Comparison of Hindcasts Anticipating the 2004 Florida Hurricane Season

JAMES B. ELSNER AND THOMAS H. JAGGER

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

(Manuscript received 12 January 2005, in final form 22 October 2005)

ABSTRACT

Advances in hurricane climate science allow forecasts of seasonal landfall activity to be made. The authors begin with a review of the forecast methods available in the literature. They then reformulate the methods using a Bayesian probabilistic approach. This allows a direct comparison to be made while focusing on a single hindcast of the 2004 season over Florida. The models, including climatology, are estimated using Gibbs sampling. Diagnostic checks verify convergence and efficient mixing of the samples from each of the models. A below average sea level pressure gradient over the eastern North Atlantic Ocean during May and June in combination with an above average tropospheric-averaged wind index associated, in part, with a strengthening of the Bermuda high pressure during July resulted in an above average probability of at least one Florida hurricane. The relatively high hindcast probabilities for 2004 were in marked contrast to the most recent 50-yr empirical probabilities for Florida, but fell short in anticipating the unprecedented level of activity that ensued. Similar results are obtained from hindcasts of total U.S. hurricane activity for 2004.

1. Introduction

Florida experienced an unprecedented number of hurricanes during the 2004 season. Hurricane Charley made landfall over southwestern Florida on 13 August with a maximum (1-min average at 10 m) wind speed estimated at 67 m s^{-1}. Hurricane Frances made landfall over southeastern Florida on 5 September with a maximum wind speed estimated at 46 m s^{-1}. A mere 21 days later, Hurricane Jeanne hit nearly the same location with winds estimated as fast as 54 m s^{-1}. Between Frances and Jeanne, Hurricane Ivan made landfall over the state of Alabama, with hurricane force winds extending over a significant portion of northwestern Florida. Thus, within a span of 6 weeks, Florida was impacted by four hurricanes with an estimated total damage of around $40 billion (U.S. dollars).

Forecasts in anticipation of the 2004 Atlantic hurricane season indicated above normal activity. By 1 June 2004 the Colorado State University (CSU) team led by W. M. Gray predicted eight hurricanes and the National Oceanic and Atmospheric Administration's (NOAA's) Climate Prediction Center predicted a 50% probability of an above normal hurricane season with only a 10% chance of a below normal season. Earlier, on 11 May 2004, the Benfield Hazard Research Center was indicating a 61% chance of an above normal hurricane season. Later, updated forecasts from these groups portended an even greater level of activity. The forecasts verified as the North Atlantic season saw nine hurricanes with six of them becoming major (having wind maxima in excess of 50 m s^{-1}).

The success of seasonal hurricane forecasts has encouraged attempts at greater geographic specificity. In particular, we are interested in forecasts of landfall activity. For example, how many hurricanes are likely to strike the United States in 2007? Skillful prediction of landfalling hurricane activity would benefit society and business through preparedness and insurance mechanisms. Although research and risk modeling groups have begun issuing hurricane landfall forecasts, there remains little peer-reviewed literature on methods and verification, which is in contrast to the extensive literature describing forecast algorithms of basin-wide activity (Gray et al. 1992, 1993, 1994; Elsner and Schmertmann 1993, 1994; Klotzbach and Gray 2003; Blake and Gray 2004). The exceptions include the work of Lehmiller et al. (1997), Elsner (2003), Elsner and Jagger (2004), and Saunders and Lea (2005). Here we aim to provide some coherence to the fledgling science and technology of seasonal landfall forecasts. The purpose

Corresponding author address: James B. Elsner, Dept. of Geography, The Florida State University, Tallahassee, FL 32306-2190.
E-mail: jelsner@garnet.fsu.edu

is twofold: 1) to compare the predictors used by the current suite of operational landfall forecasters, and 2) to continue our argument for a more probabilistic approach to the problem of seasonal hurricane prediction; an argument first articulated in Elsner and Bossak (2001) and refined in Elsner and Jagger (2004).

The present paper provides a comparison of hindcasts for the 2004 hurricane season made from current operational methods. By necessity the comparison is limited. First, the comparison is for a single season only. A comprehensive examination of the various methods would include many seasons. Second, the comparison is for Florida. Currently there are no operational seasonal hurricane models specifically for Florida. Instead, in this paper we compare the different forecast schemes by reinterpreting them as forecasts for Florida activity only. This might at first seem unfair. However, it can be argued that if a scheme is successful at forecasting entire coastal activity, then it is likely to be able to anticipate Florida activity. Another way to understand this is to consider the fact that over the period 1900–2002, inclusive, 63 of the 170 (37%) hurricanes to impact the United States did so in Florida. Third, we do not attempt to duplicate the forecast schemes. Instead, we choose a common probabilistic framework and recode the various forecast methods accordingly. In this way we provide a baseline for comparing what predictors were important in portending the Florida season of 2004.

As a consequence of the above limitations, the results presented below cannot be interpreted to mean one method is better than another at predicting Florida hurricane activity. Rather, the results provide some guidance toward improving the current suite of forecast techniques by identifying which predictors were useful in 2004. The results also argue for the utility of a Bayesian approach to seasonal hurricane modeling. The paper is organized as follows. In section 2 we provide a brief description of the forecast methodologies currently available in the scientific literature. In section 3 we expand this view to include the work of the CSU team and we categorize the methods according to type and predictors used. Because Bayesian inference is not standard practice in our discipline, in section 4 we describe the approach to predictive inference from hurricane count data. In section 5 we reformulate the methods in terms of Bayesian models for Florida counts and in section 6 we compare the hindcasts for the 2004 Florida hurricane season. Results show that a weaker than normal sea level pressure gradient over the eastern North Atlantic Ocean during May and June and an above normal tropospheric wind index during July—both associated with the strength and position of the Bermuda high pressure—indicated a higher than normal chance of at least one Florida hurricane. Hindcast probabilities for 2004 are in bold contrast to the most recent 56-yr historical probabilities. Similar results are obtained from hindcasts of total U.S. hurricane activity.

