JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 112, D17114, doi:10.1029/2007JD008385, 2007

The influence of climate state variables on Atlantic Tropical Cyclone occurrence rates

Thomas A. Sabbatelli[1] and Michael E. Mann[1]

Received 1 January 2007; revised 20 April 2007; accepted 12 June 2007; published 15 September 2007.

[1] We analyzed annual North Atlantic tropical cyclone (TC) counts from 1871-2004, considering three climate state variables—the El Niño/Southern Oscillation (ENSO), peak (August-October or 'ASO') Sea Surface Temperatures (SST) over the main development region ('MDR': 6-18°N, 20-60°W), and the North Atlantic Oscillation (NAO)—thought to influence variations in annual TC counts on interannual and longer timescales. The unconditional distribution of TC counts is observed to be inconsistent with the null hypothesis of a fixed rate random (Poisson) process. However, using two different methods, we find that conditioning TC counts on just two climate state variables, ENSO and MDR SST, can account for much or all of the apparent non-random variations over time in TC counts. Based on statistical models of annual Atlantic TC counts developed in this study and current forecasts of climate state variables, we predicted $m = 15 \pm 4$ total named storms for the 2007 season.

Citation: Sabbatelli, T. A., and M. E. Mann (2007), The influence of climate state variables on Atlantic Tropical Cyclone occurrence rates, *J. Geophys. Res.*, *112*, D17114, doi:10.1029/2007JD008385.

1. Introduction

[2] A number of past studies have examined climatic influences on variations at interannual and longer timescales in the occurrence and the intensity of North Atlantic Tropical Cyclones (TCs) [e.g., *Gray*, 1984]. The primary factor considered in past studies is the El Niño/Southern Oscillation (ENSO) [e.g., *Bove et al.*, 1998; *Landsea et al.*, 1999; *Elsner et al.*, 2000; *Elsner*, 2003; *Elsner et al.*, 2006; *Elsner and Jagger*, 2006], though the influence of the North Atlantic Oscillation ('NAO') has also been examined in some studies [*Elsner et al.*, 2000; *Elsner*, 2003; *Elsner et al.*, 2006; *Elsner and Jagger*, 2006]. Both phenomena are believed to influence TC production, development, or prevailing trajectories through their influence on storm tracks or vertical wind shear in the tropical North Atlantic. The ENSO phenomenon tends to enhance (diminish) TC counts during storm seasons coinciding with an incipient La Nina (El Niño) event, while the NAO tends to enhance (diminish) TC counts during storm seasons coinciding with an incipient negative (positive) phase winter. Influences are historically found only during the storm season *preceding* the anomaly in the index; there is no detectable impact on the following year's storm season.

[3] Sea Surface Temperatures (SST) over the main development region ('MDR': 6-18N, 20-60W) for North Atlantic TCs during the season (August-October or 'ASO') of Peak TC production [*Emanuel*, 2005a; *Webster et al.*, 2005, 2006; *Mann and Emanuel*, 2006; *Sriver and Huber*, 2006; *Elsner*, 2006] have also been argued to be an important influence on long-term North Atlantic TC behavior. MDR SSTs are considered a proxy for potential TC intensity [*Emanuel*, 2005a], with annual TC counts enhanced (diminished) in seasons associated with positive (negative) MDR SST anomalies. Related studies have argued for a significant influence of the so-called 'Atlantic Multidecadal Oscillation' ('AMO') on North Atlantic TC numbers [e.g., *Goldenberg et al.*, 2001]. However, as the procedures used to define the 'AMO' signal in terms of North Atlantic SSTs in such studies has been challenged in recent work [*Trenberth and Shea*, 2006; *Mann and Emanuel*, 2006], we have chosen in our analyses here to employ MDR ASO SSTs themselves [as in e.g., *Emanuel*, 2005a; *Mann and Emanuel*, 2006; *Elsner*, 2006], rather than an index such as the 'AMO' derived through statistical processing of the North Atlantic SST field.

[4] Previous studies have investigated long-term trends in TC statistics [e.g., *Solow and Moore*, 2000] or have used regression models employing climatic indices [*Gray*, 1984; *Elsner et al.*, 2000, 2006; *Elsner and Jagger*, 2006] and trend parameters [*Elsner*, 2003] to predict interannual variations in TC activity. In no previous studies we are aware of, however, have investigators examined whether conditioning on climatic factors can account for the entirety of non-random structure in the statistical distribution of historical North Atlantic annual TC counts. In this study we perform such an examination, employing two distinct and complementary methods to test the hypothesis that annual TC counts follow a state-dependent Poisson process against the null hypothesis of a constant rate Poisson random process.

[5] Any statistical approach to analyzing TC counts must respect the Poisson distributional nature of the underlying

[1]Department of Meteorology and Earth and Environmental Systems Institute, Pennsylvania State University, University Park, Pennsylvania, USA.

