2015 Atlantic Tropical Cyclone Outlook

Lian Xie[1], Marcela Alfaro-Córdoba[2], Bin Liu[1], and Montserrat Fuentes[2]

[1]Department of Marine, Earth, and Atmospheric Sciences

[2]Department of Statistics

North Carolina State University

Summary of 2015 Forecast Results

The 2015 Atlantic** hurricane season is forecast to be below to the long-term average (1951-2014) and the average of the past 20 years (1995-2014). Specific forecasts are described below***.

1. Expected number of tropical cyclones (tropical storms and hurricanes) developing in the Atlantic Basin: 4-6 (1950-2013 average: 10.8)
2. Expected number of hurricanes developing in the Atlantic basin: 1-3 (1950-2013 average: 6.2)
3. Expected number of major hurricanes developing in the Atlantic basin: 1 (1950-2013 average: 2.7)
4. Expected number of tropical cyclones in the Caribbean Sea: 0-1 (1950-2013 average: 2.6)
5. Expected number of hurricanes in the Caribbean Sea: 0-1 (1950-2013 average: 1.4)
6. Expected number of major hurricanes in the Caribbean Sea: 0 (1950-2013 average: 0.8)
7. Expected number of tropical cyclones in the Gulf of Mexico: 1-2 (1950-2013 average: 3.1)

*Range of expected values was obtained using a 95 percent prediction interval for the number of counts.

**Atlantic hurricane season starts on June 1, and ends on November 30. Atlantic basin includes the Gulf of Mexico and the Caribbean Sea.

1 Introduction

In this outlook, we aim to forecast the numbers of Tropical Cyclones (TCs), including Tropical Storms (TSs), Hurricanes (HRs), and Major Hurricanes (MHs), that form or pass through particular areas of the Atlantic Basin for the hurricane season (1 June to 30 November) of 2015. The three regions include the Atlantic Ocean basin, and the two sub-regions, the Caribbean Sea and the Gulf of Mexico.

A statistical log-linear regression model was utilized to achieve this goal. In this year's outlook, we expanded the candidate predictors by including more months from climate indices (August from previous year until February from the same year). In addition, a variable selection method was employed to select the relevant predictors from a large set of available candidate predictors. Details regarding the data, model, and results are given in the following sections.

2 Data

The historical TC counts were obtained by manually counting them based on the National Hurricane Center (NHC) HURDAT best track data map available at: http://www.nhc.noaa.gov/pastall.shtml.

TC counts are determined by region and then further categorized by the peak intensity within each region according to the Saffir-Simpson hurricane wind scale. The three regions are the Atlantic Ocean, the Caribbean Sea, and the Gulf of Mexico. The Atlantic Ocean regions (ATL) stands for the whole Atlantic TC basin, the Caribbean Sea region (CAR) is in enclosed by the West Indies and the coast of Central America from the East coast of the Yucatan Peninsula to Venezuela, and the Gulf of Mexico region (GOM) is bordered by the Gulf Coast of the United States (from the Southern tip of Florida to Texas) to Mexico and the Northern edge of the Yucatan Peninsula and the Northwestern coast of Cuba. This best track data for the Atlantic basin are available since 1851. However, only data after 1955 were used in building the statistical model here because of the large uncertainties in the earlier data. Subtropical and extra-tropical storms, as well as storms outside of the 1 June–31 November hurricane season were filtered out.

Various climate factors were utilized as candidate predictors to assist in the prediction of TC counts for an upcoming hurricane season. These candidate predictors include SST-related climate indices, El Nino–Southern Oscillation (ENSO) related indices, atmospheric and teleconnection indices, as well as parameters in the Main Development Region (MDR) for Atlantic hurricanes.

AMM (Atlantic Meridional Mode): The result of a maximum covariance analysis of SSTs and the zonal and meridional winds over the region 21S–32N, 74W–15E. This index is obtained from http://www.esrl.noaa.gov/psd/data/climateindices/list/.

AMO (Atlantic Multi-decadal Oscillation): An index based on North Atlantic SSTs. This index is obtained from http://www.esrl.noaa.gov/psd/data/climateindices/list/.

TNA (Tropical Northern Atlantic): Anomaly of the average of the monthly SST from 5.5N to 23.5N and 15W to 57.5W. This index is obtained from http://www.esrl.noaa.gov/psd/data/climateindices/list/.

TSA (Tropical Southern Atlantic): Anomaly of the average of the monthly SST from 0-20S and 10E-30W. This index is obtained from http://www.esrl.noaa.gov/psd/data/climateindices/list/.

DM (Atlantic Dipole Mode): The difference between TNA and TSA SSTs.

