# A Dynamic Probability Model of Hurricane Winds in Coastal Counties of the United States

THOMAS JAGGER*

*Department of Statistics, The Florida State University, Tallahassee, Florida*

JAMES B. ELSNER

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

XUFENG NIU

*Department of Statistics, The Florida State University, Tallahassee, Florida*

(Manuscript received 15 May 2000, in final form 13 September 2000)

ABSTRACT

The authors develop and apply a model that uses hurricane-experience data in counties along the U.S. hurricane coast to give annual exceedence probabilities to maximum tropical cyclone wind events. The model uses a maximum likelihood estimator to determine a linear regression for the scale and shape parameters of the Weibull distribution for maximum wind speed. Model simulations provide quantiles for the probabilities at prescribed hurricane intensities. When the model is run in the raw climatological mode, median probabilities compare favorably with probabilities from the National Hurricane Center's risk analysis program ''HURISK'' model. When the model is run in the conditional climatological mode, covariate information in the form of regression equations for the distributional parameters allows probabilities to be estimated that are conditioned on climate factors. Changes to annual hurricane probabilities with respect to a combined effect of a La Niña event and a negative phase of the North Atlantic oscillation mapped from Texas to North Carolina indicate an increased likelihood of hurricanes along much of the coastline. Largest increases are noted along the central Gulf coast.

## 1. Introduction

Landfalling hurricanes are of great social and economic concern (Pielke et al. 1999). Here in the United States, their potential for damage and loss of life rivals the potential for damage and casualties from earthquakes (Diaz and Pulwarty 1997). Knowledge of the past occurrence of hurricanes provides clues about future frequency and intensity of hurricanes at locations along the coast, which is important for land use planning, emergency management, hazard mitigation, insurance applications, and long-term derivative markets.

Empirical and statistical research (Gray et al. 1992; Elsner et al. 1999; Elsner et al. 2000a) have identified factors that contribute to conditions favorable for hurricanes over the North Atlantic basin, which includes the Caribbean Sea and the Gulf of Mexico. Research also shows that these factors influence the occurrence of hurricanes differentially depending on the particular region of the North Atlantic. For instance, the influence of an El Niño event on the frequency of hurricanes over the entire basin is significant, but its influence on the frequency of hurricanes forming over subtropical latitudes (approximately 25°–35°N) appears to be small. Additional factors are usually needed to explain the climatic variation of hurricane activity locally (Lehmiller et al. 1997).

It is demonstrated in Elsner et al. (1996) and Elsner and Kara (1999) that weaker, baroclinic-type hurricanes tend to cluster at higher latitudes (north of approximately 30°N) over the western North Atlantic, and that their overall frequency is inversely related to the frequency of stronger, deep tropical hurricanes. A recent study that combines historical and geological data (Elsner et al. 2000b), finds that climate conditions associated with strong hurricanes along the Gulf coast are associated with a negative (weak) phase of the North Atlantic oscillation (NAO). Conversely, major hurricane activity along the northeast coast is associated with a positive (strong) phase of the NAO.

* Current affiliation: MathSoft Company, Seattle, Washington.

*Corresponding author address:* James B. Elsner, Dept. of Geography, The Florida State University, Tallahassee, FL 32306.
E-mail: jbelsner@elsner.coss.fsu.edu

Results from these studies suggest a complicated relationship between hurricane activity along the coast of the United States and climate. For instance, the occurrence of a La Niña event in the eastern tropical Pacific increases the probability of tropical storms, but such an occurrence may have less of an influence on the probability of major hurricanes, especially when the NAO is in a negative phase. Thus, the annual probability distribution of hurricane winds along a stretch of coastline is a function of both hurricane intensity and climate factors.

Various techniques for estimating annual probabilities (or return periods) of tropical cyclones have been proposed in the meteorological literature (Neumann 1987; Darling 1991; Rupp and Lander 1996; Chu and Wang 1998). With the exception of Darling (1991), the basic strategy is to fit an extreme-value distribution to the historical maximum wind speeds (or minimum pressures) for tropical cyclones that have affected a particular area. Determining the area typically involves estimating the tangential wind profile radially outward from the known center of circulation to the radius of maximum wind. This can be done in a variety of statistical or theoretical ways. A series of tropical cyclones is then simulated based on local occurrence rates.

A common method is to choose a uniform area over which to consider historical storm occurrences (e.g., the milepost system used in National Oceanic and Atmospheric Administration Technical Report NWS 23). This has the advantage of allowing direct comparisons of probabilities from neighboring regions. The limitation is that the areas do not correspond to distinct political districts. Here we choose to model hurricane activity at the county level. Although this method has the disadvantage of unequal size areas, it has the advantage that results are more in line with public awareness and the political decision-making process. Moreover, county-by-county hurricane-experience-level data have already been collected.

The annual probability models are useful in establishing a baseline climatic distribution or ''climatology'' of extreme wind events, but are predicated on a static distribution of events over time. That is, the methods provide estimates of hurricane probabilities without regard to changes in climate. Thus, the annual probability of a hurricane strike along the Louisiana coast using a static climatology is the same regardless of whether there exists an El Niño event.

Because information is available on the relationship between climate patterns and hurricane activity, some of which can be used to predict activity several months in advance, it is possible to develop landfalling hurricane models to forecast probabilities against the baseline climatology. The purpose of the current paper is to describe a technique that can be used to model annual tropical cyclone intensities statistically. The model is unique in that it considers covariate information directly in assigning probabilities. The probabilities are adjusted based on the changing climate.

The paper is outlined as follows. In section 2 we describe the source of the hurricane data and our preprocessing procedures. In section 3 we detail the modeling strategy, beginning with a discussion of the Weibull distribution and explaining how the current algorithm extends the use of this distribution. We also look at model assumptions and how the parameters are estimated using a linear regression. In section 4 we show results from the model. We examine the probabilities generated from the model in the raw and conditional climatological modes and provide model intercomparisons. The probability curves for Miami-Dade County and the probability distribution for hurricane winds in coastal counties from Texas to North Carolina are shown. A summary with remarks about potential future improvements to the model is given in section 5.

## 2. Hurricane data

As mentioned, statistical models that assign a wind speed probability to locations along the coastline require an estimation of the lateral extent of the maximum winds within each storm. We circumvent this by considering the occurrence of hurricanes over an entire county. An aggregation of probability at the county level is possible by making use of the dataset of Jarrell et al. (1992). These data are a best guess of a county's hurricane history over the period of 1900–97, and are an extension of the climatologies developed by Hebert and Taylor (1975). Here we list the guidelines used in Jarrell et al. (1992) and explain how we convert this information into a range of wind speeds for the counties.

