Seasonal Prediction of North Atlantic Accumulated Cyclone Energy and Major Hurricane Activity

KYLE DAVIS AND XUBIN ZENG

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

(Manuscript received 30 July 2018, in final form 12 December 2018)

ABSTRACT

Building upon our previous seasonal hurricane prediction model, here we develop two statistical models to predict the number of major hurricanes (MHs) and accumulated cyclone energy (ACE) in the North Atlantic basin using monthly data from March to May for an early June forecast. The input data include zonal pseudo-wind stress to the $3/2$ power, sea surface temperature in the North Atlantic, and, depending on the magnitude of the Atlantic multidecadal oscillation index, the multivariate ENSO index. From 1968 to 2017, these models have a mean absolute error of 0.96 storms for MHs and 30 units for ACE. When tested over an independent period from 1958 to 1967, the models show a 22% improvement for MHs and 16% for ACE over a no-skill metric based on a 5-yr running average. Both the MH and ACE results show consistent improvements over those produced by three other centers using statistical–dynamical hybrid models and a 5-yr running average prediction over the period 2000–17 for MHs (2003–17 for ACE) in a simulated real-time prediction. These improvements vary from 25% to 37% for MHs and from 15% to 37% for ACE. While most forecasting centers called for a slightly above-average hurricane season in May/June 2017, our models predicted in June 2017 a very active season, in much better agreement with observations.

1. Introduction

The strength of the Atlantic hurricane season, which runs from 1 June to 30 November every year, can be measured in many ways, including the numbers of named tropical cyclones (wind speeds > 33 kt; $1 \text{ kt} \approx 0.51 \text{ m s}^{-1}$), hurricanes (> 63 kt), and major hurricanes (MHs; > 95 kt), as well as accumulated cyclone energy (ACE). ACE is defined as the sum of the squares of the maximum 1-min sustained surface wind speed in knots every 6 h while the storm is either tropical or subtropical in nature and at least of tropical storm strength (Bell et al. 2000). Using multiple metrics listed above allows a more complete picture of a hurricane season. For instance, according to the historical NHC “best track” hurricane database (HURDAT2; Landsea and Franklin 2013), while the 1981–2010 averages for the numbers of named storms, hurricanes, major hurricanes, and ACE are 11.9, 6.4, 2.7, and 106, respectively, the year 2013 saw 14 named storms (~ 2 higher than average), 2 hurricanes (~ 4 less than average), 0 major hurricanes (~ 2 less than average), and an ACE of 36 (~ 70 units less than average). Though an inactive year by most metrics, 2013 would

appear to be an active year if judged only by the number of named storms (14 vs the average of 11.9). As another example, 2005 had $\sim 10\%$ more ACE than 2004, yet 2005 had 28 named storms while 2004 only had 15.

The 2017 North Atlantic hurricane season was a reminder of the devastating impacts hurricanes have on countries, including the conterminous United States, which did not see any major hurricane landfalls between Wilma in October 2005 and Harvey in August 2017 (Truchelut and Staehling 2017). This period also included no hurricane landfalls of any sort between Ike in 2008 and Irene in 2011. Harvey broke the long major hurricane drought in August 2017 when it made landfall as a category 4 storm near Port Aransas, Texas, producing a deluge with accumulated rainfall exceeding 50 in. in regions from Houston to Beaumont (Blake and Zelinsky 2018). Irma followed soon after and caused massive losses in the southeastern United States, with damage estimated with 90% confidence between \$37.5 and \$62.5 billion (Cangialosi et al. 2018). Thus, the 2017 North Atlantic season was the first to feature two major hurricane landfalls over the conterminous United States since 2005. Early outlooks for the number of Atlantic hurricanes, MHs, and ACE in 2017 generally significantly underestimated the observed activity (to be further discussed in section 3).

Corresponding author: Kyle Davis, davis7000@gmail.com

DOI: 10.1175/WAF-D-18-0125.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) (www.ametsoc.org/PUBSReuseLicenses).

Predicting seasonal North Atlantic hurricane activity has been done for decades, with the first prediction from Colorado State University (CSU) issued in 1984 (Gray 1984). Today, the practice has become widely adopted by many different organizations. Hurricane predictions are made at several lead times, usually in advance of the hurricane season (prior to 1 June), at the start of the hurricane season, and right before the peak activity for the season (in late July/early August). Currently forecasts are typically issued not only for the number of hurricanes but also for the number of named storms, MHs, and ACE. Attempts have also been made to produce long-term forecasts for ACE and MHs (Murakami et al. 2016; Villarini and Vecchi 2013), as well as for many different variables besides these four (Villarini et al. 2016). Predicting all four metrics accurately—attempted with hybrid, statistical, and dynamic models—presents a challenge, even if all four metrics are highly correlated. For example, Harnos et al. (2017) showed that their model prediction ranges captured 52% of observed hurricanes, but only 34% of ACE values when their models were initialized in July. The question is: Can we develop a new set of models to improve upon the performance of existing dynamic and hybrid models in MH and ACE prediction?