2. Seasonal landfall forecast methods in the literature

Only a few papers describing methods for seasonal forecasts of landfalling hurricane activity exist in the literature. Building on the work of Ballenzweig (1959) in elucidating climate controls for hurricane steering winds, Lehmiller et al. (1997) were the first to develop a skillful statistical model for seasonal forecasts of landfall probability. Their work revealed a statistically significant model for southeastern U.S. (Key Largo, Florida, to 35°N) hurricane activity that provides a conceptual advance in seasonal forecasts by focusing on factors that are conducive to hurricane activity regionally. Their model is a discriminant analysis that includes as predictors the previous autumnal rainfall within the Sahel region of western Africa; the forward-extrapolated vertical shear magnitude between 30- and 50-mb (1 mb = 1 hPa) tropospheric winds (quasi-biennial oscillation), the 700–200-mb vertical shear in the Miami–West Palm Beach, Florida area, the July monthly sea level pressure at Cape Hatteras, North Carolina, and the July average monthly East Coast sea level pressure. They note the highest likelihood of a hurricane landfall occurs with relatively high July sea level pressures over Cape Hatteras and high vertical shears over south Florida. Skill for this model has been demonstrated only in hindcasts.

Because regional hurricane probabilities are small, it is important to use the longest available records for statistical models. Toward this end, a regression model for the southeastern United States that makes use of longer records describing the El Niño–Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO), rather than the more temporally limited upper-air sounding data, is designed in Elsner (2003). The model is a Poisson regression that uses August–October-averaged values of the Southern Oscillation index (SOI) together with May–June-averaged values of the NAO index to predict the probability of hurricanes from Texas through South Carolina. The importance of ENSO on U.S. hurricane activity is elucidated in Bove et al. (1998) and the importance of the NAO is first suggested in Liu and Fearn (2000) and Elsner et al.

(2000). Poisson regression has been successfully applied in modeling tropical cyclone activity over the Atlantic basin (Elsner and Schmertmann 1993) and elsewhere (McDonnell and Holbrook 2004). The model indicates that southeast U.S. hurricanes are more likely during La Niña conditions when the NAO is weak (Jagger et al. 2001; Elsner and Bossak 2004). ENSO is also used as a predictor in a seasonal prediction model of landfalling tropical cyclone activity along the southern China coast (Liu and Chan 2003).

Saunders and Lea (2005) offer a statistical model of total U.S. landfalling activity. They show that hurricane wind energy along the coast is predictable from 1 August using tropospheric height-averaged wind anomalies (wind index) over the North Atlantic, North America, and eastern Pacific Ocean during July. They use ordinary least squares (OLS) regression to regress normally transformed accumulated cyclone energy (ACE) onto the July averaged wind index. It is worth noting that although the height-averaged wind index from July used by Saunders and Lea (2005) is different from what was used by Lehmiller et al. (1997), it captures similar information related to the position and strength of the Bermuda high pressure system (as does the NAO). In fact, Saunders and Lea (2005) verified that the highest likelihood of southeast U.S. hurricane landfall occurs with high values of July-averaged sea level pressures over Cape Hatteras.

3. Categorization of forecast methods

The methods used for seasonal landfall forecasts can be grouped in different ways. One way to group the forecast methods is by predictors. A climatology model used as a benchmark for assessing model skill contains no predictors. In contrast, the CSU team, which now issues forecasts of landfall probabilities along with their forecasts of overall activity, uses a prediction of net tropical cyclone activity together with a measure of North Atlantic sea surface temperature (SST) anomalies as predictors. Net tropical cyclone activity (NTC) is a combined measure of six indices of hurricane activity each expressed as a percentage difference from the long-term average. The model of Elsner (2003) uses a prediction of the August–October-averaged SOI and the preseason May–June-averaged NAO index as predictors. The NAO is a variation in the sea level pressure gradient over the eastern North Atlantic (Hurrell et al. 2003). Under normal conditions pressures are low over Iceland and high over Gibraltar. Variations in this north–south gradient related to the position of the subtropical high pressure ridge have been linked to a greater threat of hurricane landfall (Jagger et al. 2001).

The model of Saunders and Lea (2005) uses a July wind index constructed as a linear combination of six vertical- and areally averaged wind anomalies from regions extending across the eastern equatorial Pacific through the North Atlantic basin. In five of the regions zonal wind is used and in the other region meridional wind is used. Vertical averaging extends from 925 to 400 mb, and the time averaging is for the month of July.

Another way to group the forecast methods is by methodology. The methodology used by the CSU team is empirical. Their landfall forecasts are based on a formula that combines a forecast of NTC activity with a measure of the North Atlantic SST anomaly (SSTA). Though not available in the peer-reviewed literature, their formula for landfall "probability" is

$$\text{landfall “probability”} = \text{forecast NTC} + \text{measured SSTA}. \tag{1}$$

Quotes are needed on the word probability because the value obtained by using this empirical methodology is not, strictly speaking, a probability.

The forecast methods of Elsner (2003) and Saunders and Lea (2005) described above are based on statistical models. In the Elsner (2003) model, the predictors are related to the logarithm of the mean hurricane rate. Values for the model coefficients are found using the method of maximum likelihood. In the Saunders and Lea (2005) model, the predictor is linearly related to the mean of the transformed values of ACE and the values for the model coefficients are found using the method of OLS. In forecasting the unknown future, statistical models have an advantage over empirical models as forecasts can be expressed in terms of uncertainty (e.g., prediction intervals).

