Predictive Skill and Predictability of North Atlantic Tropical Cyclogenesis in Different Synoptic Flow Regimes

ZHUO WANG AND WEIWEI LI

University of Illinois at Urbana–Champaign, Urbana, Illinois

MELINDA S. PENG

Marine Meteorology Division, Naval Research Laboratory, Monterey, California

XIANAN JIANG

Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, Los Angeles, California

RON MCTAGGART-COWAN

Numerical Weather Prediction Research Section, Environment and Climate Change Canada, Dorval, Quebec, Canada

CHRISTOPHER A. DAVIS

National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 29 March 2017, in final form 30 August 2017)

ABSTRACT

Practical predictability of tropical cyclogenesis over the North Atlantic is evaluated in different synoptic flow regimes using the NCEP Global Ensemble Forecast System (GEFS) reforecasts with forecast lead time up to two weeks. Synoptic flow regimes are represented by tropical cyclogenesis pathways defined in a previous study based on the low-level baroclinicity and upper-level forcing of the genesis environmental state, including nonbaroclinic, low-level baroclinic, trough-induced, weak tropical transition (TT), and strong TT pathways. It is found that the strong TT and weak TT pathways have lower predictability than the other pathways, linked to the lower predictability of vertical wind shear and midlevel humidity in the genesis vicinity of a developing TT storm. Further analysis suggests that stronger extratropical influences contribute to lower genesis predictability. It is also shown that the regional and seasonal variations of the genesis predictive skill in the GEFS can be largely explained by the relative frequency of occurrence of each pathway and the predictability differences among pathways. Predictability of tropical cyclogenesis is further discussed using the concept of the genesis potential index.

1. Introduction

Skillful prediction of tropical cyclones (TCs) is of significant socioeconomic value because of the potentially hazardous impacts of the storms. Compared to TC track and intensity forecasts, tropical cyclogenesis forecasts have received less attention. Accurate prediction of the genesis time and location, however, is critical for the extended-range forecasting of tropical cyclone track and intensity. In particular, timely and skillful prediction of tropical cyclogenesis near the U.S. coast is important for storm preparedness.

Numerical models are important tools in operational forecasting (e.g., Elsberry et al. 2009; Halperin

Corresponding author: Zhuo Wang, zhuowang@illinois.edu

Publisher's Note: This article was updated on 16 August 2018 to replace an older version of Fig. 3 that was mistakenly used when originally published.

DOI: 10.1175/JAS-D-17-0094.1

et al. 2013; Li et al. 2016). Previous studies have shown that the predictive skill of tropical cyclogenesis in numerical models varies from case to case (Komaromi and Majumdar 2015). The formation of some storms can be predicted skillfully more than one week in advance (e.g., Xiang et al. 2015a; Elsberry et al. 2011; Tsai and Elsberry 2013), while the 48-h forecasts of other storms, such as Hurricane Fay (Kimberlain 2014) and Tropical Storm Hanna (Cangialosi 2014) in 2014, can be a challenge even for the state-of-the-art operational prediction systems. The predictive skill of an operational model has substantial variations on the subseasonal and interannual time scales (Li et al. 2016; Komaromi and Majumdar 2015) and also varies from basin to basin (Halperin et al. 2016). Such variations cannot be completely attributed to changes in a prediction system or the heterogeneity of observational data assimilated in a model. Instead, the spatial and temporal variations of predictive skill often reflect variations in TC predictability: Some TCs are more predictable than others. Predictability, the extent to which the future states of a system may be predicted based on the knowledge of the current and past states of the system (American Meteorological Society 2012), depends on the available observational data and prediction models and also on the intrinsic dynamic and physical nature of the phenomenon of interest. A better understanding of the predictability of tropical cyclogenesis may not only help to identify the key processes involved in TC genesis but may also provide useful information on the reliability of operational prediction and thus aid more effective use of forecast products.

Tropical cyclone formation involves multiscale interactions among complex physical processes ranging from the convective scale to the planetary scale (Gray 1998). It is generally believed that convective processes limit the predictability of tropical cyclogenesis, while slowly varying, large-scale processes may serve as important sources of predictability. For example, the Madden–Julian oscillation (MJO), the dominant mode of intraseasonal variability in the tropics (Madden and Julian 1972; Zhang 2005), is a major source of TC predictability on the subseasonal time scale (National Research Council 2010). Owing to its quasi-periodic behavior and strong impacts on TCs, the MJO has been used as an important predictor in many statistical prediction models for tropical cyclogenesis (e.g., Leroy and Wheeler 2008; Slade and Maloney 2013). A numerical model that is skillful in MJO prediction is often found skillful in subseasonal TC prediction (Xiang et al. 2015a,b; Barnston et al. 2015). Other sources of predictability for tropical cyclogenesis include, but are not limited to, El Niño–Southern Oscillation (ENSO; e.g., Goldenberg and Shapiro 1996), the Atlantic Meridional Mode (e.g., Kossin and Vimont 2007), and the Atlantic regional Hadley circulation (Zhang and Wang 2013), which all strongly modulate Atlantic TCs on the interannual time scale. Li et al. (2016) showed that the subseasonal predictive skill of TCs tends to be higher in El Niño and La Niña years than in the ENSO-neutral years. Wang et al. (2015) attributed the skillful seasonal prediction of the Atlantic basinwide hurricane frequency in a global atmospheric model (Chen and Lin 2013) to the high predictability of the Atlantic regional Hadley circulation.

Moving to shorter temporal scales or smaller spatial scales, a preexisting, cyclonic synoptic disturbance is one of the necessary conditions for TC formation (e.g., Gray 1968). Tropical cyclones can develop from tropical easterly waves (e.g., Landsea 1993; Dunkerton et al. 2009), equatorial Rossby waves (Lussier 2010), monsoon troughs or monsoon gyres (e.g., Wu et al. 2013), disturbances resulting from the intertropical convergence zone (ITCZ) breakdown (e.g., Guinn and Schubert 1993; Wang and Magnusdottir 2005), Rossby wave dispersion from a preexisting cyclone (e.g., Li and Fu 2006), or subtropical frontal systems (the so-called tropical transition; Davis and Bosart 2003, 2004). Synoptic-scale precursor disturbances serve as the incubator for the development of a TC (e.g., Dunkerton et al. 2009; Montgomery et al. 2010; Wang et al. 2010a,b, 2012) and control the immediate environmental conditions for genesis. Different types of precursors are characterized by different synoptic-scale environmental states, and their interactions with an incipient TC vortex may be different. It is conceivable that tropical cyclogenesis in different synoptic flow regimes may be associated with different levels of predictability.

Although Lorenz (1965) demonstrated, more than 50 years ago, that predictability is dependent on flow regime, predictability of tropical cyclogenesis for different synoptic flow regimes has not been well studied. Based on the upper-level forcing and low-level baroclinicity of the genesis environment, McTaggart-Cowan et al. (2013) identified five tropical cyclogenesis pathways (or development pathways) that represent different synoptic-scale flow regimes, including nonbaroclinic, low-level baroclinic, trough-induced (TI), weak tropical transition, and strong tropical transition pathways (see more discussion in section 3). Built upon this pathway concept, we will test the following hypothesis: tropical cyclogenesis in different synoptic flow regimes is associated with different levels of predictability.

Predictability of tropical cyclogenesis will be examined in a reforecast dataset. The dataset and predictability metrics are described in section 2. A brief review of the genesis pathways is provided in section 3. Predictability of tropical cyclogenesis is examined in section 4, followed by a summary and discussion in section 5.

2. Data and methodology

a. Data

The predictability of tropical cyclogenesis over the North Atlantic was examined in the Global Ensemble Forecast System (GEFS) Reforecast, version 2 (GEFS-R2), during July–October 1985–2012. The GEFS-R2 has 11 ensemble members and was initialized once per day at 0000 UTC with the Climate Forecast System Reanalysis (CFSR; Saha et al. 2010) from 1985 to February 2011 and with the analysis from the Gridpoint Statistical Interpolation analysis system (GSI) after February 2011 (Hamill et al. 2013). The horizontal resolution of the model is T254 for first week reforecasts and is reduced to T190 beyond one week. All variables were coarsened to a 1° × 1°-resolution grid mesh before our analyses.

We focus on the North Atlantic basin. The Geophysical Fluid Dynamics Laboratory (GFDL) vortex tracker (Marchok 2002) was used to track TCs in the GEFS-R2 (Li et al. 2016). A tropical cyclone was identified as a warm-core, cyclonic vortex with the 10-m maximum wind speed exceeding 16.5 m s^{-1} (the wind speed threshold was adjusted based on the data resolution; Walsh et al. 2007). To increase the robustness of the results, short-lived TCs (lifetime less than 72 h) were excluded in both the observation and the GEFS-R2. Storms forming poleward of 40°N were regarded as extratropical cyclones and were excluded as well. Tropical cyclones were tracked in individual ensemble members and evaluated against the International Best Track Archive for Climate Stewardship (IBTrACS; Knapp et al. 2010). A hit was defined for a model TC that forms within ±120 h of the observed genesis time and within a 5° radius of the observed TC track,[1] while the other model TCs were categorized as false alarms. Additionally, a miss was flagged when the model fails to produce a hit for an observed TC, and a nonevent or correct negative was identified if neither an observed genesis event nor a predicted one occurs.

b. Metrics

Metrics of predictive skill are used to estimate predictability in this study. Predictability was often evaluated by error growth rate in early studies (e.g., Lorenz 1965, 1969; Smagorinsky 1969). As ensemble prediction becomes the common practice in recent decades, ensemble spread is often used as a metric of predictability, which indicates the error growth rate or the sensitivity of a prediction system to initial condition errors. Metrics of ensemble spread have been employed to evaluate the predictability of tropical cyclone intensity or track in some previous studies (Zhang and Tao 2013; McMurdie and Ancell 2014). An ensemble spread metric, however, is not readily applicable to tropical cyclogenesis, as the dichotomy between genesis and nongenesis cannot be determined using a single variable. A vortex tracker, such as the one described in section 2a, is often employed to identify tropical cyclogenesis in a numerical model by examining multiple variables.

Because the predictive skill of a model is dependent on the practical predictability of the phenomenon of interest, model forecasts can be used to estimate practical predictability under the assumption that higher predictive skill suggests a higher level of practical predictability. Using an idealized ensemble prediction system, Grimit and Mass (2007) showed that ensemble spread (a measure of predictability) is related to forecast accuracy (measured by predictive skill). More specifically, the probability of large forecast errors increases with increasing ensemble spread although the ensemble spread–error relation cannot be well represented by a simple linear correlation. Furthermore, practical predictability is related to model deficiencies, uncertainties in initial conditions, and intrinsic predictability (Palmer 2006). Since the GEFS-R2 was carried out using a fixed version of the GEFS and initialized primarily with the CFSR during the time period of analysis (1985–2012), different genesis predictive skills reflect the different levels of intrinsic predictability for different pathways, with the limitation that some results may be model dependent.

