Prediction Models for Annual U.S. Hurricane Counts

JAMES B. ELSNER AND THOMAS H. JAGGER

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

ABSTRACT

The authors build on their efforts to understand and predict coastal hurricane activity by developing statistical seasonal forecast models that can be used operationally. The modeling strategy uses May–June averaged values representing the North Atlantic Oscillation (NAO), the Southern Oscillation index (SOI), and the Atlantic multidecadal oscillation to predict the probabilities of observing U.S. hurricanes in the months ahead (July–November). The models are developed using a Bayesian approach and make use of data that extend back to 1851 with the earlier hurricane counts (prior to 1899) treated as less certain relative to the later counts. Out-of-sample hindcast skill is assessed using the mean-squared prediction error within a hold-one-out cross-validation exercise. Skill levels are compared to climatology. Predictions show skill above climatology, especially using the NAO + SOI and the NAO-only models. When the springtime NAO values are below normal, there is a heightened risk of U.S. hurricane activity relative to climatology. The preliminary NAO value for 2005 is -0.565 standard deviations so the NAO-only model predicts a 13% increase over climatology of observing three or more U.S. hurricanes.

1. Introduction

Coastal hurricanes are a serious social and economic concern in the United States. Strong winds, heavy rainfall, and storm surge kill people and destroy property. The devastation from Hurricane Katrina is a horrific reminder of this fact. Historical hurricane data provide clues about the future frequency and intensity of storms. Skillful seasonal forecasts of coastal hurricane activity are important for land use planning, emergency management, hazard mitigation, and (re)insurance contracts.

Empirical and statistical research (Goldenberg et al. 2001; Elsner et al. 2000a, 1999; Gray et al. 1992) identifies factors that contribute to conditions favorable for Atlantic hurricanes leading to prediction models for seasonal activity (Gray et al. 1992; Hess et al. 1995). Research shows that climate factors influence hurricane frequency differentially. The effect of the El Niño–Southern Oscillation (ENSO) on the frequency of hurricanes forming over the deep Tropics is significant, but its effect on the frequency of hurricanes over the subtropics is small. Additional factors help explain local

variations in hurricane activity (Lehmiller et al. 1997). In fact, the North Atlantic Oscillation (NAO) plays a statistically significant role in modulating coastal hurricane activity (Elsner 2003; Elsner et al. 2001; Jagger et al. 2001; Murnane et al. 2000). Numerical models are now capable of simulating hurricane seasons months in advance with some skill (Vitart and Stockdale 2001; Vitart et al. 2003).

Insights into regional hurricane activity have been used to build successful seasonal landfall models (Lehmiller et al. 1997; Saunders and Lea 2005). A limitation of these landfall models is that they are based on a restricted set of data (the last, approximately, 50 yr). Assuming similar climate conditions over time, statistical models built on longer data records would be expected to perform with greater precision. However, these earlier data are generally less reliable and more uncertain. A solution is to use a model that does not require all the data to have uniform precision. Elsner and Bossak (2001) demonstrate such a model for climatological analysis of U.S. hurricanes and Elsner and Jagger (2004) extend the model to include predictors. These models are based on Bayesian technology. In short, a Bayesian model is a conditional probability model that combines prior information with a likelihood specification to determine a posterior distribution. In the context of prediction, the posterior distri-

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

bution represents what we expect to happen along with how certain we are that it will.

There is a large body of statistical literature on the Bayesian approach to model building. A good starting place is with the contributions emphasizing practical applications in Gilks et al. (1998). Congdon (2003) provides an excellent overview of Bayesian modeling by illustrating the techniques on a wide range of problems, particularly from the health and social sciences. Only more recently have Bayesian models been applied in climate studies. Wikle (2000) gives an introduction to the hierarchical Bayesian approach to modeling atmosphere and ocean processes. Berliner et al. (2000) shows how to forecast tropical Pacific sea surface temperatures by incorporating the physical understanding of ENSO. Katz (2002) reviews how to perform uncertainty analysis within the context of integrated assessment of climate change. Wikle and Anderson (2003) demonstrate the utility of a Bayesian approach for climatological analysis of tornado occurrences that are complicated by reporting errors. Bayesian methods have also been applied in the area of climate change and detection (Solow 1988; LeRoy 1998; Berliner et al. 2000). Elsner et al. (2004) and Chu and Zhao (2004) show how to detect change points in hurricane activity with Bayesian models. A Bayesian approach to seasonal hurricane modeling is illustrated in Elsner and Jagger (2004). Comparison with a frequentist (or classical) approach demonstrates the usefulness of the Bayesian approach in focusing our beliefs on the relative importance various factors have on coastal hurricane activity. For example, the degree of belief we have in the future impact El Niño might have on the probability of a U.S. hurricane is formed from two independent information sources: the pre- and post-1900 records. Results confirm the utility of the earlier records by showing a greater precision on the model parameters when those records are included. Readers not familiar with the Bayesian approach to seasonal modeling are encouraged to examine the work of Elsner and Jagger (2004).

The purpose of the present paper is to offer a Bayesian model that can be used for actual predictions of U.S. hurricane activity by 1 July and that has been skill assessed using cross validation. Cross validation of a Bayesian model has yet to be done in the climate literature. The model makes use of the available hurricane records and accounts for the uncertainty inherent in the older data. The paper builds on our early work (Elsner and Jagger 2004) by providing a complete assessment of forecast skill. A comparison of forecast methods for predicting the 2004 Florida hurricane sea-son, reformulated as Bayesian models, is presented in Elsner and Jagger (2006).

We begin with a discussion of the hurricane counts and the selected predictors (covariates). In section 2 we explain the general idea behind the Markov chain Monte Carlo (MCMC) approach to Bayesian modeling. In section 3, we discuss the particulars of our model strategy for predicting the probability of U.S. hurricane counts. The model uses the landfall counts back to 1851, but treats the counts prior to 1899 as somewhat uncertain. Values during the nineteenth century for some of the predictors are missing so we specify them in the model as samples from a distribution. Since the strategy relies on sampling, in section 4 we discuss convergence diagnostics. The importance of the individual predictors is discussed in section 5. In section 6, we describe the procedure of cross validation in the context of Bayesian models. Results are described in section 7. Comparisons are made with a climatological model. A summary and conclusions are given in section 8. Model code along with data and initial values are provided as an appendix.

2. Data

a. Hurricane counts

A chronological list of all hurricanes that have affected the continental United States in the period 1851–2004, updated from Jarrell et al. (1992), is available from the U.S. National Oceanic and Atmospheric Administration (NOAA, online at http://www.aoml.noaa.gov/hrd/hurdat/ushurrlist.htm). A hurricane is a tropical cyclone with maximum sustained (1 min) 10-m winds of 65 kt (33 m s^{-1}) 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 in striking southeast Florida and Louisiana.

Here we only consider whether the observations indicate that the cyclone struck the continental United States at least once at hurricane intensity. The approximate length of the U.S. coastline affected by hurricanes from the Atlantic is 6000 km. We do not consider hurricanes affecting Hawaii, Puerto Rico, or the Virgin Islands. An analysis of an earlier version of the U.S. hurricane record in 50-yr intervals is presented in Elsner and Bossak (2001). Here it is assumed that the annual counts of U.S. hurricanes are certain back to 1899, but less so in the interval of 1851–98. The justifi-

cation for this cutoff is based on an increased awareness of the vulnerability of the United States to hurricanes following the Galveston tragedy of 1900 and on coastal population levels at this time.

b. Potential predictors

The approach taken here differs from other work on seasonal hurricane forecasting in that we do not search for predictors, nor are we showcasing a new predictor. We are informed about which predictors to include in our models from the research literature so instead we focus on the technology of the forecast modeling process. We begin by describing this 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. They showed that—previous autumnal rainfall within the Sahel region of western Africa, the forward extrapolated vertical shear magnitude between 30- and 50-mb tropospheric winds [quasi-biennial oscillation (QBO)], the 700–200-mb vertical shear in the Miami–West Palm Beach area, the July monthly sea level pressure at Cape Hatteras, and the July average monthly East Coast sea level pressure—are important precursors to hurricane activity over the southeastern United States. They note that the highest likelihood of a hurricane landfall occurs with relatively high July sea level pressures over Cape Hatteras and strong vertical wind shear over south Florida.

A regression model for the southeastern U.S. hurricanes that makes use of longer records describing ENSO and NAO, rather than the more temporally limited upper-air sounding data used in Lehmiller et al. (1997), is designed in Elsner (2003). The model is a Poisson regression that uses August–October averaged values of the Southern Oscillation index (SOI) as an indication of ENSO together with May–June averaged values of an index for NAO to predict the probability of hurricanes from Texas to South Carolina. The importance of ENSO on U.S. hurricane activity is elucidated in Bove et al. (1998) and the importance of NAO is suggested in Liu and Fearn (2000) and Elsner et al. (2000b). 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 southeastern U.S. hurricanes are more likely during La Niña conditions when NAO is weak in comparison to climatology (Jagger et al. 2001; Elsner and Bossak 2004). ENSO is also used as a predictor in a seasonal prediction model of tropical cyclone activity along the southern China coast (Liu and Chan 2003).

Saunders and Lea (2005) are the first to 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 over the North Atlantic, North America, and eastern Pacific during July. They use ordinary least squares regression to regress normally transformed accumulated cyclone energy onto the wind anomalies. It is worth noting that although their wind anomaly index 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 (as does NAO). In fact, Saunders and Lea (2005, manuscript submitted to *Nature*) verify that the highest likelihood of southeast U.S. hurricane landfall occurs with high values of July-averaged sea level pressures over Cape Hatteras.

The Colorado State University Tropical Meteorology Team issues forecasts of landfall probabilities along with their traditional forecasts of overall activity [see Gray et al. (1992) for a description of a model that predicts overall Atlantic hurricane activity]. Although not published in the peer-reviewed literature, their methodology for predicting U.S. landfall activity is readily accessible and 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 6 indices of hurricane activity each expressed as a percentage difference from the long-term average. SST anomalies represent those of the Atlantic multidecadal oscillation (AMO) for July. AMO is characterized by fluctuations in SSTs over the North Atlantic Ocean (Goldenberg et al. 2001) determined to some extent by temperature and density differences across the basin (thermohaline circulation).