Another approach is Bayesian modeling. Here, both the model coefficients and the observed data are treated as random variables and forecasts are expressed in terms of a posterior distribution. Bayesian models have an advantage over classical statistical models as the posterior distribution contains all the information about predictive uncertainty (not only the prediction intervals). Elsner and Jagger (2004) describe a Bayesian regression model and explain its advantage for coastal hurricane count data.

Bayesian models are implemented naturally using Markov chain Monte Carlo (MCMC) methods such as the Gibbs sampler with distributions derived from an application of Bayes's rule. The Gibbs sampler generates empirical distributions from a model using random samples from appropriate marginal and conditional distributions. Suppose the model defines probability distributions $F(X|Y)$ and $G(Y|X)$, where the vertical line is read "conditional on." The sampler starts with a random set of possible values for X, then draws a set of Y values from $G()$, which are subsequently used to update the values for X from $F()$, and so on.

Bayesian methods are applied in the area of climate change and detection (Solow 1988; LeRoy 1998; Berliner et al. 2000). An argument in support of the Bayesian approach to predictive inference is the ability to incorporate older, less reliable data. Since regional hurricane probabilities are quite small, longer records are needed to accurately assess the risk of future storm occurrences. For practical reasons that become apparent below, we adopt a Bayesian approach to modeling hurricane counts. The approach allows us to reformulate the different prediction schemes using a common framework that then makes comparison of the hindcast results straightforward. Here, the results take the form of the expected value of the predictive probability distribution for the number of hurricanes over Florida during 2004.

4. Bayesian inference for hurricane counts

The canonical model for event count data is the Poisson regression model. The model is used extensively in hurricane climate studies (Elsner and Schmertmann 1993; Solow and Moore 2000; Elsner et al. 2001; Jagger et al. 2002; Elsner 2003; McDonnell and Holbrook 2004). It is based on the Poisson distribution, which is a discrete distribution defined on the nonnegative integers. It can be derived from the distribution of waiting times between successive events. For our purposes, the Poisson regression model is used to model a set of Florida hurricane counts $h_i \in 0, 1, 2, \ldots, \infty = Z^+$ on the nonnegative integers for a set of observed years $i = 1, \ldots, N$.

Additionally, we observe a row vector of predictor variables x_i' with dimensionality $(1 \times J)$. Thus, the Poisson regression is

$$\begin{aligned} &h_i \sim \text{Poisson}(\lambda_i), \\ &\lambda_i = \exp(\beta_0 + x_i'\beta + \eta_i), \quad \text{and} \\ &\eta_i \sim \text{Normal}(0, \tau^{-1}), \end{aligned} \tag{2}$$

where λ_i is the hurricane rate for year i, β_0 is the intercept, and η_i is a random effect that has zero mean (does not contribute to the count) but adds to the variance of the counts. The symbol $\sim$ refers to a stochastic relationship and indicates that the node on the left-hand side is a sample from a distribution specified on the right-hand side. The equal sign indicates a logical relationship with the node on the left-hand side algebraically related to terms on the right-hand side.

The scalar τ captures the precision (inverse variance) of the random effect, thus indicating how much (or little) overdispersion there is (Martin 2003). Component values for the offset and parameter vector β define the specific model and are estimated using a Bayesian approach. In short, with the Bayesian approach, we assume that the parameters have a distribution and inference is made by computing the posterior probability density of the parameters conditioned on the observed data. Alternatively, in the frequentist (or classical) approach, we assume the parameters are fixed, but unknown, and we find values for the parameters most likely to have generated the observed data by maximizing the likelihood. The Bayesian approach combines the frequentist likelihood with our prior belief $f(\beta)$ using Bayes's rule so that, for the Poisson regression model, we are interested in

$$f(\beta|h) \propto f(h|\beta)f(\beta). \tag{3}$$

The distribution $f(\beta|h)$ is the posterior density indicating the probability of β conditional on the observed hurricane counts. The posterior density summarizes what we believe about the parameter values after considering the observed counts. For example, sample averages taken from the distribution approximate the posterior expectation of the parameter value. Importantly, the posterior density allows us to make probability statements, including those about whether a particular parameter value differs from zero.

In general, the posterior density $f(\beta|h)$ has no analytical solution so MCMC sampling methods are used to simulate it (Geman and Geman 1984; Gelfand and Smith 1990). Importantly, the level of precision on the posterior density can be made as high as desired by increasing the number of samples. However, Bayesian approaches require the user to formally specify prior beliefs. Here, we take the standard route and assume noninformative priors that, as the name implies, provide little information about the value of the parameters of interest. Thus, to finish the Bayesian specification of the Poisson regression model, we assume independent normally distributed priors for each component of the β vector, where the mean of the normal distribution is zero and the variance is 10^6 [$\beta_j \sim$ Normal$(0, 10^6)$]. This is a flat distribution that contributes little information. [Alternatively, we could use the dflat() distribution; see the discussion of the Bayesian inference using Gibbs sampler (BUGS) below.] We also assume a gamma distributed prior for the precision parameter [$\tau \sim G(0.001, 0.001)$], which has its mass just to the right of zero and is decreasing on R^+. The gamma distribution is often used for the precision parameter as it is bounded on the left by zero and small values reflect a lack of prior information available for this parameter (large variance). The influence of these prior specifications on results is examined in section 6.

