Improving Multiseason Forecasts of North Atlantic Hurricane Activity

JAMES B. ELSNER AND THOMAS H. JAGGER

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

MICHAEL DICKINSON AND DAIL ROWE

Accurate Environmental Forecasting, Inc., Narragansett, Rhode Island

ABSTRACT

Hurricanes cause drastic social problems as well as generate huge economic losses. A reliable forecast of the level of hurricane activity covering the next several seasons has the potential to mitigate against such losses through improvements in preparedness and insurance mechanisms. Here a statistical algorithm is developed to predict North Atlantic hurricane activity out to 5 yr. The algorithm has two components: a time series model to forecast average hurricane-season Atlantic sea surface temperature (SST), and a regression model to forecast the hurricane rate given the predicted SST value. The algorithm uses Monte Carlo sampling to generate distributions for the predicted SST and model coefficients. For a given forecast year, a predicted hurricane count is conditional on a sampled predicted value of Atlantic SST. Thus forecasts are samples of hurricane counts for each future year. Model skill is evaluated over the period 1997–2005 and compared against climatology, persistence, and other multiseasonal forecasts issued during this time period. Results indicate that the algorithm will likely improve on earlier efforts and perhaps carry enough skill to be useful in the long-term management of hurricane risk.

---

1. Introduction

Hurricanes are a serious social and economic threat to the United States. Strong winds, heavy rainfall, and storm surge kill people and destroy property. Hurricane activity over the Atlantic basin varies widely on intraseasonal to multidecadal time scales. Several factors are known to influence this activity including the El Niño–Southern Oscillation (ENSO), the North Atlantic Oscillation (NAO), and Atlantic sea surface temperature (SST).

Seasonal forecasts of hurricane activity are routinely issued by government agencies and academics. There appears to be some general skill against climatology with these forecasts, especially at lead times of a few months or less. However, the usefulness of forecasts is limited by the lack of regional specificity and by the rather short lead times (forecast horizons). Approaches are available to forecast activity near the coast (Elsner

et al. 2006; Saunders and Lea 2005; Elsner and Jagger 2004, 2006; Lehmiller et al. 1997). However, little has been done to address the limited forecast horizon. One exception is the multiseason model developed in Elsner et al. (1998). The model (with slight modifications) has been used to forecast the basin-wide hurricane count since 1997. However, as anticipated by the authors, the forecast skill has been only marginal against climatology.

The objective of this paper is to show how improvements can be made in multiseason forecasts of Atlantic hurricane activity. The idea is to make multiyear forecasts of Atlantic SST and then predictions of hurricane activity conditional on a forecast of the SST. A skillful model that spans multiple years will help with long-term emergency management planning as well as with financial planning including capital allocation and investment strategies.

We begin in section 2 by examining the track record of multiseason predictions. The skill level is put into the context of the routinely issued seasonal forecasts. In section 3 we examine time series of both Atlantic SST and North Atlantic hurricane counts. In section 4 we

---

Corresponding author address: James B. Elsner, Dept. of Geography, The Florida State University, Tallahassee, FL 32306. E-mail: jelsner@fsu.edu
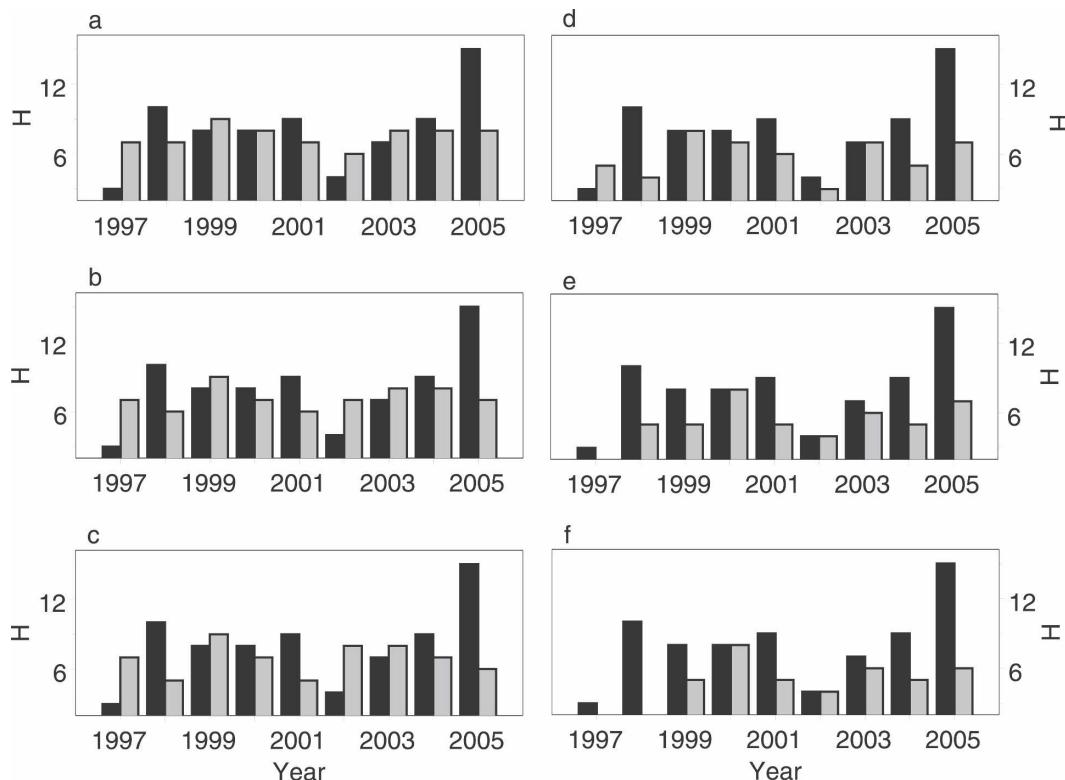
FIG. 1. Track records of the CSU seasonal hurricane forecasts and the multiseason forecasts based on Elsner et al. (1998). (a) June, (b) April, and (c) December forecasts from CSU and (d) 1-yr, (e) 2-yr, and (f) 3-yr forecasts from the multiseason model. The filled bars are the actual and the open bars are the forecast number of Atlantic basin hurricanes.

perform diagnostics on the SST and hurricane records. Time series diagnostics are run to help develop the time series model and regression diagnostics are run to help develop the regression model. In section 5 we tie the two models together into a hierarchical prediction algorithm. In section 6, the algorithm is used to retrodict hurricane activity over the period 1997–2005 and the hindcasts are compared with persistence and with actual forecasts made over this period. In section 7, we provide a summary of the method and results.

