Stratified statistical models of North Atlantic basin-wide and regional tropical cyclone counts

Michael E. Kozar[1]
Department of Meteorology, The Pennsylvania State University,
University Park, PA 16802

Michael E. Mann
Department of Meteorology, The Pennsylvania State University,
University Park, PA 16802

Suzana J. Camargo
Lamont-Doherty Earth Observatory, Columbia University,
Palisades, NY 10964

James P. Kossin
NOAA's National Climatic Data Center,
Asheville, NC 28801
Cooperative Institute for Meteorological Satellite Studies,
Madison, WI 53706

Jenni L. Evans
Department of Meteorology, The Pennsylvania State University,
University Park, PA 16802

---

[1] *Corresponding author address:* Michael Kozar, Center for Ocean-Atmospheric Prediction Studies,
Florida State University, 2035 E. Paul Dirac Dr., 200 RM Johnson Bldg., Tallahassee, FL 32306-2840
Email: mkozar@coaps.fsu.edu

Abstract

Using the historical Atlantic tropical cyclone record, this study examines the empirical

relationships between climate state variables and Atlantic tropical cyclone counts. The state

variables considered as predictors include indices of the El Niño/Southern Oscillation and

Northern Atlantic Oscillation, and both "local" and "relative" measures of Main Development

Region sea surface temperature. Other predictors considered include indices measuring the

Atlantic Meridional Mode and the West African monsoon. Using all of the potential predictors in

a forward stepwise Poisson regression, we examine the relationships between tropical cyclone

counts and climate state variables. As a further extension on past studies, both basin-wide named

storm counts and cluster analysis time series representing distinct flavors of tropical cyclones,

are modeled. A wide variety of cross validation metrics reveal that basin-wide counts or sums

over appropriately chosen clusters may be more skillfully modeled than the individual cluster

series. Ultimately, the most skillful models typically share three predictors: indices for the main

development region sea surface temperatures, the El Niño/Southern Oscillation, and the North

Atlantic Oscillation.

1. Introduction

The potential origins of interannual and longer-term variability in North Atlantic tropical

cyclones (TCs) have been investigated in numerous studies over the past decade. Anomalous

recent levels of activity, and in particular the record-breaking 2005 Atlantic Hurricane Season,

have spurred scientific interest in this topic. A number of recent studies have used statistical

regression models to examine the apparent impact of climate state variables on TC activity,

including recent trends.

The El Niño/Southern Oscillation (ENSO) has long been known to impact Atlantic TC activity, with El Niño (La Niña) tending to diminish (enhance) seasonal TC activity [e.g. Gray, 1984]. Indices of ENSO accordingly represent a predictor commonly used in statistical modeling of Atlantic TC activity [e.g. Bove et al., 1998; Elsner, 2003; Elsner and Jagger, 2006; Mann et al., 2007]. A number of studies have also considered the role of the Northern Atlantic Oscillation (NAO) in Atlantic TC activity [e.g. Elsner et al., 2000b; Elsner, 2003; Elsner and Jagger, 2006; Mann et al., 2007], which influences seasonal TC activity through an influence on large-scale storm tracks [e.g. Elsner, 2003; Kossin et al., 2010]. A warm ocean favors the formation and development of TCs [e.g. Gray, 1968], as it is closely tied to key thermodynamic quantities involved in the energetics of TCs, such as potential intensity [Emanuel, 1995]. Numerous studies modeling Atlantic TC activity thus incorporate sea surface temperatures (SST) over the Main Development Region (MDR; 6°-18°N, 20°-60°W) during the primary season (Aug-Oct) for Atlantic TC formation [e.g. Hoyos et al. 2006; Emanuel 2005; Sabbatelli and Mann 2007; Mann et al 2007]. Recently, there has been some debate within the research community as to whether MDR SSTs themselves [e.g. Emanuel, 2005] or some "relative" measure of SST that measures warmth of the MDR relative to the tropical mean [e.g. Vecchi et al., 2008; Ramsay and Sobel, 2011] is a more appropriate measure of thermodynamic influences on Atlantic TC activity.

Two recent studies [Sabbatelli and Mann, 2007; Mann et al., 2007] have modeled annual basin-wide TC counts on three predictors: ENSO, the NAO, and Aug-Oct MDR SSTs. These analyses employed Poisson regression, a tool that is appropriate for modeling a Poisson process with a rate of occurrence that is conditional on underlying state variables [e.g. Elsner et al., 2000b; Elsner, 2003]. The resulting statistical model displayed significant predictive skill, accounting in cross-validation for roughly half of the variance in annual Atlantic TC counts.

However, it may be possible to enhance the skill in TC count statistical modeling exercises by exploring a wider range of potential climate predictors. This study thus extends upon the framework established by these previous studies by employing a more exhaustive analysis of Atlantic TC counts, using a larger array of candidate predictors in the context of Poisson regression-based statistical modeling exercises.

Additional potential predictors to consider include rainfall in the Sahel region in Western Africa during the boreal summer. Studies have found that negative precipitation anomalies in this region coincide with an increase of dry African dust layers, which have the potential to inhibit TC genesis [Prospero and Lamb, 2003]. Moreover, an anomalously dry season has the potential to alter the characteristics of moist easterly waves that can eventually develop into TCs [Goldenberg and Shapiro, 1996]. Furthermore, previous studies hypothesize that droughts in Western Africa are consistent with stronger upper-level westerlies that can increase the amount of shear in the Atlantic basin [Landsea and Gray, 1992]. Recent studies use this information to correlate precipitation patterns from the African monsoon to Atlantic TC development, but this relationship is non-stationary [Bell and Chelliah, 2006; Zhang and Delworth, 2006; Fink et al., 2010]. While the stratospheric Quasi-Biennial Oscillation (QBO) was argued to influence in some early work [Gray, 1984], recent work of Camargo and Sobel [2010] questions the existence of any such relationship in the modern record. A more practical limitation is that there is no reliable long-term record of the index. It was therefore not included among the potential predictors.

Furthermore, Kossin and Vimont [2007] suggested that the Atlantic Meridional Mode (AMM), a dynamically- reproducible mode of variability, might explain the overarching variability in the tropical North Atlantic. Much like ENSO in the Pacific, the AMM is the leading

mode of a coupled air and sea feedback process in the Atlantic. Therefore, climate features such as Caribbean sea level pressure (SLP) anomalies, African rainfall amounts, and the areal extent of the Atlantic Warm Pool (AWP) are suspected to be manifestations of the AMM.

Several alternative metrics of ENSO [Barnston et al., 1997] also could be employed in this analysis (see Figure 1). In addition to the Niño3.4 index favored by, for example, Mann et al. [2007], one might alternatively employ the Niño3 or Niño1+2 indices of ENSO [e.g. Kossin et al., 2010]. Finally, some researchers have hypothesized that the influence of tropical Atlantic SSTs on Atlantic TC activity is best measured through a "relative" MDR SST index [Vecchi et al. 2008], the difference between the MDR and the global tropical mean SSTs, or through consideration of the North Atlantic SST and the mean SST over the global tropics separately [Villarini et al., 2010]. The global tropical mean SST index therefore is considered as a potential predictor, in addition to the "absolute" and "relative" measures of the Atlantic MDR SST. Incorporating all of these potential alternative climate state variables into an extensive pool of candidate predictors allows for a more comprehensive and robust exploration of the appropriate statistical models relating climate and Atlantic TC counts, building upon previous works such as those done by Villarini et al. [2010] and Sabbatelli and Mann [2007].

In addition to expanding the pool of potential predictors, it is worthwhile to decompose historical basin-wide TC counts into subgroups that cluster with respect to their path, location of genesis, and other track characteristics [Elsner et al., 1996; Nakamura et al., 2009; Kossin et al. 2010]. Previous studies and annual seasonal forecasts have displayed skill in explaining much of the nonrandom interannual variance in North Atlantic TC counts as a whole [e.g. Sabbatelli and Mann, 2007; Mann et al, 2007]. However, successful seasonal predictions of particular "flavors" of TCs in the Atlantic basin have remained elusive. In principle, the factors that govern different

flavors of TCs may differ, and additional predictive skill, as well as insight, might arise from modeling them separately, rather than collectively [e.g. Lehmiller et al., 1997].

2. Data

Sabbatelli and Mann [2007] and Mann et al. [2007] used three predictors in modeling annual Atlantic basin-wide TCs: (1) the post-season boreal winter December-February (DJF) Niño 3.4 SST index, (2) the post-season boreal winter December-March (DJFM) NAO index (Jones et al. 1997) and (3) the in-season August-October (ASO) mean MDR SSTs. The SST predictors here are derived from a consensus SST product, which is based on a blend of three individually published SST products [Rayner et al., 2003; Smith and Reynolds 2003; Kaplan et al. 1998]. Instead of focusing on how the use of each individual SST product affects the results, as was done previously [e.g. Villarini et al., 2010], this study focuses on the sensitivity between TCs and the respective climate state variables through the blended consensus products for all SST based predictors.