Several metrics were employed to evaluate the predictive skill of the GEFS-R2. Hit rate (H), false alarm rate (F), false alarm ratio (FAR), and critical success index (CSI) were computed based on the 2 × 2 contingency table as below:

$$H = \frac{a}{a + c}, \tag{1}$$

[1] The time window is chosen to include early genesis and late genesis (the predicted genesis time of a storm is too early or too late compared to the observation; see Halperin et al. 2013; Li et al. 2016). Reducing the time window to 2–3 days reduces the hit rate but does not qualitatively affect the results on different levels of predictability associated with different pathways.

TABLE 1. Characteristics and frequency of occurrence of different genesis pathways [adapted from Table 2 in McTaggart-Cowan et al. (2013)]. Also shown are the vertical shear (defined as the magnitude of the zonal wind difference between 200 and 850 hPa) and 700-hPa RH, which are averaged over a 20° × 20° box centered at the genesis location from the IBTrACS to represent the environmental state.

Pathway	Nonbaroclinic	Low-level baroclinic	Trough induced	Weak TT	Strong TT
Upper-level forcing	Low	Low	High	High	High
Low-level baroclinicity	Low	High	Low	Medium	High
Occurrence (%)	38	12	10	26	14
Vertical shear (m s^{-1})	3.8	4.7	4.2	6.8	11.7
700-hPa RH (%)	67.4	67.3	66.5	63.6	54.0

$$F = \frac{b}{b+d}, \tag{2}$$

$$\mathrm{FAR} = \frac{b}{a+b}, \tag{3}$$

$$\mathrm{CSI} = \frac{a}{a+b+c}, \tag{4}$$

where a, b, c, and d are the numbers of hits, false alarms, misses, and correct negatives, respectively. The hit rate, or the probability of detection, is the proportion of the occurrences that are correctly forecast. The false alarm ratio is the proportion of genesis forecasts that turn out to be false alarms, and the false alarm rate is the ratio of false alarms to the total number of nonoccurrences (i.e., $b + d$). The critical success index, also known as the threat score, is the number of hits divided by the total number of occasions when the event is forecast and/or observed, which is particularly useful when the occurrence is substantially less frequent than the nonoccurrence (Wilks 2006).

Selected environmental variables related to tropical cyclogenesis in the GEFS-R2 were evaluated against the CFSR. The CFSR were mapped to a grid mesh of 1.0° × 1.0° resolution in keeping with the GEFS-R2. The predictive skill of environmental variables was examined by computing the root-mean-square error (RMSE). RMSE was first derived for individual ensemble members, and then the ensemble mean RMSE was taken. A good agreement was found among ensemble members, and only the ensemble means are discussed below for brevity.

3. Tropical cyclogenesis pathways

McTaggart-Cowan et al. (2013) employed two metrics to characterize the environmental state (instead of the precursor vortex) for tropical cyclone development, and five tropical cyclogenesis pathways were categorized based on a linear discriminant analysis (LDA) of the two metrics, including nonbaroclinic (NBC), low-level baroclinic (LBC), trough-induced, weak tropical transition (TT), and strong TT pathways (Table 1 and Fig. 1). The categorization of the pathways is briefly described below, and readers are referred to McTaggart-Cowan et al. (2013, 2008) for more details.

The two chosen metrics, Q and Th, are representative of the synoptic-scale, near-storm environment and are dynamically significant with respect to tropical cyclogenesis theories (McTaggart-Cowan et al. 2008). The Q metric is defined as the average convergence of the 400–200-hPa Q vector within 6° of the point of interest. The Q vector is defined as

$$\mathbf{Q} = -\frac{R}{\sigma p}\begin{pmatrix} \frac{\partial \mathbf{V}_{\mathrm{nd}}}{\partial x}\cdot\nabla_p T \\ \frac{\partial \mathbf{V}_{\mathrm{nd}}}{\partial y}\cdot\nabla_p T \end{pmatrix}, \tag{5}$$

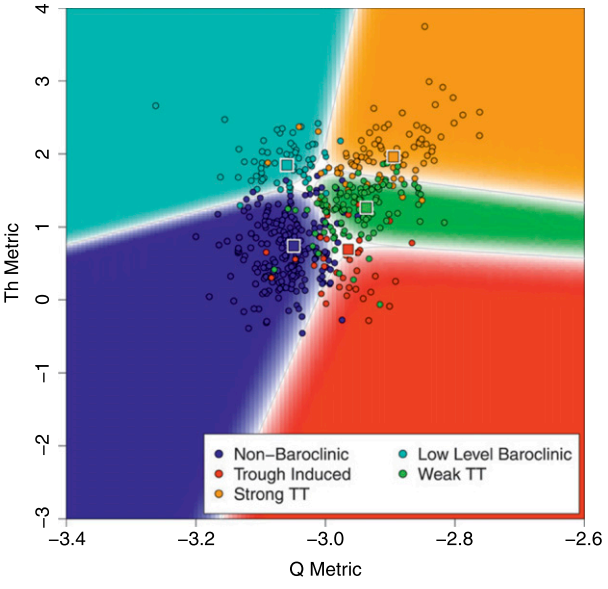

FIG. 1. Classification of North Atlantic TC pathways in the metric space. The event dots are color coded according to the latent trajectory model (LTM) class diagnosed by McTaggart-Cowan et al. (2008) as indicated in the legend. The same color scheme is used for the LDA-based background, which represents the dominant pathway in each sector of metric space scaled to white with decreasing posteriors. When a dot lies on the background of the same color, the LDA reclassification reproduces the original LTM result. The centroids of the classes are shown by color-coded squares in each panel (from McTaggart-Cowan et al. 2013).

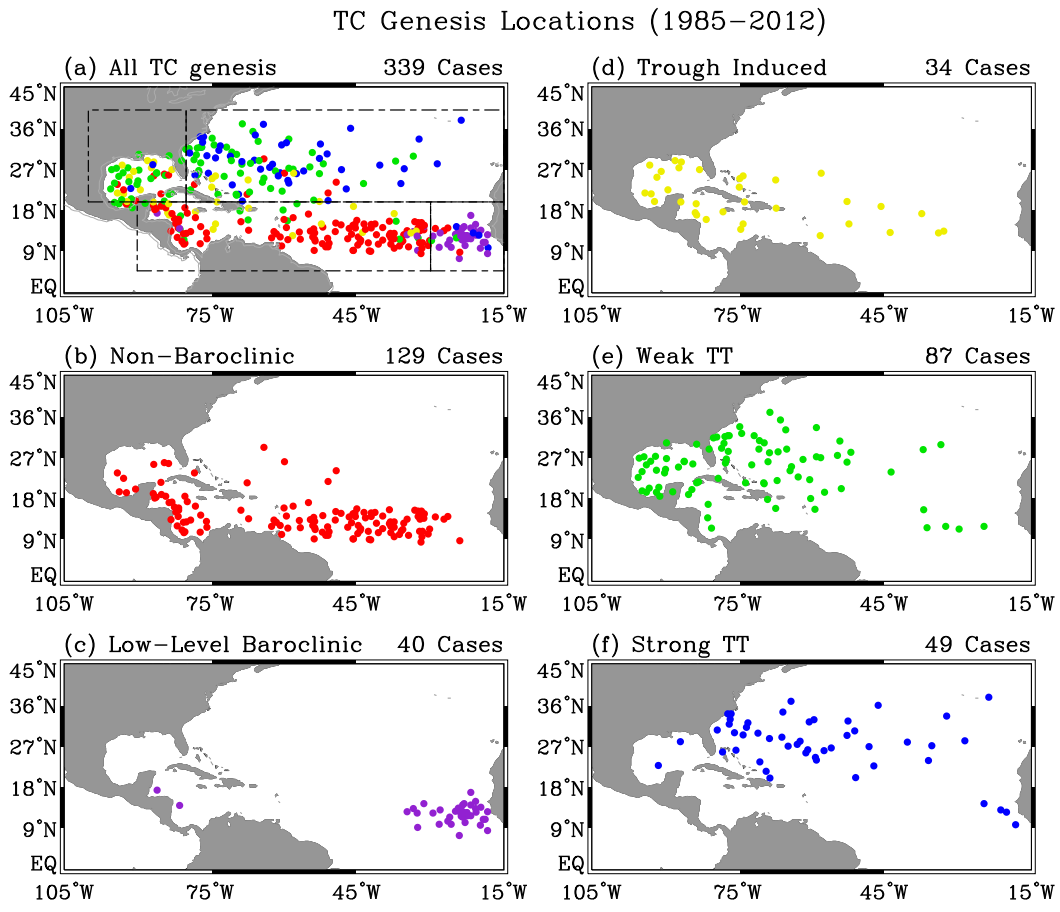

FIG. 2. Genesis distribution for (a) all analyzed Atlantic tropical cyclones and for (b)–(f) different genesis pathways. The boxes in (a) represent the different regions defined for the Atlantic basin (see text for more details).

where $\mathbf{V}_{nd}$ denotes nondivergent wind; R is the gas constant of dry air; p and T are pressure and air temperature, respectively; σ is the static stability; and ∇_P denotes horizontal gradient on a pressure surface. The Q vector was calculated using the nondivergent wind component to deal with the high–Rossby number flow in the tropics and to reduce noise associated with the irrotational wind component. The Q metric represents the synoptic-scale forcing for ascent, associated with an upper-level trough or a tropical upper-tropospheric trough (TUTT) cell. The second metric, Th, represents the low-level baroclinicity and is defined as the maximum difference in 1000–700-hPa thickness between two semicircles within 10° of the point of interest. Since the two metrics represent the environmental state and do not require the presence of a vortex, they can be evaluated at each grid point and at each analysis time of a reanalysis dataset. An LDA was employed to categorize five genesis pathways, and the pathway of a TC is determined based on the values of the metrics at the genesis location.

The NBC pathway is characterized by weak upper-level forcing and weak low-level baroclinicity (Fig. 1 and Table 1). This pathway is the most frequent genesis pathway and occurs preferentially over the Atlantic main development region (MDR; Goldenberg et al. 2001; Fig. 2b). It conforms to the typical tropical cyclogenesis scenario associated with a tropical easterly wave (e.g., Dunkerton et al. 2009).

The LBC pathway mainly occurs off the coast of West Africa (Fig. 2c), near the Cape Verde Islands (the so-called Cape Verde storms). It is characterized by weak upper-level forcing and strong low-level baroclinicity (Fig. 1 and Table 1). The low-level baroclinicity is closely related to the African easterly jet and results from the thermal contrast between the cool maritime boundary layer in the south and the hot desert air mass in the north extending from West Africa to the east Atlantic. There are two groups of waves over West Africa, one to the north of the jet and the other to the south (referred to as northern and southern waves, respectively), which have different dynamic and

thermodynamic structures (e.g., Reed et al. 1977; Pytharoulis and Thorncroft 1999). Because of the presence of the strong low-level baroclinicity, the LBC pathway is likely involved with the interaction between a southern wave and a northern disturbance (Hankes et al. 2015). As shown in Hankes et al. (2015), the merger of a southern wave and a norther wave creates a deeper and stronger wave pouch conducive to tropical cyclogenesis. It is worth pointing out that this pathway is still a "tropical only" pathway despite the low-level baroclinicity.

