Skill, Predictability, and Cluster Analysis of Atlantic Tropical Storms and Hurricanes in the ECMWF Monthly Forecasts

SUZANA J. CAMARGO,[a] FRÉDÉRIC VITART,[b] CHIA-YING LEE,[a] AND MICHAEL K. TIPPETT[c]

[a] *Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York*
[b] *European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom*
[c] *Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York*

(Manuscript received 8 April 2021, in final form 10 August 2021)

ABSTRACT: In this paper we analyze Atlantic Ocean hurricane activity in the European Centre for Medium-Range Weather Forecasts (ECMWF) monthly hindcasts for the period 1998–2017. The main climatological characteristics of Atlantic tropical cyclone (TC) activity are considered at different lead times and across the entire ECMWF ensemble using three diagnostic variables: the number of tropical cyclones, the number of hurricanes, and the accumulated cyclone energy. The impacts of changing horizontal resolution and stochastic parameterization are clear in these diagnostic variables. The model skill scores for the number of tropical cyclones and accumulated cyclone energy by lead time are also computed. Using cluster analysis, we compare the characteristics of the forecast TC tracks with observations. Although four of the ECMWF clusters have similar characteristics to observed ones, one of the ECMWF clusters does not have a corresponding one in observations. We consider the predictability of each of these clusters, as well the modulation of their frequency by climate modes, such as the El Niño–Southern Oscillation and the Madden–Julian oscillation, taking advantage of the very large sample size of TC datasets in these hindcasts.

KEYWORDS: Tropical cyclones; Hindcasts; Model evaluation/performance

1. Introduction

For many years the modulation of tropical cyclone (TC) activity by the Madden–Julian oscillation (MJO) has been known (Li et al. 2018; Nakazawa 1986; Liebmann et al. 1994). Basically, when or soon after the MJO convective phase occurs over a region, there is an increase in the occurrence of TCs. In contrast, typically fewer TCs form over that location during the suppressed phase, since the large-scale environmental is less favorable for TC genesis (Camargo et al. 2009; Klotzbach 2014). The MJO–TC connection supports the possibility of skillful TC forecasts on subseasonal time scales, since the MJO is predictable on those time scales (Robertson et al. 2020). While the MJO is not the only tropical mode of intraseasonal variability that influences TC activity (Schreck et al. 2012), it is certainly the dominant predictable one on the subseasonal time scale.

In the last few years, there has been tremendous effort in the scientific community to improve forecasts on subseasonal-to-seasonal time scales (Vitart et al. 2017; White et al. 2017; Vitart and Robertson 2018; Pegion et al. 2019; Lang et al. 2020; Mariotti et al. 2020; Merryfield et al. 2020), among them, subseasonal forecasts of tropical cyclones (Camargo et al. 2019; Robertson et al. 2020). One of the main community efforts is the Subseasonal-to-Seasonal (S2S) prediction project, established jointly by the World Weather Research Program and the World Climate Research Program in 2013 and currently in Phase II with the goal of improving forecast skill and understanding the sources of subseasonal to seasonal variability (Vitart et al. 2017). The project created an extensive database containing subseasonal-to-seasonal forecasts and reforecasts from 11 operational and research centers.

One of the modeling centers in that project is the European Centre for Medium-Range Weather Forecasts (ECMWF). The ECMWF system in the last several years has substantially improved its forecast skill on subseasonal time scales, largely due to improvements in forecasting the MJO and its teleconnections (Vitart and Molteni 2010; Vitart 2014; Vitart et al. 2017). The improvements of the MJO in the ECMWF model have been discussed in detail in Vitart (2014), Hirons et al. (2013a,b). The main source of improvement were changes in the convective parameterization that were introduced in Bechtold et al. (2008). These studies showed that increasing the sensitivity of the deep convection scheme to environmental moisture modifies the relationship between precipitation and moisture in the model, and more moisture is able to build up. This moisture control on convection is key to simulating well the MJO. The current level of MJO forecast skill in dynamical systems is a remarkable scientific achievement, given that two decades ago, statistical models had higher skill than dynamical ones (e.g., Jones et al. 2000).

Given the strong modulation of TC activity by the MJO, it is not surprising that one focus of subseasonal forecasts has been TCs. Many papers have analyzed various aspects of the skill of the ECMWF system in forecasting TCs on seasonal and subseasonal time scales. For instance, the predictability of the ensemble tracks in the ECMWF 32-day ensemble forecasts for one or two seasons in the western North Pacific or North Atlantic Ocean was evaluated in Elsberry et al. (2010, 2011, 2014), Tsai et al. (2013). Vitart (2009) examined the impact of the MJO on TCs and landfall risk, and Vitart et al. (2010) compared the skill of the weekly predictions in the ECMWF system with a statistical model in the Southern Hemisphere. The predictability of

Corresponding author: Suzana J. Camargo, suzana@ldeo.columbia.edu

DOI: 10.1175/MWR-D-21-0075.1

the ECMWF system for TCs the North Atlantic and north Indian Ocean on intraseasonal time scales was discussed in Belanger et al. (2009, 2010, 2012). A multimodel comparison of high-resolution (7–10 km) climate models simulations of the MJO and its relationship with TCs in the boreal summer during 8 years as part of the Project Athena was presented in Satoh et al. (2012). More recently, as part of the MINERVA project, Manganello et al. (2019b,a) explored the predictability in the ECMWF seasonal forecasts for the North Atlantic.

One focus of the S2S project is tropical cyclones (TCs). The forecast skill of those models was examined in Lee et al. (2018) and Lee et al. (2020). The focus of Lee et al. (2018) was the probabilistic forecast skill of basin-wide TC genesis. Lee et al. (2020) focused on TC frequency occurrence in regional subbasins, as well as intensity using the accumulated cyclone energy (ACE). The ECMWF clearly appeared as the most skillful model in the measures considered within the subset of S2S models analyzed. However, we should note that the TCs in the S2S project were not tracked using the original model resolution, but interpolated to a lower horizontal resolution common grid.

Many other studies have examined the ability of models to forecast TCs on subseasonal time scales (e.g., Camargo et al. 2019; Robertson et al. 2020). In many cases the analysis was done for specific events, as in Fu and Hsu (2011), who showed that a conventional atmosphere–ocean model was able to predict a specific event in the north Indian Ocean with a lead time of 2 weeks. Similarly, Xiang et al. (2015) showed that the Geophysical Fluid Dynamics Laboratory (GFDL) system predicted the landfall location and timing of Hurricane Sandy with 1 week of lead time and Typhoon Haiyan's landfall in the Philippines with 2 weeks of lead time. The ability of the GFDL High Resolution Atmospheric Model (HiRAM) to reproduce the subseasonal modulation of TCs in the Gulf of Mexico and western Caribbean Sea was shown in Gao et al. (2017). In a follow-up paper, Gao et al. (2019) showed that using a nested grid with enhanced resolution improved the monthly forecast skill of HiRAM in the North Atlantic.

More recently, Camp et al. (2019) showed that the Australian Bureau of Meteorology seasonal forecasting system (ACCESS-S1) has high skill for predictions of the MJO with lead times of 30 days and is able to reproduce the TC modulation by the MJO in the Southern Hemisphere. Gregory et al. (2019) evaluated the forecasts in that system for 2017–18 and showed that it skillfully predicted the formation of major cyclone events with lead times greater than 2 weeks.

In this paper, we explore the climatological characteristics, skill, and predictability of North Atlantic TCs in the ECMWF monthly hindcasts for the period 1998–2017. We consider how well the hindcasts are able to reproduce the main climatological and track characteristics of Atlantic TCs. The large number of the TCs in the hindcasts makes it ideal to explore the modulation of the TC activity by various climate modes (El Niño–Southern Oscillation, Madden–Julian oscillation, and North Atlantic Oscillation). The predictability of the system by lead time and cluster is also explored.

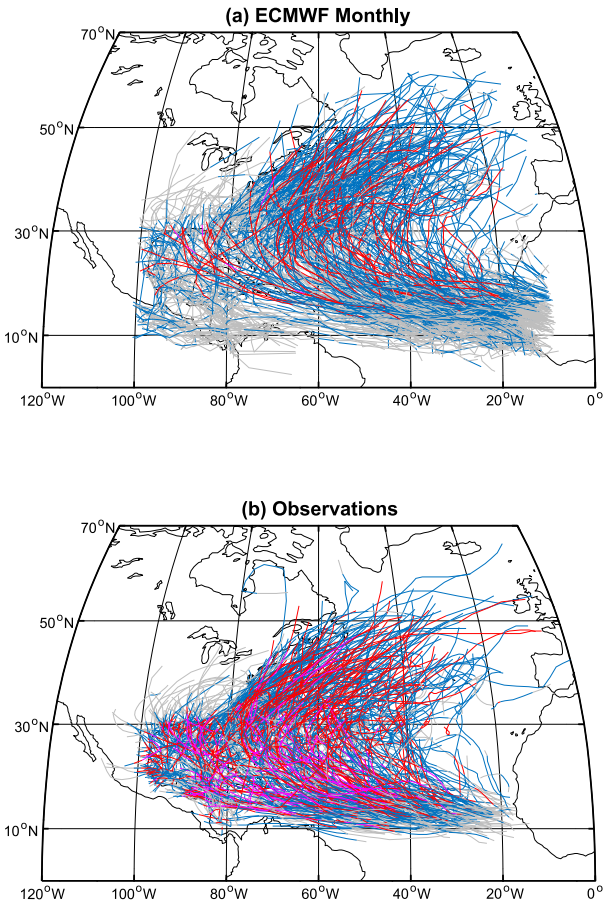

FIG. 1. Tracks of Atlantic tropical storms and hurricanes: (a) ECMWF monthly hindcasts for E2 in weeks 1 and 2 and (b) observations 1950–2018 (822 storms). In (a), 822 tracks were chosen randomly among a total of 1037 tracks so as to have the same number of tracks as in (b). Colors indicate storm intensity along track: tropical depression (gray), tropical storm (blue), hurricanes (red; categories 1 and 2), and major hurricanes (magenta; categories 3–5).

Section 2 describes the datasets and methods used in our analysis. Our results are presented in section 3, and we conclude with a discussion in section 4.

2. Data and methods

a. Atlantic tropical cyclones

The observed best track dataset for North Atlantic tropical storms and hurricanes was obtained from the National Hurricane Center (Landsea and Franklin 2013). Only storms that reach 35 kt (tropical storm strength; 1 kt $\approx 0.5\ \mathrm{m\ s^{-1}}$) are included in our analysis, including those labeled as subtropical in the NHC dataset. Storms that reach at least 64 kt are defined as hurricanes. The TC diagnostics considered in our analysis are the number of tropical cyclones (NTC), that is, the number of tropical storms and hurricanes; the number of hurricanes (NHUR); and the accumulated cyclone energy (ACE). The ACE (Bell et al. 2000) for each storm is calculated

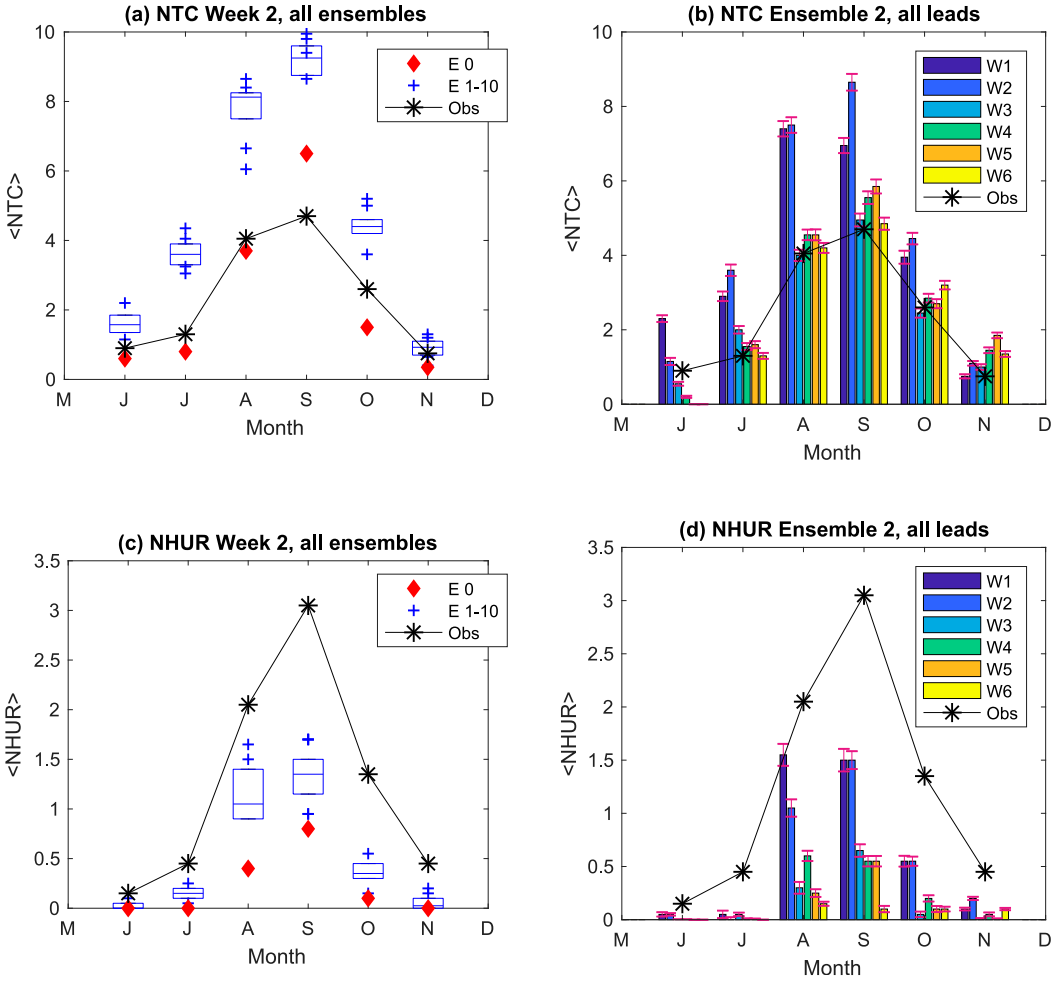

FIG. 2. (a) Number of tropical cyclones (NTC) in week 2 for all ensembles (control E0 and ensemble members 1–10: E1–E10) and in observations. (b) NTC for ensemble 2 for all leads [weeks 1–6 (W1–W6)] and in observations. (c) Number of hurricanes (NHUR) in week 2 for all ensembles and in observations. (d) NHUR for ensemble 2 for all leads and in observations. Also shown on top of each bar in (b) and (d) is the spread across the ensembles estimated using bootstrap.

by summing the square of the maximum wind speeds for the 6-hourly snapshots that reach at least 35 kt.

b. *El Niño–Southern Oscillation*

The state of the El Niño–Southern Oscillation (ENSO) is defined based on the Niño-3.4 index (Bamston et al. 1997) in August–October (ASO), the peak of the Atlantic hurricane season, available from the Climate Prediction Center (https://www.cpc.ncep.noaa.gov/) from 1950 to the present, based on Extended Reconstructed Sea Surface Temperature, version 5 [ERSST v5; Huang et al. (2017)]. Following Goddard and Dilley (2005) and Camargo et al. (2007a), in observations we define the top 25% of events of the Niño-3.4 ASO distribution in the period 1950–2018 as El Niño events and the bottom 25% of events as La Niña events.

For the ECMWF hindcast period, we consider the top five observed warm and cold ASO ENSO seasons:

- El Niño: 2002, 2004, 2006, 2009, and 2015;
- La Niña: 1998, 1999, 2007, 2010, and 2011.