The research cited above provides us with the background to select a set of predictors for forecasting U.S. hurricane activity. Based on physical and statistical relationships as well as data availability, we choose NAO, ENSO, and AMO. We do not include the tropospheric wind anomalies, as data are only available since 1948. Next, we describe in more detail the specific predictors used in our model.

c. Predictors selected

NAO index values are calculated from sea level pressures at Gibraltar and at a station over southwest Iceland (Jones et al. 1997), and are obtained from the Climatic Research Unit. The values used in this study are an average over the pre- and early-hurricane season months of May and June and are available back to 1851. Units are in standard deviations. These months are cho-
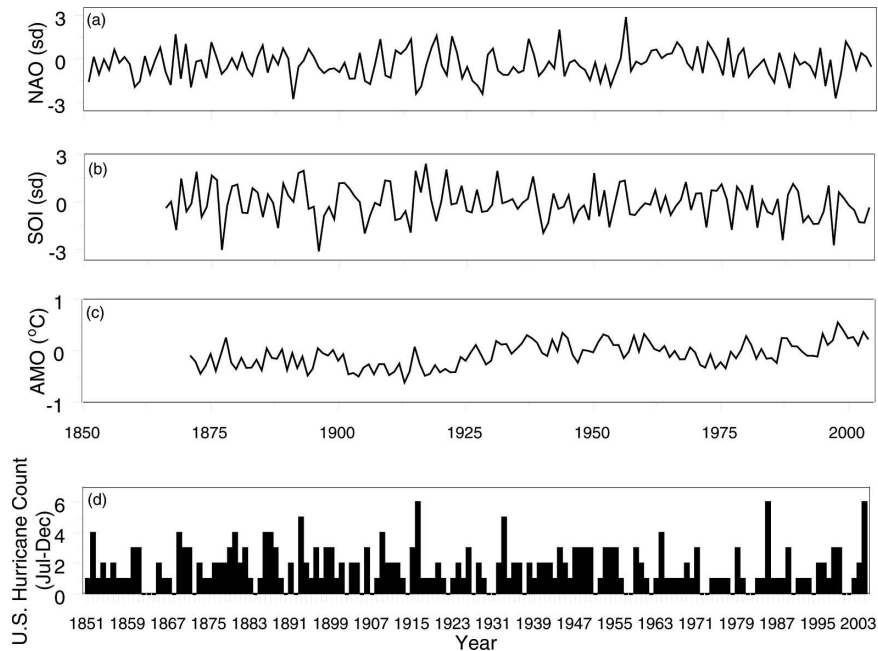
FIG. 1. Time series of the annual predictors and predictand values used in the 1 July forecast models of U.S. hurricane activity over the period 1851–2004. The predictors include the May–June averaged value of (a) NAO, (b) SOI, and (c) AMO. (d) Time series of annual U.S. hurricane counts from 1 July. A U.S. hurricane count is the number of landfalling hurricanes. There are 30 yr without a U.S. hurricane from July to November, 53 yr with exactly one, 32 yr with exactly two, 27 yr with exactly three, 7 yr with exactly four, 2 yr with exactly five, and 3 yr with exactly six.

sen as a compromise between signal strength and timing relative to the hurricane season. The signal-to-noise ratio in NAO is largest during the boreal winter and spring (see Elsner et al. 2001), whereas the Atlantic hurricane season begins in June.

SOI values are used as an indicator of ENSO. Although noisier than equatorial Pacific SSTs, values are available back to 1866. SOI is defined as the normalized sea level pressure difference between Tahiti and Darwin, Australia. SOI is strongly anticorrelated with equatorial SSTs so that an El Niño warming event is associated with a negative SOI. Units are standard deviations. Although the relationship between ENSO and hurricane activity is strongest during the hurricane season, we use a May–June average of SOI values as our predictor. The monthly SOI values are obtained from the Climatic Research Unit where they are calculated based on a method given in Ropelewski and Jones (1987). Details on sources of the early pressure data are available in Allan et al. (1991) and Können et al. (1998).

AMO values are based on a blend of model values and interpolated observations, which are used to compute Atlantic SST anomalies north of the equator (Enfield et al. 2001). As with NAO and SOI, we use a

May–June average of AMO anomalies as our predictor. The anomalies are computed by month using the climatological time period 1951–2000 and are available back to 1871. Units are in degrees Celsius. Values of AMO are obtained online from the NOAA–CIRES (Cooperative Institute for Research in Environmental Sciences) Climate Diagnostics Center (CDC).

It is possible to include other predictors. However, data on the QBO, Sahel rainfall, or tropospheric wind anomalies, for example, are limited to the past 60 yr or so. As such, the model would need to infer approximately 60% of the values for each additional predictor chosen. As explained below, one of the advantages of the approach taken here is that it can naturally handle these missing values. Practically, however, this adds a severe computational burden for the necessary cross-validation exercise, so here we limit the number of predictors to three and choose ones that have the fewest missing values.

Figure 1 shows the time series of the three predictors. The sampling interval is 1 yr. Each predictor displays high frequency (year to year) variations. Lower-frequency changes are most pronounced in AMO. Note that only NAO is available back to 1851. Pairwise correlations between the predictors (computed over years without

missing data) are not large. The highest positive correlation (+0.06) occurs between NAO and SOI. The largest negative correlation (−0.20) occurs between NAO and AMO, indicating that warm Atlantic SSTs tend to occur with a positive phase of NAO. Since the approximate standard error on the correlation is $N^{-1/2}$, this correlation is marginally significant but captures only 4% of the variance. The present work makes no attempt to model the level of uncertainty in these predictor values, although a specification in the model is made for the possibility of missing data.

d. Partial season hurricane counts

Since the predictors require data through the month of June, our hurricane count predictand excludes hurricanes making landfall before 1 July. The NOAA chronological list of U.S. hurricanes includes a total of 274 storms over the period 1851–2004. Of these, 1 occurred in May and 19 occurred in June. The 20 (7%) hurricanes striking the United States before July are eliminated in building the 1 July forecast model. The partial season counts are shown as a time series in Fig. 1. The partial hurricane season excludes the months of May and June. There is no significant trend in the counts although somewhat fewer hurricanes are noted during the most recent decades. Florida in particular has seen a significant decline in hurricanes since the middle of the twentieth century (Elsner et al. 2004). In contrast, the 2004 season featured six U.S. hurricanes after 1 July. This occurred two other times, in 1916 and in 1985. Next, we discuss the general strategy for modeling annual hurricane counts.

3. A model for annual hurricane counts

a. General specification

The canonical model for hurricane count data is the Poisson regression (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 is derived from the distribution of wait times between successive events. For our purpose, the Poisson regression model is used to model a set of partial season U.S. 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

$$
\begin{aligned}
h_i &\sim \text{Poisson}(\lambda_i) \\
\lambda_i &= \exp(\beta_0 + x'_i\beta)
\end{aligned}
\tag{1}
$$

where λ_i is the hurricane rate for year i, β_0 is the intercept, and β is the vector of predictor coefficients. The symbol $\sim$ refers to a stochastic relationship and indicates that the variable 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 variable on the left-hand side algebraically related to variables on the right-hand side.

Values for the intercept (β_0) and coefficients (β) define the specific model and are estimated using a Bayesian approach. In short, we assume that the parameters (intercept and coefficients) have a distribution and that 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 that the parameters are fixed but unknown, and we find that the values for the parameters were most likely to have generated the observed data by maximizing the likelihood. The Bayesian approach combines the most frequent likelihood with our prior belief $f(\beta)$ using Bayes's rule, so that for the Poisson regression we are interested in

$$
f(\beta|h) \propto f(h|\beta)f(\beta).
\tag{2}
$$

The distribution $f(\beta|h)$ is the posterior density and is the probability density of β, which is 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 0. In the present context, the posterior density represents the histogram of the probability of observing h hurricanes in a given year. 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.

b. Additional specifications

While the above specification mixing the Poisson distribution with the normal distribution forms the backbone of our prediction models, additional specifications are needed. As mentioned, although hurricane counts are available back to 1851, those earlier records are less precise. Therefore, we include an indicator variable

(IND) that is given a value of 1 for the years 1851–98 and a value of 0 for the years 1899–2004. The coefficient associated with this term is the logarithm of the probability, p, that a hurricane is observed given that it occurred. If one assumes that the actual hurricane rate is independent of the observation process, and that this rate is Poisson with rate λ, then conditioned on p and λ, the observations form a Poisson process with rate λp. In effect, the models allow for the possibility that a hurricane occurred but was not recorded. We discuss the influence that this term has on the other predictors in section 6.

A Bayesian approach does not require that values for the predictors start with the same year nor is it necessary that the dataset be complete. In our case, values for NAO start with 1851, but values for SOI and AMO start somewhat later. We let the Bayesian approach sample values for years in which predictor values are missing. This is done by specifying a distribution for the predictors that have missing data. Here we use the normal distribution centered on 0 with a precision of 1 for SOI and 20 for AMO[1]. These additional specifications are included in the model below (note that since there are no missing values for NAO, there is no need to specify a distribution):

$$\begin{aligned} \text{SOI}_i &\sim \text{Normal}(0, 1) \\ \text{AMO}_i &\sim \text{Normal}(0, 20) \\ h_i &\sim \text{Poisson}(\lambda_i) \\ \lambda_i &= \exp(\beta_0 + \beta_1 \text{NAOI} + \beta_2 \text{SOI}_i + \beta_3 \text{AMO}_i \\ &\quad + \beta_4 \text{IND}). \end{aligned} \quad (3)$$

The cost associated with a Bayesian approach is the requirement to formally specify our prior beliefs. Here we take the standard route and assume noninformative priors, which as the name implies, provide little information about the value of the parameters of interest. To finish our model specification, we assume independent normally distributed priors for each component of the β vector, where the mean of the normal distribution is 0 and the variance is 10^6 [$\beta_j \sim \text{Normal}(0, 10^6)$]. This is a flat distribution (very large variance) that contributes little information.