The trough-induced, weak TT, and strong TT pathways are all subject to upper-level forcing and are characterized by weak, moderate, and strong low-level baroclinicity, respectively (Fig. 1 and Table 1). The upper-level forcing may be associated with cutoff lows or upper-level troughs (Bentley et al. 2017), which are often of extratropical origin. However, it is necessary to note that the presence of the upper-level forcing does not mean that low-level tropical disturbances, such as tropical easterly waves, do not play a role in the formation of a TC (Galarneau et al. 2015). In some cases, low-level disturbances may play a dominant role in TC development (see discussion of Tropical Storm Matthew in section 4d). The trough-induced, weak TT, and strong TT pathways can thus be regarded as hybrid pathways or extratropical pathways, in contrast to the "tropical only" nature of the NBC and LBC pathways.

The trough-induced pathway is the least frequent genesis pathway, and the associated developments scatter over the central and west Atlantic (Fig. 2d). The weak TT and strong TT pathways occur more poleward than the trough-induced pathway, in keeping with the stronger low-level baroclinicity (Figs. 2e,f). The low-level baroclinicity in the weak TT and strong TT pathways may be associated with strong sea surface temperature gradient or a low-level jet. In addition, remnant cold fronts in the subtropics can enhance low-level temperature and moisture gradients as well (Davis and Bosart 2004; Zhang et al. 2017). In other words, the Q metric and the Th metric are not completely independent of each other.

4. Evaluation of tropical cyclogenesis predictability

a. *Predictability for different pathways*

The composite mean hit rate[2] for each pathway is shown in Fig. 3. Bootstrapping was used to test the significance of the differences from random sampling. For example, the total number of storms was 339, including 129 storms associated with the NBC pathway. To test the significance for the NBC pathway, 10 000 bootstrap samples were constructed, and each sample included 129 cases randomly chosen from the 339 storms. The hit rate was then calculated for each sample. If the hit rate value of the NBC pathway was above the top 5th percentile or below the bottom 5th percentile of the 10 000 bootstrap estimates, the hit rate was regarded as significantly different from random sampling. The same test was done for each pathway at different forecast lead times. As shown in Fig. 3, the LBC pathway has the highest hit rate among the five pathways, and the hit rate exceeds the top 5th percentile at all forecast lead times. The strong TT pathway has the lowest hit rate among the five pathways at all forecast lead times. The hit rates of strong TT and weak TT are significantly lower than random sampling from days 2–3 to days 8–9. The NBC and trough-induced pathways have similar skill, and neither is significantly different from random sampling.

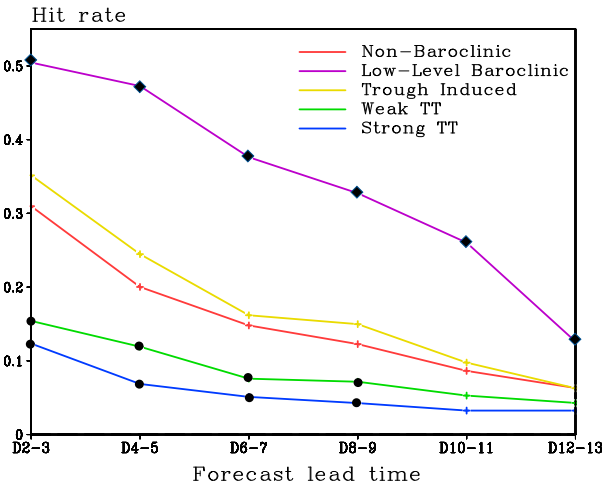

FIG. 3. Hit rate as a function of the forecast lead time for different genesis pathways (represented by different colors as shown in the legend). Closed circles (diamonds) represent hit rate falling below (above) the bottom (top) 5th percentile of random sampling.

The different predictive skills of tropical cyclogenesis associated with different pathways can partly be attributed to the different predictive skill of environmental variables. The RMSEs of vertical wind shear (defined as the magnitude of the vector wind difference between 200 and 850 hPa) and 700-hPa relative humidity (RH) were calculated with respect to the CFSR. For each observed storm in the IBTrACS, RMSE was evaluated at the observed genesis time over a 20° × 20° grid box, which is about half wavelength of a typical tropical

[2] We did not evaluate the false alarm rate or the false alarm ratio for different genesis pathways because the model environment may deviate from reality, and the pathways defined based on the observed environment cannot be simply used to categorize false alarm storms in the reforecasts.

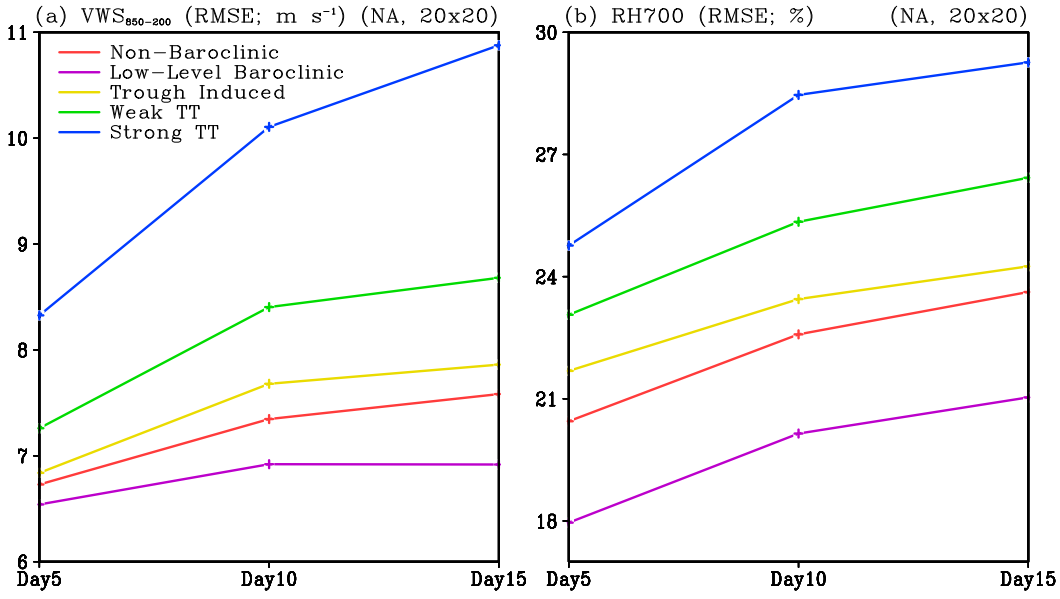

FIG. 4. RMSE of (a) vertical wind shear and (b) 700-hPa RH (see the text for details). Different colors represent different genesis pathways.

easterly wave. Since the predicted genesis can occur anywhere within a 5° radius of the observed genesis location for a "hit" event, we chose a fixed 20° × 20° box centered at the observed genesis location in the CFSR but allowed the center of the 20° × 20° box in the GEFS-R2 to move within a 10° × 10° domain centered at the observed genesis location. RMSE was calculated for each possible location. Given the 1° × 1° resolution, this yields 100 values of RMSE for a given genesis event. The minimum RMSE among the 100 values was chosen to represent a lower bound of the forecast errors. The same calculation was repeated for all GEFS ensemble members, and the ensemble mean of the minimum RMSE was taken to represent the predictive skill of the variable for a genesis event. The composite mean of the RMSE was then calculated for each genesis pathway. Because of the large volume of the dataset, only three representative forecast lead times—5, 10, and 15 days—were examined (Fig. 4).

The RMSE increases sharply from $t = 0$ to 5 days for both vertical wind shear and RH (note that the RMSE is zero at $t = 0$). Because of the different error growth rates for the different pathways, the magnitude of the RMSE, of both vertical shear and RH, shows the same sequence at different forecast lead times: strong TT > weak TT > trough induced > NBC > LBC. A larger RMSE is due to the larger error growth rate and indicates lower predictability. The predictive skill or predictability of the environmental variables is consistent with the genesis predictive skill for different pathways shown in Fig. 3 except for the relative order of the NBC and trough-induced pathways. Although the largest RMSE of vertical wind shear for the strong TT pathway may be partly attributed to the strong environmental vertical wind shear that the pathway is subject to (Table 1), the magnitude of the environmental RH does not explain the large RMSE of 700-hPa RH associated with the strong TT and weak TT pathways (Table 1). Instead, the TT pathways generally occur at higher latitudes than the other pathways (Fig. 2), and the large RMSE may be attributed to the influences of extratropical disturbances (see more discussion in section 4b).

We also examined column water vapor and 850-hPa relative vorticity. The relative magnitude of the RMSE of column water vapor for the five pathways has the same sequence as that of 700-hPa RH and vertical wind shear. The five pathways, however, are not very well separated in the RMSE of 850-hPa relative vorticity. This is probably because relative vorticity has very limited predictability for all pathways. Komaromi and Majumdar (2014) showed that variables related to large-scale, slowly evolving phenomena (such as vertical shear and the upper-level velocity potential) are more predictable that those inherently related to small-scale, rapidly evolving features (such as the low-level vorticity and upper-level divergence).

The different hit rates in Fig. 3 suggest different levels of genesis predictability for different pathways, which, from high to low, have the following order: LBC > trough induced > NBC > weak TT > strong TT. The higher predictability of the LBC pathway than the NBC pathway may be related to the role of the interaction between northern and southern waves in the

development of an LBC storm.[3] Hankes et al. (2015) showed that the merger of a northern wave and a southern wave leads to a stronger and deeper wave pouch off the coast, which is more conducive to tropical cyclogenesis. It is possible that the interaction between two synoptic waves enhances both the likelihood and predictability of tropical cyclogenesis. In addition, the interaction of southern and northern disturbances occurs preferentially near the coast in both the observation and the reforecasts. This geographic preference also contributes to a higher hit rate by reducing the errors in predicted genesis locations.

It was a surprise that the hit rate of the NBC pathway is lower than that of the trough-induced pathway, as the real-time wave tracking based on global model forecasts from multiple operational centers in the past several years (http://www.met.nps.edu/~mtmontgo/storms2008.html) gave us the impression that TCs originating from tropical easterly waves in the Atlantic MDR are more predictable than those forming farther west in the basin. It is possible that the low hit rate of the NBC pathway is due to the systematic genesis biases in the GEFS. Li et al. (2016) found that tropical easterly waves are too strong and too deep over West Africa in the GEFS reforecasts compared to the ERA-Interim. As a result, some waves develop into TCs shortly after moving over ocean. This leads to a positive bias in tropical cyclogenesis density function off the coast of West Africa and a negative bias farther downstream. The negative bias may explain the relatively low hit rate for the NBC pathway, while the positive bias near the coast may contribute to the high hit rate of the LBC pathway.

