0.05/15) to reject the null hypothesis of no difference. Thus, as with the overall coastal hurricane activity analyzed in the previous section, the 150-yr period beginning in 1851 shows no obvious temporal shifts in the level of hurricane activity on multidecadal timescales.

As we have seen for total hurricane activity, regional hurricane statistics also hint at less precision during the nineteenth century. In particular, hurricane data for the Gulf Coast and Florida indicate a substantially greater amount of interannual variance in the earliest period (1851–1900). This is consistent with the fact that records are less precise in areas that had few permanent residents. In other words, the actual level of hurricane activity during the nineteenth century may have been considerably different than what the available observations indicate. Next we show how these earlier, less reliable, records can be combined with the reliable records from the twentieth century to provide a predictive climate of U.S. hurricane activity.

4. A Bayesian approach

Observational information on past hurricane activity is available from instrumental records and historical accounts, with the historical accounts having a greater degree of uncertainty. Representing uncertainty is the province of probability theory, with its practical application of the domain of statistics (Pole et al. 1999). The Bayesian statistical approach provides a rational and coherent foundation for using all available information, while explicitly accounting for differences in uncertainty (see also, Walshaw 2000). Here we follow the formalism presented in Epstein (1985).

a. Poisson process

The arrival of hurricanes on the coast can be considered a Poisson process (Elsner et al. 2001; Solow and Moore 2000; Parisi and Lund 2000; Elsner and Kara 1999; Bove et al. 1998). The Poisson distribution is a limiting form of the binomial distribution with no upper

---

[1] Source: *Microsoft, Inc., Encarta Encyclopedia.*

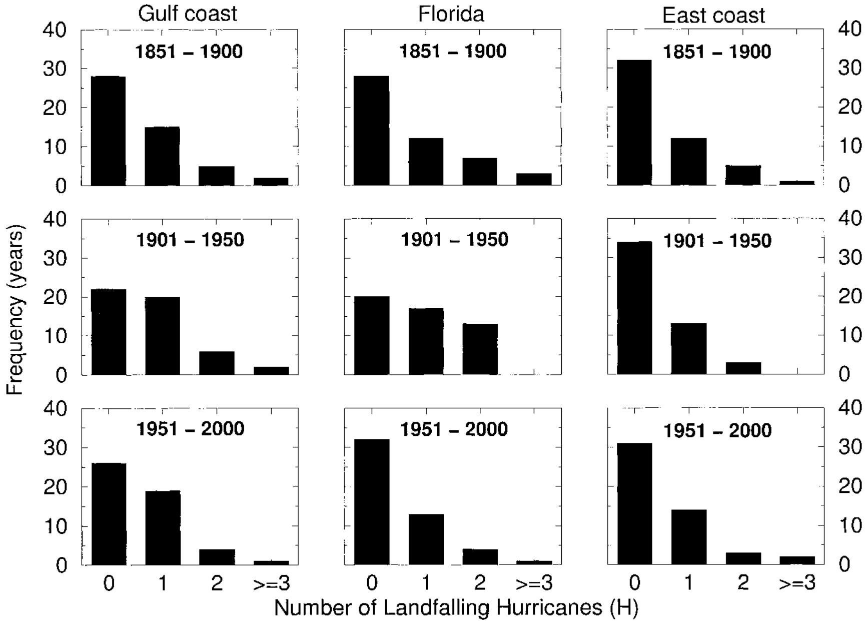

FIG. 3. Histograms of the observed annual occurrence of landfalling hurricanes within the three regions during each of the 50-yr periods.

bound on the count of occurrences. The parameter λ, the intensity, characterizes a Poisson process. Note in Tables 1 and 3 that the annual means and variances are of similar magnitude. Knowledge of λ allows statements to be made about future outcomes, but since the process is stochastic, the statements must necessarily by couched in terms of probabilities. For example, the probability of $\hat{h}$ hurricanes occurring in T years is

$$f_{\mathrm{Poisson}}(\hat{h}\,|\,\lambda, T) = \exp(-\lambda T)\frac{(\lambda T)^h}{h!}$$

$$\text{for } h = 0, 1, 2, \ldots,$$

$$\lambda > 0, \quad \text{and} \quad T > 0. \qquad (1)$$

The parameter λ and statistic T appear in the formula as a product, which is the mean and variance of the Poisson distribution. More important, knowledge of λ can come from whatever information is available and we want to use the best a posteriori knowledge of λ in making predictions about future hurricane activity (see, e.g., Epstein 1985). Therefore it is necessary to treat λ not as a single-valued parameter but as a continuous random variable that can take on any positive real number. The functional form for expressing judgement about λ is the gamma distribution (Epstein 1985).