The seasonal hurricane prediction model we developed, as described by Davis et al. (2015, hereafter DZR), predicted the number of hurricanes in the North Atlantic basin with better performance in hindcast mode than did three statistical–dynamical hybrid models produced in real time by CSU, Tropical Storm Risk (TSR; www.tropicalstormrisk.com), and the National Oceanic and Atmospheric Administration's (NOAA) Climate Prediction Center (CPC). In particular, while seasonal hurricane predictions for the 2017 season issued by 1 June by these three centers anticipated an average year with 5–7 hurricanes, our model predicted a very active season with 11 hurricanes, which agrees well with observations (10 hurricanes). To provide a more complete picture of the hurricane season and build upon the success of our seasonal hurricane number prediction model, here we expand our work to include two statistical models to predict ACE and MH numbers in the North Atlantic basin using data through the end of May.

Using a simple linear regression model, we can get a sense of the relationship between total hurricane numbers and both MHs and ACE. If we create a regression using observed hurricanes to predict MHs and ACE for the same year from 1968 to 2017, we find that the mean absolute error (MAE) is 0.68 with a correlation of 0.82 for MHs, while the MAE is 22 for ACE with a correlation of 0.89. Hence, knowing only the actual hurricane count for each year would give us enough information to create a very

accurate model for MHs and ACE, which is reasonable since MH is a subset of hurricane numbers, and hurricane numbers typically make up a significant amount of ACE.

Therefore, one straightforward approach to predicting MHs and ACE would be to develop a linear regression of MHs and ACE with the predicted hurricane numbers in DZR. Using this method yields MAEs for MH and ACE of 1.06 and 33, respectively. Results are very similar (1.08 and 34) if we use the same variables in DZR and adjust the coefficients to predict MH and ACE (using the same predictors but changing only the predictands). These results are much better than the no-skill metric based on the 5-yr running average (WMO 2008), which has an MAE of 1.4 for MH and 51 for ACE. Obviously, we have to show that our new models are an improvement over the above baseline. Section 2 presents the details of our methods, while section 3 compares our results with the above baseline and other models.

2. Data and methods

Our new models build upon our total hurricane prediction model as described in DZR. We chose a Poisson regression following previous studies (DZR; Elsner and Jagger 2006; Elsner and Schmertmann 1993) for its ability to model count data. Other options do exist, such as the gamma distribution employed by Villarini and Vecchi (2012). Input data from 1968 to 2017 are used. While our models' predictions are derived from data from March through May, the forecast is typically not available until about the second week of June as there is a time lag from when the data are published. However, this delay is likely to be insignificant as nearly all hurricane activity occurs after the first two weeks of June (Gray et al. 1993).

To build out our suite of models, our initial hope was to use the same equation from DZR since MH is a subset of hurricanes and together generated ~90% of Atlantic ACE from 1981 to 2010. However, upon inspection, we found that, though the general structure of the model would remain the same (due to the high correlation between MH, ACE, and hurricanes), each variable would need to be changed in one or both of the models to accommodate the differences inherent to both MH and ACE. For instance, 83% of major hurricanes from 1967 to 1991 formed from African easterly waves that move westward across the hurricane main development region (Landsea 1993), whereas hurricanes that form near the coastline often do not have enough time or have conducive conditions to become MHs. ACE, however, though primarily driven by MHs, also includes tropical storms, which are able to form at higher latitudes and can be a large driver of ACE in very quiet years. In addition, in using our model from DZR, we have noticed enhancements that

FIG. 1. Statistically significant correlations ($p < 0.05$) of MAM-averaged SSTs with (a) MH and (b) ACE from 1968 to 2017. Note the area used for the MH model in (a), represented by the box (64°W – 10°E and 2° – 20°N) near the hurricane main development region where most MHs form. No such area shown in (b) because the AMO index is used.

need to be made for MH and ACE prediction, particularly with how we factor in ENSO. Hence, all of these factors contributed to the use of additional models.

Input data for our models are provided by the Atlantic Oceanographic and Meteorological Laboratory (AOML) using the HURDAT2 dataset, which provides 6-hourly information on position and intensity for storms back to 1851 (Landsea and Franklin 2013). We use the March–May (MAM) averaged Atlantic multidecadal oscillation (AMO) index (Goldenberg et al. 2001), which is the detrended area-weighted average sea surface temperature (SST) over the North Atlantic, and the MAM-averaged SST in the tropical North Atlantic using the NOAA Extended Reconstructed Sea Surface Temperature (ERSST) version 4 dataset (Smith et al. 2008; Xue et al. 2003). Additionally, we utilize zonal pseudo-wind stress (ZPWS) from ICOADS (Smith et al. 2004), which is defined as the magnitude of the wind multiplied by the wind vector in the zonal direction. Finally, we use the multivariate ENSO index (MEI; Wolter and Timlin 1993), which is based on six variables over the tropical Pacific, including sea level pressure, zonal and meridional directions of the surface winds, SST, surface air temperature, and total cloudiness fraction of the sky. We condition the MEI upon the AMO, as is explained later. The rest of this section will outline each predictor and the functional form of each model. The differences from DZR will also be highlighted.

a. Sea surface temperatures

The effect of warm SSTs on hurricane development is well known (Palmen 1948; Emanuel 1986).