5. Reformulation of the landfall forecast methods

The above Bayesian specification mixing the Poisson with the normal distribution together with the prior specifications provides a framework for comparing the disparate forecast methods. The models are listed below and for completeness we include a climatological forecast:

$$
\begin{aligned}
h_i &\sim \mathrm{Poisson}(\lambda_i),\\
\text{model 0: climatology, where } \lambda_i &= \exp(\beta_0 + \eta_i),\\
\text{model 1: NTC and SSTA, where } \lambda_i &= \mathrm{NTC} \times \exp(\beta_0 + \beta_1 \times \mathrm{SSTA} + \eta_i),\\
\text{model 2: wind index, where } \lambda_i &= \exp(\beta_0 + \beta_1 \times \mathrm{WIND} + \eta_i),\\
\text{model 3: NAO, where } \lambda_i &= \exp(\beta_0 + \beta_1 \times \mathrm{NAO} + \beta_2 \times \mathrm{EPOCH} + \eta_i), \quad \text{and}\\
\eta_i &\sim \mathrm{Normal}(0, \tau^{-1}).
\end{aligned}
\tag{4}
$$

In Eq. (4) WIND is the tropospheric height–averaged wind anomoly and EPOCH is a term that accounts for changes in the level of uncertainty in the hurricane record over time (see below). The hurricane count data for Florida are obtained from the National Hurricane Center's hurricane best-track file (Neumann et al. 1999). A Florida hurricane is defined as a tropical cyclone that makes at least one landfall in the state. 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. Similarly, a U.S. hurricane is a tropical cyclone that makes at least one landfall somewhere in the United States (excluding Hawaii).

In reformulating the various forecast methods, the above models differ only in regression structure. For example, a climatological forecast (model 0) excludes predictors. Model 0 is available for predictions at any time during the year. Model 1, which represents the empirical method of the CSU team, has two predictors: NTC and SSTA, where NTC is the predicted value for the season and the SSTA is the current observed value (1 August 2004 here). Given a value for the NTC, the total number of Florida hurricanes is constrained (e.g., there cannot be more Florida hurricanes than total hurricanes); thus, we treat the logarithm of NTC as an offset. SSTAs represent those of the Atlantic multidecadal oscillation (AMO) for July. The AMO is characterized by fluctuations in SSTs over the North Atlantic Ocean driven largely by the thermohaline circulation (Goldenberg et al. 2001). The Hadley model SST and NOAA optimal interpolated SST data are used to compute Atlantic SSTAs north of the equator (Enfield et al. 2001). SSTAs are computed by month using the climatological time period 1951–2000. Data are obtained online from the NOAA–Cooperative Institute for Research in the Environmental Sciences (CIRES) Climate Diagnostics Center (CDC). Although not considered in this study, hurricane activity along the southeast coast of the United States has recently been correlated with the dipole mode of tropical Atlantic Ocean SST (Xie et al. 2005).

We note that for the present comparison, predicted values for the NTC were not available to us. Instead, we use the observed value of the number of hurricanes. Because the number of hurricanes h is part of the calculation of NTC, there is a correlation between NTC and h. However, because we used the observed h rather than a predicted h (except for 2004), model 1 is biased toward greater skill than can be expected in actual forecast situations. Assuming that a forecast of NTC can be made and that SSTAs are slowly varying, model 1 is available for predictions well in advance of the hurricane season.

Model 2, which represents the statistical formulation of Saunders and Lea (2005), uses a single predictor, WIND, which is the tropospheric height-averaged wind anomalies (wind index) over the North Atlantic, North America, and eastern Pacific during July. However, instead of forecasting normally transformed landfall ACE along the U.S. coast, we reformulate the model to forecast hurricane count probability over Florida. The u- and v-component wind data over the period 1948–2004 for the tropospheric levels 925–400 mb used to construct the index are from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis project (Kalnay et al. 1996) and are obtained online from CDC. Model 2 is available for predictions by 1 August.

Model 3 represents the statistical formulation of Elsner (2003) and Elsner and Jagger (2004) without the predicted value for the SOI. Here, we use preseason values (May–June average) of the NAO index from the period 1851–2004 (Jones et al. 1997) as the lone climate predictor. NAO values are obtained online from the Climatic Research Unit. The model also differs from the Elsner (2003) specification in that it includes the term EPOCH that effectively assigns greater uncertainty to the counts prior to 1900. This is reasonable as we are more uncertain about the actual annual count of hurricanes over Florida prior to this time (Landsea et al. 2004). Herein lies a distinct advantage of a Bayesian approach to seasonal landfall modeling. The models can quite naturally incorporate older, less reliable, data by specifying values as less precise. Model 3 is available for predictions by 1 July.

In summary, we reformulate four forecast methods into separate Bayesian models that use Gibbs sampling to determine the probability of hurricanes over Florida during 2004. The idea is straightforward. If we are transported back in time to the end of July 2004 and asked to forecast the probability of hurricane activity for Florida during the upcoming season, what scientific understanding do we have at our disposal and what forecasts would we make? It should be reiterated that we are not providing a comprehensive comparison of the available methods. This would require data that are not available to us (e.g., predicted NTC) and it would require a way to judge forecast skill across different predictands (landfall probability, ACE, probability of H number of hurricanes). Instead we opt for a limited comparison that has the advantage of making a comparable set of probability forecasts across the methods. The focus on the 2004 season is instructive, as a strict empirical climatological forecast would have put the probability at 0 of observing four Florida hurricanes, as there is no precedence for this event in the historical record. While it has been suggested that the event was simply random, we feel that this position is scientifically unsatisfying.

6. Hindcasts of hurricane activity for 2004

Posterior predictive probabilities for h Florida hurricanes, where $h = 1$, 2, 3, and 4 are generated for each of the above models using the free BUGS software developed at the Medical Research Council in the United Kingdom (Gilks et al. 1994; Spiegelhalter and Thomas 1998). All models are estimated using WinBUGS, version 1.4. The startup cost of Bayesian modeling is minimized by BUGS by eliminating the need to program in a high-level language like FORTRAN or Splus. It also chooses an appropriate MCMC sampling algorithm based on the model structure. The WinBUGS code for model 3 is given in the appendix.