The pool of candidate predictors in our study includes these indices, but the additional series discussed in the introduction are also tested: (4) the in-season June-September (JJAS) Sahel precipitation index (from the Joint Institute for the Study of the Atmosphere and the Ocean "JISAO") [2009], and alternative ENSO indices of the post-season boreal winter including the (5) Niño1+2 (from the NOAA Climate Prediction Center, NOAA/CPC) and (6) the Niño 3 SST indices. Also considered as a candidate predictor is the (7) ASO "relative" MDR SST index, which is calculated by subtracting a time series of SSTs averaged across the global tropics (23.5°N to 23.5°S) from the averaged North Atlantic MDR SST series. In addition, the pool of candidate predictors includes, the (8) ASO global tropical SSTs that were used to calculate the relative MDR index above, (9) the May-June (MJ) NAO index [Jones et al., 1997] and (10) the

in-season June-November (JJASON) AMM index (from NOAA/CPC). These datasets are all available across various time periods, ranging as far back as the mid-to-late 1800s. Plots of all time series are shown in Figure 2, and further details are provided in Table 1.

As described above and in Table 1, some of the potential predictors in this modeling study are averages taken during or after the season, e.g. the boreal DJF winter. Thus, incorporation of these predictors into statistical seasonal forecast models relies on the assumption of statistical persistence of the associated anomalies. Despite the potentially limiting nature of that assumption, seasonal predictions incorporating such predictors have proven remarkably successfully, matching or beating in their skillfulness the predictions made by other approaches (see: http://www.essc.psu.edu/essc_web/research/ Hurricane2012.html). In the current study, our emphasis and interest is not with seasonal forecasting, however, but instead a better understanding of the physical ties between TC variability and tropical climate state variables, whether they are measured before, during or after the season.

In addition to considering this expansive set of tropical predictors, it may be desirable to partition the basin-wide TC best-track database ("HURDAT") [Jarvinen et al., 1984] into clusters of like storms. Grouping storm counts by storm track using a "cluster" methodology has proven to be advantageous in several previous studies, as it allows for an analysis that can relate regional TC activity to spatially varying climate features such as ENSO, AMM, NAO, and MJO [e.g. Camargo et al., 2007 a; b; Camargo et al., 2008; Kossin et al., 2010]. Specifically, this cluster technique has been previously applied to Atlantic extratropical cyclones [Gaffney et al., 2007], western North Pacific typhoons [Camargo et al., 2007 a; b], eastern North Pacific hurricanes [Camargo et al., 2008], Fiji TCs [Chand and Walsh, 2009] southern hemisphere TCs [Ramsay et al., 2012], and the most relevant to this study—North Atlantic TCs [Kossin et al., 2010]. This

path clustering technique is accomplished by utilizing a mixture model, in which every component consists of a quadratic regression curve of TC position versus time. The model is fit to the data via a maximum likelihood estimation of the parameters [e.g. Kossin et al., 2010]. Each TC track is then assigned to one of the K different quadratic regression models, with each model being described by regression coefficients and a noise matrix. As is the case in the K-means method, the number of clusters used in this methodology is not uniquely determined in the cluster analysis. Therefore, in-sample log-likelihood values are used to obtain the optimum number of clusters, just as in Kossin et al. [2010] and Camargo et al. [2007a; 2008]. Kossin et al. [2010] found that at least four clusters are needed to capture the track types being partitioned (Figure 3). Consequently, in addition to the basin-wide Atlantic TC series, our analysis uses an extension of the four-cluster decomposition of Atlantic TCs of Kossin et al. [2010], dating back to 1878 in this case instead of 1950. Further, by combining multiple clusters we can also analyze an inherently larger subset of TCs, focusing on, for example, clusters containing TCs of primarily tropical origin, or clusters that contain the majority of societally impactful TCs (i.e. intense hurricanes and landfalling storms).

Due to improvements over time in the detection of TCs from technological advances such as aircraft reconnaissance and satellites, there is likely a bias in estimates of basin-wide TC counts in earlier decades of HURDAT [see e.g. Landsea, 2007; Chang and Guo, 2007; Mann et al., 2007]. Therefore, a recent published adjustment to HURDAT [Vecchi and Knutson, 2008] is used in our modeling exercises. This adjustment, which dates back to 1878, provides an estimate of potential missing TCs in HURDAT prior to advent of weather satellites by using historical ship tracks in the pre-satellite era in combination with modern storm track information. All analyses in this article of the total basin-wide Atlantic TC counts, therefore, make use of the

Vecchi and Knutson [2008] ("VK08") estimates. Clearly, since individual cluster series are dependent on storm track information, these count adjustments cannot be translated to the individual cluster series. As a result, since the four individual cluster series are derived from counts and tracks within HURDAT, they likely contain an uncorrected undercount bias, especially early in the record (prior to the advent of aircraft reconnaissance).

Ultimately, this leads to a total of six target series—the VK08 adjusted basin-wide TC counts, the four individual TC cluster series defined in Kossin et al. [2010] and a combination of those clusters representing the majority of societally significant storms that are more likely to become intense hurricanes and/or make landfalls (clusters 2–4 in Kossin et al. 2010). We seek to derive objective, optimal statistical models in terms of the potential underlying climate factors (TC series shown in Figure 4) for each of these series.

A potential downside of modeling the individual cluster TC series is that the sample size contributing to seasonal mean counts is often greatly diminished relative to basin-wide counts. More specifically, the relative sampling uncertainty goes as the square root of the total number of storms divided by the total number of storms (i.e. $\sqrt{n_{TC,tot}} / n_{TC,tot} = n_{TC,tot}^{-1/2}$). Therefore, when analyzing exclusive quantities such as major hurricane landfalls or in this instance, individual clusters of TCs, the sample sizes may be prohibitively small for establishing statistical skill or significance in any underlying statistical model, which again underscores the utility of considering combinations of multiple clusters. Discussion of sampling sizes in this manner is not referring to the sample size of the individual time series, which is equivalent to the number of years in annual time series.

3. Methods

Using the expanded pool of ten predictors discussed in section 2 (Table 1; Figure 2), and the set of six target predictands (adjusted basin-wide TC counts, each of the four cluster TC count series, and the combination of multiple clusters; Figure 4), Poisson regression is applied to model climate influences on annual TC counts. This approach assumes TC counts can be appropriately represented by a Poisson process, in which the probability of observing a certain number of TCs (n_{TC}) in any given year is defined as

$$P(n_{TC}) = \left(\frac{1}{n_{TC}!}\right) \lambda^{n_{TC}} \exp(-\lambda) \qquad (1).$$

In a Poisson distribution, the mean occurrence rate, λ, is the sole free parameter, and the unconditional case has a maximum likelihood value of the mean annual count. The null hypothesis is that λ is a constant, while Poisson regression tests the alternative hypothesis that λ is a function of other variables, e.g. climate state variables. Poisson regression models are governed by the following equation

$$E = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots) \qquad (2),$$

where in our case, E is the expected number of TCs in a single year as predicted by the Poisson regression model, X represents a vector of the various climate predictors, and β contains the regression coefficients. Some recent studies suggest that alternative forms of regression models might also be suited to model the influence of climate state variables on TC activity [e.g. Villarini et al. 2010; Mestre and Hallegatte, 2007], but we choose to build on the rich body of work applying Poisson regression to modeling TC count data [e.g. Solow and Nicholls, 1990; Solow and Moore, 2000; Solow and Moore, 2002; Elsner et al., 2000b; 2001; Elsner, 2003; Elsner and Jagger, 2006; Villarini et al., 2010; Tippett et al., 2011].

A standard forward stepwise (Poisson) regression is applied to each of the six available predictands (Figure 4) using all possible combinations of n_x predictors (n_x=1,2,…,$n_{x,max}$; where $n_{x,max}$ is the maximum available number of predictors available). Not all predictors extend back over the full interval to 1878 (Table 1), so models are tested over three possible time intervals, the shortest of which allows testing of all predictors, and the longest of which has a smaller pool of candidate predictors: $n_{x,max}$= 7 for the analyses from 1878, $n_{x,max}$= 8 for the analyses from 1900, and $n_{x,max}$= 10 for the analyses from 1950. Since, each of the potential predictors measure tropical climate features, they overlap by varying degrees (Table 2). Therefore, to minimize redundancy among the predictors, at most one NAO index, one ENSO index, and one MDR or tropical mean SST index is used in any modeling exercise.

3.1. Development of the Poisson model via a stepwise forward regression approach

The Poisson regression model is constructed using all possible choices of a single predictor from the pool of $n_{x,max}$ available predictors. In the first iteration, the predictor that yields the lowest mean-squared error (MSE) is chosen. From the remaining pool of $n_{x,max}$-1 candidates, a second predictor is considered, now selecting the *bivariate* combination with the lowest MSE (i.e. the choice of second predictor is dependent on its interaction with the first predictor chosen). This procedure is repeated until all $n_{x,max}$ predictors are used. Cross-validation statistics (see below) are used to select the optimal order, $n_{x,opt}$, of the statistical model (with the goal that $n_{x,opt} << n_{x,max}$). The statistical model with $n_{x,opt}$ variables and the lowest averaged cross-validated MSE is selected.