The warm ocean creates a tremendous amount of energy that, under the right atmospheric conditions, fuels storms. Warm SSTs enhance latent and sensible heat fluxes and are usually associated with lower surface pressure over the tropical Atlantic, enhanced midlevel moisture, and reduced vertical wind shear (Knaff 1997).

Figure 1 shows three main areas where MAM-averaged SSTs are highly correlated with ACE and MH: the far North Atlantic, an area off the western coast of Europe and Africa, and an area off of the coast of South America stretching across the equatorial Atlantic. Though these areas have the highest correlations, most of the North Atlantic has a degree of positive correlation with ACE and MH, except for a region off of the East Coast of the United States. This resembles the AMO pattern as given in Goldenberg et al. (2001), or the first rotated empirical orthogonal function of non-ENSO global SST variability. The high correlations in the tropical Atlantic also resemble the Atlantic meridional mode (Chiang and Vimont 2004). Using the average MAM AMO values produces a better prediction for ACE than using tropical Atlantic area-averaged SSTs. The AMO may be a better predictor for ACE because it takes into account the whole North Atlantic, thus better capturing storms that may form in subtropical zones or outside of the hurricane main development region. MH, which is the major driver for ACE, is predicted using the MAM SSTs in the tropical Atlantic between 64°W – 10°E and 2° – 20°N , the same as in DZR, as the area used in the tropical Atlantic also correlates highly with MH. This is expected since, as

FIG. 2. Relationship between the average peak hurricane season MEI values against annual (a) MH numbers and (b) ACE. The MEI values from August through October were used. Activity is noticeably suppressed when the average MEI is above 0.5 and even more when the MEI is above 1. Otherwise, there is very little correlation between MEI and MH/ACE.

already noted, most major hurricanes form from African easterly waves.

b. ENSO

It has been well documented that El Niño reduces hurricane activity in the tropical North Atlantic through increased wind shear in the Caribbean and the tropical North Atlantic (Gray 1984), as well as increased static stability over the Atlantic main development region (Tang and Neelin 2004). However, this relationship is not entirely linear. Figure 2 shows suppressed activity in terms of MH and ACE for every year when the average MEI during the peak of the hurricane season (August–October) is greater than 0.5, and very inactive years when the MEI is greater

than 1. However, once the average MEI is less than 0.5, the relationship is much less clear. There seems to be a spike in activity when the average MEI is around zero, but otherwise there is almost no relationship.

Using the April–May MEI values before the start of hurricane season from 1968 to 2017, ACE and MEI correlate very poorly (0.01) with similar results for MH and MEI (−0.03) compared to −0.34 and −0.36, respectively, for August/September MEI values. This suggests that the April/May MEI values have limited ability to predict the MEI values during the peak hurricane season (which is related to MH and ACE to a certain degree, as shown in Fig. 2). This is due to the ENSO springtime predictability barrier, a time

FIG. 3. The relationship between April–May MEI and ACE/MH for (a) years when the May AMO is < -0.1 , (b) years when May AMO is between -0.1 and 0.1 , (c) May AMO > 0.1 , and (d) the new relationship between AMO and MEI as described in the text. Conditioned in this manner, the correlation between the April–May MEI and MH (ACE) goes from -0.03 (0.02) to -0.23 (-0.48).

when ENSO undergoes rapid changes (Webster and Yang 1992). Thus, a better method must be found to incorporate the effects of the MEI before the start of hurricane season.

As a starting point, DZR mentioned one way of approaching this issue: using the MEI only when the AMO is negative. This approach yielded better statistical relationships as documented in DZR. This was also found in Klotzbach (2011), who compared ACE and MH by the phase of the AMO and the sign of ENSO. He stated that, when the AMO is in its negative phase, background conditions are less favorable for development, and prohibitive when combined with an El Niño event. However, when the AMO phase is positive, a weak or moderate phase of El Niño may not be enough to suppress hurricane activity. DZR's weakness in applying the ENSO variable lies in the jump in prediction when the AMO is very close to zero. For instance, in 2015, despite the potential for the onset of a strong El Niño, the AMO jumped from a small negative value to a small positive value, leading to the omission of the El Niño effect in our equation. This caused an overprediction of hurricanes. Thus, to create a smoother transition in considering the ENSO effect and to better capture years that show

potential for a strong El Niño–La Niña, we have implemented a new set of rules as follows:

- When the AMO is $< -0.1^{\circ}\text{C}$ in May, there is a greater correlation between MEI and ACE/MH, and thus we use the April/May MEI directly.
- If the AMO is neutral (between -0.1° and 0.1°C , inclusive), we only use the MEI if its absolute value is > 1 , which may signify that a strong event is coming.
- When the AMO is $> 0.1^{\circ}\text{C}$, we do not use the MEI, though a strong El Niño event may yet suppress activity. Attempts to estimate the strength of the ENSO event at the start of hurricane season seem to do more harm than good for our models when the May AMO is $> 0.1^{\circ}\text{C}$.