The strong TT pathway has the lowest predictability, which can be partly attributed to the high environmental vertical wind shear associated with the pathway (Davis and Bosart 2004). Zhang and Tao (2013), using idealized simulations, showed that environmental vertical shear affects the predictability of TC formation and intensity. However, vertical shear does not explain the predictability differences among the other pathways because the environmental vertical wind shear, derived from the CFSR, is similar among the LBC, NBC, trough-induced, and weak TT pathways (Table 1). On the other hand, it is instructive to recall that the strong TT and weak TT pathways, with relatively low predictability, are of hybrid or extratropical nature as they are associated with upper-level troughs or cutoff lows originating from the extratropics. At the short time scale (within the forecast lead time of 1–2 days), forecast errors grow faster in the tropics than in the extratropics probably because of active moist convection and the lack of quasigeostrophic constraint in the tropics. However, at the longer time scales, the strong coupling between the atmosphere and ocean in the tropics limits the error growth (Charney and Shukla 1981; Palmer 1996), while error growth continues in the extratropics through baroclinic and barotropic processes and upscale energy cascade (e.g., Métais et al. 1994; Straus and Paolino 2009). The extratropical atmosphere thus has a large error growth rate and hence lower predictability than the tropical atmosphere at forecast lead times beyond 1–2 days [see Fig. 10 in Davis et al. (2016)], and it is conceivable that TCs developing under stronger extratropical influences (as the strong TT and weak TT pathways) are intrinsically less predictable than those developing in a purely tropical environment at this time scale.

b. *Possible extratropical influences on predictability*

In this section, we further address the question whether stronger extratropical influence necessarily implies lower genesis predictability (or lower predictive skill). The Q metric can be used as a proxy for extratropical influences. Although the strong TT, weak TT, and trough-induced pathways are separated primarily based on the strength of the low-level baroclinicity (Fig. 5a; also see McTaggart-Cowan et al. 2013), Fig. 5b shows that stronger upper-level forcing occurs most frequently in the strong TT pathway and least frequently in the trough-induced pathway, consistent with the co-dependency of the two metrics mentioned in section 3. As noted by Zhang et al. (2017), a strong upper-level potential vorticity feature may induce or enhance the low-level baroclinicity. Bentley et al. (2017) showed that the strong TT pathway is often associated with a stronger upper-level disturbance, such as a cutoff low or a meridional trough. Such upper-level disturbances may either extend to the lower troposphere and initiate the low-level cyclonic circulation or interact with low-level precursor disturbances (Davis and Bosart 2004; Galarneau et al. 2015).

To examine further the possible impacts of upper-level disturbances of extratropical origin on tropical cyclogenesis predictability, we combined the strong TT and weak TT storms and separated them into the strong-Q

[3] Hankes et al. (2015) showed that merger developers, compared to nonmerger developers, are associated with a stronger heat low extending from West Africa to the east Atlantic (i.e., enhanced low-level baroclinicity). We also found that 31% of the LBC storms develop out of the merger of a southern wave and a northern wave from 1990 to 2010, while only 7% of NBC storms are merger developers. Hankes et al. (2015) used a rather stringent definition for the merger of two disturbances and did not consider stationary or weak vorticity disturbances north of the jet. It is possible that a large fraction of the LBC storms were involved with the interaction between a southern wave and a northern disturbance if a looser definition were used.

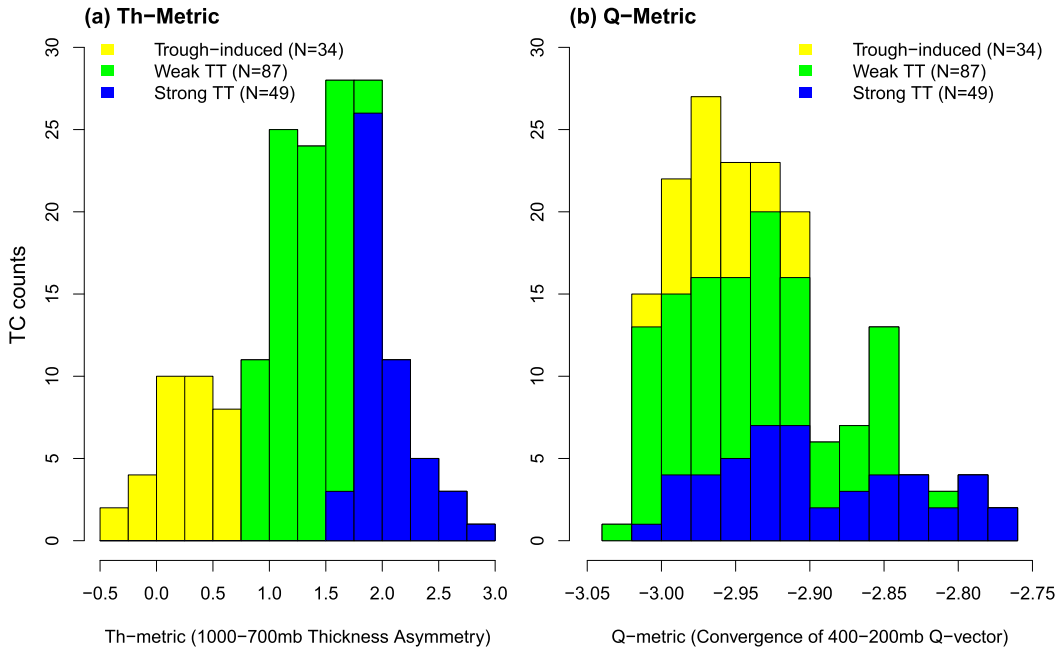

FIG. 5. Histograms of the low-level baroclinicity (Th metric) and the upper-level forcing (Q metric) for the TI, weak TT, and strong TT pathways. Note that these are stacked bar charts in which one bar is stacked above another and does not go behind other bars.

and weak-Q groups by the median of the Q metric. The two groups had the same sample size of 68. The hit rate was calculated for each group, and the significance of the hit rate difference between the two groups was examined using a permutation test. The permutation test consists of 10 000 drawings. In each drawing, a group of 68 storms were randomly chosen from the pool of 136 storms, and the remaining storms were treated as another group. The hit rate for each group was calculated, and the difference in the hit rate between the two groups was recorded. The 10 000 drawings thus produced 10 000 outcomes of the hit rate difference. The difference in the hit rate between the strong-Q and weak-Q groups is regarded significant if it exceeds the top 5th percentile of the 10 000 drawings. The same test was repeated for all forecast lead times.

As show in Fig. 6a, the strong-Q group indeed has a lower hit rate than the weak-Q group, and the difference is significant at all forecast lead times except at day 12–13. We repeated the same calculation by defining two groups after combining the strong TT, weak TT, and trough-induced pathways (Fig. 6b). The three-pathway combination has the advantage of a larger sample size (170), but it should be borne in mind that the trough-induced pathway differs from the strong TT and weak TT pathways in many aspects (McTaggart-Cowan et al. 2013). Again, the strong-Q group has a lower hit rate than the weak-Q group, and the difference is significant at all forecast lead times. We also examined the strong TT and weak TT pathways separately. Although a strong-Q group always has a lower hit rate than a weak-Q group, the difference does not exceed the 5th-percentile threshold at some forecast lead times because of the small sample sizes of the pathways (not shown). Overall, the calculations suggest that stronger extratropical influences imply lower predictability of tropical cyclogenesis.

c. *Geographic and seasonal variations of predictability*

Previous studies have reported the geographic variations of model forecasting skill for tropical cyclogenesis (e.g., Halperin et al. 2013, 2016). In fact, the spatial variations of the model forecast skill are so large in operational models that longitude and latitude, despite of their lack of explicit physical meaning, were selected as two major predictors in a hybrid prediction scheme to reduce the biases of a numerical model by Halperin et al. (2017). Figure 7 shows the hit rate, false alarm rate, false alarm ratio, and critical success index for four regions over the Atlantic: the eastern MDR (EMDR), the central and western MDR (CWMDR), the Gulf of Mexico (Gulf), and the subtropical Atlantic (SubAtl; delineated in Fig. 2a), which are selected based on the geographic distribution of the pathways. The hit rates are clearly separated among the four regions: EMDR > CWMDR > Gulf > SubAtl

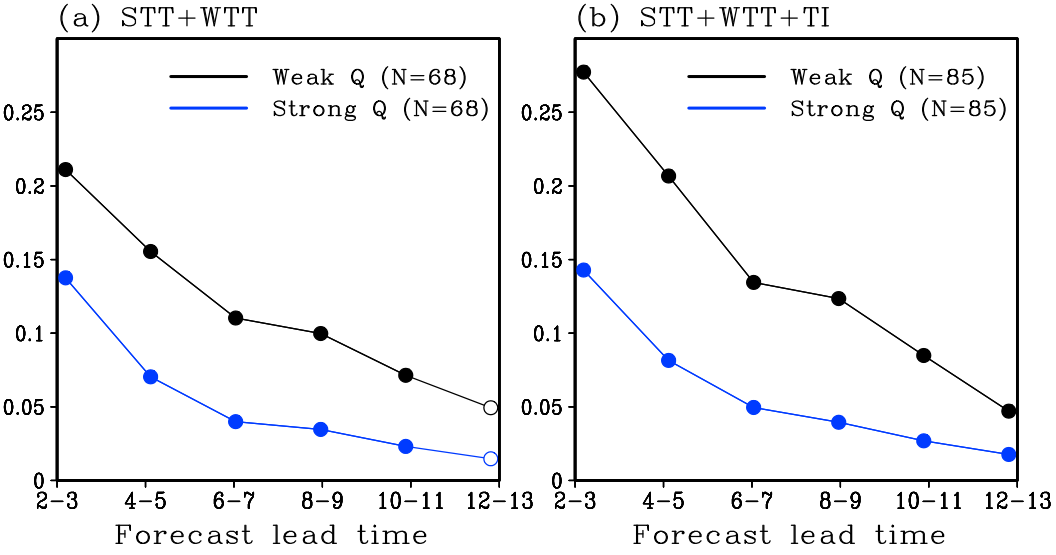

FIG. 6. Hit rate as a function of the forecast lead time for strong-Q (blue) and weak-Q groups (black) based on (a) the combination of strong TT and weak TT pathways and (b) the combination of strong TT, weak TT, and TI pathways. The sample size N for each group is indicated in the figure legends. Closed circles represent that the hit rates between the two groups are significantly different from each other based on a permutation test (see text for more information).

(Fig. 7a). The false alarm rate in the EMDR and CWMDR is much higher than that in the other two regions (Fig. 7b), and the EMDR hit rate is even larger than the hit rate of the LBC pathway. This is likely due to the tropical cyclogenesis biases in the GEFS mentioned earlier (Li et al. 2016). However, the higher hit rate over the EMDR and CWMDR does not occur at the expense of a larger false alarm ratio. The four regions have a similar false alarm ratio beyond 3 days (Fig. 7c), or the proportion of forecast events that fail to materialize is about the same in the different regions. The critical success index shows a sequence similar to the hit rate, but the EMDR and CWMDR are less well separated (Fig. 7d).

If evaluated based on the hit rate, tropical cyclogenesis over the EMDR is most predictable, followed by the CWMDR, and tropical cyclogenesis over the subtropical Atlantic region is least predictable. The differences in the hit rate between different regions can be explained by the different contributions of the pathways (Fig. 8a). Over the EMDR, the LBC pathway accounts for about 70% of tropical cyclogenesis; over the CWMDR, about 70% of TCs develop via the NBC pathway. In contrast, weak TT makes the largest contribution over the Gulf (~55%), and strong TT and weak TT each contribute more than 40% over the subtropical Atlantic. If the hit rate differences between the regions result from the differences in pathway frequency, then one could make an estimate of the hit rate for each region based only on the knowledge of the pathway frequency (Fig. 8a) and the per-pathway hit rate (Fig. 3): no knowledge of the region itself is required. The estimated hit rate (in solid lines) and the actual hit rate for each region (in dashed lines) are shown in Fig. 8c. To the extent that these curves are similar (including ordering), the relative frequency of the pathways explains regional differences in predictability. In particular, the relatively low predictive skill of tropical cyclogenesis over the subtropical Atlantic can be attributed to strong extratropical influences.