The numbers that describe the outcome of a Poisson process for seasonal hurricane activity are the length of the time interval sampled T', and the number of hurricanes that occur over this interval h'. For instance, during the first 10 years of the record (1851–60), observations indicate 15 U.S. hurricanes, so $T' = 10$ and $h' = 15$.

The gamma distribution of possible future values for λ is given by

$$f_\gamma(\hat{\lambda}\,|\,h', T') + \frac{T'^{h'}\lambda^{h'-1}}{\Gamma(h')}\exp(-\lambda T'), \qquad (2)$$

with the expected value $\mathrm{E}(\hat{\lambda}) = h'/T'$, and the gamma function $\Gamma(x)$ given by

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt. \qquad (3)$$

Of importance here is the fact that, if the probability density on $\hat{\lambda}$ is a gamma distribution, with initial numbers (prior parameters) h' and T', and the statistics h and T are later observed, then the posterior density of $\hat{\lambda}$ is also gamma with parameters $h + h'$ and $T + T'$. In Bayesian terminology, the gamma density is the conjugate prior for the intensity of the Poisson process, λ.

b. Posterior density for λ

The additive nature of the prior parameters h' and T' with the sample statistics h and T indicate how to combine the earlier, unreliable hurricane records with the later, reliable records to obtain a posterior density on the annual hurricane rates λ. Since the earlier records have greater uncertainty we must incorporate this lack of precision into our estimates of the prior parameters. To quantify our prior judgement about λ we use a bootstrap procedure to estimate quantiles on the annual counts of hurricanes during the uncertain period.

The record of U.S. hurricanes is uncertain before

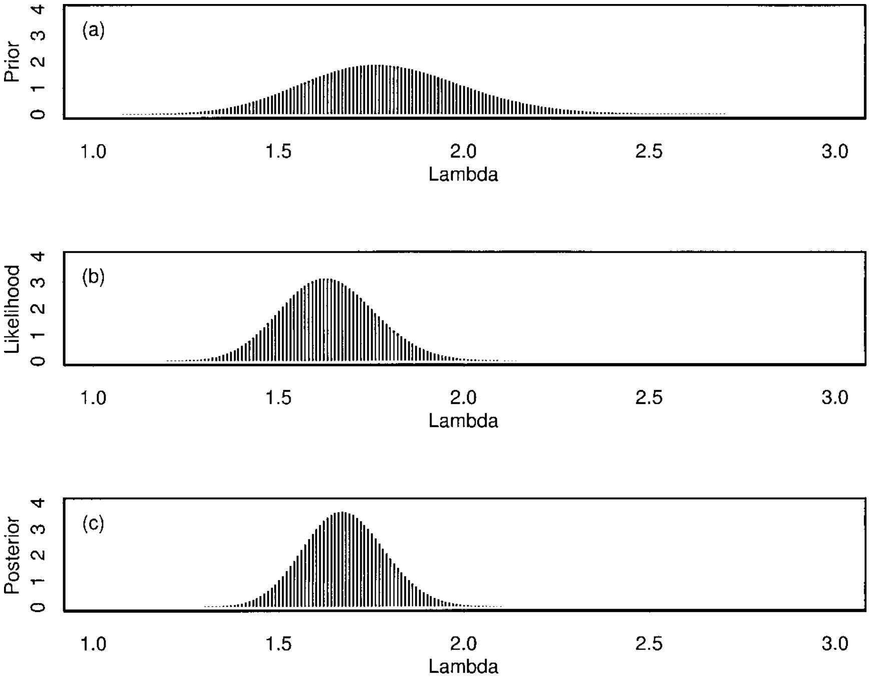

FIG. 4. Gamma densities on the Poisson intensity λ (lambda) for annual U.S. hurricane rates based on (a) prior, (b) likelihood, and (c) posterior parameters.