Figure 3 shows the relationship between MH/ACE and MEI for the three different divisions shown above, as well as how the relationships change when we implement the above rules. The main improvement here we have made over DZR is adding a third category (analyzing the AMO when it is between -0.1 and 0.1). We found this method to work better than using Pacific SSTs directly or using May wind shear values. If we only consider years when we use the MEI, the correlation coefficients of MH and ACE with MEI

improve to -0.23 and -0.48 , respectively, compared to -0.19 and -0.24 if we had used the method in DZR.

c. Zonal pseudo-wind stress

We use the ZPWS raised to the $3/2$ power, which is defined as the magnitude of the wind multiplied by the wind vector in the zonal direction, making it proportional to the surface wind stress (Smith et al. 2004; Zeng et al. 1998). Because the ZPWS can be either positive or negative, we raise the absolute values to the $3/2$ power and then multiply it by its original sign, either $+1$ or -1 . This variable affects the horizontal and vertical distributions of ocean temperature and correlates highly with sea level pressure. Low sea level pressure is needed for hurricane development and is associated with more midlevel moisture, deeper convection, and less subsidence (Knaff 1997). It is also related to vertical wind shear (i.e., the vertical gradient of the horizontal wind vector), which is widely recognized as an important factor affecting hurricane development over the tropical Atlantic (e.g., Gray 1984; Gray et al. 1993). However, vertical wind shear's memory at the seasonal time scale is limited. Therefore we chose the May-averaged ZPWS (rather than the May vertical wind shear) in DZR and this study, as wind stress is related to the horizontal gradient of SSTs and hence has a longer memory.

For our new models, we use the May-averaged ZPWS across 25° – 35° N and 87° – 47° W raised to the $3/2$ power (Figs. 4a,b), as the turbulent dissipation rate in the ocean mixed layer is proportional to the $3/2$ power of the wind stress (Kraus and Businger 1994). The area we use for our new models is much smaller than in DZR. The area does not include the Gulf of Mexico, stays north of the Yucatan Peninsula, goes north to around Cape Hatteras, and stretches east about halfway across the Atlantic; all of which is different from DZR. When we average the boxed region shown in Figs. 4a and 4b and then raise it to the $3/2$ power, its area-averaged correlation with MH is 0.25 (significant with $p < 0.1$), and its correlation with ACE is 0.37 (significant with $p < 0.01$).

The use of this new area is partly motivated by the correlation analysis below. Using August–October NCEP reanalysis zonal wind shear (WS) (Kalnay et al. 1996), calculated as the zonal wind difference between 200 and 700 hPa, we correlated ZPWS to the $3/2$ over the new area and that in DZR with these WS values over the Atlantic. Using the area in DZR yields correlations with WS south and southeast of the Dominican Republic and Puerto Rico. On the other hand, using the new area yields correlations with WS in the same region and farther to the northeast into the main Atlantic hurricane development region as defined by Goldenberg and Shapiro (1996) (Fig. 4c). This correlation farther north provides

additional information to the model, as WS in the deep tropics is already picked up by the ENSO variable (Gray 1984).

Figure 4d shows the correlation between May WS (from the region shown in the box) and August–October WS. Though May WS over the boxed region in Fig. 4d does correlate even better with MH and ACE than ZPWS (Fig. 4c), May WS does not reduce as much error in our model. This is likely for the same reason why our new area works better than that in DZR: WS in the deep tropics is partly captured by our ENSO variable and hence is a partially redundant variable.

Figures 4e and 4f shows how ZPWS to the $3/2$ is correlated with SSTs. This correlation shows a north–south gradient across the deep tropics to the subtropics over the Atlantic (Fig. 4e). Higher ZPWS to the $3/2$ (relative to zero, not absolute strength) correlates, in May, with lower SSTs in the deep tropics and higher SSTs in the subtropics over the Atlantic, as well as lower SSTs over the tropical Pacific related to ENSO. In fact, ZPWS to the $3/2$ has a weak negative correlation (-0.22) with the tropical SSTs we use in the MH model. This partly explains why this variable has an effect on our models: higher May ZPWS indicates warmer SSTs in the Gulf of Mexico and subtropics just prior to hurricane season (Fig. 4e), areas where storms are more likely to form earlier in the hurricane season (though not as likely to become MHs). The correlation of ZPWS with ACE for June/July is 0.34 , similar to the correlation for the whole season (0.37). This may also explain the stronger correlation of this variable with ACE than MH. By peak hurricane season, correlations over the Atlantic become weak, but a moderate correlation still exists with the ENSO region (Fig. 4f).

d. Functional forms of the ACE and major hurricane models

Using the variables as defined above, the University of Arizona (UA) MH model takes on the following functional form to compute the expected value of MH using a Poisson regression,

$$E(\text{MH}|\mathbf{x}) = \exp(-23.60 + 0.94\text{SST}_{\text{MAM}} + 0.004\text{ZPWS}^{3/2} - 0.28\text{MEI}|\text{AMO}), \quad (1)$$

and for ACE numbers,

$$E(\text{ACE}|\mathbf{x}) = \exp(4.73 + 1.81\text{AMO}_{\text{MAM}} + 0.003\text{ZPWS}^{3/2} - 0.32\text{MEI}|\text{AMO}), \quad (2)$$

where AMO_{MAM} equals the MAM-averaged AMO index; SST_{MAM} equals the MAM SSTs over the region from 64° W to 10° E and from 2° to 20° N; ZPWS equals

a) MH **b) ACE** **c) May ZPWS Correlated with Aug-Oct WS** **d) May WS Correlated with Aug-Oct WS** **e) May ZPWS Correlated with May SST** **f) May ZPWS Correlated with Aug-Oct SST**