The leading contender model for anticipating U.S. hurricane counts is climatology. We are interested therefore in comparing the skill of our models against the skill available from a climatology model. To make

honest comparisons, we keep the specifications identical with the exception that the climatology model involves no predictors. To do this we replace the equation specifying λ_i [see Eq. (3)] with $\lambda_i = \exp(\beta_0 + \beta_4 \text{IND})$. The climatology model includes the same 154-yr dataset of hurricane counts as well as the uncertainty associated with the earlier records. The climatology model is our baseline and thus, skill differences between our models and climatology are attributable to the predictors.

4. Convergence of model samples

Samples of model parameters and data are generated using the software WinBUGS (Windows version of Bayesian inference using Gibbs Sampling) developed at the Medical Research Council in the United Kingdom (Gilks et al. 1994; Spiegelhalter et al. 1996). WinBUGS chooses an appropriate MCMC sampling algorithm based on the model structure. The WinBUGS code for our U.S. hurricane forecast model is given in the appendix. Prediction is achieved by setting the hurricane count for the year of interest to the missing data flag [not available (NA)]. In this way, counts for the year are sampled conditional on the model coefficients and the available data.

We check for mixing and convergence of the models by examining successive samples of the parameters. Although convergence diagnostics are a good idea in general, they are particularly important prior to cross validation to establish a minimum number of samples to discard as burn-in. Typically a large number will all but guarantee convergence, but since it is necessary to rerun the sampler N times (where N is the number of years) for hold-one-out cross validation, a small number is preferred.

Figure 2 shows 3000 samples of β_1 from the full predictor model using two different sets of initial conditions. Samples from the posterior distribution of β_1 indicate relatively good mixing and quick settling as both sets of initial conditions result in samples that fluctuate around a fixed median value. Autocorrelation values for lags greater than three samples are generally less than 0.07 (in absolute value), indicating that the samples mix efficiently starting from either set of initial conditions. Posterior densities computed using a kernel smoother are nearly identical for the two starting conditions. The average value of β_1 over the last 1000 samples is -0.217 using the first set of initial conditions and is -0.210 using the second set. Similar convergence and mixing occurs with the other model parameters.

---

[1] The precision is the inverse of the variance, here estimated from the available data.
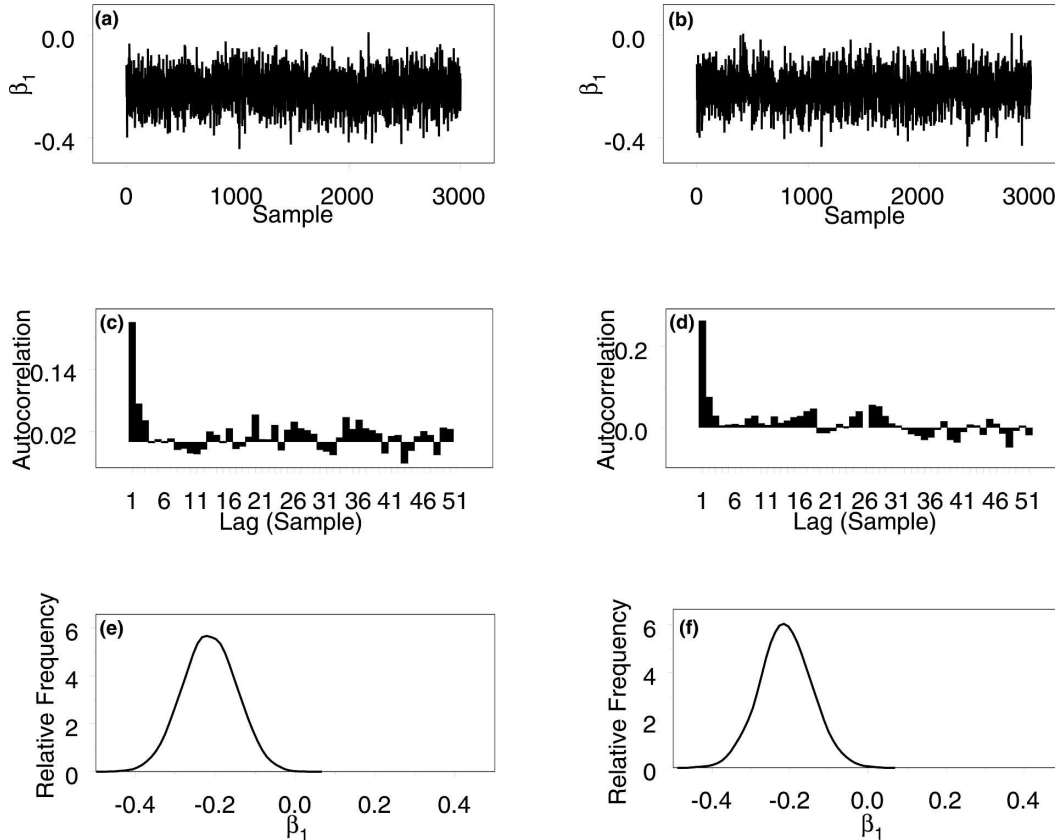
FIG. 2. Diagnostic plots of the β_1 parameter for the prediction model. Successive sample values starting with (a) $\beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ and (b) $\beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = 1$. Three thousand samples are generated from each set of initial conditions. (c), (d) Corresponding autocorrelation functions of the sample history. (e), (f) Corresponding kernel densities of the posterior distributions of β_1 for the two sets of initial conditions for samples 1–3000.

Based on these diagnostics, we choose to discard the first 2000 samples in each of our model runs.

5. Predictor importance

We examine the importance of the predictors by generating posterior samples for the respective coefficients specified in Eq. (3). Based on the convergence diagnostics results of the previous section, we generate 12K samples and discard the first 2×10^3 as burn-in. The coefficients β_0, β_1, β_2, and β_3 are stochastically specified as having a normal distribution with a mean of 0. The value of p is specified using a uniform distribution between 0.8 and 0.95. Values of the posterior samples indicate the relative influence that the associated predictor has in forecasting seasonal U.S. hurricane counts. Density plots of the 10×10^3 samples are shown in Fig. 3.

The ratio of the areas under the curve on one side of 0 to the total provides an estimate of the significance of the associated predictor to the model. Sample values of the coefficient β_0 are greater than 0 with a mean of 0.445. We also note that the β_1 sample values tend to be less than 0 with a posterior mean of -0.214. This implies that when NAO is weak, the probability of U.S. hurricanes increases. This is consistent with our earlier results (Elsner 2003; Jagger et al. 2001). Out of the 10K samples, only 4 (0.04%) were greater than 0, which can be interpreted as a p value of 0.0004, indicating little support for the null hypothesis that the NAO term is not important. The 95% credible interval on the value of β_1 is $(-0.342, -0.085)$. In contrast, β_2 and β_3 sample values lie on both sides of the zero line. This indicates that SOI and AMO predictors tend to be less important in the model. The posterior mean of the β_2 samples is 0.063, indicating that when SOI is positive (La Niña conditions), the probability of U.S. hurricanes increases. However, the 95% credible interval for this coefficient $(-0.056, +0.184)$ includes 0. Nearly 15% of the β_2 samples are less than 0. The posterior mean of the β_3 samples is $+0.226$, indicating that when AMO is
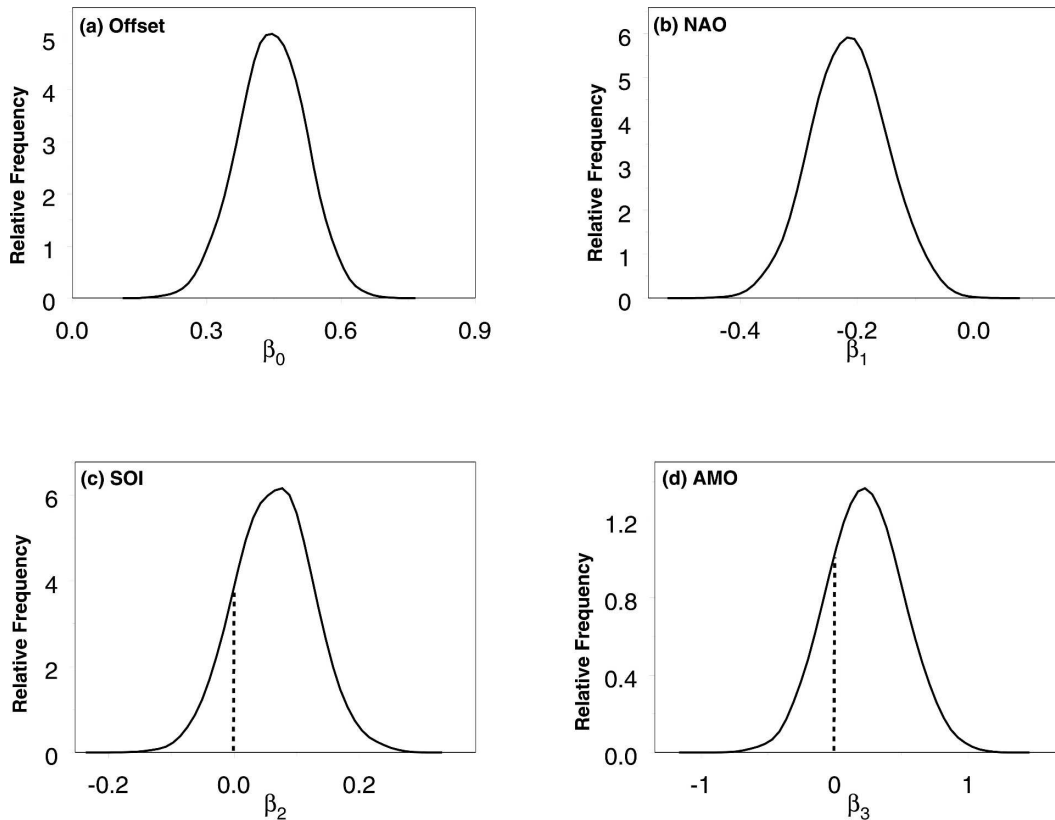
FIG. 3. Kernel density of the posterior distribution of (a) β_0, (b) β_1, (c) β_2, and (d) β_3. The distributions are based on 10×10^3 samples after the first 2×10^3 samples are removed.

positive, the probability of U.S. hurricanes increases (Goldenberg et al. 2001). Like the SOI coefficient, the 95% credible interval on the AMO coefficient (-0.313, $+0.773$) includes 0. Approximately 21% of the β_3 samples are less than 0.