0148-0227/07/2007JD008385$09.00

process (that is, that TC counts are characterized by a point process with a low occurrence rate). Our first approach employs Poisson regression [see e.g., *Elsner et al.*, 2000, 2001; *Elsner*, 2003; *Elsner and Jagger*, 2006], a variant on linear regression which is appropriate for modeling a conditional Poisson process in which the expected occurrence rate co-varies with some set of state variables (e.g., indices of ENSO, the NAO, and MDR SST). The second approach categorizes the data with respect to the climate state variables using a binary classification scheme, testing both for the statistical significance of differences in occurrence rates between the resulting data subgroups, and examining the resulting subgroup distributions for consistency with a Poisson random process. The two methods are complementary in that the latter method avoids the restrictive linearity assumptions implicit in regression, while the former method accounts for continuous variations in expected TC occurrence rates as a function of the underlying state variables (e.g., distinguishing between the impacts of strong vs. weak El Nino events).

2. Data

[6] Our analysis employed four data sets including (1) historical annual North Atlantic TC counts, (2) the December-February (DJF) Niño3.4 SST ENSO index, (3) the December-March (DJFM) NAO index, and (4) Aug-Oct (ASO) seasonal SST means over the main development region ('MDR') of 6°–18°N, 20°–60°W. Our analysis was confined to the 135 year interval 1870-2004 over which all three primary data sets of interest were available. The more recent seasons of 2005 and 2006 for which preliminary data are available, are subsequently interpreted in the context of these analyses, while forecasts for the 2007 season are made based on projected values of the climate indices. Data are available at the supplementary website: http://www.meteo.psu.edu/~mann/TC_JGR07.

[7] Historical estimates of the annual TC counts are available back to 1850 [*Jarvinen et al.*, 1984]. The reliability of these data, particularly prior to the late 20th century in which satellite and aircraft reconnaissance are available, has been vigorously debated in recent studies [e.g., *Landsea*, 2005; *Emanuel*, 2005b]. *Emanuel* [2005b] nonetheless makes a credible argument for why long-term TC count data should be reliable, even if TC intensity estimates are not. As *Emanuel* [2005b] notes, prior to aircraft reconnaissance, ships crossing the Atlantic would not have been warned off from a developing or approaching storm, and were likely to encounter either the storm or evidence of its existence. Combined with other impacts on islands or coastal localities, the existence of an Atlantic tropical cyclone was therefore likely to have been known, even prior to aircraft reconnaissance.

[8] Various alternative indices of the El Nino/Southern Oscillation (ENSO) are available. We employed the boreal winter (DJF) Niño3.4 index (SST averaged over the region 5°S-5°N, 120°-170°W) favored by many investigators [e.g., *Trenberth*, 1997]. Use of alternative (e.g., Niño3) ENSO indices yielded similar conclusions. The Niño3.4 index was taken from the *Kaplan et al.* [1998] data set and updated with subsequent values available through NCEP. The boreal winter (DJFM) NAO index was taken from *Jones et al.* [1997], updated with more recent values from the University of East Anglia/CRU. For simplicity, the 'year' was defined to apply to the preceding storm season for both indices (e.g., the 1997/1998 El Nino and winter 1997/1998 NAO value were assigned the year 1997).

[9] The MDR SST index was taken from the HadISST2 observational SST data set [*Rayner et al.*, 2003] and updated with more recent values from the UK Met Office. The data were averaged over the season most relevant to tropical cyclone formation (August-September-October, or 'ASO'). Estimated uncertainties in the observational SST data are relatively small back to 1870 for both the Nino3.4 and North Atlantic regions of interest in this study [see, e.g., *Kaplan et al.*, 1998].

3. Methods

[10] As in previous studies [e.g., *Elsner et al.*, 2000], we assumed that annual TC counts n can be modeled as a (Poisson) point process, viz.

$$P_i(n) = (1/n!)\mu^n \exp(-\mu) \qquad (1)$$

where the mean occurrence rate μ, is the sole free parameter of the distribution, and in the unconditional case has a Maximum Likelihood value equal to the mean annual count. While the appropriate null hypothesis holds the rate parameter μ to be constant over time, it is of interest to investigate the alternative hypothesis that μ may vary with respect to some set of governing factors or 'state variables' [e.g., time—*Solow and Moore*, 2000; *Elsner*, 2003 and/or climate state indices—e.g., *Elsner*, 2003; *Elsner and Jagger*, 2006].

[11] For the purposes of our study, μ was conditioned on the three climate state variables discussed above (ENSO as measured by the DJF Niño3.4 index, NAO as measured by, the DJFM NAO index, and MDR SST as measured by the MDR ASO SST index). Two distinct statistical approaches were taken, as described below. We note that here is room for further development of the methods presented below. For example, one could extend the approaches used in the present study to account explicitly for the increased uncertainty in TC counts back in time, and in particular the impact of unreported events [e.g., as in *Solow and Moore*, 2000; *Elsner and Jagger*, 2006].

3.1. Binary Classification Approach

[12] In this approach, each year is classified as belonging to one of two possible binary states (positive or negative) with respect to each state variable, depending on the sign of the anomaly in that variable (relative to the 1870-2004 mean). An alternative tertiary classification procedure was tested in which a third neutral category was introduced (defined by absolute anomalies within one standard deviation). The choice of binary vs. tertiary classification schemes represents a tradeoff between the level of discrimination (two vs. three states) and resulting sample sizes. While similar results were obtained using the tertiary categorizations scheme, we preferred the binary classification scheme due to the larger sizes of the data sub-samples. For similar reasons, only the two most significant (see section 4 for further discussion) of the three state variables, MDR SST and Niño3.4 were used.

