A New Statistical Model for Predicting Seasonal North Atlantic Hurricane Activity

KYLE DAVIS, XUBIN ZENG, AND ELIZABETH A. RITCHIE

Department of Atmospheric Sciences, The University of Arizona, Tucson, Arizona

(Manuscript received 25 November 2014, in final form 6 March 2015)

ABSTRACT

Statistical, dynamical, and statistical–dynamical hybrid models have been developed in past decades for the seasonal prediction of North Atlantic hurricane numbers. These models' prediction skills show considerable decadal variability, with particularly poor performance in the past few years. Here, environmental factors that affect hurricane activities are reevaluated to develop a new statistical model for seasonal prediction by 1 June of each year. The predictors include the April–May multivariate ENSO index (MEI) conditioned upon the Atlantic multidecadal oscillation (AMO) index, the $3/2$ power of the average zonal pseudo–wind stress across the North Atlantic in May, and the average March–May tropical Atlantic sea surface temperature. When compared to the actual number of hurricanes each year from 1950 to 2013, this model has a root-mean-square error (RMSE) of 1.91 with a correlation coefficient of 0.71. It shows a 39% improvement in RMSE over a no-skill metric (based on the 5-yr running mean of seasonal hurricane counts) for the period 2001–13. It also outperforms three statistical–dynamical hybrid models [CPC, Colorado State University (CSU), and Tropical Storm Risk (TSR)] by more than 25% for the same period. Furthermore, two approaches are developed to provide the uncertainty ranges around the predicted (deterministic) hurricane number per season that better encompass the range of uncertainty than does the standard method of adding/subtracting a standard deviation of the errors.

1. Introduction

Of all hazards afflicting the United States, including both human caused and natural, Atlantic hurricanes are among the most damaging. These storms are defined as tropical cyclones in the North Atlantic basin (including the Gulf of Mexico and Caribbean Sea) whose maximum sustained wind speeds exceed 63 knots (kt; 1 kt = 0.51 m s^{-1}). From 1970 to 2002, it is estimated that these hurricanes cost the United States, in 2002 values, $44 billion in damage (Zanetti et al. 2003, 34–35; Murnane 2004). That figure is much larger than the $17 billion caused by earthquakes and the $24 billion in human-caused disasters, which includes the events of 11 September 2001. Insurance firms, in particular, need to monitor hurricane activity, as expensive natural disasters can put great strain on the industry. For example, the period 1991–94 was notable for costly natural disasters, as nine smaller insurance companies became insolvent (Changnon et al. 1997). Among these disasters is Hurricane Andrew, the most financially damaging tropical cyclone up to that point (Zanetti et al. 2003). More recently, both Hurricanes Katrina and Sandy have produced enormous insured losses. When comparing all three storms in terms of 2013 values, insured losses were estimated at about $26 billion for Andrew, at $49 billion for Katrina (http://www.iii.org/sites/default/files/docs/pdf/hurricane_sandy_fact_file_2014.pdf), and at $35 billion for Sandy (Bevere et al. 2013). Firms can be better prepared to deal with the aftermath of disastrous hurricanes by having better Atlantic hurricane activity predictions. Since insurance companies typically enter into their reinsurance contracts long before the start of hurricane season, they prefer to have hurricane forecasts by 1 January of each year, although skilled forecasts at any lead time could prove potentially useful (Murnane 2004).

Since 1984, Colorado State University (CSU) has produced Atlantic basin seasonal tropical activity forecasts, including the number of hurricanes (http://hurricane.atmos.colostate.edu/). More recently, other groups have also produced seasonal Atlantic hurricane forecasts, which employ a wide variety of methods. Examples include the linear statistical model employed

Corresponding author address: Kyle Davis, Dept. of Atmospheric Sciences, Rm. 542, The University of Arizona, P.O. Box 210081, 1118 E 4th St., Tucson, AZ 85721-0081.
E-mail: davis7000@gmail.com

DOI: 10.1175/WAF-D-14-00156.1

by Gray et al. (1992) and the nonlinear statistical model used by Elsner and Schmertmann (1993). A dynamical model was used by Vitart et al. (2007), while hybrid models were used by Vecchi et al. (2013) and Kim and Webster (2010). Now, several private companies, government agencies, and universities produce predictions of Atlantic hurricane numbers each year, both before the season starts (including at the conclusion of the previous year's hurricane season) and partway through the season.

Seasonal hurricane forecasting receives considerable media attention prior to the Atlantic hurricane season, and there is considerable criticism when the forecasts are not accurate. Though the number of hurricanes is not necessarily the best way to classify the activity of a tropical cyclone season [a better metric might be Accumulated Cyclone Energy (Bell et al. 2000)], hurricanes receive considerable attention as they are the most destructive component to the season. A blog by the American Meteorological Society (AMS) mentioned a series of failed seasonal hurricane forecasts in recent years by the NOAA Climate Prediction Center (CPC), Colorado State University (CSU), and Tropical Storm Risk (TSR) (http://blog.ametsoc.org/news/time-to-heed-the-hurricane-season-forecast/). For example, 2012 was predicted to be a quiet year, yet it was active (though the tropics were quiet as predicted, the extratropics had highly anomalous activities). The 2013 season was predicted to be an active year, but the season was actually very quiet. Vecchi and Villarini (2014) argued the importance of reviewing inaccurate seasonal forecasts to better understand the physical factors that are important but not accounted for in the prediction. Thus CSU, CPC, and TSR issue end-of-season verifications that review how the predictions compared to what was observed and how the atmospheric conditions affected what actually happened. Furthermore, at some point, seasonal forecasts may simply fail because there are limits to the predictability of the climate system, which is nonlinear. Thus, relationships may work well over a certain period of time but will, eventually, fail. Nevertheless, forecasts have shown improvements over climatology, but more skill is needed further in advance of the hurricane season. Forecasts in August (already 2 months into the season) have shown skill compared with a 5-yr running average climatology, and these predictions are still useful as 90% of Atlantic tropical cyclone activity occurs after 1 August (Gray et al. 1993). However, forecasts made before the start of hurricane season have shown little skill (Blake et al. 2010), with improvements ranging between 9% and 20% for mean absolute error (MAE) and between 10% and 17% for root-mean-square error (RMSE) for the period 2001–13 over the 5-yr running average climatology (see Fig. 4, described in greater detail below).

Owing to the decadal variability and predictable long-range signals of hurricanes, the purpose of this study is to develop a new statistical model (henceforth referred to as the University of Arizona, or simply UA, model) to forecast seasonal Atlantic hurricane activity by the start of the hurricane season (1 June) each year that significantly improves the MAE and RMSE over a 5-yr running mean climatology. This research considers new predictors as well as commonly used ones in different ways to produce a unique and innovative model. In particular, because of the observation that sea surface temperatures in the Pacific Ocean can be more or less correlated with the number of hurricanes depending on the phase of the Atlantic multidecadal oscillation (AMO) (see section 2), this study uses the multivariate El Niño–Southern Oscillation (ENSO) index (MEI) conditioned on the AMO as a predictor. The performance of the UA model will be compared against existing models for seasonal hurricane predictions.