The other 11 seasons are considered to be neutral. Since this is a small number of ENSO events with which to analyze observed TCs, we also consider the top 17 warm and cold ASO ENSO seasons in the period 1950–2018:

- El Niño: 1957, 1963, 1965, 1969, 1972, 1982, 1986, 1987, 1991, 1994, 1997, 2002, 2004, 2006, 2009, 2015, and 2018;
- La Niña: 1950, 1954, 1955, 1956, 1961, 1964, 1970, 1971, 1973, 1975, 1988, 1995, 1998, 1999, 2007, 2010, and 2011.

The other 35 seasons are considered to be neutral.

Besides the top observed ENSO events for the period 1998–2017, listed above, we also considered the Niño-3.4 values simulated by the ECMWF hindcasts to determine the ENSO phase in our analysis. As the ECMWF hindcasts are not available for the whole ASO season, we consider the Niño-3.4

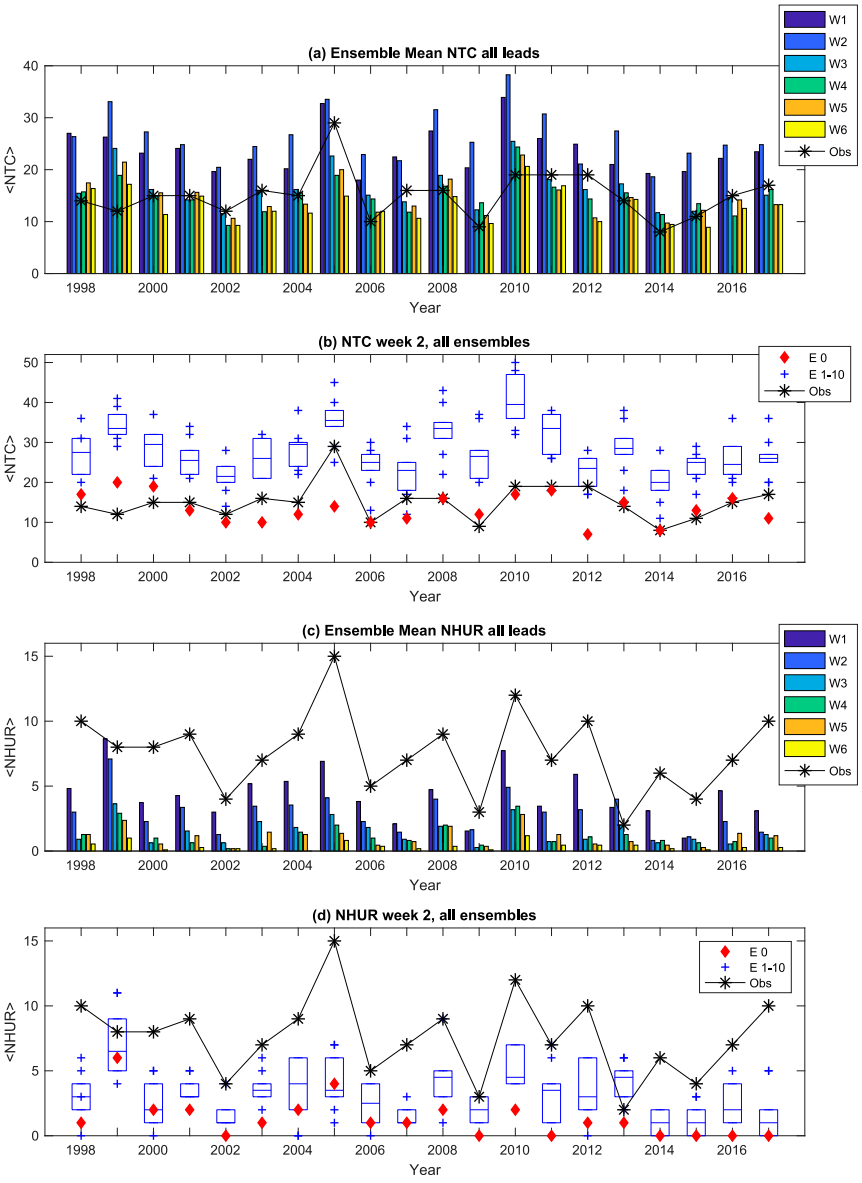

FIG. 3. (a) Ensemble mean NTC per season for all leads [weeks 1–6 (W1–W6)] and in observations. (b) NTC per season for week 2 for all ensembles (E0; E1–E10) and in observations. (c) Ensemble mean NHUR per season for all leads and in observations. (d) NHUR per season for week 2 for all ensembles and in observations.

hindcast values for the date of the storm formation, with the thresholds of the El Niño and La Niña phases determined by the 25th and 75th percentiles of the distribution of the model Niño-3.4 for that month. We do not expect large differences between the two ENSO definitions, given the lead time, but they will not be the same given the monthly versus seasonal definitions, in particular in the late-spring and early-summer months.

c. *North Atlantic Oscillation*

The monthly values of the North Atlantic Oscillation (NAO) index are obtained from the Climate Prediction Centre (https://www.cpc.ncep.noaa.gov/) and are available from 1950 to the present. The NAO index is calculated following the procedure of Barnston and Livezey (1987) and standardized by the 1981–2010 climatology. Similarly to ENSO phases, NAO positive (NAO+) and NAO negative (NAO−) observed phases are defined as the top and bottom quartiles of the distribution during the 1950–2018 period. Here we consider NAO for the months of May and June (MJ), following Kossin et al. (2010) who found a relationship between the NAO in those months with the Atlantic hurricane activity using the same cluster analysis considered here.

The NAO analysis for the ECMWF system was performed in two different ways. The first considers the NAO value at

TABLE 1. Statistics of observations and the ECMWF hindcasts for tropical cyclones (NTC), number of hurricanes (NHUR), and ACE ($\times 10^{-5}$) per season. For the ECMWF ensemble mean and for the control simulation (E0), the mean, standard deviation, and minimum and the maximum values across all years for each lead time are shown. For observations we show the mean, standard deviation, minimum, and maximum values for the period 1950–2018, as well as for the same period of the hindcasts (1998–2017).

	NTC				NHUR				ACE			
Lead	Mean	Std dev	Min	Max	Mean	Std dev	Min	Max	Mean	Std dev	Min	Max
ECMWF ensemble mean												
Week 1	23.7	7.0	6	44	4.3	2.8	0	14	23.0	10.6	2.2	51.6
Week 2	26.4	7.9	7	50	2.9	2.2	0	11	21.3	9.4	3.9	56.9
Week 3	16.3	5.8	3	35	1.5	1.6	0	7	15.6	7.4	2.2	40.5
Week 4	14.9	5.2	3	28	1.2	1.3	0	6	13.6	6.2	2.5	33.2
Week 5	14.7	5.8	2	35	1.1	1.2	0	4	12.6	5.7	1.4	28.7
Week 6	13.0	4.9	3	27	0.4	0.6	0	3	8.1	3.4	1.1	19.4
ECMWF control (E0)												
Week 1	10.9	4.1	6	21	1.8	2.0	0	8	11.0	6.5	2.2	27.5
Week 2	13.5	3.7	7	20	1.3	1.5	0	6	13.3	5.6	4.0	23.1
Week 3	9.0	3.0	3	15	0.8	0.8	0	3	10.2	5.1	2.2	22.9
Week 4	8.6	3.2	3	14	0.5	0.8	0	3	9.5	4.8	2.5	20.7
Week 5	7.2	4.4	2	18	0.5	0.8	0	2	7.2	5.4	1.4	21.0
Week 6	7.2	3.6	3	14	0.2	0.4	0	1	4.4	2.4	1.1	9.4
Observations												
1998–2017	15.1	4.6	8	29	7.6	3.1	2	15	15.2	6.5	5.2	28.1
1950–2018	11.9	4.2	4	29	6.4	2.7	2	15	12.0	6.2	1.9	28.1

the time of TC genesis. The NAO+ and NAO− phases are defined based on the 25th and 75th percentiles, respectively, of all NAO values at the time of TC genesis in the hindcast, i.e., the NAO will correspond to values varying from June to November, as predicted by the system, depending on the TC genesis time. The second considers the observed MJ NAO values, with NAO+ and NAO− defined on the basis of the 25th and 75th percentiles, respectively, of the NAO values in the period 1998–2017. Seasons with NAO+ in MJ are 1999, 2000, 2002, 2009, and 2013; seasons with NAO− in MJ are 1998, 2008, 2010, 2012, and 2014. We only obtained statistically significant results using the second option, and we show only results from that case below.

d. Madden–Julian oscillation

The MJO is defined based on the phase of the real-time multivariate MJO (RMM) index developed by Wheeler and Hendon (2004) using observed wind and outgoing longwave radiation data. The observed RMM data are available online through the Australian Bureau of Meteorology (http://www.bom.gov.au/climate/mjo/) for the period 1979 to the present. The MJO phase in the ECMWF system was obtained by applying the same method to the model data.

e. ECMWF monthly hindcasts

The ECMWF monthly forecasting system has been run routinely since October 2004 (Vitart 2003, 2004), with regular updates since then. This system fills the gap between the two other operational forecasting systems at ECMWF, namely, the medium-range weather and seasonal forecast forecasting systems. The current version of the ECMWF monthly forecasting system is described in detail in Vitart et al. (2008). Although the ECMWF monthly forecast system was originally set up differently from the ECMWF medium-range ensemble system, the current version was developed with the goal of "seamless" merging the medium-range and monthly coupled forecasting systems. A variable resolution approach reduces the computational cost of extending the medium-range forecasts and reforecasts (hindcasts) to monthly lead times (Buizza et al. 2007). From day 0 to day 10, the atmospheric model is run with a T399 horizontal resolution and L62 vertical resolution, forced with persisted SST anomalies. From day 11 to day 46, the horizontal resolution is T255 and the atmospheric model is coupled to the ocean model.

Hindcasts matching the forecast start dates are produced 2 times per week on Thursdays and Sundays. Here we analyze 11 ensemble members of the hindcasts for the period 1998–2017. The hindcasts were used to calibrate the forecasts for 2018 and are for the period June–November of each year. The first hindcast had a start date of 4 June 1998, and the last hindcast had a start date of 1 November 2017. Only target periods for the months June–November (hurricane season) of each year were included in our analysis. These hindcasts were produced in 2018, with the version of the system from cycle 45r1 (https://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model). The first ensemble member (E0) is the control simulation, and ensemble members 1–10 are produced using stochastic perturbations of the model physics (Leutbecher et al. 2017; Ollinaho et al. 2017) to represent the model uncertainty. The impact of stochastic perturbations on TC forecasts has been discussed in Lang et al. (2012). They showed that the stochastic parameterization leads to a large TC track and intensity spread that matches well the ensemble mean average error. More recently, Vidale et al. (2021) compared the impact of the stochastic physics with that of model resolution in two global climate

TABLE 2. Correlations (Corr) and rank correlations (Rank) between ECMWF and observations, as well as the root-mean-square error (RMS) (normalized by the observed standard deviation) of the ECMWF hindcasts for NTC, NHUR, and ACE per season. Significance is indicated at the 99% level with boldface type and an asterisk and at the 95% level with boldface.

	NTC			NHUR			ACE		
Lead	Corr	Rank	RMS	Corr	Rank	RMS	Corr	Rank	RMS
Week 1	**0.76***	**0.71***	37.8	**0.68***	**0.70***	20.9	**0.73***	**0.74***	6.46
Week 2	**0.53**	0.40	49.5	0.41	**0.50**	29.9	**0.48**	**0.52**	6.33
Week 3	**0.59***	**0.59***	5.6	0.44	0.42	39.1	**0.52**	**0.58***	4.04
Week 4	**0.49**	**0.48**	0.4	**0.56**	**0.61***	40.9	**0.50**	**0.58***	4.06
Week 5	**0.51**	**0.47**	1.3	**0.59***	**0.61***	41.5	**0.58***	**0.62***	3.95
Week 6	0.44	**0.46**	8.8	**0.53***	**0.47**	46.1	**0.49**	**0.53***	6.17

models, one of which, EC-Earth3, is based on the ECMWF system. The physics parameterization is not the only aspect of the model that affects TC forecasts—horizontal resolution and ocean coupling are also fundamental for skillful TC forecasts (Magnusson et al. 2019). Relevant for these time scales, ensemble perturbations based on stochastic physics have been shown to improve the MJO skill in the ECMWF system (Lock et al. 2019).

The tropical storms and hurricanes in the ECMWF system are tracked using the tracking algorithm developed by Vitart (Vitart et al. 1997, 2001; Vitart and Anderson 2001). The simulated storms are required to satisfy specific criteria to be tracked: maximum relative vorticity above a threshold of $3.5 \times 10^{-5}\,s^{-1}$, a local minimum sea level pressure and a warm core in the upper troposphere, and duration of at least 2 days. From all of the storms tracked in the hindcasts, we only consider storms whose maximum wind speed reaches at least $17\,m\,s^{-1}$ in weeks 1 and 2 and at least $15\,m\,s^{-1}$ in weeks 3–6. This differing threshold reflects the change in model resolution after day 10.

f. *Cluster analysis*

The cluster analysis applied here was developed by Gaffney (2004) and is described in detail in Gaffney et al. (2007). It has been applied to observed tropical cyclone tracks in various regions: the western North Pacific (Camargo et al. 2007b,c), the eastern North Pacific (Camargo et al. 2008), and the Southern Hemisphere (Ramsay et al. 2012), as well as extensively to climate model TC tracks (Camargo 2013; Daloz et al. 2015;

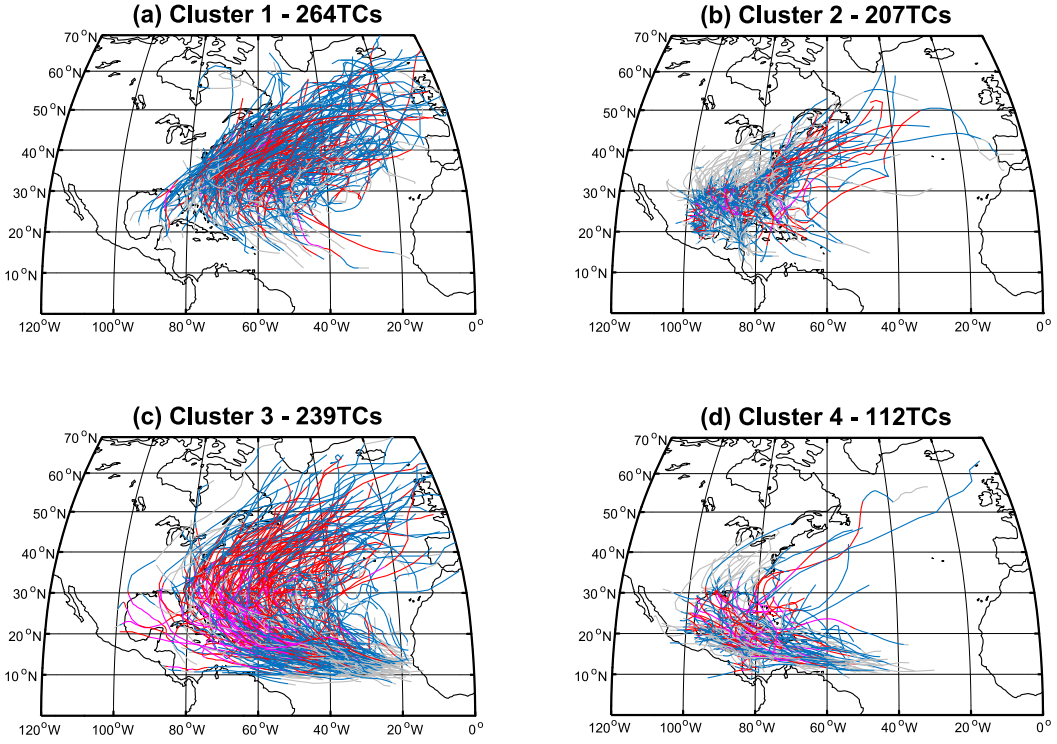

FIG. 4. Observed tracks of Atlantic tropical storms and hurricanes for the period 1950–2018 for four clusters. Colors indicate storm intensity along track: tropical depression (gray), tropical storm (blue), hurricanes (red; categories 1 and 2), and major hurricanes (magenta; categories 3–5).