### *a. Guidelines used in Jarrell et al. (1992)*

The dynamic probability model developed here makes use of the hurricane climatology prepared in Jarrell et al. (1992). We begin with a description of the guidelines used by the authors in preparing this climatology.

1) First the authors assigned a Saffir–Simpson-scale (Simpson 1974) number (1–5) to hurricanes in the North Atlantic best-track Hurricane Dataset (HURDAT) based primarily on estimated central pressure values at the time of landfall. The best-track dataset is a compilation of the 6-hourly positions and intensities of tropical cyclones back to 1886 (Neumann et al. 1999). Some subjectivity is inherent in this classification, particularly with hurricanes during earlier years and with storms moving inland over a sparsely populated area. Thus some hurricanes at the borderline between two Saffir–Simpson-scale numbers could be classified either way. Intensity values were sometimes modified by storm surge estimates, in which case the central pressure may not agree with

TABLE 1. Hurricane data. Values are the Saffir–Simpson category by county for Hurricane Alicia in 1983. The data are reproduced from a portion of appendix C of Jarrell et al. (1992). The raw numbers indicate a direct hit, and the numbers in parentheses indicate an indirect hit. Note that Alicia was considered a category-3 hurricane at landfall but a category-2 hurricane when it hit Harris County on the other side of Galveston Bay.

| Year | Matagorda | Brazoria | Galveston | Harris | Chambers | Jefferson |
|---|---|---|---|---|---|---|
| 1983 | (3) | 3 | 3 | 2 | 3 | (3) |

the scale assignment. Beginning with the 1996 hurricane season, scale assignments are based on maximum winds.

2) Second, the authors determined which coastal counties received direct hits and which received indirect hits. A direct hit was defined as the innermost core regions, or ''eye,'' moving over the county. Each hurricane was judged individually, but a general rule of thumb was applied in cases of greater uncertainty. That is, a county was regarded as receiving a direct hit when all or part of a county fell within $R$ to the left of a storm's landfall and $2R$ to the right (with respect to an observer at sea looking toward shore). The radius to maximum winds $R$ is defined as the distance from the storm's center to the circumference of maximum winds around the center. The determination of an indirect hit was based on a hurricane's strength and size and on the configuration of the coastline. In general, it was determined that the counties on either side of the direct-hit zone that received hurricane-force winds or tides of at least 1–2 m above normal were considered to be indirectly hit. Subjectivity was also necessary here because of storm paths and coastline geography.

### *b. Converting to a wind speed range*

Appendix C of Jarrell et al. (1992) is a matrix of county strikes by Saffir–Simpson category based on the above guidelines. The 175 columns of the matrix are the coastal counties from Cameron, Texas, to Washington, Maine, and the 91 rows are the years from 1900 to 1990, inclusive. Table 1 shows the entries of the matrix for counties affected by Hurricane Alicia in 1983. The values in the table indicate that Brazoria, Galveston, and Chambers Counties felt the direct impact of a category-3 hurricane, and Harris County, located across the bay, felt the direct impact of a category-2 hurricane. Matagorda County to the south and Jefferson County to the north experienced the indirect impact of Alicia at full fury. Parentheses are used to indicate an indirect hit. Note that these data contain information on the general size of the hurricane as it made landfall (i.e., Alicia's impact stretched from Matagorada County to Jefferson County). Thus, in using these data to model wind, it is not necessary to model the size of the storm as is the case when using center positions only.

TABLE 2. Wind speed ranges. The first column lists the codes used in appendix C of Jarrell et al. (1992). Values in the second column are the interpreted range of possible maximum Saffir–Simpson category experienced in the county, and values in the third column are the corresponding maximum sustained (1 min) near-surface (10 m) wind speed ranges (m s$^{-1}$). In developing the model, we use speeds corresponding to the minimal wind speeds of the Saffir–Simpson categories.

| Symbol | Saffir–Simpson bounds | Wind speed range (m s$^{-1}$) |
|---|---|---|
| (1) | [0,1) | <33 |
| 1 | [1,2) | 33–42 |
| (2) | [1,2) | 33–42 |
| 2 | [2,3) | 42–50 |
| (3) | [1,3) | 33–50 |
| 3 | [3,4) | 50–58 |
| (4) | [1,4) | 33–58 |
| 4 | [4,5) | 58–69 |
| (5) | [1,5) | 33–69 |
| 5 | [5,∞) | ≥69 |

The climatological data described above consist of values that should be interpreted as ranges in wind speed experienced somewhere in the county. Table 2 shows ranges used in the current study. First the Saffir–Simpson category is interpreted over an interval using open and closed bounds. These bounds are then used to interpret a range of wind speeds. For example, the symbol ''(3)'' is interpreted as a range of 1-min, 10-m wind speeds between 33 and 58 m s$^{-1}$ experienced over at least a part of the county, and a ''3'' is interpreted as a range of winds between 50 and 58 m s$^{-1}$.

The climatological range of hurricane wind speeds is a useful reference of the hurricane history of individual coastal counties, a level at which land use planning, hazard mitigation, and emergency management decisions are frequently made. The hurricane-experience record is extended to 1997 based on written reports by the National Hurricane Center's hurricane specialists who examined the storms' impact with regard to the criteria listed in Jarrell et al. (1992).

## 3. Modeling strategy

### *a. The Weibull distribution*

Batts et al. (1980) suggest that the maximal wind speed over an area in a given year be modeled using a Weibull distribution. The survival function (1.0 minus the cumulative distribution function) for the Weibull distribution is an exponential curve. Let $V$ be the unknown yearly maximum wind speed, and $\upsilon$ some known value, then the survival function for the Weibull distribution is

$$S(\upsilon) = \Pr(V > \upsilon) = e^{-(\upsilon/b)^a},$$

where $a$ is the shape parameter and $b$ is the scale parameter.

To apply the formula, Johnson and Watson (1999) use wind speeds (they also consider wave and surge heights separately) from tropical cyclones over an area and use the maximum likelihood estimator (MLE) to calculate both the estimated values $\hat{\boldsymbol{\theta}} = (\hat{a}, \hat{b})$ along with a covariate matrix of the parameters $\boldsymbol{\Sigma}$. The goal of Johnson and Watson (1999) is to estimate many-year return periods. For example, to estimate 50-yr return periods, they sample the parameters as if they were normally distributed with a mean of $\boldsymbol{\theta}$ and covariance $\boldsymbol{\Sigma}$. Next they generate simulated 50-yr wind speeds. Last, they compute the return period as $n/m$ where $n$ is the total number of samples and $m$ is the number of samples whose wind speeds exceed a certain value.

