J3.14 PREDICTIVE UNITED STATES' HURRICANE CLIMATE

James B. Elsner[1] and Brian H. Bossak
Florida State University, Tallahassee, Florida

1. INTRODUCTION

Hurricanes are of important social and economic concern in the United States. Strong winds and storm surge accompanying hurricanes over land kill people and destroy property. 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.

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 20th century. Moreover, less reliable but still useful information is available back through at least the second half of the 19th 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 past climate that not only tells a story about the past, but can be used to gauge future activity. This paper is an extended abstract of Elsner and Bossak (2001).[1]

2. DATA

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 data set consists of the six-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 non-traditional archives (Ludlum 1963). These studies update and add information about hurricane landfalls during the period 1851 through 1900. For instance, the New York Times' 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 pre-industrial era (Elsner and Kara 1999; Elsner et al. 2000). Recently, the National Oceanic and Atmospheric Administration (NOAA) embarked on a three-year 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 19th century. A concatenated data set consisting of landfalling hurricane accounts from historical archives and modern direct measurements is carefully analyzed here.

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 1901–1950 to a low of 1.4 during the period 1951–2000. Each subperiod has at least one year of six 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 re-samples, with replacement, the annual counts. The number of bootstrap samples is

---

[1] Corresponding author address: James B. Elsner, Florida State Univ., Geography Dept., Tallahassee, FL 32306-2190; e-mail: jelsner@garnet.fsu.edu.

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 re-sampling (number of bootstrap samples is 1000).

Years	total no.	mean	var	quantiles of the mean 5%	95%
		All U.S. hurricanes			
1851–00	88	1.7	2.35	1.42	2.14
1901–50	92	1.8	1.77	1.54	2.16
1951–00	72	1.4	1.56	1.16	1.72
		Major U.S. hurricanes			
1851–00	26	0.5	0.50	0.36	0.68
1901–50	36	0.7	0.70	0.54	0.92
1951–00	27	0.5	0.46	0.38	0.70

1000. The independence assumption is valid since the lag-1 temporal autocorrelation value is a negligible -0.002.

We test for differences in mean rates between the three periods using a Wilcoxon signed rank test. The p-values based on the large-sample approximation for the null hypothesis of no difference against the two-sided alternative 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 inter-annual activity decreases from 2.3 during the last half of the 19th century to 1.6 during the last half of the 20th century, indicating a potential bias during the earliest 50 year period. This is because information is measured in terms of precision, which is the inverse of the variance. Larger inter-annual variance (lower precision) during the 19th 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. 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 1901–1950. 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. 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.

3. 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 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. Here we follow the formalism presented in Epstein (1985).

3.1 Poisson Process

The arrival of hurricanes on the coast can be considered a Poisson process (Elsner et al. 2001; Solow and Moore 2000; Elsner and Kara 1999). The Poisson distribution is a limiting form of the binomial distribution with no upper bound on the count of occurrences. The parameter λ, the intensity, characterizes a Poisson process. Note in 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

$$f_{Poisson}(\hat{h}|\lambda, T) = \exp(-\lambda T)\frac{(\lambda T)^h}{h!}. \qquad (1)$$

The parameter λ and statistic T appear in the formula as a product, which is the mean and variance of the Poisson distribution. More importantly, 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 ten years of the record (1851–1860), observations indicate 15 U.S. hurricanes, so $T' = 10$ and $h' = 15$.

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

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

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

$$\Gamma(x) = \int_0^{\infty} t^{x-1}\mathrm{e}^{-t}\mathrm{d}t. \tag{3}$$

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.

3.2 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. Here, 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 1900. However, a bootstrap of the annual mean from the available observations over the period 1851–1899 indicate a 90% confidence interval of (1.45, 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 χ^2(0.05), where $\chi^2(0.05)$ 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 $\chi^2(0.95)$, where $\chi^2(0.95)$ 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. 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 1 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.

3.3 Predictive Distribution

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 of the annual rate is contained in the gamma density with

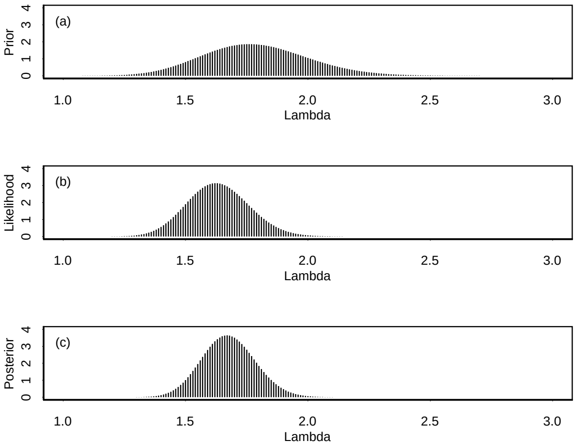

Figure 1: Gamma densities on the Poisson intensity λ (Lambda) for annual U.S. hurricane rates based on (a) prior, (b) likelihood, and (c) posterior parameters.