As a diagnostic check on each of the four models, we graphically and statistically check successive sample values. Figures 1a and 1b show 5000 samples of β_1 for model 3 using two different sets of initial conditions. Samples from the posterior distribution of β_1 indicate quick settling (convergence) as both sets of initial conditions result in samples that fluctuate around a fixed median value. Convergence is important as it implies that the model is not sensitive to the prior specifications. Also, small autocorrelation values for lags greater than four samples (Figs. 1c and 1d) indicate that the successive samples are reasonably effective at moving through the posterior distribution (mixing) rather than getting stuck in one part of the distribution or another. Posterior densities (Figs. 1e and 1f) estimated from samples 2001–5000 are smooth and the expected value of β_1 is -0.1864 over these later 3000 samples, regardless of initial conditions. Other models show similar convergence and mixing properties so we are comfortable with the robustness of the procedure.

We compare model hindcasts of the exceedence probability (probability that Florida will be hit by at least k hurricanes) in 2004 by generating 10^5 samples and counting the proportion of times h_{2004} equals or exceeds k number of hurricanes after ignoring the first 2000 samples as "burn-in." Burn-in is the descriptive term used in the Bayesian literature that refers to the first set of samples that may contain information about the initial conditions and therefore are discarded from the analysis. The simplest model (model 0) took 132 s to generate all samples and the most complex model (model 3) took 549 s on a Xeon$^{\text{TM}}$ 3.4-GHz processor. The expected value of the posterior density for β_1 is -0.1376 in model 1, $+0.1193$ in model 2, and -0.1997 in model 3. The probability that β_1 is greater than 0 in model 1 is 0.3109 (SSTA term), the probability that it is less than 0 in model 2 is 0.0005 (WIND term), and the probability that it is greater than 0 in model 3 is 0.0308 (NAO term). Hindcast results are shown in Table 1.

Examining the first row of values, the models predict varying probabilities of at least one Florida hurricane in the 2004 season. Model 0 forecasts the lowest probability at 37% and model 3 forecasts the highest probability at 47%. Examining other rows, model 2 predicts the highest probability, 3.6%, of observing at least three

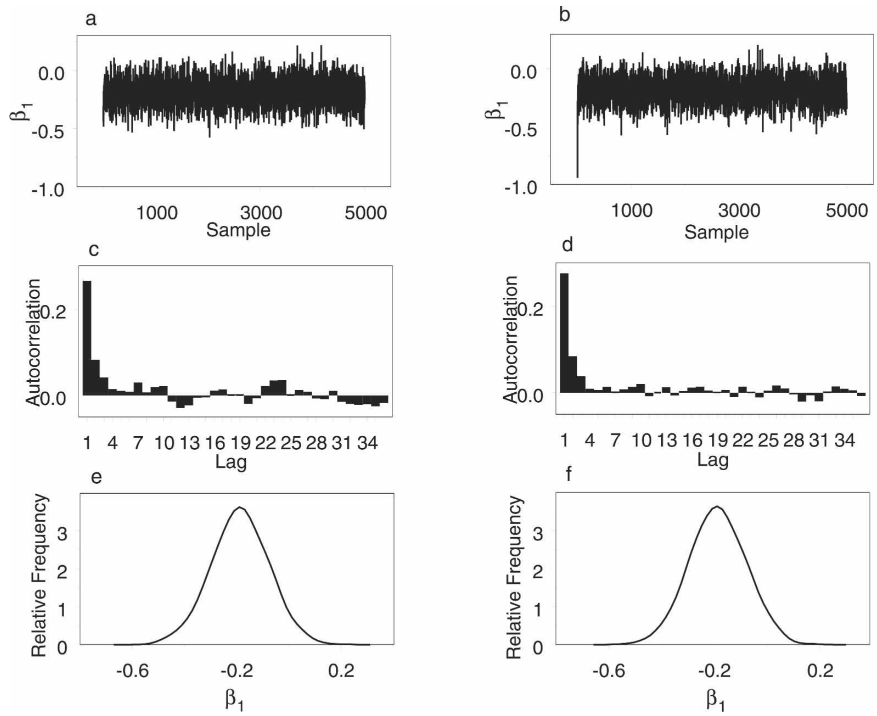

FIG. 1. Diagnostic plots of the β_1 parameter for model 3. Successive sample values starting with (a) $\beta_1 = 0$ and (b) $\beta_1 = 2$. (c), (d) Corresponding autocorrelation functions of the sample history. (e), (f) Corresponding kernel density estimates of the posterior distributions of β_1 for the two different sets of initial conditions.

hurricanes, which compares with 2.1% for model 0, 2.7% for model 1, and 3.1% for model 3. Model 2 also predicts the highest probability of any of the models for what actually occurred (four or more). Note that models 1–3 indicate an increased risk (over climatology) of a Florida hurricane in 2004. As previously noted a naïve empirical probability forecast would have given Florida a 0% chance of having four hurricanes. Another comparison is made in Fig. 2, which shows bar graphs of the 2004 forecast probabilities from each of the models. Models 1–3 show an improvement over climatology with model 2 portending the greatest risk of an extreme season ($h \geq 4$), albeit with probabilities only near 1%. The results clearly show that although the models indicated an above average probability of a Florida hurricane, they all fell short in anticipating the extremely active year that ensued.

TABLE 1. Model comparisons. Listed are the hindcast exceedence probabilities of hurricane activity over FL for the 2004 season. Model 0 is Bayesian climatology, model 1 is a reformulation of the CSU team, model 2 is a reformulation of Saunders and Lea (2005), and model 3 is a reformulation of Elsner (2003).

Exceedance probability	Model 0	Model 1	Model 2	Model 3	Empirical climatology
$\Pr(h \geq 1)$	0.3656	0.3889	0.4421	0.4698	0.3750
$\Pr(h \geq 2)$	0.0941	0.1071	0.1359	0.1399	0.1250
$\Pr(h \geq 3)$	0.0216	0.0266	0.0360	0.0309	0.0178
$\Pr(h \geq 4)$	0.0061	0.0075	0.0102	0.0055	0.0000

For comparison, we reran each of the models this time predicting the probability of hurricanes along the entire U.S. coast. There were six U.S. hurricanes in 2004. Similar to the Florida results we find all three of the models predicting higher than climatologically averaged probabilities (Fig. 3). In this case, model 2 outperforms both models 1 and 3, predicting a 13.4% chance of four or more U.S. hurricanes compared with 9.5% for model 1 and 10.1% for model 3. Similar to the results from Florida, the models indicated an increased likelihood of an above normal U.S. hurricane season, but they fell short in anticipating the extremely active year that ensued.