1900. However, a bootstrap of the annual mean from the available observations over the period 1851–99 indicates a 90% confidence interval of 1.45 to 2.16 hurricanes per year. Although one cannot say for certain what the true rate of U.S. hurricanes was over this earlier period, we make a sound judgement that we are 90% confident that the credible (confidence) interval contains it. In other words, we admit a 5% chance that the true rate is less than 1.45 and a 5% chance that it is greater than 2.16.

To capture this information, we make use of the close relationship between the gamma and χ^2 distributions. Specifically, if $\hat{\lambda}$ is gamma with parameters h' and T', then the random variable $\hat{Z} = 2\hat{\lambda}T'$ is χ^2 with $2h'$ degrees of freedom (Epstein 1985). The probability that $\lambda < 1.45$ is 0.05 implies that $\hat{Z} = 2(1.45)T' = 2.9T'$ is $\chi^2_n(0.05)$, where $\chi^2_n(0.05)$ is the lower 0.05 quantile of a χ^2 distribution with n degrees of freedom. Similarly, the probability that $\lambda < 2.16$ is 0.95 means that $\hat{Z} = 2(2.16)T' = 4.32T'$ is $\chi^2_n(0.95)$, where $\chi^2_n(0.95)$, is the upper 0.95 quantile. Thus the ratio of the upper to lower quantiles of the χ^2 variable must be $2.16/1.45 = 1.49$, and the degrees of freedom when the χ^2 ratio is 1.49 are 138. Since h' is one-half the degrees of freedom, $h' = 69$. Moreover, T' is $\chi^2_{138}(0.05)/2.9 = 38.6$. This procedure formally quantifies the prior information.

After quantifying our prior judgement we have two distinct pieces of information for obtaining a posterior distribution on λ. We have our likelihood statistics based on the reliable period of record (1900–2000) which gives $h = 165$ and $T = 101$ and we have the prior parameters $h' = 69$ and $T' = 38.6$. Note that the reliable period includes 1900 and thus h and T are slightly different from those presented in Table 1. The posterior parameters are thus $h'' = h + h' = 234$ and $T'' = T + T' = 139.6$. Note that although the likelihood parameters h and T must be integers, the prior parameters can take on any real value depending on our degree of belief (Epstein 1985). Figure 4 shows the prior, likelihood, and posterior gamma densities for the Poisson parameter λ based on (2). Of note is the relatively broad distribution for the prior estimate indicative of both the uncertainty and relative short length of the unreliable period. The distribution of the likelihood estimate is narrow and centered on a mean rate of 1.6 hurricanes per year. Combining the prior and likelihood results in a posterior distribution that represents the best information about λ. The posterior distribution has flatter tails representing less uncertainty than both the prior and likelihood distributions.

c. *Predictive distribution*

The knowledge we obtain about λ from the posterior distribution is codified in the two numbers $h'' = h + h'$ and $T'' = T + T'$ of the gamma density. Of practical interest is information about future hurricane activity. Therefore, the question becomes how to obtain this future information when the posterior annual rate is known only in terms of a probability distribution. The answer lies in the fact that the *predictive density* for obtaining $\hat{h}$ U.S. hurricanes over the next $\hat{T}$ years when knowledge

TABLE 4. Bayesian statistics. Values of the statistics are used in determining the predictive distributions of U.S. hurricanes (''NA'' indicates not applicable).

Statistic	All U.S.	Major U.S.	Gulf Coast	Florida	East Coast
Period of record	1851–2000	1851–2000	1851–2000	1851–2000	1851–2000
Likelihood statistics					
Reliable period	1900–2000	1900–2000	1880–2000	1900–2000	1851–2000
h (No. hurricanes)	165	64	82	67	70
T (yr)	101	101	121	101	150
Prior statistics					
Unreliable period	1851–99	1851–99	1851–79	1851–99	None
Confidence interval	(1.45, 2.16)	(0.35, 0.67)	(0.41, 0.86)	(0.51, 0.92)	NA
dof	138	54	40	64	NA
h' (No. hurricanes)	69	27	20	32	NA
T' (yr)	38.6	54.5	32.3	45.7	NA
Posterior statistics					
h'' (No. hurricanes)	234	91	102	99	70
T'' (yr)	139.6	155.5	153.3	146.7	150
A sample of prediction values					
$\hat{T}$ (yr)	10	10	10	10	10
$\hat{p}$ (probability)	0.933	0.940	0.939	0.936	0.938

of the annual rate is contained in the gamma density with parameters h'' and T'' is a negative binomial distribution, with parameters h'' and $T''/(\hat{T} + T'')$ (see Epstein 1985)

$$f_{\mathrm{nb}}\left(\hat{h} \mid h'', \frac{T''}{\hat{T} + T''}\right) = \frac{\Gamma(\hat{h} + h'')}{\Gamma(h'')\hat{h}!}\left(\frac{T''}{\hat{T} + T''}\right)^{h''}\left(\frac{\hat{T}}{\hat{T} + T''}\right)^{\hat{h}}. \qquad (4)$$