Goodness of fit of the resulting statistical models is measured by a suite of metrics, including mean-squared error (MSE) and coefficient of determination (R^2) [e.g. Wilks, 2005] and

χ^2 statistics measuring both the goodness of fit of the statistical model and the "adequacy" of the fit. Adequacy is defined here by the χ^2 test of independence as the consistency of the residual variance with purely random Poisson statistical behavior. The χ^2 statistics are defined as:

$$\alpha_{fit} = \sum_{i=0}^{1} \sum_{j=0}^{1} \frac{(e_{ij} - o_{ij})^2}{o_{ij}} \quad (2) \qquad\qquad \alpha_{adequacy} = \sum_{i=1}^{n_{yrs}} \frac{(E_i - O_i)^2}{O_i} \quad (3),$$

where E once again is the expected TC counts from the Poisson regression model's predictions, O is the observed TC count, n_{yrs} is the number of years in the time interval, e_{ij} is a cell in the 2x2 contingency table constructed from expected model-predicted counts, and o_{ij} is a cell in a 2x2 contingency table constructed from the observed counts. Ultimately, as values of α_{fit} approach zero, the probability that such a skillful model would arise from chance alone becomes increasingly low. As values of $\alpha_{adequacy}$ approach one, the probability that residual unresolved variance is consistent with purely random Poisson process behavior becomes increasingly high.

Independent cross-validation experiments are used to evaluate the predictive skill of the underlying statistical models. In these experiments, the model is calibrated over one-half of the data set, and an independent prediction of TC counts is made for the other half, and the goodness of fit of the prediction is evaluated. The procedure is repeated alternatively using both the first and second half of the data for calibration/validation, and an average set of validation scores are obtained. A variety of statistical measures are favored in the climate literature for evaluating the cross-validation skill of statistical models. Among these are the validation MSE, and various forms of the coefficient of determination R^2 (which measures the fraction of variance resolved by the statistical model). Calculating a validation R^2 score adopting the outside-sample baseline mean yields what is referred to as the "reduction of error" ("*RE*"), while calculating a validation R^2 score adopting the out-of-sample baseline mean yields a somewhat more challenging metric,

that is sometimes referred to as the "coefficient of efficiency" or "*CE*". These two metrics are calculated as:

$$RE = 1 - \frac{\|E_{out} - O_{out}\|^2}{\|O_{out} - \overline{O_{in}}\|^2} \quad (4) \qquad CE = 1 - \frac{\|E_{out} - O_{out}\|^2}{\|E_{out} - \overline{O_{out}}\|^2} \quad (5),$$

where E again is the expected TC counts based on the prediction from the Poisson regression model, O is the number of TCs from the observed predictand, and the subscripts designate the in-sample or out-of-sample half of the model or predictand data. Additionally, the squared linear correlation coefficient (r^2) [e.g. Wilks, 2005] can be used to measure the fraction of resolved variance, though it should be noted that r^2 is insensitive to both the mean and variance of the estimate, and so is a somewhat less rigorous validation measure. In principle, the most skillful models should out-perform the others with respect to most, if not all, of these alternative cross-validation skill metrics.

In addition to the primary cross-validation methods detailed above, we adapt a one-year validation methodology to assess the skill of our basin-wide TC count models. To accomplish this we predict the number of total TCs in a single year, training the model on the data (predictors and TC counts) from all of the previous years. A table with the results of these tests, which verified to within one standard deviation of a Poisson process 80% of the time from 1983-2007, is included in a supplementary data file.

Overall, the principle behind our approach is to make the determination of statistical models linking climate and TCs as objective as possible. This goal is achieved by employing (i) a large pool of predictors that includes all or nearly all predictors that have been suggested in previous statistical modeling exercises, and (ii) an objective stepwise screening process both for

selecting predictors from a larger pool of candidate predictors, and for independently evaluating the skillfulness of competing statistical models.

4. Results

4.1. Basin-wide Tropical Cyclone Counts

Time series resulting from the three Poisson models (spanning the three data intervals) for the adjusted basin-wide TC counts are shown in Figure 5a, while key results from the calibration and cross-validation exercises are summarized in Table 3. The cross-validation results provide some support for recent work by Mann et al. [2007] favoring the use of *absolute* (rather than "relative") MDR SST, and the use of ENSO and NAO indices as additional predictors. The Niño 3.4 index is statistically favored over the other two (Niño 3 and Niño 1+2) ENSO indices considered in all but the middle time interval (1900-2007). NAO shows less robustness as a predictor of total counts, with no consistency in the choice of NAO predictor across the three models. Based on χ^2 tests on each interval (Table 3), the statistical models are highly significant ($\alpha_{fit} << 0.05$) and provide no compelling evidence for unresolved structure ($\alpha_{adequacy} \sim 0.70\text{-}0.90$).

As in Mann et al. [2007], just under half of the total variance ($R^2 = 0.43\text{-}0.49$) in annual basin-wide TC counts is resolved in calibration using the each of the training intervals, with cross-validation tests indicating similar levels of skill, depending on the precise skill metric used ($RE = 0.37\text{-}0.55$, $CE = 0.21\text{-}0.39$; $r^2 = 0.30\text{-}0.41$; Table 3). The models also indicate significant skill in reproducing lower frequency variability when a 5 and 10 year filter is applied (calibration and validation $r^2 \approx 0.5$). Overall, the validation scores in Table 3 give somewhat mixed results with respect to whether the adjusted "VK08" TC series or uncorrected TC series yields a more

skillful model, with MSE and *CE* favoring the VK08 series but R*E* and r^2 favoring the uncorrected series (Table 3).

Somewhat surprisingly, dependence of total-basin TC counts on the additional predictors tested here, such as the AMM index and Sahel precipitation, is not objectively supported by the forward regression exercises. It is possible that the shorter training intervals that are available for using these predictors (Sahel rainfall extends only back to 1900, and the AMM series extends only back to 1950) are simply not sufficient to identify potential additional useful information in these series. Redundancies between these unsupported predictors with the ENSO, NAO, and/or SST indices identified in the model development (Table 2) also contribute to their lack of inclusion in the model. Indeed, one complication in comparing the relative merit of competing predictors over relatively short (i.e. less than 60 year) time intervals is the fact that the competing predictors differ substantially in their frequency-domain attributes [see e.g. Mann and Emanuel, 2006]. For example, the AMM is closely correlated with Atlantic TC activity on both interannual and decadal timescales [Kossin and Vimont, 2007; Vimont and Kossin, 2007] while MDR SST is most strongly correlated with Atlantic TC activity on decadal and longer timescales [e.g. Mann and Emanuel, 2006]. Furthermore, we must also consider the possibility that climate state variables utilized in these models may be subject to biases due to evolving observation and data sampling techniques. In combination with the previously discussed TC undercount bias, these factors may explain why the correlation between the adjusted counts and the statistical model is much higher in the modern era than it is in the 19th and early 20th centuries (Figure 5a).

4.2. Tropical Cyclone Cluster Time Series

The results for the models of the four individual TC cluster counts (Figures 5b-e) are more mixed than those obtained for models of basin-wide TC counts. Overall, a larger mix of

predictors emerge in the forward stepwise screening regression approach, but in most cases either the statistical skill, regression adequacy, or both are called into question by the statistical results obtained.

The first cluster (Figure 4b) contains TCs that originate primarily over the north and eastern parts of the basin (Figure 3a); these storms tend to have significant curvature in their paths that is modulated by subtropical high variability, placing these TCs into a more baroclinic environment (Kossin et al. 2010). Overall, this cluster is the most populated of the four clusters, accounting for 31% of all basin-wide TCs from 1878 to 2010. Models trained on this cluster (Figure 5b) tend to be less skillful than those found above for basin-wide TC counts (Table 3), though interestingly, two of the same predictors from the total basin-wide count models (MDR SST and winter NAO) are nonetheless chosen in the forward stepwise screening regression procedure for cluster 1 TCs using the full (1878-2007) interval. The calibration and cross-validation results suggest that less than 20% of the interannual variance is resolved by the statistical model, though CE scores are close to zero, indicating that much of the out-of-sample skill comes from predicting the changes in mean counts. The level of skill in the statistical models using shorter training intervals is further reduced and there is substantial evidence (i.e. low values of $\alpha_{adequacy}$) for unresolved structure in these models. For the shortest training interval (1950-2007), the selected model does not pass statistical significance ($\alpha_{fit} > 0.05$). Interestingly, this first cluster is the only one for which there is some hint of the predictive value of Sahel rainfall, however the cross-validation results in this 1900-2007 case indicate limited statistical skill.

The second cluster (Figure 4c) contains TCs that typically originate in the Gulf of Mexico or the Caribbean Sea and have a northward component in their tracks (Figure 3b). TCs within

this cluster are strongly modulated in intraseasonal time-scales by the Madden-Julian Oscillation [Kossin et al., 2010] and account for 29% of total basin-wide observed TCs from 1878-2007. Cross-validation results for this cluster are uniformly poor. Since Caribbean genesis events form a substantial portion of this cluster, the differing TC variability with ENSO between the western and eastern Caribbean may be confounding this signal. For example, phases of eastern Caribbean TC activity correspond better with decadal trends in ENSO than do phases of western Caribbean TCs [e.g. Evans et al., 2011; Giannini et al., 2000, 2001]. It is also possible that the standard climate factors considered here simply do not have any decisive relationship with this particular family of Atlantic TCs.

