Statistical–Dynamical Seasonal Forecast of North Atlantic and U.S. Landfalling Tropical Cyclones Using the High-Resolution GFDL FLOR Coupled Model

HIROYUKI MURAKAMI

National Oceanic and Atmospheric Administration/Geophysical Fluid Dynamics Laboratory, and Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, New Jersey

GABRIELE VILLARINI

IIHR–Hydroscience and Engineering, The University of Iowa, Iowa City, Iowa

GABRIEL A. VECCHI AND WEI ZHANG

National Oceanic and Atmospheric Administration/Geophysical Fluid Dynamics Laboratory, and Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, New Jersey

RICHARD GUDGEL

National Oceanic and Atmospheric Administration/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

(Manuscript received 4 September 2015, in final form 15 February 2016)

ABSTRACT

Retrospective seasonal forecasts of North Atlantic tropical cyclone (TC) activity over the period 1980–2014 are conducted using a GFDL high-resolution coupled climate model [Forecast-Oriented Low Ocean Resolution (FLOR)]. The focus is on basin-total TC and U.S. landfall frequency. The correlations between observed and model predicted basin-total TC counts range from 0.4 to 0.6 depending on the month of the initial forecast. The correlation values for U.S. landfalling activity based on individual TCs tracked from the model are smaller and between 0.1 and 0.4. Given the limited skill from the model, statistical methods are used to complement the dynamical seasonal TC prediction from the FLOR model. Observed and predicted TC tracks were classified into four groups using fuzzy *c*-mean clustering to evaluate the model's predictability in observed classification of TC tracks. Analyses revealed that the FLOR model has the highest skill in predicting TC frequency for the cluster of TCs that tracks through the Caribbean and the Gulf of Mexico.

New hybrid models are developed to improve the prediction of observed basin-total TC and landfall TC frequencies. These models use large-scale climate predictors from the FLOR model as predictors for generalized linear models. The hybrid models show considerable improvements in the skill in predicting the basin-total TC frequencies relative to the dynamical model. The new hybrid model shows correlation coefficients as high as 0.75 for basinwide TC counts from the first two lead months and retains values around 0.50 even at the 6-month lead forecast. The hybrid model also shows comparable or higher skill in forecasting U.S. landfalling TCs relative to the dynamical predictions. The correlation coefficient is about 0.5 for the 2–5-month lead times.

1. Introduction

Tropical cyclones (TCs) were the most costly natural disaster to affect the United States over the period 1980–2011 (Pielke et al. 2008; Smith and Katz 2013). According to Smith and Katz (2013), TCs were responsible for over $400 billion in damages over the period based on the Consumer Price Index, which correspond to 47% of all the damage caused by all natural disasters responsible for more than $1 billion combined. Smith and Katz (2013) also reported apparent increasing trends in both the annual frequency of billion-dollar events and in the annual aggregated loss from these events. Therefore, predicting TC activity at

Corresponding author address: Hiroyuki Murakami, NOAA/GFDL, 201 Forrestal Rd., Princeton, NJ 08540-6649.
E-mail: hir.murakami@gmail.com

DOI: 10.1175/MWR-D-15-0308.1

seasonal time scales is a topic of great scientific and socioeconomic interest.

Since Gray (1984a,b) first attempted seasonal forecasting of TC activity for the North Atlantic, tremendous effort has been devoted to construct and improve statistical models in which observed large-scale climate indices ahead of the hurricane season are used to predict subsequent basin-total TC frequency (Gray et al. 1992, 1993, 1994; Klotzbach and Gray 2004, 2009; Elsner and Jagger 2006; Klotzbach 2008) and landfalling TCs (Lehmiller et al. 1997; Klotzbach and Gray 2003, 2004; Saunders and Lea 2005; Elsner et al. 2006; Klotzbach 2008; Jagger and Elsner 2010). The most widely used model for the statistical seasonal forecasts is based on Poisson regression, which was first used by Elsner and Schmertmann (1993). Most of the current statistical seasonal forecasts show skill for forecasts starting from April and later for the TC season in June–November (e.g., Elsner and Jagger 2006), and prediction skill is limited when the lead time increases and the target region is smaller than the entire North Atlantic (Lehmiller et al. 1997; Klotzbach and Gray 2012).

Recent advances in dynamical modeling and computational resources have enabled prediction using high-resolution dynamical models [see review in Camargo et al. (2007)]. These models showed significant skill in predicting seasonal total basin-total TCs (e.g., Vitart and Stockdale 2001; Vitart 2006; Vitart et al. 2007; LaRow et al. 2008; Camargo and Barnston 2009; LaRow et al. 2010; Zhao et al. 2010; Alessandri et al. 2011; Chen and Lin 2011, 2013; Vecchi et al. 2014; Camp et al. 2015), with correlation values up to 0.96 between observed and predicted North Atlantic TC counts over the 2000–10 period (Chen and Lin 2011, 2013). However, predicting U.S. landfall frequency using dynamical models remains challenging even though it is of paramount societal and scientific importance (Vecchi and Villarini 2014). Vecchi et al. (2014) and Camp et al. (2015) found some predictive skill for TC landfall in the Caribbean, but limited skill for U.S. landfall frequency.

Some of the limitations of dynamically forecasting TCs can be alleviated using so-called ''hybrid predictions'' or ''statistical–dynamical predictions.'' In the hybrid predictions, a statistical model is constructed using the empirical relationship between observed TC activity and predicted large-scale parameters simulated by a dynamical model. Using the statistical model, future TC activity is then predicted given the large-scale parameters predicted by a dynamical model (e.g., Zhao et al. 2010; Wang et al. 2009; Vecchi et al. 2011, 2013, 2014). For example, previous studies showed that basin-total North Atlantic TC activity substantially correlated with relative sea surface temperature (SST) anomalies (i.e., local SST anomaly relative to tropical mean anomaly) in observations (e.g., Latif et al. 2007; Swanson 2008; Vecchi et al. 2008; Villarini et al. 2010, Villarini and Vecchi 2012) and dynamical models (Zhao et al. 2010; Villarini et al. 2011; Murakami et al. 2012; Knutson et al. 2013; Ramsay and Sobel 2011). Using these simulated–predicted SST anomalies as predictors, previous studies achieved substantial skill in predicting basin-total TC frequency (Zhao et al. 2010; Vecchi et al. 2011, 2013, 2014), basin-total power dissipation index (PDI), and accumulated cyclone energy (ACE; Villarini and Vecchi 2013) compared to dynamical models. However, while these studies showed that it is possible to skillfully forecast North Atlantic TC activity at the basin scale, little is known about the applicability of hybrid systems at regional scales.

Vecchi et al. (2014) reported that the high-resolution dynamical model that will be used in this study has higher skill in predicting TCs near the coastline of the Gulf of Mexico and Caribbean Sea relative to those near the coastline of the northeastern United States (see their Fig. 13). This indicates that models may have higher skill in predicting–simulating one or more groups of TC tracks. If any hybrid models could improve predictions for the groups with poor forecasting skill, we could improve the prediction skill for landfall TC frequency as well as basin-total TC frequency. Moreover, finding predictors in the way of constructing a hybrid model will help the understanding of the potential physical mechanisms responsible for U.S. landfalling TCs. Elsner (2003) pointed out the importance of grouping TC tracks to improve the understanding and the prediction of regional TC activity such as landfalling TCs. Kossin et al. (2010) classified all North Atlantic TC tracks into four clusters, revealing distinct characteristics for each cluster in terms of their tracks and genesis locations, seasonality, and relationship between frequency of TCs and climate variability. Colbert and Soden (2012) classified TC tracks into three groups (straight moving, recurving landfall, or recurving ocean) highlighting differences in the climate conditions associated with each one of them. However, there is no information about how predictable these TC clusters are.

In this study, we first examine the predictability of observed basin-total TC frequency of observed TC clusters. Second, we attempt to construct a hybrid model to improve the prediction skill in TC frequency for each cluster, which in turn leads to the improvements in predicting basin-total TC frequency. Third, we examine observed and predicted TC landfall ratios and construct a hybrid model to improve the prediction skill in TC landfall ratio. Finally, we show the forecasting skill in landfall TC frequency over the United States predicted

by Forecast-oriented Low Ocean Resolution (FLOR) and compare the results from our newly developed hybrid model.

The remainder of this paper is organized as follows. Section 2 describes the models and data used in this study. Section 3 assesses the performance of the hybrid models in predicting TCs within each considered TC cluster, TC landfall ratio, and TC landfall frequency over the United States and compares these results to the dynamical model. Finally, section 4 provides a summary of the results.

2. Methods

Throughout this study, we focus on the prediction of North Atlantic TCs during July–November because about 92% of all storms occurred during these months over the 1980–2014 period. We focus on tropical storms or more intense cyclones (wind speed $\geq$ 34 kt; 1 kt = 0.51 m s^{-1}), and these storms are defined as TCs. The targeted prediction is the frequency of basin-total TCs and landfalling TCs along the U.S. coastline. In this section, dynamical models, observed data, and TC detection algorithms are described.

a. Dynamical model

The dynamical model used for retrospective seasonal forecasts is the Forecast-oriented Low Ocean Resolution (FLOR; Vecchi et al. 2014; Jia et al. 2015) of the Geophysical Fluid Dynamics Laboratory (GFDL) Coupled Model, version 2.5 (CM2.5; Delworth et al. 2012). FLOR comprises 50-km mesh atmosphere and land components, and 100-km mesh sea ice and ocean components. The number of vertical levels for the atmosphere component is 32 layers (top at 1.0 hPa). For each year and each month in the period 1980–2014, 12-month duration predictions were performed after initializing the model to observationally constrained conditions. Here, we defined forecasts from July, June, . . . , January, December initial conditions as the lead month (L) 0, 1, . . . , 6, 7 forecasts for the predictions of TC activity in the subsequent season (July–November).

The 12-member initial conditions for ocean and sea ice components were built through a coupled ensemble Kalman filter (EnKF; Zhang and Rosati 2010) data assimilation system developed for the GFDL Coupled Model, version 2.1 (CM2.1; Delworth et al. 2006; Wittenberg et al. 2006; Gnanadesikan et al. 2006), whereas those for atmosphere and land components were built from a suite of SST-forced atmosphere–land-only simulations to the observed values using the components in FLOR. Therefore, the predictability comes entirely from the ocean and sea ice and may be thought of as a lower bound on the potential prediction skill of a model because predictability could also arise from atmospheric (particularly stratospheric) and land initialization. During the simulation using FLOR, simulated temperature and wind stress are adjusted using so-called "flux adjustment" in which the model's momentum, enthalpy, and freshwater fluxes from atmosphere to ocean are adjusted to bring the model's long-term climatology of SST and surface wind stress closer to observations and improve simulations of TCs and precipitation (Vecchi et al. 2014; Delworth et al. 2015).

b. Observational datasets and detection algorithm for tropical cyclones

The observed TC "best track" data were obtained from the National Hurricane Center best track hurricane database (HURDAT2; Landsea and Franklin 2013) as archived in the International Best Track Archive for Climate Stewardship (IBTrACS; Knapp et al. 2010) and used to evaluate the TC simulations in the retrospective seasonal predictions. We also use the Met Office Hadley Centre SST product (HadISST1.1; Rayner et al. 2003) as observed SST.

Model-generated TCs were detected directly from 6-hourly output using the tracking scheme documented in Murakami et al. (2015). In the detection scheme, the flood fill algorithm is applied to find closed contours of some specified negative sea level pressure (SLP) anomaly with a warm core (1-K temperature anomaly). The detection scheme also requires that the TC lasts for 36 consecutive hours while maintaining a warm core as well as a specified wind speed criteria (15.75 m s^{-1}).