The seasonal variations of the pathway occurrences help to explain the differences in genesis predictive skill or predictability between the peak season and the early/late seasons. Because the tropical transition pathways make a larger relative contribution in June–July (JJ) and October–November (ON) than in August–September (AS; Fig. 8b), the hit rate and critical success index are higher in the peak season (AS) than in the early season (JJ) or late season (ON) (Figs. 9a,d). The hit rates during different seasonal periods are also estimated based on the relative frequency occurrence of each pathway (Fig. 8b) and the per-pathway hit rate (Fig. 3). The estimations are very close to the actual hit rates. This suggests that the seasonal difference in tropical cyclogenesis predictability can be explained by the predictability differences among pathways. It is also worth noting that the false alarm ratio in AS is close to that in

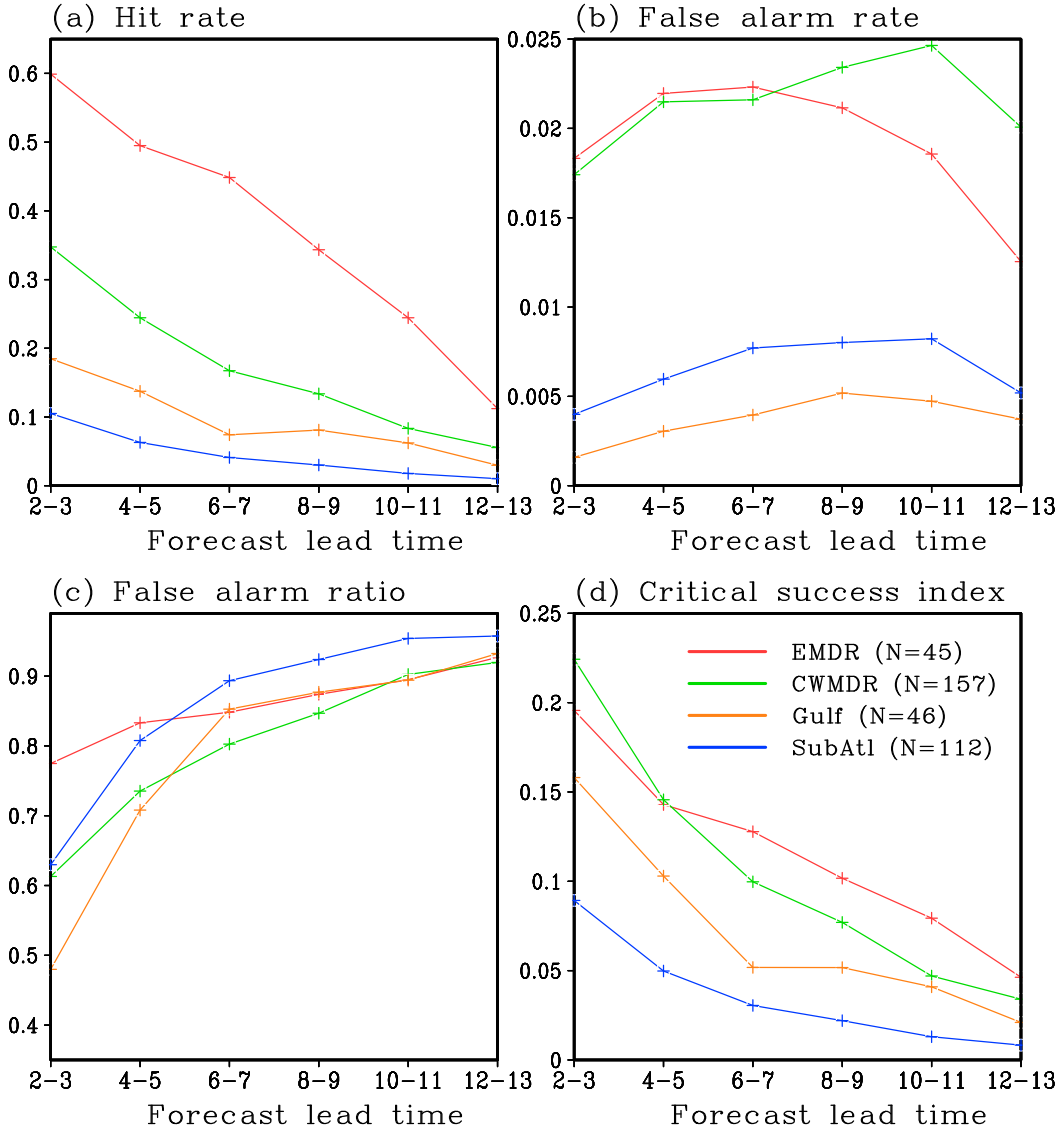

FIG. 7. (a) Hit rate, (b) false alarm rate, (c) false alarm ratio, and (d) critical success index as a function of the forecast lead time for different subregions of the Atlantic: the EMDR, the CWMDR, Gulf, and SubAtl.

ON and smaller than that in JJ (Fig. 9c) even though the GEFS produces a larger number of false alarm storms in AS than in JJ or ON (Fig. 9b),

d. Illustrative examples

An example of each genesis pathway is selected from the hurricane season of 2010 and shown in Fig. 10. Superimposed on the observed track and predicted genesis locations in each plot is an inset showing the ensemble hit rates at different forecast lead times. The ensemble hit rate is defined as the ratio of the number of the ensemble forecasts that produce a hit to the total ensemble size (i.e., 11). The ensemble hit rate is then averaged over different forecast lead periods (i.e., days 2–5, 6–9, and 10–13).

Hurricane Igor (Fig. 10a) is a low-level baroclinic case. It originated from an African easterly wave and formed near the coast of West Africa on 8 September 2010 (Pasch and Kimberlain 2011). It was not a "merger developer" according to the definition in Hankes et al. (2015), but the formation of the storm resulted from the interaction between the primary African easterly wave south of the jet and vorticity disturbances north of the jet (not shown). The ensemble hit rate of the storm is ~0.6 for days 2–5 and ~0.5 for days 6–9 and drops to ~0.2 for days 10–13. The predicted genesis locations are close to

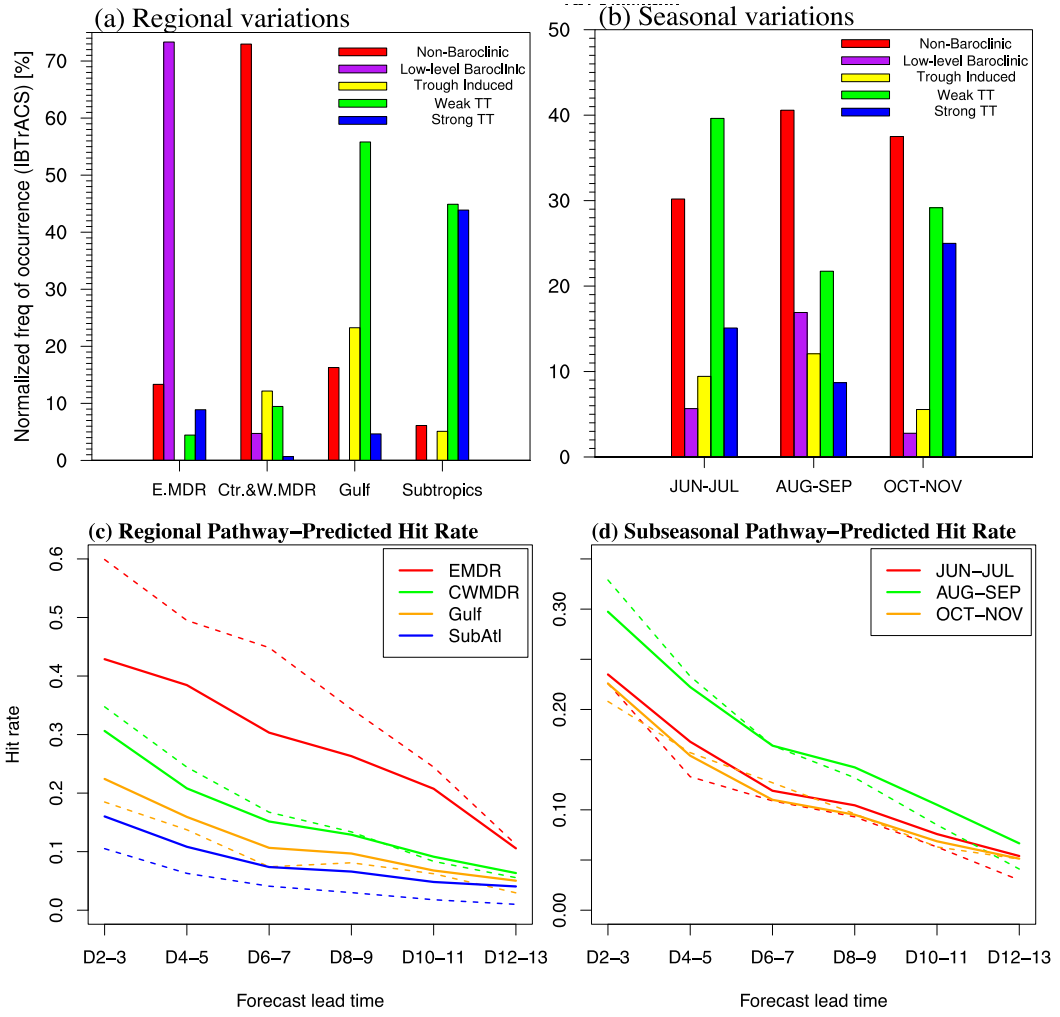

FIG. 8. (a) The relative contributions of different genesis pathways in different regions of the Atlantic. The relative contribution of a pathway in a region is defined as the number of storms for the pathway in the region normalized by the total number of storms in the region. (b) As in (a), but for different seasonal periods over the North Atlantic basin. (c),(d) The solid lines show the regional and subseasonal hit rates estimated based on the mean hit rate of each pathway and their relative frequencies of occurrence in a region or during a subseasonal period (see text for more details). To facilitate comparison, the actual hit rates from Figs. 7a and 9a are shown in (c) and (d), respectively, in dashed lines. Note that the TI pathway does not take place over the EMDR and the low-level baroclinic pathway is absent over Gulf and SubAtl.

the observed genesis location with a couple of exceptions. Hurricane Danielle (Fig. 10b) is a nonbaroclinic case, and genesis occurred on 21 August 2010. Its development involved interaction between an easterly wave and an ITCZ disturbance (Kimberlain 2010). The ensemble hit rate is slightly lower than that of Hurricane Igor during days 2–5 and days 6–9 but slightly higher during days 10–13.