parameters h'' and T'' is a negative binomial distribution, with parameters h'' and $\frac{T''}{\hat{T}+T''}$ (see Epstein 1985)

The mean and variance of the negative binomial are $\hat{T}\frac{h''}{T''}$ and $\hat{T}\frac{h''}{T''}\left(\frac{\hat{T}+T''}{T''}\right)$, respectively. Note that the variance of the predictive distribution is always larger that 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 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 years then it is important.

4. 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. The start year of reliability is assigned based on U.S. census of “settled regions" defined as at least two inhabitants per square mile. 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 re-sampling 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 time scales.

Figure 2 show the predictive densities for coastal hurricanes and all major hurricanes. Here the reliable period begins in 1900. The top plots 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. The middle panels show the cumulative probability distributions of observing no more than a specified number of storms over the next 10 years. The bottom panels 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.

5. SUMMARY & CONCLUSIONS

Predictive climate distributions of U.S. landfalling hurricanes are estimated from availabe records. A Bayesian approach combines the reliable records of hurricane activity during the 20th century with the less precise accounts of activity during the 19th 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-year periods beginning in 1851. However, evidence indicates that the records during the 19th 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.

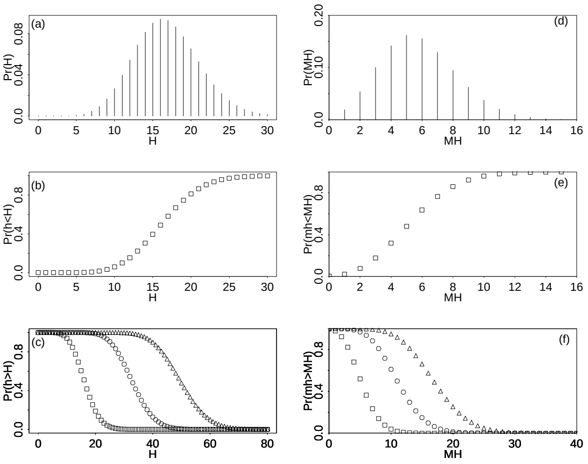

Figure 2: Predictive densities for the likelihood of U.S. hurricanes (H) and U.S. major hurricanes (MH) over the next 10 to 30 years. For U.S. hurricanes (a) is the probability of observing H hurricanes in 10 years, (b) is the cumulative probability of $h \leq H$ in 10 years, and (c) is the probability of $h \geq H$ in 10 (squares), 20 (circles), 30 (triangles) years. Plots (d–f) are the same, except for major U.S. hurricanes.

The main conclusions of this paper are:

- The statistics of the observed counts of U.S. hurricanes are similar across three consecutive 50-year periods beginning in 1851.
- Similar statistical distributions are noted across regional hurricane activity.
- Evidence suggests that hurricane records from the 19th 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 to 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.

Details concerning this analysis are provided in Elsner and Bossak (2001).

Acknowledgements. 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).

5. REFERENCES

- Efron, B., and R. J. Tibsharini, 1993: An Introduction to the Bootstrap. Chapman & Hall, 436 pp.
- Elsner, J. B., and A. B. Kara, 1999: Hurricanes of the North Atlantic: Climate and Society. Oxford University Press, 488 pp.
- Elsner, J. B., and B. H. Bossak, 2001: Baysian analysis of U.S. hurricane climate. J. Climate, in press.
- Elsner, J. B., K.-b. Liu, and B. Kocher, 2000: 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. 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, Meteorological Monographs, **20**, no. 42, Amer. Meteor. Soc., 199 pp.
- 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.
- Ludlum, D. M., 1963: Early American Hurricanes, 1492–1870. Amer. Meteor. Soc., 198 pp.
- Neumann, C. J., B. R. Jarvinen, C. J. McAdie, G. R. Hammer, 1999: Tropical Cyclones of the North Atlantic Ocean, 1871–1998. National Oceanic and Atmospheric Administration, 206 pp.
- Pole, A., M. West, and J. Harrison, 1999: Applied Bayesian Forecasting and Time Series Analysis. Chapman & Hall, 409 pp.
- Solow, A. R., and L. Moore, 2000: Testing for a trend in a partially incomplete hurricane record. J. Climate, **13**, 3696–3699.