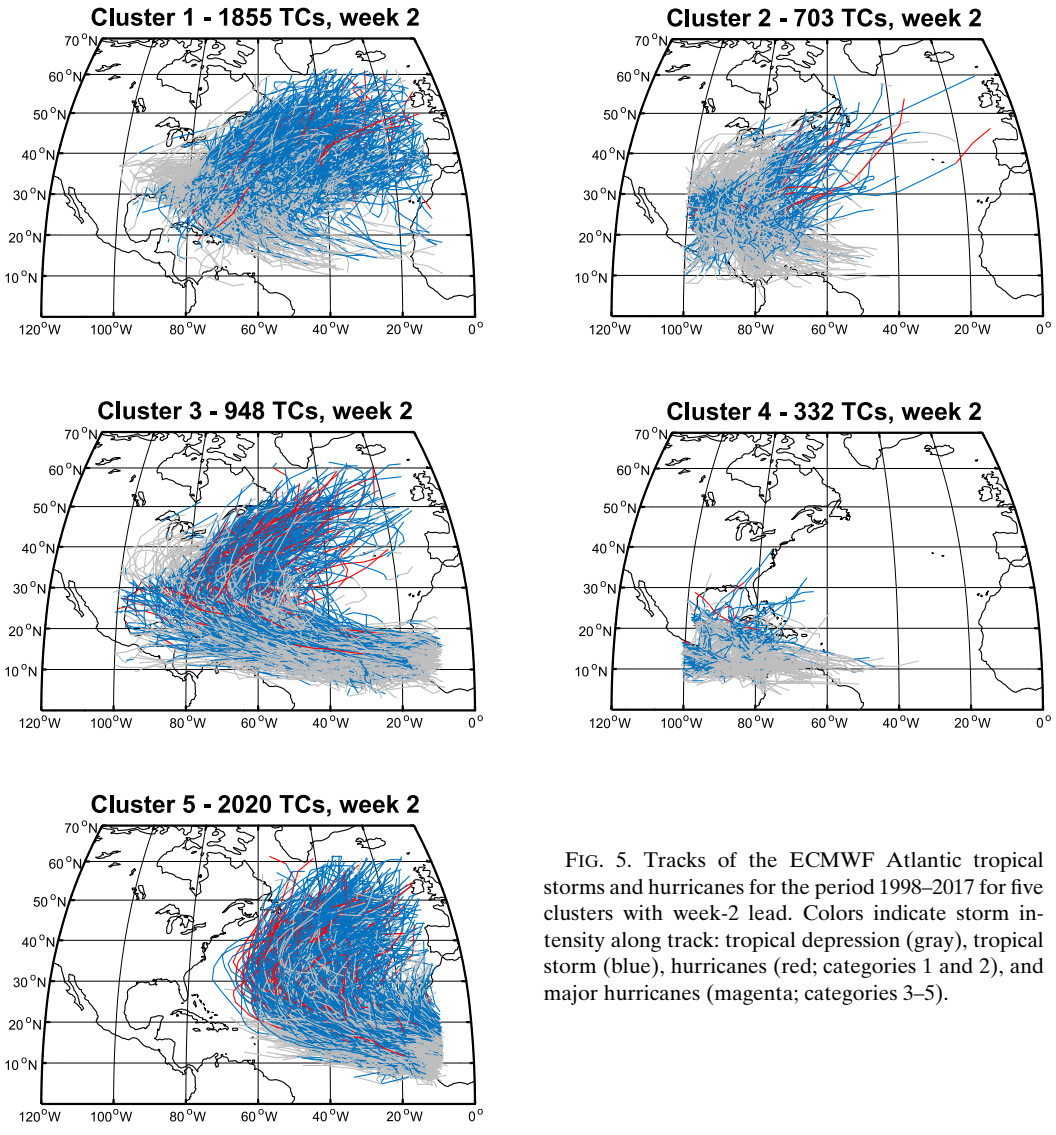

FIG. 5. Tracks of the ECMWF Atlantic tropical storms and hurricanes for the period 1998–2017 for five clusters with week-2 lead. Colors indicate storm intensity along track: tropical depression (gray), tropical storm (blue), hurricanes (red; categories 1 and 2), and major hurricanes (magenta; categories 3–5).

Nakamura et al. 2017; Ramsay et al. 2018; Patricola et al. 2018; Bell et al. 2019). In the case of the Atlantic, the cluster analysis of observed tracks was discussed in depth in (Kossin et al. 2010; Kozar et al. 2012; Boudreault et al. 2017). The method is a probabilistic curve-clustering, which groups together tracks with similar characteristics, using linear regression mixture models of second-order polynomials that are fitted to the tracks. The cluster analysis was run for 101 times (one time with the tracks input in the original order, and 100 times with tracks input in random order), for both observed and model tracks to ensure the stability of the cluster analysis, as discussed in (Camargo et al. 2008). The case with the smallest log-likelihood value was selected to assign each track to a particular cluster.

The cluster analysis of observed TC tracks is an extension of Kossin et al. (2010) and Boudreault et al. (2017) to the period 1950–2018. As in these previous studies, we consider four clusters ($k = 4$) for the observed TC tracks. However, for the TC tracks from the ECMWF hindcasts, we consider five clusters, four of which are very similar to the observed ones, and an additional cluster, which is not present in observations and is associated with systematic model biases, as will be discussed below. Cluster analysis is a powerful tool to identify biases in model tracks, as well as the occurrence of nonobserved track types (Camargo 2013; Daloz et al. 2015; Nakamura et al. 2017; Ramsay et al. 2018).

g. *Skill scores and model bias*

To examine how the hindcasts reproduce the observed interannual variability of NTC, NHUR, and ACE, we calculate the Pearson correlation, the Spearman rank correlation, and the root-mean-square error between the model and observations by lead time and by cluster.

Following Lee et al. (2018) and Lee et al. (2020), we use the Brier skill score (BSS) to determine the skill of NTC and the ranked probability skill score (RPSS) to evaluate the ACE

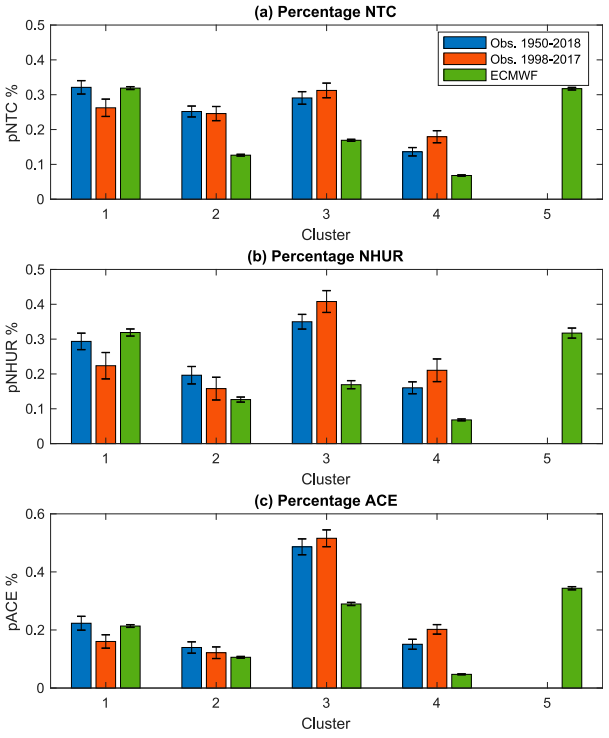

FIG. 6. Percentage of NTC, NHUR, and ACE per cluster in observations (1950–2018 and 1997–2018), and the ensemble mean ECMWF monthly hindcasts for all leads. Errors estimated using a bootstrap procedure with 1000 bootstrap samples.

skill. Also similarly to Lee et al. (2020), we will consider two reference forecasts for the skill scores. Our first reference is a constant mean climatology for the months from June to November (BSS_c), and the second one is a monthly varying climatology (BSS_m). In both cases the observational data used to construct the climatology match the forecast targets over the entire period and thus include data that would not available at the time of forecast issuance. Skill scores based on the first reference forecast can be interpreted as a measure of the model skill in forecasting the value of NTC or ACE, and credit is given for correctly representing the annual cycle. Skill scores based on the second reference forecast measure the model skill in forecasting anomalies with respect to the monthly climatology, and models do not receive credit for correctly representing the annual cycle. When computing RPSS for ACE, a quantile matching is applied to ECMWF storm intensity to address its underestimation of storm intensity, which will be discussed in the next section.

When examining the regional skill of TC occurrence forecasts, we use the threat score, the Brier score and the Brier skill score, using both climatologies (constant and monthly varying) as reference forecasts (BSS_c and BSS_m).

We calculate the model biases using the difference between the model and reanalysis climatology in the period 1998–2017 in the months from June to November. The reanalysis used was the ERA5 reanalysis produced by ECMWF (Hersbach et al. 2020). We examined the biases of many fields but only show here 850-hPa winds and 500-hPa geopotential winds, which are relevant to our results.

3. Results

a. *Monthly ECMWF Atlantic TC climatology*

We first examine the tracks of the observed and ECMWF Atlantic TCs. Given the very large number of TCs in the ECMWF monthly dataset, we only show the tracks of one of the ECMWF ensemble members (ensemble 2), chosen randomly, for week 2 in Fig. 1a. Figure 1b shows the tracks of all

TABLE 3. Statistics of observations and the ECMWF hindcasts for NTC, NHUR, and ACE ($\times 10^{-6}$) in each cluster per season. For the ECMWF, the total and the percent of TCs in each cluster, as well as the mean and the standard deviation across all years, ensembles, and leads for each cluster are shown. For observations, we show the total, percent, mean, and standard deviation across all years for each cluster for two time periods.

	NTC				NHUR				ACE			
Cluster	Total	%	Mean	Std dev	Total	%	Mean	Std dev	Total	%	Mean	Std dev
						ECMWF						
1	8743	35%	6.6	0.6	470	19%	0.4	0.1	448.9	21%	0.34	0.05
2	3158	13%	2.4	0.9	252	10%	0.2	0.1	222.7	11%	0.17	0.07
3	4145	17%	3.1	1.2	863	34%	0.6	0.5	608.6	29%	0.46	0.23
4	1318	5%	1.0	0.44	42	2%	0.03	0.03	99.1	5%	0.07	0.03
5	7426	30%	5.6	1.2	880	35%	0.7	0.4	722.5	34%	0.55	0.20
						Observations for 1998–2017						
1	79	26%	3.9	1.7	34	22%	1.7	1.3	4.9	16%	0.24	0.16
2	74	25%	3.7	2.2	24	16%	1.2	1.1	3.7	12%	0.18	0.17
3	94	31%	4.7	1.7	62	41%	3.1	1.5	15.6	52%	0.78	0.50
4	54	18%	2.7	1.6	32	21%	1.6	1.3	6.1	20%	0.31	0.27
						Observations for 1950–2018						
1	264	32%	3.8	1.7	130	29%	1.9	1.2	18.4	22%	0.27	0.17
2	207	25%	3.0	1.9	87	20%	1.3	1.2	11.5	14%	0.17	0.16
3	239	29%	3.5	2.1	155	35%	2.2	1.6	40.1	49%	0.58	0.47
4	112	14%	1.6	1.4	71	16%	1.0	1.1	12.4	15%	0.18	0.21

TCs in observations in the period 1950–2018. The first thing that is apparent is that the ECMWF storms are much weaker than the observed storms, with too few category-1–2 hurricanes and too many tropical storms in comparison with observations. These intensities biases are present despite the relatively high model resolution in weeks 1–2. Furthermore, there are also some spatial biases, which will be explored using cluster analysis.

The climatological number of TCs (NTC) and hurricanes (NHUR) per month are shown in Fig. 2. The left panels show the mean number of TCs and hurricanes separately for the first ensemble member (control) and for ensemble members 1–10 at week 2, while the right panels show the mean numbers by lead time. Some interesting patterns are clear from this figure. First, in Figs. 2a and 2b, there is large difference in the mean number of TCs and hurricanes between the control and the ensemble members that were generated using stochastic perturbations of the model physics. For instance in August and September, the mean number of TCs for the control is approximately 4, while the value ranges between about 6 and 8 for the other ensemble members. This difference is clear for both TCs and hurricanes, especially during the peak season (August–October, ASO). While the number of TCs for the deterministic control run (E0) at week 2 is close to the observed number, the perturbed ensemble members clearly overproduce TCs. In the case of hurricanes, the perturbed ensemble members produce more hurricanes than the control, but all ensemble members produce too few hurricanes in comparison with observations, especially during the peak season in September, when none of the ensemble members has a mean number of hurricanes close to the observed. For hurricanes, the bias in the deterministic ensemble member is much larger than in the stochastic ones. Recently, Vidale et al. (2021) showed that the use of stochastic parameterization enhances the simulation of TCs in a way similar to the increase in model horizontal resolution. They showed that the use of stochastic parameterization increased the TC frequency by approximately 30%, removed some of the climatological model biases and improved the simulation of the TC seasonal cycle. The main reason for these changes is that the stochastic parameterization generates a higher number of seeds, as well as a higher genesis rate, leading to a more efficient transition of seeds into TCs.

Figures 2b and 2d showed a clear decrease in the number of TCs and hurricanes in the model with longer lead times. There is a distinct difference between weeks 1–2 and 3–6. This difference is probably associated with the degradation of the model horizontal resolution at day 11. There is well-known relationship between model resolution and the ability of the models to simulate intense storms (Davis 2018; Moon et al. 2020a,b), and the global climate model TC climatology (genesis and intensity) is known to improve with model resolution (Murakami and Sugi 2010; Shaevitz et al. 2014; Camargo et al. 2020, e.g.).

There is a seasonal dependence on the reduction of TC numbers by lead time at Week 3. While the number of TCs reduces with lead time in June and July, that is not the case during the peak season (ASO), when the number of TCs remains constant with lead time, and slightly increases with lead time in November. In the case of hurricanes, besides the clear decrease after week 3, there are no other clear signals because of the very low number of occurrences for longer leads.