2. Past research

The present work is motivated by Elsner et al. (1998; hereafter ENT98) who use a time series model for the annual North Atlantic hurricane counts. The ENT98 model is based on a singular spectrum analysis (SSA) performed in Elsner et al. (1999). The results show important variations in North Atlantic hurricanes on sub-decadal time scales that to some degree are predictable. Here we choose a different approach that uses a time series model for the SST variation and a regression of hurricane activity onto the predicted SST values. The

main difference between the present approach and the approach of ENT98 is the parametric time series model for SST that replaces the empirical model for annual hurricane counts.

With the exception of the El Niño years of 1997 and 2002, hurricane activity over the Atlantic basin has been above normal every year since 1995. As expected, predictive skill of the ENT98 model has been marginally better than climatology. Interestingly, the overall track record of the multiseasonal ENT98 model is better than the track record of forecasts issued by Gray et al. at Colorado State University (CSU) over this time period. (Fig. 1). The correlation between predicted and actual hurricane counts using the CSU June, April, and December forecasts are +0.30, −0.12, and −0.47, respectively. The correlation between predicted and actual hurricane counts using the ENT98 one-, two-, and 3-yr forecasts are +0.40, +0.31, and +0.47, respectively.

Hurricane counts are used directly in building the ENT98 model. It is our hypothesis that improvements on this strategy can be made by forecasting rates rather than counts. Thus our perspective is more statistical in

that we assume the counts to be random conditional on the underlying rate. Another limitation to the ENT98 model is the lack of covariate information. Here we make use of the fact that, in general, hurricane activity responds to changes in SST. We thus develop a time series model of the SST, rather than a time series model of the raw counts as in ENT98, and then we use the predicted values as covariate information to forecast the hurricane rates. It is anticipated that, with both of these improvements, the skill level of multiseason projections of Atlantic hurricane activity can be increased.

The change from a period of below-normal activity to a period of above-normal activity is considered the result of low-frequency climate variations driven in part by SST over the North Atlantic (Goldenberg et al. 2001). There is also a perspective that since SSTs are influenced by climate change, the next lull in hurricane activity will not be as low as in the past (Emanuel 2005; Webster et al. 2005).

3. Atlantic SST and hurricanes

Hurricanes derive their power from the ocean through moist enthalpy and instability. Other things being equal, the amount of energy available depends on the SST and the depth of the ocean mixed layer. Cold water below a surface layer of warm SST (shallow mixed layer) will limit the energy available for hurricane maintenance. In the absence of information about the mixed layer depth and on interannual and longer time scales, we use SST as a proxy for surface forcing of hurricane activity.

Statistical analysis confirms that over the Atlantic basin warmer SST tends to result in more and stronger hurricanes. Saunders and Harris (1997) using linear regression showed that SST over the tropical Atlantic is the dominating influence behind the interannual variance in Atlantic hurricane numbers. Kimberlain and Elsner (1998) showed that the increase in hurricane activity since 1995 could be related to warmer SSTs to the east of the Lesser Antilles. Goldenberg et al. (2001) show that the number and strength of Atlantic hurricanes have a multidecadal variation which they suggest is related to SST changes resulting from North Atlantic Ocean currents.

Modeled SST and National Oceanic and Atmospheric Administration (NOAA) optimal interpolated SST datasets are used to compute Atlantic SST anomalies north of the equator in Enfield et al. (2001). Anomalies (in °C) are computed by month using the base period 1951–2000. Data are obtained from the NOAA/Cooperative Institute for Research in Environmental Sciences (CIRES) Climate Diagnostics Center

back to 1871. For this study we average the SST anomalies over the hurricane-season months of August through October.

Figure 2 displays the August–October averaged value of the Atlantic SST by year over the period 1871–2005. The data show large year-to-year fluctuations, multidecadal variability, and a positive trend. The exact cause of the low-frequency variation is open to debate, but it has been suggested that is part of a multidecadal mode [Atlantic Multidecadal Oscillation (AMO)] that influences hurricane formation through changes in local atmospheric boundary layer processes and vertical wind shear, particularly in the main development region of the Atlantic (Bell and Chelliah 2006; Enfield and Mestas-Nuñez 1999).

The distribution of the individual values is also displayed with a kernel density plot. The distribution appears symmetric and normal. Formally we check for normality using the Shapiro–Wilks statistic and find that the test statistic has a value of 0.985 corresponding to a P value of 0.719 and providing no evidence against the null hypothesis of normally distributed SST data.

A series of the annual hurricane counts over the same time period is also displayed in Fig. 2. Similar to the record of SST, the annual hurricane counts show large interannual fluctuations superimposed on low-frequency variations. The case for a trend in the counts is less compelling than for the SST, especially when considering the fact that data prior to 1946, when reconnaissance flights began, is generally accepted as somewhat incomplete. The distribution of annual counts is discrete and follows a Poisson distribution. A χ^2 goodness-of-fit test with 3 degrees of freedom gives a P value of 0.662 providing no evidence against the null hypothesis of Poisson distributed counts.