7. Discussion and conclusions

As early as May of 2004, there were indications that Florida might be in for an active hurricane season. The May–June-averaged NAO index value was −0.52 standard deviations, putting it in the second quartile of the distribution (1851–2003). While not extreme, it did in-

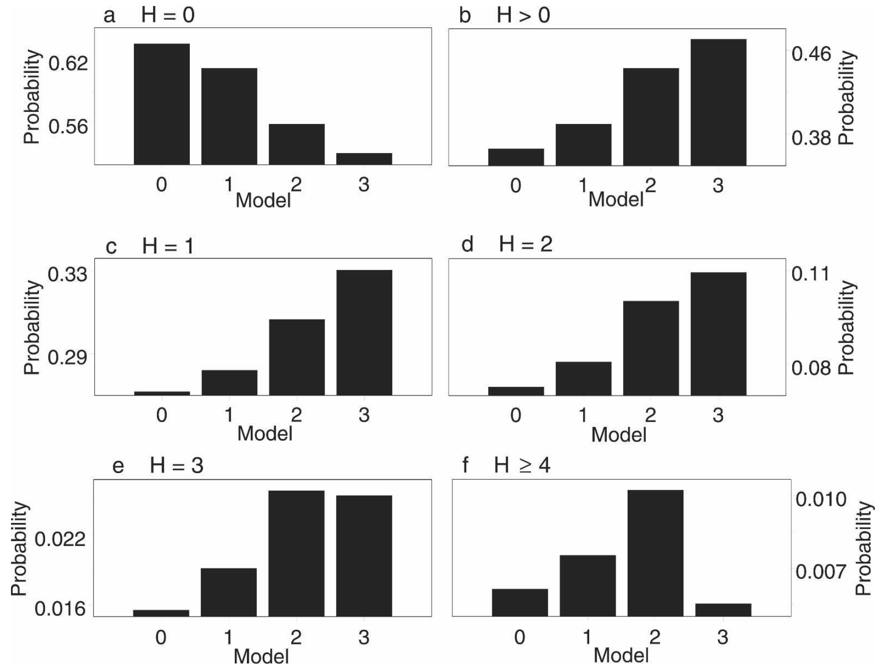

FIG. 2. Forecast probabilities for hurricane activity over Florida during 2004 by model type: Probability of (a) no hurricanes, (b) at least one hurricane, (c) exactly one hurricane, (d) exactly two hurricanes, (e) exactly three hurricanes, and (f) four or more hurricanes.

dicate a heightened probability of Florida hurricanes. Model 3, available for predictions by 1 July, would have predicted a 47% chance of at least one hurricane strike to Florida. This compares with a 37% chance based on climatology (1948–2003) alone. Moreover, the model predicted a 14% chance of two or more Florida hurricanes. This is 5% above climatology. Signals suggesting an active Florida season continued into July. By 1 August, models 1 and 2 would have predicted an above normal season. Model 1, combining a prediction of above average hurricane activity (forecast of seven hurricanes made on 6 August 2004) with observed warmer than normal Atlantic SSTs (AMO), would have predicted an 11% chance of two or more Florida hurricanes in 2004. Similarly, model 2, using a July wind index, would have predicted a 13% chance of two or more hurricanes.

Understanding regional hurricane variation requires not only understanding hurricane origin and development mechanisms, but also what influences where they will track. In this regard, the NAO and the position of the subtropical high are the leading candidates. The July wind index developed in Saunders and Lea (2005) captures, to some extent, the strength and position of the subtropical Bermuda high pressure. Tropical cyclones that form and remain equatorward of the subtropical high tend to intensify at low latitudes en route to Florida. Other factors that are likely relevant to seasonal forecasts of regional hurricane risk include the ENSO (also captured in the wind index to some extent) and Atlantic SSTs.

The physical relationship between the springtime NAO and summer/autumn hurricanes is the subject of ongoing research. We speculate that it might be related to how the NAO is related to the summer climate over the continents of Europe and North America. A weak NAO during May–June is associated with weaker midlatitude weather systems (and thus less rainfall) over Europe during spring. This creates a feedback (perhaps through soil moisture) with a tendency for greater summer/autumn middle-tropospheric ridging and an enhancement of the dry conditions. Ridging over the eastern and western sides of the North Atlantic basin during the hurricanes season tends to keep the middle-tropospheric trough, responsible for hurricane recurvature, farther to the north.

Evidence in support of this hypothesis comes from examining the correlation between the May–June NAO index and 500-hPa heights over western Europe (35°–45°N, 20°W–0°) averaged over August–October for the period 1948–2004. Using NCEP–NCAR reanalysis data (Kalnay et al. 1996), we find a linear correlation of −0.37, indicating a significant relationship whereby ridging (high heights) occurs with low values of the NAO. The ridging is associated with weaker upper-level winds and less vertical wind shear. Additional evi-