Our algorithm extends this approach in three directions:

1) it makes use of ranged data,
2) it examines coastal counties in the United States (Texas to North Carolina), and
3) it conditions the parameters using covariates.

We consider the parameters to be variables, with values changing from year to year. In the simplest case we determine the parameters from a linear regression onto two covariates including the El Niño–Southern Oscillation (ENSO) and the NAO. This allows us to study the model and ask ''what if'' questions. Also, given that some covariates can be forecast, they may be used to make predictions of upcoming hurricane activity along the coast. Because we regress on the scale and shape parameters, results may show an increase in hurricane activity but a decrease in major hurricane activity. Without regressing on the shape parameter, the resulting change in probability is uniform across the wind speeds. Moreover, because we consider each coastal county separately we can tailor our forecasts for that county. That is, the regression parameters (and thus the change in probabilities) can vary in sign and magnitude from county to county. In this paper we consider only the section of U.S. coastline most susceptible to hurricanes (Texas to North Carolina).

### *b. Assumptions*

The algorithm we use is based on the assumption that the yearly maximal wind speed due to tropical storms in a given area is a Weibull distribution. Note that if $V$ follows a Weibull distribution, then so does $V^x$ for any power $x$. Thus it makes no difference if we model wind speed, its force, or its power. As discussed in section 5, future improvements to the algorithm could consider modeling the maximum intensity of each storm as a Weibull distribution.

We further assume that the estimated parameters are the true parameters, and that these quantities are random and normally distributed. From this assumption, we derive confidence regions for the parameters. This approach works fine for the scale parameter of normally distributed quantities but is not true for other distributions. Also, this method assumes that the confidence region is an ellipsoid, which is an additional restriction. However, for large datasets, this assumption is valid because the MLE is approximately normally distributed.

Also, we model the parameters using a linear regression. In fact we could use a generalized linear regression or even a nonlinear regression. However, for this study, we consider only a simple linear regression of the parameters onto the covariates.

### *c. Determining the parameters using linear regression*

Our approach to estimating annual exceedence probabilities is to use the maximum likelihood estimator to determine a linear regression for the parameters $a$ and $b$ based on some covariate information. We will assume that we are using different covariates for both $a$ and $b$ but in practice they are usually the same.

Now because we are using linear regression,

$$a = \sum_{j=1}^{p_a} \theta_j X_j^{(a)}, \quad \text{and} \quad b = \sum_{j=1}^{p_b} \theta_{p_a+j} X_j^{(b)},$$

where $X_i^{(a)}$ and $X_i^{(b)}$ are the $p_a$ shape and $p_b$ scale predictors, respectively. We denote the $n$ observations from predictor $X_j^{(a)}$ as $x_{1,j}^{(a)}, \ldots, x_{n,j}^{(a)}$ and likewise for $X_j^{(b)}$.

The $i$th observation of the predictors is denoted $x_i^{(a)}$ with $p_a$ components $[x_{i,1}^{(a)}, \ldots, x_{i,p_a}^{(a)}]$ and $x_i^{(b)}$ with $p_b$ components $[x_{i,1}^{(b)}, \ldots, x_{i,p_b}^{(b)}]$; we associate one value for our shape and scale parameters, $a_i$ and $b_i$. These are the predicted values of the parameters based on covariate information.

Using matrix notation, we have $\mathbf{a}$ and $\mathbf{b}$ being column vectors, $\mathbf{X}^{(a)}$ and $\mathbf{X}^{(b)}$ as $n \times p_a$ and $n \times p_b$ covariate matrices, with $\boldsymbol{\theta} = (\boldsymbol{\theta}_a, \boldsymbol{\theta}_b)$ being a column vector of parameters, divided into two smaller column vectors of size $p_a$ and $p_b$, where

$$a_i = x_i^{(a)}\boldsymbol{\theta}_a, \qquad \mathbf{a} = \mathbf{X}^{(a)}\boldsymbol{\theta}_a, \qquad b_i = x_i^{(b)}\boldsymbol{\theta}_b, \quad \text{and}$$
$$\mathbf{b} = \mathbf{X}^{(b)}\boldsymbol{\theta}_b.$$

Further let us assume that our wind speed data for each observation are a closed interval $[l_i, u_i)$ for $i = 1, \ldots, n$. Assuming that the yearly maximal wind speeds are from the Weibull distribution we can derive the MLE for $\boldsymbol{\theta}$. First let

$$\begin{aligned} f(a, b, l, u) &= \log[\Pr(l \le V < u \mid a, b)] = S(l) - S(u) \\ &= \log[e^{-(l/b)^a} - e^{-(u/b)^a}]. \end{aligned}$$

then

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} f(a_i, b_i, l_i, u_i) = \sum_{i=1}^{n} f(x_i^{(a)}\boldsymbol{\theta}_a, x_i^{(b)}\boldsymbol{\theta}_b, l_i, u_i).$$

In the initial estimator we are using a linear link function, separately for each parameter. This may be relaxed in the future; one may choose to use a log link function, for example, where $\log(a_i) = x_i^{(a)}\boldsymbol{\theta}_a$.

To find the maximum likelihood estimator, we determine the gradient (first derivative) and Hessian (second derivative) of the log likelihood function with respect to $\boldsymbol{\theta}$. This task is made simpler by first deriving the gradient and Hessian of the log likelihood ratio for a single observation, $f(a, b, l, u)$ with respect to $(a, b)$. Let

$$g(a, b, l, u) = \begin{bmatrix} \dfrac{\partial f(a, b, l, u)}{\partial a} \\ \dfrac{\partial f(a, b, l, u)}{\partial b} \end{bmatrix} \quad \text{and}$$

$$h(a, b, l, u) = \begin{bmatrix} \dfrac{\partial f(a, b, l, u)}{\partial a^2} & \dfrac{\partial f(a, b, l, u)}{\partial a \partial b} \\ \dfrac{\partial f(a, b, l, u)}{\partial b \partial a} & \dfrac{\partial f(a, b, l, u)}{\partial b^2} \end{bmatrix}.$$