4. Data analysis and modeling

In this section we perform analysis and modeling of the data to help us decide on the overall structure of the prediction algorithm, which is outlined in the next section. We begin with time series analysis of the Atlantic SST record and continue with a regression analysis of the hurricane counts.

a. Time series model of Atlantic SST

We employ standard tools from time series analysis to help us model Atlantic SST. Our goal is a prediction model for late summer/early fall Atlantic SST values. Regardless of how well we understand the nature of the phenomenon behind the identified time series patterns, if we properly model them, we can extrapolate the patterns to predict future values.
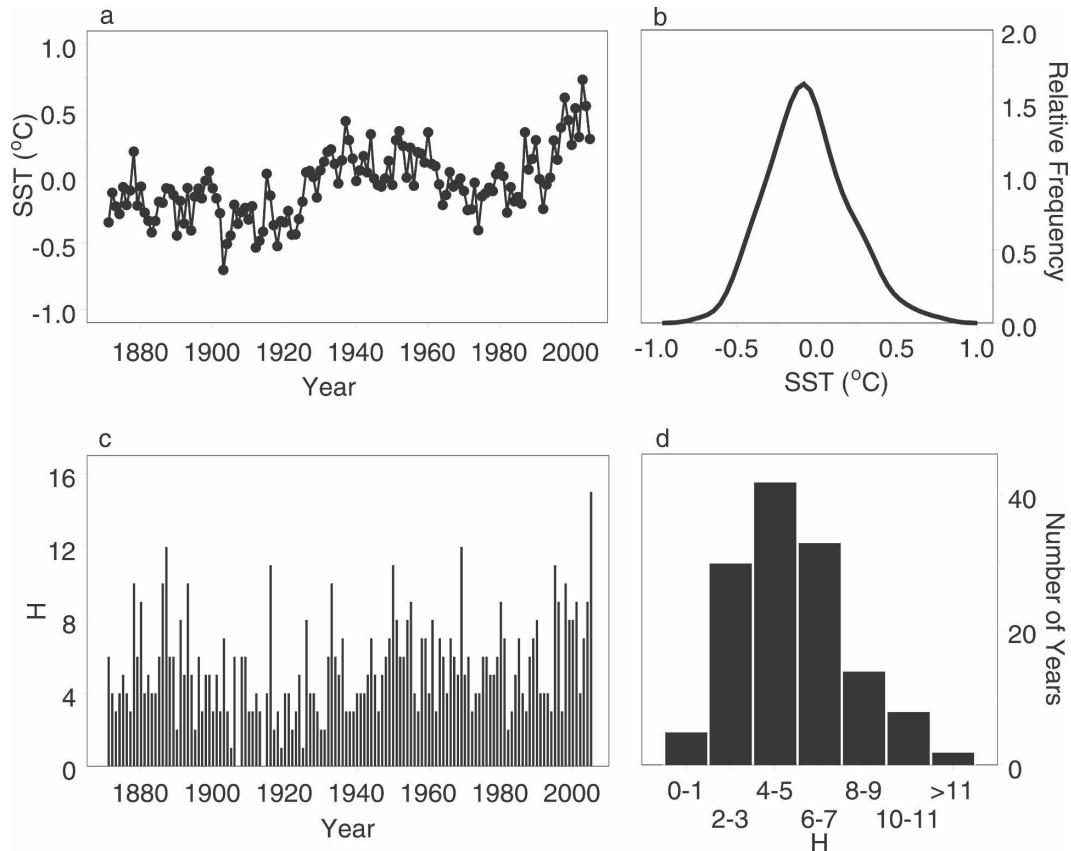
FIG. 2. (a) Time series of August–October averaged Atlantic basin SST anomalies (base: 1950–2000) over the period 1871–2005. (b) Kernel density plot of the time series values. (c) Time series of annual Atlantic hurricane counts over the period 1871–2005. (d) Distribution of annual hurricane counts.

The first step in developing a time series model is to check for stationarity. Stationarity refers to the property that the mean, variance, and autocorrelation of the time series are roughly constant. Stationarity can be checked with various statistics. The behavior of the autocorrelation function can also be used. For example, a slow decay of the autocorrelation values with increasing lag is often evidence of nonstationarity. We begin with the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test.

The KPSS test is a commonly used direct method for checking stationarity in time series (Kwiatkowski et al. 1992). It is a one-sided right-tailed test so that one rejects the null hypothesis of stationarity at the $100\alpha\%$ level if the KPSS statistic is greater than the corresponding quantile value. For the late summer/early fall SST data the KPSS statistic is 1.249 under the null hypothesis of no trend. The corresponding 99th percentile value is 0.762, so we reject the null hypothesis of stationarity about a constant value (no trend). However, if we test for stationarity under the null hypothesis of a linear trend, the KPSS statistic drops to 0.125 (the cor-

responding 99th percentile value is 0.229), so we fail to reject the null hypothesis and we proceed as if the time series is stationary after a linear trend removal.

Trend removal is typically performed by fitting a curve and subtracting the fitted values from the original data. However, Box and Jenkins (1976) recommend differencing the series to achieve stationarity and that is what is done here. The time series of the first difference ($SST_t - SST_{t-1}$) is shown in Fig. 3. Note that the long-term trend is removed with this procedure. A stationary time series has time invariant first and second statistical moments. Thus, given a sample time series y_t (for $t = 1, \ldots, T$), the lag j sample autocorrelation is defined as

$$\hat{\rho}_j = \frac{\sum_{t=j+1}^{T} (y_t - \bar{y})(y_{t-j} - \bar{y})}{\sum_{t=1}^{T} (y_t - \bar{y})^2}, \qquad (1)$$