2. Methodology

Data for the number of Atlantic hurricanes per season are obtained from the Hurricane Research Division of the Atlantic Oceanographic and Meteorological Laboratory (AOML) (http://www.aoml.noaa.gov/hrd/tcfaq/E11.html). Only hurricanes from 1950 to 2013 are considered. In the presatellite era (before 1966), the possibility exists of missed hurricanes in the record. Vecchi and Knutson (2011) estimated that from 1950 to 1965, the average number of hurricanes missed per year was close to one in 1950 and was one-half in the early 1960s. Because of the relatively small number of predicted missed hurricanes each year and the uncertainty associated with it, no adjustments will be made to the data.

As mentioned, the UA scheme is designed to produce its prediction by 1 June as the inputs to the model do not require data measured after 31 May. However, it is possible that there may be a short lag of a few days between when the data recording finishes and when the required input data are published by the responsible organization. However, this delay in acquiring the input data would cause only a short delay in producing the prediction and is considered inconsequential, as nearly all hurricane activity happens later in the season (Gray et al. 1993).

The UA model uses a Poisson regression (i.e., the regression between the logarithm of the number of hurricanes and various predictors), following previous studies (Elsner and Schmertmann 1993; Elsner and Jagger 2006). This model ensures that the predicted

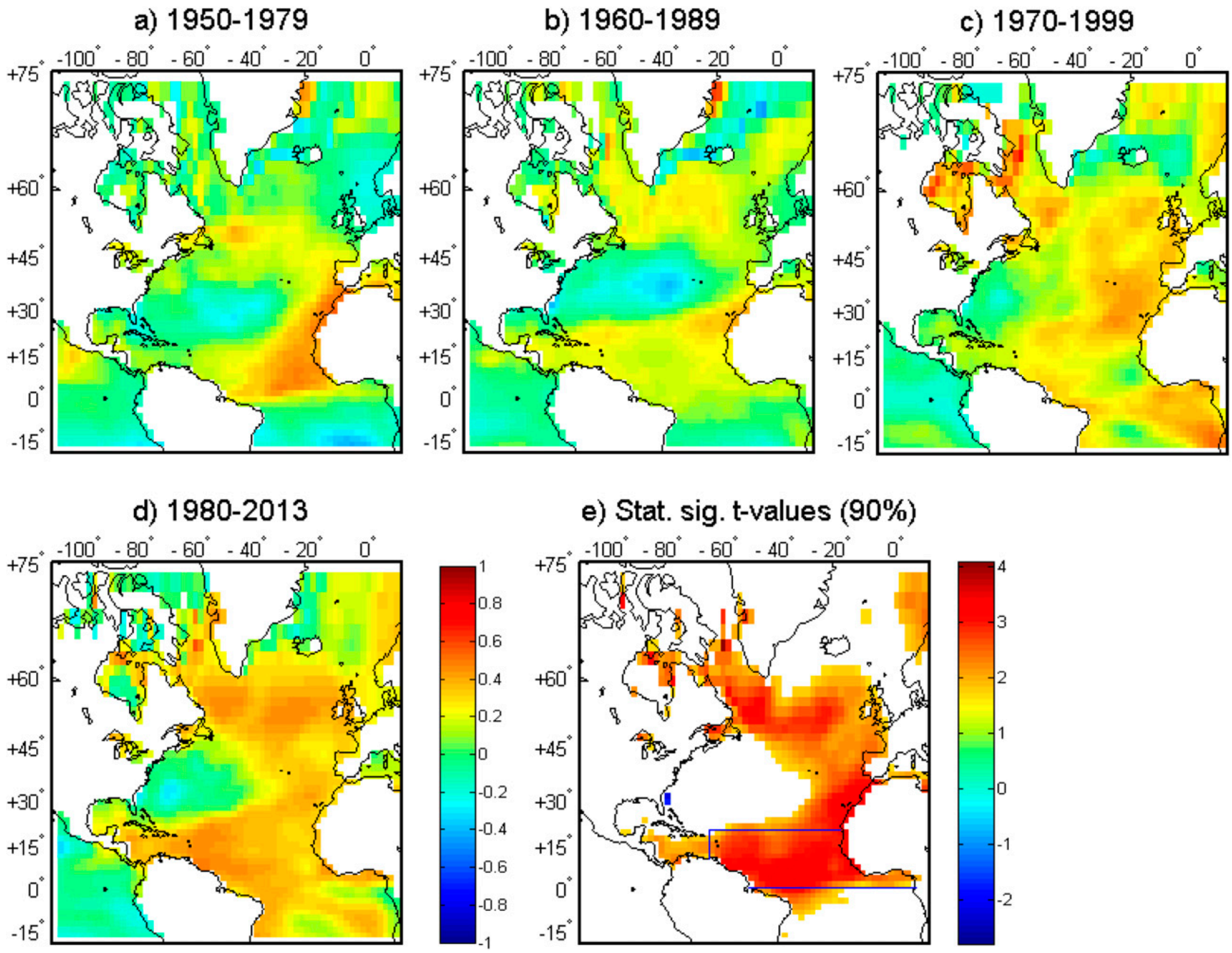

FIG. 1. (a)–(d) Correlations of the MAM SSTs with the number of hurricanes for several approximately 30-yr periods and (e) the Student's t test values that are statistically significant at the 90% or higher level for the whole 64-yr period (without considering the year-to-year autocorrelation of the time series). In (e), the box denotes the region (0°–20°N, 64°W–10°E) used for averaging MAM SSTs.

number of hurricanes is greater than zero, which is an advantage over using a linear regression.

The remainder of this section will describe the predictors and then finally present the model. The three predictors specify conditions for Atlantic sea surface temperatures (SSTs), the MEI conditioned upon the AMO (denoted as MEI_AMO), and zonal pseudo–wind stress (denoted as PWS).

a. SSTs

High SSTs are very important for hurricane formation (Palmen 1948). The average of the March–May (MAM) SSTs used for this model is taken from the Extended Reconstructed SST, version 3b (ERSST.v3b; Smith et al. 2008; Xue et al. 2003). Figure 1 shows the correlation of SSTs with the number of North Atlantic hurricanes for several approximately 30-yr periods, starting with 1950–79. The correlations are positive in every period between northeastern South America and western Africa where easterly waves can develop into hurricanes. This area of high correlation often extends around the North Atlantic in a crescent shape to southeastern Canada. In the middle of the aforementioned crescent is a persistent area of negative correlation that typically lies near the eastern coast of the United States. The results resemble the AMO pattern as given in Goldenberg et al. (2001), or the first rotated EOF of non-ENSO global SST variability.

The results from Fig. 1 show the complexity of the North Atlantic SSTs in relation to hurricane development, even in regions that do not directly influence hurricane activity. Smirnov and Vimont (2012) showed that SSTs in the subtropical Atlantic propagate, during peak hurricane season, into the tropical Atlantic.