The mean and variance of the negative binomial are $\hat{T}h''/T''$ and $\hat{T}h''/T''[(\hat{T} + T'')/T'']$, respectively. Note that the variance of the predictive distribution is always larger than it would be if λ were known precisely. If we are only interested in the climatological probability of a hurricane next year, then $\hat{T}$ is one and small when compared with T'' so it makes little difference, but if we are interested in the distribution of likely hurricane activity over the next 10, 20, or 30 yr then it is important.

5. Results

Here we present results of the Bayesian approach to generating predictive climate distributions of U.S. hurricanes. We examine hurricanes and major hurricanes along the entire U.S. coast as well as hurricanes affecting the Gulf Coast, Florida, and the East Coast separately. Table 4 lists the values of the Bayesian statistics used in determining the predictive distributions. The start year of reliability is assigned based on U.S. census of ''settled regions'' defined as at least two inhabitants per square mile (Landsea 2001, unpublished manuscript). We use the latest reliable year for the region. For the entire U.S. coast, the record is not reliable before 1900 because historical records from sparsely populated regions like southern Florida are missing before this time. For the Gulf Coast (excluding Florida), reliable records extend back to 1880 before which they are unreliable for south Texas. For the East Coast reliable records extend back to at least 1851. The likelihood parameters (h and T) are determined from annual counts over the reliable period and the 90% confidence intervals are determined from a bootstrap resampling of the mean annual rate over the unreliable period. The prior parameters are then estimated from the ratio of the upper to lower bounds on the confidence interval as explained in the previous section. Posterior parameters are the sum of the prior and the likelihood statistics, except for the East Coast where only likelihood information is used. Predictive values are representative of climate timescales.

Figure 5 show the predictive densities for all coastal hurricanes and all major hurricanes. Here the reliable period begins in 1900. Figures 5a,d show the probability of observing a specific number of hurricanes and major hurricanes over the next 10 years. Note the tails are fatter on the right side of the distributions. Figures 5b,e show the cumulative probability distributions of observing no more than a specified number of storms over the next 10 years. Figures 5c,f show the cumulative probability distributions of observing at least the specified number of storms over the next 10, 20, and 30 years. The expected number of U.S. hurricanes over the next 30 years is 50 of which 18 are anticipated to be intense. These probability distributions represent the best estimates of the future climate of U.S. hurricanes.

Figure 6 shows the predictive densities for regional hurricane activity along the Gulf Coast, Florida, and the East Coast. The figure is arranged as previously, with Figs. 6a,d,g showing the probability of a specified num-