FIG. 4. ZPWS correlated with (a) MH numbers and (b) ACE from 1968 to 2017. The boxes show the area used for the models, 25°–35°N and 87°–47°W, and the gray areas over the ocean represent areas where we do not have data for the full period. The area average of the boxed area is raised to the 3/2 power. (c) The correlation of May area-averaged ZPWS as used in (a) and (b) with August–October zonal WS. (d) May area-averaged WS over the boxed area correlated with August–October WS. (e), (f) May area-averaged ZPWS correlated with May SST and peak hurricane season SST.

the zonal pseudo-wind stress in May over the North Atlantic within the area from 25° to 35°N and from 87° to 47°W; MEI|AMO equals the MEI conditioned upon the AMO, as discussed in [section 2b](#); and $\mathbf{x}$ is the vector of these predictors. Note that the coefficients in these two equations are based on data from 1968 to 2017 and will be adjusted each year based upon newly available data. While the above equations are not constrained to integer values, we round the output to the nearest integer. The MH model variables contain the following p values: <0.001 for SST_{MAM}, 0.066 for ZPWS^{3/2}, and 0.084 for MEI|AMO, and for ACE all p values are <0.001 . Because we use different models for MH [Eq. (1)] and hurricane numbers (in [DZR](#)), we need to ensure their consistency (e.g., avoid a potential prediction of more MH than hurricanes). Therefore, we compute the ratio of observed MHs to observed hurricane numbers for every year since 1968 and take the maximum ratio (which is 0.66). We then apply this to the predicted hurricane number from [DZR](#) to constrain the predicted MH value (and take it as the new expected value in the Poisson distribution). This helps prevent the MH prediction from unreasonably high values.

To find the most parsimonious model, we use Akaike's information criteria (AIC) to determine if any variables should not be kept due to insignificant or redundant contributions to the model. As we experimented with the use of different combinations of variables, we find Eqs. (1) and (2) have the lowest AIC values (i.e., are most parsimonious).

The two most unique aspects of our models compared to others include 1) using the AMO and ENSO together to improve consistency of performance over different decades and 2) using ZPWS (rather than upper- or lower-level winds). In contrast to [DZR](#), we have a different treatment of the MEI conditioned by the AMO; adopt different latitudes/longitudes for box averages of ZPWS; use the AMO, rather than SST over the tropical North Atlantic for ACE in Eq. (2); use different methods to create a probabilistic range (to be explained in [section 3](#)); and provide more discussions on the physical basis of our selected predictors. Most importantly, compared with other centers that combine their statistical models with state-of-the-art dynamic seasonal prediction models, can our statistical models [Eqs. (1) and (2)] using observational data from March–May outperform their hybrid models in predicting MH and ACE from June to November? This will be addressed next.

3. Results

[Figure 5](#) compares our predicted MH numbers and ACE with observations from 1968 to 2017. The correlation is 0.65 for MH (significant at $p < 0.01$) and 0.75

for ACE (significant also at $p < 0.01$). Our MH model has a mean average error (MAE) of 0.96, which is 80% of the standard deviation of the year-to-year variation in observed MH, and for the ACE model, an MAE of 30, or 66% of the standard deviation of observed ACE. These results improve upon those (1.1 for MH and 33 for ACE) using the baseline model established by [DZR](#) (see [section 1](#)). In fact, using our new ACE model to predict hurricane numbers like [DZR](#) (after adjusting the coefficients) does yield a small improvement (MAE of 1.46 vs 1.60 hurricanes in [DZR](#)). Comparing again to the baseline model, from 1968 to 2017, our new model performs better in MH prediction in 13 years, the baseline outperforms in 7 years, and their performance is the same for the other 30 years. Results are similar for ACE, with the new model performing better four more times (vs 13 – 7 = 6 more times for MH). The new models also have lower maximum errors (3 vs 4 for MH, 117 vs 133 for ACE). The better performance over the baseline model represents the value of our new models over [DZR](#).

[Figure 5](#) also shows that our models capture the decadal trends in both MH and ACE. Looking specifically at the MH model, it captures the general trend of low year-to-year variation pre-1995 with no years having more than three MHs, except 1969. After 1994, there is significantly more variation in the observed MH number, which our model generally follows. Our ACE model shows similar trends, with less variation in ACE before 1995 than since 1995.

The forecast range can be explored using the Poisson distribution with a mean and variance of the predicted values from Eqs. (1) and (2) and defined as the difference between the 75th and 25th percentiles. This range is 3.0 on average for MH and captures the observed value 78% of the time. In contrast, the range based on the ± 1 standard deviations of the model prediction errors is 3.8 on average and captures approximately 86% of the observations.