Figure 1 shows that MAM SSTs in the high-latitude North Atlantic that have no direct influence on hurricane development correlate relatively highly with the number of hurricanes, which may reflect the Atlantic meridional mode (AMM; Chiang and Vimont 2004). However, the SSTs in the middle North Atlantic are frequently negatively correlated with the number of hurricanes. Figure 1 also shows that these correlations change from period to period: locations that are positively correlated in one period may be negatively correlated in another. For example, between 1970 and 1999, SSTs across most of the North Atlantic were positively correlated with the number of hurricanes except for a small area just off of the eastern coast of the United States, but from 1960 to 1989 the negatively correlated area extended much farther east. Thus, the correlation does not stay constant from period to period in each region.

The fact that the correlation strength and sign change for a region from decade to decade makes choosing the right area for hurricane forecasting more difficult. However, the region where SSTs are most consistently correlated with the number of hurricanes lies in 0°–20°N, 64°W–10°E (indicated by the box in Fig. 1e), which is along the track that easterly waves follow as they make their way from North Africa. This area is contained within the region defined for the AMM, consistent with the results of Vimont and Kossin (2007). They found that during the positive phase of the AMM, the conditions are more conducive to tropical cyclone activity, because of warmer SSTs, lower sea level pressures, and less vertical wind shear.

The average MAM SSTs of this area have a moderate correlation with the number of hurricanes each year (0.41) from 1950 to 2013. To ensure that the UA model uses the most appropriate predictor, SSTs for individual months were also compared with the MAM SSTs. Using the average SSTs for any of the MAM months individually yielded MAE values all within 0.11 of each other, but using the average of all MAM months together yielded the most predictive power by minimizing the MAE.

b. ENSO

Another common predictor is ENSO. There are fewer hurricanes during El Niño years than La Niña years because the vertical wind shear over the Caribbean and equatorial Atlantic is much higher during El Niño conditions (Gray 1984), as well as generating increased static stability over the hurricane main development region (Tang and Neelin 2004). One index used to represent this oscillation is the MEI, calculated as the first principal component of six variables over the tropical Pacific: sea level pressure, zonal and meridional components of the surface wind, sea surface temperature, surface air temperature, and cloud fraction (Wolter and Timlin 1993). The MEI was chosen over other ENSO indices, such as Niño-3.4, because it reduced both the MAE and RMSE the most in the model. The April–May value of the MEI is used because it is the latest value available if a seasonal forecast is to be made by 1 June. Overall, the April–May value shows a weak negative correlation with the number of hurricanes (−0.10) between 1950 and 2013; however, this correlation changes through time. Figures 2a and 2b show that there are periods when the MEI and the number of hurricanes are positively correlated (1955–80 and from 2000 to present), but also years when they are negatively correlated (1950–55 and 1980–2000). The 10-yr running correlation between April–May MEI and the number of hurricanes clearly shows this multidecadal variation (Fig. 2b). Hence, MEI is not strongly correlated for the entire period from 1950 to 2013 because its correlation with the number of hurricanes is sometimes positive and sometimes negative.

To make use of this multidecadal variation to predict seasonal Atlantic hurricane activity in the UA model, a condition that tempers or magnifies ENSO effects on hurricanes must be found. Klotzbach (2011) showed that the AMO, which measures the North Atlantic SSTs and has its own multidecadal oscillation, can have this effect because background conditions are more favorable for hurricane formation with a positive AMO. Taking only the years when the May unsmoothed AMO [http://www.esrl.noaa.gov/psd/data/timeseries/AMO/; Enfield et al. (2001) for the smoothed version] is positive, the number of hurricanes and April–May MEI are weakly and positively correlated (0.13). When the May AMO is negative, however, there is a stronger negative correlation between April–May MEI and the number of hurricanes (−0.46). Hence, to incorporate these effects into our model, whenever the May AMO is positive, MEI_AMO is set to zero, and when the May AMO is negative, MEI_AMO is set to the April–May MEI value. This adjustment increases the negative correlation between MEI and the number of hurricanes from −0.10 to −0.26 for the period 1950–2013.

c. Zonal pseudo–wind stress

Zonal pseudo–wind stress is defined as the magnitude of the wind multiplied by the wind vector in the zonal direction and is proportional to the surface wind stress (Smith et al. 2004). It is used here for two reasons: it affects the horizontal and vertical distribution of ocean temperature, and it is also highly correlated with sea level pressure—and low sea level pressure is needed for hurricane development (Knaff 1997). For instance, using the NCEP–NCAR reanalyses (Kalnay et al. 1996), the average sea level pressure and average zonal

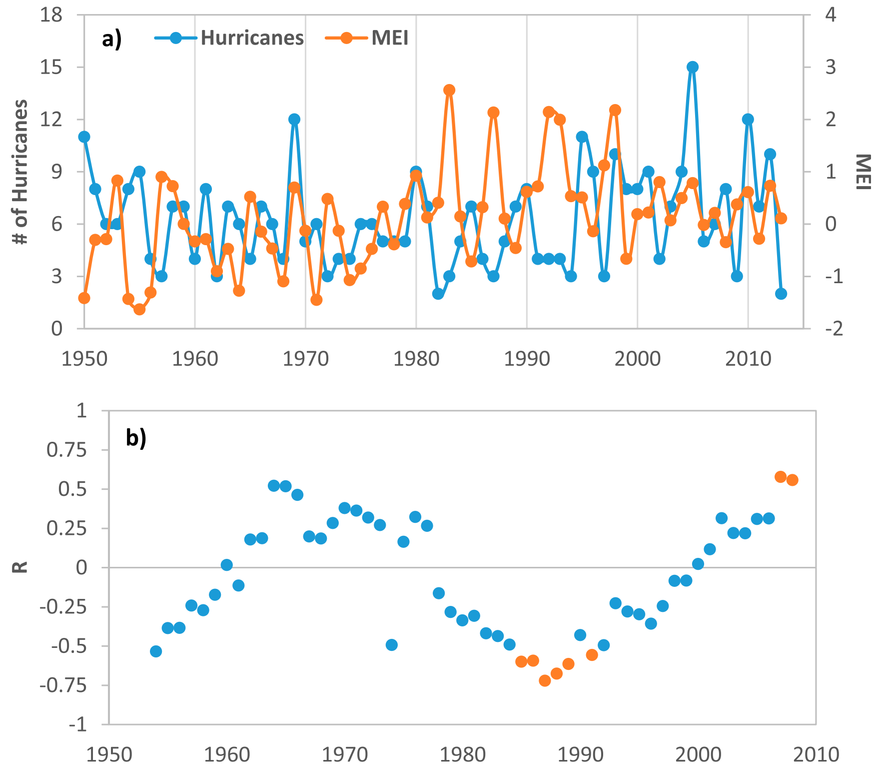

FIG. 2. (a) The MEI and number of hurricanes and (b) their running 10-yr correlations using all data from 1950 to 2013 (correlation computed by taking the 5 yr after and the 4 yr before the year shown). The brown-colored points are statistically significant at the 90% level.

pseudo–wind stress have a correlation coefficient of −0.66 over the domain in Fig. 3 for the month of May.