3. Results

a. Clustering TC tracks and forecasting skill by a dynamical model

We first applied a clustering algorithm to observed TC tracks (Fig. 1, green tracks). The cluster technique used here is the fuzzy c-means clustering developed by Kim et al. (2011). Fuzzy clustering has been known to produce more natural classification results for datasets such as TC tracks that are too complex to determine their boundaries by looking for distinctive patterns (Kim et al. 2011). Following Kossin et al. (2010), the final number of clusters is equal to 4, yielding early recurving TCs (CL1), Gulf of Mexico and Caribbean TCs (CL2), subtropical (or extratropical transition)-type TCs (CL3), and classic "Cape Verde hurricanes" (CL4). Each cluster comprises a comparable number of the total storms as shown in the fractional ratio between 20% and 28%. When compared to Kossin et al. (2010), we obtained similar TC clusters.

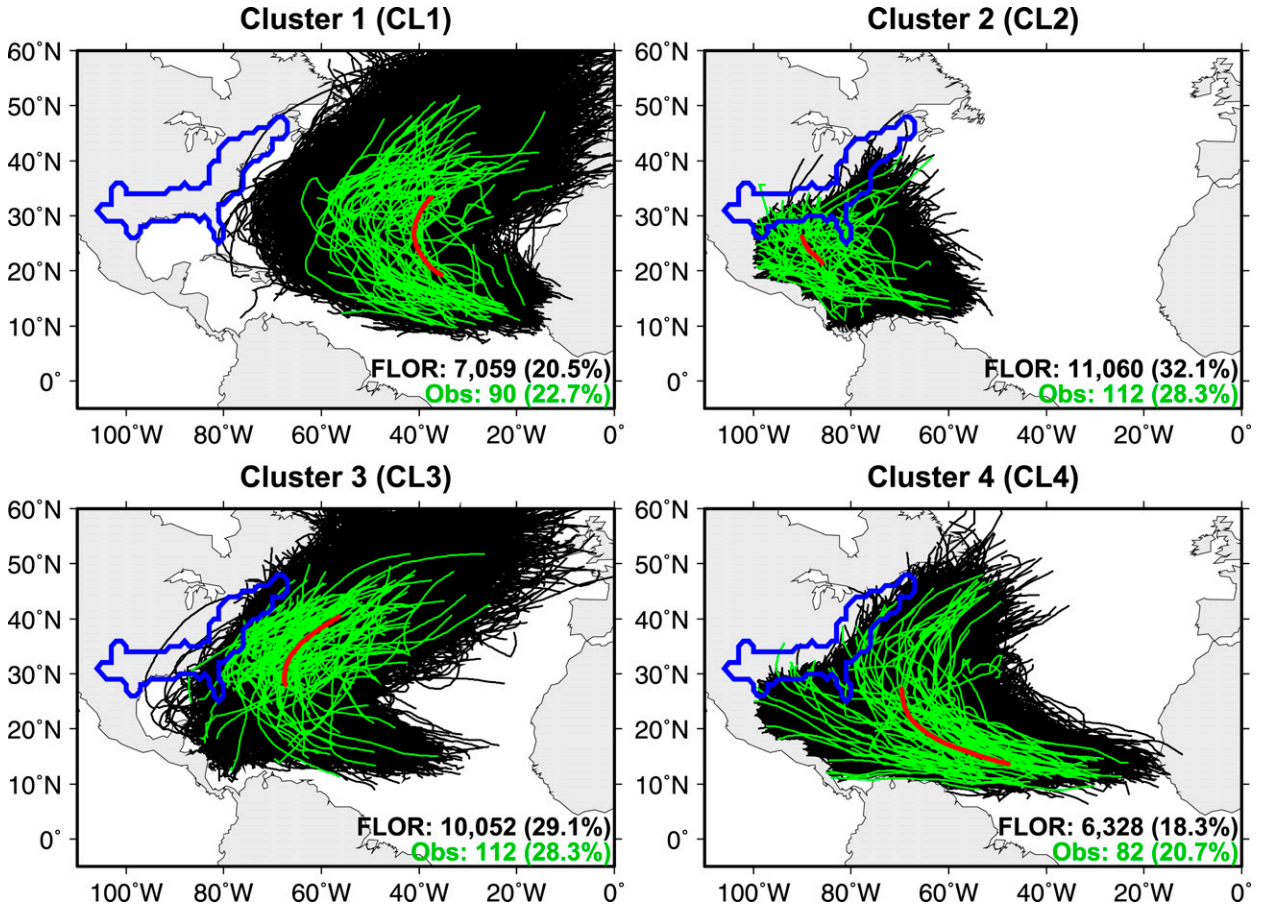

FIG. 1. TC tracks during the period 1980–2014 as separated by the cluster analysis yielding: Clusters 1–4. The fuzzy *c*-mean clustering is first applied to the observed TC tracks (green), which yields mean tracks (red) for each cluster. All predicted TC tracks by FLOR (black) are assigned to an observed TC cluster, regardless of any lead months and ensemble members, based on the RMSE between the predicted and mean TC track. The numbers in the panel bottom-right corners indicate the sample size with the fractional ratio in parentheses. The blue domain in each panel shows the region of the United States considered for landfall.

The TC cluster CL1, CL2, CL3, and CL4 as shown in Fig. 1 correspond to CL3, CL2, CL1, and CL4 in Kossin et al. (2010), respectively.

Second, we assigned predicted TCs to one of the observed TC clusters. Regardless of the lead-month forecasts and ensemble members, we computed the root-mean-square error (RMSE) between the predicted (black track) and observed mean (red track) TC track for each TC cluster. To compute RMSE, we interpolate every TC track into 20 segments with equal length following Kim et al. (2011). We assign the predicted TC track to the TC cluster with the minimum RMSE. An alternative way is to conduct the cluster analysis using the combined data of observed and predicted TC tracks. However, because we obtained similar results to the method above (figure not shown), we will use the RMSE for the assignment. The results for assigned TC tracks are shown in Fig. 1 as black tracks. Although the dynamical model predicts fractional ratios of TC frequency for CL1 and CL3 similar to the observations (about 20%), it slightly overestimates (underestimates) the fractional ratio for CL2 (CL4). Figure 2 shows forecast skill in predicting TC frequency for each cluster and for each lead month by the dynamical model in terms of rank correlation (Fig. 2a) and RMSE (Fig. 2b). For the sample size of 35 years (i.e., 1980–2014), correlations of 0.33 and 0.43 are statistically significant at the 5% and 1% levels if the data are assumed to be independent in each year. Rank correlation for the basin-total TC counts (black line in Fig. 2a) is about 0.6 for lead month $L = 0$–2, and decreases to about 0.4 for $L =$ 5–7. Vecchi et al. (2014) also reported similar results for the correlations for the basin-total frequency. RMSE for the basin-total TC counts (black line in Fig. 2b) is about 5–7. This large RMSE is mainly because of the underestimation in predicting TC frequency as also

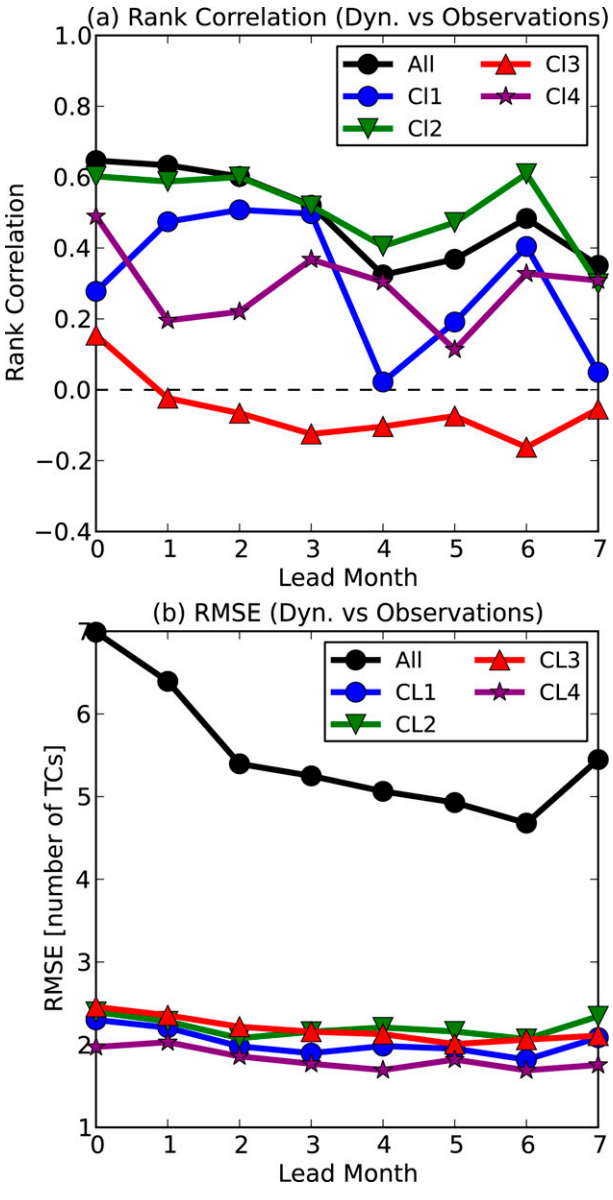

FIG. 2. Forecast skill in predicting TC frequency for each cluster and for each lead month predicted by the dynamical model. (a) Rank correlation between observed and predicted and (b) RMSE. Black line shows the basin-total TC frequency as defined as the total TC frequency among the TC clusters.

reported in Murakami et al. (2015). Shorter lead-month predictions show larger RMSE (black line in Fig. 2b). Although further investigation is required, this may be related to initial spinups due to the cold SST bias in the initial condition in the tropical North Atlantic, which is inherited from CM2.1 for the initialization. During the predictions for the first few months, FLOR tries to adjust the cold bias through the flux adjustments; however, it may take a few months to adjust the cold biases.

Among the four TC clusters, the dynamical model shows relatively higher skill in predicting CL2, followed by CL1 and CL4. The higher skill in predicting CL2 is consistent with Vecchi et al. (2014), who reported that FLOR has higher skill in predicting TCs near the coastline of the Gulf of Mexico and Caribbean Sea. On the other hand, the FLOR predictions for CL3 show the lowest skill, indicating that the prediction of TCs that undergo extratropical transition remains challenging for dynamical models (see also Jones et al. 2003); it is unclear whether this reflects a deficiency in the models and initialization, or an inherent limit to the predictability of the year-to-year variations of CL3 storms. In fact, this type of storm often forms from nontropical areas of low pressure that undergo tropical transition. The environments for these storms, especially prior to development, may have relatively high levels of vertical shear. Consequently, this type of storm formation may have lower levels of predictability.

b. *Correlations between observed TC frequency and predicted large-scale parameters*

Section 3a showed that the dynamical model has the lowest skill in predicting CL3 and CL4 TCs. If these biases could be improved, prediction of basin-total frequency and landfalling TC frequency could potentially be improved as well. For this purpose, we start by constructing a hybrid model in which observed TC frequency is regressed and predicted using some key large-scale parameters simulated by the dynamical model for each cluster. To identify the key parameters, we first investigate correlations between observed TC frequency and large-scale parameters. The large-scale parameters considered are relative SST (RSST), geopotential height at 500 hPa (Z_{500}), and zonal component of vertical wind shear (200–850 hPa, WS). The RSST is defined as the local SST anomaly subtracted from the tropical mean (30°S–30°N) SST anomaly. We have performed a preliminary investigation including other parameters such as midlevel relative humidity, low-level relative vorticity, SLP, steering flow among others; however, these parameters did not show significant correlations with TC frequency for any clusters (figure not shown). Also, we prefer parsimonious models that incorporate a smaller number of predictors to help avoid an overfitting problem when a hybrid model is constructed.