For ACE, the range based on the 25th and 75th percentiles of the Poisson distribution is 14 on average and captures only 22% of the observations. One way to address this issue is to use the negative binomial distribution. This distribution gives nearly the same mean predictions as the Poisson distribution in Eq. (2) but the prediction range based on the 25th and 75th percentiles captures 48% of the observations with an average range of 55. The range based on ± 1 standard deviations of the prediction errors using the Poisson distribution captures the observations 48% of the time with an average range of 53, results that are nearly the same as those based on the negative binomial distribution. Therefore, to be consistent with the hurricane and MH prediction, we still use the Poisson distribution for the ACE prediction

FIG. 5. Hindcast vs observed MH numbers and ACE from 1968 to 2017. The predictions from the leave-one-out approach (LOO) are also shown.

in Eq. (2) with the prediction range based on the standard deviations of the prediction errors.

Next we use a leave-one-out approach as in prior studies (e.g., Kim and Webster 2010) to compare our results with a no-skill metric based on the 5-yr running average prediction (using the most recent 5 years of observed MH or ACE as the prediction). This method uses all data for the period (1968–2017) with the exception of the year being predicted. A prediction is then made for the year left out, and this process is followed for every year in the period. It should be noted this method artificially increases skill (DelSole and Shukla 2009) since all years were used to develop the original model. Figure 5 shows that for the UA MH model, the MAE is 1.1 using this method versus 1.4 for the 5-yr running average (representing a 21% improvement), and the root-mean-square error (RMSE) is 1.6 versus 1.8 (an 11% improvement). We see greater improvements over the 5-yr running average prediction for ACE

with an MAE of 33 for our model versus 51 for the 5-yr running average (representing a 34% improvement), and an RMSE of 43 versus 62 (30% improvement).

Again, using a leave-one-out approach and comparing our results with the complete history of CSU's June predictions for both MH and ACE (www.tropical.colostate.edu/forecast-verification; not with CSU's current models), we see promising results. The MAE from 1990 to 2017 for MH is 1.5 for CSU versus 1.3 for UA, and for ACE, using years 1984–2017, the MAE is 49 for CSU versus 35 for UA.

To provide the most unbiased assessment of our model possible, we use our model to predict both MH numbers and ACE using data not used to create nor train the model. We tested the model's skill on 10 years of data, from 1967 to 1958, predicting backward as if in real time (e.g., using data from 1968 to 2017 for prediction in 1967, from 1967 to 2017 for 1966, etc.). It should be noted that in the late 1950s and early 1960s,

FIG. 6. A comparison of (a) MH number predictions (from 2000 to 2017) and (b) ACE predictions (from 2003 to 2017) from TSR, CSU, NOAA, UA, and the 5-yr running average baseline. The coefficients in the UA model were determined with data only available up to the time of prediction to best simulate a real-time result.

there may be a slight underestimate in ACE and MH before the availability of weather satellites (Vecchi and Knutson 2011). We then compare the results of a no-skill 5-yr running average (using the average MH and ACE results of the years 1968–72 to predict for 1967, for example). We see encouraging results, with our MH model producing an average error of 1.4 for the period, compared to the no-skill metric's 1.8. Similarly for ACE, our MAE of 39 is smaller than the no-skill metric's 46. Thus, our model performed better than the no-skill metric for years not used to select variables.

We also compare the results of our models with a diverse group of entities, including a university (CSU), a private company (TSR), and a government agency (NOAA, <http://www.cpc.ncep.noaa.gov/products/outlooks/hurricane-archive.php>). Using data available from 1968 only up to the year predicted as if it was a real-time prediction, we do a comparison for both MH and ACE (Fig. 6). Our MH model shows an improvement over all other models with an MAE of 1.06 from 2000 to 2017 (Fig. 6a), varying from a 25% improvement over NOAA to 37% over the 5-yr running average. Similarly the UA model also has the lowest MAE of the five models evaluated for ACE when comparing years from 2003 to 2017, with an MAE of 41 (Fig. 6b). This represents an improvement from 15% over TSR to 37% over the 5-yr average. Note that we do expect some inflation of skill using this method since those years were used to choose the variables for our models. Also, these results do not necessarily reflect the models these organizations currently use, but compare with what they used at the time when the forecasts were released.

Finally, as a test in real-time prediction, for the 2017 and 2018 hurricane seasons, we issued a forecast in early June using the MH and ACE models to complement our annual hurricane number prediction based on DZR, though the MH and ACE models were slightly different from those presented here for the 2017 forecast (we believe the adjustments made to the model used in 2017

will enhance the model to perform better on average). Seasonal hurricane predictions for 2017 issued in early June by many forecasting centers called for an average to a slightly above average year. For example, the predicted MH numbers were two (CSU), three (TSR), and from two to four (NOAA). We made a public prediction of six major hurricanes in early June 2017 using our models as constructed at the time, and if we used the UA model in this paper predicting with data from 1968 to 2016, our prediction would have been four. In fact, for MH (and ACE, discussed later), the UA models predicted higher activity than about 20 groups that submitted Atlantic forecasts online (<http://seasonalhurricanepredictions.org/>) in 2017. There were six MHs observed in 2017. In 2018, we predicted two major hurricanes using the same model shown in this paper, which agrees exactly with observations. This was identical to the accurate May/June predictions of CSU and NOAA, with TSR predicting one. However, this year (2018) was special in that the May/June predictions were better than the July/August predictions updated later by these groups, as all centers except for UA (we do not issue an August prediction) lowered their predictions.