Figure 3 shows the spatial correlation between zonal pseudo–wind stress and the number of hurricanes for 1950–2013 using the International Comprehensive Ocean–Atmosphere Data Set (ICOADS; http://www.esrl.noaa.gov/psd/) 2° zonal pseudo–wind stress data for May. The area of greatest correlation lies in 18°–39°N, 4°–98°W, and this area (indicated by the box in Fig. 3) is used for our model.

However, when used as a predictor, the zonal pseudo–wind stress is raised to the 3/2 power because the turbulent dissipation rate in the ocean mixed layer is proportional to the 3/2 power of wind stress (Kraus and Businger 1994). Since zonal pseudo–wind stress can be positive or negative, this is computed by multiplying the absolute value of the zonal pseudo–wind stress raised to the 3/2 power by the original sign of the zonal pseudo–wind stress. The correlation between zonal pseudo–wind stress and the number of hurricanes increases from 0.29 to 0.31 when raised to the 3/2 power. Also, with just the zonal pseudo–wind stress itself for PWS in Eq. (1), MAE is 1.56 and RMSE is 1.99 from 1950 to 2013. However, when the stress is raised to the 3/2 power for PWS, MAE and RMSE decrease to 1.48 and 1.91, respectively.

d. *Functional form of the UA model*

With all variables defined as above, the UA model for the period 1950–2013 is

$$\text{Hurricane No.} = \exp(-14.86 + 0.65\text{SST} - 0.23\text{MEI\_AMO} + 0.01\text{PWS}), \quad (1)$$

where SST represents the average MAM SST (°C) over the region defined by 0°–20°N and 64°W–10°E (indicated by the box in Fig. 1e), MEI_AMO represents the April–May MEI conditioned upon the AMO index (as described in section 2b), and PWS is the average May zonal pseudo–wind stress for the region defined by 18°–39°N and 4°–98°W (indicated by the box in Fig. 3) raised to the 3/2 power. The coefficients are determined by fitting a Poisson regression using all data available from 1950 to 2013 and are all statistically significant at the 0.01 level.

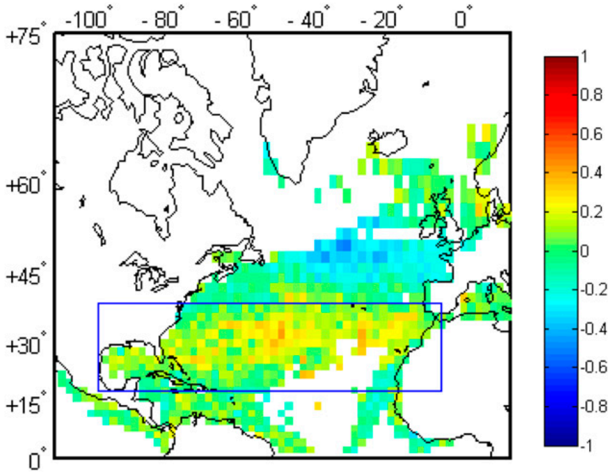

FIG. 3. The correlation of zonal pseudo–wind stress with the number of hurricanes from 1950 to 2013. The white areas in the ocean are locations that do not have data for the whole period. The box denotes the region (18°–39°N, 4°–98°W) used for computing the average stress.

For the different tests below, these coefficients will be different as we change the time period tested. In addition, each seasonal hurricane prediction made is rounded to the nearest integer.

Overall, Eq. (1) predicts the number of Atlantic hurricanes with an MAE of 1.48 and an RMSE of 1.91. The variable with the biggest impact is SST. When it is withheld and all other values are retained, MAE and RMSE increase by 0.52 and 0.64, respectively. The next most impactful variable is PWS. When it is left out, MAE and RMSE increase by 0.33 and 0.38, respectively. Finally, MEI_AMO is the least impactful, with a difference of 0.22 in MAE and of 0.20 in RMSE.

3. Results

For the 1950–2013 period, the UA model prediction using Eq. (1) by 1 June each year is found to be exactly the same as the observed hurricane number for 19% of the years, and off by +1 or −1 for 44% of the years. The absolute errors are ≥2 for 38% of the years. The fitted values have a correlation coefficient of 0.71 with the observed number of hurricanes.

a. Comparison with other models

The UA model shares many of the same predictors as other models (e.g., the CSU model), such as SST and ENSO, but still remains distinct in its exact model structure. Most importantly for maintaining multidecadal skill, the UA model conditions ENSO on AMO. First, we compare our model with three others (TSR, CSU, and CPC) for the common period of 2001–13, a challenging period for hurricane prediction. Note that we are comparing official predictions issued by early June, regardless of when the prediction was actually prepared. It is possible that some of these organizations prepare their predictions several weeks in advance but do not officially issue them until early June.

The TSR model uses the forecasted July–September trade wind speed over the Caribbean and tropical North Atlantic as well as the forecasted August–September tropical North Atlantic SSTs (http://www.tropicalstormrisk.com/). The CSU model uses four predictors for their June forecasts: SSTs in the equatorial Pacific forecasted for September by the European Centre for Medium-Range Weather Forecasts (ECMWF) on 1 May, April–May SSTs in the eastern North Atlantic, April–May 200-mb zonal wind in the tropical Pacific, and May sea level pressure in the central North Atlantic (http://hurricane.atmos.colostate.edu/Forecasts/). CPC issues its prediction range (with a 70% chance) based on climate factors that influence hurricane activity, along with the predictions of models (http://www.noaanews.noaa.gov/stories2014/20140807_hurricaneoutlook_atlantic_update.html). Also, since CPC only issues a range, we used the midpoint of their range to compare with the official predictions of other groups. This could mean the prediction is a noninteger if their range was an odd number. Because the TSR, CSU, and CPC models all use model predictions and observations in their predictions, it is probably appropriate to classify them as statistical–dynamical hybrid models. For a fair comparison with other models, we recalibrate the coefficients in Eq. (1) each year using data only available prior to that year's hurricane season. Note that besides the 1 June hurricane number evaluated here, CPC, CSU, and TSR also make many other hurricane-related predictions (e.g., Accumulated Cyclone Energy, number of major hurricanes, and number of named storms at several different lead times) for the North Atlantic.