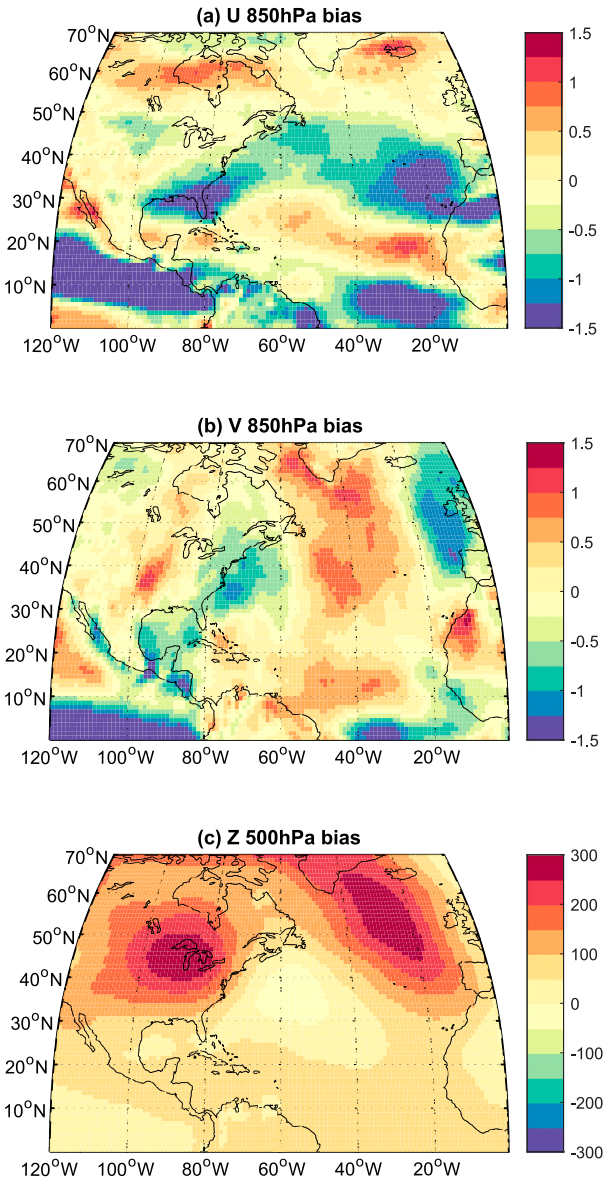

FIG. 7. Biases in the climatological fields of low-level (850 hPa) (a) zonal and (b) meridional winds (m s^{-1}) and for (c) 500-hPa geopotential heights (m^2 s^{-2}).

Figure 3a shows the ensemble mean number of TCs per season for different lead times, with the week-2 ensemble spread shown in Fig. 3b. Similar figures for the mean number of hurricanes are shown in Figs. 3c,d. The observed number of TCs and hurricanes are also shown in the figure. The statistics of NTC and NHUR per season across all years for the ensemble mean and observations are given in Table 1. As noted above for specific months, there is a large difference in the mean number of TCs and hurricanes per season for weeks 1 and 2 as

TABLE 4. Statistics of observations and ECMWF for NTC, NHUR, and ACE ($\times 10^{-6}$) in each cluster per season for El Niño and La Niña. The ENSO state is based on the value of Niño-3.4 in the ECMWF ensemble at the time of storm genesis. The percentages shown for the cluster are based on the total value for each cluster. Also shown are the statistics for 5 and 17 ENSO events in observations. Boldface type indicates statistical significance of the sum of all years in each cluster for the variable. An asterisk indicates statistical significance of the total per seaon in each cluster for the variable. Statistical significance was done using the Kolmogorov–Smirnov hypothesis test. For observations, statistical significance was determined by correlation with Niño-3.4 in ASO at the 95% significance level.

	NTC				NHUR				ACE			
	Total		%		Total		%		Total		%	
Cluster	EN	LN	EN	LN	EN	LN	EN	LN	EN	LN	EN	LN
ECMWF monthly—Niño-3.4 model												
Total	**5938***	**7514***	0.24	0.30	**520**	**885**	0.21	0.35	**469.8***	**679.1***	0.22	0.32
1	**2278**	**2473**	0.26	0.28	121	153	0.26	0.33	**117.0**	**131.3**	0.26	0.29
2	**624***	**1157***	0.20	0.37	**44**	**99**	0.17	0.39	**41.2***	**82.8***	0.19	0.37
3	**858***	**1327***	0.21	0.32	**143**	**337**	0.17	0.39	**116.1**	**211.5**	0.19	0.35
4	**276***	**514***	0.21	0.39	9	16	0.21	0.38	**21.1***	**39.3***	0.21	0.40
5	1902	2043	0.26	0.28	**203**	**208**	0.23	0.32	**174.3**	**214.1**	0.24	0.30
ECMWF monthly—5 ENSO events												
Total	**4639***	**6808***	0.21	0.31	**380***	**836***	0.17	0.37	**361.1***	**632.7***	0.19	0.34
1	**1753***	**2096***	0.24	0.28	**77***	**134***	0.19	0.34	**87.2***	**113.4***	0.23	0.30
2	**462***	**1083***	0.17	0.39	**34***	**88***	0.15	0.39	**30.9***	**77.2***	0.16	0.39
3	**661***	**1256***	0.18	0.34	**106***	**329***	0.13	0.41	**87.9***	**203.5***	0.16	0.37
4	**202***	**451***	0.18	0.40	8	16	0.20	0.40	**16.0***	**35.5***	0.18	0.41
5	**1561***	**1922***	0.23	0.29	**155***	**269***	0.19	0.33	**139.1***	**203.1***	0.21	0.31
Observations—5 ENSO events												
Total	57	80	0.19	0.27	**25**	**44**	0.16	0.29	57.3	83.6	0.19	0.28
1	20	15	0.25	0.19	8	4	0.24	0.12	14.4	6.0	0.30	0.12
2	10	21	0.14	0.28	2	8	0.08	0.33	5.5	11.2	0.15	0.30
3	**17**	**30**	0.18	0.32	**10**	**23**	0.16	0.37	27.8	51.7	0.18	0.33
4	10	14	0.19	0.26	5	9	0.16	0.28	9.6	14.7	0.16	0.24
Observations—17 ENSO events												
Total	**167**	**233**	0.20	0.28	**82**	**129**	0.19	0.29	**155.7**	**253.4**	0.19	0.31
1	68	63	0.26	0.24	32	25	0.25	0.19	47.3	39.3	0.26	0.21
2	38	64	0.18	0.31	13	29	0.15	0.33	21.3	37.8	0.18	0.33
3	**43**	**71**	0.18	0.30	**26**	**51**	0.17	0.33	**68.7**	**135.1**	0.17	0.34
4	**18**	**35**	0.16	0.31	**11**	**24**	0.15	0.34	18.5	41.2	0.15	0.33

compared with longer lead times. The model generates too many TCs per season in weeks 1 and 2 (Fig. 3a), with the observations typically below or at the minimum of the ensemble spread for week 2 (Fig. 3b). For longer lead times (weeks 3–6), there is a much better match between the mean observed NTC and the model ensemble mean. While the number of hurricanes produced by the model is larger in weeks 1 and 2 (Fig. 3c) than in longer leads, even at week 2, the number of hurricanes across the full ensemble is much lower than the observed values. Table 1 shows the statistics of NTC and NHUR for the ensemble mean, as well as the control simulation. The stark differences between the ensemble mean and the control simulation statistics are very clear, for both variables, emphasizing the high number of storms generated using the stochastic parameterization.

Table 1 also shows the statistics of ACE per lead time in the ECMWF system, as well as in observations. ACE is an integrated quantity, including number, duration and intensity. For the control run (E0), the mean ACE is lower than the observed ACE for all leads, decreasing with lead time. In contrast, the ensemble mean ACE is much higher than the observed ACE in weeks 1 and 2, due to excessive number of TCs in the model at those leads. Similarly to NTC, there is a sharp decrease in ACE at week 3, with the higher grid spacing used in the model. The mean ACE values are close to the observed values for leads 3–5, with another large decrease in week 6.

To capture the interannual variability of the hindcasts, the Pearson correlation, Spearman rank correlation and the normalized root-mean-square error of the ensemble mean NTC and NHUR are given in Table 2 and correspond to the seasonal values shown in Figs. 3a and 3c. For NTC both correlations are statistically significant for NTC for leads 1–5, with the highest correlations occurring for week 1 and week 3. The RMS for NTC is highest in the first two leads, decreasing an order of magnitude for leads 3–6.

In the case of NHUR, both correlations are highest for week 1, decrease for weeks 2 and 3 and increase in week 4, with the same or higher value for week 5. For NHUR the RMS is lowest for weeks 1 and 2, increasing substantially at week 3 and remaining with similar values for weeks 4–6. While the changes in RMS from weeks 1 and 2 to week 3 are clear from Figs. 2 and 3 and probably associated with the increase in model grid

spacing, the reason behind the increase in correlation at the longer leads is not clear.

Similar to NTC and NHUR, the highest correlation between the ensemble mean ACE and observed ACE occurs at week 1 (Table 2), and the correlations are statistically significant for all leads. The lowest ACE correlations values occur at week 2—they increase at week 3 and remain with similar values for weeks 4 and 5, before decreasing again for week 6. The RMS highest values occur for weeks 1 and 2, when the ensemble mean ACE is typically too high—ACE RMS reduces and is very stable for weeks 3–5, before increasing again for week 6.

The analysis presented so far shows the important influence of both model resolution and use of the stochastic parameterization in determining the ability of the ECMWF model in reproducing the climatological and interannual variability of TC activity in the Atlantic.

b. *Cluster analysis*

To consider the climatological characteristics of the ECMWF tracks, we use cluster analysis, which is described in detail in section 2f. The application of this cluster method to observed North Atlantic tropical cyclones is described in Kossin et al. (2010), with additional aspects discussed in Kozar et al. (2012) and Boudreault et al. (2017).

Following Kossin et al. (2010), here we use four clusters to describe the Atlantic tropical cyclone tracks as shown in Fig. 4. Cluster 1 mostly consists of storms that form in the subtropics and mainly affect the U.S. East Coast. Gulf of Mexico TCs dominate cluster 2. Cluster-3 and cluster-4 TCs form in the tropical Atlantic main development region (MDR) and are differentiated by the longitude of the formation, with cluster-3 TC genesis closer to Africa (MDR–East/Cape Verde storms), and cluster-4 TC genesis closer to the Caribbean (MDR–West/Caribbean storms). Most of the major hurricanes belong to cluster 3 (Cape Verde storms). The highest rate of landfalling storms occur for the Gulf of Mexico (cluster 2) and the Caribbean (cluster 4) TCs. Additional discussion of the characteristics of the observed clusters can be found in Kossin et al. (2010), Kozar et al. (2012), Boudreault et al. (2017).

We applied the same cluster method to the ECMWF Atlantic TC tracks. However, we used five clusters instead of 4. The reason for this choice is clear in Fig. 5, which shows the tracks for storms in week 2. While clusters 1–4 in the ECMWF monthly tracks can be clearly identified with the observed clusters, there are tracks in the model that belong to an additional cluster (cluster 5) with characteristics that do not match in observations. This additional cluster consists of storms that form in the tropics near Africa and recurve toward Africa and/or Europe in locations that are much more eastward than in the observations. Typically, these tracks are absent in observations because of the subtropical high, which forces the storms to go around that region. The tracks in cluster 5 are associated with systematic model biases, as discussed below.

The contributions of each cluster to the mean NTC, NHUR, and ACE per season in observations and the model

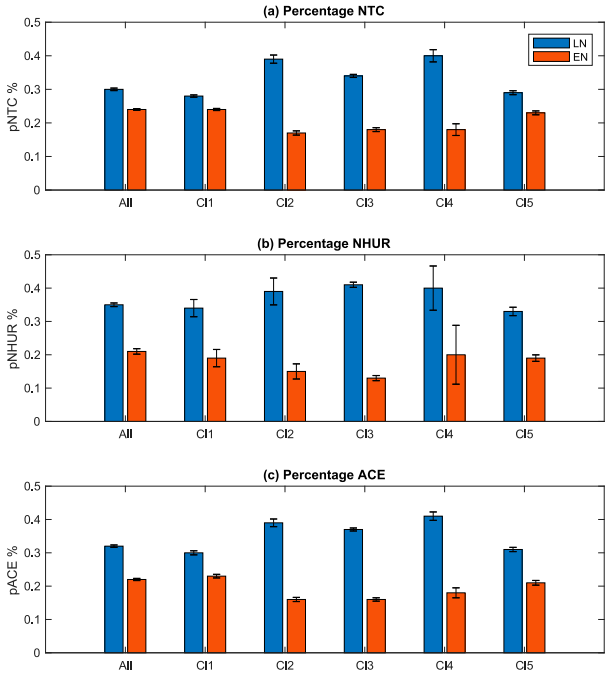

FIG. 8. Percentage of NTC, NHUR, and ACE in El Niño (EN) and La Niña (LN) phases for the ECMWF monthly hindcasts for all storms (All) and by cluster (C1–C5) for all leads. Error bars are 95% confidence intervals estimated using a bootstrap procedure with 1000 bootstrap samples.

are given in Fig. 6. The statistics of the clusters are also given in Table 3. The total number of storms and percentages in observations varies with the period considered, because of sampling. For simplicity, in our discussion we will refer to the observed values for the longer period (1950–2018), which are more robust. The observed results for both periods, with the shorter one (1998–2007) matching the model hindcasts, are shown in Fig. 6 and Table 3.

In observations, cluster 1 has the highest number of TCs (32%) and cluster 4 has the lowest (14%). The largest difference between the model and observations is the high number of TCs (30%) in cluster 5, which has no counterpart in observations. The percentage of TCs in cluster 5 varies little by lead time, with a minimum of 28% for week 1, and a maximum of 34% in week 2. While the percentage of TCs in cluster 1 (subtropical) is similar in the model and observations, the percentage of TCs in the other clusters is substantially lower, in particular in cluster 4. The reduction of TCs in the western part of the basin (clusters 2, 3, and 4) can be explained by the tendency of the model to generate too many TCs near the African coast, which results in a much more active eastern Atlantic than in observations.