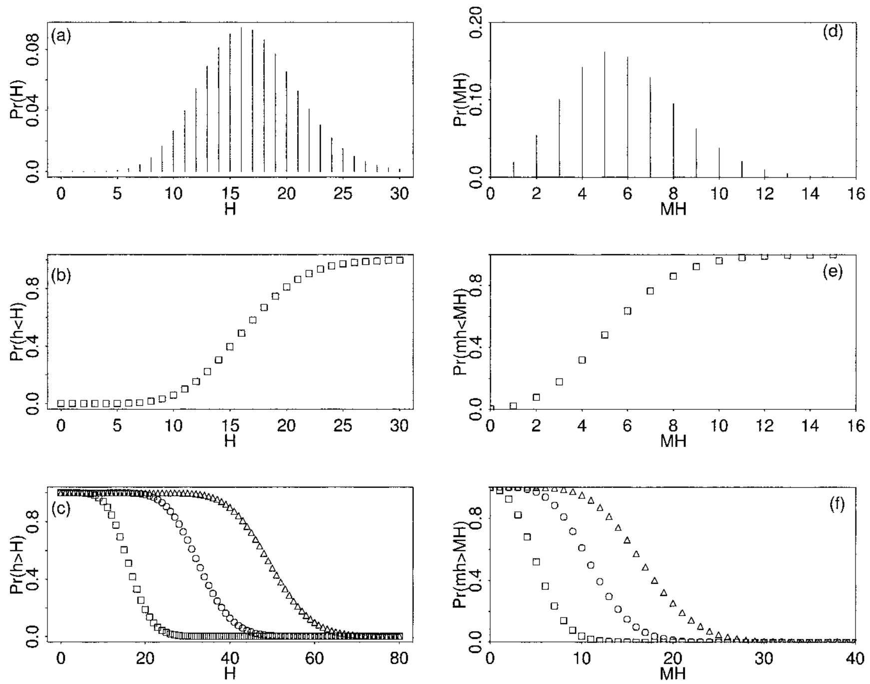

FIG. 5. Predictive densities for the likelihood of U.S. hurricanes (H) and U.S. major hurricanes (MH) over the next 10–30 yr. For U.S. hurricanes (a) is the probability of observing H hurricanes in 10 yr, (b) is the cumulative probability of $h \leq H$ in 10 yr, and (c) is the probability of $h \geq H$ in 10 (squares), 20 (circles), or 30 (triangles) yr. Plots (d)–(f) are the same, but for major U.S. hurricanes.

ber of hurricanes and Figs. 6b,e,h,c,f,i showing the cumulative probability distributions. The probability of hurricanes along the Gulf Coast and Florida are similar and higher than the probability along the East Coast. The probability of observing precisely 7 hurricanes during the next 10 years is close to 14% for both the Gulf Coast and Florida, but is less than 10% for the East Coast, and the probability of 10 hurricanes exceeds 5%, except along the East Coast where it is approximately 2%. The probability of observing more than 20 hurricanes over the next 30 years exceeds 40% for the Gulf Coast and Florida, but is less than 10% for the East Coast.

6. Refinements

Results from the previous section represent a raw climatology of U.S. hurricane activity from the available database. The predictive distributions are useful in defining a long-term climatological prediction of future activity predicated on the past. The climatology can be adjusted by conditioning on knowledge of teleconnections (Elsner et al. 2001; Jagger et al. 2001). The Bayesian approach is to use a ‘‘dynamic’’ model whereby the predictor parameters (θ_t) are distinct but stochastically related through the system equation

$$\theta_t = \mathbf{G}_t \theta_{t-1} + \omega_t, \tag{5}$$

where $\mathbf{G}_t$ is a matrix of known coefficients and ω_t is an unobservable stochastic term (Pole et al. 1999). The system equation has the general form of a first-order Markov process, where the matrix $\mathbf{G}_t$ defines a known deterministic functional relationship of the parameter vector at one time with its value at the immediately preceding time. The Bayesian approach can be used also to account for the influence (if any) of global warming on hurricanes by discounting the older information. That is, records of landfalling storms from the most recent years, possibly related to current trends, are given more weight than records from earlier decades.

Additional refinements are possible. For one thing, the hurricane records of Ludlum (1963) could be incorporated as a separate prior using a similar approach. Moreover, geological records of overwash deposits associated with storm surge (see Liu and Fearn 2000; Donnelly et al. 2001) could be included. Also, the prior distribution obtained from the NOAA reanalysis period could be determined differently. For instance, additional data on the level of uncertainty and likely bias on the landfall intensity estimate of each storm are provided. A Monte Carlo sampling procedure could be devised to estimate the upper and lower bounds of the confidence interval using this information. This can be achieved by assigning a subjective (or empirical) probability distribution to the wind speeds at landfall for each storm.