[13] Using the binary classification scheme, we categorized years with respect to each of the two factors separately, and further, into three distinct sub-groupings, defined as (1) 'favorable': years in which both factors are favorable to TC production (positive MDR SST and negative Niño3.4 anomalies), (2) 'unfavorable': years in which both factors are unfavorable to TC production (negative MDR SST and positive Niño3.4 anomalies), and (3) 'neutral': years in which the two factors tend to offset in terms of their favorability to TC formation, i.e., anomalies in MDR SST and Niño3.4 that are of the same sign.

[14] We used a χ^2 test to evaluate the goodness-of-fit of a Poisson distribution for both the unconditional (i.e., all 135 years grouped together) and conditional (i.e., 'favorable, 'neutral', and 'unfavorable') data categorizations. We assumed χ^2 to have $\nu = B - 2$ degrees of freedom, where B is the number of occupied bins, and 2 degrees of freedom are subtracted based on constraints provided from the data (normalization of the distribution, and estimation of the rate parameter μ). The bin bandwidth was chosen using the objective criterion cited by *Wilks* [2005],

$$h \approx cIQR/N^{1/3} \tag{2}$$

where N is the sample size, IQR is the inter-fourth quartile range of the data, and c = 2 is taken for relatively skew distributions such as the Poisson. h was rounded to the nearest integer value.

[15] The t statistic was then used to evaluate the statistical significance of the differences in TC rate parameter estimates μ_i between any two data sub-samples. The t statistic reduces to

$$t = (\mu_1 - \mu_2)/(\mu_1/\phi_1 + \mu_2/\phi_2)^{1/2} \tag{3}$$

using the expression for the sample variance of a Poisson distribution, $\sigma^2 = \mu$, where ϕ_1 and ϕ_2 denote the degrees of freedom in the respective sub-samples, and the degrees of freedom in the t statistic is min(ϕ_1, ϕ_2) -1. When only Niño3.4—which is serially uncorrelated—is used as a conditioning variable, ϕ_1 and ϕ_2 reduce to simply N_1 and N_2, the nominal sizes of the respective sub-samples. However, significant serial correlation in the MDR SST series (the lag one autocorrelation coefficient $\rho = 0.55$ yields a decorrelation timescale τ = 1.67 years) decreases the effective number of independent climate states sampled when conditioning on MDR SST as, e.g., two neighboring years are not statistically independent with respect to the enhanced likelihood of elevated TC counts. Reduced degrees of freedom (ϕ) were therefore taken into account in estimating the statistical significance of t scores when conditioning fully or partly on the MDR SST series. In such cases, only events spaced more than two decorrelation timescales (i.e., 3 years) apart were considered to constitute statistically independent samples.

[16] Finally, we used a cross-validation procedure to evaluate the predictive skill in the binary conditional Poisson model approach. One could [see, e.g., *Elsner and Jagger*, 2006] leave each year out one at a time, forming conditional TC rate parameter estimates based on the remaining years and evaluating the skill of the resulting classifications applied to each choice of missing year. However, when serial correlation is present in the state variables, which as discussed above is the case here, the results of such a cross-validation procedure are likely to give too liberal an estimate of skill. We therefore employed an alternative split calibration/validation procedure. Conditional TC rate parameter estimates were obtained using the first half (i.e., years 1870-1937) of the data, and subsequently used to categorize the subsequent TC count data based on the climate state variable anomalies (measured relative to the calibration period baseline) over the latter half (i.e., years 1943-2004). This procedure was then repeated with the role of the first and last half of the data sets reversed. The average of the mean squared error (MSE) between the predicted and observed TC count data obtained for both sub-intervals was used as an estimate of cross-validated MSE, which was compared to the MSE obtained over the full (1870-2004) model development interval.

3.2. Poisson Regression

[17] Poisson regression is a variant on linear regression appropriate for data such as TC counts for which the null hypothesis of a Poisson distribution is appropriate [see *Elsner et al.*, 2000, 2001; *Elsner*, 2003; *Elsner and Jagger*, 2006 for further discussion]. Given a count series Y with unconditional mean rate μ believed to follow a state-dependent Poisson distribution, Poisson regression estimates a generalized linear model for the conditional expected rate of occurrence λ = E (Y) as a function of a set of state variables $X_1, X_2, \ldots, X_M$, of the form,

$$\log \lambda = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots. + \beta_M X_M \tag{4}$$

or alternatively,

$$\lambda = \exp[\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots. + \beta_M X_M] \tag{5}$$

where the residuals are assumed to be Poisson distributed.

[18] Unlike ordinary linear regression, a closed-form analytical solution to equation (5) is not possible. However, it is straightforward to numerically estimate maximum likelihood values for the regression parameters β_i, and thus obtain estimates for the conditional expected occurrence rates λ_i. The residual series $\varepsilon_i = Y_i - \lambda_i + \mu$ can be analyzed for consistency with a Poisson distribution based on a χ^2 test, as described in section 3.1 above.