Figure 3 shows a correlation map between the time series of observed TC frequency for each cluster and the three large-scale parameters computed for each $1° \times 1°$ grid box within the global domain during the peak season. For RSST (Figs. 3a–d), all clusters except for CL3 show higher positive correlations in the tropical North

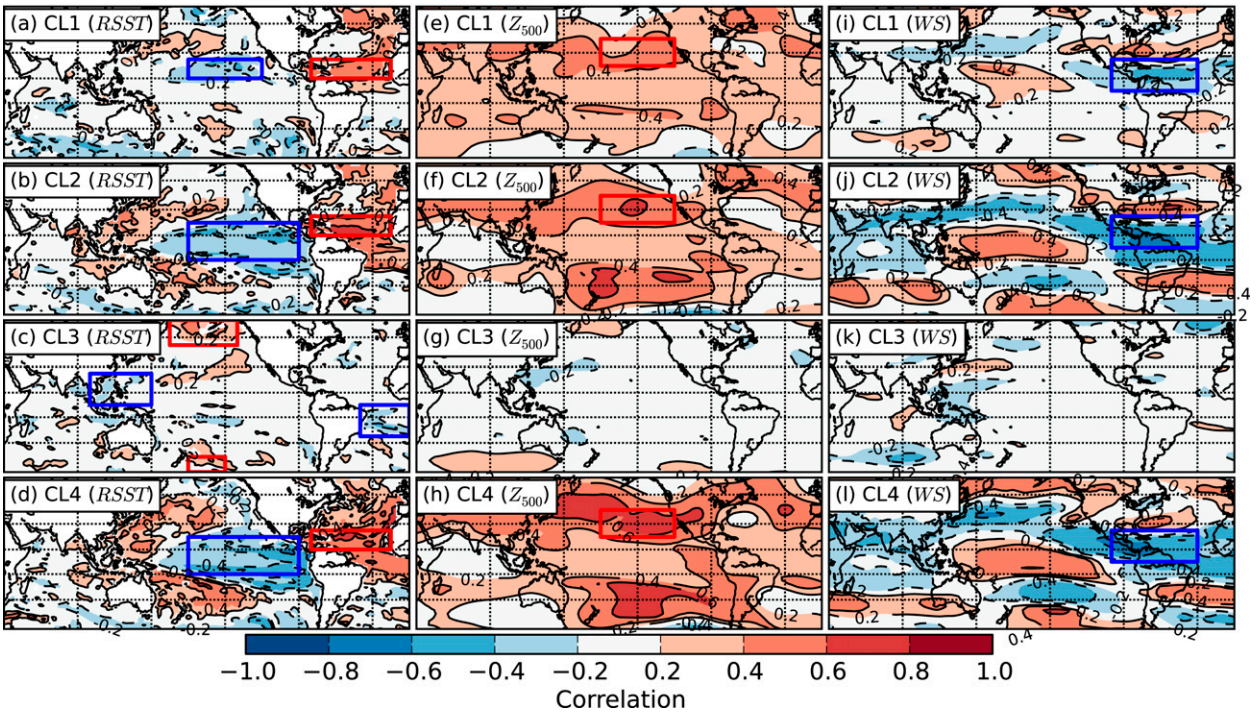

FIG. 3. Correlation map between the time series of observed TC frequency for (top to bottom) each cluster and (left to right) simulated mean large-scale parameters during July–November for each 1° × 1° grid box. (a)–(d) RSST, (e)–(h) Z_{500}, and (i)–(l) WS. Rectangles indicate domains for predictors with red (blue) rectangles showing positive (negative) signs.

Atlantic, which is consistent with previous studies (e.g., Villarini et al. 2010; Vecchi et al. 2011; Villarini and Vecchi 2012). CL2 and CL4 also show the La Niña–like pattern in the Pacific, indicating that TC frequency for these clusters increases during La Niña years. Although CL1 does not show the La Niña–like pattern clearly, the cluster shows a negative correlation in the subtropical Pacific. CL3 is unique with respect to the other clusters because there is no significant pattern in the correlation even in the tropical North Atlantic, indicating that CL3 is insensitive to local SST anomalies. As for Z_{500} (Figs. 3e–h), there are higher positive correlations in the subtropical central Pacific for CL1, CL2, and CL4. A preliminary investigation implies that this correlation is related to the Pacific–North American (PNA) pattern. We found that when the anomaly of Z_{500} is positive in the box of Figs. 3e, 3f, and 3h, Z_{500} in the subtropical North Atlantic (30°–50°N, 55°–75°W) is negative through a series of wave trains along the subtropical westerly jet. Moreover, during the positive phase, convection is active in the tropical North Atlantic and along the West African coast, leading to more robust easterly waves and TC development associated with the enhanced convection. On the other hand, CL3 shows no correlation with Z_{500}, which is again largely different from other clusters. For WS (Figs. 3i–l), there are higher negative correlations in the tropical North Atlantic for CL1, CL2, and CL4, which is reasonable because TC activity is unfavorable under strong vertical wind shear. However, it is intriguing that CL3 is not sensitive to vertical wind shear. This is probably because CL3 comprises a mixture of TCs and extratropical transition cyclones. The vertical wind shear and baroclinicity may provide energy to extratropical cyclones. The frequency of extratropical cyclones may show a positive correlation with WS, whereas that of TCs shows a negative correlation. Therefore, they cancel each other out, resulting in no correlation over the Atlantic as shown in Fig. 3k.

Kossin et al. (2010) showed how the frequency of each cluster responds uniquely to the El Niño–Southern Oscillation (ENSO), the Atlantic meridional mode (AMM, Servain et al. 1999; Xie and Carton 2004; Chiang and Vimont 2004; Vimont and Kossin 2007), and the North Atlantic Oscillation (NAO; Jones et al. 1997). To elucidate the potential influence of the natural variability on the frequency of TCs for each cluster, we revisited these responses for the TC cluster in this study. Correlations between the observed TC frequency for each cluster and the five climate indices are shown in Table 1. First, we used the Niño-3.4 index to represent ENSO. The Niño-3.4 index is obtained from the mean SST

TABLE 1. Rank correlation coefficients between the observed interannual variability of climate indices and the observed TC number for each cluster. The climate indices are Niño-3.4 (July–November mean), AMM (July–November mean), AMO (July–November mean), NAO (May–June mean), and SNAO (July–August mean). The observed TC number is the mean of July–November for each year. Statistical significance is highlighted according to the level of significance.

Index	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Niño-3.4 (July–November mean)	−0.06	−0.47[a]	+0.15	−0.35[b]
AMM (July–November mean)	+0.51[a]	+0.36[b]	−0.13	+0.36[b]
AMO (July–November mean)	+0.43[a]	+0.37[b]	−0.13	+0.26
NAO (May–June mean)	−0.15	−0.09	−0.14	+0.04
SNAO (July–August mean)	+0.24	+0.21	−0.24	+0.13

[a] Statistically significant at 99% level.
[b] Statistically significant at 95% level.

anomaly in the region bounded by 5°N and 5°S and between 170° and 120°W. CL2 shows the highest negative correlation with the Niño-3.4 index in addition to a relatively higher and statistically significant correlation for CL4. This result is consistent with Kossin et al. (2010) who argued that the tropical TC clusters (i.e., CL2 and CL4) show relatively higher correlations with ENSO than the midlatitude TC clusters (i.e., CL1 and CL3) do. Moreover, this result suggests that if FLOR were able to predict El Niño (or La Niña) accurately, it should show good skill in predicting TC frequency for the CL2 and CL4. In fact, FLOR shows higher skill in predicting CL2 (highest correlation) and CL4 (smallest RMSE) than the other TC clusters (Fig. 2).

Kossin et al. (2010) also examined the effect of AMM on TC frequency for each cluster. The AMM is the internal coupled mode of ocean subsurface temperature and lower atmospheric wind fields in the tropical and subtropical North Atlantic (Chiang and Vimont 2004). The AMM index is the standardized first expansion coefficient of the singular value decomposition (SVD) mode for the SST and zonal and meridional components of the 10-m wind field. The input data are defined over the tropical to subtropical region (21°S–32°N, 0°–74°W), and seasonal cycle, Niño-3.4 index, and linear trend are removed for each grid cell. During a positive phase of the AMM, the Atlantic intertropical convergence zone (ITCZ) is displaced northward. Warmer-than-normal SSTs and weaker-than-normal vertical wind shear during positive phases of the AMM tend to enhance TC development in the Atlantic (Vimont and Kossin 2007). The slow variation of the AMM generally reflects the decadal variability that is described by the Atlantic multidecadal oscillation (AMO; Delworth and Mann 2000; Kossin et al. 2010). We calculated the AMO index following Deser et al. (2010). The AMO index is defined as the area-average SST anomaly over the North Atlantic (0°–70°N, 0°–90°W) minus the global mean SST anomaly. Table 1 shows correlations between the AMM or AMO index versus TC frequency for each cluster. The CL1, CL2, and CL4 (CL1 and CL2) show significant correlation with the AMM (AMO) index. These results are consistent with Kossin et al. (2010), pointing to the potential use of the SST anomaly over the tropical Atlantic as a predictor in a hybrid model to predict TC frequency for these clusters.

Kossin et al. (2010) also showed that the May–June NAO index (Hurrell et al. 2015), which is defined as sea level pressure difference between Gibraltar and Reykjavik, Iceland (Jones et al. 1997), relates significantly with midlatitude storms during the summer months (i.e., our CL3 storms). Here, we revisit the relationship between the NAO index and TC frequency for each cluster. However, we could not find a significant relationship between them (Table 1). The NAO tends to be weaker during the summer months because of weaker equator–pole temperature gradients. Alternatively, Folland et al. (2009) redefined the summer NAO (SNAO) considering the difference in spatial patterns of NAO between summer and winter. In this study, the SNAO is defined as the second empirical orthogonal function (EOF)[1] of summertime (July–August) mean sea level pressure over the extratropical North Atlantic (25°–70°N, 70°W–50°E) for which the spatial pattern of the SNAO mode is similar to the ''mobile NAO'' reported by Portis et al. (2001) (Fig. 4). Table 1 shows correlations between the SNAO index and TC frequency for each cluster. The CL3 shows a negative correlation, which is consistent with Kossin et al. (2010), indicating that the CL3 TCs tend to be more frequent during a negative phase of SNAO. However, the correlation is still small and not statistically significant, as for the NAO. The clear relationship between the NAO and TC frequency of CL3 reported by Kossin et al. (2010) was not obtained for this study.

[1] When EOF is applied to observed July–August mean SLP data during the period 1980–2014, the second mode, as shown in Fig. 4, explains 21.1% of the total variance.