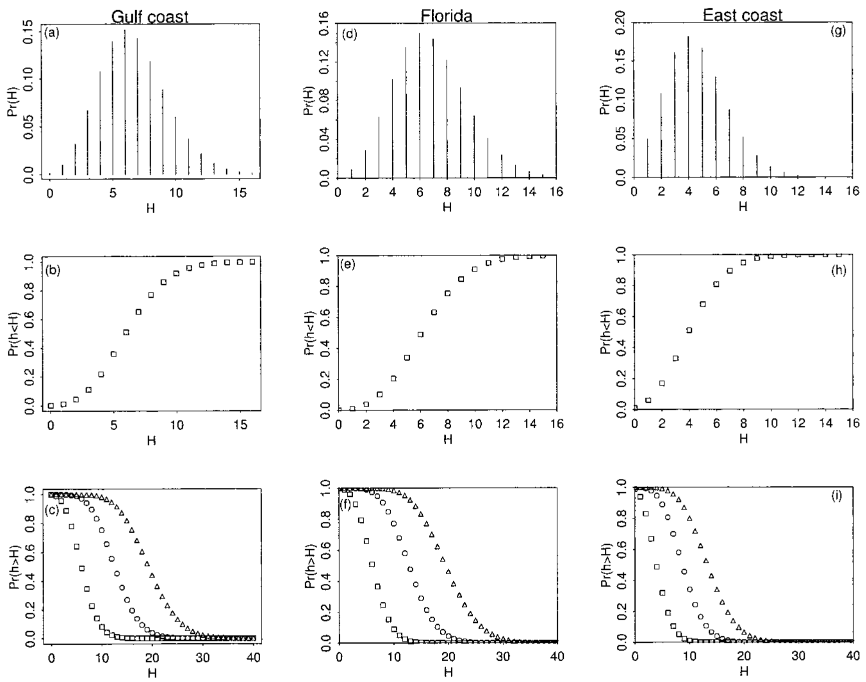

FIG. 6. Predictive densities for the likelihood of hurricanes along the Gulf Coast, Florida, and the East Coast over the next 10–30 yr. For Gulf Coast hurricanes (a) is the probability of observing H hurricanes in 10 yr, (b) is the cumulative probability of $h \leq H$ in 10 yr, and (c) is the probability of $h \geq H$ in 10 (squares), 20 (circles), or 30 (triangles) yr. Plots (d)–(f) are the same, but for Florida hurricanes, and plots (g)–(i) are the same, but for East Coast hurricanes.

The Monte Carlo sampler then repeatedly chooses a possible set of annual hurricane counts over the uncertain period. The resulting Monte Carlo distribution gives the desired bounds on the credible (confidence) interval. Work in these areas might make the results presented here still more precise.

7. Summary and conclusions

Hurricanes cause significant social and economic disruption within the United States. Here we examine the historical record of U.S. hurricanes back to 1851. Data come from the hurricane reanalysis project made possible by the meritorious works of Ludlum (1963) and Fernández-Partagás and Diaz (1996). The present analysis provides a comprehensive comparison of the modern hurricane record with the record from early industrial times.

First we examined variations in hurricanes over three consecutive 50-yr subperiods. Results show a picture of homogeneity in the level of coastal hurricane activity. The distributions of hurricanes during each subinterval are indistinguishable, indicating a stationary record of hurricanes since early industrial times. Stationarity is found for all hurricanes and major hurricanes as well as for regional activity, including the Gulf Coast, Florida, and the East Coast. However, evidence of a bias in the earlier records exists. As such, we determine the predictive distribution of hurricane activity using a Bayesian approach.

The Bayesian statistical approach provides a rational and coherent foundation for incorporating all available information while explicitly accounting for the differences in the degree of uncertainty. Here we followed the work of Epstein (1985) to determine predictive climatological distributions of hurricane activity over the next 10–30 years. The method is based on the fact that the gamma density is the conjugate prior for the intensity of the Poisson process (see also Krzysztofowicz 1983). Better understanding of hurricane occurrences over time provides a sound basis for assessing the likely losses associated with a catastrophic reinsurance contract (Michaels et al. 1997). It also might provide an instrument for trading futures options on the long-term derivative market.

The main conclusions of this paper are the following:

- The statistics of the observed counts of U.S. hurricanes are similar within each of the three consecutive 50-yr periods beginning in 1851.
- Similar statistical distributions are noted across regional hurricane activity.
- Evidence suggests that hurricane records from the nineteenth century are less precise, with the level of precision depending on region.
- Bayesian theory provides a framework to define a

predictive hurricane climate that uses all the available records, and that can be used as a benchmark against which future activity is gauged.

- According to this climatology, the expected number of U.S. hurricanes over the next 30 years is 50 with 18 of these hurricanes anticipated to become intense. The mean number of Florida hurricanes over this period is anticipated to be 20.