Tropical Storm Matthew (Fig. 10c), falling into the trough-induced category, formed over the Caribbean Sea on 23 September 2010. Although the storm developed in the presence of an upper-level trough (not shown), a low-level tropical easterly wave likely played a dominant role in the development of Matthew, and the storm formed near the center of the wave pouch (Brennan 2010; Montgomery et al. 2012). The ensemble hit rate from the GEFS-R2 is about 0.7 for days 2–5, and it drops quickly with the forecast lead time. The predicted genesis locations have a large spread, indicating a high fraction of early genesis or late genesis predictions (i.e., incorrect genesis time). The ensemble hit rate would be lower if the radius threshold (5°) or the time window threshold (±120 h) were reduced in the hit rate calculation.

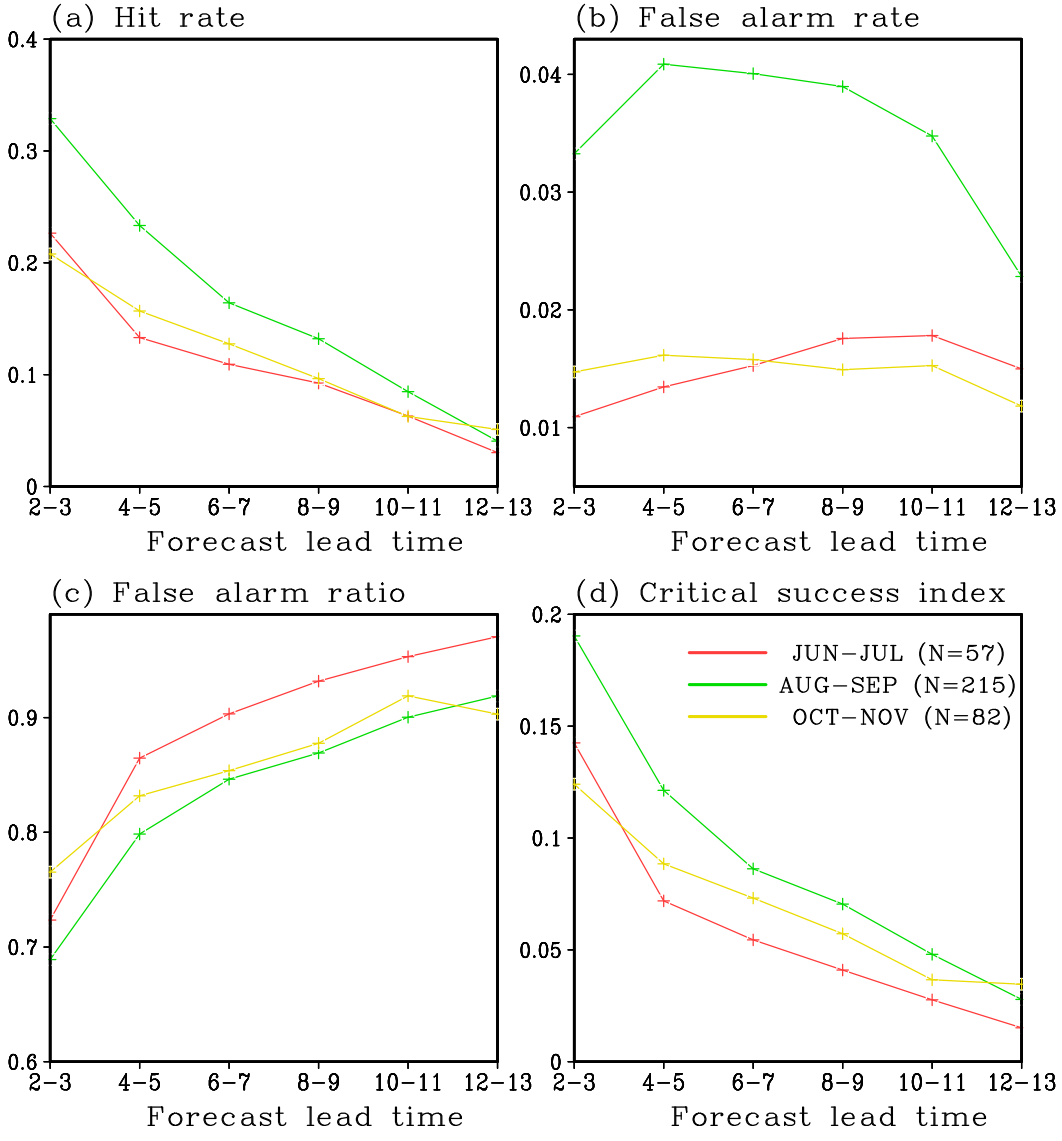

FIG. 9. As in Fig. 7, but for different subseasonal time periods over the whole North Atlantic basin.

Tropical Storm Bonnie (Fig. 10d) and Hurricane Shary (Fig. 10e) are weak TT and strong TT storms, respectively. Bonnie developed on 22 July 2010, from the interaction between an African easterly wave and an upper-level low to the north of Hispaniola in late July (Stewart 2010). Shary formed on 28 October 2010, along the southern portion of a stationary frontal system under the influence of an upper-level low associated with the midocean trough (Avila 2011). While the ensemble hit rate is low for Bonnie, the predicted genesis locations are very close to that observed. As for Shary, only one ensemble member predicts the formation of a TC during days 2–5, and the ensemble hit rate is nearly zero, indicating a serious forecast challenge.

5. Summary and discussion

The practical predictability of tropical cyclogenesis over the North Atlantic was examined in different synoptic flow regimes using the NCEP GEFS reforecasts. Flow regimes were identified objectively by five tropical cyclogenesis pathways that are categorized based on the upper-level forcing and low-level baroclinicity of the environmental atmospheric state (McTaggart-Cowan et al. 2013). Among them, the NBC and LBC pathways have negligible upper-level forcing and can be regarded as purely tropical pathways; trough-induced, weak TT, and strong TT pathways are associated with strong upper-level forcing and have weak, moderate, and

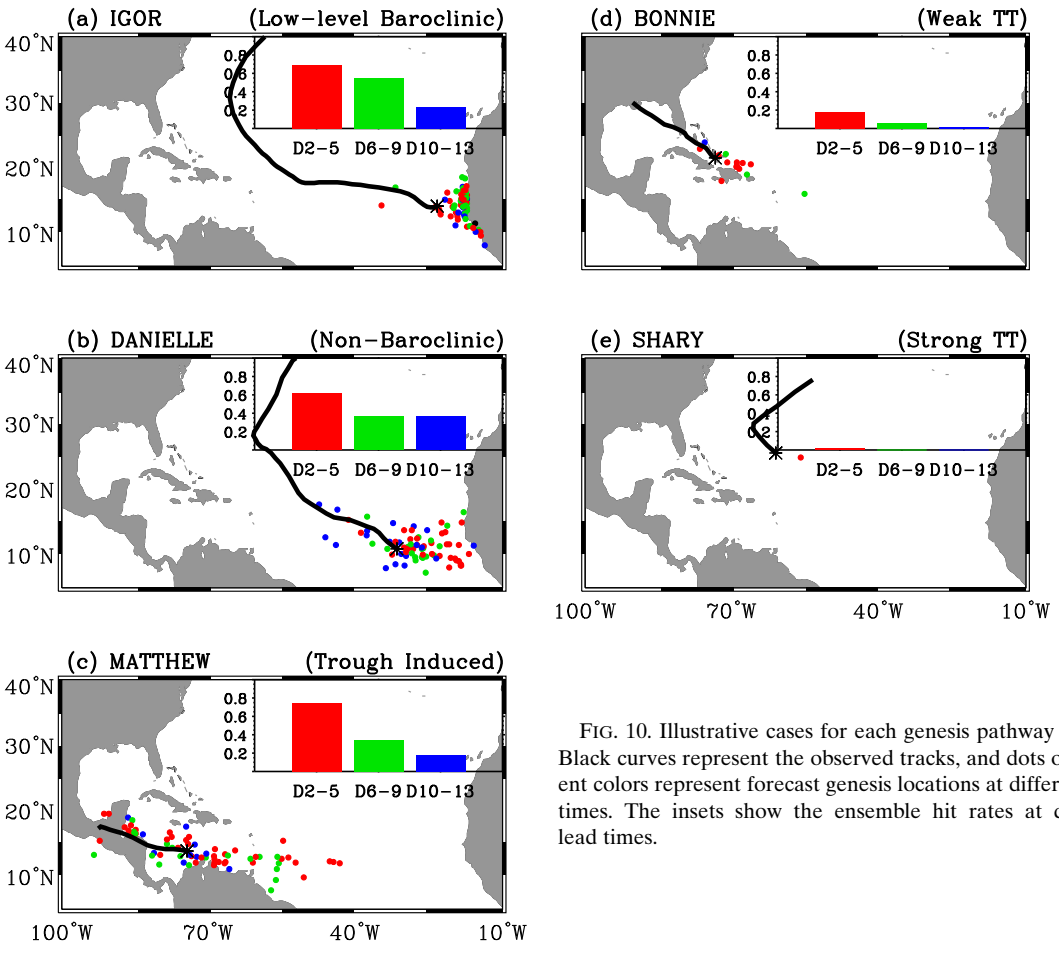

FIG. 10. Illustrative cases for each genesis pathway in 2010. Black curves represent the observed tracks, and dots of different colors represent forecast genesis locations at different lead times. The insets show the ensemble hit rates at different lead times.

strong low-level baroclinicity, respectively, and can be regarded as hybrid or extratropical pathways. Since the conventional ensemble spread metric cannot be readily applied to the dichotomy between genesis and nongenesis, the dependence of tropical cyclogenesis predictability on flow regimes was estimated by comparing the hit rate for different genesis pathways in the way that a lower hit rate represents lower predictability. The hit rate of the pathways has the following order: LBC > trough induced > NBC > weak TT > strong TT. Further analysis showed that the RMSEs of vertical wind shear and 700-hPa relative humidity help to explain the different predictive skills for different pathways.

The low predictability of tropical transition pathways is consistent with the general perception that the extratropical atmosphere is less predictable than the tropical atmosphere for forecast lead time beyond a few days. We further examined whether stronger extratropical influences imply lower genesis predictability using the Q metric (i.e., upper-level forcing) as a proxy for extratropical influences. Although the strong TT, weak TT, and trough-induced pathways are categorized based on the low-level baroclinicity, they are also associated with upper-level forcing of different strengths: strong upper-level forcing occurs most frequently in the strong TT pathway and least frequently in the trough-induced pathway. Once the genesis events classified into these pathways are stratified by the Q metric, it shows that the strong-Q group always has a lower hit rate than the corresponding weak-Q group, suggesting that stronger extratropical influences may lead to lower genesis predictability.

Previous studies have reported that the predictive skill of a model in tropical cyclogenesis varies from region to region and from month to month (see the introduction). Large spatial variations of the hit rate were also found in this study: the MDR region has a higher hit rate than the Gulf of Mexico or the subtropical Atlantic (the false alarm ratio is similar and very large in all the regions). These spatial variations can be largely explained by the different relative contributions of the pathways in different regions. The relatively frequent occurrence of TT

pathways over the Gulf of Mexico and the subtropical Atlantic contributes to the lower genesis predictability in these regions. In addition, the purely tropical pathways (NBC and LBC) occur more frequently in August and September, and this contributes to a relatively high hit rate in the peak hurricane season compared to the early and late seasons.