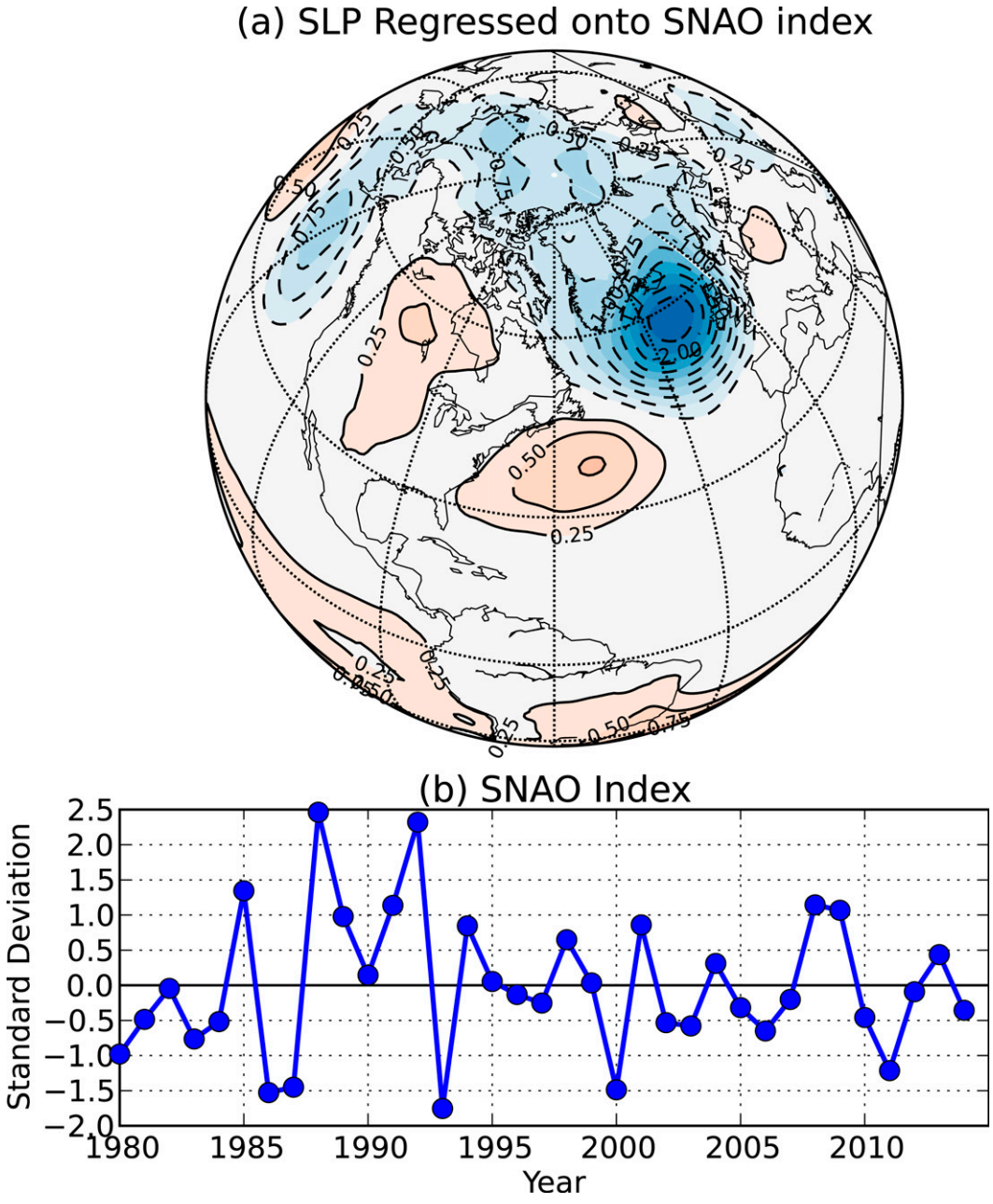

FIG. 4. (a) July–August mean SLP regressed onto the SNAO index (hPa σ^{-1}) with orange (blue) showing positive (negative) values. (b) Time series of SNAO index for the period 1980–2014 (σ).

CL3 does not show significant correlations with any parameters and climate indices, suggesting that CL3 may have a substantially larger stochastic element to its variability than the other clusters, and may thus be inherently less predictable. However, we want to identify any key large-scale parameters to construct a hybrid model. When the correlation is computed between observed TC frequency and RSST, observations (Fig. 5b) and dynamical models (Fig. 5a) show relatively higher correlations in the four domains. Although physical mechanisms explaining the relationship between Atlantic CL3 TCs and the remote SSTs are difficult to interpret, we utilized RSST in these four domains as a predictor for CL3 TCs.

The domains of the predictors used for the hybrid model are shown in the rectangles in Fig. 3. The dynamical model should also have significant forecast skill in predicting these large-scale parameters for each

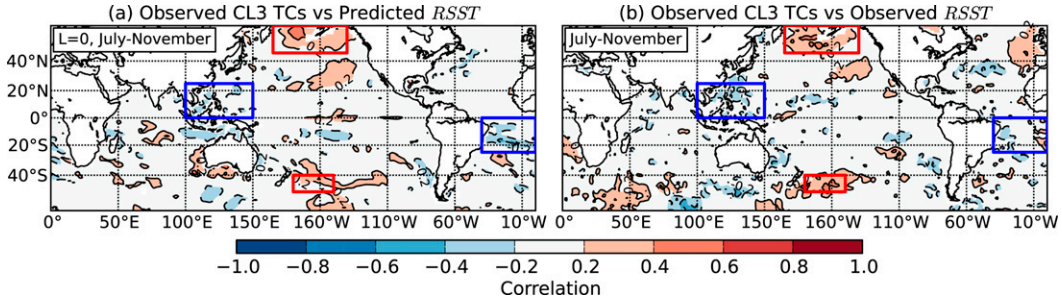

FIG. 5. Correlation map between RSST and CL3 TC frequency during July–November. (a) Correlation between observed RSST and predicted TC frequency by the model for CL3 at lead month 0. (b) As in (a), but for correlation between observed RSST and observed TC frequency for CL3. Red (blue) rectangles indicate domains for predictors showing a positive (negative) sign.

domain. Figure 6 shows anomaly correlations for the lead months 0, 3, and 6, respectively, for each parameter. The red shaded area is the region where the anomaly correlation exceeds 0.5, revealing that the dynamical model has skill in predicting the large-scale parameters for each domain used for predictors, even for the lead month 6. The skill in predictions and the correlation with respect to the observations justifies the use of the large-scale parameters in the domains as the predictors for the hybrid model.

c. *Hybrid Poisson regression model*

Using the predictors discussed in section 3b, a Poisson regression model (e.g., Villarini et al. 2010; Elsner and Jagger 2013) is constructed to predict TC frequency for each cluster using large-scale parameters predicted by the dynamical model of FLOR. First of all, the probability of any TC frequency ($=y$) in the ith year is obtained when the mean frequency (i.e., rate) λ_i is given as follows.

$$p(Y_i = y) = \frac{e^{-\lambda_i}\lambda_i^y}{y!}, \tag{1}$$

where $y = 0, 1, 2, \ldots, \infty$. The Poisson regression model is expressed as

$$\log(\lambda_i) = \beta_0 + \sum_{j=1}^{p} \beta_j x_{i,j}. \tag{2}$$

There are p predictors (indicated by the x_j values) and $p + 1$ parameters (β_j values). The model uses the logarithm of the rate as the response variable, but it is linear in the regression structure. First, we determine β_j given the observed Y_i and predicted $x_{i,j}$ (regression or training). Then, the cross validation is performed to evaluate

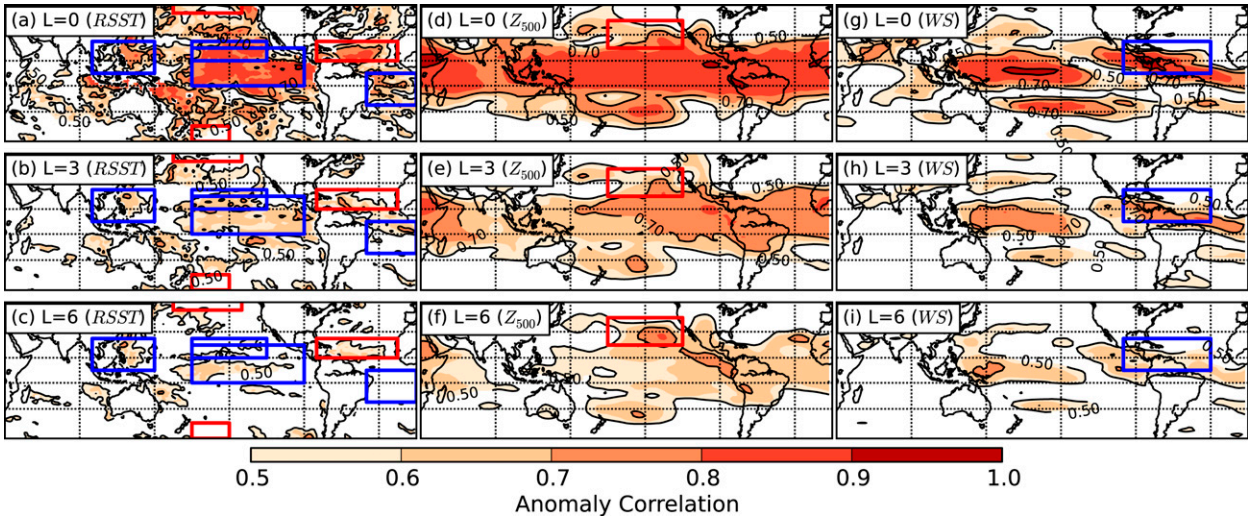

FIG. 6. Anomaly correlation for (left to right) simulated large-scale parameters for (top to bottom) each lead month $L = 0$, 3, and 6: (a)–(c) RSST, (d)–(f) Z_{500}, and (g)–(i) WS. The rectangles in these domains are the same as in Figs. 3 and 5.

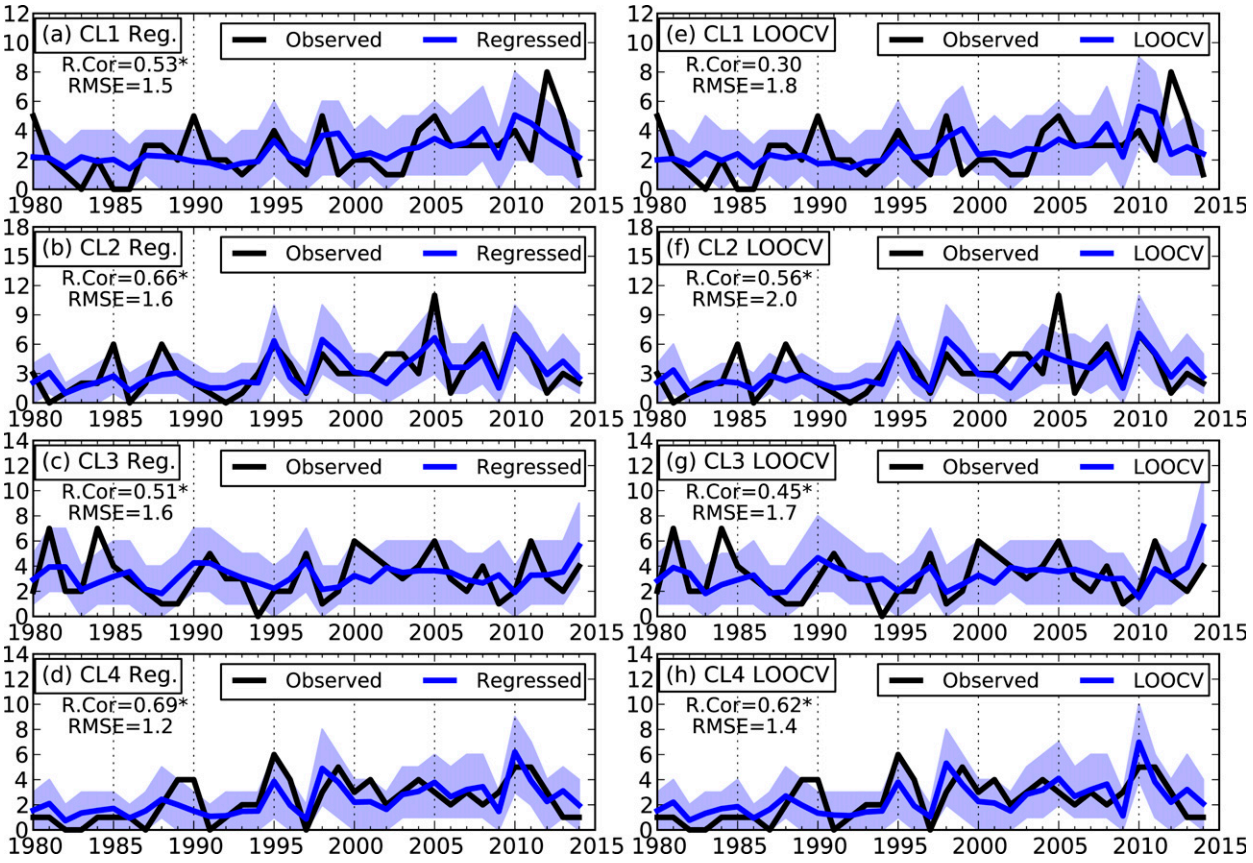

FIG. 7. Results of the interannual variation of TC frequency by (a)–(d) the regression and (e)–(h) the LOOCV for (top to bottom) each cluster at lead month 0. Observed (regressed or cross validated) TC frequency is shown in black (blue). Blue regions indicate 10% bottom range and 90% top range computed from random resampling based on the Poisson distribution. Numbers shown in each panel show rank correlation and RMSE between the black and blue lines. The asterisks indicate statistical significance of correlations at the 99% level.

the model skill. Here we apply so-called leave-one-out cross validation (LOOCV; Elsner and Jagger 2013). In the LOOCV, we first exclude a single year of observations and predictors; then, we determine the coefficients of the Poisson regression model using the remaining years. Using the model, the rate of TC frequency for the excluded year is predicted. This is done for 35 years, removing each year's data point successively. The use of the predictors from the fixed domains yielded from all of the years, including the prediction year, may lead to underestimations of prediction errors in LOOCV (e.g., chapter 7.10.2 in Hastie et al. 2009). Therefore, the hybrid model should be built completely from scratch for each of the cross-validation evaluations. Here, for each evaluation year we perform the clustering analysis, and we select the predictors by making correlation maps between the observed TC frequency and simulated large-scale fields using all of the data except for the evaluation year. Because leaving one year does not change correlation maps significantly from those obtained using all the years, the selected predictors for each evaluation step are identical to the original predictors; however, the domains of predictors are slightly shifted considering the maximum correlations.