The differences between the model and observations are greater for NHUR and ACE, since the nonexistent coast of Africa cluster 5 is the dominant contributor to those variables. If we consider only the four other ECMWF model clusters, their relative contributions to NHUR and ACE are similar to observations, with cluster 3 being the dominant cluster, with

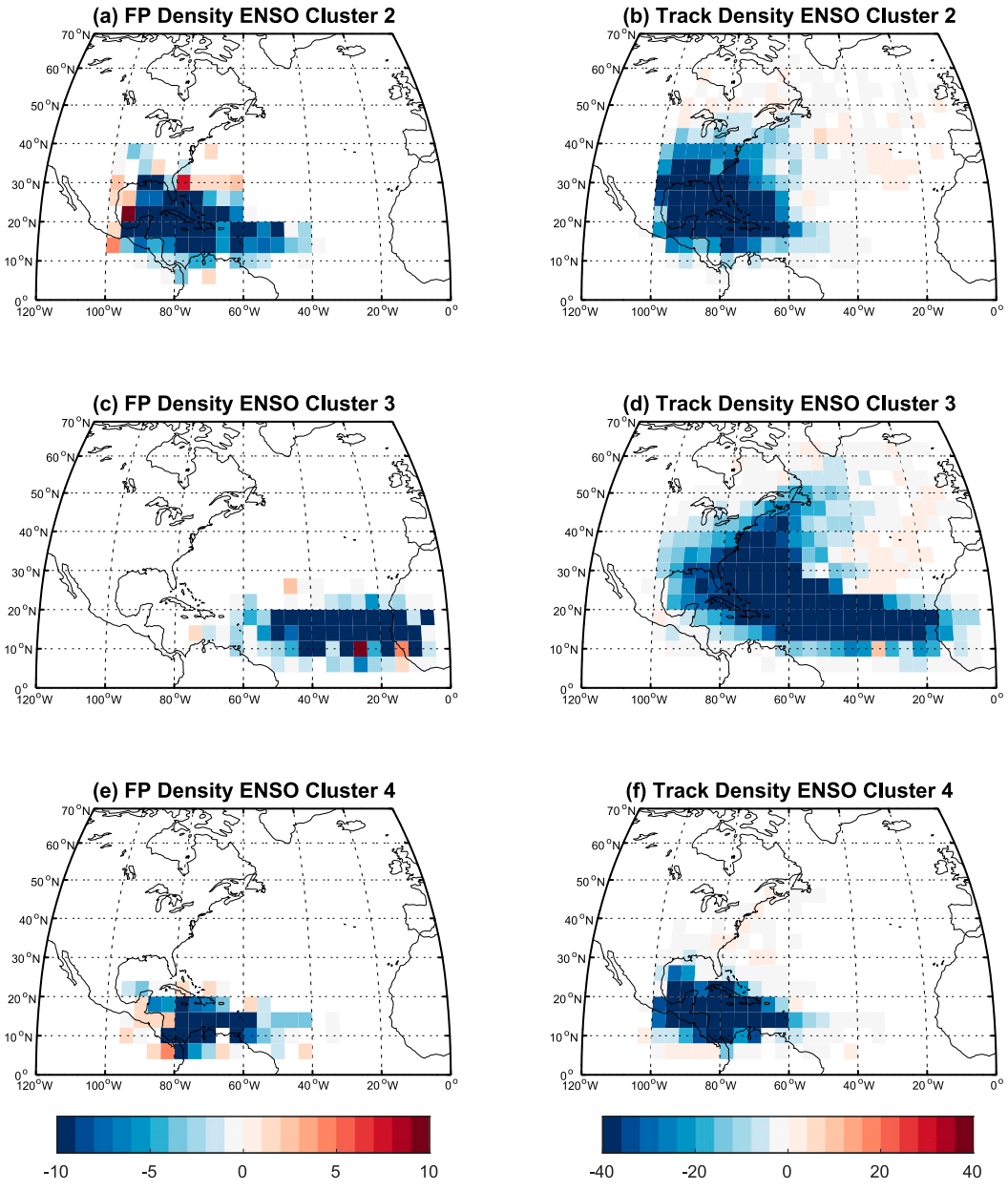

FIG. 9. Difference in (a),(c),(e) first position and (b),(d),(e) track density between El Niño and La Niña seasons in the ECMWF monthly hindcasts for clusters (top) 2, (middle) 3, and (bottom) 4 and all leads.

the highest number of hurricanes and major hurricanes. Slightly lower than cluster 3 for NHUR and ACE is the contribution of cluster 1 (subtropical storms), which is also reproduced in the model. Similarly to NTC, the biggest difference between models and observations occurs for cluster 4, which has a very low contribution to the totals in the model with too few TCs intensifying to hurricanes and major hurricanes in the Caribbean.

To better understand the reasons behind the occurrence of this extra cluster, we calculated the model bias of various environmental fields, by comparing the model climatology with the ERA5 climatology. Here we show in Fig. 7, the environmental fields that have biases associated with the additional cluster near Africa in the period June to November. The ECMWF model has a positive bias in the 850-hPa zonal winds over the tropical eastern Atlantic, reducing the amplitude of the low-level winds (easterlies) (Fig. 7a). Furthermore, the meridional winds at 850 hPa have a positive bias in the Atlantic main development region (Fig. 7b), whereas the 500-hPa geopotential height has a positive bias in the mid and high latitudes of the North Atlantic, on an elongated axis from southeast to northwest (Fig. 7c). The combination of these biases is conducive to tracks with an east-northward trajectory, therefore closer to Africa and avoiding the region of high

TABLE 5. Statistics of observations and ECMWF for NTC, NHUR, and ACE ($\times 10^{-6}$) in each cluster per season for NAO+ and NAO−. Following (Kossin et al. 2010), the NAO phases are defined as the bottom and top quartile of the NAO mean value in May and June before the hurricane season. Five NAO+ seasons and five NAO− seasons were considered. The percentages shown for each cluster are based on the total value of each variable for each cluster. Also shown are the statistics for 5 and 17 NAO seasons in the observations. Significance is indicated at the 99% level with boldface type and an asterisk and at the 95% level with boldface type only.

	NTC				NHUR				ACE			
	Total		%		Total		%		Total		%	
Cluster	NAO+	NAO−	NAO+	NAO−	NAO+	NAO−	NAO+	NAO−	NAO+	NAO−	NAO+	NAO−
ECMWF monthly—5 NAO events MJ												
Total	**6060**	**6684**	0.24	0.27	**613***	**752***	0.24	0.30	513	591.45	0.24	0.28
1	2145	2263	0.25	0.26	114	142	0.24	0.30	108.97	118.81	0.24	0.26
2	**683***	**903***	0.22	0.29	47	71	0.19	0.28	**44.56***	**63.82***	0.20	0.29
3	**940***	**1218***	0.23	0.29	**208***	**283***	0.24	0.33	**142.68***	**187.92***	0.23	0.31
4	**281***	**424***	0.21	0.32	5	9	0.12	0.21	**19.72**	**29.90**	0.20	0.30
5	**2011**	**1876**	0.27	0.25	239	247	0.27	0.28	197.08	190.98	0.27	0.26
Observations—5 NAO events												
Total	62	76	0.21	0.25	**25**	**47**	0.16	0.31	51.83	80.72	0.17	0.27
1	17	16	0.21	0.20	**3**	**13**	0.09	0.38	6.97	11.75	0.14	0.24
2	14	16	0.19	0.22	4	3	0.17	0.12	5.32	7.22	0.14	0.20
3	19	29	0.20	0.31	12	20	0.19	0.32	29.67	42.72	0.19	0.27
4	12	15	0.22	0.28	6	11	0.19	0.34	9.87	19.03	0.16	0.31
Observations—17 NAO events												
Total	202	213	0.25	0.26	**106**	**124**	0.24	0.28	179.96	223.88	0.22	0.27
1	64	62	0.24	0.23	31	34	0.24	0.26	38.64	42.27	0.21	0.23
2	54	48	0.26	0.23	23	17	0.26	0.19	29.23	22.72	0.25	0.20
3	**60**	**71**	0.25	0.30	38	50	0.24	0.32	89.02	118.45	0.22	0.29
4	24	32	0.21	0.29	14	23	0.20	0.32	**23.06**	**40.44**	0.18	0.32

geopotential height, which is consistent with the additional cluster that is present in the model.

c. *Relationship with climate modes*

Since the number of hindcasts is large in comparison with observations, we are able to robustly determine whether the observed modulation of Atlantic TCs by various climate modes is reproduced in the model. Moreover, the large sample size allows us to identify characteristics of these modulations that are not easily discerned in observations due to the shortness of the reliable TC record and the relatively small number of climate modes events.

1) ENSO

The El Niño–Southern Oscillation (ENSO) has a well-known relationship with Atlantic TC activity (Gray 1984; Camargo et al. 2007a). The ENSO modulation of Atlantic TC activity, with typically lower activity in El Niño years and higher activity in La Niña years, is the main source of predictability of TC seasonal forecasts in the Atlantic (e.g., Vitart et al. 2007; Camargo and Barnston 2009; Klotzbach et al. 2019, 2020; Caron et al. 2020). However, the basin-wide modulation is not the only response of North Atlantic TC activity to ENSO.

Table 4 presents the statistics of the hindcasts and observations by ENSO phase for NTC, NHUR, and ACE. For the hindcasts, we considered the ENSO phase defined by the Niño-3.4 values in the model, as well as the observed ENSO events in the period. In the case of observations, we show the statistics for two periods, a shorter period coinciding with the hindcast period and a longer period (1950–2018). Because there are relatively few ENSO events in the shorter period, we only discuss here the results for the longer period.

The hindcasts clearly reproduce the observed basin-wide reduction in the occurrence of TC events during El Niño seasons, and the increase during La Niña seasons, as is also shown in Fig. 8 with the label "All." This modulation is also present in NTC and ACE in both model and observations. While the reduction is statistically significant for NTC and ACE in the model independently of the ENSO definition, the reduction in NHUR is only statistically significant using the observed ENSO values, probably due to the difference in how we define the ENSO events in observations and the model, as discussed above. These differences could be particularly important in the late spring and early summer, when ENSO is typically transitioning from one phase to the other and Niño-3.4 has smaller anomalies than in boreal winter.

Kossin et al. (2010) analyzed the modulation of the Atlantic clusters by ENSO. In observations, the clusters with a statistically significance modulation by ENSO are the two tropical clusters (clusters 3 and 4), as shown in Table 4. They found a statistically significant modulation by ENSO for NTC, NHUR for both clusters, and ACE for cluster 3. Note that the period analyzed in Kossin et al. (2010) was 1950–2007, while here the cluster analysis was extended to 2018.

The ECMWF model has a strong ENSO modulation in individual clusters. When considering the model Niño-3.4 to define ENSO, clusters 1–4 show a statistically significant ENSO modulation for NTC and ACE, with fewer clusters being statistically significant for NHUR, namely clusters 2 and

3. Cluster 5 has a statistically significant ENSO modulation for NHUR and ACE.

If we consider the observed ENSO, instead of the model ENSO, all clusters show a statistically significant ENSO modulation for all variables (NTC, NHUR, and ACE), with exception of cluster 4 for NHUR, which has a very small sample size (see Table 4 and Fig. 8). The strong ENSO modulation of the clusters in the ECMWF model is clear in the first position and track density differences between El Niño and La Niña for clusters 2–4 (Fig. 9). The increase in genesis and occurrence in La Niña seasons in comparison with El Niño seasons is apparent not only in the genesis locations (left panels), but across basin subregions. While it is well known that the flavor of ENSO influences the degree to what the TCs are suppressed in the Atlantic (Patricola et al. 2016), there are not enough ENSO events in the hindcast period to examine that.

2) AMM

The Atlantic meridional mode (AMM) is another climate mode that has a substantial influence on the interannual variability of Atlantic hurricane activity (Kossin and Vimont 2007; Vimont and Kossin 2007), with higher activity than usual during the positive phase of the AMM and suppressed activity in seasons when the AMM is in its negative phase. Furthermore using model simulations, Patricola et al. (2016) demonstrated that constructive combinations of ENSO and AMM phases can lead to either extremely active or suppressed Atlantic seasons. Kossin et al. (2010) singled out the two MDR tropical clusters (3 and 4) as being most affected by the AMM, which was confirmed by Boudreault et al. (2017).

During the ECMWF hindcast period, the AMM is positive in most of the Atlantic hurricane seasons. Considering months with positive/negative values above/below 1 standard deviation of the data, there are only 8 months with negative AMM, as compared with 45 months with positive AMM. Moreover, there are only six ASO seasons with negative AMM, and from those only one is below 1 standard deviation. Therefore, unfortunately it does not make sense to analyze the influence of the AMM in this ECMWF hindcast period, as there are not enough seasons with a negative AMM for the analysis to be meaningful and statistically significant in both AMM phases.

3) NAO

The relation between the NAO and Atlantic hurricanes was first discussed by Elsner and collaborators in various publications (Elsner and Kara 1999; Elsner and Kocher 2000; Elsner 2003). Subsequent publications have documented the complexity of the relationship. For instance, Kossin et al. (2010) noted that the annual rate of storms in cluster 1 (subtropical) is associated with the May–June NAO index. Specifically when the NAO is in its negative phase during those months, there are more storms in cluster 1 during the hurricane season (June–November). The link between the NAO and Atlantic TC activity is thought to be through the position and strength of the Atlantic subtropical high. Interestingly, Kozar et al. (2012) found a relationship between Atlantic TC counts and the boreal winter postseason NAO index, in particular to cluster 4 (Caribbean). More recently, Angus and Leckerbusch (2020) showed that there is an inverse relationship between the TC activity and the following European winter season and explained this connection through ENSO and the winter NAO. However, other studies do not find a relationship between Atlantic hurricane tracks and the NAO (Colbert and Soden 2012). Therefore, the topic was revisited by Boudreault et al. (2017), who pointed out that while there is a statistically significant relationship between May–June NAO and basin-wide TC activity, NAO does not have any predictive skill for landfalling Atlantic hurricanes.

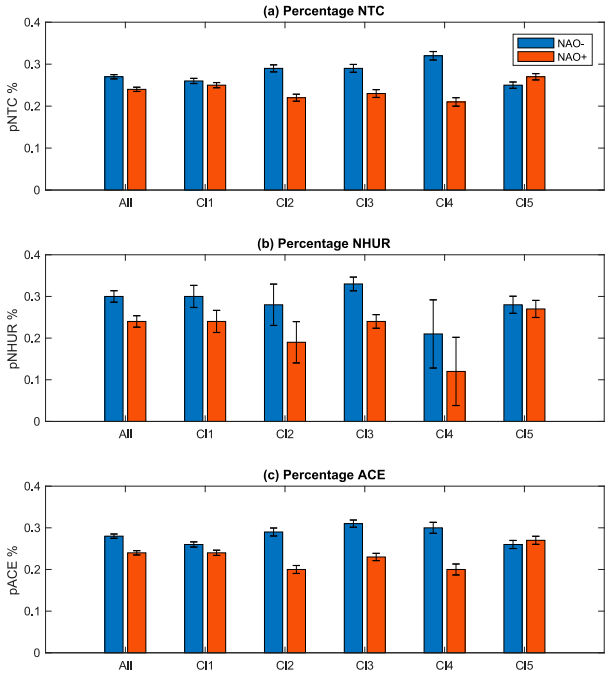

FIG. 10. Percentage of NTC, NHUR, and ACE in the positive and negative NAO phases (NAO+ and NAO−) by cluster for the ECMWF monthly hindcasts for all leads. Errors are estimated using a bootstrap procedure with 1000 bootstrap samples.

Here we reexamine the relationship of TC activity with the NAO using the ECMWF hindcast and with the expectation that its large sample will help to clarify this issue. Table 5 shows the statistics of NAO and Atlantic TCs in observations and the ECMWF monthly hindcasts. For observations, there is a statistically significant increase in NHUR in the basin for NAO− when compared with NAO+. There are also statistically significant differences in NTC for cluster 3 and ACE for cluster 4. We note that the analysis performed here considers a different period than that used in Kossin et al. (2010) and that the modulation of cluster 1 activity by NAO is not apparent.