Figure 4 compares the predictions using the UA model and these three models as well as the climatology prediction (based on the average number of hurricanes from the previous 5 years). Note that predictions given by CSU, TSR, and CPC represent their predictions for those years, not necessarily what their current models would predict. The 2001–13 period is used because all three organizations posted numeric forecasts beginning in 2001 (before 2001, CPC simply predicted how the season should compare to climatology without issuing any actual numbers). Since the variables of the UA model were chosen using data for the whole period, including 2001–13, it is possible there may be some degradation of skill for future predictions. Likewise, it is

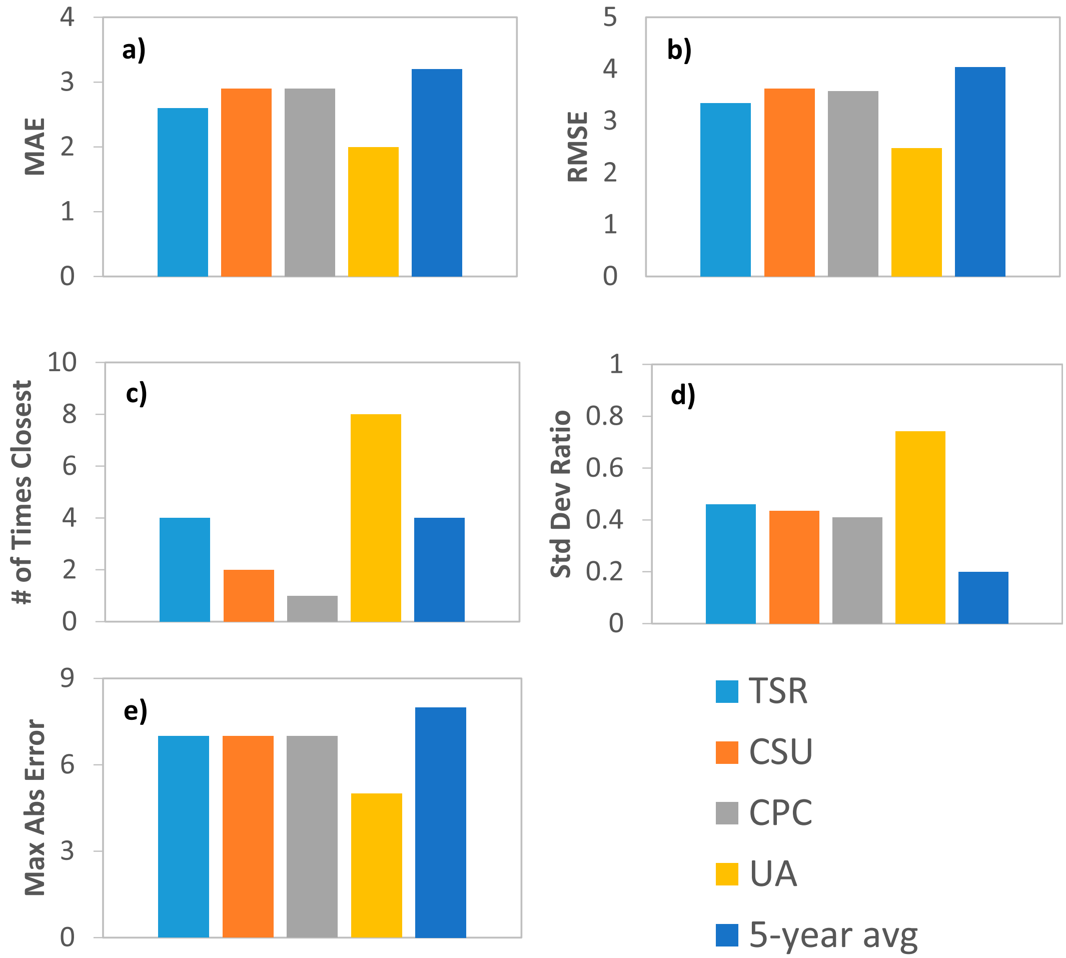

FIG. 4. Comparison of four models with the 5-yr average (the no-skill metric) for the period 2001–13: (a) MAE, (b) RMSE, (c) the number of times when the model prediction is closest to the observation (note that if two or more predictions tie on any given year, they are all deemed the closest), (d) ratio of the prediction std devs to the std dev of the actual number of hurricanes, and (e) max absolute error. Note that the coefficients in the UA model were recalibrated each year using data from prior years only to offer a fair comparison with other models.

important to mention that hindcast skill does not always translate to forecast skill in the future. Figure 4a (4b) shows that MAE (RMSE) for the UA model improves by 38% (39%) over the 5-yr climatology prediction, and MAE (RMSE) for the UA model improves by more than 23% (more than 25%) over the other three models. The number of years when the UA prediction is closest to the observed number as compared to the other four models (TSR, CSU, CPC, and climatology) is twice as high as the number from climatology or TSR, while the numbers from CSU and CPC are even lower (Fig. 4c). The ratio of standard deviations of predicted versus standard deviations of actual hurricane numbers is closest to 1 (i.e., 0.74) from UA, indicating the most realistic interannual variability of predicted hurricane numbers from UA compared with the underdispersion from climatology and the other three models (Fig. 4d). Finally, the maximum absolute error is also lowest (i.e., 5) for UA compared with the value of 7 for TSR, CSU, and CPC, and 8 for climatology (Fig. 4e). These results demonstrate the consistent improvements of the UA model for June hurricane number forecasting over the TSR, CSU, CPC, and climatology predictions based on all five metrics.

The UA model is also compared to the full record of June hurricane predictions produced by CSU. Like the previous analysis, this again is a comparison with what CSU predicted, not what their current models would predict. The UA model again uses only data available up to the hurricane season being predicted (from 1950 through May of year predicted). The MAE for CSU for all predictions between 1984 and 2013 is 2.3, and for the UA model it is 1.93, representing a 16% improvement.

Kim and Webster (2010) created a hybrid model and compared its hurricane number prediction with the CFS hybrid forecast (Wang et al. 2009) and the ECMWF forecast for the period 2002–09. The MAE for its June forecasts (1.88) is much better than that (3.0) of the ECMWF June forecasts and is even better than that (2.0) of the CFS forecasts in July–August. The UA model is found to yield exactly the same MAE as the hybrid model of Kim and Webster (2010) for this period.

b. Robustness of the UA model

North Atlantic hurricanes show strong decadal variability. Figure 2a shows a period of high activity around 1950, a decrease in activity through the early 1990s, an increase in activity until the 2000s, and finally another decrease in activity through the most recent years. Before 1980, there was also relatively little year-to-year variability (with a standard deviation $\sigma = 2.2$ for hurricane numbers per year). However, the variance has been much higher since 1980, with $\sigma = 3.1$ for hurricane numbers. Between 1990 and 2005, hurricanes varied quite widely from the mean climatological value of six.

TSR, CSU, and CPC have been able to improve upon a 5-yr climatology prediction, especially for August forecasts (Blake et al. 2010). However, the June forecasts have at times had large errors since the late 1990s. A robust model should be able to show skill for both the less variable pre-1980 period and the more variable post-1980 period using each period to initialize the model for the other.

To test the robustness of the UA model, we first divide the whole period into a 1950–80 period and a 1981–2013 period because of the difference in the dispersion of the number of hurricanes for these two periods. For the year 1980, the 1981–2013 data were used to train the model and thus issue a prediction for 1980 using the data from the spring of 1980. Next, recalibrating the model with data from 1980 to 2013, a prediction was made for 1979, and so forth going backward until 1950. For predictions for 1981, the 1950–80 period was used to train the model. Again, data from the spring of 1981 were used to predict the number of Atlantic hurricanes for 1981, and the model was recalibrated using data from 1950 to 1981 to predict for 1982, and so forth. This method, though still a hindcast, simulates real-time predictions to a large degree because it only uses data available up to that point (or after that point in terms of the backward-focused 1950–80 period, which was done simply as a sensitivity test and has no practical application). There are still some differences between a real-time prediction and a hindcast, since the variables of the UA model were chosen using data for the whole period from 1950 to 2013.