We also made a similar prediction for ACE in early June 2017 and 2018. TSR predicted an ACE of 98, CSU predicted 100, and NOAA predicted between 69 and 143 for 2017. The early version of the model from which we issued the forecast predicted an ACE of 181 and the UA model presented here would predict 177, significantly higher than the other groups on the same seasonal prediction website just mentioned. We observed 226 units in 2017. For 2018, we predicted ACE of 96, and 129 units were observed, meaning an error of 33, very similar to the average error of 30 for our model. Once again, our prediction for ACE was nearly identical to the May/June predictions from CSU and NOAA, with TSR substantially lower at 43. However, like MH, this year (2018) was unique in that the July/August

predictions from these centers were worse than their May/June predictions.

4. Conclusions

Building on our previous seasonal prediction model for North Atlantic hurricane numbers, we have developed two new statistical models to predict both MH numbers and ACE each year, using data from March to May, for the North Atlantic basin. These models use SSTs, ZPWS to the $3/2$ power, and the MEI conditioned on the AMO. The similarities to, and differences from, DZR have also been highlighted.

Compared with observations from 1968 to 2017, our models have an MAE of 0.96 storms and 30 units for MH and ACE, respectively. These results beat the baseline model, which employs a linear regression of MH and ACE with our predicted hurricane numbers from DZR, demonstrating the value of our models over a simple extension of DZR for hurricanes. When compared to other centers using statistical–dynamical hybrid models, including TSR, CSU, and NOAA, our models perform better from 2000 (or 2003 in the case of ACE) through 2017 in terms of MAE in hindcast mode. For MH, our model's improvements vary from 25% over NOAA to 37% over the 5-yr running average. For ACE, the improvements vary from 15% over TSR to 37% over the 5-yr average. For the 2017 hurricane season, while NOAA, CSU, and TSR called for an average to a slightly above average year with their early June outlooks, we predicted a very active hurricane season, in much better agreement with observations. Our models also performed well for the MH and ACE prediction for the 2018 Atlantic hurricane season.

While the results using our observation-based statistical models are better than those based on hybrid models, this does not imply that the hybrid modeling approach itself is not good. In fact, with continued improvement in global dynamical seasonal prediction, various methods still need to be explored to innovatively combine statistical models with dynamical prediction models. We are currently working in this direction, which will also allow us to issue forecasts at different times of the year.

Acknowledgments. This work was supported by the Agnese Nelms Haury Program in Environment and Social Justice and the NASA MAP program (NNX14AM02G). The authors thank AOML for the historical hurricane data, ICOADS for the zonal pseudo–wind stress data, ESRL for the AMO and MEI data, and NCDC for the SST data. We also thank Thomas Galarneau, Anton Beljaars, Phil Klotzbach, and two anonymous reviewers for their helpful comments.