The β_4 coefficient is used to express our belief about how accurate the earlier records are relative to the later records. The value of β_4 is deterministically specified as the logarithm of p, where p is the probability that the earlier records are known perfectly. The higher the value of p, the stronger our belief is that the earlier records are as reliable as the later records and the less influence this term has on the prediction. Experimentation shows that SOI and AMO predictors are most affected by the inclusion of the earlier records compared with the NAO predictor. The relative change in the posterior mean value of the β_1 coefficient is only $+8.6\%$ if we exclude the earlier records (by setting p to 0), as compared to a $+50.6\%$ change in the mean value of β_2 if the earlier records are excluded. Most striking is the mean value of β_3, which changes from $+0.226$ to -0.181 upon exclusion of the earlier records. The results presented in this section suggest that we perform a cross-validated hindcast comparison among four com-

peting modeling strategies including climatology, the full predictor model (all three predictors), the reduced predictor model (NAO and SOI), and the NAO-only model.

6. Cross validation

To assess how well each of the four modeling strategies can be expected to perform in actual forecast situations, we perform hold-one-out cross validations (Michaelsen 1987; Elsner and Schmertmann 1994). Cross validation results in a relatively accurate measure of a prediction scheme's ability to produce a useful prediction rule from historical data. In short, the method works by successively omitting an observation from the modeling procedure and then measuring the error that results from using the model to predict the left out observation. The idea is to remove the information about the omitted observation that would be unavailable in an actual forecast situation. Since we assume that one hurricane season is largely independent of the next, we can remove one year at a time rather than multiple years (block removal).

Although necessary for getting an honest measure of
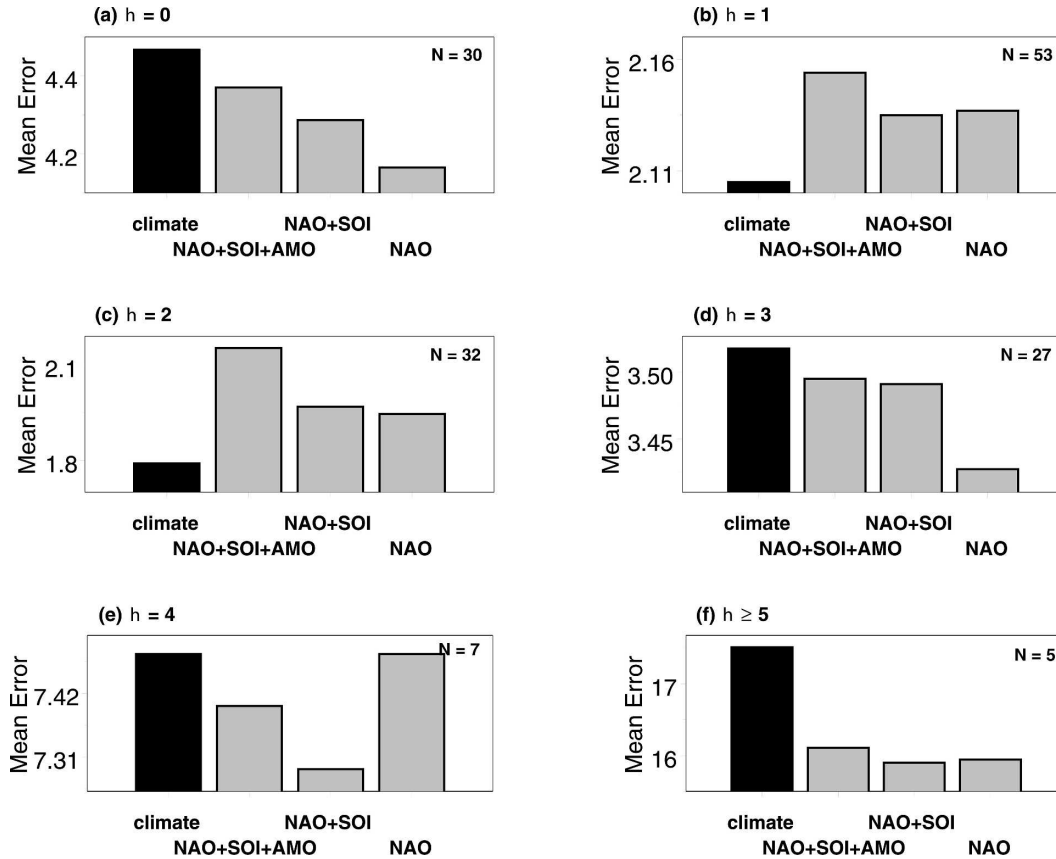
FIG. 4. Cross-validated MSE for years with (a) $h = 0$, (b) $h = 1$, (c) $h = 2$, (d) $h = 3$, (e) $h = 4$, and (f) $h \geq 5$ hurricanes. Errors are computed for four modeling strategies including climatology (climate), a full predictor model (NAO + SOI + AMO), a reduced predictor model (NAO + SOI), and a model with NAO as the only predictor (NAO).

forecast skill, the procedure of cross validation poses special challenges in the context of Bayesian modeling. First, the model specification requires conditional sampling to generate the posterior predictive distributions. As such, it is necessary to assure that solutions starting from different initial conditions adequately converge. Second, there is no analytical solution to the cross validation. This requires us to rerun the sampler $N = 154$ times once for each year withheld. We note that cross validation with Bayesian models can be done using a procedure called "importance sampling." Here we take advantage of fast computation that allows us the direct calculation of posterior predictive densities.

To automate the procedure in WinBUGS, we make 154 copies of our datasets. Within each copy of the hurricane dataset (predictand), we replace an actual count with the missing value code (NA). Furthermore, it is necessary to duplicate the predictor values for which a distribution is specified. Thus we obtain 154 copies of the 154 values (including the NAs) of the SOI

and AMO datasets. The WinBUGS code shown in the appendix is modified by including an outer cross-validation loop. This requires an additional index on the predictors and the predictand. Furthermore, it is necessary to add an index on the coefficients and on the probability that the hurricane is observed.

7. Results

A hold-one-out cross validation as described above is performed on four modeling strategies including climatology (no predictors), full model (NAO, SOI, and AMO as predictors), reduced model (NAO and SOI as predictors), and single predictor (NAO only). Model skill is assessed using the posterior expected error [also called the mean-squared error (MSE)] given as the squared difference between the posterior predicted probability and the observation, where the observation is either a 0 or 1 for each hurricane count:
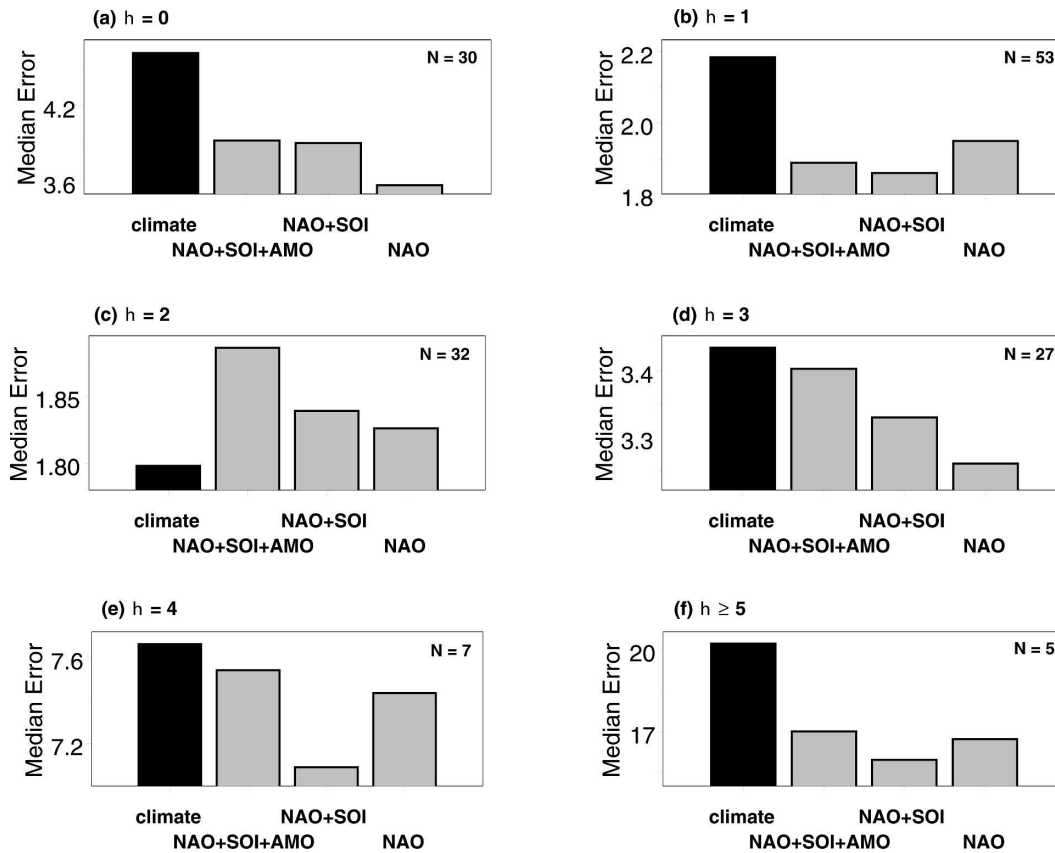
FIG. 5. Same as in Fig. 4, but for median squared error.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} \sum_{k=0}^{\infty} p_i(k)[k - o_i(k)]^2 \quad \text{with} \quad n = 154.$$

(4)

The posterior predicted probability is based on 10×10^3 samples. For each year ($i = 1, \ldots, 154$) and each hurricane count ($k = 1, \ldots, 10$), there is a predicted probability [$p_i(k)$] and an observed count [$o_i(k)$]. The squared differences are summed over all hurricane counts (up to 10) and over all forecast years. Using the MSE is useful for linear models. Linear nested models in the Bayesian context choosing the model that minimizes the squared prediction error is equivalent to choosing the median probability model from all possible models (Barbieri and Berger 2002).