[19] Poisson regression was performed for various combinations of climate state variables as discussed in more detail in section 4. Cross-validation was performed using the split calibration/validation procedure discussed in section 3.1 wherein the regressions were performed alternatively using the first and last half of the full data set, with TC counts predicted and compared with observed counts over the remaining independent half of the data set. Quality of regression fit was measured by both the coefficient of determination R^2 and mean square error (MSE).

4. Results

[20] Certain relationships between annual TC counts and the Niño3.4 and MDR SST time series are evident by

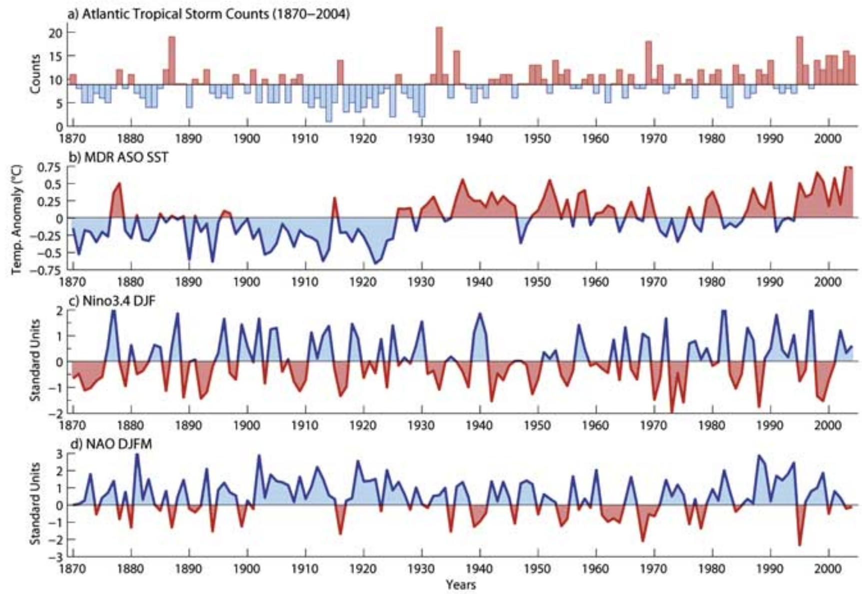

Figure 1. Time Series (1870-2004) of (a) annual Atlantic TC counts, (b) MDR ASO SST time series, (c) Niño3.4 DJF SST index, and (d) NAO DJFM SLP index. Red (blue) indicates positive (negative) anomalies in TC counts and Hurricane-favorable (unfavorable) conditions in the three indices (MDR SST, Niño3.4 and NAO). Note that year convention applies to the 'D' in DJF and DJFM for both 'c' and 'd'.

inspection alone (Figure 1). The clear increase in TC counts subsequent to the 1920s, and the positive trend over roughly the past decade, closely coincide with corresponding tendencies for positive MDR SST anomalies. Anomalously low TC counts in certain years (e.g., 1982 and 1997) correspond to prominent El Niño years, and the low TC counts of the early 1990s correspond to general tendency for El Niño-like conditions. The NAO has a weaker, but nonetheless statistically significant impact on TC counts, with a tendency for elevation of counts during negative NAO years. The Pearson correlation coefficients between the TC counts and the three predictors ($r = 0.48$ for MDR SST, $r = -0.32$ for Niño3.4, and $r = -0.25$) are statistically significant at the $p < 0.0001$, $p = 0.0001$, and $p = 0.003$ levels respectively for a two-sided hypothesis test, taking into account the serial correlation in each series. The extent to which these state variables can account for the non-random structure in long-term TC counts is investigated below using each of the two methods discussed in section 3.

4.1. Binary Classification Approach

[21] We first note that the unconditional distribution of TC counts is highly inconsistent with the null hypothesis of a random Poisson process. Based on a χ^2 test (Table 1) we reject at the $p < 0.05$ level the null hypothesis of a Poisson process for the entire TC count record 1870-2004. By inspection (Figure 2, panel *a*), it is clear that there is bimodality in the distribution which cannot be captured by the model of a constant mean Poisson process.

[22] Conditioning on ENSO influences (i.e., on Niño3.4) alone does not ameliorate this problem, as the conditional distributions for negative Niño3.4 values (i.e., 'La Nina'-like behavior) is still observed (Table 1) to be inconsistent ($p < 0.05$) with a Poisson distribution. Conditioning on MDR SST provides significant improvement, though the p values ($p = 0.79$ and $p = 0.25$ for +MDR SST and $-$MDR SST respectively) average only just above the median ($p = 0.5$) level between acceptance and rejection of the null hypothesis. However, when TC counts are simultaneously conditioned on both Niño3.4 and MDR SST, we find that the null hypothesis can likely not be rejected. The resulting three separate distributions ('favorable', 'neutral', and 'unfavorable', as defined in section 3.1) are generally well captured by a Poisson distribution (Figure 2, panels *b*-*d*). While in one of the three cases ('favorable') the p value ($p = 0.27$) indicates a moderate 27% chance of falsely rejecting the null hypothesis, the χ^2 tests yield an average value $p = 0.70$ for the three cases, well above the median expected level for false rejection of the null hypothesis. The results of the analysis are therefore consistent with the hypothesis that the annual TC counts are produced by a state-dependent Poisson process, with the occurrence rate being dictated by two state variables (Niño3.4 and MDR SST).