Acknowledgments. We thank S. Kavlakov and S. Elsner for fruitful discussions on this topic and E. Fogarty for help with the data. The study was funded by the National Science Foundation (ATM-0086958) and the Risk Prediction Initiative of the Bermuda Biological Station for Research (RPI-99-001).

REFERENCES

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.

Donnelly, J. P., and Coauthors, 2001: A 700-year sedimentary record of intense hurricane landfalls in southern New England. *Geol. Soc. Amer. Bull.,* **113,** 714–727.

Efron, B., and R. J. Tibsharini, 1993: *An Introduction to the Bootstrap.* Chapman and Hall, 436 pp.

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

——, and B. H. Bossak, 2001: Changes in hurricane landfall probabilities in response to variations in climate. *Hurricanes and Typhoons: Past, Present, and Future,* R. J. Murnane and K.-B. Liu, Eds., Columbia University Press, in press.

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

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

——, 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.

——, B. H. Bossak, and X. Niu, 2001: Secular changes to the ENSO–U.S. hurricane relationship. *Geophys. Res. Lett.,* in press.

Epstein, E. S., 1985: *Statistical Inference and Prediction in Climatology: A Bayesian Approach. Meteor. Monogr.,* No. 42, Amer. Meteor. Soc., 199 pp.

Fernández-Partagás, J., and H. F. Diaz, 1996: Atlantic hurricanes in the second half of the nineteenth century. *Bull. Amer. Meteor. Soc.,* **77,** 2899–2906.

Gray, W. M., C. W. Landsea, P. W. Mielke Jr., and K. J. Berry, 1992: Predicting Atlantic seasonal hurricane activity 6–11 months in advance. *Wea. Forecasting,* **7,** 440–455.

Heilbron, J. L., and W. F. Bynum, 2001: 1901 and all that. *Nature,* **409,** 14.

Jagger, T., J. B. Elsner, and X. 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, 1886–1983: Contents, limitations, and uses. NOAA Tech. Memo. NWS NHC-22, Coral Gables, FL, 21 pp.

Krzysztofowicz, R., 1983: Why should a forecaster and a decision maker use Bayes theorem. *Water Resour. Res.,* **19,** 327–336.

Landsea, C. W., 2001: The Atlantic hurricane data (HURDAT) reanalysis project: Results for the 19th century. *Hurricanes and Typhoons: Past, Present, and Future,* R. J. Murnane and K.-B. Liu, Eds., Columbia University Press, in preparation.

Lehmiller, G. S., T. B. Kimberlain, and J. B. Elsner, 1997: Seasonal prediction models for North Atlantic basin hurricane location. *Mon. Wea. Rev.,* **125,** 1780–1791.

Liu, K.-B., and M. L. Fearn, 2000: Reconstruction of prehistoric landfall frequencies of catastrophic hurricanes in northwestern Florida from lake sediment records. *Quat. Res.,* **54,** 238–245.

Ludlum, D. M., 1963: *Early American Hurricanes, 1492–1870.* Amer. Meteor. Soc., 198 pp.

Michaels, A., D. Malmquist, A. Knap, and A. Close, 1997: Climate science and insurance risk. *Nature,* **389,** 225–227.

Murnane, R. J., and Coauthors, 2000: Model estimates hurricane wind speed probabilities. *Eos, Trans. Amer. Geophys. Union,* **81,** 433, 438.

Neumann, C. J., B. R. Jarvinen, C. J. McAdie, and G. R. Hammer, 1999: *Tropical Cyclones of the North Atlantic Ocean, 1871–1998.* National Oceanic and Atmospheric Administration, 206 pp.

Parisi, F., and R. Lund, 2000: Seasonality and return periods of landfalling Atlantic basin hurricanes. *Aust. New Zeal. J. Stat.,* **42,** 271–282.

Pole, A., M. West, and J. Harrison, 1999: *Applied Bayesian Forecasting and Time Series Analysis.* Chapman and Hall, 409 pp.

Simpson, R. H., 1974: The hurricane disaster potential scale. *Weatherwise,* **27,** 169–186.

Solow, A. R., and L. Moore, 2000: Testing for a trend in a partially incomplete hurricane record. *J. Climate,* **13,** 3696–3699.

Walshaw, D., 2000: Modeling extreme wind speeds in regions prone to hurricanes. *Appl. Stat.,* **49,** 51–62.