　　MSE over the 154 yr is 3.493 for climatology, 3.481 for the full model, 3.439 for the reduced model, and 3.410 for the NAO-only model. The smaller the error, the better is the forecast. MSE equals the model error plus the statistical error. The statistical error is equal to the Poisson rate. For example, if the rate is 2 h yr^{-1} (hurricanes per year), then the minimum MSE is

2 (h yr^{-1})^2. Thus, with a Poisson model, the minimum MSE is not 0. With a Poisson distribution, the mean is equal to the variance so that the minimum MSE will only approach 0 if the mean approaches 0. This is different from the normal distribution where the mean and variance are not related. It is also worth noting that although forecasts of the mean hurricane count fluctuate around 2, there can be a sizeable change in the forecast probability of a large number of hurricane landfalls with a small change in the forecast mean; changes in the tails of these probability distributions are of practical importance to catastrophe reinsurers and risk managers.

　　The climatology model will perform well on years in which the count is close to the mean rate. Therefore, it is interesting to compute the MSE for years grouped by the observation count. For example, we can compute MSE for years in which no hurricanes were observed. In this case, the outer summation in Eq. (4) is over the 30 yr of no U.S. hurricanes. Results are shown in Fig. 4. The number of years is shown in the upper-right corner of each panel. The predictor models capture the variation in annual counts better than does the climatology

TABLE 1. Mean- and median-squared errors for two groups of years. The first group is years in which there were one or two U.S. hurricanes, and the second group had 0 or more than two.

Group	Climate		NAO + SOI + AMO		NAO + SOI		NAO	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
$h = 1$ or 2	1.987	1.863	2.096	1.886	2.072	1.840	2.065	1.859
$h = 0$ or 3+	5.348	3.997	5.186	3.870	5.122	3.689	5.067	3.625

model. Of the three modeling strategies that include predictors, the NAO-only strategy appears to have a slight edge although the NAO + SOI strategy is somewhat better at predicting years with a large number of hurricanes. The median-squared prediction errors are shown in Fig. 5. The plots are similar, but we note that the models outperform climatology for years in which only one hurricane is observed.

In general, the model strategies show skill above climatology for years in which there are no hurricanes or more than two hurricanes. Table 1 shows the mean and median errors for two groups of years where group 1 corresponds to years in which there are 1 or 2 hurricanes and group 2 to years in which there are 0 or 3+ hurricanes. The median error for the years of group 1 is 1.840 (1.863) for the NAO + SOI (climatology) model. This compares with an error for the years of group 2 as 3.689 (3.997) for the NAO + SOI (climatology) model.

Since climatology is the benchmark for this kind of seasonal forecast, it is interesting to examine hindcasts
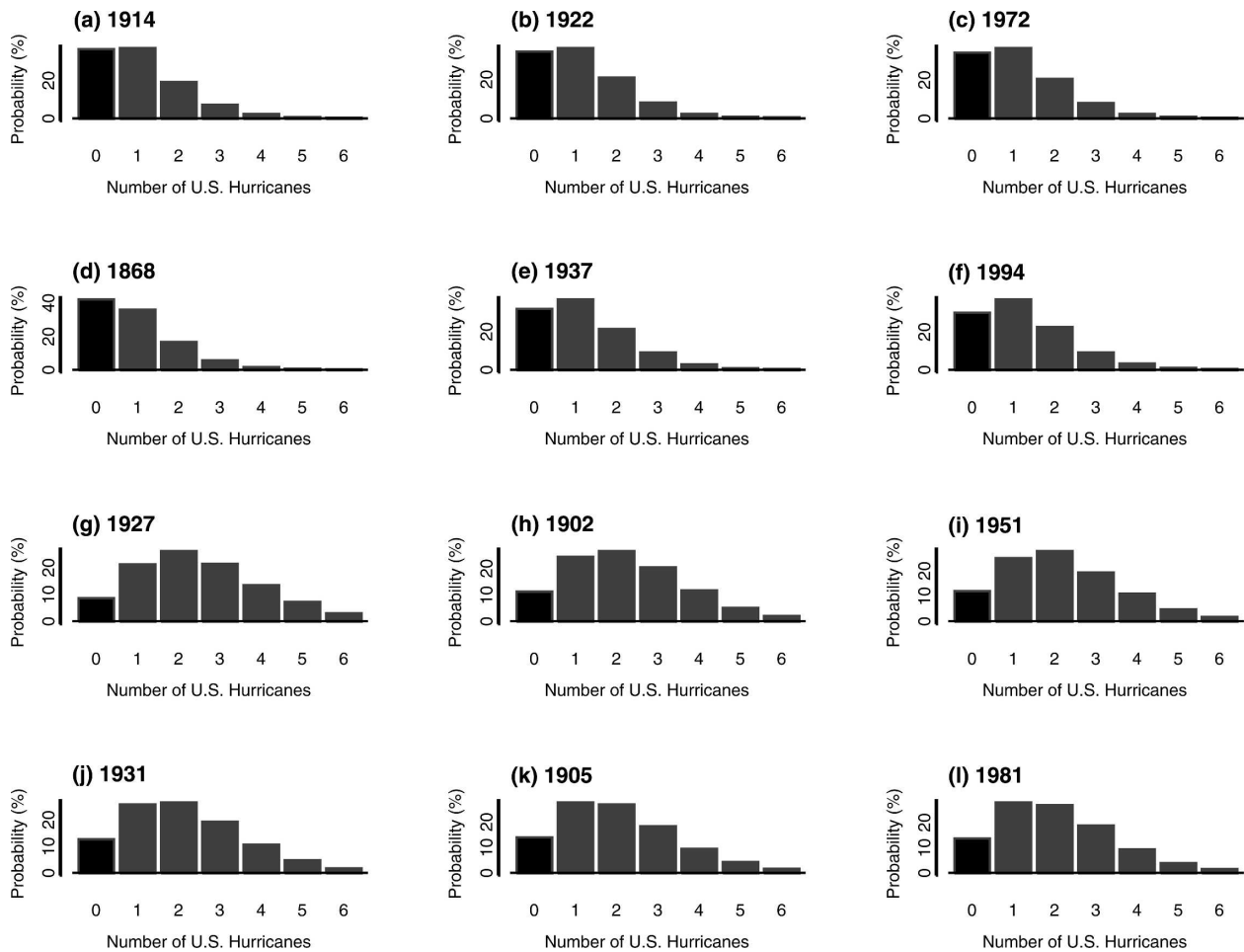
FIG. 6. Posterior predictions showing the probability of observing h hurricanes when the observed hurricane count is zero. The probabilities are based on the NAO + SOI model; (a)–(f) six best and (g)–(l) six worst predictions (in order) relative to climatology.
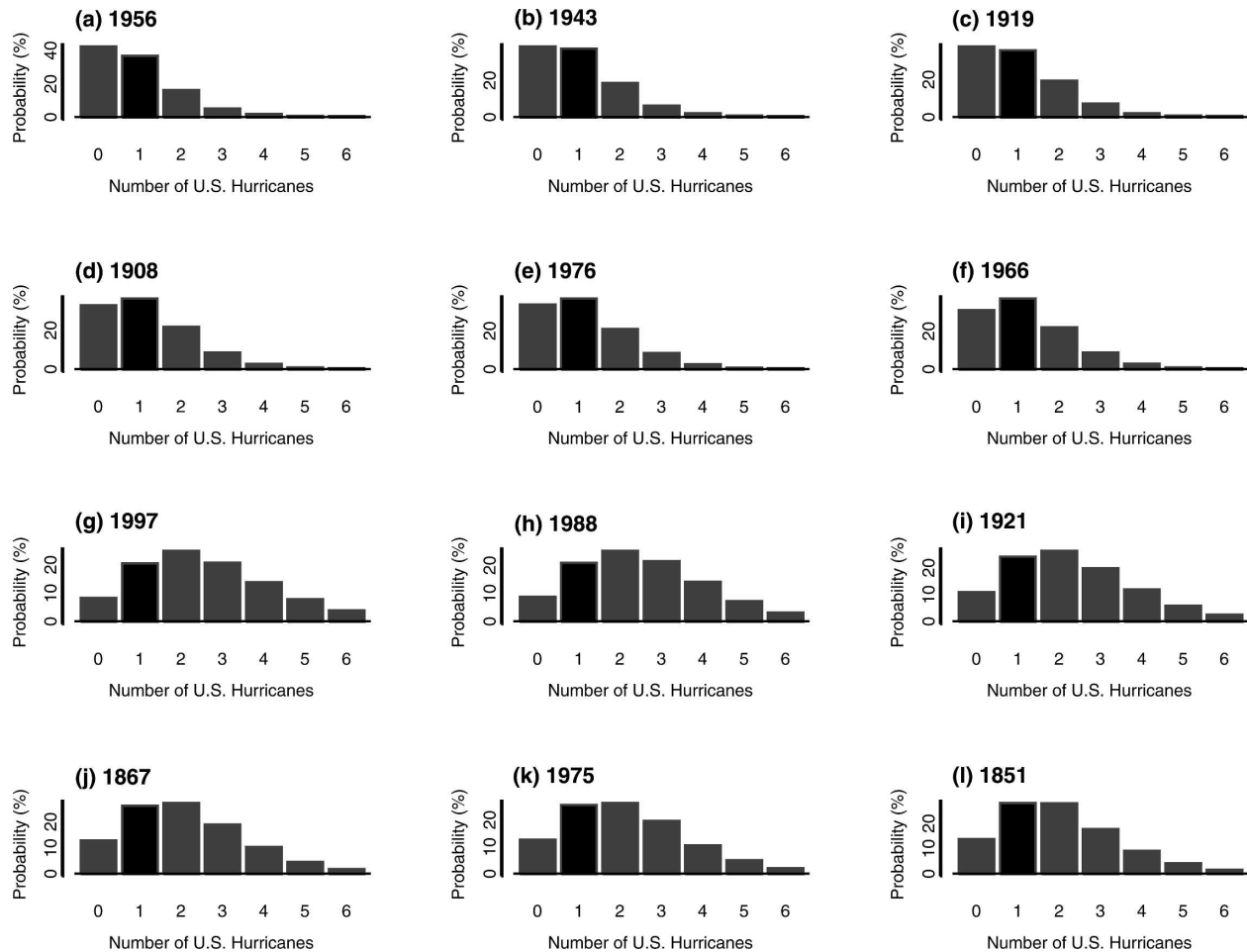
FIG. 7. Same as in Fig. 6, but when the observed count is one.

made in years in which the predictor models were better and worse than climatology. The rationale for doing this is to understand the limitations of the model. The first step is the identification of years in which the model does poorly. Are there any systematic failures (e.g., all presatellite-era years)? The second step is to determine how it failed. Did the model over- or underpredict? The third step is to try to determine why it performed poorly and what might need to be added to improve performance. In addition, by including best years with worst years, we can directly compare differences.