From this we can write both the gradient vector $\boldsymbol{\nabla}\ell(\boldsymbol{\theta})$ and the negative of the Hessian or the observed Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$ using the chain rule for differentials, that is, $d\ell(\boldsymbol{\theta}) = d\boldsymbol{\theta}' g(a, b, l, u)$, and $dg(a, b, l, u) = h(a, b, l, u) d\boldsymbol{\theta}$ as

$$\boldsymbol{\nabla}\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \begin{bmatrix} & 0 \\ x_i^{(a)\prime} & \vdots \\ & 0 \\ 0 & \\ \vdots & x_i^{(b)\prime} \\ 0 & \end{bmatrix} g(a_i, b_i, l_i, u_i)$$

$$\mathbf{I}(\boldsymbol{\theta}) = \sum_{i=1}^{n} \begin{bmatrix} & 0 \\ x_i^{(a)\prime} & \vdots \\ & 0 \\ 0 & \\ \vdots & x_i^{(b)\prime} \\ 0 & \end{bmatrix} h(a_i, b_i, l_i, u_i) \begin{bmatrix} x_i^{(a)} & 0 \cdots 0 \\ 0 \cdots 0 & x_i^{(b)} \end{bmatrix},$$

where $x_i^{(a)}$ and $x_i^{(b)}$ are $1 \times p_a$ and $1 \times p_b$ matrices (row vectors), respectively.

From these equations, $\hat{\boldsymbol{\theta}}$ is found by solving the nonlinear equation that results from setting the score equations to zero, that is, solving $\boldsymbol{\nabla}\ell(\boldsymbol{\theta}) = 0$. The covariance matrix $\boldsymbol{\Sigma}$ can be estimated from the observed Fisher information evaluated at $\hat{\boldsymbol{\theta}}$.

Once $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\Sigma}}$ are determined, probabilities can be estimated. First generate $N$ independent samples of the regression parameters using the multivariate normal distribution, MVN($\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}}$). Next, for each sampled regression parameter and a given set of predictors, calculate the Weibull scale and shape parameters. Now for every wind speed of interest, say $v_1, \ldots, v_k$, the probability from each sample is $S(v_i)$ for $i = 1, \ldots, k$, generating $N \times k$ data points. For each wind speed, order the points to determine a confidence interval for the probabilities. For our purpose we use speeds corresponding to the minimal wind speeds of the Saffir–Simpson categories.

Assuming that the predictors are known, the above procedure gives a reasonable set of confidence intervals for any hurricane intensity of interest in a given county. One must be careful and note that the probabilities assume a continuance of the conditions stated by the predictors. This also allows us to make single-year probability forecasts, based on the expected values for the predictors, using $S(v_i)$ for the probabilities. If the predictors themselves are random, one can sample both the parameters and predictors in determining the Weibull shape and scale parameters.

## 4. Results and validations

The above algorithm provides a way to model the annual probabilities at values of wind speeds corresponding to various hurricane intensity levels. It does this as follows.

1) First it selects a set of wind speeds and quantiles.
2) Then, for each wind speed, it generates probability samples by evaluating the Weibull survival curve using $10^4$ randomly assigned parameter pairs and determines the quantiles.
3) Last, for each quantile, it generates values that can be plotted on a velocity versus probability graph.

The dynamic probability model resulting from this algorithm can be used in two ways, and results are shown here for both. We first show results using the model in the raw climatological mode. This means that the model provides annual exceedence probabilities of experiencing winds from a hurricane somewhere in the county at various intensities without regard to climate factors. Model probabilities in the raw climatological mode are compared with probabilities generated from other models. We then show results using the model in the conditional climatological mode. This means that the model provides exceedence probabilities conditioned on climate factors. Because the model can be run in a conditional climatological mode, it is referred to as a ''dynamic'' probability model. The probabilities are updated from their raw climatological values depending on the strength of the climate anomalies. The difference in probabilities between the raw climatology and the conditional climatology provides information that can be used to adjust a forecast of future hurricane activity.

### *a. Raw climatological probabilities*

Here we present output from the model using the raw climatological mode. Figure 1 shows the annual exceedence probabilities for Miami-Dade County as a function of wind speed. At the wind speeds plotted, the value of the horizontal tick is the median of the Weibull survival function using $10^4$ sample pairs of the scale and shape parameters. Wind speeds are plotted in units

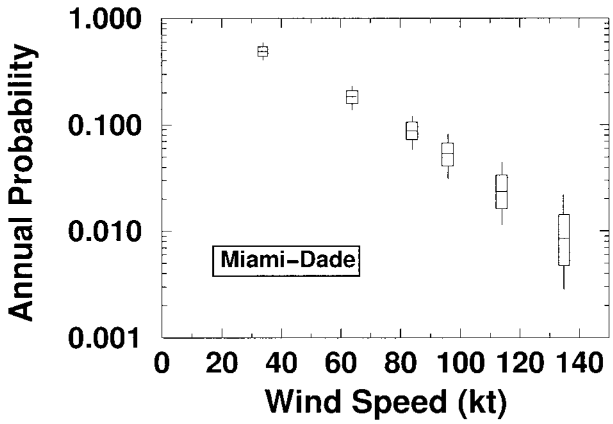


FIG. 1. Annual exceedence probabilities for tropical cyclone winds over Miami-Dade County using the dynamic probability model. The bar and whisker symbol indicates the median, quartile, and decile values from the model simulations. Wind speeds are given in knots as is used in operational meteorology (1 kt = 0.515 m $s^{-1}$).

of knots (kt) as is the convention in operational meteorology (1 kt = 0.515 m $s^{-1}$). Fifty percent of the simulated probabilities for winds exceeding 80 kt from a hurricane are less than 0.1. In other words, the median probability of a hurricane affecting at least a portion of the county with 80-kt winds or higher in any one year is approximately 10%. The individual quartile and decile intervals around this probability are shown using a bar and whiskers. The median probability of Miami-Dade experiencing a 100-kt wind somewhere in the county during any year is approximately 4%. As expected, the quartile and decile intervals are larger for stronger winds.

To examine the geographic distribution of annual exceedence probabilities, we run the model for hurricane winds (64 kt or greater) for coastal counties from Texas through North Carolina (Fig. 2). In counties without a sufficient number of hurricanes during the 98-yr period, the maximum likelihood estimator does not converge for the scale and shape parameters, and thus no probabilities are given. As expected, largest annual probabilities in the range of 15%–25% are found in southern Florida and along portions of the western Gulf coast and eastern North Carolina. Moderate probabilities are noted for the central Gulf coast and portions of North Carolina. Lowest probabilities, generally less than 10%, are noted over portions of South Carolina and along the northern stretch of peninsular Florida. The magnitude of the wind probability gradient along the west coast of the Florida peninsula increases with wind speed [see Elsner and Kara (1999) for a discussion]. The gradient is weak for tropical storm intensities but large for major hurricane intensities.