[23] Having established the viability of a state-dependent Poisson random model for the observed TC count data, we

Table 1. Results of Reduced χ^2 Tests Described in Text[a]

Scenario (1870-2004)	χ^2/ν	ν	p
All Years	2.09	9	0.027
+MDR SST	0.59	8	0.79
$-$MDR SST	1.32	3	0.25
+Nino3.4	1.02	8	0.42
$-$Nino3.4	2.29	7	0.025
+MDR/$-$Nino ('Favorable')	1.27	6	0.27
$-$MDR/+Nino ('Unfavorable')	0.28	9	0.98
+MDR/+Nino or $-$MDR/-Nino ('Neutral')	0.49	7	0.85

[a]Indicated are reduced χ^2 value (χ^2/ν), degrees of freedom ν and the p value for rejection of the null hypothesis of a poisson distribution.

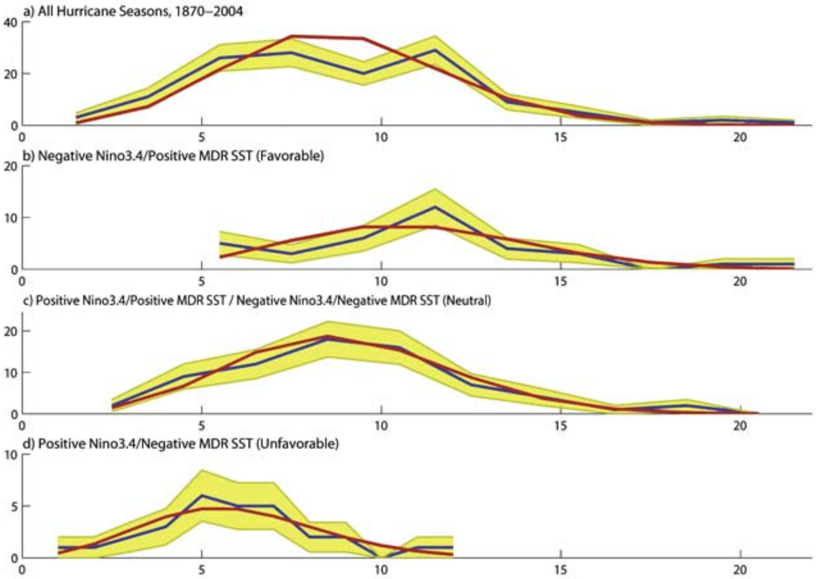

Figure 2. Histograms of TC counts n vs. bin centers (blue) with associated one standard deviation uncertainties ($\pm\sqrt{n}$, yellow shading) and best fit Poisson distributions (red). Results are shown for unconditional case (all data—panel *a*) and the 'favorable', 'neutral', and 'unfavorable' sub-groupings discussed in the text (panels *b-d*). Bin bandwidths were determined as discussed in text.

assessed the statistical significance of differences in the estimated conditional occurrence rates μ. There is a clear dependence of μ both on each of the two state variables separately and on the sub-categorization into the three 'favorable', 'neutral', and 'unfavorable' cases (Table 2). The highest average annual TC count is found for the 'favorable' state ($\mu \approx 11$), while the lowest ($\mu \approx 6$) is found for the 'unfavorable' state, with all other sub-groupings yielding intermediate values of μ. While differences in occurrence rate (Table 3) are highly significant conditioning on either one of the two state variables (Niño3.4 or MDR SST) alone, the most significant difference (i.e., lowest p value) is observed conditioning on both state variables (i.e., the 'unfavorable' vs. 'favorable' categories). Partitioning into the 'favorable', 'neutral', and 'unfavorable' categories yields both individual distributions that as noted earlier are on average consistent with Poisson, and mean TC occurrence rates that differ significantly between any two categories (Table 3). The MSE (Table 4) using the conditional means from the binary classification approach (MSE = 10.80 for the full 1870-2004 model development interval, and MSE = 11.79 in cross-validation) represents a significant improvement over climatology (MSE = 13.75) or persistence (MSE = 19.89). The cross-validation results, however, suggest that the binary classification approach gives moderately less predictive skill than the Poisson regression approach, as discussed in more detail below.