WHWP (Western Hemisphere Warm Pool): Monthly anomaly of the ocean surface area warmer than 28.5 ℃ in the Atlantic and eastern North Pacific. This index is obtained from http://www.esrl.noaa.gov/psd/data/climateindices/list/.

GGST, NGST, and SGST: Global, north-hemisphere, and south-hemisphere monthly Land-Surface Air and Sea-Surface Water Temperature Anomalies based on the GISS Surface Temperature Analysis (GISTEMP). Data is obtained from http://data.giss.nasa.gov/gistemp/.

NINO12: SST index in the Nino1+2 region. This index is obtained from http://www.esrl.noaa.gov/psd/data/climateindices/list/. To represent ENSO impacts, NINO12 values for the July/August/September average during the hurricane season were used in building the model. And the forecast values for NINO12 obtained from the National Center for Environmental Protection (NCEP) coupled forecast system (CFS) model were used for forecasts of the upcoming hurricane season from 1 June–31 November 2015.

SOI (Southern Oscillation Index): This index is obtained from http://www.esrl.noaa.gov/psd/data/climateindices/list/.

NAO (Northern Atlantic Oscillation): This index is obtained from http://www.esrl.noaa.gov/psd/data/climateindices/list/.

EPO (East Pacific/North Pacific Oscillation): This index is obtained from http://www.esrl.noaa.gov/psd/data/climateindices/list/.

PNA (Pacific North American Index): This index is obtained from http://www.esrl.noaa.gov/psd/data/climateindices/list/.

AO (Arctic Oscillation): This index is obtained from http://www.cpc.ncep.noaa.gov/products/precip/CWlink/daily_ao_index/ao.shtml

QBO (Quasi-Biennial Oscillation): This index is obtained from http://www.esrl.noaa.gov/psd/data/climateindices/list/.

SFI (Solar Flux): observed monthly solar flux (10.7cm) at Ottawa/Penticton. Data available from http://www.esrl.noaa.gov/psd/data/climateindices/list/.

MDRSLP, MDRSST, MDRVWS: Monthly averaged sea-level pressure (MDRSLP), SST (MDRSST), and vertical wind shear (MDRVWS, wind shear between 200 and 850 mb), over the main development region (MDR, 10-20N, 80-20W). Data derived from the NCEP/NCAR

Reanalysis I dataset by http://www.esrl.noaa.gov/psd/cgi-bin/data/timeseries/timeseries1.pl.

3 Methods

Our goal is to estimate the expected TS, HR and MH counts (predictands) to form in a particular area, represented by λ_i for i=1,…,9. Forecasts are made for TS, HR and MH counts for each region as listed in Table 1. TS includes tropical storms, hurricanes (categories 1-2), and major hurricanes (categories 3-5); HR includes hurricanes and major hurricanes; MH includes major hurricanes only.

Table 1**.** Predictands including the TS, HR and MH counts for each region during the forecast hurricane season.

Region	Category	Predictand
ATL	TS	λ_1
	HR	λ_2
	MH	λ_3
CAR	TS	λ_4
	HR	λ_5
	MH	λ_6
GOM	TS	λ_7
	HR	λ_8
	MH	λ_9

The statistical model of log-linear regression is used, which assumes the log(λ_i) to be linearly related to the selected climate indices. Two different groups of models are evaluated based on an analysis of the quality of hindcast predictions using only data up to the hindcasted year. The first group is constructed using the January and February values for the indices using best predictive ability. Once the months for each index are selected, the averages of the values are calculated to create a single representative value from each index. The second group consists of different combinations of values for the indices from months ranging from August until February.

Before implementing the regression in all models, we apply a variable selection technique to identify the combination of indices that has the best predictive ability for each region and strength category. Many of the indices show correlation to each other, so we use the Lasso methodology, which allows for variable selection in the presence of correlation among the predictors. Table 2 shows the selected predictor indices for each category.

Table 2. Selected predictor indices and months for each predictand.

Region	Category	Selected predictors
ATL	TS	AMO (Sept, Nov), MDROLR (Nov), QPO (Aug), TNA (Ago, Oct, Nov, Jan), TSA (Nov, Feb), WHWP (Ago, Oct).
	HR	AMO (Nov, Feb), MDRSLP (Jan), MDRVWS (Dec), TNA (Ago, Jan), WHWP (Ago, Nov)
	MH	AMM (Jan), AMO (Feb), MDRVWS (Sept), TSA (Ago)
CAR	TS	AMO (Aug, Nov), MDRSST (Jan), WHWP (Oct, Nov)
	HR	AMO (Aug), MDRSLP (Aug, Feb), TNA (Dec), WHWP (Nov)
	MH	AMO (Nov, Jan, Feb), NAO (Aug, Sept)
GOM	TS	AMO (Jan-Feb), TSA (Jan-Feb)