The third cluster (counts: Figure 4d; tracks: Figure 3c) includes many of the more intense TCs and ones that largely form in the eastern part of the North Atlantic Basin. This cluster contains a large fraction of recurving TCs, and the highest percentage of hurricanes compared to the other clusters [Kossin et al., 2010]. From 1878-2007, this cluster contains 22% of all storms in the basin. The statistical models for this particular cluster (Figure 5d) show among the greatest apparent skill of all four clusters considered, though the results are quite variable with respect to the time interval considered. The chosen predictors are again similar—MDR SST, ENSO, and NAO, though the flavors of the indices chosen are different (Niño1+2 for ENSO and the pre/within-season MJ NAO). Interestingly, the mean Tropical SSTs are preferred over either the Atlantic MDR or the MDR anomaly index for the two longest intervals. The adequacy tests fail over these two intervals, however, ($\alpha_{adequacy} \ll 0.05$) suggesting that there is substantial unresolved structure in the residuals. Moreover, the cross-validation results are both inferior and highly variable with respect to the particular metric used for the two longer training intervals. Over the most recent interval (1950-2007), the cross-validation results indicate skill that is

competitive with that obtained above for basin-wide storms, suggesting that 30-50% of the interannual variance can be skillfully resolved. In this case, it is possible that the shortest interval was able to yield the most skill, since the cluster series is unadjusted for potential undercount biases and is particularly dependent on the observed storm tracks, which have been more accurately observed in the modern era, especially in the eastern part of the basin.

The fourth cluster of TCs (counts: Figure 4e; tracks: Figure 3d) corresponds to storms that develop primarily in the southern part of the basin and typically have relatively east to west tracks. Of the four clusters, this cluster is comprised of the fewest number of storms, accounting for only 19% of the total storm counts from 1878-2007. Cross-validation scores generally indicate that the statistical models (Figure 5e) can skillfully resolve more than 20% of the interannual variance, though the precise skill varies with the time interval considered. The most skillful model is derived for the most recent (1950-2007) interval, where, unlike any other cluster, the AMM (only available back to 1950) is chosen as one of the three predictors (the other two are MDR SST and Niño 1+2). Given the rather small area of the Niño 1+2 region, its inclusion both here and in models for cluster 3 is interesting. However, one interpretation of this correlation between the Niño 1+2 and these southernmost clusters may be that Atlantic TCs in the southern part of the basin are influenced by a larger area of tropical Pacific SSTs, possibly further to the north along the South American coast or even west into the eastern portions of the Niño 3 region.

Adequacy tests for the 1950-2007 interval, give an indeterminate result ($\alpha_{adequacy} \sim 0.3$) suggesting the possibility of unresolved residual structure. For the two longer intervals (1900-2007 and 1878-2007) a different set of predictors are chosen, including in both cases the "relative" MDR SST series, a series that is not selected for any of the other cluster series or for

basin-wide TC counts. In both cases, there is substantial unresolved structure in the residuals, in agreement with the adequacy test results ($\alpha_{adequacy}$<0.05).

Overall, models targeting the rather small cluster subsets proved to be less skillful than the models targeted on basin-wide activity as a whole. Therefore, we also consider a larger subset comprised of the sum of the last three clusters (Figure 4f). This combination is selected in order to remove the weaker and mostly open ocean storms in cluster 1, which will allow the analysis to focus on storms of greater societal impacts (i.e. landfalling and more intense hurricanes). The resulting cluster combination accounts for roughly two-thirds of all observed TCs and 85% of all landfall events since 1950. It is of great interest that the vast majority of the most destructive storms in recent history are also included, with the only notable exception being Hurricane Bob (1991).

Not surprisingly, the same three predictors preferred in modeling basin-wide activity— absolute MDR SSTs, Niño 3.4, and DJFM NAO—were utilized in the models trained on this subset of TCs of primarily tropical origin. Overall, these Poisson models for the three-cluster subset of storms (Figure 5f) exhibit much more skill, based on cross-validation tests, than do the models that targeted on individual clusters. In fact, models across all three training intervals often match and in some cases exceed the calibration and cross-validation scores of the models trained on total TC counts. For all three training intervals, the models resolve more than 40% of the interannual variance, and χ^2 tests indicate high significance (α_{fit}<< 0.05). However, models trained on the longer two time intervals indicate the possibility of unresolved residual structure ($\alpha_{adequacy}$ < 0.5), possibly at least in part due to heteroscedasticity in the storm track errors through time.

A more restrictive combination of clusters 3 and 4, which focuses on TCs of mainly tropical origins in the southern and eastern parts of the basin, might also be considered as a potential target TC series. Such a partition focused on "tropical" type storms might be better represented by the tropical predictors used in this study. The resulting models for this partition (on each of the three time intervals) were consistent with, albeit less skillful than, those created for the less restrictive cluster combination (clusters 2,3,and 4). Once again, the models for in this partition favored the use of the "absolute" MDR SST index, the Niño 3.4 index, and the NAO index. Adequacy tests fail over all three of the intervals in this case, however, ($\alpha_{adequacy}$< 0.10) suggesting that there is substantial unresolved structure in the residuals.

5. Conclusions

The results of the statistical analyses with HURDAT, a recent set of adjusted counts, and observed climate predictors confirm previous findings [e.g. Mann et al., 2007], that Atlantic basin-wide TC counts can be skillfully modeled in terms of the MDR SST, ENSO, and the boreal winter post-season NAO index, resolving roughly half of the total interannual variance in both calibration and cross-validation. In the present analysis, the model predictors are chosen from a larger set of ten potential predictors (including other alternative predictors that have been argued for in some previous studies) through a forward stepwise screening regression approach. Thus, the recovery of the three predictors identified by Mann and collaborators provides confirmation of the significance of these climate factors on North Atlantic basin TC counts. Adequacy tests suggest little evidence of any residual unresolved structure in the resulting statistical model, hinting at the possibility that these three climate predictors account for essentially all non-random year-to-year variability in Atlantic basin-wide TC counts.

The results presented here do not support the use of a "relative" MDR SST index in place of the "absolute" MDR SST as a predictor of TC activity [Vecchi et al., 2008; Ramsay and Sobel, 2011]. The most skillful models for basin-wide storm counts, and the majority of the clusters, include the actual MDR SSTs as a predictor, rather than either the relative SSTs or the global tropical mean SSTs. Despite the high correlations between the different SST indices, simply substituting relative SSTs for "absolute" MDR SSTs in our basin-wide TC count models results in an increased mean-squared error of more than 25% for each of the target time intervals.

When observed Atlantic TCs are partitioned into four distinct "flavors" of TCs through a cluster analysis of Atlantic TCs [Kossin et al., 2010], the findings are neither as consistent nor as significant as they are with the total number of TCs in the adjusted historical record. For example, the second cluster — which largely includes storms of Gulf of Mexico and western Caribbean origin — shows no clear statistical relationship with any of the ten candidate climate indices considered. However, other clusters (associated for example with storms of tropical Atlantic origin) do appear to evidence some degree of skillfulness in being modeled in terms of climate predictors. Most commonly, the same predictors emerge – MDR SST, indices of ENSO, and indices of the NAO, usually in this order – in the statistical models for the TC cluster series that emerged in the model for basin-wide TC counts. Considering the upward ocean heat fluxes that drive TCs, it is not surprising physically that a measure of SST in the basin during the most active period of the season is the most significant predictor. Likewise, the consistent dependence of the models in this study on measures of ENSO and the NAO are expected due to their relationships to inhibiting wind shear and large scale steering respectively. It is interesting nonetheless, that in most cases the indexes for ENSO and the NAO are post-season averages, not in-season measurements. This is likely due to the fact that NAO and ENSO both exhibit a

stronger signal in the boreal winter than the spring or early summer. In addition, under certain

limited circumstances, predictors such as Sahel rainfall, the AMM, or "relative" MDR SST are

selected in the statistical models.

Models for combinations of clusters yielded higher skill than the individual clusters

alone, and are comparable with the full TC count models. Statistical tests for this series showed

consistent predictive skill across multiple training periods, similar to the models for basin-wide

activity. Encouragingly, the same three predictors—MDR SSTs, Niño 3.4, and DJFM NAO

emerged for the "cluster combinations" used here. Ultimately, the three clusters (2, 3, and 4)

included in the larger partition of storms contain 100% of all category 5 storms, 88% of all major

hurricanes, and 85% of all landfall events since 1950. Therefore, additional work on these types

of TCs is clearly warranted to explain the remaining unexplained variance within these models of

highly societally significant storms. Models of these somewhat skillful clusters might be

improved further by excluding subtropical storms that are likely explained by a modified set of

climate state variables that also captures variability in the midlatitudes.

While the analyses indicate that basin-wide TC counts or large (judiciously chosen)

subsets thereof can be more skillfully modeled than can the TC counts for the individual clusters,

the total number of TCs across the North Atlantic basin does not necessarily correlate with the

destructiveness of particular hurricane seasons. Therefore, in many applications, it is important

to know not just *how many* TCs may form in a particular season, but *what types* of TCs are likely

to form and *where* they are likely to track. There is evidence in several of the analyses of TC

cluster series of substantial unresolved, non-random structure in the residual unexplained

variability. One possible interpretation of this result is that there are yet other climatic processes

that may condition TC behavior in a given season that are not captured by even the somewhat

expansive pool of candidate predictors considered in this study. Further correction of remaining biases, both in best-track TC records and the climate predictors themselves, could potentially lead to improved results in modeling climate influences on Atlantic TC behavior, and even more so to modeling improvements for the subset of Atlantic TC having the greatest impact on the residents of that region.