The impact of genesis pathways on predictability may be understood through a thought experiment using a genesis potential index (GPI). GPI is typically expressed as a set of multiplicative factors based on physically relevant predictors of storm formation (e.g., Emanuel and Nolan 2004; Camargo et al. 2007). Denoting these factors by F_k, where k is one of the N predictors included in the GPI, we propose that the genesis predictability, $P(G)$, is directly related to the joint predictability of the individual factors,

$$P(G) = \prod_{k=1,N} P(F_k). \tag{6}$$

Because of the prevalence of NBC and LBC pathways in the MDR, it is reasonable to expect that the current GPI formulations are most directly applicable to these development types. The Emanuel and Nolan (2004) GPI has $N = 4$, but the number and form of optimal predictors may change for different pathways. In particular, the GPI index mainly focuses on the lower- to mid-tropospheric conditions, such as the 850-hPa vorticity and 700-hPa RH, and the upper-tropospheric dynamical conditions are not taken into account other than the 200–850-hPa vertical shear. Tropical cyclone formation following strong TT and weak TT pathways, with their reliance on upper-level extratropical precursors in a baroclinic environment, might reasonably be assumed to be sensitive to additional factors such that $N > 4$. Because $P(G)$ decreases as N increases unless the predictability of the additional factors is unity, genesis pathways that rely on more ingredients are expected to have lower predictability.

Additionally, a change in $P(G)$ can be realized through changes in the marginal predictability, $P(F_k)$, for one or more of the factors. Tropical cyclogenesis is involved with interaction between processes of different spatiotemporal scales. The establishment of some favorable conditions, such as enhanced low-level vorticity and elevated midlevel humidity, results from the positive feedback between the parent wave and the enclosed moist convection in a typical NBC storm (Dunkerton et al. 2009), and the stochastic nature of convection limits the predictability of this process. In the case of the LBC pathway, the interaction of northern and southern waves leads to a stronger and deeper wave pouch conducive to tropical cyclogenesis (Hankes et al. 2015). The tropical, synoptic nature of this interaction and its strong geographic preference likely enhance the predictability of genesis. In contrast, the development of a TT storm is dependent on upper-level disturbances of extratropical origin. Small forecast errors tend to grow faster outside the tropics than within the tropics for forecast lead time beyond a few days. In addition, the apparent requirement of near saturation over a deep layer is more difficult to achieve in strong vertical shear (it is only through the finite amplitude structure of the pre-TC disturbance that this can occur and only for a limited time given the evolving structure characteristic of a baroclinic wave). The resultant low predictability of vertical shear and humidity contributes to the poorer prediction of tropical transition (Fig. 4). An investigation of the relative contributions of this reduced factor predictability [decreased $P(F_k)$] and an increased number of factors (increased N) to the overall genesis predictability for the TT pathways may yield additional insight for the development of a pathway-conditional GPI, which may provide improved genesis guidance with a qualified level of predictive uncertainties.

This study aims to provide a systematic investigation of tropical cyclogenesis predictability in different synoptic flow regimes. However, there are some admitted limitations. First, the dependence of predictability on flow regimes was estimated based on the hit rate. A single metric may not provide a complete picture of how predictable TCs are. For example, the false alarm ratio is likely high for all the five pathways, and the false alarm rate may be highest for the LBC pathway because of a large number of false alarm storms over the eastern MDR in the GEFS-R2. Second, this study was based on only one reforecast dataset. It is possible that some results are model dependent. For example, the GEFS has a negative bias in tropical cyclogenesis frequency over the central and west Atlantic because of a systematic bias of the model (Li et al. 2016), which may lead to the lower predictive skill of the NBC pathway than the trough-induced pathway. On the other hand, given the generally lower predictability of extratropical atmosphere than tropical atmosphere at the synoptic time scale and beyond and the impacts of vertical shear on predictability, it is likely a robust finding that the tropical transition pathways have lower intrinsic predictability than the other pathways. This suggests that storms developing near the southeast coast of North America, which are more likely subject to strong extratropical influences, may pose a particular challenge for operational forecasts and emergency management. Knowledge of the predictability of such coastal storms can assist decision-making by providing useful estimates

of predictive skill before forecast validation becomes available.

Acknowledgments. This work is supported by NOAA Grant NA16OAR4310080, NOAA Grant NA15NWS4680007, and NRL Grant N00173-15-1-G004. X.J. acknowledges support by NOAA MAPP Program under Award NA12OAR4310075. The authors are grateful to Dr. Mike Montgomery and two anonymous reviewers for constructive comments on a first version of this manuscript. The GEFS reforecasts are available online (http://esrl.noaa.gov/psd/forecasts/reforecast2).

REFERENCES

American Meteorological Society, 2012: "Predictability." Glossary of Meteorology, http://glossary.ametsoc.org/wiki/Predictability.

Avila, L. A., 2011: Tropical cyclone report: Hurricane Shary (AL202010) 28-30 October 2010. National Hurricane Center Rep., 10 pp.

Barnston, A. G., N. Vigaud, L. N. Long, M. K. Tippett, and J.-K. E. Schemm, 2015: Atlantic tropical cyclone activity in response to the MJO in NOAA's CFS model. *Mon. Wea. Rev.*, **143**, 4905–4927, https://doi.org/10.1175/MWR-D-15-0127.1.

Bentley, A., L. Bosart, and D. Keyser, 2017: Upper-tropospheric precursors to the formation of subtropical cyclones that undergo tropical transition in the North Atlantic basin. *Mon. Wea. Rev.*, **145**, 503–520, https://doi.org/10.1175/MWR-D-16-0263.1.

Brennan, M. J., 2010: Tropical cyclone report: Tropical Storm Matthew, 23-26 September 2010. National Hurricane Center Rep., 15 pp.

Camargo, S. J., K. A. Emanuel, and A. H. Sobel, 2007: 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.

Cangialosi, J. P., 2014: Tropical cyclone report: Tropical Storm Hanna, 22-28 October 2014. National Hurricane Center Rep., 14 pp.

Charney, J. G., and J. Shukla, 1981, Predictability of monsoons. *Monsoon Dynamics*, J. Lighthill and R. Pearce, Eds., Cambridge University Press, 735 pp.

Chen, J.-H., and S.-J. Lin, 2013: Seasonal predictions of tropical cyclones using a 25-km-resolution general circulation model. *J. Climate*, **26**, 380–398, https://doi.org/10.1175/JCLI-D-12-00061.1.

Davis, C. A., and L. F. Bosart, 2003: Baroclinically induced tropical cyclogenesis. *Mon. Wea. Rev.*, **131**, 2730–2747, https://doi.org/10.1175/1520-0493(2003)131<2730:BITC>2.0.CO;2.

——, and ——, 2004: The TT problem: Forecasting the tropical transition of cyclones. *Bull. Amer. Meteor. Soc.*, **85**, 1657–1662, https://doi.org/10.1175/BAMS-85-11-1657.

——, D. A. Ahijevych, W. Wang, and W. C. Skamarock, 2016: Evaluating medium-range tropical cyclone forecasts in uniform- and variable-resolution global models. *Mon. Wea. Rev.*, **144**, 4141–4160, https://doi.org/10.1175/MWR-D-16-0021.1.

Dunkerton, T. J., M. T. Montgomery, and Z. Wang, 2009: Tropical cyclogenesis in a tropical wave critical layer: Easterly waves. *Atmos. Chem. Phys.*, **9**, 5587–5646, https://doi.org/10.5194/acp-9-5587-2009.

Elsberry, R. L., W. M. Clune, and P. A. Harr, 2009: Evaluation of global model early track and formation prediction in the western North Pacific. *Asia-Pac. J. Atmos. Sci.*, **45**, 357–374.

——, M. S. Jordan, and F. Vitart, 2011: Evaluation of the ECMWF 32-day ensemble predictions during 2009 season of 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.

Emanuel, K. A., and D. S. Nolan, 2004: Tropical cyclone activity and global climate. *26th Conf. on Hurricanes and Tropical Meteorology,* Miami, FL, Amer. Meteor. Soc., 10A.2, https://ams.confex.com/ams/26HURR/techprogram/paper_75463.htm.

Galarneau, T. J., R. McTaggart-Cowan, L. F. Bosart, and C. A. Davis, 2015: Development of North Atlantic tropical disturbances near upper-level potential vorticity streamers. *J. Atmos. Sci.*, **72**, 572–597, https://doi.org/10.1175/JAS-D-14-0106.1.

Goldenberg, S. B., and L. J. Shapiro, 1996: Physical mechanisms for the association of El Niño and West Africa rainfall with Atlantic major hurricanes. *J. Climate*, **9**, 1169–1187, https://doi.org/10.1175/1520-0442(1996)009<1169:PMFTAO>2.0.CO;2.

——, C. W. Landsea, A. M. Mestas-Nuñez, and W. M. Gray, 2001: The recent increase in Atlantic hurricane activity: Causes and implications. *Science*, **293**, 474–479, https://doi.org/10.1126/science.1060040.

Gray, W. M., 1968: Global view of the origin of tropical disturbances and storms. *Mon. Wea. Rev.*, **96**, 669–700, https://doi.org/10.1175/1520-0493(1968)096<0669:GVOTOO>2.0.CO;2.

——, 1998: The formation of tropical cyclones. *Meteor. Atmos. Phys.*, **67**, 37–69, https://doi.org/10.1007/BF01277501.

Grimit, E. P., and C. F. Mass, 2007: Measuring the ensemble spread–error relationship with a probabilistic approach: Stochastic ensemble results. *Mon. Wea. Rev.*, **135**, 203–221, https://doi.org/10.1175/MWR3262.1.

Guinn, T. A., and W. H. Schubert, 1993: Hurricane spiral bands. *J. Atmos. Sci.*, **50**, 3380–3403, https://doi.org/10.1175/1520-0469(1993)050<3380:HSB>2.0.CO;2.

Halperin, D. J., H. E. Fuelberg, R. E. Hart, J. H. Cossuth, P. Sura, and R. J. Pasch, 2013: An evaluation of tropical cyclone genesis forecasts from global numerical models. *Wea. Forecasting*, **28**, 1423–1445, https://doi.org/10.1175/WAF-D-13-00008.1.

——, ——, ——, and ——, 2016: Verification of tropical cyclone genesis forecasts from global numerical models: Comparisons between the North Atlantic and eastern North Pacific basins. *Wea. Forecasting*, **31**, 947–955, https://doi.org/10.1175/WAF-D-15-0157.1.

——, R. E. Hart, H. E. Fuelberg, and J. H. Cossuth, 2017: The development and evaluation of a statistical–dynamical tropical cyclone genesis guidance tool. *Wea. Forecasting*, **32**, 27–46, https://doi.org/10.1175/WAF-D-16-0072.1.

Hamill, T. M., G. T. Bates, J. S. Whitaker, D. R. Murray, M. Fiorino, T. J. Galarneau Jr., Y. Zhu, and W. Lapenta, 2013: NOAA's second-generation global medium-range ensemble reforecast dataset. *Bull. Amer. Meteor. Soc.*, **94**, 1553–1565, https://doi.org/10.1175/BAMS-D-12-00014.1.

Hankes, I., Z. Wang, G. Zhang, and C. L. Fritz, 2015: Merger of African easterly waves and formation of Cape Verde storms. *Quart. J. Roy. Meteor. Soc.*, **141**, 1306–1319, https://doi.org/10.1002/qj.2439.