In the case of the ECMWF hindcasts, there is a much clearer signal of the NAO modulation of the Atlantic TC activity, as shown in Fig. 10 and Table 5. The total NTC and NHUR are modulated by the NAO, with an increase in activity for NAO−. Furthermore, clusters 2, 3, and 4 have a statistically significant modulation of NTC and ACE with the NAO phase. Modulation of NHUR in Cluster 3 is also statistically significant. The lack of a signal in NHUR for clusters 2 and 4 could be

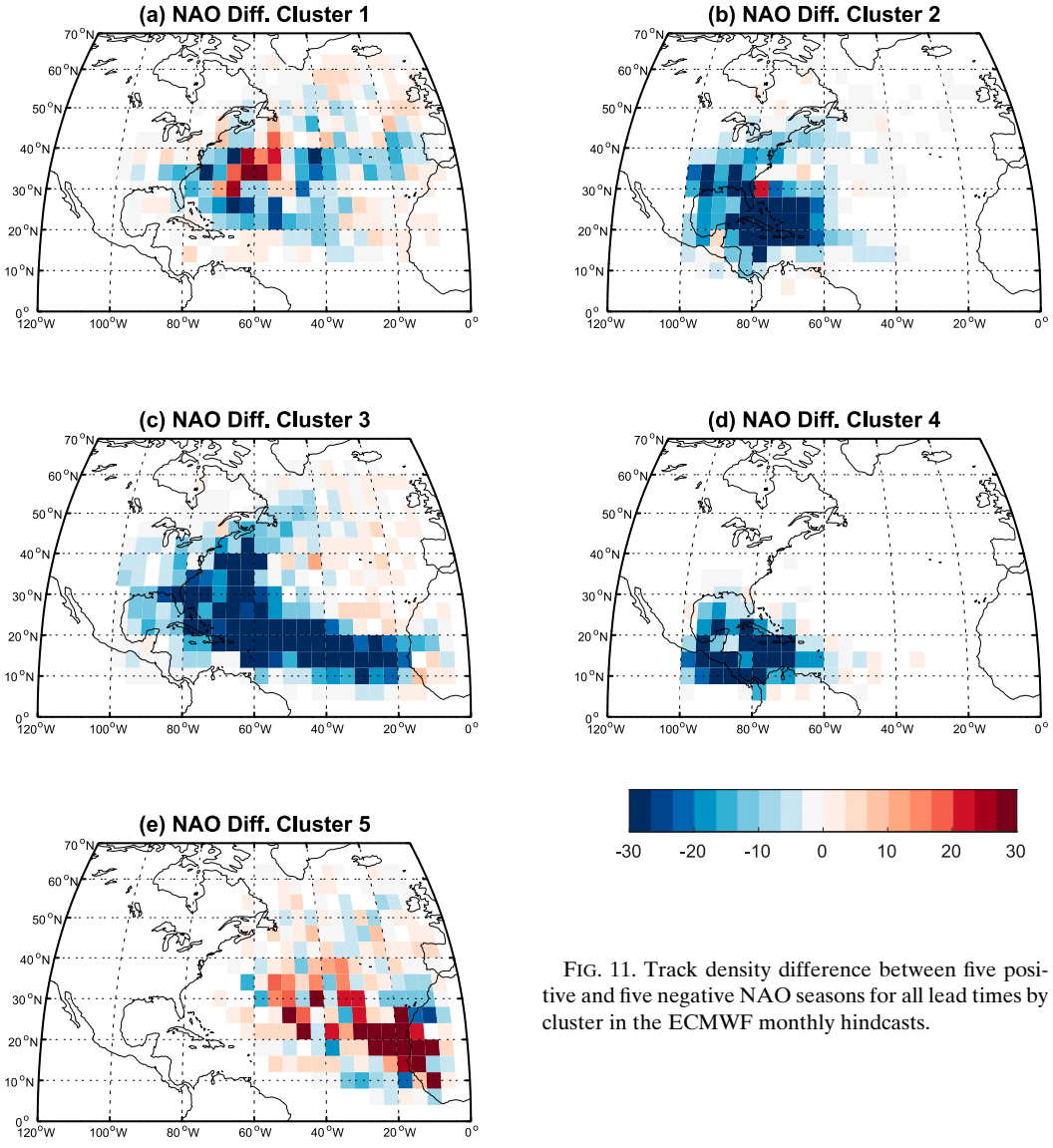

FIG. 11. Track density difference between five positive and five negative NAO seasons for all lead times by cluster in the ECMWF monthly hindcasts.

due to the low number of hurricanes in those clusters, in particular cluster 4. Interestingly, cluster 5 has a modulation in NTC that is the opposite of the other clusters, that is, a higher value of NTC for NAO−. Furthermore, there is no modulation by NAO in the subtropical cluster 1 in the hindcasts. These results appear clearly in the track density differences between NAO+ and NAO− (Fig. 11), which show a clear increase in TC occurrence for NAO− for clusters 2, 3, and 4, the opposite in cluster 5, and the lack of a clear signal in cluster 1.

4) MJO

As mentioned above, there is a strong modulation of TCs globally by the MJO. In the case of the Atlantic, this modulation was first noticed in Gulf of Mexico storms (Maloney and Hartmann 2000). Kossin et al. (2010) found a statistically significant MJO modulation of Gulf of Mexico storms in cluster 2, but not in the other clusters. The observed increase in TC activity occurs in the combined RMM phases 8 + 1 and 2 + 3 with a suppression in the other phases (4 + 5, 6 + 7). This signal is also clear in a genesis index (Camargo et al. 2009), which elucidates the role of environmental fields in the MJO modulation of TC activity. Klotzbach (2014) found that the statistical significance of the MJO modulation depends on the variable considered, with the strongest signals being an increase in TC activity in phases 1 + 2 for NTC, NHUR, and ACE when compared with phases 6 + 7.

Here we examine the modulation of the TC activity by the MJO in the ECMWF hindcast applying the same type of visualization used in Lee et al. (2018), the "candy plot," as shown in Fig. 12. The color of each circle indicates the percentage of TCs occurring in each MJO phase in the basin. In observations (see Fig. 6 in Lee et al. 2018), the Atlantic has a statistically significant increase in TC activity in phases 2, 3, and 4 and a statistically significant decrease in phases 6, 7, and 8.

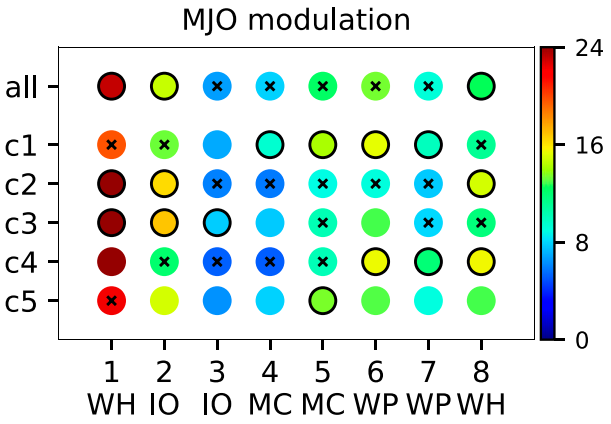

FIG. 12. Candy plot of MJO modulation using RMM following Lee et al. (2020). The color of each candy indicates the probability distribution function in percent in the corresponding MJO phase in the basin. The sum of the circles across the MJO phases in each basin is 100%. A black circle at the edge of the disk indicates that the value is above the 90th percentile of the distribution, and a cross symbol at the center of the disk means that the value is below the 10th percentile of the distribution. Definitions: "all" means not separated by clusters, the clusters are identified as c1–c5, and the locations of the RMM phases 1–8 as defined in Wheeler and Hendon (2004) are also indicated by their location: western Hemisphere (WH), Indian Ocean (IO), Maritime Continent (MC), and western Pacific (WP).

The modulation in the Atlantic in the ECMWF hindcasts shows a statistically significant increase in genesis in the Atlantic for RMM phases 1, 2, and 8, and a suppression in phases 3–7 (Fig. 12). The modulation is robust across lead times. The high TC activity in phase 1 is statistically significant for all lead times and for phase 8 for weeks 2–6. Similarly, the suppression of TC activity in phases 3–6 is significant for weeks 2–6, with phase 5 being also statistically significant for week 1. The MJO modulation is not robust across lead times for phases 2 and 7, with the signal being statistically significant for only a few cases and in the case of phase 7, the modulation changes sign for different lead times. The hindcasts are clearly strongly modulated by the MJO phase even at long leads, indicating the possibility of skillful forecasts at subseasonal time scales at long leads for specific MJO phases.

d. *Skill scores*

To examine the skill of the ECMWF monthly system in forecasting Atlantic TC activity, we use the Brier skill score (BSS) for the genesis forecasts. The BSS measures skill relative to a reference forecast, and similarly to Lee et al. (2018), we consider two climatologies as reference forecasts: a constant climatology and a monthly varying climatology. The resulting scores are shown in Fig. 13. The BSS for all leads combined are given in the diamonds, and the BSS values for different lead times are shown in the curves. The seasonal varying climatology leads to lower values of BSS, and the genesis forecasts have negative skill for all leads, except week 1. In contrast, the BSS for the constant climatology is positive for all lead times. The highest skill values for both cases occur for week 1 and week 6, with very similar BSS values from week 2 to week 5.

We examined the skill of ACE forecasts using the rank probability skill score (RPSS). Similar to the case of genesis the RPSS is negative for all leads when comparing the ECMWF's prediction skill with the monthly varying climatology, while it is positive when comparing with a constant climatology. The RPSS is slightly higher for week 1 and is almost constant afterward using a constant climatology, whereas with a varying climatology there is a minimum at week 3. While there is some skill in the ECMWF system for basin wide genesis and ACE, the skill is low and not greater than that of a monthly varying climatology.

Next we examined the regional skill for predicting TC occurrence in the Atlantic basin. The threat score is shown Fig. 14a. The higher the threat score the more accurate the forecast is (a perfect forecast has a threat score equal to 1), measuring the fraction of events that were correctly predicted. The highest threat score values occur in the middle of the basin, while the lowest ones are near the African coast and north of South America. The Brier score, which measures the magnitude of the probability forecast error, is shown in Fig. 14b, and in this case a perfect score is zero. The region with the highest values for the Brier score (largest forecast errors) occur near Florida and the southeastern United States. The two bottom panels of Figs. 14c and 14d represent the Brier skill score

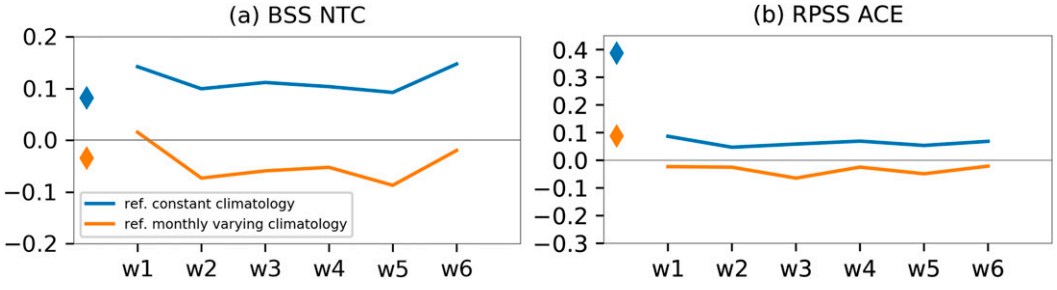

FIG. 13. (a) Brier skill score (BSS) of TC genesis forecasts and (b) ranked probability skill score (RPSS) of ACE forecasts. Curves show the skill score by lead time (weeks 1–6), and diamonds show the skill score for all leads. Skill scores are computed using as reference forecast a seasonal constant climatology (blue) and a monthly varying climatology (orange), following Lee et al. (2020).

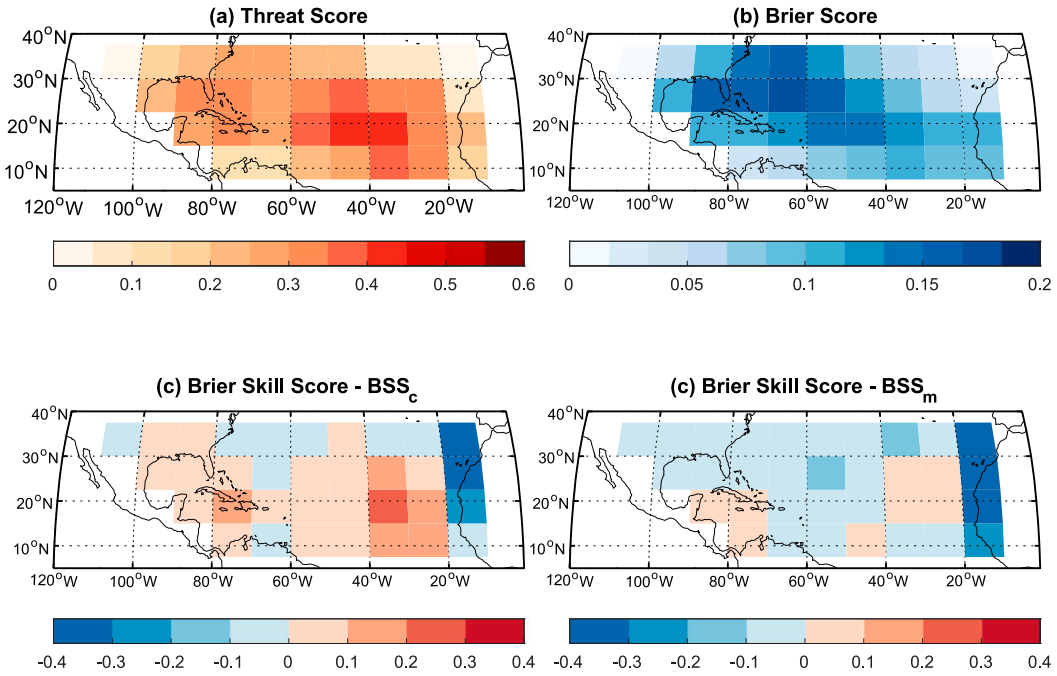

FIG. 14. (a) Threat score, (b) Brier score, and BSS of (c) seasonal constant climatology and (d) monthly varying climatology for TC occurrence hindcasts in week 2. BSS$_c$ and BSS$_m$ are computed following Lee et al. (2020).

relative to a constant climatology (BSS$_c$) and a monthly varying climatology (BSS$_m$). In both cases, negative BSS values can be found near Africa, where the model forms too many storms, as noted above. Positive BSS can be found in the middle of the basin and in the Caribbean (relative to both types of climatology), as well as in the Gulf of Mexico (for a constant climatology only). The region close to the U.S. Southeast Coast has negative BSS in both cases.

e. *Predictability*

Given the low skill scores of the ECMWF system, we would like to explore its potential predictability. We estimated the predictability by designating one ensemble member as "observations" for NTC and NHUR, and then calculated the Brier skill score using the other ensemble members, repeating the process across all ensembles. We excluded the control simulation (E0) from this process, given the different characteristics of that ensemble member from the others. The top panels of Fig. 15 show the predictability of the ensemble for basin-wide values of NTC and NHUR. As expected, the highest BSS occurs for week 1 and decreases with lead times. In the case of NTC, the BSS for weeks 1 and 2 is substantially higher than for weeks 3–6, which have very similar values. In contrast, for NHUR, BSS for week 2 is similar to that at longer leads times. By week 6, the NHUR BSS value reaches 0, indicating lack of predictability, which is not the case for NTC.

The bottom panels of Fig. 15 consider the predictability of NTC and NHUR for each cluster. Different clusters, have different levels of predictability. The cluster with most predictable NTC and NHUR is cluster 5, which has no observed counterpart and is the second most populated cluster. The least predictable (NTC and NHUR) cluster is cluster 4 (MDR–Caribbean), which has a bias of very low frequency in the ECMWF system. For NTC at week 1, clusters 1, 2, and 5 have approximately the same high BSS values. However, while the BSS of clusters 1 (subtropical) and 2 (Gulf of Mexico) strongly decrease at week 2, the BSS decrease of cluster 5 is slower. By week 4, the only clusters with positive BSS are clusters 5 and 3 (MDR–Cape Verde), both with genesis on the eastern part of the basin. In the case of NHUR, clusters 5, 3, 1, and 4 (in decreasing order of BSS) have positive BSS values at week 1, but by week 3, only clusters 5 and 3 still have predictability. Probably the lower predictability of the system for NHUR is associated with the low-intensity bias of the model, especially for higher lead times, when NHUR is very low. It is encouraging that the cluster with the most intense and dangerous hurricanes (cluster 3) has a positive BSS for NHUR up to week 3, indicating the potential for skillful forecasts for those storms.