The geographic distribution of exceedence probabilities shown in Fig. 2 matches closely the variation of tropical cyclone frequencies given in Fig. 12 of Neu-

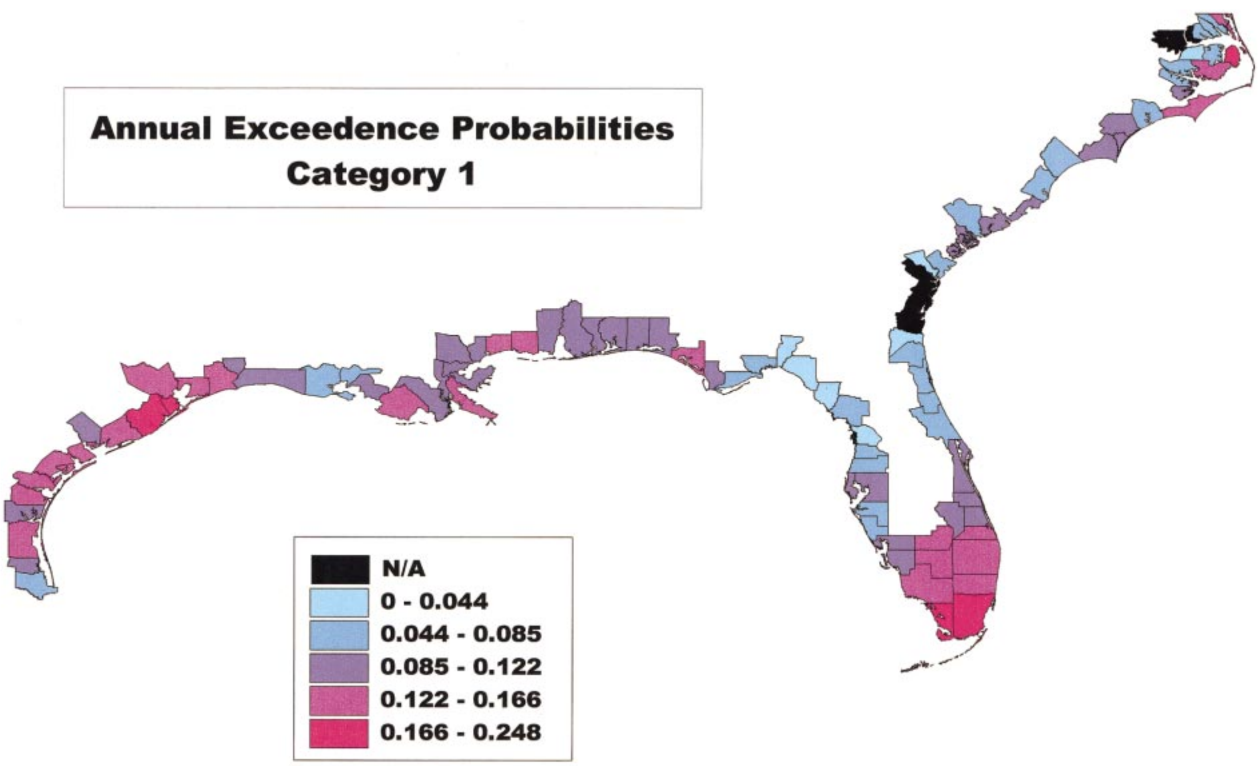


FIG. 2. Annual exceedence probabilities for category-1 hurricane winds in coastal counties from Texas to North Carolina using the dynamic probability model run in the raw climatological mode.

TABLE 3. Hurricanes affecting Galveston County, Texas (1900–97). Values are the Saffir–Simpson category with codes used in Jarrell et al. (1992).

| Year | Code | Year | Code | Year | Code |
|---|---|---|---|---|---|
| 1900 | 4 | 1909 | 3 | 1915 | 4 |
| 1921 | (2) | 1932 | (4) | 1941 | 3 |
| 1942 | 1 | 1943 | 2 | 1945 | (2) |
| 1947 | 1 | 1949 | 2 | 1959 | 1 |
| 1961 | (4) | 1963 | 1 | 1971 | (1) |
| 1983 | 3 | 1989 | 1,1 | | |

mann et al. (1999). They note the largest frequencies of hurricanes over eastern North Carolina, southern Florida, central Texas, and southeastern Louisiana as does the current approach. In contrast, however, they indicate that the local maximum over southeastern Louisiana is larger than the local maximum over central Texas. This discrepancy is partly explained by the differences in methodology. Neumann et al. (1999) determine the frequency of tropical cyclones whose center passes within 75 nautical miles of the sampling point. For southeastern Louisiana, they use a sampling point near the eastern tip of Plaquemines Parish, which extends well into the Gulf of Mexico. Thus they likely include hurricanes that are not in our direct landfall dataset (e.g., Chouinard and Liu 1997). The discrepancies between results using the two methodologies are expected to be smaller for coastal locations that do not extend seaward.

Although counties with a larger area can anticipate more frequent hits, the historical occurrence of hurricane impacts in counties along the coast indicates this factor is not very important (Whitehead 1999) because a hurricane's impact is often relatively large when compared with the size of an individual county, so that a hurricane rarely affects only a single county. Furthermore, most counties are close to average size. A region of the coast where there is a significant positive correlation between the number of hurricane landfalls and county size is along Florida's east coast. Here the largest counties are in the southern part of the state—a part that extends eastward toward the hurricane breeding grounds. Thus, along the east coast of Florida, the effect of county size is compounded by the effect of latitude and by the geometry of the coastline. All else being equal (coastline geometry, size of area, etc.), counties located closer to the hurricane breeding grounds will get hit with a greater frequency.

### *b. Validation*

First we compare model probabilities with empirical probabilities using the same climatological data used by the algorithm to choose the model. Table 3 lists hurricane experience levels for Galveston County, Texas, by Saffir–Simpson scale using the codes of Jarrell et al. (1992) over a 98-yr period (1900–97). The Galveston tragedy of 1900 was caused by a direct hit of a category-4 hurricane. Galveston was indirectly hit by a category-4 hurricane in 1961 (Carla) and directly hit by two category-1 hurricanes in 1989 (Chantal and Jerry). From this table, the empirical probability of experiencing hurricane winds somewhere in Galveston is 0.173 (17/98), and the probability of experiencing major hurricane winds is 0.051 (5/98). Note that we count only one hit per year, and for major hurricanes we exclude indirect hits based on the definitions of ''direct'' and ''indirect'' given in section 2. These values compare favorably to model probabilities of 0.185 and 0.045 for hurricanes and major hurricanes, respectively.