TABLE 1. A comparison between climatology (based on the average hurricane number of the prior 5 years) and the UA model based on data available up to that point for predictions after 1980, or data after that year if before 1981 (as described in the text). The model is also compared using Eq. (1) for the whole period.

		Model	Climatology	Model improvement (%)
1950–80	MAE	1.55	1.97	21
	RMSE	1.92	2.57	25
1981–2013	MAE	1.85	2.55	27
	RMSE	2.28	3.24	30
Full period	MAE	1.48	2.27	35
using Eq. (1)	RMSE	1.91	2.93	35

Table 1 shows the UA model prediction capabilities during the 1950–80 and 1981–2013 periods using MAE and RMSE. It also compares the UA model for the fitted values [i.e., using Eq. (1) for the whole period]. As expected, both the UA and the climatology approach have smaller MAEs and RMSEs for the less volatile 1950–80 period. For instance, the MAE ratio of the climatology approach for 1950–80 versus for 1981–2013 is $1.97/2.55 = 0.77$. In other words, the same climatology approach performs 23% better for 1950–80 than for 1981–2013. For the 1950–80 period, the UA model shows an MAE improvement of 21% relative to the climatology. For the more variable 1981–2013 period when the climatology approach works less well, the UA model shows a greater improvement over climatology, as its MAE is 27% lower than the climatology.

The robustness of the UA model was further tested by training the predictors during the period 1950–2013 and then performing an independent test on the period 1900–49. This is necessary because a model would be biased if all the data are used to screen predictors (DelSole and Shukla 2009). It should be noted that the uncertainty in the hurricane record increases in this presatellite era. Hurricanes not close to land or weather-reporting ships would not be included in the presatellite historical record. Vecchi and Knutson (2011) estimated that on average two hurricanes were missed per year around 1900, decreasing to about one by 1949. A local maximum occurred in the early 1940s, where three hurricanes may have been missed in 1941 because of a predicted active period.

As explained, the UA model was initialized with data from 1950 to 2013 and then predicted the number of hurricanes from 1949 to 1900, but the test was done twice, once using the data as it appears in the AOML dataset, and a second time using the corrected data provided by Vecchi and Knutson (2011). This method simulates a real-time forecast (albeit going from 1949 to

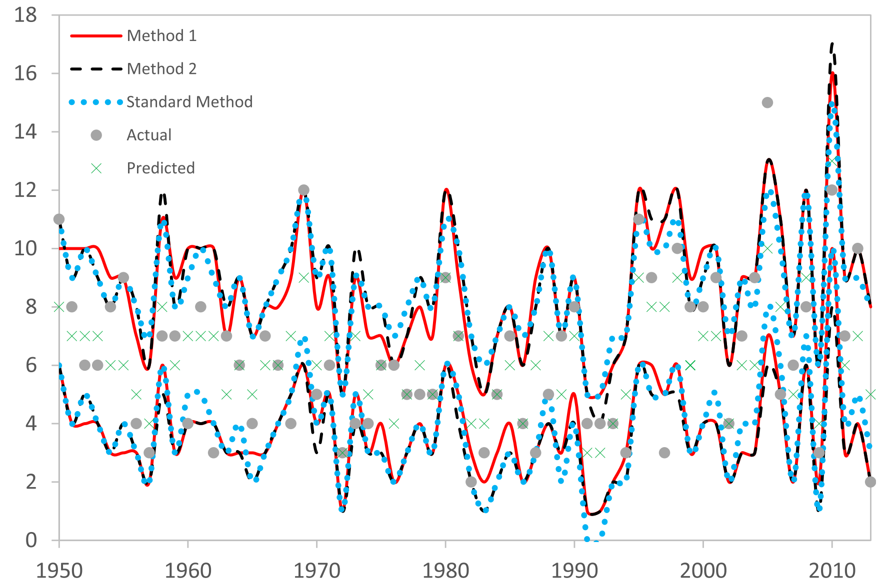

FIG. 5. The upper and lower limits of the prediction ranges, the actual yearly predictions (green crosses), and the actual number of hurricanes (gray circles) for the year. Method 1 was created by adding and subtracting one std dev based on the sign of the May AMO. Method 2 was created by using separate regressions for the upper and lower bounds (see more discussion in the text). The standard method was created by adding/subtracting one std dev of the forecasting errors.

1900, instead of from 1900 to 1949) because the model was recalibrated after each prediction, incorporating each year after the prediction was made (using the same method described above). The results were then compared to a running 5-yr average in terms of MAE and RMSE.

Using the uncorrected data, the MAEs are 2.04 for the UA model and 1.90 for the 5-yr average, respectively, while the RMSEs are 2.51 and 2.56. When the corrected data from Vecchi and Knutson (2011) are used, the UA model's performance is improved, with an MAE of 1.84 and an RMSE of 2.39. This is also better than the 5-yr average (with an MAE of 2.00 and an RMSE of 2.57).

c. *Forecasting range*

Because of the chaotic nature of the atmosphere and oceans, it is useful to provide a range for hurricane counts to complement the deterministic predictions. The standard approach would be to add ±1.97 (with $\sigma = 1.97$ being the standard deviation of forecasting errors from 1950 to 2013) to the unrounded deterministic predictions. The upper bound is then rounded up and the lower bound is rounded down. The range calculated in this way captures 84% of the actual hurricane counts (Fig. 5).

The question is then: can we compute the range that varies from year to year, with an average range similar to that in the above approach, but which captures at a higher rate the true number of hurricanes? Motivated by our prior study (Zeng et al. 2012), here we explore two new methods. The first method is the same as the above approach but using the standard deviation of model errors for years with positive May AMO ($\sigma = 2.39$) and $\sigma = 1.45$ with negative May AMO. The ranges using this method vary between 3 and 6 from 1950 to 2013 (Fig. 5), which is comparable with other models, as shown by Vecchi and Villarini (2014). Its average range of 4.9 is the same as the simple standard method, but this method captures 89% of the actual hurricane counts (compared with 84% for the standard approach).

In the second method, years (from 1950 to 2013) when the model forecasting error is greater than one are identified. Another Poisson regression is repeated for those years only, and it is then applied to the entire period to define the upper bound for the model after the values are rounded up. To define the lower bound, years with errors less than negative one are grouped together, a regression is performed, and the coefficients are applied to the whole period with the values rounded down. The accuracy of this method is higher than the standard method, as the actual number of hurricanes falls within the computed range 94% of the time (Fig. 5). The average range using this method (5.3) is slightly higher than that of the standard method (4.9) and the first method (4.9), because the range reached 7–9 for 4 out of the 64 years. If the average range (4.9) of the standard method is increased to match that of our second method (5.3), it would capture the actual number of hurricanes for 92% of the time, which is still lower than that of our second method (94%).

Each method has certain advantages. For the standard method, it is the simplest and reasonably accurate (as indicated by the percentage of actual hurricane numbers covered by the predicted range). The first new method remains simple but slightly more difficult than the standard method, maintains relatively small ranges, and is more accurate. The second new method (triple regression) is the most accurate, but has a slightly larger average range than the above two methods.