4. Conclusions

Here, we examined the characteristics of Atlantic hurricane activity in the ECMWF monthly system and compared them with observations. Both model resolution and the stochastic parameterization used in 10 of the 11 ensemble members strongly influence how well the ECMWF hindcast system can reproduce the observed characteristics of Atlantic TC activity. In the case of NTC, the control ensemble produces a mean value close to the observed, while the perturbed ensemble members produce too many storms. The number of hurricanes is low in all ensemble members, in particular in the control,

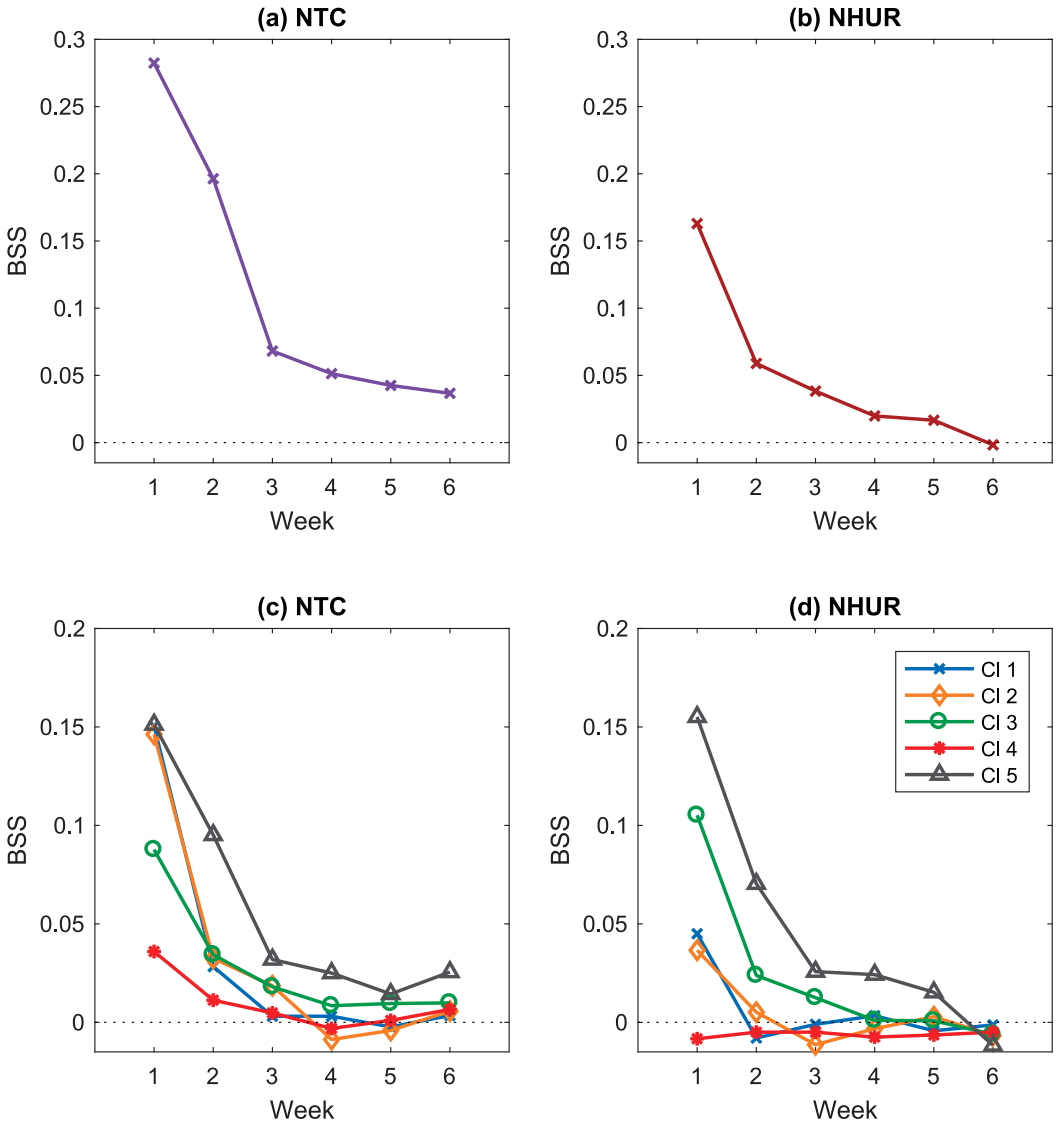

FIG. 15. Estimate of predictability for NTC and NHUR using BSS among 10 ensemble members (excluding control): (a) NTC, (b) NHUR, (c) NTC per cluster, and (d) NHUR per cluster.

when compared with observations. There is a clear decline in NTC and NHUR at longer lead times (weeks 3–6), when the model horizontal resolution is degraded. Overall, the ECMWF monthly system has a low-intensity bias in hurricane intensity, with very few occurrences of major hurricanes at all leads.

Using cluster analysis, we compared the track characteristics of the ECMWF Atlantic TCs with observations. There are clear spatial differences, with the ECMWF system having an additional cluster with recurving tracks near the coast of Africa with characteristics that do not correspond to the observations. Furthermore, the cluster with Caribbean storms (cluster 4) has too few TCs, indicating an eastward bias in the model genesis. Nonetheless, despite the additional cluster, the other clusters tracks have strong similarities with observations. In particular, the Cape Verde cluster (cluster 3), with the most intense and devastating hurricanes is well represented in the ECMWF system. Similarly, the Gulf of Mexico (cluster 2) and East Coast (cluster 1) clusters have many characteristics similar to the observed ones.

The recurving nonobserved tracks near Africa are shown to be associated with systematic biases in low-level winds in geopotential heights in the ECMWF model. These biases lead to lower BSS values near the coast of Africa. The model skill is highest in the mid-Atlantic and near the Caribbean. The skill score results depend on choice of the reference forecast (monthly varying or constant climatology).

The modulation of the ECMWF Atlantic TCs by climate modes was examined. Strong modulation by ENSO, NAO and MJO was clear for the full basin, as well as for individual clusters. In particular, the MJO modulation strongly supports the possibility of skillful subseasonal TC forecasts, as the MJO is the main source of predictability on those time scales.

While in the current system, there are only positive skill scores in certain cases, our predictability analysis indicates that there is possibility of developing skillful subseasonal Atlantic TC forecasts, not only for the full basin, but on regional scales, as indicated by the predictability of the different clusters.

Acknowledgments. This work was supported by NOAA Grant NA16OAR4310079. Author Camargo also acknowledges the Vetlesen Foundation for their generous and sustained support of climate science at the Lamont-Doherty Earth Observatory.

Data availability statement. The data underlying the figures are available at the Columbia Academic Commons (https://academiccommons.columbia.edu/doi/10.7916/d8-htpz-gq16). The full hindcast dataset for Atlantic hurricanes is available from Dr. Frédéric Vitart at ECMWF by request (frederic.vitart@ecmwf.int).

REFERENCES

Angus, M., and G. C. Leckerbusch, 2020: On the dependency of Atlantic hurricane and European windstorm hazards. *Geophys. Res. Lett.*, **47**, e2020GL090446, https://doi.org/10.1029/2020GL090446.

Bamston, A. G., M. Chelliah, and S. B. Goldenberg, 1997: Documentation of a highly ENSO-related SST region in the equatorial Pacific. *Atmos.–Ocean*, **35**, 367–383, https://doi.org/10.1080/07055900.1997.9649597.

Barnston, A. G., and R. E. Livezey, 1987: Classification, seasonality, and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.*, **115**, 1083–1126, https://doi.org/10.1175/1520-0493(1987)115<1083:CSAPOL>2.0.CO;2.

Bechtold, P., M. Köhler, T. Jung, F. Doblas-Reyes, M. Leutbecher, M. J. Rodwell, F. Vitart, and G. Balsamo, 2008: Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Quart. J. Roy. Meteor. Soc.*, **134**, 1337–1351, https://doi.org/10.1002/qj.289.

Belanger, J. I., J. A. Curry, and C. D. Hoyos, 2009: Variability in tornado frequency associated with U.S. landfalling tropical cyclones. *Geophys. Res. Lett.*, **36**, L17805, https://doi.org/10.1029/2009GL040013.

——, ——, and P. J. Webster, 2010: Predictability of North Atlantic tropical cyclone activity on intraseasonal time scales. *Mon. Wea. Rev.*, **138**, 4362–4374, https://doi.org/10.1175/2010MWR3460.1.

——, P. J. Webster, J. A. Curry, and M. T. Jelinek, 2012: Extended prediction of north Indian Ocean tropical cyclones. *Wea. Forecasting*, **27**, 757–769, https://doi.org/10.1175/WAF-D-11-00083.1.

Bell, G. D., and Coauthors, 2000: Climate assessment for 1999. *Bull. Amer. Meteor. Soc.*, **81**, S1–S50, https://doi.org/10.1175/1520-0477(2000)81[s1:CAF]2.0.CO;2.

Bell, S. S., S. S. Chand, S. J. Camargo, K. J. Tory, C. Turville, and H. Ye, 2019: Western North Pacific tropical cyclone tracks in CMIP5 models: Statistical assessment using a model-independent detection and tracking scheme. *J. Climate*, **32**, 7191–7208, https://doi.org/10.1175/JCLI-D-18-0785.1.

Boudreault, M., L.-P. Caron, and S. J. Camargo, 2017: Reanalysis of climate influences on Atlantic tropical cyclone activity using cluster analysis. *J. Geophys. Res. Atmos.*, **122**, 4258–4280, https://doi.org/10.1002/2016JD026103.

Buizza, R., J.-R. Bidlot, N. Wedi, M. Fuentes, M. Hamrud, G. Holt, and F. Vitart, 2007: The new ECMWF VAREPS (Variable Resolution Ensemble Prediction System). *Quart. J. Roy. Meteor. Soc.*, **133**, 681–695, https://doi.org/10.1002/qj.75.

Camargo, S. J., 2013: Global and regional aspects of tropical cyclone activity in the CMIP5 models. *J. Climate*, **26**, 9880–9902, https://doi.org/10.1175/JCLI-D-12-00549.1.

——, and A. G. Barnston, 2009: Experimental seasonal dynamical forecasts of tropical cyclone activity at IRI. *Wea. Forecasting*, **24**, 472–491, https://doi.org/10.1175/2008WAF2007099.1.

——, K. A. Emanuel, and A. H. Sobel, 2007a: Use of a genesis potential index to diagnose ENSO effects on tropical cyclone genesis. *J. Climate*, **20**, 4819–4834, https://doi.org/10.1175/JCLI4282.1.

——, A. W. Robertson, S. J. Gaffney, P. Smyth, and M. Ghil, 2007b: Cluster analysis of typhoon tracks. Part I: General properties. *J. Climate*, **20**, 3635–3653, https://doi.org/10.1175/JCLI4188.1.

——, ——, ——, ——, and ——, 2007c: Cluster analysis of typhoon tracks. Part II: Large-scale circulation and ENSO. *J. Climate*, **20**, 3654–3676, https://doi.org/10.1175/JCLI4203.1.

——, ——, A. G. Barnston, and M. Ghil, 2008: Clustering of eastern North Pacific tropical cyclone tracks: ENSO and MJO effects. *Geochem. Geophys. Geosyst.*, **9**, Q06V05, https://doi.org/10.1029/2007GC001861.

——, M. C. Wheeler, and A. H. Sobel, 2009: Diagnosis of the MJO modulation of tropical cyclogenesis using an empirical index. *J. Atmos. Sci.*, **66**, 3061–3074, https://doi.org/10.1175/2009JAS3101.1.

——, and Coauthors, 2019: Tropical cyclone prediction on subseasonal time-scales. *Trop. Cyclone Res. Rev.*, **8**, 150–165, https://doi.org/10.1016/j.tcrr.2019.10.004.

——, and Coauthors, 2020: Characteristics of model tropical cyclone climatology and the large-scale environment. *J. Climate*, **33**, 4463–4487, https://doi.org/10.1175/JCLI-D-19-0500.1.

Camp, J., and Coauthors, 2019: The western Pacific subtropical high and tropical cyclone landfall: Seasonal forecasts using the Met Office GloSea5 system. *Quart. J. Roy. Meteor. Soc.*, **145**, 105–116, https://doi.org/10.1002/qj.3407.

Caron, L.-P., F. Massonnet, P. J. Klotzbach, T. J. Philp, and J. Stroeve, 2020: Making seasonal outlooks of Arctic sea ice and Atlantic hurricanes valuable—not just skillful. *Bull. Amer. Meteor. Soc.*, **101**, E36–E42, https://doi.org/10.1175/BAMS-D-18-0314.1.

Colbert, A. J., and B. J. Soden, 2012: Climatological variations in North Atlantic tropical cyclone tracks. *J. Climate*, **25**, 657–673, https://doi.org/10.1175/JCLI-D-11-00034.1.

Daloz, A. S., and Coauthors, 2015: Cluster analysis of downscaled and explicitly simulated North Atlantic tropical cyclone tracks. *J. Climate*, **28**, 1333–1361, https://doi.org/10.1175/JCLI-D-13-00646.1.

Davis, C. A., 2018: Resolving tropical cyclone intensity in models. *Geophys. Res. Lett.*, **45**, 2082–2087, https://doi.org/10.1002/2017GL076966.

Elsberry, R. L., M. S. Jordan, and F. Vitart, 2010: Predictability of tropical cyclone events on intraseasonal timescales with the ECMWF monthly forecast model. *Asia-Pac. J. Atmos. Sci.*, **46**, 135–153, https://doi.org/10.1007/s13143-010-0013-4.

——, ——, and ——, 2011: Evaluation of the ECMWF 32-day ensemble predictions during the 2009 season of the western North Pacific tropical cyclone events on intraseasonal timescales. *Asia-Pac. J. Atmos. Sci.*, **47**, 305–318, https://doi.org/10.1007/s13143-011-0017-8.

——, H.-C. Tsai, and M. S. Jordan, 2014: Extended-range forecasts of Atlantic tropical cyclone events during 2012 using the ECMWF 32-day ensemble predictions. *Wea. Forecasting*, **29**, 271–288, https://doi.org/10.1175/WAF-D-13-00104.1.

Elsner, J. B., 2003: Tracking hurricanes. *Bull. Amer. Meteor. Soc.*, **84**, 353–356, https://doi.org/10.1175/BAMS-84-3-353.

——, and A. B. Kara, Eds., 1999: Hurricane cycles and trends. *Hurricanes of the North Atlantic: Climate and Society*, Oxford University Press, 240–278.

——, and B. Kocher, 2000: Global tropical cyclone activity: A link to the North Atlantic Oscillation. *Geophys. Res. Lett.*, **27**, 129–132, https://doi.org/10.1029/1999GL010893.

Fu, X., and P.-C. Hsu, 2011: Extended-range ensemble forecasting of tropical cyclogenesis in the northern Indian Ocean: Modulation of Madden–Julian Oscillation. *Geophys. Res. Lett.*, **38**, L15803, https://doi.org/10.1029/2011GL048249.

Gaffney, S. J., 2004: Probabilistic curve-aligned clustering and prediction with regression mixture models. Ph.D. thesis, Dept. of Information and Computer Science, University of California, 281 pp.

——, A. W. Robertson, P. Smyth, S. J. Camargo, and M. Ghil, 2007: Probabilistic clustering of extratropical cyclones using regression mixture models. *Climate Dyn.*, **29**, 423–440, https://doi.org/10.1007/s00382-007-0235-z.

Gao, K., J.-H. Chen, L. M. Harris, S.-J. Lin, B. Xiang, and M. Zhao, 2017: Impact of intraseasonal oscillations on the tropical cyclone activity over the Gulf of Mexico and Western Caribbean Sea in GFDL HiRAM. *J. Geophys. Res. Atmos.*, **122**, 13 125–13 137, https://doi.org/10.1002/2017JD027756.