Figure 7 reveals results of training (Figs. 7a–d) and LOOCV (Figs. 7e–h) for each cluster at lead month 0. To compare skill in these hybrid models with the dynamical model, Fig. 8 shows comparisons of rank correlations (Figs. 8a,e) and RMSE (Figs. 8b,f) between the dynamical model (solid lines) and LOOCV (dashed lines). Predictions for all clusters using the hybrid approach are improved in LOOCV in terms of RMSE for every lead month. Although CL1 was not improved in terms of correlation, most of the clusters show improvements in simulating observed interannual variations. When these predicted TC frequencies are summed up, we derive the basin-total TC frequency. The basin-total frequency also shows higher skill in the hybrid than in the dynamical model (Figs. 8e–h). We obtain a maximum correlation coefficient of 0.75 at lead month 1 and the correlation remains relatively high at 0.55 at lead month 7. On the other hand, the values of the correlation

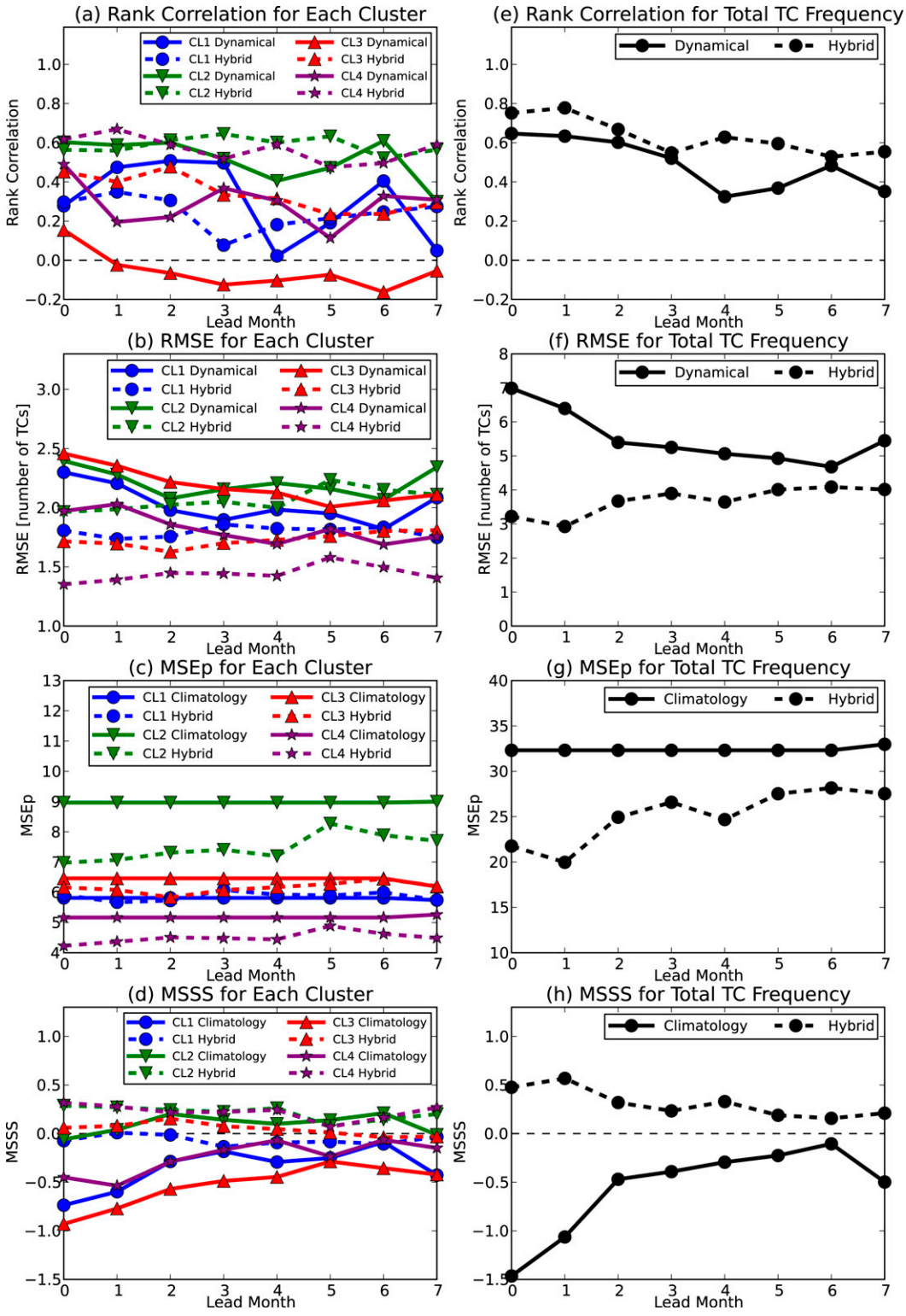

FIG. 8. (a)–(d) Comparisons of forecast skills between dynamical model or climatology (solid line) and hybrid model (dashed line) in predicting TC frequency for each cluster and for each lead month. (a) Rank correlation between observations and models. (b) RMSE between observations and models. (c) MSEp between climatology and hybrid model. (d) MSSS between observations and models. (e)–(h) As in (a)–(d), but for basin-total TC frequency.

coefficient for the dynamical model are 0.65 and 0.35 at lead month 1 and lead month 7, respectively, highlighting the improvements introduced by the hybrid model in predicting basin-total TC frequency.

Model skill is also assessed using the probabilistic form of the mean-square error (MSEp; Elsner and Jagger 2013) as defined in the following equation:

$$\mathrm{MSEp} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=0}^{\infty} p_i(j)(j-o_i)^2, \tag{3}$$

where i is index of year, j is number of TCs, n is number of years (=35), p_i is the probability of j TCs in the ith year given the predicted rate λ_i rate [i.e., Eq. (1)], and o_i is the observed TC count. Because the dynamical model has no probabilistic form of p_i, we compared MSEp with the statistical model based on the climatology in which Eq. (2) is replaced with the following equation:

$$\log(\lambda) = \beta_0. \tag{4}$$

Figures 8c and 8g compare MSEp between the hybrid and climatological models, highlighting the usefulness of the hybrid model, with the exception of the hybrid model for CL1, which is equivalent to the climatological forecast.

Another metric for model skill, the mean-square skill score (MSSS; Kim et al. 2012; Li et al. 2013) is defined with the following equation:

$$\mathrm{MSSS} = 1 - \frac{\frac{1}{n}\sum_{i=1}^{n}(o_i - f_i)^2}{\frac{1}{n}\sum_{i=1}^{n}(o_i - \lambda)^2}, \tag{5}$$

where n is the total number of years; o_i and f_i are the number of TCs from observations and predictions for the ith year, respectively; and λ is the observational mean number of TCs. The MSSS is a metric for the skill comparison of the model and climatological forecasts, with high values indicating a good model (Kim et al. 2012; Li et al. 2013). Figures 8d and 8h compare MSSS between the hybrid and climatological model, revealing higher scores for the hybrid model relative to the climatological model in predicting TC frequency for each cluster and whole basin.

The statistical–dynamical model developed in this study is further compared with previous available forecasts. The statistical forecasts issued by Colorado State University (CSU; Klotzbach and Gray 2015) reported that forecast skill in predicting named storms from June during 1982–2014 is +0.71 in terms of correlation coefficient, which is equivalent to the skill of our new model. Although the focus was on hurricanes (not for named storms), Wang et al. (2009) examined the forecast skill of a statistical–dynamical model using the vertical wind shear predicted by the Climate Forecast System (CFS) during August–October over the main development region as a predictor. They showed correlation coefficients of about 0.7 and 0.4 between observed and predicted hurricane counts for the lead months 0 and 6, respectively. Moreover, Vecchi et al. (2011) showed a high correlation coefficient of about +0.8 between observed and predicted hurricane counts for the lead-month-0 forecasts for the period 1982–2009 using simple predictors of tropical Atlantic and global mean SST anomalies. Because statistical modeling of seasonal hurricanes generally shows higher prediction skill than that of named storms (P. J. Klotzbach 2015, personal communication), constructing a statistical–dynamical model for hurricanes is an interesting topic for future studies.

d. *Correlations between observed TC landfall ratio and predicted large-scale parameters*

Section 3c revealed that the skill in predicting basin-total TC frequency is higher in the hybrid than in the dynamical model. Therefore, skillful forecasting of the fraction of total TCs making landfall in the United States could lead to accurate predictions of TC landfall activity when combined with predictions of basin-total TC frequency. In this section, we first investigate the physical drivers for the observed landfall ratio. The landfall domain defined in this study is the coastal region of the United States as identified in the blue region in Fig. 1. In this study, once a TC propagates into the blue region in Fig. 1, we count one for TC landfall frequency regardless of multiple landfall events for the same TC. Figure 9 shows the interannual variation of basin-total TC frequency (red), landfall TC frequency in the United States (blue), landfall ratio (black), and Niño-3.4 index (green) in the observations. The rank correlation between basin-total TC frequency and landfall ratio is 0.08, indicating that there is no strong linear relationship between the two variables (Holland 2007; Kossin et al. 2015). Indeed, while there were 18 TCs in 2010, which was the second largest TC frequency during the period 1980–2014, only one of them made landfall in the United States that year. The lack of relationship between basin-total and landfall TC frequencies is mainly due to systematic and significant relationships between climate and TC track variability (Holland 2007; Kossin et al. 2010, 2014, 2015; Villarini et al. 2012).