4.2. Poisson Regression

[24] We performed univariate Poisson regression alternatively using (i) MDR SST and (ii) Niño3.4 as state variables, (iii) bivariate regression using both MDR SST and Niño3.4 as state variables, and (iv) multivariate regression using all three climate state variables MDR SST, Niño3.4, and NAO (Figure 3a). Cross-validated resolved variance R^2 and MSE scores were similar to the scores obtained from the full model development interval 1870-2004, and far superior to either climatology or persistence, indicating significant skill in each of the regression models. Interestingly, the predictive skill systematically increases while the consistency of residuals (see Figure 3b) with a Poisson distribution decreases as additional state variables are added to the

Table 2. Estimates of Occurrence Rate μ for the Various TC Data Sub-Groupings Discussed in Text[a]

Scenario (1870-2004)	μ	N	ϕ
All Years	8.85	135	
+MDR SST	10.33	64	28
−MDR SST	7.52	71	31
+Nino3.4	7.78	58	
−Nino3.4	9.66	77	
+Nino/+MDR ('Favorable')	10.94	35	20
−MDR/+Nino ('Unfavorable')	5.97	29	20
+MDR/+Nino or −MDR/-Nino ('Neutral')	9	71	33

[a]Provided are the sample sizes N and, where appropriate, the effective sample size φ accounting for temporal autocorrelation in state variables.

Table 3. Results of t Tests for Differences of Occurrence Rates μ Among the Different Sub-Groupings Discussed in Text[a]

Scenario (1870-2004)	t	Φ	P
+MDR SST vs. −MDR SST	3.59	27	0.0006
+Nino3.4 vs. −Nino3.4	3.70	57	0.0002
Favorable vs. Unfavorable	5.41	19	<0.0001
Favorable vs. Neutral	2.15	19	0.02
Neutral vs. Unfavorable	4.02	19	0.0004

[a]Indicated are the effective degrees of freedom in the t statistic $\Phi = \min(\varphi_1, \varphi_2) - 1$, and the one-tailed p value for rejection of the null hypothesis of equal means.

Table 4. Assessments of Predictive Skill for Competing Statistical Models Considered in This Study[a]

Model/Predictors	R^2 full	MSE full	R^2 valid.	MSE valid	p resid.
Climatology	0.00	13.75			
Persistence	0.07	19.89			
Binary Cond: MDR, Nino		10.80		11.79	
Poisson Reg: MDR	0.24	10.81	0.16	10.47	0.83
Poisson Reg: Nino	0.10	12.51	0.12	12.31	0.08
Poisson Reg: MDR, Nino	0.33	9.37	0.26	9.95	0.35
Poisson Reg: MDR, Nino, NAO	0.38	8.70	0.32	9.02	0.00

[a]Mean square error (MSE) over the full model development period (1870-2004) is indicated for each case. The MSE for simple (i) climatological mean and (ii) persistence predictions is provided for comparison. In the case of Poisson regression models, the coefficient of determination (R^2) is also provided. Validation MSE and R^2 scores are based on the split calibration/validation procedures described in the text.

regression—i.e., first MDR only, then MDR and Niño3.4, and finally MDR, Niño3.4 and NAO (Table 4). Improved skill thus appears to come at a cost of increased bias in the conditional TC rate estimates.

[25] Each of the Poisson regression models are seen to improve significantly (as measured by both full 1870-2004 model development interval and cross-validation MSE scores) over climatology (Table 4). Moreover, both bivariate and three variable Poisson regression models yield significant improvements (as measured by MSE scores) over the binary classification approach with MDR SST and Niño3 outlined in section 4.1. This further suggests a tendency for a tradeoff between resolved variance (as determined from regression and validation R^2 and MSE scores) and bias (as determined from the distribution of residuals) in modeling TC counts. While the binary classification approach yielded the greatest consistency with a pure state-dependent Poisson process (as conditional distributions were consistent with Poisson at a mean level $p = 0.70$), it also produced the least resolved variance in modeling annual TC counts by conditioning on two or more climate state variables.

4.3. Predictions

[26] The binary classification approach to modeling TC numbers yields a simple forecasting scheme for seasonal TC counts. Depending on the forecast values for the two state variables (MDR ASO SST and DJF Niño3.4 anomalies) at the start of the tropical cyclone season (June 1st), the predicted TC total would be $\mu = 6 \pm 3$ (i.e., between 3 and 9) for 'unfavorable' anomaly combinations, $\mu = 9 \pm 3$ (between 6 and 12) for 'neutral' anomaly combinations, and $\mu = 11 \pm 3$ (between 8 and 14) for 'favorable' anomaly combinations. It is instructive to interpret the two most recent (2005 and 2006) Atlantic tropical storm seasons in this context. The TC count for the 2006 season ($n = 10$) was consistent with the predicted count ($m = 9 \pm 3$) given the observed 'neutral' conditions (positive MDR SST anomaly and positive 2006/2007 DJF Niño3.4 anomaly—see Table 5). The 2005 TC count ($n = 28$) is considerably more difficult to explain, even given the 'favorable' (positive 2005 MDR SST and negative 2005/2006 DJF Niño3.4) observed conditions, for which the predicted count is $m = 11 \pm 3$. Given a mean expected rate $\mu = 11$, the probability of equaling or exceeding a TC count of $n = 28$ is $\approx 0.01\%$, i.e., implausible.