References

Barnston, A.G., M. Chelliah, and S.B. Goldenberg (1997), Documentation of a highly ENSO-related SST region in the equatorial Pacific. *Atmos.-Ocean*, *35*, 367-383.

Bell, G.D., and M. Chelliah (2006), Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes. *J. Climate,19*, 590-612.

Bove, M.C., Elsner, J. B., Landsea C.W., Nui X. and J. O'Brien (1998), Effect of El Niño on U.S. Landfalling Hurricanes Revisited. *Bull. Amer. Meteor. Soc.*, *79*, 2477-2482

Camargo, S. J., and A. H. Sobel (2010), Revisiting the Influence of the Quasi-Biennial

    Oscillation on Tropical Cyclone Activity. *J. Climate*, *23*, 5810–5825.

——, A.W. Robertson, A.G. Barnston, and M. Ghil (2008), Clustering of eastern North Pacific

    tropical cyclone tracks: ENSO and MJO effects. *Geochem. Geophys. Geosys.*,*9*,

    Q06V05, doi: 10.1029/2007GC001861.

——, ——, S.J. Gaffney, P. Smyth, and M. Ghil (2007a), Cluster analysis of typhoon tracks.

    Part I: General properties. *J. Climate*, *20*, 3635– 3653.

——, ——, S.J. Gaffney, P. Smyth, and M. Ghil (2007b), Cluster analysis of typhoon tracks.

    Part II: Large-scale circulation and ENSO. *J. Climate*, *20*, 3654 – 3676.

Chand, S.S., and K.J.E. Walsh (2009), Tropical cyclone activity in the Fiji region: Spatial

    patterns and relationship to large-scale circulation. *J. Climate*, *22***,** 3877-3893.

Chang, E. K. M. and Y. Guo (2007), Is the number of North Atlantic tropical cyclones

    significantly underestimated prior to the availability of satellite observations?, *Geophys.*

    Res. Lett.,34, L14801, doi:10.1029/2007GL030169

Elsner, J. B. (2003), Tracking hurricanes. Bull. Amer. Meteor. Soc., *84*, 353-356.

——, G.S. Lehmiller, and T. B. Kimberlain (1996), Objective Classification of Atlantic

    hurricanes. *J. Climate*, *9*, 2880-2889

——, B.H. Bossak, and X.-F. Niu (2001), Secular changes in the ENSO-U.S. Hurricane

Relationship, *Geophys. Res. Lett.,28*, 4123-4126

——, K-b Liu, B. Kocher (2000a), Spatial Variations in Major U.S. Hurricane Activity:

Statistics and a Physical Mechanism. *J. Climate*, *13*, 2293–2305.

——, and T. Jagger (2006), Prediction models for annual U.S. hurricane counts. *J. Climate*, *19*,

2935-2952

——, ——, and X.–F. Nui (2000b), Changes in the rates of North Atlantic major hurricane

activity during the 20th century. *Geophys. Res. Lett.,27,* 1743-1746

Emanuel, K.A. (1995), Sensitivity of tropical cyclones to surface exchange coefficients and a

revised steady-state model incorporating eye dynamics, *J. Atmos. Sci.*, *52*, 3969-3976.

—— (2005), Increasing destructiveness of tropical cyclones over the past 30 years. *Nature, 436*,

686-688.

Evans, J.L., Fuentes J.D., Hu X. –M. and H. Hamilton (2011), Earth-Atmosphere interactions:

Tropical storm and hurricane activity in the Caribbean and their consequent health

impacts. *J. Race and Policy Special issue "Disaster Preparedness and Health Care:

U.Va in St. Kitts and Nevis"*, *7*(1).

Fink, A.H., J. M. Schrage, and S. Kotthaus (2010), On the potential causes of the non-stationary correlations between West Africa precipitation and Atlantic hurricane activity. *J. Climate*, early online.

Gaffney, S.J., A.W. Robertson, P. Smyth, S.J. Camargo and M. Ghil (2007), Probabilistic clustering of extratropical cyclones using regression mixture models. *Clim. Dyn.*, *29*, 423–440.

Giannini, A., Y. Kushnir, and M. A. Cane (2000), Interannual variability of Caribbean rainfall, ENSO and the Atlantic Ocean. *J. Clim.*, *13,* 297–311.

Giannini, A., M. A. Cane, and Y. Kushnir (2001), Interdecadal changes in the ENSO teleconnection to the Caribbean region and North Atlantic Oscillation. *J. Clim.*, *14*, 2867–2879.

Goldenberg, S. B., and L. J. Shapiro (1996), Physical mechanisms for the association of El Niño and West African rainfall with Atlantic major hurricanes. *J. Climate*, *9,* 1169–1187.

Gray, W. M. (1968), Global view of the origins of tropical disturbances and storms. *Mon. Wea.   Rev.*, *96*, 669-700.

——— (1984), Atlantic seasonal hurricane frequency, Part 1: El Niño and 30 mb quasi-biennial oscillation influences, *Mon. Wea. Rev., 115*, 1669-1668.

Hoyos, C. D., P. A. Agudelo, P.J. Webster, and J. A. Curry (2006), Deconvolution of the Factors Contributing to the Increase in Global Hurricane Intensity. *Science*, *312***,** 94-97.

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 Technical Memorandum NWS NHC 22, Coral Gables Florida, 21 pp.

Joint Institute for the Study of the Atmosphere and the Ocean, (cited 2009), "Sahel rainfall index (20-10N, 20W-10E), 1900 - May 2009," [available online at http://jisao.washington.edu/ data/sahel/]

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. Geophys. Res*.,103(C9), 18,567-18,589

Kossin, J.P., and Daniel J. Vimont (2007), A More General Framework for Understanding Atlantic Hurricane Variability and Trends, *Bull. Amer. Meteor. Soc.*, *88*, 1767-1781.

——, S. J. Camargo, and M. Sitkowski (2010), Climate modulation of North Atlantic hurricane tracks. *J. Climate*, *23*, 3057-3076.

Landsea, C.W. (2007), Counting Atlantic tropical cyclones back to 1900, *Eos Trans. AGU, 88(18)*, 197-208.

——, and W. M. Gray (1992), The strong association between western Sahel monsoon rainfall and intense Atlantic hurricanes. *J. Climate, 5,* 435–453.

Lehmiller, G.S., Kimberlain, T.B., and J.B. Elsner (1997), Seasonal prediction models for North Atlantic basin hurricane location, *Monthly Weather Review*, *125*, 1780-1791.

Mann, M.E., Emanuel, K.A. (2006), Atlantic Hurricane Trends linked to Climate Change, *Eos, 87(24)*, 233-241.

——, Sabbatelli, T.A., Neu, U. (2007), Evidence for a Modest Undercount Bias in Early Historical Atlantic Tropical Cyclone Counts, *Geophys. Res. Lett., 34*, L22707, doi:10.1029/2007GL031781.

Mestre, O. and S. Hallegatte (2009), Predictors of tropical cyclone numbers and extreme hurricane intensities over the North Atlantic using generalized additive and linear models. *J. Climate*, *22*, 633–648.

Nakamura, J., U. Lall, Y. Kushnir, and S.J. Camargo (2009), Classifying North Atlantic tropical
cyclone tracks by mass moments. *J. Climate*, *22*, 5481-5494.

NOAA/National Weather Service's National Centers for Environmental Prediction/Climate
Prediction Center (cited 2009), "El Niño Regions," [available online at
http://www.cpc.noaa.gov products/analysis_ monitoring/ensostuff/ñino_regions.shtml].

Prospero, J.M., Lamb PJ. (2003), African droughts and dust transport to the Caribbean: climate
change implications. *Science, 302*, 1024–1027.

Ramsay, H.A. and A.H. Sobel (2011), The effects of relative and absolute sea surface
temperature on tropical cyclone potential intensity using a single column model. *J.
Climate*, *24*,183-193.

——, S.J. Camargo, and D. Kim (2012), Cluster analysis of tropical cyclone tracks in the
Southern Hemisphere. *Clim. Dyn.,* early online, doi: 10.1007/s00382-011-1225-8.

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 temperatures, sea ice, and night
marine air temperature since the late nineteenth century, *J. Geophys. Res.*, *108(D14),*
4407, doi:10.1029/2002JD002670

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.

Smith, T.M., and R.W. Reynolds (2003), Extended reconstruction of global sea surface

    temperatures based on COADS data (1854-1997). *J. Clim.,16,* 1495-1510.

Solow, A. R., and N. Nicholls (1990), The relationship between the Southern Oscillation and

    tropical cyclone frequency in the Australian region, *J. Climate*, *3*, 1097– 1101.

——, and L. Moore (2000), Testing for trend in a partially incomplete hurricane record, *J.

    Climate*, *13,* 3696– 3699.

——, and —— (2002), Testing for trend in North Atlantic hurricane activity, 1900– 1998,

    J. Climate, *15*, 3111–3114.

Tippett, M.K., S.J. Camargo, and A.H. Sobel (2011), Poisson regression index for tropical

    cyclone genesis and the role of large-scale vorticity in genesis. *J. Climate*, *24,* 2335-2357.