A number of previous studies relate U.S. landfalling frequency to NAO and Southern Oscillation index (SOI; Elsner and Kocher 2000; Elsner et al. 2000; Elsner

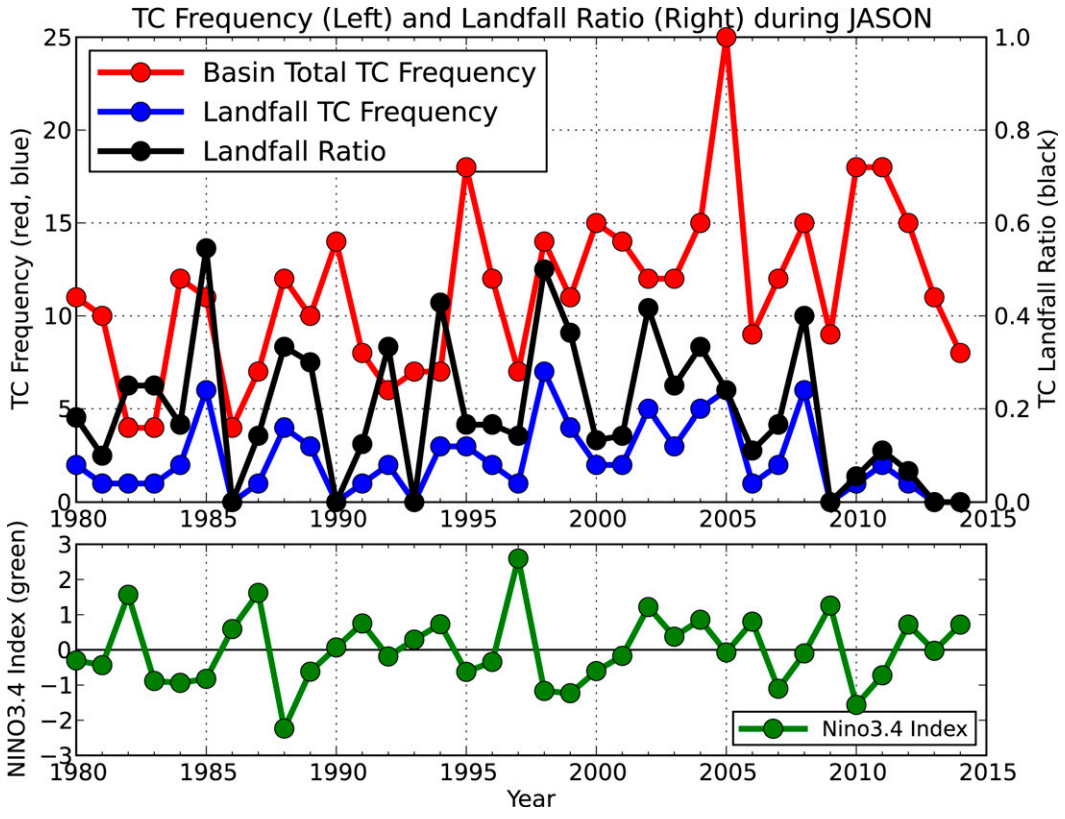

FIG. 9. (top) Observed time series of TC frequency and landfall ratio in the North Atlantic. The red line indicates observed basin-total TC frequency, whereas the blue line indicates observed landfall TCs over the United States. The black line shows the landfall ratio. (bottom) The green line shows the Niño-3.4 index that is obtained from the mean SST anomaly in the region bounded by 5°N and 5°S and between 170° and 120°W.

et al. 2006; Elsner and Jagger 2006). However, little is known for landfall ratio. The observed relationship between TC landfall ratio and climate indices was first analyzed by Villarini et al. (2012). They constructed a statistical model to predict landfall ratio using three predictors [May–June mean NAO, the SOI, and tropical mean SST (30°S–30°N)]. Here we computed the correlation coefficient between observed TC landfall ratio and observed climate indices (Table 2). Although statistical significance is low, we found that the observed TC landfall ratio is moderately and negatively correlated with the observed Niño-3.4 index, indicating that TC landfall ratio tends to be higher (lower) during La Niña (El Niño) years. The increased landfall ratio during La Niña years is consistent with Bove et al. (1998), who examined the effects of El Niño on U.S. landfalling hurricanes and found that the probability of U.S. hurricanes increased from 28% during El Niño years to 66% during La Niña years. The higher probability of U.S. landfall in La Niña years relative to El Niño years is likely due to differences in the steering flow as well as an increase in storm formations. In addition, the favorable large-scale conditions in La Niña years tend to be more conducive across the basin, allowing storms to persist at hurricane strength and reach coastal regions of the United States. However, although the rank correlation between observed landfall ratio (black line in Fig. 9) and the Niño-3.4 index (green line in Fig. 9) is negative (i.e., −0.24; Table 2), the correlation is not statistically significant at the 90% level (p value = 0.17), indicating

TABLE 2. Rank correlation coefficients and their p value (in parentheses) between the observed interannual variability of climate indices and the observed landfall ratio. The climate indices are the same as in Table 1. The observed landfall ratio is the mean of July–November for each year.

Index	Rank correlation (p value)
Niño-3.4 (July–November mean)	−0.24(0.170)
AMM (July–November mean)	+0.07(0.682)
AMO (July–November mean)	+0.19(0.287)
NAO (May–June mean)	+0.12(0.507)
SNAO (July–August mean)	+0.40(0.019)[a]

[a] Statistically significant at 95% level.

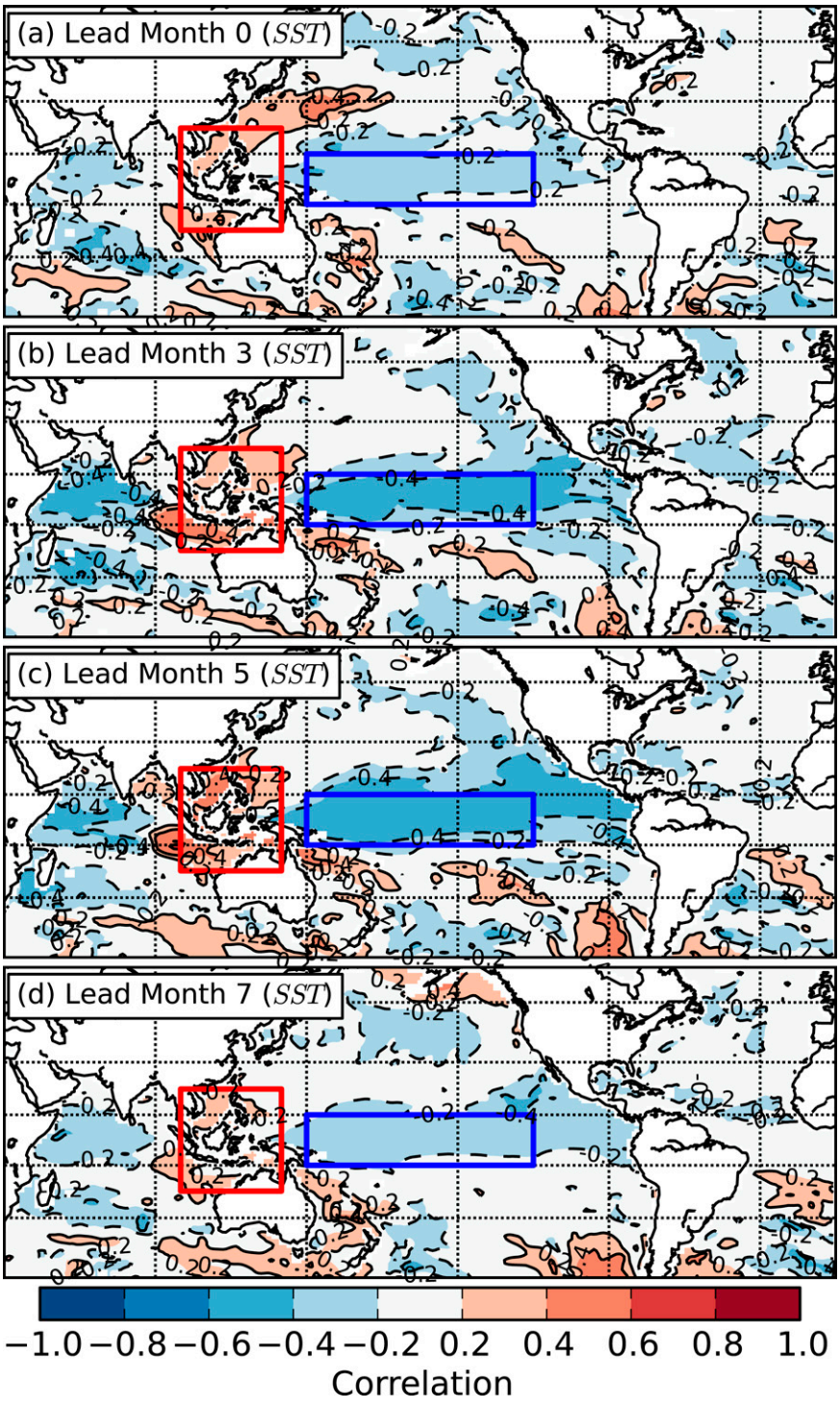

FIG. 10. Correlation map between observed TC landfall ratio over the United States and simulated SST anomalies for (a)–(d) each lead month: 0, 3, 5, and 7.

that landfall ratio is only slightly higher during La Niña. We also found a La Niña–like pattern in the correlation between the observed TC landfall ratio and predicted SST for each grid cell (Fig. 10). This highlights the model's predictability of ENSO even for the lead-month-7 forecasts, which in turn leads to the predictability of TC landfall ratio through the accurate prediction of ENSO. Other natural variability of AMM and AMO show positive

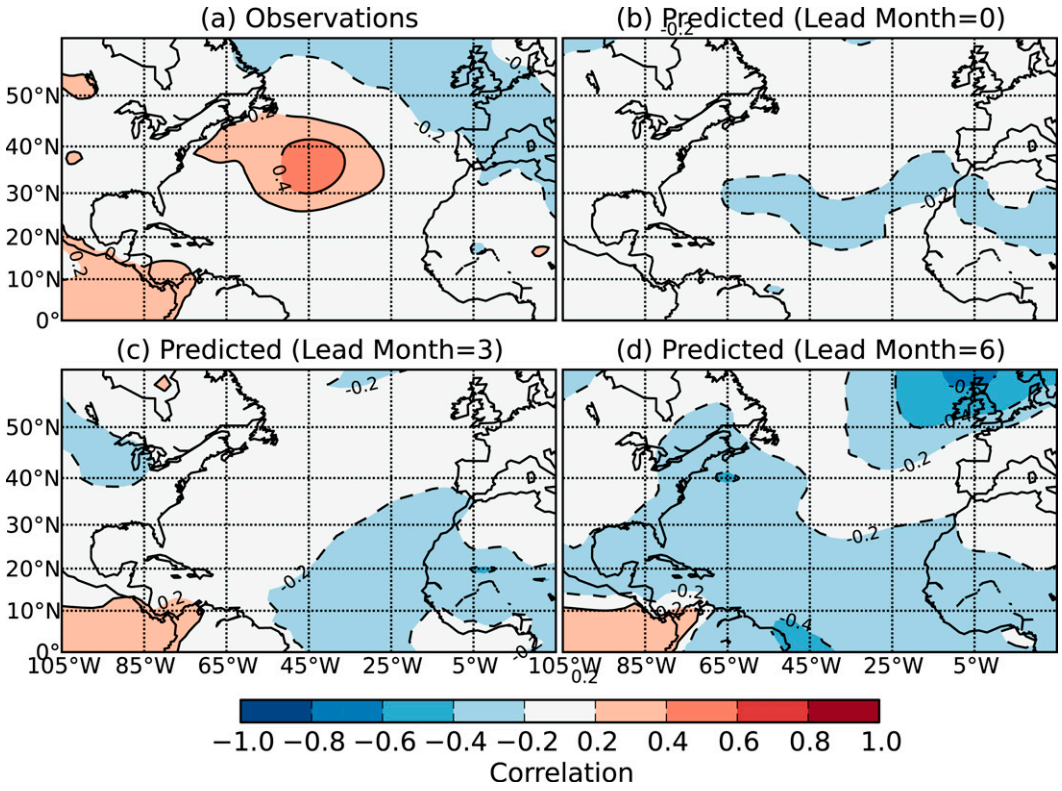

FIG. 11. Correlation map between observed mean TC landfall ratio over the United States during July–November and observed–simulated mean SLP anomalies during July–August: (a) observations, (b)–(d) predictions from lead month 0, 3, and 6, respectively, with orange (blue) showing positive (negative) values.

correlations with the TC landfall ratio, even though they are not statistical significant (Table 2). We also preliminarily computed correlations between TC landfall ratio and large-scale parameters of steering flows at 700 hPa and geopotential height at 700 hPa because TCs are steered approximately at the 700-hPa level. Although observations show marginal correlations in the North Atlantic, FLOR does not show significant correlations for these parameters. This is probably because of the difficulties in predicting large-scale flows even a month in advance partially because the atmosphere is not initialized. Further comparisons with the forecast systems in which atmosphere is initialized are required.