where $\bar{y} = (1/T)\Sigma_{t=1}^{T} y_t$ is the sample mean. The sample autocorrelation function is a plot of $\hat{\rho}_j$ against lag j.
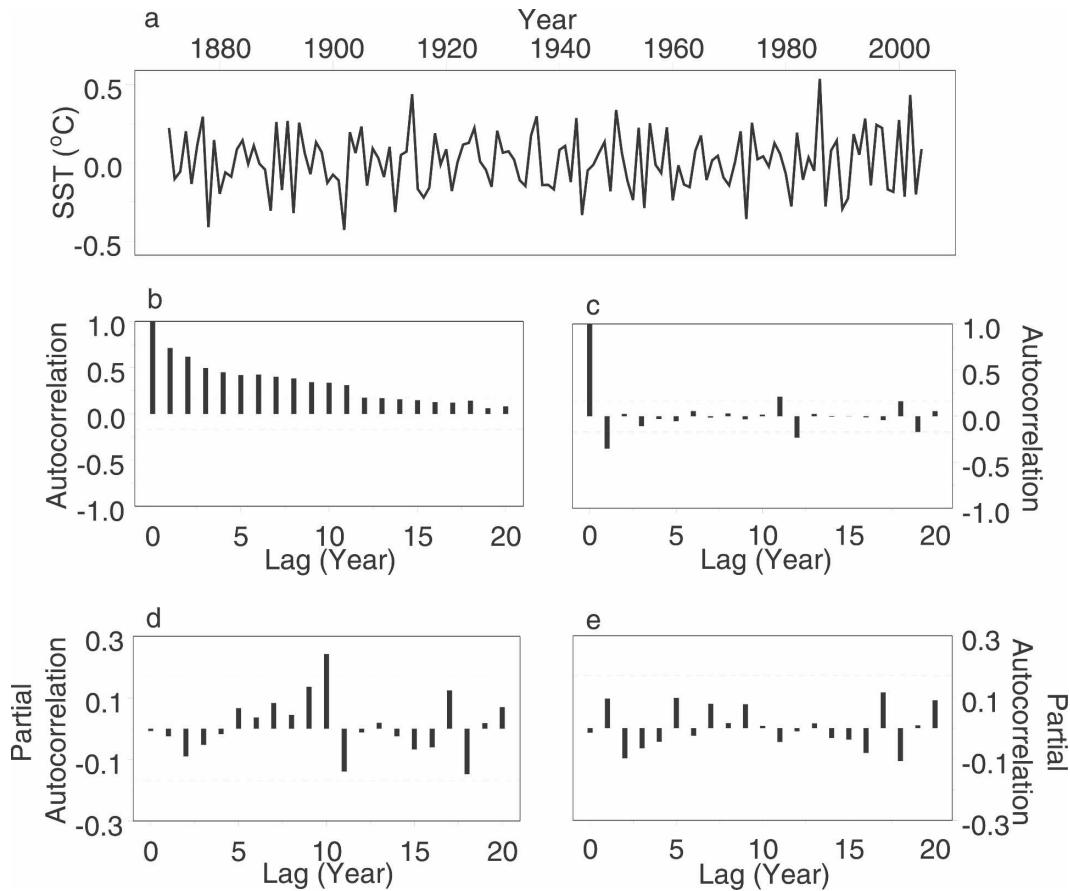
FIG. 3. (a) Time series of the first difference in August–October Atlantic SST. (b) Autocorrelation function of the raw (not differenced) time series, and (c) autocorrelation function of the differenced time series. (d) Partial autocorrelation function of the raw and (e) differenced time series.

The autocorrelation function provides information that can be used to model a time series. For comparison we plot the values of the autocorrelation function for the original series and for the differenced series in Fig. 3. The dashed lines indicate the 95% confidence levels. Large autocorrelation values (above the 95% level) extend to lag 11 for the raw time series. In contrast, the difference series shows only significant correlation at lags 1, 11, and 12. The nonstationarity in the raw series as evidenced by the slow decay in the autocorrelation function makes it difficult to choose the order of the time series model as the trend contributes to the short lag autocorrelations.

A common way to specify a univariate time series is to use an autoregressive (AR) model. Let y_t be a zero-mean time series, then

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t. \quad (2)$$

Thus an autoregressive model is a linear regression of the current value of the series against one or more prior values of the series. The value of p is called the order of

the AR model, and is denoted AR(p). The AR integrated moving average [ARIMA(p, d, q)] models extend the scope of the AR models to include moving average (MA) terms (to order q) and differencing (to order d), where the "I" in ARIMA stands for "integrated" since the predictions must be integrated (anti-differenced) to forecast the original time series. Once the orders are determined, the model coefficients are estimated using the method of maximum likelihood.

For comparisons we entertain two time series models. The first is an ARIMA(3, 0, 0) model [or equivalently, a AR(3) model on the raw time series] and the second is a ARIMA(12, 1, 1) model with coefficients for all lags except 1, 11, and 12 set to zero. The models can be expressed as

$$SST_t = \phi_1 SST_{t-1} + \phi_2 SST_{t-2} + \phi_3 SST_{t-3} + \varepsilon_t, \quad (3)$$

$$SST_t = \phi_1 SST_{t-1} + \phi_{11} SST_{t-11} + (\phi_{12} - \phi_{11}\phi_1) SST_{t-12}$$
$$- \phi_{12}\phi_{11} SST_{t-23} + \varepsilon_{t-1}, \quad (4)$$

where $\varepsilon_t = z_t + m_1 z_{t-1}$.

The ARIMA model is somewhat more complicated than the AR model. Note that the ARIMA model contains a component corresponding to an 11–12-yr cycle in the differenced series. We offer no physical explanation for this, only noting its frequency correspondence with the 11-yr solar cycle. Elsner and Kara (1999) provide a historical summary of studies related to the sun–hurricane connection and describe a few mechanisms for a possible extraterrestrial influences on Atlantic hurricane activity, including cosmic rays (Elsner and Kavlakov 2001). We note that August–October SST values are generally lower (higher) during years of low (high) sunspot numbers, which correspond to a "cooler" (warmer) sun. For example, the periods 1963–66, 1971–77, 1983–87, and 1992–97 featured reduced sunspot activity and an average hurricane-season North Atlantic SST anomaly of -0.066 which compares with an anomaly of $+0.146$ for the years not included in the above list during the period 1957–2003. Under the assumption of independence (not entirely valid), the p value for the difference in mean anomalies is 0.0013 ($N = 46$) indicating a statistical relationship.

A qualitative comparison of the two competing models is made by examining the partial autocorrelation functions of the model residuals (Fig. 3). The partial autocorrelation at lag j is the autocorrelation between SST_t and SST_{t-j} that is not accounted for by lags 1 through $j - 1$. In general, partial correlation values are larger for the AR model. Specifically, a significant partial correlation value (above the 95% line) at lag 10 is noted for the AR model residuals indicating that the model might be inadequate in describing all the important serial correlation in the time series. In contrast, no significant lag correlations are noted in the partial autocorrelation function of the residuals from the ARIMA model. Another comparison is made by examining the Akaike information criteria (AIC) given by

$$\text{AIC}(p, q) = \ln[\hat{\sigma}^2(p, q)] + \frac{2}{T}(p + q). \quad (5)$$

The information criterion captures the trade-off between model fit and parsimony. Models with too many parameters are penalized by lack of parsimony. A model with a lower AIC value is the better model accounting for fit and number of parameters. For the AR model the AIC is -86 which compares with -96 for the ARIMA model. Based on the autocorrelation analysis and the AIC values we choose the ARIMA(12, 1, 1) model for predicting the SST out to 5 yr.