[27] The Poisson regression models all successfully predict the 2006 TC count within estimated uncertainties, but like the binary classification approach, all significantly under-predict the historic 2005 TC total of $n = 28$ storms (Table 5, and also Figure 3a). However, the most skillful of the Poisson regression models as judged by cross-validation results (i.e., Table 3)—the three state variable model—comes closest to the observed total with a predicted TC count of $m = 18 \pm 4$ The high predicted total in this case is a result of simultaneously favorable conditions in all three state variables (anomalously warm MDR ASO SSTs, La Nina conditions in the tropical Pacific, and a substantially negative phase NAO). Given a conditional expected mean rate $\mu = 18$, the probability of observing or exceeding $n = 28$ storms is approximately 2%. In other words, for every 50 years with conditions similar to those observed for 2005, a TC count as high or higher than that observed might be expected given the three variable Poisson regression model. In this case, the 2005 TC total is still observed to be improbable, but not entirely implausible. It is of course

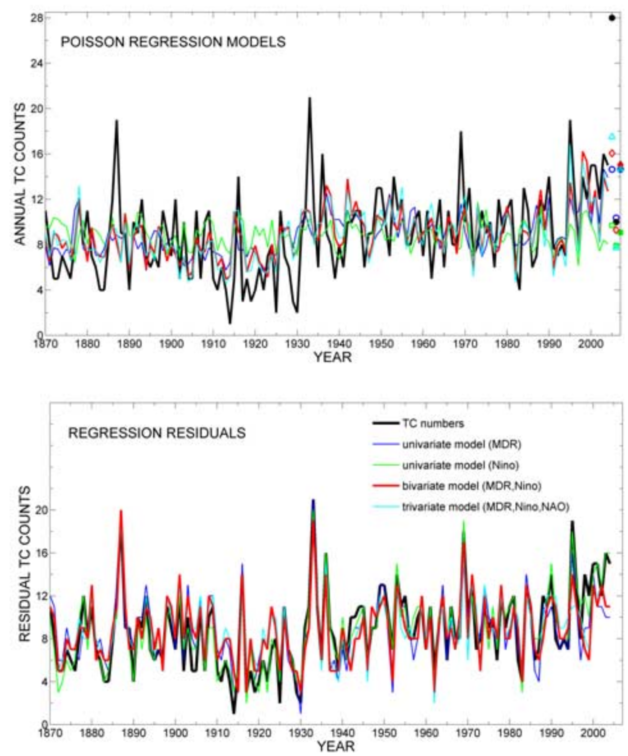

Figure 3. Poisson regression models of annual Atlantic TC counts using the MDR ASO SST, Niño3.4, and NAO series as predictors. Shown are (a) the statistical model fits over 1870-2004 based on the two univariate, bivariate and three-variable Poisson regressions (colored curves) along with the observed TC counts for 1870-2004 (black curve), observed TC counts for 2005 and 2006 (filled black circles), predicted TC counts for 2005 and 2006 (unfilled colored symbols) and 2007 (filled colored symbols). (b) Poisson regression residuals as defined in text (colored curves) along with the observed TC counts for 1870-2004 (black curve).

Table 5. Climate State Variable Values and Associated Annual TC Count Predictions m and Associated One Standard Error Uncertainties $\pm\sqrt{m}$ for 2005-2007[a]

Year	Model	MDR	Nino3.4	NAO	Predicted (n)	Observed (m)
2005	Binary conditioning	+	-	x	***11 ± 3***	***28***
	Poisson regression	x	−0.65	x	***10 ± 3***	
		28.87C	x	x	***15 ± 4***	
		28.87C	−0.65	x	***16 ± 4***	
		28.87C	−0.65	−0.82	***18 ± 4***	
2006	Binary conditioning	+	+	x	***9 ± 3***	***10***
	Poisson regression	x	0.72	x	***8 ± 3***	
		28.35C	x	x	***10 ± 3***	
		28.35C	0.72	x	***9 ± 3***	
		28.35C	0.72	2.43	***8 ± 3***	
2007	Binary conditioning	+	-	x	***11 ± 3***	***To be determined***
	Poisson regression	x	−0.2	x	***10 ± 3***	
		27.9C[b]	x	x	***15 ± 4***	
		27.9C[b]	−0.2[b]	x	***15 ± 4***	
		27.9C[b]	−0.2[b]	0.47[b]	***15 ± 4***	

[a]2007 climate variables are forecast based on the procedure described in the text.
[b]Predicted value.

possible that the true distribution of TC occurrence is heavy-tailed, in which case the probability of very large counts might be substantially greater than estimated under the assumption of conditional Poisson statistics. One could conceivably also argue that biases in the earlier data [e.g., *Landsea*, 2005] leads to an underestimation of the frequency of very large annual counts such as observed in 2005. However, our finding in section a that long-term TC data are essentially consistent with random Poisson statistics after controlling for dependence on two climate state variables, would seem to argue against the proposition that systematic biases compromise the reliability of the earlier data [*Landsea*, 2005].

[28] Finally, we use the statistical models developed above to forecast Atlantic TC counts for the 2007 tropical storm season. At the time this manuscript was finalized, weak La Nina conditions (Nino3.4 = −0.2) were predicted by NCEP for winter 2007/2008. MDR SST anomalies were currently similar to those observed for the 2005 season, so we infer by persistence ASO MDR SST anomalies equal to those for the 2005 season. As there is no basis for forecasting the winter 2007/2008 NAO value, we assume climatological mean DJFM conditions (NAO index = 0.47). Given these assumed values, the binary classification approach yields the 'favorable' forecast m = 11 ± 3, while each of the Poisson regression models (with the exception of the Niño3.4-only regression which yields a forecast m = 11 ± 3) predict a total of m = 15 ± 4 storms for the 2007 tropical storm season.