Vecchi, G.A., K.L. Swanson, and B.J. Soden (2008), Whither Hurricane Activity?. *Science,

    322(687)*, DOI: 10.1126/science.1164396.

——, and T.R. Knutson (2008), On Estimates of Historical North Atlantic Topical Cyclone

Activity. *J. Climate*, *21,* 3580-3600.

Villarini, Gabriele, G.A. Vecchi, J.A. Smith (2010), Modeling the Dependence of Tropical Storm

Counts in the North Atlantic Basin on Climate Indices. *Mon. Wea. Rev.*, *138,* 2681–2705.

Vimont, D. J., and J. P. Kossin (2007), The Atlantic meridional mode and hurricane activity.

Geophys. Res. Lett.,*34,* L07709, DOI:10.1029/2007GL029683.

Wilks, D.S. (2005), Statistical Methods in the Atmospheric Sciences: An Introduction, 2nd

Edition, Academic Press.

Zhang, R. and T.L. Delworth (2006), Impact of Atlantic multidecadal oscillations on India/Sahel

rainfall and Atlantic hurricanes. *Geophys. Res. Lett.*,*33*, L17712.

Table and Figure Captions

Table 1. List of all state variables that are considered as candidate predictors, as explained in

section 2. The longest time interval of each predictor is also listed.

Table 2. Correlation Matrix showing the Pearson linear correlation coefficients (r) between each

of the ten candidate predictors that are listed in Table 1. In order to include all ten of the

predictors, each correlation is calculated for the shortest time interval included in this study,

1950-2007.

Table 3. Results of calibration and cross validation tests employed in the statistical modeling

exercises. The various statistics are tabulated as defined in the text. Predictors are indicated in

the order they are selected in the forward stepwise screening regression. The results of the χ^2 tests are measured with respect to the probability (α) of rejecting the relevant null hypothesis. As values of α_{fit} approach zero, the probability that such a skillful model would arise from chance alone becomes increasingly low. As values of $\alpha_{adequacy}$ approach one, the probability that residual unresolved variance is consistent with purely random Poisson process behavior becomes increasingly high.

Figure 1. Locations of the four Niño regions. The indices measured in the Niño 3, 3.4, and 1+2 regions are used in this study. Source: NOAA/NWS/NCEP/CPC 2009.

Figure 2. Time series (1878-2007) of potential predictors, indicating conditions that are favorable (red) and unfavorable (blue) for TC activity.

Figure 3. TC tracks from 1950-2007 for each of the four clusters, as separated by the cluster analysis methods detailed within the text. (adapted from Kossin et al. 2010)

**Figure 4.** Time series (1878-2007) of the primary TC predictands that are analyzed in this study. Red (blue) indicates positive (negative) TC count anomalies with respect to the mean of each individual time series. Basin-wide TC counts have been adjusted as described in text. Note the wide variation in range between the ordinate axes.

Figure 5. Statistical models for Atlantic basin-wide and cluster TC count series using the predictors specified in Table 1. Colors in each panel correspond to: TC counts (black), and the models trained on the interval 1878-2007 (red), 1900-2007 (blue), and 1950-2007 (green). Total TC counts (a) are adjusted as discussed in text.

Figures and Tables

Table 1. List of all state variables that are considered as candidate predictors, as explained in section 2. The longest time interval of each predictor is also listed.

Candidate Predictor Information	
Predictors	**Time Interval Used**
1 post-season Dec-Feb Niño3.4 index	1878-2007
2 post-season Dec-Mar NAO index	1878-2007
3 in-season Aug-Oct mean Atlantic MDR SSTs	1878-2007
4 in-season Jun-Sep Sahel precipitation index	1900-2007
5 post-season Dec-Feb Niño1+2 index	1950-2007
6 post-season Dec-Feb Niño3 index	1878-2007
7 in-season Aug-Oct relative Atlantic MDR SSTs	1878-2007
8 in-season Aug-Oct mean global Tropical SSTs	1878-2007
9 pre/in-season May-June NAO index	1878-2007
10 in-season Jun-Nov AMM index	1950-2007

Table 2. Correlation Matrix showing the Pearson linear correlation coefficients (r) between each of the ten candidate predictors that are listed in Table 1. In order to include all ten of the predictors, each correlation is calculated for the shortest time interval included in this study, 1950-2007.

Pearson Linear Correlation Coefficients (1950-2007)		ENSO Indexes			NAO Indexes		Other SST Indexes			Other Indexes	
		3.4	3	1+2	DJFM	MJ	MDR	Rel.	Trop.	AMM	Sahel
ENSO Indexes	3.4	1.00	0.97	0.69	0.10	-0.02	-0.03	-0.41	0.40	-0.26	-0.25
	3	0.97	1.00	0.77	0.14	-0.08	0.01	-0.42	0.47	-0.26	-0.29
	1+2	0.69	0.77	1.00	0.14	-0.16	0.07	-0.34	0.46	-0.17	-0.31
NAO Indexes	DJFM	0.10	0.14	0.14	1.00	0.03	-0.09	-0.30	0.20	-0.20	-0.16
	MJ	-0.02	-0.08	-0.16	0.03	1.00	-0.19	0.02	-0.28	-0.17	0.08
Other SST Indexes	MDR	-0.03	0.01	0.07	-0.09	-0.19	1.00	0.67	0.62	0.62	0.13
	Rel.	-0.41	-0.42	-0.34	-0.30	0.02	0.67	1.00	-0.16	0.84	0.64
	Trop.	0.40	0.47	0.46	0.20	-0.28	0.62	-0.16	1.00	0.17	-0.51
Other Indexes	AMM	-0.26	-0.26	-0.17	-0.20	-0.17	0.79	0.84	0.17	1.00	0.42
	Sahel	-0.25	-0.29	-0.31	-0.16	0.08	0.13	0.64	-0.51	0.42	1.00

Table 3. Results of calibration and cross validation tests employed in the statistical modeling exercises. The various statistics are tabulated as defined in the text. Predictors are indicated in the order they are selected in the forward stepwise screening regression. The results of the χ^2 tests are measured with respect to the probability (α) of rejecting the relevant null hypothesis. As values of α_{fit} approach zero, the probability that such a skillful model would arise from chance alone becomes increasingly low. As values of $\alpha_{adequacy}$ approach one, the probability that residual unresolved variance is consistent with purely random Poisson process behavior becomes increasingly high.