Table 2 shows a correlation between interannual variation of the observed NAO index and the observed TC landfall ratio. It reveals that there is no significant correlation between them in contrast to the significant relationship reported by the previous literature (e.g., Villarini et al. 2012). Although further investigation is required, there are a number of uncertainties for the inconsistency among the studies, such as the difference in analysis periods or the domain defined for the TC landfall. Nevertheless, the observed SNAO index shows the highest correlation (+0.40) with the observed TC landfall ratio, indicating that SNAO has the potential to be a good predictor for TC landfall ratio. Figure 11 shows correlation maps between observed TC landfall ratio during July–November and observed–predicted SLP anomaly during (July–August). Although the observations (Fig. 11a) show a marked positive anomaly over the central North Atlantic (30°–40°N, 30°–50°W) associated with a positive phase of SNAO, FLOR (Figs. 11b–d) does not show the positive correlation over the region even for the lead-month-0 forecast (Fig. 11b). These results indicate poor skill in predicting NAO–SNAO by FLOR. We assume that this poor skill is probably due to the lack of vertical resolution to resolve the stratosphere (L. Jia 2015, personal communication) in addition to the lack in atmospheric initialization. Scaife et al. (2005) suggest a source of NAO variability is within the lower stratosphere, pointing to the need to reproduce the stratospheric variability in the models to fully simulate surface climate variations. In fact, a vertically high-resolution model shows high skill in predicting NAO (Scaife et al. 2014), although the lead month for the skillful predictions is still short (one month). Because FLOR does not show any useful information on

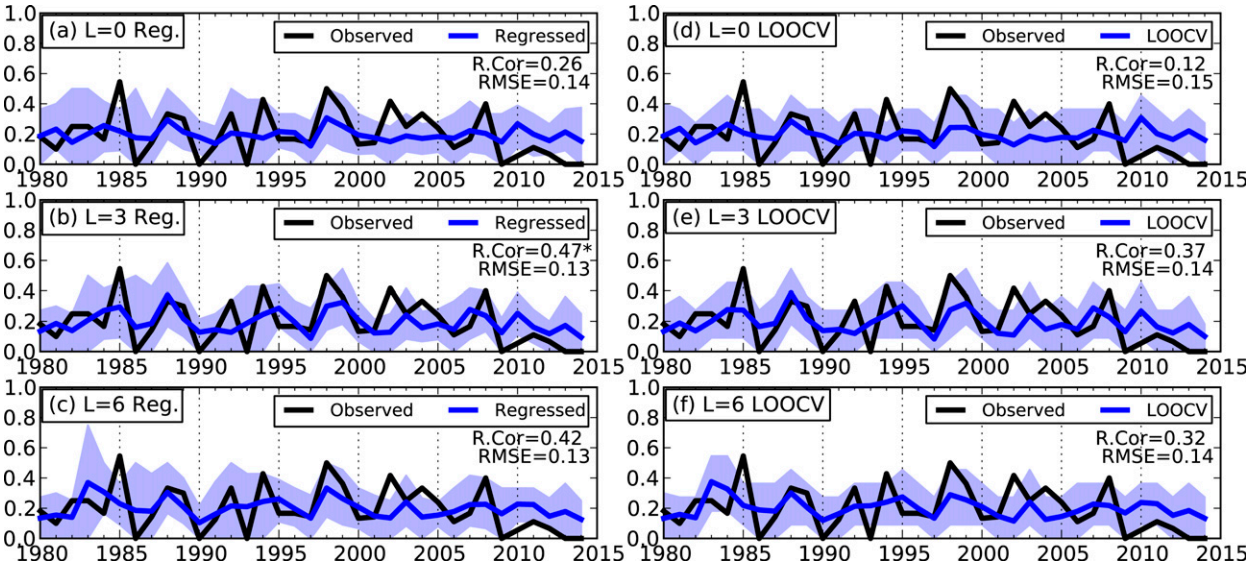

FIG. 12. Results of interannual variation of TC landfall ratio by (a)–(c) the regression and (d)–(f) the LOOCV for each lead month: (top) 0, (middle) 3, and (bottom) 6. Blue areas indicate the range between 10% and 90% computed from random resampling based on the binomial distribution. Numbers shown in each panel indicate rank correlation and RMSE between the black and blue lines. The asterisks indicate statistical significance of the correlations at the 99% level.

SNAO, we use only SST anomaly as a predictor in the hybrid binomial regression model, as will be shown in the following subsection.

e. *Hybrid binomial regression model*

As shown in section 3d, SST anomalies in the tropical Pacific are correlated with the TC landfall ratio for the United States as an indication of La Niña years. Although the correlation is not high, we use the SST anomalies in the domain shown in Fig. 10 as a predictor for the hybrid model for predicting TC landfall ratio using a binomial regression model (Villarini et al. 2012). Following Villarini et al. (2012), let us define Y_1 and Y_2 as two Poisson random variables with means of μ_1 and μ_2. Let us define m as their sum ($m = Y_1 + Y_2$), which also follows a Poisson distribution with mean equal to $\mu_1 + \mu_2$. In this study, m represents the basin-total TC frequency, whereas Y_1 represents the frequency of landfall TCs over the United States and Y_2 represents frequency of nonlandfall TCs. Given m, the distribution of Y_1 can be written as

$$f(Y_1 = y \mid \mu) = \frac{\Gamma(m+1)}{\Gamma(y+1)\Gamma(m-y+1)}\mu^{y}(1-\mu)^{(m-y)}, \tag{6}$$

where $\mu = \mu_1/(\mu_1 + \mu_2)$. The mean and the variance of Y_1/m are μ and $\mu(1-\mu)$, respectively. Similar to what is described in Eq. (2), we can relate the parameter μ to a vector of p predictors:

$$\log\left(\frac{\mu}{1-\mu}\right) = \beta_0 + \sum_{j=1}^{p}\beta_j x_j. \tag{7}$$

The dependence of μ on the predictors can be written explicitly as

$$\mu = \frac{\exp\left(\beta_0 + \sum_{j=1}^{p}\beta_j x_j\right)}{1 + \exp\left(\beta_0 + \sum_{j=1}^{p}\beta_j x_j\right)}. \tag{8}$$

Here we consider one predictor (i.e., $p = 1$) of SST as discussed above. Similar to the procedure described in section 3c, we first determine β_p given the observed μ and simulated x_p. Then, the LOOCV is performed to evaluate the hybrid model. As in section 3c for each evaluation year during the LOOCV process, we select the predictors by making correlation maps between the observed TC landfall ratio and simulated large-scale fields using all of data except for the evaluation year.

Figure 12 reveals results of training (Figs. 12a–c) and LOOCV (Figs. 12d–e) for lead months 0, 3, and 6, respectively. The overall correlation is relatively low for LOOCV (i.e., at most 0.37), indicating that landfall ratio remains difficult to predict even when using the hybrid model. To compare skill in the hybrid model with the dynamical model, Fig. 13 shows comparisons of rank correlations (Fig. 13a), RMSE (Fig. 13b), and MSSS (Fig. 13c) between the dynamical model (solid line) and

LOOCV (dashed line). Although rank correlation is lower when compared with the basin-total TC frequency (Fig. 8e), the hybrid model shows higher skill in predicting landfall ratio than the dynamical model does. RMSE (Fig. 13b) looks similar between the hybrid and dynamical models, although the hybrid model shows slightly lower RMSE than the dynamical model. MSSS (Fig. 13c) shows slightly higher skill in the hybrid model than the dynamical model.

f. Synthesized hybrid model for predicting landfall TCs over the United States

Here we have two hybrid models: the Poisson regression model to predict TC frequency for each cluster, yielding basin-total TC frequency by summing all TC clusters (section 3c), and the binomial regression model to predict TC landfall ratio over the United States (section 3e). By combining the two hybrid models, we can make predictions of TC landfall frequency over the United States. A schematic diagram is shown in Fig. 14 for the synthesized hybrid model. Given the key large-scale parameters for a specific year (step 1), we predict the mean TC frequency for each cluster ($\lambda_{1,2,3,4}$; step 2). Given the predicted mean λ, random resampling of TC frequency is performed for k times based on the Poisson distribution as shown in Eq. (1),[2] thereby yielding k samples of TC frequency (n) for each cluster (step 3). For each iteration, TC frequency for all the clusters is summed up, providing a sample of basin-total TC frequency (N) (step 3). A similar resampling procedure is performed for the TC landfall ratio, yielding k samples of landfall ratio (U; steps 4–6). For each sample, the landfall TC frequency over the United States (X) is computed by multiplying N and U (step 7). Based on the k samples for X, we can compute a probabilistic range (e.g., range of 10% bottom bound or 90% top bound) of the predicted TC landfall frequency as well as a mean X value for each year.

Figure 15 shows comparisons between the dynamic model and the synthesized hybrid model in terms of landfall TC frequency over the United States. First of all, the dynamical model systematically underestimates landfall TC frequency (Figs. 15a–c), whereas this underestimation is improved in the hybrid model (Figs. 15d–f). Moreover, the amplitude of the interannual variation is much larger in the hybrid than the dynamical model. For example, the anomalous year of 1998 is well predicted by the hybrid model. On the other hand, the hybrid model significantly overestimates TC landfall frequency

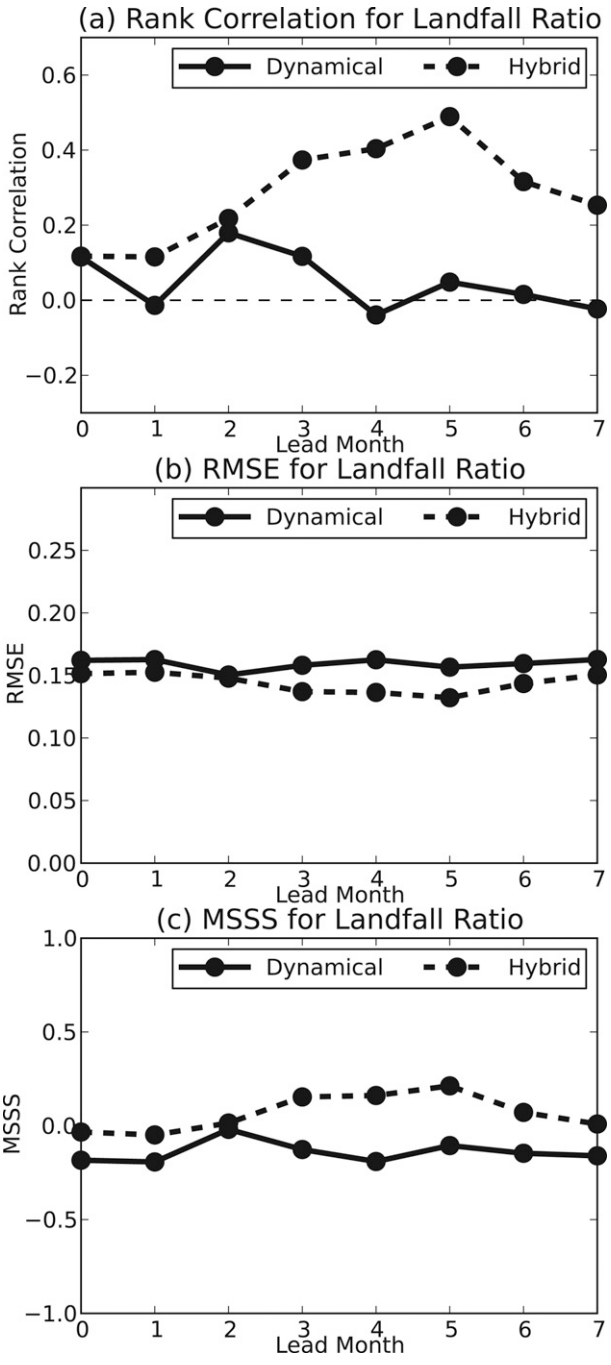

FIG. 13. (a)–(c) As in Figs. 8(e),(f),(h), but for the TC landfall ratio over the United States.

in 2010. In 2010, most TCs formed in the far eastern Atlantic and tended to recurve before they made landfall over the United States. That year was characterized by La Niña conditions. From Figs. 3a–d and Fig. 10, both TC frequency for each cluster and TC landfall ratio are expected to be large for 2010. That year is also characterized by negative SNAO (Fig. 9) that is unfavorable to landfall

[2] We used the function "rpois" in the statistics tool R, which is based on Ahrens and Dieter (1982).