Other models also produce range estimates. TSR uses the standard deviation of errors in replicated real-time forecasts from 1980 to the year before the prediction. Their June range has lately been about six (http://www.tropicalstormrisk.com/). The range of the CPC prediction is small, typically about three, and it is supposed to capture 70% of observations. However, the CPC range has only captured 31% of the actual hurricane numbers since 2001 (http://www.noaanews.noaa.gov/stories2014/20140807_hurricaneoutlook_atlantic_update.html). CSU issues their forecast uncertainties as one standard deviation of the 1982–2010 cross-validated hindcast errors (http://hurricane.atmos.colostate.edu/Forecasts/2014/june2014/jun2014.pdf).

d. UA model forecasting in 2014

The first actual prediction using the UA model was for the 2014 hurricane season. It was widely hypothesized in preseason forecasts that 2014 would have relatively little tropical cyclone activity in the Atlantic basin. The CPC projected a 70% chance of a below-normal season (having between three and six hurricanes) as a result of 1) the likelihood of an El Niño developing, 2) below-average SSTs in the Atlantic, and 3) increased atmospheric stability (http://www.noaanews.noaa.gov/stories2014/20140807_hurricaneoutlook_atlantic_update.html). CSU also predicted that the season would be below average, at first stating there would be three hurricanes, but then raising the number up to four for their June forecast (http://hurricane.atmos.colostate.edu/Forecasts/). TSR projected the season to be more active and predicted five plus/minus three hurricanes (http://www.tropicalstormrisk.com/).

Using the UA model as described above, a prediction was made for the 2014 North Atlantic hurricane season. The conditions for 2014 show relatively average SSTs for the period; a positive May AMO, which eliminates the ENSO variable from the model; and higher-than-average pseudo–wind stress. Thus, the UA prediction for 2014 was five hurricanes. The predicted range is from three to eight using the standard method, from two to eight using the AMO-conditioned standard deviation method, and from two to seven using the triple-regression method. The deterministic number (five) is slightly below the long-term average of six hurricanes. The actual number of hurricanes for 2014 was six, which is covered by the predicted ranges from all three methods.

4. Conclusions

We have developed a statistical model to make more accurate predictions of the number of seasonal North Atlantic hurricanes by 1 June of each year. The model uses three different predictors: average MAM SSTs for the region (0°–20°N, 64°W–10°E) over the Atlantic, the April–May value of the MEI when the May AMO value is negative and zero when the value is positive, and the average zonal pseudo–wind stress for the region (18°–39°N, 4°–98°W) over the Atlantic raised to the $3/2$ power.

Atlantic SSTs that are most consistently positively correlated with the number of hurricanes lie in the region 0°–20°N and 64°W–10°E and the best predictions come from using the average of MAM. This is the most significant predictor of seasonal hurricane activity in our method.

The April–May value of the MEI correlates in an oscillatory manner with the number of hurricanes. When the May AMO value is negative, MEI more strongly correlates with the number of hurricanes, but when the AMO is positive, the correlation is only slightly positive. The conditioning of the MEI on the AMO value is another important reason for the success of our method.

Finally, the zonal pseudo–wind stress is used because of its ability to affect the distribution of SSTs around the ocean and its high correlation with sea level pressure when both are taken over the region 18°–39°N, 4°–98°W for the month of May. Raising the average zonal pseudo–wind stress to the $3/2$ power helps the model to better capture the interannual variability of the hurricane numbers.

Using a fitted regression line for the period of 1950–2013, our model yields an RMSE of 1.91 and a correlation coefficient of 0.71 when compared with the observed number of hurricanes. Performing simulated real-time hindcasts, compared with a no-skill metric (or climatology prediction) based on a 5-yr running average, our model improves the MAE by 38% and the RMSE by 39% for the period 2001–13. Compared with predictions from TSR, CSU, and CPC, we see an improvement of at least 23% in MAE and at least 25% in RMSE from 2001 to 2013. It is important to mention, however, that hindcast skill does not always translate to forecast skill in the future.

The robustness of our model has been tested by initializing the model using data from 1950 to 1980 (or from 1981 to 2013) for prediction during 1981–2013 (or 1950–80). For these two periods, the UA model shows an MAE improvement of 21% for 1950–80 and 27% for 1981–2013 over the climatology prediction. Our model also does a reasonable job of prediction using an independent dataset (years 1900–49), beating the 5-yr running average climatology prediction based on the corrected hurricane number data from Vecchi and Knutson (2011).

We have also created two uncertainty ranges of the predicted hurricane number for every season from 1950 to 2013 that better predicts the uncertainty than the standard method of adding/subtracting a standard deviation of the errors. The first is using the standard deviation of forecast errors for positive (negative) May AMO years and adding/subtracting the value to predictions with corresponding May AMO signs. The second is rerunning a regression for errors greater (less) than one (negative one) and using the new coefficients to fit an upper (lower) bound. The first new method improves upon the standard method by capturing the actual number of hurricanes 89% of the time (versus 84%) from 1950 to 2013 while maintaining the same average range. The second method further increases the percentage to 94% but has a slightly larger average range (5.3) than the standard or first new method (4.9).

Future research should explore not only how to further improve the accuracy of the predictions, but also how to improve upon them with greater lead time. This would be beneficial for those responsible for mitigating human casualties and property damage, such as community groups, government agencies, and insurance companies. Furthermore, improvement upon other tropical cyclone intensity categories, such as all named storms, tropical storms, and major hurricanes, in all basins of the world, would also be beneficial.

Acknowledgments. This work was supported by Science Foundation Arizona, the NSF (AGS-0944101), and NASA (NNX14AM02G). The authors thank AOML for the historical hurricane data, ICOADS for the wind stress data, NCEP for sea level pressure data, ESRL for the AMO and MEI data, and NCDC for the SST data. The authors appreciate the help of Mike Brunke and Patrick Broxton for their editing support. Chris Landsea, Phil Klotzbach, and an anonymous reviewer are thanked for their constructive and helpful comments.

REFERENCES

Bell, G. D., and Coauthors, 2000: Climate assessment for 1999. *Bull. Amer. Meteor. Soc.,* **81** (6), S1–S50, doi:10.1175/1520-0477(2000)081<1328:CAF>2.3.CO;2.

Bevere, L., T. Seiler, P. Zimmerli, and H. Feyen, 2013: Natural catastrophes and man-made disasters in 2012: A year of extreme weather events in the US. Sigma Archive 2/2013, Swiss Reinsurance Company, 39 pp. [Available online at http://media.swissre.com/documents/sigma2_2013_EN.pdf.]

Blake, E. S., R. J. Pasch, and G. D. Bell, 2010: Verification of 12 years of NOAA seasonal hurricane forecasts. Preprints, *29th Conf. on Hurricanes and Tropical Meteorology,* Tucson, AZ, Amer. Meteor. Soc., 3A.1. [Available online at https://ams.confex.com/ams/pdfpapers/167939.pdf.]