| Accepted Poisson Regression Model Statistics									
Predictand	Time Interval	Predictors Used	Calibration Statistics				Cross Validation Scores			
			MSE	R^2	α_{fit}	$\alpha_{adequacy}$	MSE	RE	CE	r^2
Unadjusted TC Counts	1878-2007	MDR SST, Nino 3.4, DJFM NAO	8.82	0.49	0.00	0.64	11.06	0.55	0.29	0.41
VK08 Adjusted TC Counts	1878-2007	MDR SST, Nino 3.4, DJFM NAO	9.00	0.44	0.00	0.85	9.84	0.47	0.39	0.39
	1900-2007	MDR SST, Nino 3	9.42	0.43	0.00	0.71	11.09	0.43	0.32	0.34
	1950-2007	MDR SST, Nino 3.4, MJ NAO	8.84	0.47	0.00	0.90	12.24	0.37	0.21	0.30
Cluster One TCs	1878-2007	MDR SST, DJFM NAO	2.73	0.14	0.00	0.32	3.47	0.23	-0.23	0.08
	1900-2007	MDR SST, Sahel Precip, DJFM NAO	2.68	0.19	0.00	0.39	3.29	0.29	-0.09	0.07
	1950-2007	MJ NAO, Nino 3	2.61	0.14	0.10	0.87	2.81	0.15	0.10	0.13
Cluster Two TCs	1878-2007	Nino 3.4, MDR SST	2.69	0.08	0.02	0.27	2.79	0.07	0.06	0.07
	1900-2007	MDR (relative), Nino 3.4	2.83	0.08	0.03	0.24	2.92	0.08	0.07	0.07
	1950-2007	DJFM NAO, MDR SST	2.40	0.11	0.16	0.43	2.81	0.01	0.00	0.04
Cluster Three TCs	1878-2007	Tropical SST, Nino 3.4, DJFM NAO	2.52	0.35	0.00	0.00	4.55	0.33	-0.42	0.12
	1900-2007	Tropical SST, Nino 3.4, DJFM NAO	2.66	0.37	0.00	0.00	4.97	0.21	-1.97	0.07
	1950-2007	MDR SST, Nino 1+2, MJ NAO	2.32	0.48	0.00	0.64	2.66	0.53	0.31	0.37
Cluster Four TCs	1878-2007	MDR (relative), Nino 3, MJ NAO	2.49	0.21	0.00	0.04	2.74	0.20	0.12	0.21
	1900-2007	MDR (relative), Nino 3.4, MJ NAO	2.49	0.28	0.00	0.04	2.83	0.27	0.18	0.23
	1950-2007	AMM, MDR SST, Nino 1+2	1.18	0.34	0.00	0.31	1.61	0.19	0.11	0.26
Sum of Last Three Clusters	1878-2007	MDR SST, Nino 3.4, DJFM NAO	6.73	0.43	0.00	0.43	7.56	0.45	0.34	0.36
	1900-2007	MDR SST, Nino 3.4	7.43	0.41	0.00	0.37	9.69	0.34	0.22	0.29
	1950-2007	MDR SST, Nino 3.4, DJFM NAO	5.94	0.56	0.00	0.73	9.14	0.38	0.20	0.40
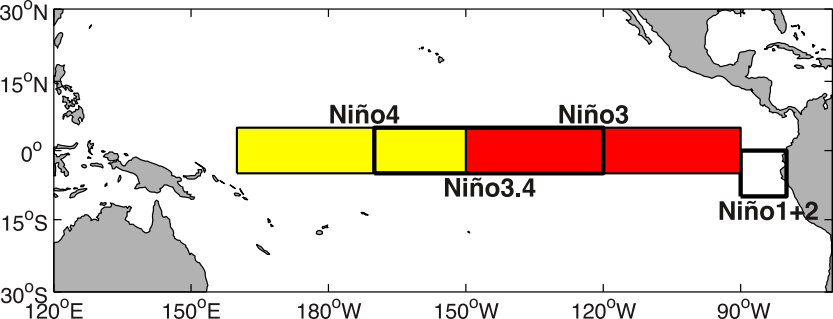
(a) MDR SST ASO
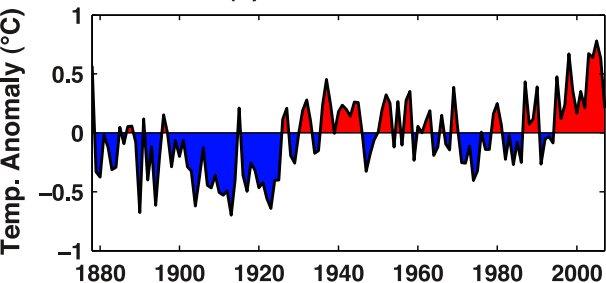
(b) Global Tropical SST ASO
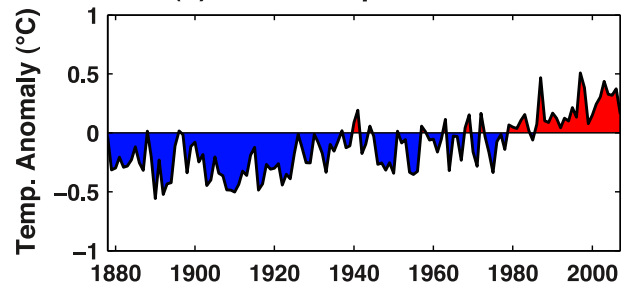
(c) Relative MDR SST ASO
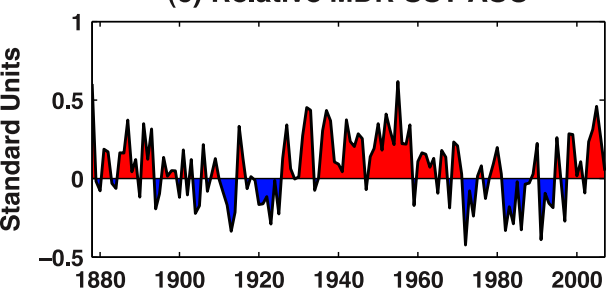
(d) NAO MJ
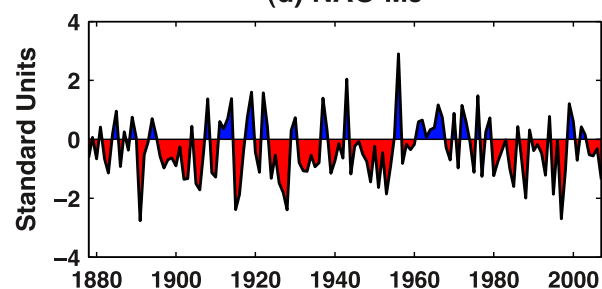
(e) NAO DJFM
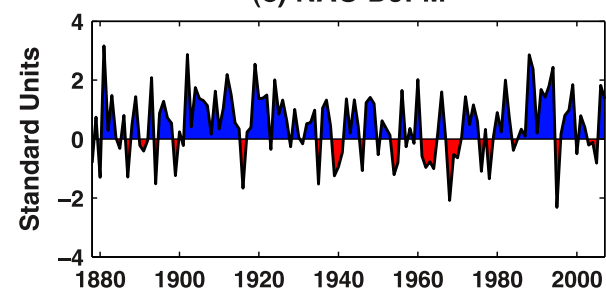
(f) Niño3.4 DJF
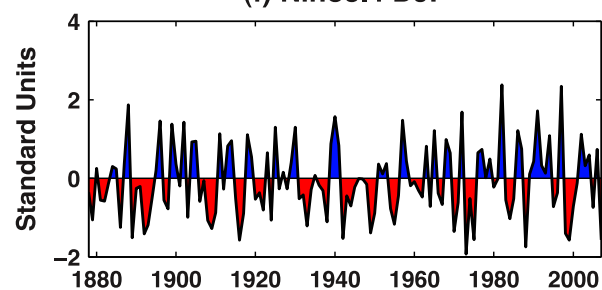
(g) Niño3 DJF
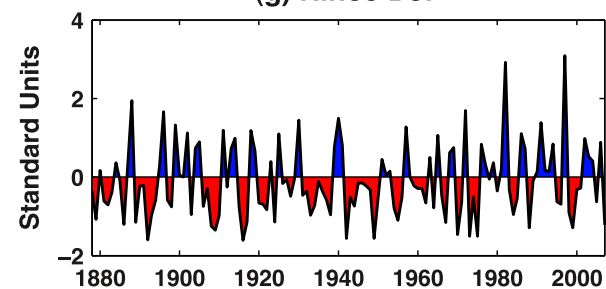
(h) Niño1+2 DJF
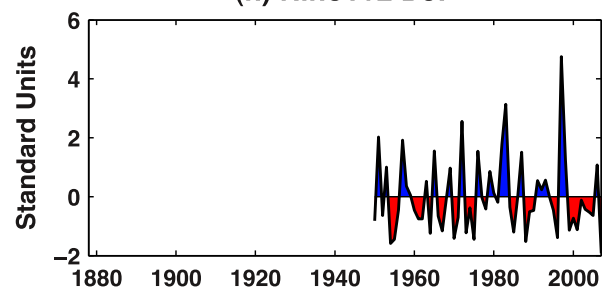
(i) Sahel Precipitation JJAS