——, ——, L. Harris, Y. Sun, and S.-J. Lin, 2019: Skillful prediction of monthly major hurricane activity in the North Atlantic with two-way nesting. *Geophys. Res. Lett.*, **46**, 9222–9230, https://doi.org/10.1029/2019GL083526.

Goddard, L., and M. Dilley, 2005: El Niño: Catastrophe or opportunity? *J. Climate*, **18**, 651–665, https://doi.org/10.1175/JCLI-3277.1.

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.

Gregory, P. A., J. Camp, K. Bigelow, and A. Brown, 2019: Subseasonal predictability of the 2017–2018 Southern Hemisphere tropical cyclone season. *Atmos. Sci. Lett.*, **20**, e886, https://doi.org/10.1002/asl.886.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, https://doi.org/10.1002/qj.3803.

Hirons, L. C., P. Inness, F. Vitart, and P. Bechtold, 2013a: Understanding advances in the simulation of intraseasonal variability in the ECMWF model. Part I: The representation of the MJO. *Quart. J. Roy. Meteor. Soc.*, **139**, 1417–1426, https://doi.org/10.1002/qj.2060.

——, ——, ——, and ——, 2013b: Understanding advances in the simulation of intraseasonal variability in the ECMWF model. Part II: The application of process-based diagnostics. *Quart. J. Roy. Meteor. Soc.*, **139**, 1427–1444, https://doi.org/10.1002/qj.2059.

Huang, B., and Coauthors, 2017: Extended Reconstructed Sea Surface Temperature, version 5 (ERSSTv5): Upgrades, validations, and intercomparisons. *J. Climate*, **30**, 8179–8205, https://doi.org/10.1175/JCLI-D-16-0836.1.

Jones, C., D. E. Waliser, J.-K. E. Schemm, and W. K. M. Lau, 2000: Prediction skill of the Madden–Julian Oscillation in dynamical extended range forecasts. *Climate Dyn.*, **16**, 273–289, https://doi.org/10.1007/s003820050327.

Klotzbach, P. J., 2014: The Madden–Julian oscillation impacts on worldwide tropical cyclone activity. *J. Climate*, **27**, 2317–2330, https://doi.org/10.1175/JCLI-D-13-00483.1.

——, and Coauthors, 2019: Seasonal tropical cyclone forecasting. *Trop. Cyclone Res. Rev.*, **8**, 134–149, https://doi.org/10.1016/j.tcrr.2019.10.003.

——, L.-P. Caron, and M. M. Bell, 2020: A statistical/dynamical model for the North Atlantic seasonal hurricane prediction. *Geophys. Res. Lett.*, **47**, e2020GL089357, https://doi.org/10.1029/e2020GL089357.

Kossin, J. P., and D. J. Vimont, 2007: A more general framework for understanding Atlantic hurricane variability and trends. *Bull. Amer. Meteor. Soc.*, **88**, 1767–1782, https://doi.org/10.1175/BAMS-88-11-1767.

——, S. J. Camargo, and M. Sitkowski, 2010: Climate modulation of North Atlantic hurricane tracks. *J. Climate*, **23**, 3057–3076, https://doi.org/10.1175/2010JCLI3497.1.

Kozar, M. E., M. E. Mann, S. J. Camargo, J. P. Kossin, and J. L. Evans, 2012: Statistical modeling of Atlantic tropical cyclone counts. *J. Geophys. Res.*, **117**, D18103, https://doi.org/10.1029/2011JD017170.

Landsea, C. W., 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.

Lang, A. L., K. Pegion, and E. A. Barnes, 2020: Introduction to special collection: Bridging weather and climate: Subseasonal-to-Seasonal (S2S) prediction. *J. Geophys. Res. Atmos.*, **125**, e2019JD031833, https://doi.org/10.1029/2019JD031833.

Lang, S. T. K., M. Leutbecher, and S. C. Jones, 2012: Impact of perturbation methods in the ECMWF ensemble prediction system on tropical cyclone forecasts. *Quart. J. Roy. Meteor. Soc.*, **138**, 2030–2046, https://doi.org/10.1002/qj.1942.

Lee, C.-Y., S. J. Camargo, F. Vitart, A. H. Sobel, and M. K. Tippett, 2018: Subseasonal tropical cyclone genesis prediction and MJO in the S2S dataset. *Wea. Forecasting*, **33**, 967–988, https://doi.org/10.1175/WAF-D-17-0165.1.

——, ——, ——, ——, J. Camp, S. Wang, M. K. Tippett, and Q. Yang, 2020: Subseasonal predictions of tropical cyclone occurrence and ACE in the S2S dataset. *Wea. Forecasting*, **35**, 921–938, https://doi.org/10.1175/WAF-D-19-0217.1.

Leutbecher, M., and Coauthors, 2017: Stochastic representations of model uncertainties at ECMWF: State of the art and future vision. *Quart. J. Roy. Meteor. Soc.*, **143**, 2315–2339, https://doi.org/10.1002/qj.3094.

Li, T., L. Wang, M. Peng, B. Wang, C. Zhang, W. Lau, and H. Kuo, 2018: A paper on the tropical intraseasonal oscillation published in 1963 in a Chinese journal. *Bull. Amer. Meteor. Soc.*, **99**, 1765–1779, https://doi.org/10.1175/BAMS-D-17-0216.1.

Liebmann, B., H. H. Hendon, and J. D. Glick, 1994: The relationship between tropical cyclones of the western Pacific and Indian Oceans and the Madden–Julian Oscillation. *J. Meteor. Soc. Japan*, **72**, 401–412, https://doi.org/10.2151/jmsj1965.72.3_401.

Lock, S.-J., S. T. K. Lang, M. Leutbecher, R. J. Hogan, and F. Vitart, 2019: Treatment of model uncertainty from radiation by the Stochastically Perturbed Parametrization Tendencies (SPPT) scheme and associated revisions in the ECMWF ensembles. *Quart. J. Roy. Meteor. Soc.*, **145**, 75–89, https://doi.org/10.1002/qj.3570.

Magnusson, L., and Coauthors, 2019: ECMWF activities for improved hurricane forecasts. *Bull. Amer. Meteor. Soc.*, **100**, 445–458, https://doi.org/10.1175/BAMS-D-18-0044.1.

Maloney, E. D., and D. L. Hartmann, 2000: Modulation of hurricane activity in the Gulf of Mexico by the Madden-Julian Oscillation. *Science*, **287**, 2002–2004, https://doi.org/10.1126/science.287.5460.2002.

Manganello, J. V., B. A. Cash, K. I. Hodges, and J. L. Kinter, 2019a: Seasonal forecasts of North Atlantic tropical cyclone activity in the North American Multi-Model Ensemble. *Climate Dyn.*, **53**, 7169–7184, https://doi.org/10.1007/s00382-017-3670-5.

——, ——, E. T. Swenson, and J. L. Kinter, 2019b: Assessment of climatology and predictability of Mid-Atlantic tropical cyclone landfalls in a high-atmospheric-resolution seasonal prediction system. *Mon. Wea. Rev.*, **147**, 2901–2917, https://doi.org/10.1175/MWR-D-19-0107.1.

Mariotti, A., and Coauthors, 2020: Windows of opportunity for skillful forecasts subseasonal to seasonal and beyond. *Bull. Amer. Meteor. Soc.*, **101**, E608–E625, https://doi.org/10.1175/BAMS-D-18-0326.1.

Merryfield, W. J., and Coauthors, 2020: Current and emerging developments in subseasonal to decadal prediction. *Bull. Amer. Meteor. Soc.*, **101**, E869–E896, https://doi.org/10.1175/BAMS-D-19-0037.1.

Moon, Y., D. Kim, S. J. Camargo, A. A. Wing, K. A. Reed, M. F. Wehner, and M. Zhao, 2020a: A horizontal resolution-dependent wind speed adjustment factor for tropical cyclones in climate model resolutions. *Geophys. Res. Lett.*, **46**, e2020GL087528, https://doi.org/10.1029/2020GL087528.

——, and Coauthors, 2020b: Wind and thermodynamic structures of tropical cyclones in global climate models and their sensitivity to horizontal resolution. *J. Climate*, **33**, 1575–1595, https://doi.org/10.1175/JCLI-D-19-0172.1.

Murakami, H., and M. Sugi, 2010: Effect of model resolution on tropical cyclone climate projections. *SOLA*, **6**, 73–76, https://doi.org/10.2151/sola.2010-019.

Nakamura, J., and Coauthors, 2017: Western North Pacific tropical cyclone model tracks in present and future climates. *J. Geophys. Res. Atmos.*, **122**, 9721–9744, https://doi.org/10.1002/2017JD027007.

Nakazawa, T., 1986: Intraseasonal variations of OLR in the tropics during the FGGE year. *J. Meteor. Soc. Japan*, **64**, 17–34, https://doi.org/10.2151/jmsj1965.64.1_17.

Ollinaho, P., and Coauthors, 2017: Towards process-level representation of model uncertainties: Stochastically perturbed parametrizations in the ECMWF ensemble. *Quart. J. Roy. Meteor. Soc.*, **143**, 408–422, https://doi.org/10.1002/qj.2931.

Patricola, C. M., P. Chang, and R. Saravanan, 2016: Degree of simulated suppression of Atlantic tropical cyclones modulated by flavour of El Niño. *Nat. Geosci.*, **9**, 155–160, https://doi.org/10.1038/ngeo2624.

——, S. J. Camargo, P. J. Klotzbach, R. Saravanan, and P. Chang, 2018: The influence of ENSO flavors on western North Pacific tropical cyclone activity. *J. Climate*, **31**, 5395–5416, https://doi.org/10.1175/JCLI-D-17-0678.1.

Pegion, K., and Coauthors, 2019: The Subseasonal Experiment (SubX): A multimodel subseasonal prediction experiment. *Bull. Amer. Meteor. Soc.*, **100**, 2043–2060, https://doi.org/10.1175/BAMS-D-18-0270.1.

Ramsay, H. A., S. J. Camargo, and D. Kim, 2012: Cluster analysis of tropical cyclone tracks in the Southern Hemisphere. *Climate Dyn.*, **39**, 897–917, https://doi.org/10.1007/s00382-011-1225-8.

——, S. S. Chand, and S. J. Camargo, 2018: A statistical assessment of Southern Hemisphere tropical cyclone tracks in climate models. *J. Climate*, **31**, 10 081–10 104, https://doi.org/10.1175/JCLI-D-18-0377.1.

Robertson, A. W., F. Vitart, and S. J. Camargo, 2020: Sub-seasonal to seasonal prediction of weather to climate with application to tropical cyclones. *J. Geophys. Res. Atmos.*, **125**, e2018JD029375, https://doi.org/10.1029/2018JD029375.

Satoh, M., and Coauthors, 2012: The Intra-Seasonal Oscillation and its control of tropical cyclones simulated by high-resolution global atmospheric models. *Climate Dyn.*, **39**, 2185–2206, https://doi.org/10.1007/s00382-011-1235-6.

Schreck, C. J., J. Molinari, and A. Aiyyer, 2012: A global view of equatorial waves and tropical cyclogenesis. *Mon. Wea. Rev.*, **140**, 774–788, https://doi.org/10.1175/MWR-D-11-00110.1.

Shaevitz, D. A., and Coauthors, 2014: Characteristics of tropical cyclones in high-resolution models of the present climate. *J. Adv. Model. Earth Syst.*, **6**, 1154–1172, https://doi.org/10.1002/2014MS000372.

Tsai, H.-C., R. L. Elsberry, M. S. Jordan, and F. Vitart, 2013: Objective verifications and false alarm analyses of western North Pacific tropical cyclone event forecasts by the ECMWF 32-day ensemble. *Asia-Pac. J. Atmos. Sci.*, **49**, 409–420, https://doi.org/10.1007/s13143-013-0038-6.

Vidale, P. L., K. Hodges, B. Vannière, P. Davini, M. J. Roberts, K. Strommen, and A. Weisheimer, 2021: Impact of stochastic physics and model resolution of tropical cyclones in climate GCMs. *J. Climate*, **34**, 4315–4341, https://doi.org/10.1175/JCLI-D-20-0507.1.

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

Vitart, F., 2003: Monthly forecasting system. ECMWF Tech. Memo. 424, https://doi.org/10.21957/idipp7l0x.

——, 2004: Monthly forecasting at ECMWF. *Mon. Wea. Rev.*, **132**, 2761–2779, https://doi.org/10.1175/MWR2826.1.

——, 2009: Impact of the Madden–Julian Oscillation on tropical storms and risk of landfall in the ECMWF forecast system. *Geophys. Res. Lett.*, **36**, L15802, https://doi.org/10.1029/2009GL039089.

——, 2014: Evolution of ECMWF sub-seasonal forecast skill scores. *Quart. J. Roy. Meteor. Soc.*, **140**, 1889–1899, https://doi.org/10.1002/qj.2256.

——, and J. L. Anderson, 2001: Sensitivity of Atlantic tropical storm frequency to ENSO and interdecadal variability of SSTs in an ensemble of AGCM integrations. *J. Climate*, **14**, 533–545, https://doi.org/10.1175/1520-0442(2001)014<0533:SOATSF>2.0.CO;2.

——, and F. Molteni, 2010: Simulation of the Madden–Julian Oscillation and its teleconnections in the ECMWF forecast system. *Quart. J. Roy. Meteor. Soc.*, **136**, 842–855, https://doi.org/10.1002/qj.623.

——, and A. W. Robertson, 2018: The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events. *npj Climate Atmos. Sci.*, **1**, 3, https://doi.org/10.1038/s41612-018-0013-0.

——, J. L. Anderson, and W. F. Stern, 1997: Simulation of interannual variability of tropical storm frequency in an ensemble of GCM integrations. *J. Climate*, **10**, 745–760, https://doi.org/10.1175/1520-0442(1997)010<0745:SOIVOT>2.0.CO;2.

——, ——, J. Sirutis, and R. E. Tuleya, 2001: Sensitivity of tropical storms simulated by a general circulation model to changes in

cumulus parameterization. *Quart. J. Roy. Meteor. Soc.*, **127**, 25–51, https://doi.org/10.1002/qj.49712757103.

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

——, and Coauthors, 2008: The new VarEPS-monthly forecasting system: A first step towards seamless prediction. *Quart. J. Roy. Meteor. Soc.*, **134**, 1789–1799, https://doi.org/10.1002/qj.322.

——, A. Leroy, and M. C. Wheeler, 2010: A comparison of dynamical and statistical predictions of weekly tropical cyclone activity in the Southern Hemisphere. *Mon. Wea. Rev.*, **138**, 3671–3682, https://doi.org/10.1175/2010MWR3343.1.

——, and Coauthors, 2017: The Subseasonal to Seasonal (S2S) prediction project database. *Bull. Amer. Meteor. Soc.*, **98**, 163–173, https://doi.org/10.1175/BAMS-D-16-0017.1.

Wheeler, M. C., and H. H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, **132**, 1917–1932, https://doi.org/10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2.

White, C. J., and Coauthors, 2017: Potential applications of Subseasonal-to-Seasonal (S2S) predictions. *Meteor. Appl.*, **24**, 315–325, https://doi.org/10.1002/met.1654.

Xiang, B., and Coauthors, 2015: Beyond weather time-scale prediction for Hurricane Sandy and Super Typhoon Haiyan in a global climate model. *Mon. Wea. Rev.*, **143**, 524–535, https://doi.org/10.1175/MWR-D-14-00227.1.