Kimberlain, T. B., 2010: Tropical cyclone report: Hurricane Danielle, 21-30 August 2010. National Hurricane Center Rep., 16 pp.

——, 2014: Tropical cyclone report: Hurricane Fay, 10-13 October 2014. National Hurricane Center Rep., 21 pp.

Knapp, K. R., M. C. Kruk, D. H. Levinson, H. J. Diamond, and C. J. Neumann, 2010: The International Best Track Archive for Climate Stewardship (IBTrACS) unifying tropical cyclone data. *Bull. Amer. Meteor. Soc.*, **91**, 363–376, https://doi.org/10.1175/2009BAMS2755.1.

Komaromi, W. A., and S. J. Majumdar, 2014: Ensemble-based error and predictability metrics associated with tropical cyclogenesis. Part I: Basinwide perspective. *Mon. Wea. Rev.*, **142**, 2879–2898, https://doi.org/10.1175/MWR-D-13-00370.1.

——, and ——, 2015: Ensemble-based error and predictability metrics associated with tropical cyclogenesis. Part II: Wave-relative framework. *Mon. Wea. Rev.*, **143**, 1665–1686, https://doi.org/10.1175/MWR-D-14-00286.1.

Kossin, J. P., and D. J. Vimont, 2007: A more general framework for understanding Atlantic hurricane variability and trends. *Bull. Amer. Meteor. Soc.*, **88**, 1767–1781, https://doi.org/10.1175/BAMS-88-11-1767.

Landsea, C., 1993: A climatology of intense (or major) Atlantic hurricanes. *Mon. Wea. Rev.*, **121**, 1703–1713, https://doi.org/10.1175/1520-0493(1993)121<1703:ACOIMA>2.0.CO;2.

Leroy, A., and M. C. Wheeler, 2008: Statistical prediction of weekly tropical cyclone activity in the Southern Hemisphere. *Mon. Wea. Rev.*, **136**, 3637–3654, https://doi.org/10.1175/2008MWR2426.1.

Li, T., and B. Fu, 2006: Tropical cyclogenesis associated with Rossby wave energy dispersion of a preexisting typhoon. Part I: Satellite data analyses. *J. Atmos. Sci.*, **63**, 1377–1389, https://doi.org/10.1175/JAS3692.1.

Li, W., Z. Wang, and M. Peng, 2016: Evaluating tropical cyclone forecasts in the NCEP Global Ensemble Forecasting System (GEFS) Reforecast version 2. *Wea. Forecasting*, **31**, 895–916, https://doi.org/10.1175/WAF-D-15-0176.1.

Lorenz, E. N., 1965: A new study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**, 321–333, https://doi.org/10.3402/tellusa.v17i3.9076.

——, 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289–307, https://doi.org/10.3402/tellusa.v21i3.10086.

Lussier, L. L., III, 2010: A multi-scale analysis of tropical cyclogenesis within the critical layer of tropical easterly waves in the Atlantic and western North Pacific sectors. Ph.D. dissertation, Naval Postgraduate School, 217 pp.

Métais, O., J. J. Riley, and M. Lesieur, 1994, Numerical simulations of stably-stratified rotating turbulence. *Stably Stratified Flows: Flow and Dispersion over Topography*, I. P. Castro and N. J. Rockliff, Eds., Clarendon Press, 139–151.

Madden, R. A., and P. R. Julian, 1972: Description of global circulation cells in the tropics with a 40–50-day period. *J. Atmos. Sci.*, **29**, 1109–1123, https://doi.org/10.1175/1520-0469(1972)029<1109:DOGSCC>2.0.CO;2.

Marchok, T. P., 2002: How the NCEP tropical cyclone tracker works. Preprints, *25th Conf. on Hurricanes and Tropical Meteorology*, San Diego, CA, Amer. Meteor. Soc., P1.13, https://ams.confex.com/ams/25HURR/techprogram/paper_37628.htm.

McMurdie, L. A., and B. Ancell, 2014: Predictability characteristics of landfalling cyclones along the North American west coast. *Mon. Wea. Rev.*, **142**, 301–319, https://doi.org/10.1175/MWR-D-13-00141.1.

McTaggart-Cowan, R., G. D. Deane, L. F. Bosart, C. A. Davis, and T. J. Galarneau, 2008: Climatology of tropical cyclogenesis in the North Atlantic (1948–2004). *Mon. Wea. Rev.*, **136**, 1284–1304, https://doi.org/10.1175/2007MWR2245.1.

——, T. J. Galarneau Jr., L. F. Bosart, R. W. Moore, and O. Martius, 2013: A global climatology of baroclinically influenced tropical cyclogenesis. *Mon. Wea. Rev.*, **141**, 1963–1989, https://doi.org/10.1175/MWR-D-12-00186.1.

Montgomery, M. T., L. L. Lussier III, R. W. Moore, and Z. Wang, 2010: The genesis of Typhoon Nuri as observed during the Tropical Cyclone Structure 2008 (TCS-08) field experiment—Part 1: The role of the easterly wave critical layer. *Atmos. Chem. Phys.*, **10**, 9879–9900, https://doi.org/10.5194/acp-10-9879-2010.

——, and Coauthors, 2012: The Pre-Depression Investigation of Cloud-systems in the Tropics (PREDICT) Experiment: Scientific basis, new analysis tools, and some first results. *Bull. Amer. Meteor. Soc.*, **93**, 153–172, https://doi.org/10.1175/BAMS-D-11-00046.1.

National Research Council, 2010: Assessment of intraseasonal to interannual climate prediction and predictability. National Research Council Rep., 192 pp., https://doi.org/10.17226/12878.

Palmer, T. N., 1996: Predictability of the atmosphere and oceans: From days to decades. *Decadal Climate Variability: Dynamics and Predictability*, D. L. T. Anderson and J. Willebrand, Eds., NATO ASI Series, Vol. 44, Springer, 83–155, https://doi.org/10.1007/978-3-662-03291-6_3.

——, 2006: Predictability of weather and climate: From theory to practice. *Predictability of Weather and Climate*. T. Palmer and R. Hagedorn, Eds., Cambridge University Press, 1–29.

Pasch, R. J., and T. B. Kimberlain, 2011: Tropical cyclone report: Hurricane Igor, 8-21 September 2010. National Hurricane Center Rep., 20 pp.

Pytharoulis, I., and C. Thorncroft, 1999: The low-level structure of African easterly waves in 1995. *Mon. Wea. Rev.*, **127**, 2266–2280, https://doi.org/10.1175/1520-0493(1999)127<2266:TLLSOA>2.0.CO;2.

Reed, R. J., D. C. Norquist, and E. E. Recker, 1977: The structure and properties of African wave disturbances as observed during phase III of GATE. *Mon. Wea. Rev.*, **105**, 317–333, https://doi.org/10.1175/1520-0493(1977)105<0317:TSAPOA>2.0.CO;2.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, https://doi.org/10.1175/2010BAMS3001.1.

Slade, S. A., and E. D. Maloney, 2013: An intraseasonal prediction model of Atlantic and east Pacific tropical cyclone genesis. *Mon. Wea. Rev.*, **141**, 1925–1942, https://doi.org/10.1175/MWR-D-12-00268.1.

Smagorinsky, J., 1969: Problems and promises of deterministic extended range forecasting. *Bull. Amer. Meteor. Soc.*, **50**, 286–311.

Stewart, R. S., 2010: Tropical cyclone report: Tropical Storm Bonnie, 22-24 July 2010. National Hurricane Center Rep., 17 pp.

Straus, D. M., and D. Paolino, 2009: Intermediate time error growth and predictability: Tropics versus midlatitudes. *Tellus*, **61A**, 579–586, https://doi.org/10.1111/j.1600-0870.2009.00411.x.

Tsai, H.-C., and R. L. Elsberry, 2013: Opportunities and challenges for extended-range predictions of tropical cyclone impacts on hydrological predictions. *J. Hydrol.*, **506**, 42–54, https://doi.org/10.1016/j.jhydrol.2012.12.025.

Walsh, K., M. Fiorino, C. Landsea, and K. McInnes, 2007: Objectively determined resolution-dependent threshold criteria for the detection of tropical cyclones in climate models and reanalyses. *J. Climate*, **20**, 2307–2314, https://doi.org/10.1175/JCLI4074.1.

Wang, C.-C., and G. Magnusdottir, 2005: ITCZ breakdown in three-dimensional flows. *J. Atmos. Sci.*, **62**, 1497–1512, https://doi.org/10.1175/JAS3409.1.

Wang, Z., M. T. Montgomery, and T. Dunkerton, 2010a: Genesis of pre–Hurricane Felix (2007). Part I: The role of the wave critical layer. *J. Atmos. Sci.*, **67**, 1711–1729, https://doi.org/10.1175/2009JAS3420.1.

——, ——, and ——, 2010b: Genesis of pre–Hurricane Felix (2007). Part II: Warm core formation, precipitation evolution, and predictability. *J. Atmos. Sci.*, **67**, 1730–1744, https://doi.org/10.1175/2010JAS3435.1.

——, ——, and C. Fritz, 2012: A first look at the structure of the wave pouch during the 2009 PREDICT–GRIP dry runs over the Atlantic. *Mon. Wea. Rev.*, **140**, 1144–1163, https://doi.org/10.1175/MWR-D-10-05063.1.

——, G. Zhang, M. S. Peng, J.-H. Chen, and S.-J. Lin, 2015: Predictability of Atlantic tropical cyclones in the GFDL HiRAM model. *Geophys. Res. Lett.*, **42**, 2547–2554, https://doi.org/10.1002/2015GL063587.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences.* 2nd ed. Academic Press, 627 pp.

Wu, L. G., H. J. Zong, and L. Jia, 2013: Observational analysis of tropical cyclone formation associated with monsoon gyres. *J. Atmos. Sci.*, **70**, 1023–1034, https://doi.org/10.1175/JAS-D-12-0117.1.

Xiang, B., and Coauthors, 2015a: 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.

——, M. Zhao, X. Jiang, S.-J. Lin, T. Li, X. Fu, and G. Vecchi, 2015b: The 3–4-week MJO predictive skill in a GFDL coupled model. *J. Climate*, **28**, 5351–5364, https://doi.org/10.1175/JCLI-D-15-0102.1.

Zhang, C., 2005: Madden–Julian oscillation. *Rev. Geophys.*, **43**, RG2003, https://doi.org/10.1029/2004RG000158.

Zhang, F., and D. Tao, 2013: Effects of vertical wind shear on the predictability of tropical cyclones. *J. Atmos. Sci.*, **70**, 975–983, https://doi.org/10.1175/JAS-D-12-0133.1.

Zhang, G., and Z. Wang, 2013: Interannual variability of the Atlantic Hadley circulation in boreal summer and its impacts on tropical cyclone activity. *J. Climate*, **26**, 8529–8544, https://doi.org/10.1175/JCLI-D-12-00802.1.

——, ——, M. Peng, and G. Magnusdottir, 2017: Characteristics and impacts of extratropical Rossby wave breaking during the Atlantic hurricane season. *J. Climate*, **30**, 2363–2379, https://doi.org/10.1175/JCLI-D-16-0425.1.