Changnon, S. A., D. Changnon, E. R. Fosse, D. C. Hoganson, R. J. Roth Sr., and J. M. Totsch, 1997: Effects of recent weather extremes on the insurance industry: Major implications for the atmospheric sciences. *Bull. Amer. Meteor. Soc.,* **78,** 425–435, doi:10.1175/1520-0477(1997)078<0425:EORWEO>2.0.CO;2.

Chiang, J. C. H., and D. J. Vimont, 2004: Analogous Pacific and Atlantic meridional modes of tropical atmosphere–ocean variability. *J. Climate,* **17,** 4143–4158, doi:10.1175/JCLI4953.1.

DelSole, T., and J. Shukla, 2009: Artificial skill due to predictor screening. *J. Climate,* **22,** 331–345, doi:10.1175/2008JCLI2414.1.

Elsner, J. B., and C. P. Schmertmann, 1993: Improving extended-range seasonal predictions of intense Atlantic hurricane activity. *Wea. Forecasting,* **8,** 345–351, doi:10.1175/1520-0434(1993)008<0345:IERSPO>2.0.CO;2.

——, and T. H. Jagger, 2006: Prediction models for annual U.S. hurricane counts. *J. Climate,* **19,** 2935–2952, doi:10.1175/JCLI3729.1.

Enfield, D. B., A. M. Mestas-Nunez, and P. J. Trimble, 2001: The Atlantic multidecadal oscillation and its relationship to rainfall and river flows in the continental U.S. *Geophys. Res. Lett.,* **28,** 2077–2080, doi:10.1029/2000GL012745.

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, doi:10.1126/science.1060040.

Gray, W. M., 1984: Atlantic seasonal hurricane frequency. Part I: El Niño and 30 mb quasi-biennial oscillation influences. *Mon. Wea. Rev.,* **112,** 1649–1668, doi:10.1175/1520-0493(1984)112<1649:ASHFPI>2.0.CO;2.

——, C. W. Landsea, P. W. Mielke Jr., and K. J. Berry, 1992: Predicting Atlantic seasonal hurricane activity 6–11 months in advance. *Wea. Forecasting,* **7,** 440–455, doi:10.1175/1520-0434(1992)007<0440:PASHAM>2.0.CO;2.

——, ——, ——, and ——, 1993: Predicting Atlantic basin seasonal tropical cyclone activity by 1 August. *Wea. Forecasting,* **8,** 73–86, doi:10.1175/1520-0434(1993)008<0073:PABSTC>2.0.CO;2.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kim, H.-M., and P. J. Webster, 2010: Extended-range seasonal hurricane forecasts for the North Atlantic with a hybrid dynamical–statistical model. *Geophys. Res. Lett.,* **37,** L21705, doi:10.1029/2010GL044792.

Klotzbach, P. J., 2011: El Niño–Southern Oscillation's impact on Atlantic basin hurricanes and U.S. landfalls. *J. Climate,* **24,** 1252–1263, doi:10.1175/2010JCLI3799.1.

Knaff, J. A., 1997: Implications of summertime sea level pressure anomalies in the tropical Atlantic region. *J. Climate,* **10,** 789–804, doi:10.1175/1520-0442(1997)010<0789:IOSSLP>2.0.CO;2.

Kraus, E. B., and J. A. Businger, 1994: *Atmosphere–Ocean Interaction.* Oxford University Press, 362 pp.

Murnane, R. J., 2004: Climate research and reinsurance. *Bull. Amer. Meteor. Soc.,* **85,** 697–707, doi:10.1175/BAMS-85-5-697.

Palmen, E., 1948: On the formation and structure of tropical hurricanes. *Geophysica,* **3,** 26–38.

Smirnov, D., and D. J. Vimont, 2012: Extratropical forcing of tropical Atlantic variability during boreal summer and fall. *J. Climate,* **25,** 2056–2076, doi:10.1175/JCLI-D-11-00104.1.

Smith, S. R., J. Servain, D. M. Legler, J. N. Stricherz, M. A. Bourassa, and J. J. O'Brien, 2004: In situ-based pseudo–wind stress products for the tropical oceans. *Bull. Amer. Meteor. Soc.,* **85,** 979–994, doi:10.1175/BAMS-85-7-979.

Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore, 2008: Improvements to NOAA's historical merged land–ocean temperature analysis (1880–2006). *J. Climate,* **21,** 2283–2296, doi:10.1175/2007JCLI2100.1.

Tang, B. H., and J. D. Neelin, 2004: ENSO influence on Atlantic hurricanes via tropospheric warming. *Geophys. Res. Lett.,* **31,** L24204, doi:10.1029/2004GL021072.

Vecchi, G. A., and T. R. Knutson, 2011: Estimating annual numbers of Atlantic hurricanes missing from the HURDAT database (1878–1965) using ship track density. *J. Climate,* **24,** 1736–1746, doi:10.1175/2010JCLI3810.1.

——, and G. Villarini, 2014: Next season's hurricanes. *Science,* **343,** 618–619, doi:10.1126/science.1247759.

——, and Coauthors, 2013: Multiyear predictions of North Atlantic hurricane frequency: Promise and limitations. *J. Climate,* **26,** 5337–5357, doi:10.1175/JCLI-D-12-00464.1.

Vimont, D. J., and J. P. Kossin, 2007: The Atlantic meridional mode and hurricane activity. *Geophys. Res. Lett.,* **34,** L07709, doi:10.1029/2007GL029683.

Vitart, F., and Coauthors, 2007: Dynamically-based seasonal forecasts of Atlantic tropical storm activity issued in June by EUROSIP. *Geophys. Res. Lett.,* **34,** L16815, doi:10.1029/2007GL030740.

Wang, H., J.-K. E. Schemm, A. Kumar, W. Wang, L. Long, M. Chelliah, G. D. Bell, and P. Peng, 2009: A statistical forecast model for Atlantic seasonal hurricane activity based on the NCEP dynamical seasonal forecast. *J. Climate,* **22,** 4481–4500, doi:10.1175/2009JCLI2753.1.

Wolter, K., and M. S. Timlin, 1993: Monitoring ENSO in COADS with a seasonally adjusted principal component index. *Proc. 17th Climate Diagnostics Workshop,* Norman, OK, NOAA/NMC/CAC, 52–57.

Xue, Y., T. M. Smith, and R. W. Reynolds, 2003: Interdecadal changes of 30-yr SST normals during 1871–2000. *J. Climate,* **16,** 1601–1612, doi:10.1175/1520-0442-16.10.1601.

Zanetti, A., R. Enz, I. Menzinger, J. Mehlhorn, and S. Suter, 2003: Natural catastrophes and man-made disasters in 2002: High flood loss burden. Sigma Archive 2/2003, Swiss Reinsurance Company, 37 pp. [Available online at http://reliefweb.int/sites/reliefweb.int/files/resources/365140A8A2E57BCCC1256CF3002EA021-swissre-annual-2002.pdf.]

Zeng, X., K. L. Kiviat, K. Sakaguchi, and A. M. A. Mahmoud, 2012: A toy model for monthly river flow forecasting. *J. Hydrol.,* **452–453,** 226–231, doi:10.1016/j.jhydrol.2012.05.053.