More explanation is needed on the chosen model. The SST_{t-23} term arises from the product expansion of an ARIMA(12, 1, 1) and an ARIMA(11, 1, 1) model

resulting in autoregressive terms with 1, 11, 12, and 23 lags. The coefficient on the SST $_{t-12}$ term is based on examining another model where the lag 11 and lag 12 terms are added simultaneously (separate coefficients). The chosen model, where the coefficient on the lag 12 term contains the lag 1, lag 11, and lag 12 coefficients, had a smaller AIC while still removing the significant correlations and partial coefficients at lags 11 and 12. The model also contains a moving average term ε_t containing a first-order coefficient m_1 where z_t is a random uncorrelated value (innovation) at year t.

In summary, our time series model for Atlantic late summer/early fall SST contains an autoregressive component to capture the interannual persistence in SST values, a moving average component to capture the possibility that a random fluctuation may propagate to future values through feedbacks, a cyclic component possibly related to changes in solar insolation, and a trend component to capture the low-frequency variability possibly related to climate change. The model is from a family of models referred to as Box–Jenkins models after Box and Jenkins (1976).

b. Regression model of hurricanes on SST

Since hurricanes derive energy from the heat and moisture of the underlying ocean, theoretically the maximum potential intensity of a hurricane increases with increasing SST, all else being equal (Emanuel 1987; Holland 1997). Thus, as a first approximation, a season with a uniform distribution of maximum intensities over an ocean of uniform temperature will result in more hurricanes if the ocean warms even if the number of tropical cyclones remains constant. In fact, at least over the Atlantic, the year-to-year variation in hurricane activity is statistically linked (positively correlated) to Atlantic SST (Saunders and Harris 1997; Kimberlain and Elsner 1998; Goldenberg et al. 2001). More realistically the thermal structure of the ocean's mixed layer (top several hundred meters) is a better representation of the available energy for hurricanes against mixing and overturning generated by the strong winds. Thus, SST serves as a proxy for oceanic heat content. Figure 4 shows time series' plots of the August–October Atlantic SST and annual Atlantic hurricane counts.

The interannual variation in the pattern of hurricane activity appears to coincide with the interannual variation in the pattern of SST. This association is confirmed quantitatively in the scatterplot. The regression line indicates that SST explain a statistically significant (P value <0.001) 17% of the interannual variability in hurricane counts. Considering the complexity of hurri-
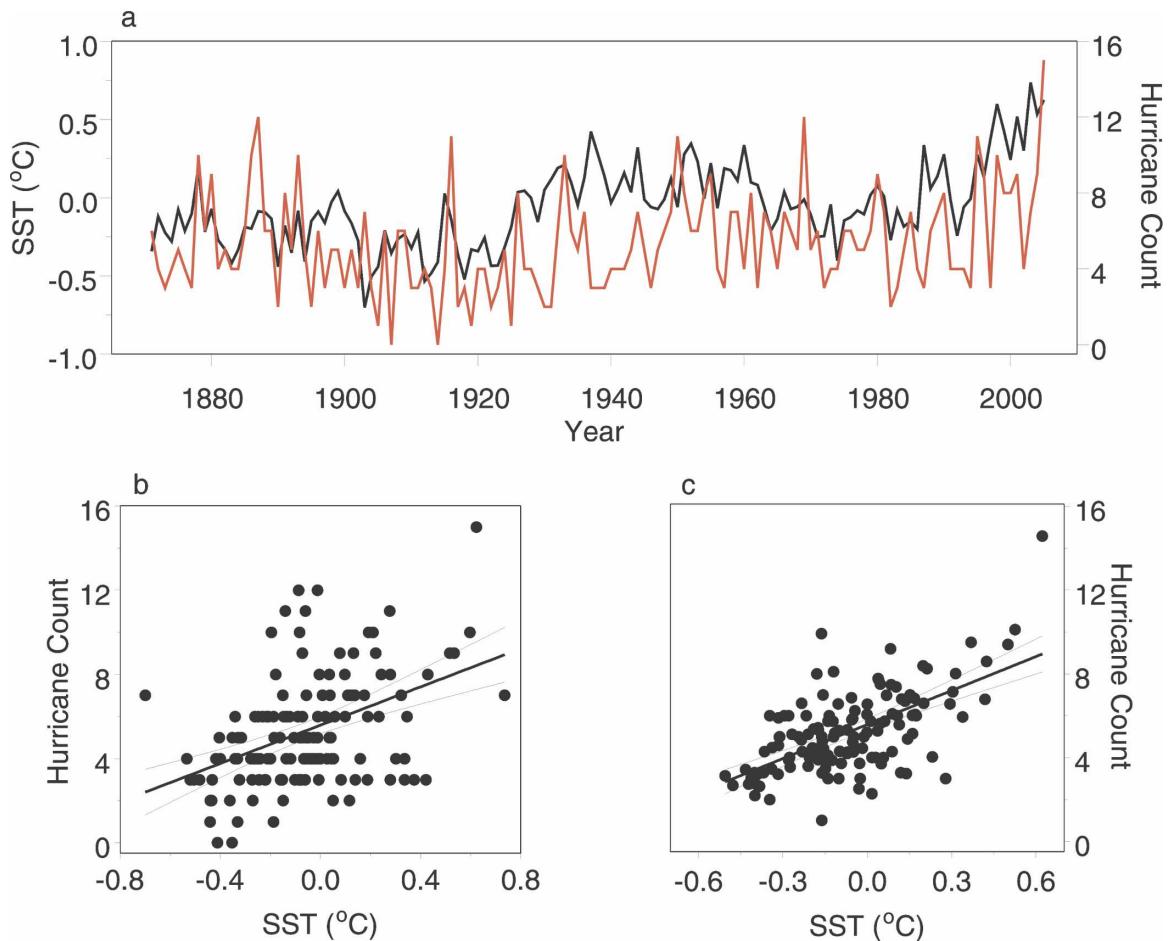
FIG. 4. (a) Time series of the August–October Atlantic SST (black) and annual hurricane count (red). (b) Scatterplot of the SST and hurricane counts, and (c) scatterplot of the 5-yr smoothed SST and hurricane counts. The straight lines in the scatterplots are the least squares regression line of hurricanes on SST. The regression slope is significantly different from zero for both raw and smoothed data as can be seen by the 95% confidence lines on the regression line.

cane genesis and the fact that SST is only a proxy for oceanic heat content, the result is encouraging. Indeed it is well understood that, regardless of surface warmth, atmospheric winds creating a shearing environment can inhibit hurricane activity. This is why seasonal forecast models can be improved by including a variable characterizing ENSO, which tends to modulate Atlantic hurricane activity through its effect on environmental wind shear and subsidence.