An advantageous aspect of the Lasso methodology is that the contribution of each climate index on the predicted counts is easily interpretable compared to dimension reduction techniques such as principle component analysis (PCA). Once the predictors are selected, the log of our response is modeled as

$$\log(\lambda) = \beta_0 + \sum \beta_k \, (Index_k) + \varepsilon,$$

where β_0 is the intercept; β_k and $Index_k$ are the regression coefficients and the selected indices, respectively, which are specific to each region and intensity category; and ε is the random error. Using the data from previous years, the coefficients are estimated using maximum likelihood methods. With these estimates, we then use the current climate index values (listed in Table 3) to predict the values of $\log(\lambda_i)$ for year 2015. All values are normalized based on the 1955 – 2015 values.

Table 3. Values of climate indices used in forecast of year 2015

Predictor	Months Averaged Over	Value
AMM	January 2015	-0.40
AMO	January and February 2015	0.36
AMO	February 2015	0.30
AMO	August 2014	1.35
AMO	September 2014	1.33

AMO	November 2014	0.65
AMO	January 2015	0.42
AMO	February 2015	0.30
MDROLR	November 2014	-0.64
MDRSLP	January 2015	1.51
MDRSLP	August 2014	0.53
MDRSLP	February 2015	0.59
MDRSST	January 2015	0.58
MDRVWS	September 2014	1.18
MDRVWS	December 2014	1.04
NAO	August 2014	-1.54
NAO	September 2014	1.54
QBO	August 2014	-1.12
TNA	August 2014	0.39
TNA	October 2014	1.30
TNA	November 2014	0.77
TNA	December 2014	0.72
TNA	January 2015	0.31
TSA	August 2014	0.57
TSA	November 2014	-1.01
TSA	February 2015	0.98
TSA	January and February 2015	1.05
WHWP	August 2014	1.44
WHWP	October 2014	1.65
WHWP	November 2014	1.43

4 Model validation

We have conducted hindcast (re-forecast) analyses to evaluate the predictive ability of our models. The same set of predictors selected by Lasso for the entire dataset are used, and the coefficients are determined using the same methodology as described in Section 3, but using only those predictors and the data up to the year before the predicted year. So, for example, the prediction for the year 2002 uses only the data up to year 2001 to determine the coefficients.

Yearly observations were classified in three categories, according to the average number of storms in the respective region and level of activity: normal, above normal and below normal. Table 4 shows the definition of each category.

Table 4. Description for hit rate validation categories

Category	Description
Below Normal	All observations that are less than the smallest observation included in the normal category
Normal	Average number of storms (TC, HU or MH) plus – minus 30% of the data. Average was defined in two ways: last 20 years average and 1950 – 2014 average (long-term).
Above Normal	All observations that are greater than the biggest observation included in the normal category

Predictions were classified using the same categories, and then a hit was defined as the event where prediction and observation were in the same category for a specific year. A hit rate validation was implemented and models with more than 75% of hit rate using either the 20-year average or the long-term average were selected. Table 5 shows the hit rate for each region and intensity category.

Table 5. Hit rate percentage for each region and activity category. Forecast categories in bold are used to create outlook summary

Region	Category	20-year	Long-term
ATL	TS	**83%**	**93%**
	HR	**77%**	**97%**
	MH	63%	63%
CAR	TS	**74%**	**90%**
	HR	67%	**90%**
	MH	**93%**	**93%**
GOM	TS	**70%**	**86%**

5 Results for 2015

The values of λ are calculated using the methodology described in Section 3. Using the normal approximation, we also create 95% prediction intervals for the $\log(\lambda_i)$. We transform back to the scale of the data by exponentiating the estimates and bounds. This does result in non-symmetric prediction limits, but is still a 95% interval since the exponential function is monotone. The estimated number of storms (λ_i) and the lower and upper bounds are listed in Table 6.

Table 6**.** Estimates and 95% prediction intervals for the expected TC counts for each region and intensity category.

Region	Category	Forecasts for year 2015		
		Lower Limit	Estimate	Upper Limit
ATL	TS	3.4	4.7	6.1
	HR	1.7	2.4	3.0
	MH	0.8	1.0	1.3
CAR	TS	0.6	0.9	1.2
	HR	0.4	0.6	0.9
	MH	0.0	0.2	0.3
GOM	TS	1.0	1.3	1.6
	HR			
	MH			

Acknowledgments. We thank Dorit Hammerling and Danny Modlin for sharing the data and experience from conducting the forecast in previous years.