REFERENCES

- Bell, G. D., and Coauthors, 2000: Climate Assessment for 1999. *Bull. Amer. Meteor. Soc.*, **81** (6), S1–S50, [https://doi.org/10.1175/1520-0477\(2000\)81\[s1:CAF\]2.0.CO;2](https://doi.org/10.1175/1520-0477(2000)81[s1:CAF]2.0.CO;2).
- Blake, E. S., and D. A. Zelinsky, 2018: Hurricane Harvey. National Hurricane Center Tropical Cyclone Rep., 76 pp., https://www.nhc.noaa.gov/data/tcr/AL092017_Harvey.pdf.
- Cangialosi, J. P., A. S. Latta, and R. Berg, 2018: Hurricane Irma. National Hurricane Center Tropical Cyclone Rep., 111 pp., https://www.nhc.noaa.gov/data/tcr/AL112017_Irma.pdf.
- 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, <https://doi.org/10.1175/JCLI4953.1>.
- Davis, K., X. Zeng, and E. A. Ritchie, 2015: A new statistical model for predicting seasonal North Atlantic hurricane activity. *Wea. Forecasting*, **30**, 730–741, <https://doi.org/10.1175/WAF-D-14-00156.1>.
- DelSole, T., and J. Shukla, 2009: Artificial skill due to predictor screening. *J. Climate*, **22**, 331–345, <https://doi.org/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, [https://doi.org/10.1175/1520-0434\(1993\)008<0345:IERSP0>2.0.CO;2](https://doi.org/10.1175/1520-0434(1993)008<0345:IERSP0>2.0.CO;2).
- , and T. H. Jagger, 2006: Prediction models for annual U.S. hurricane counts. *J. Climate*, **19**, 2935–2952, <https://doi.org/10.1175/JCLI3729.1>.
- Emanuel, K. A., 1986: An air–sea interaction theory for tropical cyclones. Part I: Steady-state maintenance. *J. Atmos. Sci.*, **43**, 585–605, [https://doi.org/10.1175/1520-0469\(1986\)043<0585:AASITF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1986)043<0585:AASITF>2.0.CO;2).
- Goldenberg, S. B., and L. J. Shapiro, 1996: Physical mechanisms for the association of El Niño and West African rainfall with Atlantic major hurricane activity. *J. Climate*, **9**, 1169–1187, [https://doi.org/10.1175/1520-0442\(1996\)009<1169:PMFTAO>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<1169:PMFTAO>2.0.CO;2).
- , C. W. Landsea, A. M. Mestas-Núñez, and W. M. Gray, 2001: The recent increase in Atlantic hurricane activity: Causes and implications. *Science*, **293**, 474–479, <https://doi.org/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, [https://doi.org/10.1175/1520-0493\(1984\)112<1649:ASHFPI>2.0.CO;2](https://doi.org/10.1175/1520-0493(1984)112<1649:ASHFPI>2.0.CO;2).
- , C. W. Landsea, P. W. Mielke, and K. J. Berry, 1993: Predicting Atlantic basin seasonal tropical cyclone activity by 1 August. *Wea. Forecasting*, **8**, 73–86, [https://doi.org/10.1175/1520-0434\(1993\)008<0073:PABSTC>2.0.CO;2](https://doi.org/10.1175/1520-0434(1993)008<0073:PABSTC>2.0.CO;2).
- Harnos, D. S., J.-K. E. Schemm, H. Wang, and C. A. Finan, 2017: NMME-based hybrid prediction of Atlantic hurricane season activity. *Climate Dyn.*, <https://doi.org/10.1007/s00382-017-3891-7>, in press.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471, [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/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, <https://doi.org/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, <https://doi.org/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, [https://doi.org/10.1175/1520-0442\(1997\)010<0789:IOSSLP>2.0.CO;2](https://doi.org/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.
- Landsea, C. W., 1993: A climatology of intense (or major) Atlantic hurricanes. *Mon. Wea. Rev.*, **121**, 1703–1713, [https://doi.org/10.1175/1520-0493\(1993\)121<1703:ACOIMA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1993)121<1703:ACOIMA>2.0.CO;2).
- , and J. L. Franklin, 2013: Atlantic hurricane database uncertainty and presentation of a new database format. *Mon. Wea. Rev.*, **141**, 3576–3592, <https://doi.org/10.1175/MWR-D-12-00254.1>.
- Murakami, H., and Coauthors, 2016: Seasonal forecasts of major hurricanes and landfalling tropical cyclones using a high-resolution GFDL coupled climate model. *J. Climate*, **29**, 7977–7989, <https://doi.org/10.1175/JCLI-D-16-0233.1>.
- Palmen, E., 1948: On the formation and structure of tropical hurricanes. *Geophysica*, **3**, 26–38.
- 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, <https://doi.org/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, <https://doi.org/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, <https://doi.org/10.1029/2004GL021072>.
- Truchelut, R. E., and E. M. Staehling, 2017: An energetic perspective on United States tropical cyclone landfall droughts. *Geophys. Res. Lett.*, **44**, 12 013–12 019, <https://doi.org/10.1002/2017GL076071>.
- 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, <https://doi.org/10.1175/2010JCLI3810.1>.
- Villarini, G., and G. A. Vecchi, 2012: North Atlantic power dissipation index (PDI) and accumulated cyclone energy (ACE): Statistical modeling and sensitivity to sea surface temperature changes. *J. Climate*, **25**, 625–637, <https://doi.org/10.1175/JCLI-D-11-00146.1>.
- , and —, 2013: Multiseason lead forecast of the North Atlantic power dissipation index (PDI) and accumulated cyclone energy (ACE). *J. Climate*, **26**, 3631–3643, <https://doi.org/10.1175/JCLI-D-12-00448.1>.
- , B. Luitel, G. A. Vecchi, and J. Ghosh, 2016: Multi-model ensemble forecasting of North Atlantic tropical cyclone activity. *Climate Dyn.*, <https://doi.org/10.1007/s00382-016-3369-z>, in press.
- Webster, P. J., and S. Yang, 1992: Monsoon and ENSO: Selectively interactive systems. *Quart. J. Roy. Meteor. Soc.*, **118**, 877–926, <https://doi.org/10.1002/qj.49711850705>.
- WMO, 2008: Report from expert meeting to evaluate skill of tropical cyclone seasonal forecasts. World Meteorological Organization Tech. Doc. 1455, 27 pp.
- 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, <https://doi.org/10.1175/1520-0442-16.10.1601>.
- Zeng, X., M. Zhao, and R. E. Dickinson, 1998: Intercomparison of bulk aerodynamic algorithms for the computation of sea surface fluxes using the TOGA COARE and TAO data. *J. Climate*, **11**, 2628–2644, [https://doi.org/10.1175/1520-0442\(1998\)011<2628:IOBAAF>2.0.CO;2](https://doi.org/10.1175/1520-0442(1998)011<2628:IOBAAF>2.0.CO;2).