On longer time scales, the effect of ENSO can be considered a "random" component to hurricane activity, whereas the autocorrelative structure of the SST data as described in the previous section suggests that the relationship with hurricane activity likely will be stronger. The scatterplot of the smoothed hurricane counts against the smoothed SST verifies this. Here we find that SST explains 32% of the variability in hurri-

cane activity. The series' are smoothed using a local regression with a span of 5 yr. These diagnostics provide statistical validation for our approach to modeling hurricane counts from SST. However, since interest is in future hurricane counts we use a regression model appropriate for discrete data.

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 basin-wide hurricane counts $h_t \in 0, 1, 2, \ldots, \infty = Z^+$ on the nonnegative integers for a set of observed years $t = 1, \ldots, T$. Ad-

ditionally we have a set of SST values one for each year. Thus the Poisson regression is

$$h_t \sim \text{Poisson}(\lambda_t)$$

$$\lambda_t = \exp(\beta_0 + \text{SST}_t\beta_1), \qquad (6)$$

where λ_t is the hurricane rate for year t, β_0 is the intercept, and β_1 is SST coefficient. The symbol $\sim$ refers to a stochastic relationship and indicates that the variable on the left-hand side is a random draw (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.

5. Prediction algorithm

Diagnostic results from the previous section provide us with the background to build a multiyear prediction algorithm for annual hurricane counts. In short, the algorithm generates a predictive distribution of annual hurricane counts out to 5 yr by generating predictive samples of SST values and predictive samples of regression model coefficients and combining them mathematically through Eq. (6). Specifically the algorithm uses Monte Carlo (MC) sampling within a hierarchy. A schematic of the hierarchy is shown in Fig. 5. At the top are the SST and hurricane data records. The SST data are used alone to build a univariate time series model and are used together with the hurricane data to build a regression model. Details of the time series and regression models were given in the previous section.

Given values for the coefficients of the ARIMA model and random uncorrelated innovations for each year, the time series model together with the historical series of SST values produces a single sample of predicted SST values with a single value for each year out to 5 yr. Using 100 different innovation series provides a sample consisting of 100 5-yr SST predictions. Given the historical series of SST values and corresponding hurricane counts, the regression model produces a single set of model coefficients (offset and slope), representing expected values from a Student's t distribution. The regression model specifies a Poisson rate for a single year conditional on a value of SST. Choosing a random set of model coefficients from their respective distributions together with a single SST value produces a single rate for a single year. The procedure is repeated for each forecast year and for 100 sampled SST predictions and 1000 sampled regression coefficients for a total of 500 000 hurricane counts. To get a single predicted rate we average over all samples for a given forecast year. This MC approach combines the uncer-
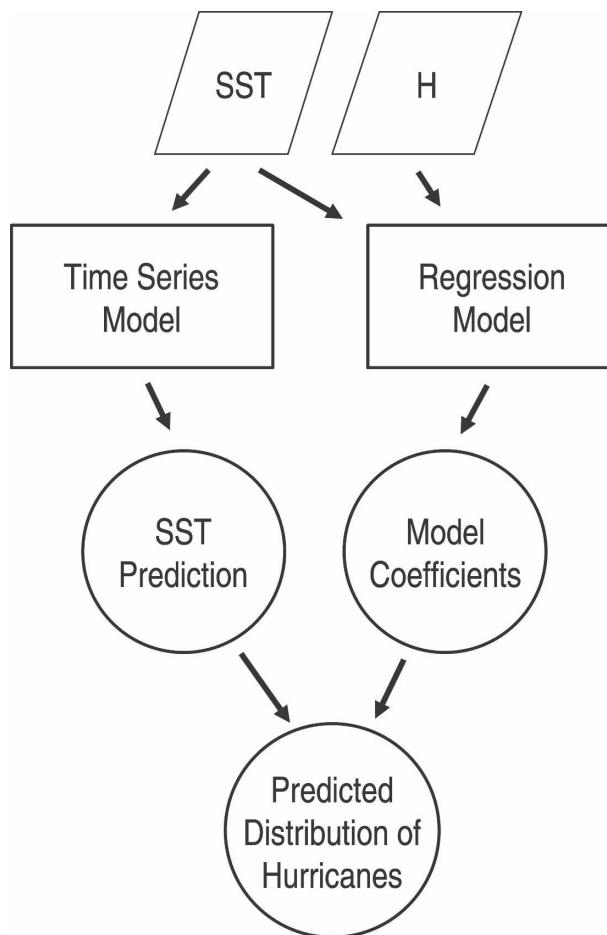

FIG. 5. Graph illustrating the algorithm for generating a predictive distribution of annual Atlantic basin hurricane counts out to 5 yr. The data are the time series of Atlantic SST and hurricane counts (H). Samples of the SST predictions are generated from the time series model and samples of the regression coefficients are generated from the regression model.

tainty inherent in the time series predictions with the uncertainty associated with the regression of hurricanes on SST to give an assessment of overall predictive uncertainty.