For each of the 154 yr, we compute the difference in MSE between climatology and the NAO + SOI model. We then rank the errors from largest to smallest and choose the six best and six worst years grouped by the observed count. Figure 6 shows the hindcasts for years when no hurricanes occurred. The top six panels are the best relative to climatology and the bottom six are the worst. The black bars indicate the observed number for

that year and the bar heights are the hindcast probability. The best hindcast relative to climatology was made in the 1914 season. Other good years are 1922, 1972, and 1994. The worst hindcast relative to climatology for years in which no hurricanes were observed was made in the 1927 season. The most recent bad year was 1981 when the predictor model indicated a somewhat active season.

Figure 7 shows the hindcasts for years with one U.S. hurricane. The best years are those in which the model indicated an inactive season while the worst years are those in which the model indicated an active season. Note, however, that even for years in which the model underperforms climatology, the hindcast probability of observing exactly one hurricane is relatively high. This is more pronounced for years in which two hurricanes hit the United States (Fig. 8). Here we see that the mode of the hindcast distributions for the bottom six performing years relative to climatology corresponds to what actually occurred. The reason the model scores
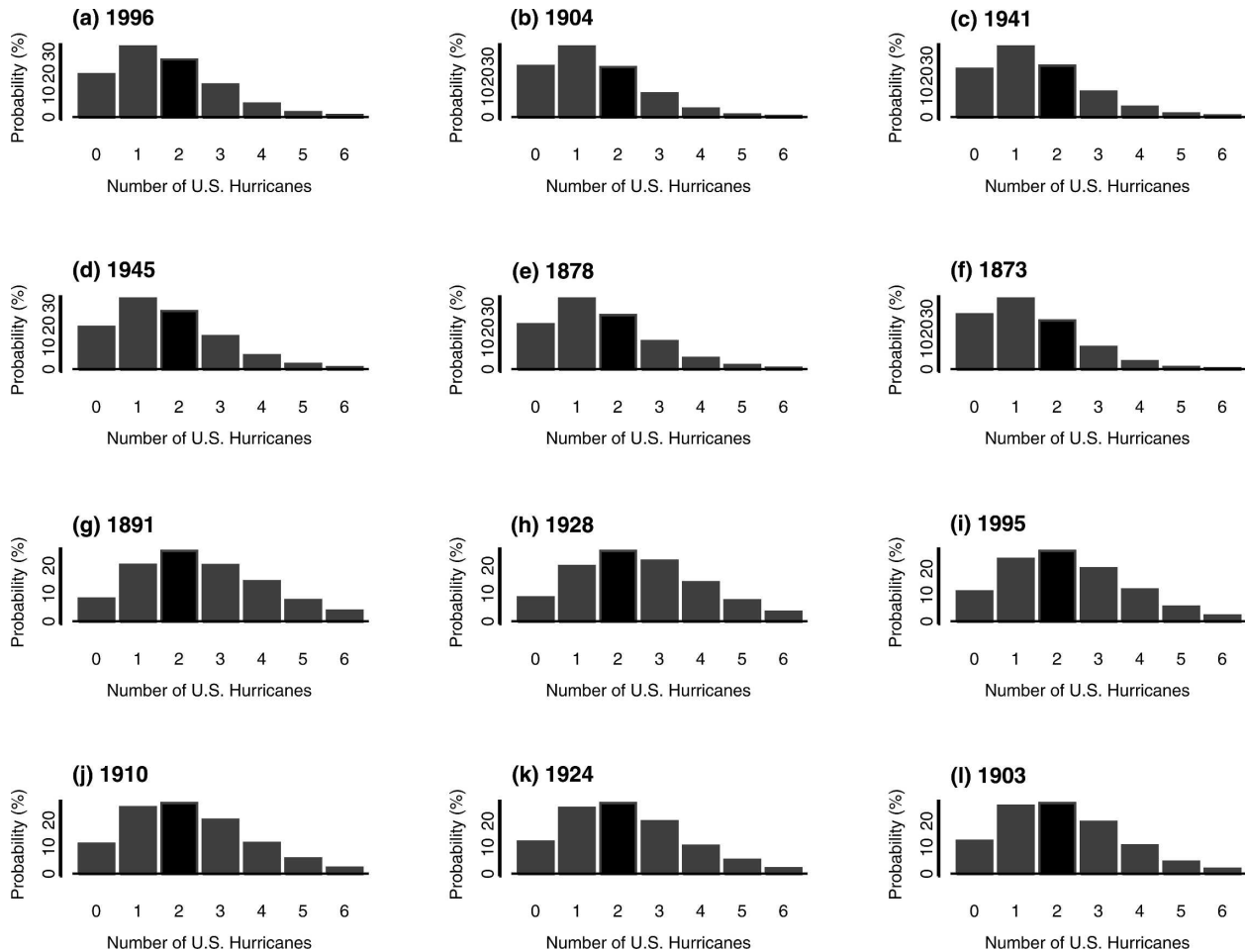
FIG. 8. Same as in Fig. 6, but when the observed count is two.

better than climatology in years like 1996 is that climatology assigns a higher probability to three or more hurricanes. Likewise, the reason climatology outperforms the predictor model in years like 1995 when the observed count matches the hindcast mode is that climatology assigns a lower probability to three or more hurricanes.

Figure 9 shows the hindcasts for years with three U.S. hurricanes. The best years (like 1998) are those in which the model predicts an active season and the worst years (like 1999) are those in which the model predicts an inactive year. Similar results are seen in years in which there are four or more U.S. hurricanes (Fig. 10). Since there are only 12 yr in this grouping, this figure shows all cases. Although the hindcast for 2004 using the NAO + SOI model is slightly worse than climatology, we note that the NAO-only model hindcast is better than climatology.

We find no systematic bias (clustering) for years in which climatology outperforms the model or for years

in which the model outperforms climatology. We note that during the modern era (since 1950), 1951 and 1981 were years in which the model seriously overpredicted U.S. hurricane activity and 1979 and 1999 were years in which the model seriously underpredicted activity. We speculate that part of the reason for the model's poor performance during these years is connected to the presence of baroclinically initiated (BI) hurricanes. Since BI hurricanes are less likely to make landfall in the United States (Elsner et al. 1996), their occurrence tends to reduce the conditional probability of a landfall given a hurricane. We find two BI hurricanes in 1951 (Able and Jig) and one in 1981 (Emily), but none in 1979 and 1999. Clearly, additional studying in understanding why the model fails to accurately predict hurricane activity for some years is needed.

8. Summary and conclusions

Seasonal landfall forecasting is relatively new. More work is needed to understand the physical mechanisms
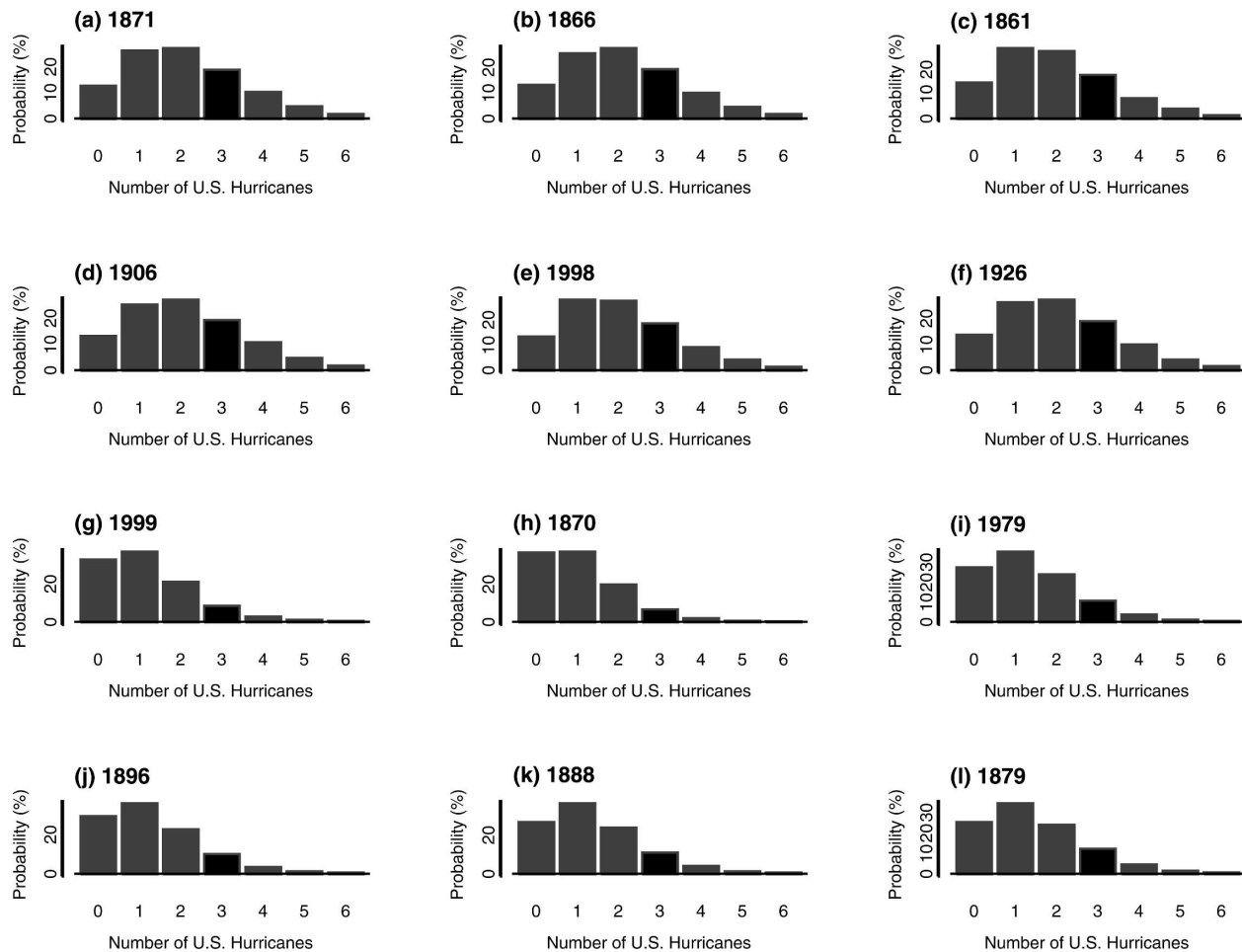
FIG. 9. Same as in Fig. 6, but when the observed count is three.

responsible for the frequency of particular storm tracks. The topic is relevant to business, government, and society. In fact, risk-modeling companies like Accurate Environmental Forecasting are beginning to offer products that make use of the science and technology of landfall forecasting. Statistical models of infrequent events must rely on the longest available data records. Here we develop prediction models for U.S. hurricane activity that take advantage of the historical record extending back to 1851. The models use a log-linear specification for the annual hurricane rate and they include three predictors previously identified as important in modulating hurricane activity. Predictors include May–June averaged values representing NAO, SOI, and AMO.