More important, we compare probabilities from our dynamical probability model with those from two different models. The values we use for comparison are taken from Fig. 3 of Darling (1991). It is necessary to make a couple of adjustments to our values so they can be compared directly with the values in this figure. The probabilities obtained from the dynamical probability model represent exceedence levels for the county as a whole. Thus an annual hurricane probability represents the probability of experiencing winds of hurricane intensity or greater *somewhere* in the county. This approach is different from other models. For example, the probabilities plotted in Darling (1991) represent the annual exceedance probabilities for winds at a specific location (Turkey Point Power Plant in southern Miami-Dade County). Furthermore, the values in Darling (1991) are expressed as 10-min, 10-m wind speeds, whereas our values are expressed as 1-min, 10-m speeds based on the definitions used by the National Hurricane Center (NHC). The conversion from a 1-min wind speed to a 10-min wind speed is obtained by multiplying the 1-min speed by 1.15 (Darling 1991).

The reduction of our area probability to a point probability is less straightforward. To do this for Turkey Point we utilize the climatology data from the best-track dataset over the period of 1886–1999. The best-track file, maintained by the NHC, contains latitude and longitude of the estimated center position of each tropical cyclone at 6-h intervals (Jarvinen et al. 1984). Maximum wind speeds in knots are given along with central pressures for recent storms. The radius to maximum winds $R$ defines the swath of hurricane winds across the county. The average $R$ is a function of wind speed and latitude (Neumann 1987). For Turkey Point, the average $R$ for hurricane intensity (64 kt) is 21 nautical miles.

Historically, the number of hurricanes passing within an average $R$ distance of Turkey Point is 14; the number of hurricanes passing within this distance of Miami-Dade County is 29. The ratio (0.48) of these two frequencies is taken as the scaling factor for reducing the area probability to a probability at Turkey Point. Taking the wind speed conversion and scaling factor into account, we determine that our 64-kt, 1-min wind speed probability of 0.185 for Miami-Dade translates to a 0.089 probability ($0.185 \times 0.48$) of a 55.6-kt, 10-min wind at Turkey Point. This point probability compares favorably with a probability of 0.079 from Neumann (1987) and 0.089 from Darling (1991). For 114-kt hurricanes, our 1-min wind speed probability of 0.024 for the county translates into a 0.0120 probability (0.024 × 0.5) of a 100-kt, 10-minute wind at Turkey Point. This value is in comparison with a probability of 0.0079 from Neumann (1987) and 0.0040 from Darling (1991). Overall our model gives probabilities that are very close to those of Neumann (1987), who also uses a Weibull distribution, but the probabilities are too high for stronger wind speeds when compared with Darling (1991).

As noted in Darling (1991), the problem with the statistical approaches is likely related to fitting a distribution to data that are clumped at lower speeds. The effect of this is to lift the tail of the distribution, resulting in an overestimation of rare events. Thus, although it is likely that the current approach produces an overestimation of the probability at the highest winds, the model is useful for moderate hurricane intensities (categories 1–3 on the Saffir–Simpson scale). Moreover, the distributional assumption allows us to condition the probabilities on climate factors, as discussed next.

### *c. Conditional climatological probabilities*

The novelty of the dynamical probability model is that it can be run in a conditional climatological mode (see also Murnane et al. 2000). In short, the model generates exceedence probabilities that are conditioned on climate factors by modeling the parameters with linear regression. Here we examine the influence of the ENSO and the NAO on hurricane wind speed probabilities. A La Niña event over the Pacific Ocean is associated with an increase in the annual probability of one or more hurricanes reaching the United States (Bove et al. 1998; Elsner and Kara 1999), whereas a negative NAO is associated with a greater number of strong hurricanes along the Gulf coast (Elsner et al. 2000b).

During a mature La Niña episode (cold ocean conditions in the eastern tropical Pacific) the sea level pressure (SLP) pattern features positive anomalies across the central and eastern Pacific and negative anomalies over Australia and Indonesia. This pattern results in positive values of the Southern Oscillation index. Monthly SLPs at Darwin, Australia, serve as an indication of the maturity of the La Niña episode. SLP values are available online from the Climatic Research Unit (CRU) of the University of East Anglia (Ropelewski and Jones 1987). Here we average the monthly values over the calendar year for the period of 1900–97.

Negative anomalies of the NAO are associated with an SLP pattern that features lower pressures across the subtropical North Atlantic and higher pressures centered over Iceland. Monthly SLPs over Reykjavik, Iceland, (Jones et al. 1997) serve as an indication of the strength of the NAO (also available online from the CRU). Here we average the monthly values over the calendar year for the period of 1900–97. Thus, annual values of Darwin and Reykjavik SLP provide covariate information for determining the scale and shape parameters using linear regression as part of the dynamic probability model. The annual average provides a less noisy signal of the NAO as compared with monthly or seasonal values. Future work will focus this relationship between the NAO and hurricane landfalls by examining monthly indices.

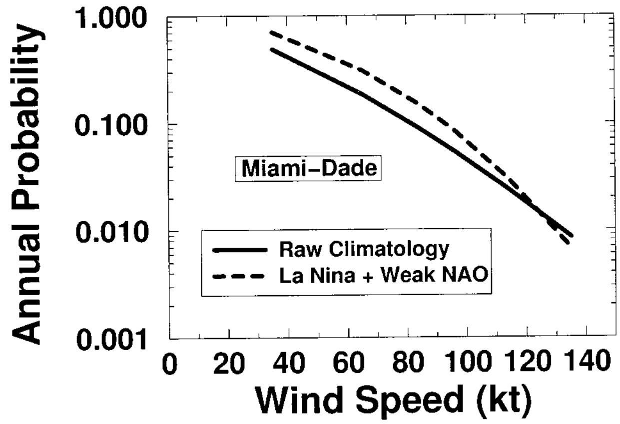


FIG. 3. Median annual exceedence probability curves for hurricane winds over Miami-Dade County using the dynamic probability model in the raw (solid line) and conditional (dashed line) climatological modes. The conditional probability curve is based on a mature La Niña episode and a weak (negative) phase of the NAO.