6. Validation

Model skill is examined by estimating the time series and regression parameters on data over the period 1871–1996 (training period), then applying the prediction algorithm to forecast hurricane counts from 1997–2005 (forecast period). The length of the training period (initially 116 yr) is dictated by a trade-off between having enough data to produce an accurate model and leaving enough data to get reliable validation statistics. The choice of years for the training period is made
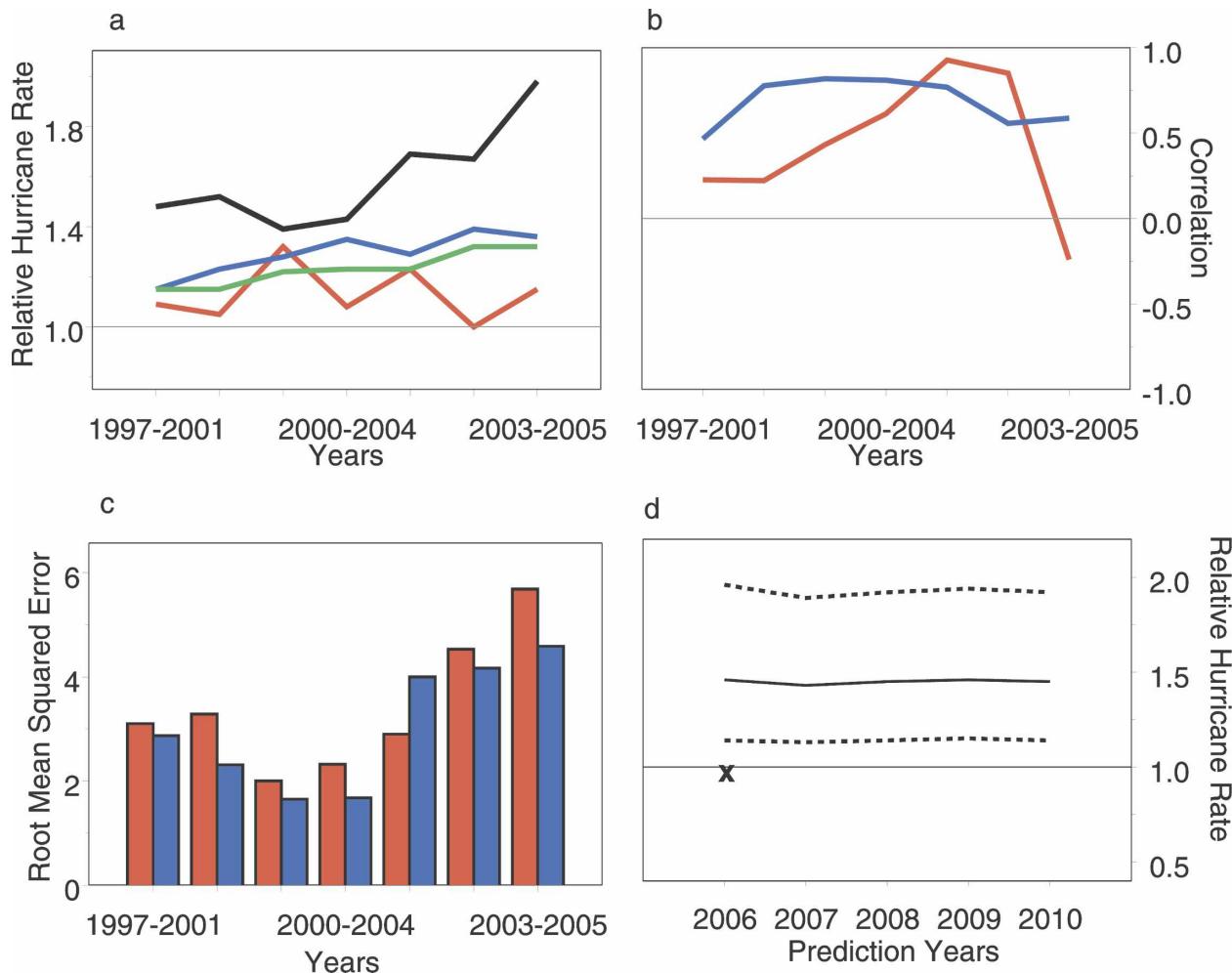
FIG. 6. Forecast validation statistics and a forecast for the 5 yr beginning with 2006. (a) 5-yr observed and predicted hurricane rates relative to a climatological base period. The black line indicates the observed rates, the blue line the present SST model, the green line a persistence forecast, and the red line the ENT98 model. Note that all three models predict above the long-term climatological mean represented by a relative rate of 1 with the SST model providing the best forecasts, although all fall short of the observed rates. (b) Correlation between observed and predicted counts for each of the 7 forecast periods using the SST (blue line) and the ENT98 (red line) model. (c) Same as (b) except RMSE. (d) Forecast relative hurricane rate for the period 2006–10 using the SST model. The mean relative rate is the solid line and the 95% confidence lines are the dashed lines. The X marks the observed relative hurricane rate for 2006.

partly based on the fact that comparisons can be made with actual forecasts from the ENT98 model. The first forecast in the validation period is issued for 5 consecutive seasons 1997–2001 and is initialized in December of 1996. The training period is then updated to include the observed number of hurricanes for 1997 and the model parameters reestimated to make a forecast for the period 1998–2002. This is repeated 7 times so that the last forecast period is 2003–05.

Figure 6 shows various results from the validation exercise. First, we compare the observed relative hurricane rate with forecast relative rates. Relative rates are computed by taking the ratio of the short-term (5 yr) mean hurricane count to the long-term mean hurricane count. A relative rate of 1 indicates a long-term climatological mean number of hurricanes. The long-term mean is the average number of hurricanes starting with 1871 and ending with the year before the first forecast year. For example, in comparing the 5-yr forecast issued beginning with the 1997 season, we use a long-term mean of 5.14 hurricanes based on the period 1871–1996. The observed short-term mean is 7.60 hurricanes so the observed relative rate is 1.48. Forecasts made with the present model are converted to relative rates by taking the ratio of the 5-yr forecast mean to the observed long-term mean. The same procedure is used

to convert the forecasts from the ENT98 model and a persistence model that uses the previous 10 yr as a forecast for the next 5 yr. Note that with forecasts made starting with the 2002 and 2003 seasons, the number of forecast years is 4 and 3, respectively.

The observed short-term hurricane rates have exceeded the long-term mean rate in each of the 7 forecast periods. The three forecast models do a good job in predicting the above-normal activity with the SST model performing the best, although all three forecast models fall short in predicting the magnitude of the recent high level of hurricane activity. The correlation between the observed and predicted number of hurricanes over each of the 5-yr prediction intervals for the SST model and the ENT98 model is also shown. The correlation are generally above 0.5 for the SST model and better than the ENT98 model, although the ENT98 model proved superior with its 2001 and 2002 forecasts. Note that the persistence model provides a 5-yr mean rate forecast rather than forecasts of annual counts, so correlation as a measure of skill is meaningless. The root-mean-squared error statistics show similar results, with the SST model providing more skill than the ENT98 model with the exception of the 2001 forecast when the ENT98 model predicted 11 hurricanes (observed $= 15$) for the 2005 season 5 yr in advance.