Models are Bayesian and implemented using the freely available WinBUGS software. Noninformative priors are specified for the model coefficients, missing predictor values, and hurricane count precision. Posterior samples of the model coefficients are examined for

convergence. We generate 12×10^3 samples for each model parameter and discard the first 2×10^3 as burn-in. Predictions are achieved by setting the hurricane count for the year being hindcast to the missing data flag (NA) and allowing the sampler to generate samples of hurricane counts for the hindcast year that are conditional on the model parameters. A hold-one-out cross-validation exercise is used to examine the skill of the modeling strategies against climatology.

We examine three modeling strategies: a full model that includes all three predictors (NAO, SOI, and AMO), a reduced model that includes NAO and SOI, and a single-predictor model that includes only NAO. We compute the MSE for forecasts that could be issued by 1 July. The models provide the predicted probability of observing h hurricanes during the balance of the hurricane season. Results show that all three models capture the interannual variation in hurricane counts better than does climatology. Of the three, the NAO + SOI model and the NAO-only model perform the best.
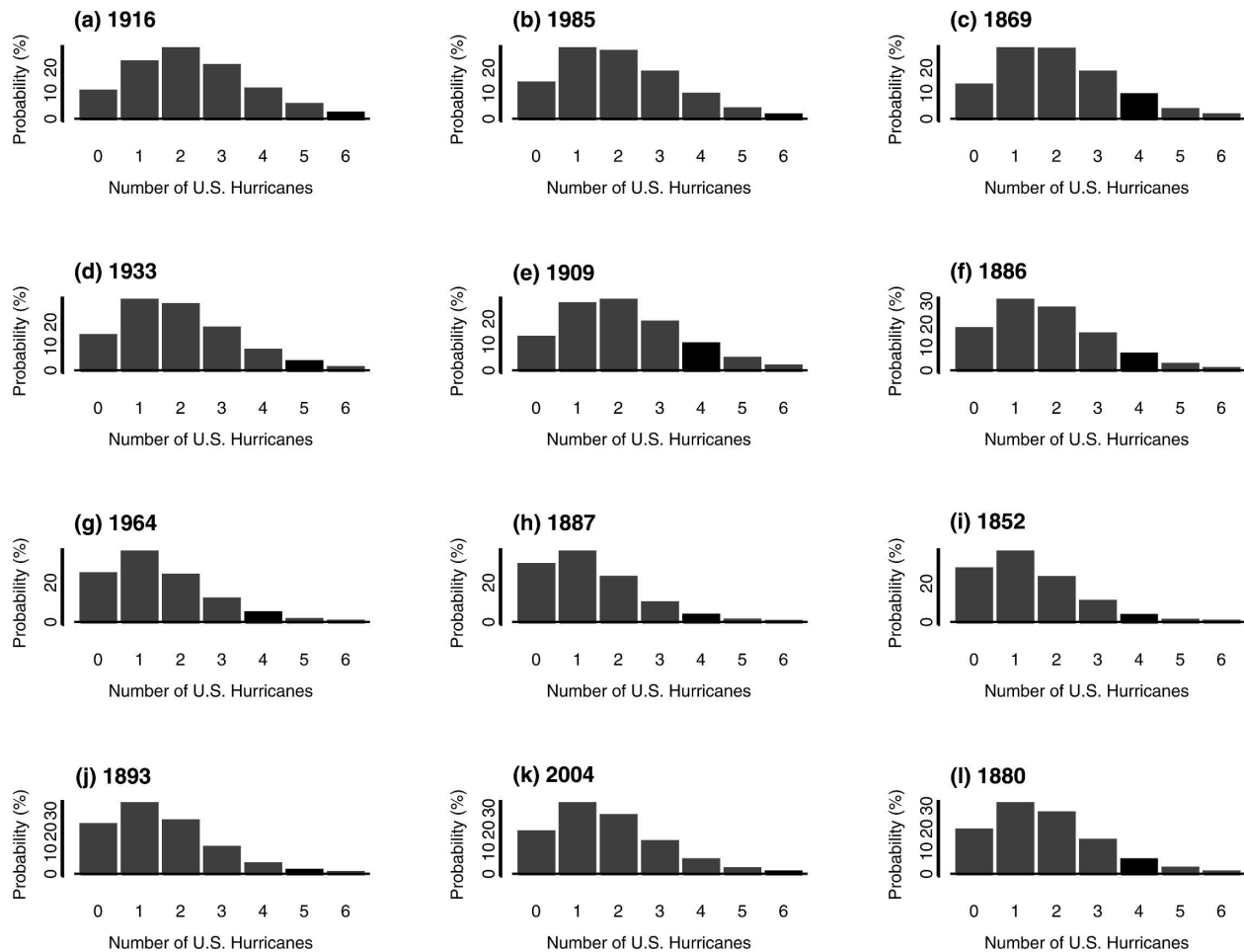
FIG. 10. Same as in Fig. 6, but when the observed count is four or more.

The importance of NAO in portending U.S. hurricane activity is further highlighted by considering early-nineteenth-century activity. The May–June averaged NAO values for 1837 and 1842 were −2.0 and −2.6 standard deviations respectively, which correspond to the lower 97th and 98th percentiles of the 1851–2004 NAO values. Based on hurricane records collated in the Historical Hurricane Information Tool (Bossak and Elsner 2004), we estimate six and four U.S. hurricanes during these 2 yr, respectively. Although there is considerable uncertainty as to the precise number of hurricanes and to the value of NAO, it is likely that activity during these years was above normal when NAO was negative. We use the model to retrodict U.S. hurricane activity for these 2 yr and find that the relative risk of observing five or more hurricanes (three or more hurricanes) during 1837 (1842) is 3 (2) times more likely with NAO at the observed value compared to climatology. The preliminary May–June averaged NAO value for 2005 is −0.565 standard deviations, indicating a 13% increase over climatology of observing three or more U.S. hurricanes this year, assuming the model is correct.

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

Forecast models of landfall activity can be improved by including spatial information. For example, a model that predicts activity regionally (e.g., Gulf Coast would

be able to exploit the spatial correlation structure arising from the space–time nature of hurricane tracks. 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 and 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 Iceland. The SLP variables indicate the state of NAO. Although the Jagger model has yet to be implemented operationally, our aim is to have it probabilistically predict the likely near-coastal paths of hurricanes for an entire season. Toward this end, it is possible to reformulate the model using a Bayesian approach.

APPENDIX

Forecast Model and Data

Use the forecast model below together with the observed data and the initial values to predict the probability of coastal hurricanes during the upcoming season. Download the WinBUGS software. Cut and paste the model, data, and initial conditions into a WinBUGS document (e.g., from a pdf). Add in the current values for the predictors in the data statement. Set the last hurricane count to NA. Use the specification tool, sample monitor tool, and update tool to generate the posterior predictive probabilities for h (the number of coastal hurricanes

```
model; {
for(i in 1: N) {
AMO[i] ~ dnorm(0, 20)
SOI[i] ~ dnorm(0, 1)
h[i] ~ dpois (lambda [i])
log(lambda [i]) <- beta0 + beta1*NAO[i] +
   beta2*SOI[i] + beta3*AMO[i] + beta4*IND[i]
}
beta0 ~ dnorm(0.0, 1.0E-6)
beta1 ~ dnorm(0.0, 1.0E-6)
beta2 ~ dnorm(0.0, 1.0E-6)
beta3 ~ dnorm(0.0, 1.0E-6)
beta4 <- log(p)
p ~ dunif(lower, upper)
}
```

```
# Initial conditions
list(beta0 = 0, beta1 = 0, beta2 = 0, beta3 = 0, p =
   0.9)
data;
list(N = 154, lower=0.80, upper=0.95,
```

Hurricane counts by year, the NA is for the season to be forecast. Counts are for the period July–November.

```
H=c(1, 4, 1, 2, 1, 2, 1, 1, 1, 3, 3, 0, 0, 0, 2, 0, 1, 0, 4, 2,
3, 0, 2, 1, 1, 2, 2, 2, 3, 4, 2, 3, 1, 0, 1, 3, 3, 3, 1, 0, 1,
0, 5, 2, 1, 2, 1, 3, 3, 1, 1, 0, 2, 2, 0, 3, 0, 0, 3, 2, 2, 2,
1, 0, 3, 3, 1, 1, 1, 2, 1, 0, 1, 2, 1, 2, 0, 2, 1, 0, 0, 2, 5,
0, 2, 1, 0, 2, 1, 2, 2, 2, 0, 3, 2, 1, 3, 3, 3, 3, 0, 1, 3, 3,
3, 1, 0, 0, 1, 2, 1, 0, 1, 4, 1, 1, 1, 1, 2, 1, 3, 0, 0, 1, 1,
1, 1, 0, 2, 1, 0, 0, 1, 1, 5, 1, 1, 1, 3, 0, 1, 1, 1, 0, 2, 1,
0, 3, 3, 0, 0, 1, 2, 6, NA),
```