(j) AMM JJASON

a. Cluster 1
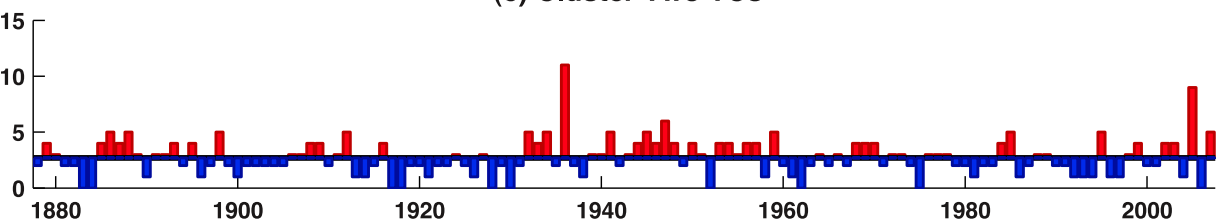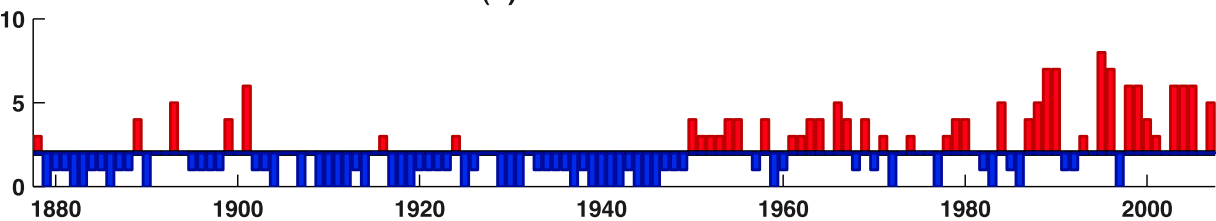b. Cluster 2
c. Cluster 3
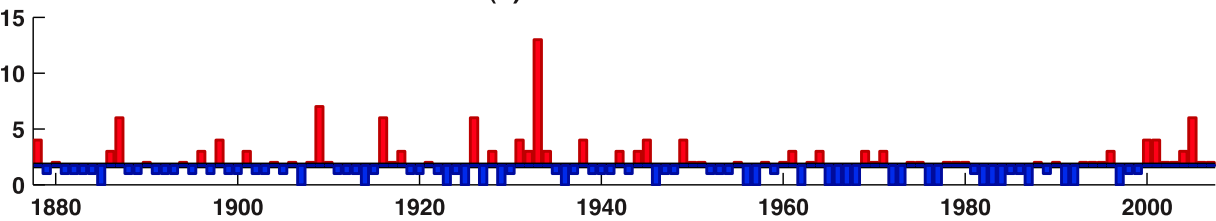d. Cluster 4
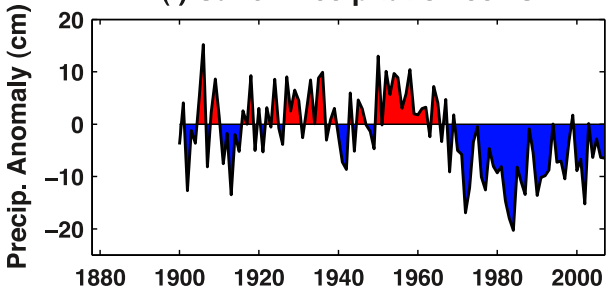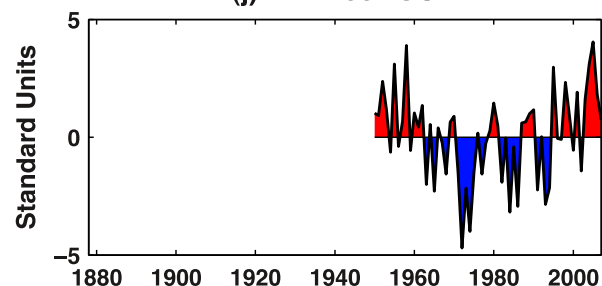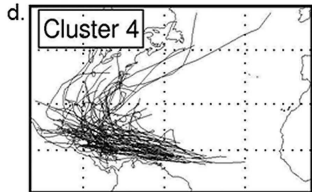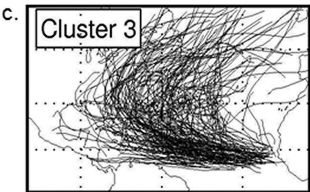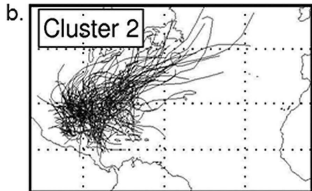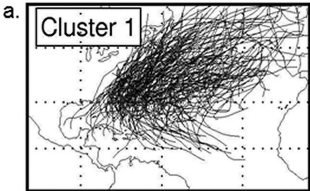
(a) VK08 TCs
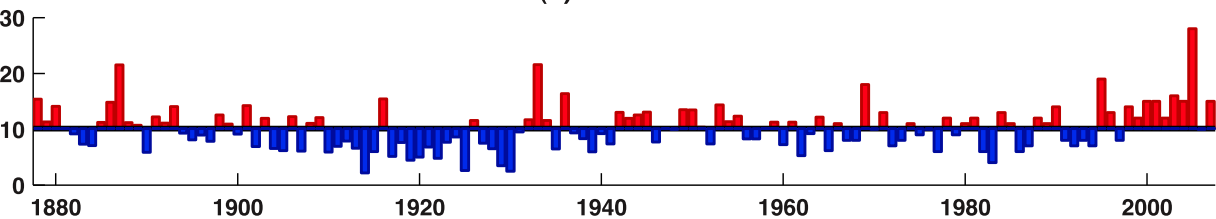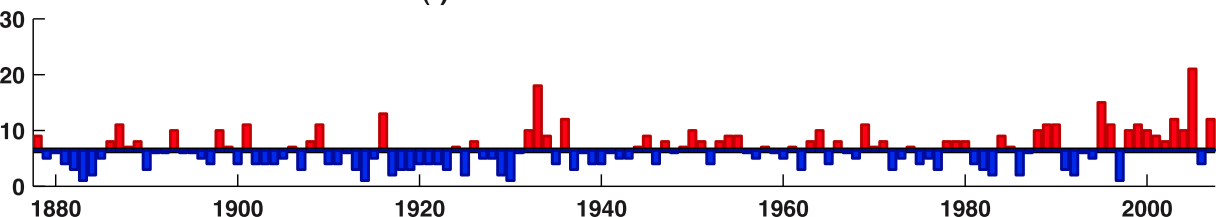
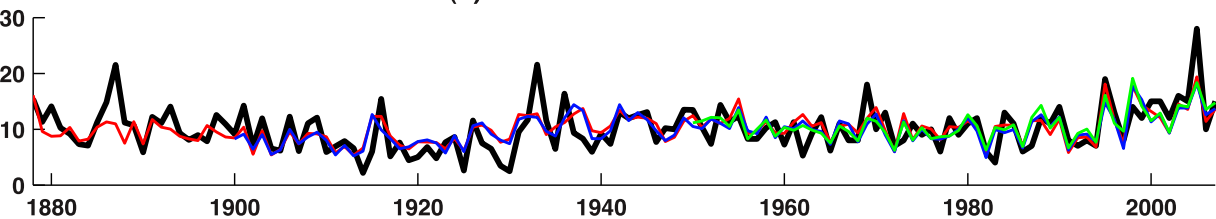**(b) Cluster One TCs**
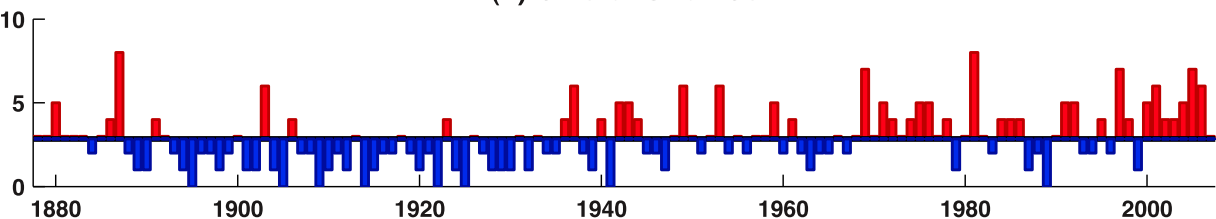
(c) Cluster Two TCs
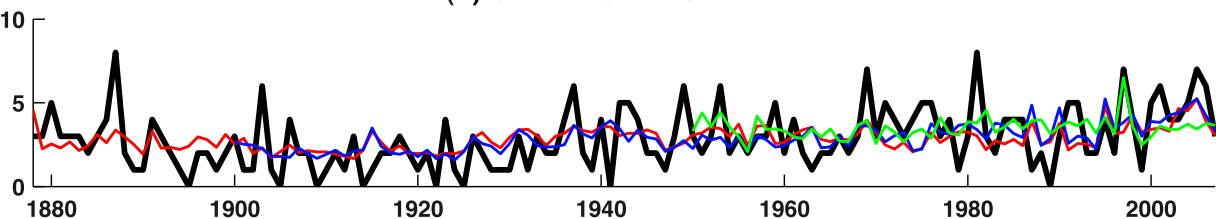
(d) Cluster Three TCs
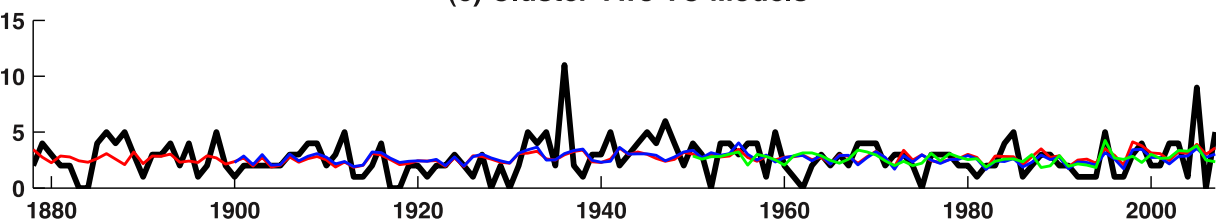
(e) Cluster Four TCs

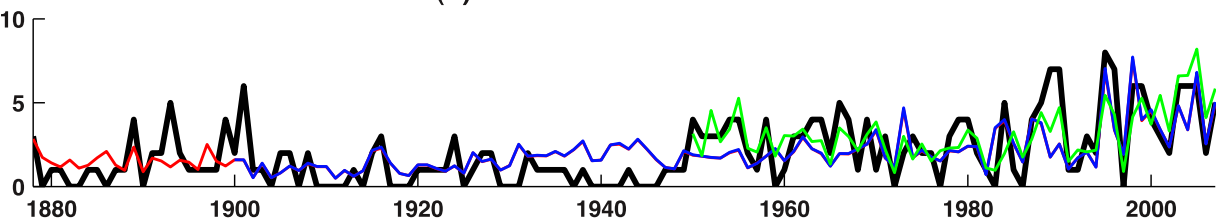**(f) Sum of TCs in Clusters 2–4**

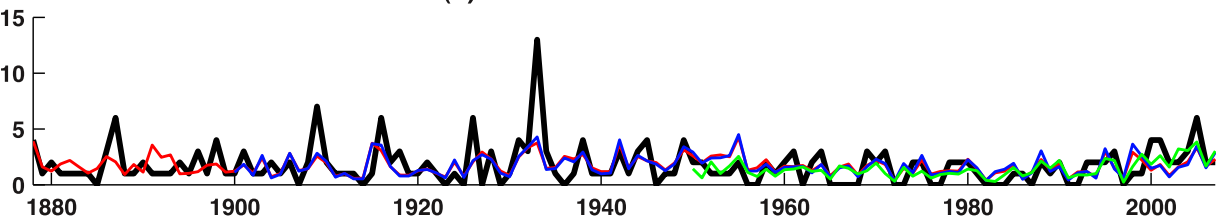**(a) Basin−wide TC Models**

(b) Cluster One TC Models

(c) Cluster Two TC Models

(d) Cluster Three TC Models

(e) Cluster Four TC Models

(f) Sum of Clusters 2−4 TC Models