To show the statistical effect of these two covariates on annual hurricane probabilities, we choose values from a distribution of the Weibull parameters that correspond to two standard deviations of the covariates. Because we are interested in the influence of a La Niña event and a negative phase of the NAO, we use $-1.43$ and $+4.84$ hPa for the SLP anomalies at Darwin and Reykjavik, respectively. Figure 3 compares the median annual exceedence probabilities for Miami-Dade County using the raw and conditional climatological modes. A mature La Niña event coupled with a relaxed NAO is associated with above-average hurricane and above-average major hurricane probabilities. In particular, the annual probability of a 1-min, 65-kt-or-greater wind increases from 0.185 to 0.307 and from 0.054 to 0.083 for 100-kt-or-greater winds. The increases are consistent with the research cited above that indicates increased hurricane activity over the United States during La Niña episodes and a greater frequency of Gulf-coast major hurricanes during a relaxed NAO. Note that the model indicates a nonlinear change in probabilities as a function of wind speed, and, for strongest winds ($\geq$120 kt), the conditional probabilities are smaller than the raw probabilities.

To examine the geographic distribution of the changes in hurricane probabilities based on the covariates, we run the model in the conditional climatological mode for the counties shown in Fig. 2. The raw climatological probabilities are then subtracted from these conditional

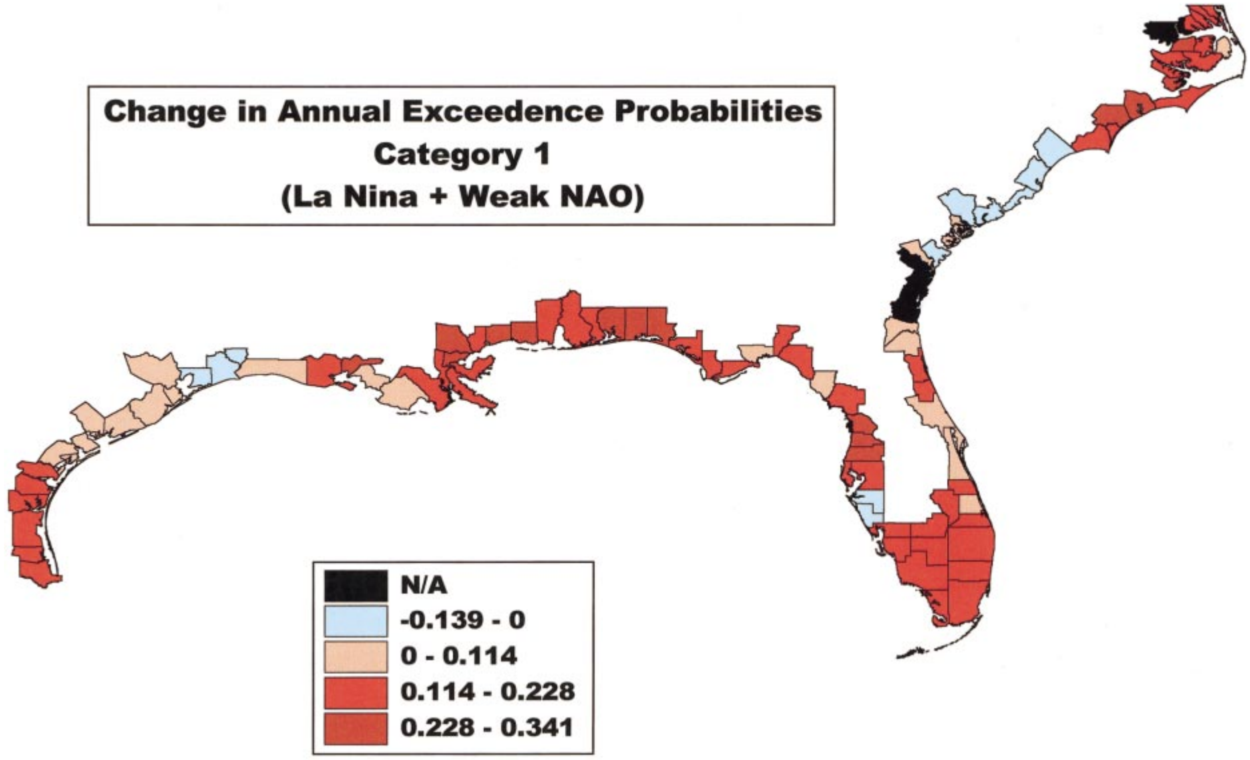


FIG. 4. Change in annual exceedence probabilities for category-1 hurricane winds in coastal counties from Texas to North Carolina using the dynamic probability model. The change is computed by subtracting the raw climatological probabilities from the conditional probabilities. The conditional probabilities are based on a mature La Niña episode plus a weak (negative) phase of the NAO.

values and are displayed in Fig. 4. Positive values (red shading) indicate an increase in probability, and negative values (blue shading) indicate a decrease in probability, over the raw climatological values. As expected, probability increases are noted over much of the coast, although there is considerable spatial variability. The largest increases are found along the central Gulf coast, the southern Texas coast, and portions of Florida and North Carolina. Probability decreases are noted in South Carolina. A recent independent study by Saunders et al. (2000), investigating the occurrence of land falling hurricanes, corroborates these findings by showing that landfall probabilities along the central Gulf and south Texas coastlines are significantly influenced by La Niña conditions. Of interest, our map shows important geographic variation in probability changes, suggesting that the influence of large-scale climate anomalies on hurricane activity is regional.

## 5. Summary and future work

Our current understanding of hurricane climate allows for seasonal predictions of hurricane activity. This paper combines knowledge of climate teleconnections with historical records of hurricane impacts to produce annual probabilities of hurricane winds. It does this by developing a dynamic probability model. The algorithm that produces the model uses the hurricane climatology of Jarrell et al. (1992) and the assumption that the annual maximum tropical cyclone wind speed follows a Weibull distribution with a scale and shape parameter. The model can be run in a raw climatological mode in which it produces annual probabilities of hurricane wind speeds occurring somewhere in the county.

In general, raw climatological probabilities from the dynamic probability model scaled to estimate a point probability at Turkey Point in south Florida compare favorably with output from other models at hurricane intensities. At higher intensities (categories 4 and 5), the dynamic probability model values are likely too high, resulting from fitting the data to a distribution. The dynamic probability model indicates the greatest threat to hurricane activity is over southern Florida, the central Texas coast, and the eastern counties of North Carolina. Probabilities are lower along the Big Bend region of Florida and over South Carolina.