NAO index values averaged over the preseason months of May–June.

```
NAOI=c(-1.575, 0.170, -1.040, -0.010, -0.750,
0.665, -0.250, 0.145, -0.345, -1.915, -1.515,
0.215, -1.040, -0.035, 0.805, -0.860, -1.775,
1.725, -1.345, 1.055, -1.935, -0.160, -0.075,
-1.305, 1.175, 0.130, -1.025, -0.630, 0.065,
-0.665, 0.415, -0.660, -1.145, 0.165, 0.955,
-0.920, 0.250, -0.365, 0.750, 0.045, -2.760,
-0.520, -0.095, 0.700, 0.155, -0.580, -0.970,
-0.685, -0.640, -0.900, -0.250, -1.355, -1.330,
0.440, -1.505, -1.715, -0.330, 1.375, -1.135,
-1.285, 0.605, 0.360, 0.705, 1.380, -2.385, -1.875,
-0.390, 0.770, 1.605, -0.430, -1.120, 1.575, 0.440,
-1.320, -0.540, -1.490, -1.815, -2.395, 0.305,
0.735, -0.790, -1.070, -1.085, -0.540, -0.935,
-0.790, 1.400, 0.310, -1.150, -0.725, -0.150,
-0.640, 2.040, -1.180, -0.235, -0.070, -0.500,
-0.750, -1.450, -0.235, -1.635, -0.460, -1.855,
-0.925, 0.075, 2.900, -0.820, -0.170, -0.355,
-0.170, 0.595, 0.655, 0.070, 0.330, 0.395, 1.165,
0.750, -0.275, -0.700, 0.880, -0.970, 1.155, 0.600,
-0.075, -1.120, 1.480, -1.255, 0.255, 0.725,
-1.230, -0.760, -0.380, -0.015, -1.005, -1.605,
0.435, -0.695, -1.995, 0.315, -0.385, -0.175,
-0.470, -1.215, 0.780, -1.860, -0.035, -2.700,
-1.055, 1.210, 0.600, -0.710, 0.425, 0.155, -0.525,
-0.565),
```

SOI values averaged over the preseason months of May–June.

```
SOID=c(NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, 0.38, 0.06, -0.94,
-0.02, -0.28, -0.78, -0.95, 2.33, 1.43, 1.24, 1.26,
-0.75, -1.5, -2.09, 1.01, -0.05, 2.48, 2.48, 0.46,
0.46, -0.2, -1.11, 0.52, -0.37, 0.58, 0.86, 0.59,
-0.12, -1.33, 1.4, -1.84, -1.4, -0.76, -0.23,
-1.78, -1.43, 1.2, 0.32, 1.87, 0.43, -1.71, -0.54,
```

−1.25, −1.01, −1.98, 0.52, −1.07, −0.44, −0.24,
−1.31, −2.14, −0.43, 2.47, −0.09, −1.32, −0.3,
−0.99, 1.1, 0.41, 1.01, −0.19, 0.45, −0.07, −1.41,
0.87, 0.68, 1.61, 0.36, −1.06, −0.44, −0.16, 0.72,
−0.69, −0.94, 0.11, 1.25, 0.33, −0.05, 0.87, −0.37,
−0.2, −2.22, 0.26, −0.53, −1.59, 0.04, 0.16, −2.66,
−0.21, −0.92, 0.25, −1.36, −1.62, 0.61, −0.2, 0,
1.14, 0.27, −0.64, 2.29, −0.56, −0.59, 0.44, −0.05,
0.56, 0.71, 0.32, −0.38, 0.01, −1.62, 1.74, 0.27, 0.97,
1.22, −0.21, −0.05, 1.15, 1.49, −0.15, 0.05, −0.87,
−0.3, −0.08, 0.5, 0.84, −1.67, 0.69, 0.47, 0.44,
−1.35, −0.24, −1.5, −1.32, −0.08, 0.76, −0.57,
−0.84, −1.11, 1.94, −0.68),

#AMO values averaged over the preseason months
of May–June.

AMO=c(NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,
NA, −0.117, −0.211, −0.333, −0.229, −0.272,
−0.243, −0.148, 0.191, −0.263, −0.239, −0.168,
−0.381, −0.512, −0.338, −0.296, 0.067, 0.104,
−0.254, −0.167, −0.526, −0.096, −0.43, 0.013,
−0.438, −0.297, −0.131, −0.098, −0.046, −0.063,
−0.194, −0.155, −0.645, −0.603, −0.374, −0.214,
−0.165, −0.509, −0.171, −0.442, −0.468, −0.289,
−0.427, −0.519, −0.454, 0.046, −0.275, −0.401,
−0.542, −0.488, −0.52, −0.018, −0.551, −0.444,
−0.254, −0.286, 0.048, −0.03, −0.015, −0.219,
−0.029, 0.059, 0.007, 0.157, 0.141, −0.035, 0.136,
0.526, 0.113, 0.22, −0.022, −0.173, 0.021, −0.027,
0.261, 0.082, −0.266, −0.284, −0.097, 0.097, −0.06,
0.397, 0.315, 0.302, −0.026, 0.268, −0.111, 0.084,
0.14, −0.073, 0.287, 0.061, 0.035, −0.022, −0.091,
−0.22, −0.021, −0.17, −0.184, 0.121, −0.192,
−0.24, −0.283, −0.003, −0.45, −0.138, −0.143,
0.017, −0.245, 0.003, 0.108, 0.015, −0.219, 0.09,
−0.22, −0.004, −0.178, 0.396, 0.204, 0.342, 0.079,
−0.034, −0.122, −0.24, −0.125, 0.382, 0.072, 0.294,
0.577, 0.4, 0.213, 0.359, 0.074, 0.388, 0.253, 0.167),

Indicator predictor for reliable (0) versus less reli-
able records (1)

IND=c(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0))

REFERENCES

Allan, R. J., N. Nicholls, P. D. Jones, and I. J. Butterworth, 1991: A further extension of the Tahiti–Darwin SOI, early SOI results and Darwin pressures. *J. Climate,* **4,** 743–749.

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

Barbieri, M. M., and J. O. Berger, 2002: Optimal predictive model selection. Discussion Paper 02-02, Institute of Statistics and Decision Sciences, Duke University, 28 pp.

Berliner, L. M., C. K. Wikle, and N. Cressie, 2000: Long-lead prediction of Pacific SSTs via Bayesian dynamic modeling. *J. Climate,* **13,** 3953–3968.

Bossak, B. H., and J. B. Elsner, 2004: Plotting early nineteenth century hurricane information. *Eos, Trans. Amer. Geophys. Union,* **85.**

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

Chu, P. S., and X. Zhao, 2004: Bayesian change-point analysis of tropical cyclone activity: The central North Pacific case. *J. Climate,* **17,** 4893–4901.

Congdon, P., 2003: *Applied Bayesian Modelling.* John Wiley and Sons, 530 pp.

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

——, and C. P. Schmertmann, 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.

——, and ——, 2006: Comparison of hindcasts anticipating the 2004 Florida hurricane season. *Wea. Forecasting,* **21,** 184–194.

——, G. S. Lehmiller, and T. B. Kimberlain, 1996: Objective classification of Atlantic basin hurricanes. *J. Climate,* **9,** 2880–2889.

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

——, T. H. 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.-F. Niu, 2001: Secular changes to the ENSO–U.S. hurricane relationship. *Geophys. Res. Lett.,* **28,** 4123–4126.

——, X.-F. Niu, and T. H. Jagger, 2004: Detecting shifts in hurricane rates using a Markov chain Monte Carlo approach. *J. Climate,* **17,** 2652–2666.

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 lan-

guage and program for complex Bayesian modelling. *Statistician,* **43,** 169–178.

——, S. Richardson, and D. J. Spiegelhalter, Eds., 1998: *Markov Chain Monte Carlo in Practice.* Chapman and Hall/CRC, 486 pp.

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.

Hess, J. C., J. B. Elsner, and N. E. LaSeur, 1995: Improving seasonal hurricane predictions for the Atlantic basin. *Wea. Forecasting,* **10,** 425–432.

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.

Jarrell, J. D., P. J. Hebert, and M. Mayfield, 1992: Hurricane experience levels of coastal county populations from Texas to Maine. NOAA Tech. Memo. NWS NHC-46, 152 pp.

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.

Katz, R. W., 2002: Techniques for estimating uncertainty in climate change scenarios and impact studies. *Climate Res.,* **20,** 167–185.

Können, G. P., P. D. Jones, M. H. Kaltofen, and R. J. Allan, 1998: Pre-1866 extensions of the Southern Oscillation index using early Indonesian and Tahitian meteorological readings. *J. Climate,* **11,** 2325–2339.

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 northwestern 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.

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.

Michaelsen, J., 1987: Cross-validation in statistical climate forecast models. *J. Climate Appl. Meteor.,* **26,** 1589–1600.

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

Ropelewski, C. F., and P. D. Jones, 1987: An extension of the Tahiti–Darwin Southern Oscillation index. *Mon. Wea. Rev.,* **115,** 2161–2165.

Saunders, M. A., and A. S. Lea, 2005: Seasonal prediction of hurricane activity reaching the coast of the United States. *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., N. G. Best, W. R. Gilks, and H. Inskip, 1996: Hepatitis B: A case study in MCMC methods. *Markov Chain Monte Carlo in Practice,* W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, Eds., Chapman and Hall/CRC, 21–43.

Vitart, F., and T. N. Stockdale, 2001: Seasonal forecasting of tropical storms using coupled GCM integrations. *Mon. Wea. Rev.,* **129,** 2521–2537.

——, D. Anderson, and T. Stockdale, 2003: Seasonal forecasting of tropical cyclone landfall over Mozambique. *J. Climate,* **16,** 3932–3945.

Wikle, C. K., 2000: Hierarchical space–time dynamic models. *Lecture Notes in Statistics: Studies in the Atmospheric Sciences,* L. M. Berliner, D. Nychka, and T. Hoar, Eds., Springer-Verlag, 45–64.

——, and C. J. Anderson, 2003: Climatological analysis of tornado report counts using a hierarchical Bayesian spatiotemporal model. *J. Geophys. Res.,* **108,** 9005, doi:10.1029/2002JD002806.