The amount by which the model underestimates the observed activity over the verification period might be partly due to a low bias in hurricane counts in the years prior to about 1944. Recent work in trying to quantify this bias (Landsea 2007) could lead to a more accurate forecast model. This can be done by using the bias-corrected set of hurricane counts to recalibrate the Poisson regression.

In general the validation results indicate skill for the SST model. Thus we use the model to predict the relative rates out to the year 2010. The rates are integrated over the forecast period. The forecast calls for a continuation of an active hurricane period with expected relative rates at levels about 50% above the long-term mean. The 95% confidence interval ranges between 10% and 100% of the long-term rate.

The 2006 Atlantic hurricane season, which featured 5 hurricanes, came and went during the course of the research and review process. This is slightly below the long-term (1871–2005) mean number of 5.34 and well below the forecast mean relative rate of 1.46. Model failure is attributed to El Niño. The El Niño produced greater wind shear and subsidence across the tropical Atlantic that is not accounted for in the model. However, to put this forecast into context, a mean relative rate of 1.46 translates into 7.8 hurricanes. For comparison, Klotzbach and Gray (2006) forecast 9 hurricanes

for 2006 as late as 1 June 2006, while Saunders and Lea (2006) forecast 8.2 hurricanes for 2006 as late as 4 April 2006.

7. Summary

Hurricanes are capable of generating large financial losses to the insurance industry. Predictions of the level of hurricane activity have been around since the mid-1980s, and currently there are at least a half dozen groups regularly issuing such forecasts, but with the exception of ENT98 outlooks covering a longer time horizon of several years have not been routinely attempted.

Here we develop a statistical algorithm to predict basin-wide North Atlantic hurricane counts out to 5 yr that improves on the strategy of ENT98. The algorithm has two components: an ARIMA time series model to forecast average hurricane-season Atlantic SST, and a regression model to forecast the annual hurricane counts conditional on the predicted SST. The algorithm is developed based on results from diagnostic tests for goodness of fit, stationarity, component order, and significance of the regression coefficients. The algorithm uses Monte Carlo sampling to generate predictive samples of SST and samples of the regression coefficients. In this way forecasts are samples of hurricane counts that combine uncertainty in the predictive SST values with uncertainty in the regression model of hurricanes on SST.

Forecast skill is evaluated over the period 1997–2005 and compared with skill from a persistence forecast and the skill from the original multiseasonal forecast model. Results indicate the algorithm has sufficient skill to be useful in the long-term guidance of hurricane risk. The model contains no term to account for higher-frequency fluctuations associated with, for example, the inhibiting or enhancing conditions of ENSO, so forecasts will fail in years when such conditions dominate. A forecast to the year 2010 indicates a continuation of the recent above-normal activity. The connection between regional North Atlantic SSTs and hurricanes could lead to the development of space–time multiseason forecast models.

REFERENCES

Bell, G. D., and M. Chelliah, 2006: Leading tropical modes associated with interannual and multidecadal fluctuations in North Atlantic hurricane activity. *J. Climate,* **19,** 590–612.

Box, G. E. P., and G. M. Jenkins, 1976: *Time Series Analysis: Forecasting and Control.* Rev. ed. Holden-Day, 575 pp.

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

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

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

——, and S. P. Kavlakov, 2001: Hurricane intensity changes associated with geomagnetic variation. *Atmos. Sci. Lett.,* **2,** 86–93.

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

——, and ——, 2006: Prediction models for annual U.S. hurricane counts. *J. Climate,* **19,** 2935–2952.

——, X. Niu, and A. A. Tsonis, 1998: Multi-year prediction model of North Atlantic hurricane activity. *Meteor. Atmos. Phys.,* **68,** 43–51.

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

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

——, R. J. Murnane, and T. H. Jagger, 2006: Forecasting U.S. hurricanes 6 months in advance. *Geophys. Res. Lett.,* **33,** L10704, doi:10.1029/2006GL025693.

Emanuel, K. A., 1987: The dependence of hurricane intensity on climate. *Nature,* **326,** 483–485.

——, 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature,* **436,** 686–688.

Enfield, D. B., and A. M. Mestas-Nuñez, 1999: Multiscale variabilities in global sea surface temperatures and their relationships with tropospheric climate patterns. *J. Climate,* **12,** 2719–2733.

——, ——, 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.

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,* **293,** 474–479.

Holland, G. J., 1997: The maximum potential intensity of tropical cyclones. *J. Atmos. Sci.,* **54,** 2519–2541.

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

Kimberlain, T. B., and J. B. Elsner, 1998: The 1995 and 1996 North Atlantic hurricane seasons: A return of the tropical-only hurricane. *J. Climate,* **11,** 2062–2069.

Klotzbach, P. J., and W. M. Gray, cited 2006: Extended range forecast of Atlantic seasonal hurricane activity and U.S. landfall strike probability for 2006. [Available online at http://hurricane.atmos.colostate.edu/Forecasts/2006/june2006.]

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin, 1992: Testing the null hypothesis of stationarity against the alternative of a unit root. *J. Econ.,* **54,** 159–179.

Landsea, C. W., 2007: Counting Atlantic tropical cyclones back to 1900. *Eos, Trans. Amer. Geophys. Union,* **88,** 197.

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.

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.

Saunders, M. A., and A. R. Harris, 1997: Statistical evidence links exceptional 1995 Atlantic hurricane season to record sea warming. *Geophys. Res. Lett.,* **24,** 1255–1258.

——, and A. S. Lea, 2005: Seasonal prediction of hurricane activity reaching the coast of the United States. *Nature,* **434,** 1005–1008.

——, and ——, cited 2006: April forecast update for Atlantic hurricane activity in 2006. [Available online at http://tsr.mssl.ucl.ac.uk.]

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

Webster, P. J., G. J. Holland, J. A. Curry, and H.-R. Chang, 2005: Changes in tropical cyclone number, duration, and intensity in a warming environment. *Science,* **309,** 1844–1846.