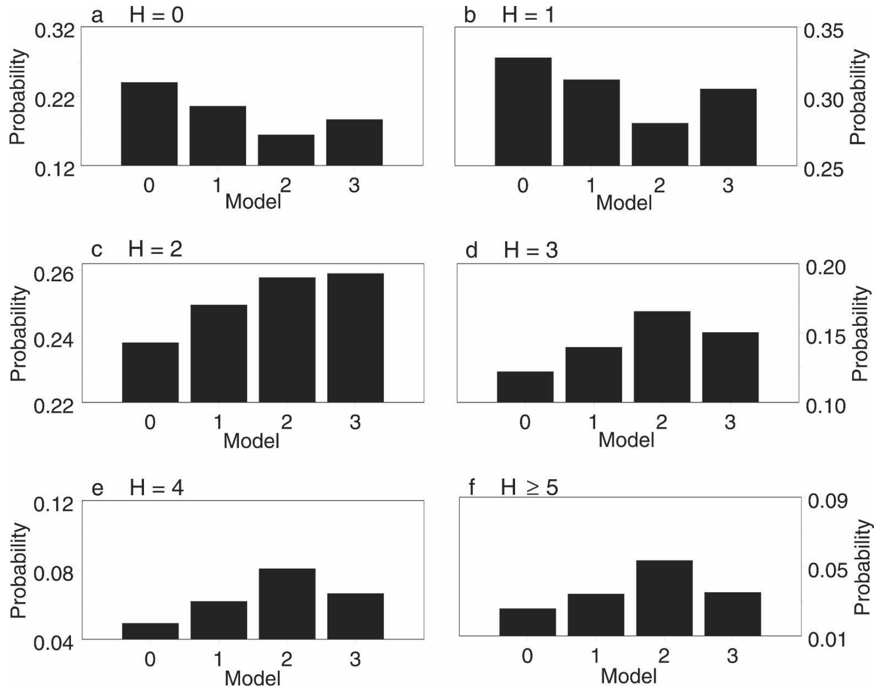

FIG. 3. Forecast probabilities for hurricane activity along the entire U.S. coast during 2004 by model type: Probability of (a) no hurricanes, (b) exactly one hurricane, (c) exactly two hurricanes, (d) exactly three hurricanes, (e) exactly four hurricanes, and (f) five or more hurricanes.

dence comes from a composite difference map of mean August–October precipitation rates for the five lowest and five highest (lowest minus highest) May–June NAO years (Fig. 4). Here, we see a teleconnection between a wet southeastern United States and dry conditions over both western Europe and over the mid-Atlantic extending to New England. Based on the feedbacks involved in maintaining the ridging, we suggest that it might be possible to find a longer lead relationship between the NAO and U.S. hurricanes. Moreover, the dry conditions northeast of the Greater Antilles are consistent with a persistence of a southwesterly displaced subtropical high. As a postscript, we note that, like in 2004, the 2005 value of the NAO was negative (−0.56 standard deviations) indicating an increased likelihood of Florida hurricane activity.

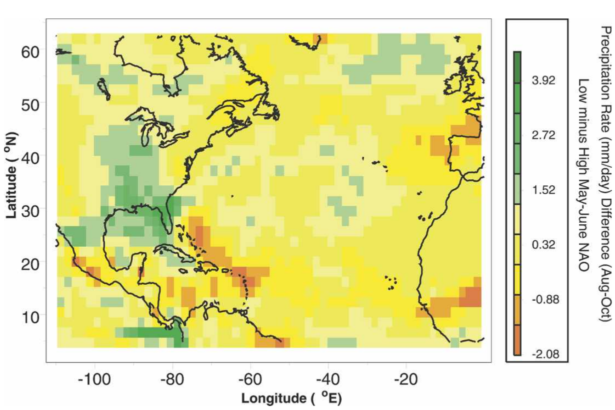

FIG. 4. Composite difference in August–October precipitation rates for the five years of lowest (and highest) May–June NAO values (low minus high). Data period is 1948–2004. Data are courtesy of CDC.

Forecast models of landfall activity will likely improve by taking into account relevant spatial information. A step in this direction is taken in Jagger et al. (2002). The model is a statistical space–time specification based on a truncated Poisson count process that includes neighborhood response values (hurricane counts in adjacent grid boxes), local offsets, and climate variables as predictors. The climate variables include a factor for the state of ENSO; rainfall over Dakar, Senegal; and sea level pressures (SLPs) over the Azores and over Iceland. The SLP variables indicate the state of the NAO. Although this model has yet to be implemented operationally, the aim is to have it predict the most likely near-coastal paths of hurricanes prior to the start of the season. Toward this end, it should be possible to reformulate the Jagger model using a Bayesian approach.

Seasonal landfall forecasting is relatively new. More work is needed to understand the physical mechanisms responsible for the frequency of particular storm tracks. Here, we have summarized and compared the relevant studies identifying the climate factors important for seasonal forecasts of regional hurricane activity. In doing so, we have described a methodology for building useful seasonal forecast models. The topic is relevant to business, government, and society. In fact, risk modeling companies are beginning to offer products that make use of the science and technology of landfall forecasting.

Acknowledgments. The model code is available upon request. The NCEP–NCAR reanalysis data are provided by NOAA–CIRES, Climate Diagnostics Center, Boulder, Colorado. Partial support for this study was provided by the National Science Foundation (BCS-0213980 and ATM-0435628) and the Risk Prediction Initiative (RPI-05001). The views expressed within are those of the authors and do not reflect those of the funding agencies.