5. Conclusions

[29] Two different methods, a binary classification scheme and Poisson regression, are used to condition expected annual TC counts on climate state variables. Modeling annual Atlantic TC counts as a state-dependent Poisson process using the binary classification approach, we find that two climatic factors, ENSO and tropical North Atlantic MDR SST, are adequate to explain the apparent non-random variability in historical variations in Atlantic TC numbers. Modeling TC counts instead using Poisson regression, we find that the most skillful statistical model employs all three state variables considered in the study, ENSO, tropical North Atlantic MDR SST, and the NAO, as predictors. This three variable statistical model also comes closest to predicting the historic 2005 TC count of 18, ascribing unlike the other statistical models developed in this study, a non-trivial probability for that event given the climate state of 2005. However, analysis of residuals also indicates some evidence of bias, implying the need for cautious use of the model. Three of the four Poisson regression models developed in the study predict 15 ± 4 storms for the 2007 Atlantic tropical storm season.

[30] **Acknowledgments.** We thank two anonymous reviewers for their helpful comments. We also acknowledge generous financial support (for T.S.) provided by the Pennsylvania State University Schreyer Honours College.

References

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 Am. Meteorol. Soc*, *79*, 2477–2482.
Elsner, J. B. (2003), Tracking hurricanes, *Bull Am. Meteorol. Soc*, *84*, 353–356.
Elsner, J. B. (2006), Evidence in support of the climate change: Atlantic hurricane hypothesis, *Geophys. Res. Lett.*, *33*, L16705, doi:10.1029/2006GL026869.
Elsner, T. H., and J. B. Jagger (2006), Prediction models for annual U.S. hurricane counts, *J. Climate*, *19*, 2935–2952.
Elsner, J. B., T. Jagger, and X.-F. Niu (2000), Changes in the rates of North Atlantic major hurricane activity during the 20th century, *Geophys. Res. Lett.*, *27*, 1743–1746.
Elsner, J. B., B. H. Bossak, and X.-F. Niu (2001), Secular changes in the ENSO-U.S. hurricane relationship, *Geophys. Res. Lett.*, *28*, 4123–4126.
Elsner, J. B., 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. (2005a), Increasing destructiveness of tropical cyclones over the past 30 years, *Nature*, *436*, 686–688.
Emanuel, K. (2005b), Emanuel replies, *Nature*, *438*, E13.
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.
Gray, W. M. (1984), Atlantic seasonal hurricane frequency, Part I: El Niño and 30 mb quasi-biennial oscillation influences, *Mon. Wea. Rev.*, *115*, 1649–1668.
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 Natl. Hurricane Cent., Miami, Fla.

Jones, P. D., T. Jónsson, and D. Wheeler (1997), Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and South-West Iceland, *Int. J. Climatol.*, *17*, 1433–1450.
Kaplan, A., M. A. Cane, Y. Kushnir, A. C. Clement, M. B. Blumenthal, and B. Rajagopalan (1998), Analyses of global sea surface temperatures 1856-1991, *J. Geo. Res.*, *103*(C9), 18,567–18,589.
Landsea, C. W. (2005), Hurricanes and global warming, *Nature*, *438*, E11–E13.
Landsea, C. W., R. A. Pielke Jr., A. M. Maestas-Nunez, and J. A. Knaff (1999), Atlantic basin hurricanes, *Clim. Change*, *42*, 89–129.
Mann, M. E., and K. A. Emanuel (2006), Atlantic hurricane trends linked to climate change, *Eos*, *87*, 24, p. 233, 238, 241.
Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res.*, *108*(D14), 4407, doi:10.1029/2002JD002670.
Solow, A. R., and L. Moore (2000), Testing for a trend in a partially incomplete hurricane record, *J. Climate*, *13*, 3696–3699.
Sriver, R., and M. Huber (2006), Low frequency variability in globally integrated tropical cyclone power dissipation, *Geophys. Res. Lett.*, *33*, L11705, doi:10.1029/2006GL026167.
Trenberth, K. E. (1997), The definition of El Niño, *Bull Am. Meteorol. Soc.*, *78*, 2771–2777.
Trenberth, K. E., and D. J. Shea (2006), Atlantic hurricanes and natural variability in 2005, *Geophys. Res. Lett.*, *33*, L12704, doi:10.1029/2006GL026894.
Webster, P. J., G. J. Holland, J. A. Curry, and H. R. Chang (2005), Changes in tropical cyclone number, duration, and intensity, in warming environment, *Science*, *309*, 1844–1846.
Wilks, D. S. (2005), Statistical Methods in the Atmospheric Sciences: An Introduction, 2nd Edition, Academic Press, 520 pp.

M. E. Mann and T. A. Sabbatelli, Department of Meteorology and Earth and Environmental Systems Institute, Pennsylvania State University, 523 Walker Building, University Park, PA 16802, USA. (mann@psu.edu)