The model is unique in that it can be run in a conditional climatological mode. Regression equations for the scale and shape parameters of the Weibull distribution allow climate anomalies to influence the probabilities locally. Comparisons of the raw versus conditional probabilities indicate the influence of ENSO and the NAO on hurricane activity consistent with recent statistical findings. In particular, a mature La Niña episode combined with a weak NAO increases the likelihood of a hurricane along the central Gulf coast.

We can improve the algorithm by modeling the maximum intensity of each tropical cyclone as a Weibull distribution. If we use a Poisson distribution to model the occurrence of cyclones, then we have a two-step algorithm. For any given year let $X$ be the number of tropical cyclones of any wind speed that affect a given county. This is a Poisson distribution with $\Pr(X = x) = e^{-\lambda}\lambda^x/x!$ for a given rate $\lambda$. Let $V_i$, $i = 1, \ldots, x$ denote the maximum wind speed for each tropical cyclone. Let us assume that these random quantities, $V_i$, $i = 1, \ldots, x$, and $X$ are independent. That is, the intensity of a cyclone does not depend on the number of cyclones that affect the county in a given year nor on the intensities of other tropical cyclones that have occurred. Given this, we can model the distribution of $V_i$, $i = 1, \ldots, X$, in two stages. First the number of tropical cyclones is chosen, and second the wind speed for each of these cyclones is chosen.

We take the maximum value of the wind in a given year, $V_X = \max_{i=1,\ldots,X} V_i$, with $V_X = 0$ if there are no tropical cyclones. In this case we can determine the probability that the maximum wind speeds from a tropical cyclone exceed $\upsilon$ as follows:

$$S(\upsilon) = \Pr(V_X \geq \upsilon) = \sum_{x=1}^{\infty} \Pr(X = x)\Pr\left(\max_{i=1,\ldots,x} V_i > \upsilon\right)$$

$$= \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!}\{1 - [1 - e^{-(\upsilon/b)^a}]^x\} = 1 - e^{-\lambda e^{-(\upsilon/b)^a}}.$$

We can look at this result by assuming that we count only tropical cyclones whose wind speed exceeded $\upsilon$. Then, we have a Poisson process $X^*$ with rate $\lambda^* = \lambda e^{-(\upsilon/b)^a}$ so $S(\upsilon) = 1 - \Pr(X^* > 0)$.

In the two stage model, we can regress three parameters $\lambda$, $a$, and $b$ onto the predictors. We could also consider a four-parameter model by using a cutoff wind speed value $\upsilon_0$ and replacing $\upsilon$ with $\upsilon - \upsilon_0$ in our Weibull distribution. In this case, $\lambda$ is not the rate for the number of tropical cyclones of any velocity affecting the county, but is only the rate for cyclones whose wind speeds are $\upsilon_0$ or higher. Another improvement might be to incorporate information from adjacent counties into the model for a particular county. A Bayesian approach will work for adjusting the Weibull parameters in this case.

We note that our use of La Niña and NAO covariates in the dynamical probability model is based on previous research indicating their importance in modulating landfall activity along the U.S. coastline. If the model is used to test the influence of additional covariates on landfall probabilities, then issues of statistical significance must be addressed. Work on these areas is in progress.

*Acknowledgments.* The corresponding author thanks Ethan Gibney for his help with model validation and map preparation. The National Science Foundation (ATM-9618913) and the Risk Prediction Initiative of the Bermuda Biological Station for Research (RPI-99-001) provided funding for this research.

## REFERENCES

Batts, M. E., M. R. Cordes, L. R. Russell, J. R. Shaver, and E. Simiu, 1980: *Hurricane Wind Speeds in the United States.* NBS Building Science Series, No. 124, U.S. Department of Commerce, 42 pp.

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.

Chouinard, L. E., and C. Liu, 1997: Model for recurrence rate of hurricanes in Gulf of Mexico. *J. Waterway, Port, Coast. Ocean Eng.,* **123,** 113–119.

Chu, P.-S., and J. Wang, 1998: Modeling return periods of tropical cyclones intensities in the vicinity of Hawaii. *J. Appl. Meteor.,* **37,** 951–960.

Darling, R. W. R., 1991: Estimating probabilities of hurricane wind speeds using a large-scale empirical model. *J. Climate,* **4,** 1035–1046.

Diaz, H. E., and R. S. Pulwarty, Eds., 1997: *Hurricanes: Climate and Socioeconomic Impacts.* Springer, 292 pp.

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

——, G. S. Lehmiller, and T. B. Kimberlain, 1996: Objective classifications of Atlantic 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. 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.

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.

Hebert, P. J., and G. Taylor, 1975: Hurricane experience levels of coastal county populations from Texas to Maine. U.S. Department of Commerce, NOAA, NWS, Community Preparedness Staff and Southern Regions, 153 pp. [Available from NOAA Miami Regional Library, 4301 Rickenbacker Cswy., Miami, FL 33149.]

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.

Jarvinen, B. R., C. J. Neumann, and M. A. S. Davis, 1984: A tropical cyclone data tape for the North Atlantic Basin, 1886–1983: Contents, limitations, and uses. NOAA Tech. Memo. NWS NHC-22, 21 pp.

Johnson, M. E., and C. C. Watson Jr., cited 1999: Return period estimation of hurricane perils in the Caribbean. USAID-OAS Caribbean Disaster Mitigation Project, April, 1999. [Available online at http://www.oas.org.]

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

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.

Murnane, R. J., and Coauthors, 2000: Hurricane landfall probabilities along the east and Gulf coasts of the United States. *Eos, Trans. Amer. Geophys. Union,* **81,** 433, 438.

Neumann, C., 1987: The National Hurricane Center risk analysis program (HURISK) (reprinted with corrections 1991). NOAA Tech. Memo. NWS NHC-38, 57 pp.

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

Pielke, R. A., Jr., C. W. Landsea, R. T. Musulin, and M. Downton, 1999: Evaluation of catastrophe models using a normalized historical record. *J. Insur. Regul.,* **18,** 177–194.

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

Rupp, J. A., and M. A. Lander, 1996: A technique for estimating recurrence intervals of tropical cyclone-related high winds in the tropics: Results for Guam. *J. Appl. Meteor.,* **35,** 627–637.

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

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

Whitehead, D., 1999: An evaluation of coastal county hurricane vulnerability. M.S. thesis, Department of Meteorology, Florida State University, 64 pp. [Available from The Florida State University, Tallahassee, FL 32306.]