APPENDIX

WinBUGS Code for Model 3

```
#Model
model; {
  for(i in 1: N) {
  h[i] ~ dpois(lambda[i]) #Hurricane
    counts as Poisson
  log(lambda[i]) <- beta0 + beta1 *
    NAO[i] + beta2 * epoch[i] + eta[i]
    #Likelihood
  tau1[i]<-tau[ind[i] + 1] #Precision on
    the counts
  eta[i] ~ dnorm(0, tau1[i]) #Random
    effect
  }
tau[1] ~ dgamma(0.001, 0.001)
tau[2] ~ dgamma(0.001, 0.001)
beta0 ~ dnorm(0.0, 1.0E-6) #Fixed effect
beta1 ~ dnorm(0.0, 1.0E-6) #Coefficient
beta2<-log(p)
p ~ dunif(lower, upper) # Probability of
    observing a hurricane, used when
    epoch=1
  prob1<-step(h[N]-1) # probability of
    at least 1 hurricane
  prob2<-step(h[N]-2) # probability of
    at least 2 hurricanes
  prob3<-step(h[N]-3) # probability of
    at least 3 hurricanes
  prob4<-step(h[N]-4) # probability of
    at least 4 hurricanes
}
# Initial Conditions
list(beta0=0, beta1=0, tau=c(0.1, 0.1),
    p=0.9)
# Data
list(N=154, lower=0.80, upper=0.95,
# Florida hurricanes by season, the NA
    (not available) is for the 2004 season
h=c(1, 2, . . . , 0, 0, 0, NA)
# NAO index values averaged over the
    preseason months of May-June
NAO=c(-1.575, 0.170, . . . , -0.710,
    0.425, 0.155, -0.525),
# Epoch variable for reliable (0) versus
    less reliable records (1)
epoch=c(1, 1, . . . , 0, 0, 0, 0)
)
```

REFERENCES

Ballenzweig, E. M., 1959: Relation of long-period circulation anomalies to tropical storm formation and motion. *J. Meteor.*, **16,** 121–139.

Berliner, L., R. Levine, and D. Shea, 2000: Bayesian climate change assessment. *J. Climate,* **13,** 3805–3820.

Blake, E. S., and W. M. Gray, 2004: Prediction of August Atlantic basin hurricane activity. *Wea. Forecasting,* **19,** 1044–1060.

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.

Elsner, J. B., 2003: Tracking hurricanes. *Bull. Amer. Meteor. Soc.,* **84,** 353–356.

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

——, and ——, 1994: Assessing forecast skill through cross validation. *Wea. Forecasting,* **9,** 619–624.

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

——, and ——, 2004: Hurricane landfall probability and climate. *Hurricanes and Typhoons: Past, Present, and Future,* R. J. Murnane, and K.-b. Liu, Eds., Columbia University Press, 333–353.

——, and T. H. Jagger, 2004: A hierarchical Bayesian approach to seasonal hurricane modeling. *J. Climate,* **17,** 2813–2827.

——, 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.

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

Enfield, D. B., A. M. Mestas-Nunez, and P. J. Trimble, 2001: The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S. *Geophys. Res. Lett.,* **28,** 2077–2080.

Gelfand, A. E., and A. F. M. Smith, 1990: Sampling-based approaches to calculating marginal densities. *J. Amer. Stat. Assoc.,* **85,** 398–409.

Geman, S., and D. Geman, 1984: Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.,* **6,** 721–741.

Gilks, W. R., A. Thomas, and D. J. Spiegelhalter, 1994: A language and program for complex Bayesian modelling. *Statistician,* **43,** 169–178.

Goldenberg, S. B., C. W. Landsea, A. M. Mestas-Nuñez, and W. M. Gray, 2001: The recent increase in Atlantic hurricane activity: Causes and implications. *Science,* **239,** 474–479.

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.

——, ——, ——, and ——, 1993: Predicting Atlantic basin seasonal tropical cyclone activity by 1 August. *Wea. Forecasting,* **8,** 73–86.

——, ——, ——, and ——, 1994: Predicting Atlantic basin seasonal tropical cyclone activity by 1 June. *Wea. Forecasting,* **9,** 103–115.

Hurrell, J. W., Y. Kushnir, G. Ottersen, and M. Visbeck, 2003: An overview of the North Atlantic Oscillation. *The North Atlantic Oscillation: Climate Significance and Environmental Impact, Geophys. Monogr.,* No. 134, Amer. Geophys. Union, 1–36.

Jagger, T. H., 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.

——, X. Niu, and J. B. Elsner, 2002: A space–time model for seasonal hurricane prediction. *Int. J. Climatol.,* **22,** 451–465.

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.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Klotzbach, P. J., and W. M. Gray, 2003: Forecasting September Atlantic basin tropical cyclone activity. *Wea. Forecasting,* **18,** 1109–1128.

Landsea, C. W., C. Anderson, N. Charles, G. Clark, J. Dunion, J. Fernández-Partagás, P. Hungerford, C. Neumann, and M. Zimmer, 2004: The Atlantic hurricane database re-analysis project documentation for 1851–1910: Alterations and additions to the HURDAT database. *Hurricanes and Typhoons: Past, Present, and Future,* R. J. Murnane and K.-b. Liu, Eds., Columbia University Press, 177–221.

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.

LeRoy, S., 1998: Detecting climate signals: Some Bayesian aspects. *J. Climate,* **11,** 640–651.

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

Liu, K.-S., and J. C. L. Chan, 2003: Climatological characteristics and seasonal forecasting of tropical cyclones making landfall along the South China coast. *Mon. Wea. Rev.,* **131,** 1650–1662.

Martin, A. D., 2003: Bayesian inference for heterogeneous event counts. *Sociol. Methods Res.,* **32,** 30–63.

McDonnell, K. A., and N. J. Holbrook, 2004: A Poisson regression model of tropical cyclogenesis for the Australian–southwest Pacific Ocean region. *Wea. Forecasting,* **19,** 440–455.

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.

Saunders, M. A., and A. S. Lea, 2005: Seasonal prediction of U.S. landfalling hurricane activity from 1 August. *Nature,* **434,** 1005–1008.

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

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

Spiegelhalter, D. J., and A. Thomas, 1998: Graphical modeling for complex stochastic systems: The BUGS project. *IEEE Intell. Syst. Appl.,* **13,** 14–15.

Xie, L., T. Yan, and L. Pietrafesa, 2005: The effect of Atlantic sea surface temperature dipole mode on hurricanes: Implications for the 2004 Atlantic hurricane season. *Geophys. Res. Lett.,* **32,** L03701, doi:10.1029/2004GL021702.