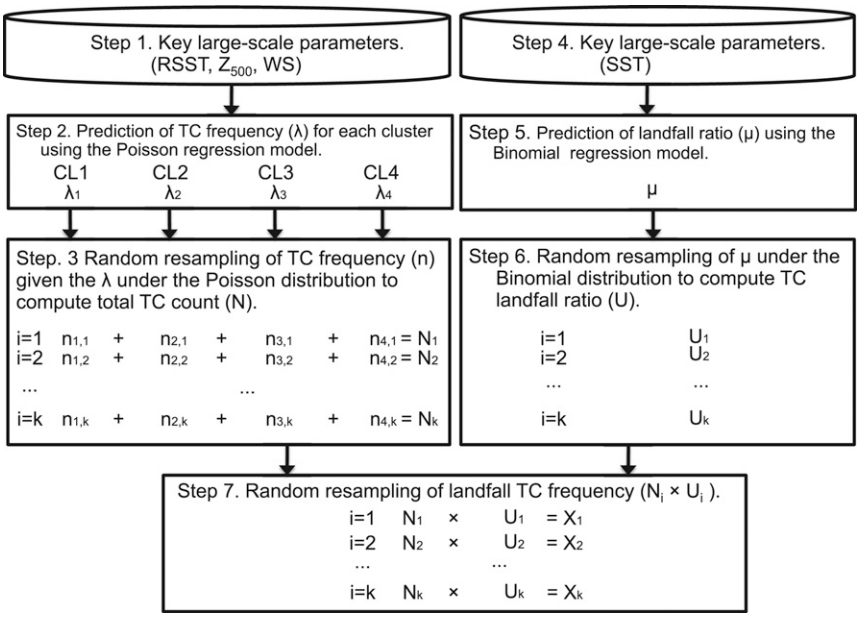

FIG. 14. Schematic diagram showing synthesized hybrid model to predict landfall TC frequency. Details are explained in the main text.

TCs. Had FLOR been able to predict the negative SNAO, it would have been possible to avoid the large number of landfall TCs in 2010 by the hybrid and the dynamical models.

Figure 16 summarizes the comparisons of rank correlations, RMSE, and MSSS for each lead month between the dynamical and hybrid models. The hybrid model shows comparable or higher skill for most of the lead months in terms of rank correlations relative to the dynamical forecasts, although correlations show no significant differences in the first three and the last two lead months (Fig. 16a). The hybrid model shows smaller RMSE (higher MSSS) for all lead months relative to the dynamical forecasts. We can conclude that our new

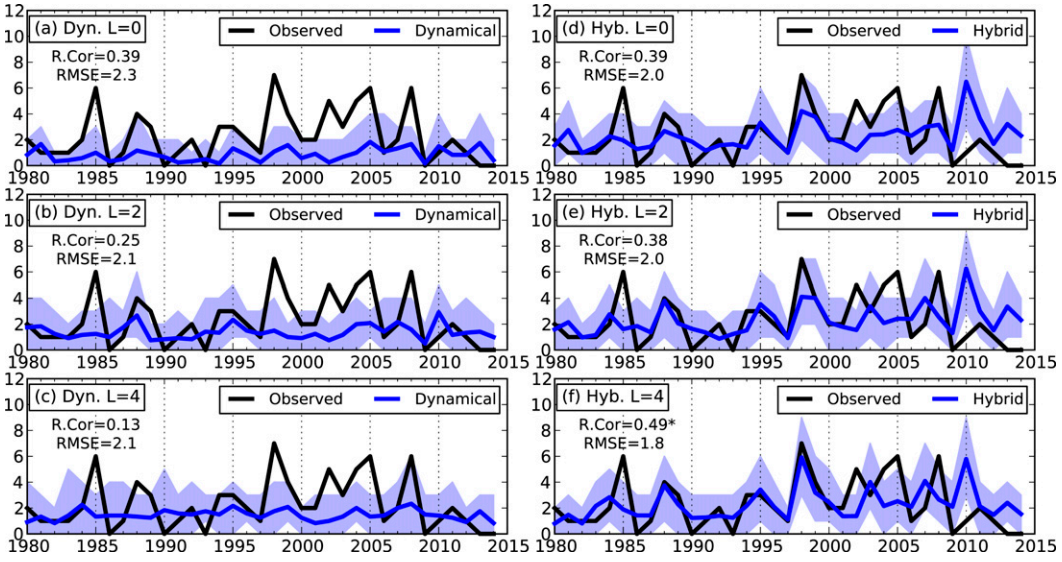

FIG. 15. Results of the interannual variation in the frequency of landfall TCs by (a)–(c) the dynamical model and (d)–(f) the synthesized hybrid model for each lead month of (top) 0, (middle) 2, and (bottom) 4. Blue areas indicate 10%–90% range computed from random resampling. Numbers in each panel show rank correlation and RMSE between the black and blue lines. The asterisks indicate statistical significance of the correlations at the 99% level.

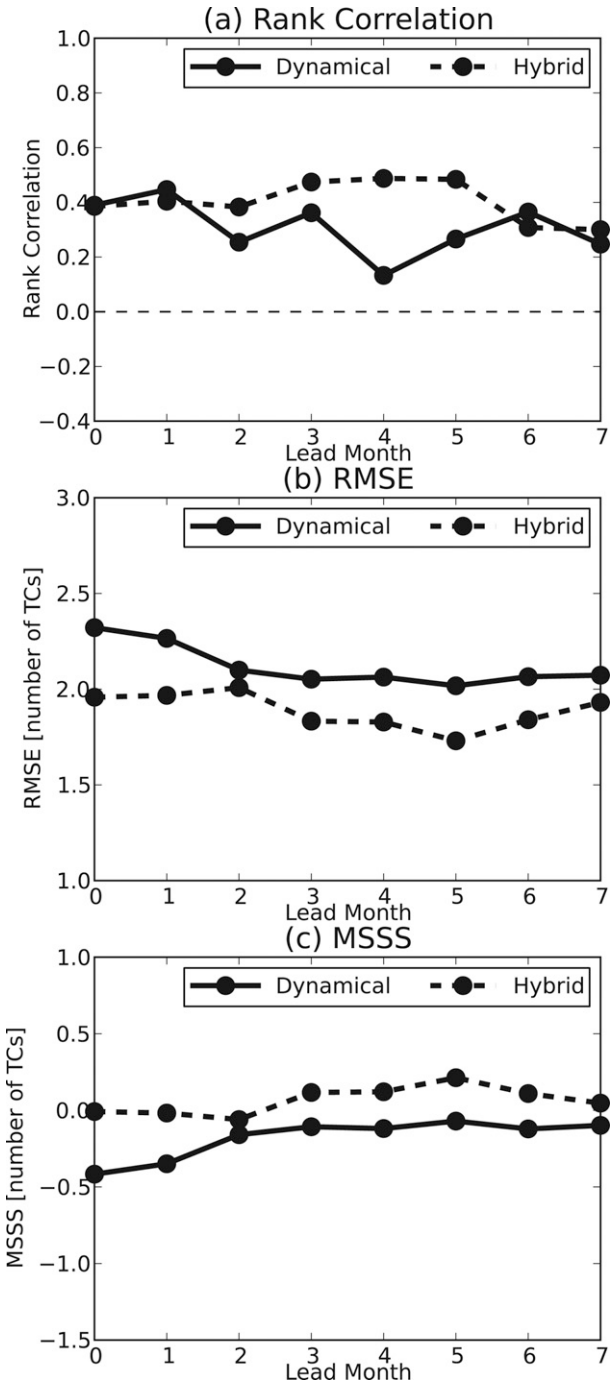

FIG. 16. (a)–(c) As in Figs. 8(e),(f),(h), but for frequency of landfall TCs over the United States.

hybrid model retains forecast skill up to lead month 5 with a correlation coefficient of +0.5 and forecast RMSE of 2.0 storms per year for U.S. landfalling TCs.

The forecast skill in predicting landfall TC count is further compared with previous studies using statistical and statistical–dynamical models, although direct comparisons are difficult because of difference in analyzed period and targeted domain. Our statistical–dynamical model shows comparable skill to the statistical model developed by Elsner et al. (2006), which shows a correlation coefficient of +0.35 between observed and predicted U.S. landfalling hurricane counts from February (i.e., lead month 6) forecasts for the period 1992–2004. In addition to the different time periods for both training and verification, it is also worth highlighting that the focus of Elsner et al. (2006) was on landfalling hurricanes rather than TCs more generally as done in this study. Kim et al. (2015) reported a correlation coefficient of +0.56 (+0.57) between the observed TC number affecting New York State and their statistical (and statistical–dynamical) model predictions from May during 1979–2013. Although it is a statistical model, Yan et al. (2015) showed a correlation coefficient of +0.60 between observed and predicted landfall TC counts on the eastern seaboard of the United States after August for the period 1950–2012 using May–July climate indices of AMM, AMO, and ENSO as predictors. Camp et al. (2015) used a dynamical model and showed limited skill (+0.12 for correlation coefficient) in predicting U.S. landfalling TCs from May. Overall, although seasonal forecasting of U.S. landfalling TCs has been improving over the years, the forecast skill is at most +0.60 among the cited studies, highlighting that the forecasting of landfalling activity is much more challenging than the forecasting of the overall TC activity.

We also preliminarily checked the performance of an alternative statistical method in which TC landfall frequency is computed using the constant climatological mean landfall ratio based on observations along with the basin-total TC frequency predicted from the Poisson regression model. Although the method shows some improvements in terms of RMSE for the lead month 0 and 1 predictions relative to the synthesized hybrid model, the scheme does not show improvements in terms of rank correlation.

The lead month 5 shows the highest skill in predicting landfall TC frequency (Fig. 16). This is mainly because the prediction skill in landfall ratio is the highest for the lead month 5 (Fig. 13) because of the highest correlations between the observed landfall ratio and predicted SST among all the lead months (Fig. 10). Although it is not still clear why lead month 5 shows the highest correlations for the SST in the key domains, accurate predictions for landfall TC frequency seem to be critically dependent on the accurate prediction for landfall ratio in which the present study shows limited skill (Fig. 13).

4. Summary

In this study, we evaluated retrospective seasonal forecasts based on a GFDL high-resolution coupled climate model (FLOR), and we constructed new hybrid models to improve the forecast skill in predicting the frequency of basin-total and U.S. landfalling TCs. First, we classified observed TCs into four groups of TCs using the fuzzy *c*-means clustering algorithm. Predicted TCs by FLOR are assigned to one of the observed clusters. We found that FLOR has high skill in predicting the Gulf of Mexico and Caribbean TCs (CL2), whereas it has low skill in predicting the subtropical TCs (CL3). The CL3 storms also exhibited limited statistical relationships to large-scale climate conditions, suggesting that the limited prediction skill may reflect limited underlying predictability.

Second, we constructed a hybrid model to predict TC frequency for each cluster using the empirical relationship between observed TC frequency and predicted large-scale parameters by a dynamical model. The hybrid model shows equivalent or higher skill in predicting TC frequency for each cluster relative to the dynamical model. The improvements for each cluster result in improved prediction of basin-total TC frequency. We obtained maximum and minimum values of the correlation coefficient equal to 0.75 and 0.55 at lead months 1 and 7, whereas those for counting TCs directly from the dynamical model are 0.65 and 0.35, respectively.

Third, we evaluated the retrospective prediction skill for the TC landfall ratio over the United States, revealing that the dynamical predictions have no skill in predicting the landfall ratio when looking at simulated storms directly. Meanwhile, the observed TC landfall ratio is analyzed, revealing that the landfall ratio has no correlation with basin-total TC frequency. However, the observed interannual variation in landfall TC ratio has a moderate correlation with the SST anomaly in the tropical Pacific. This is associated with a La Niña–like pattern, indicating that TC landfall ratio is higher during La Niña years. A binomial hybrid model was constructed for better prediction of U.S. landfalling TC ratios using the simulated SST anomaly in the tropical Pacific. The hybrid model predicts interannual variations of TC landfall ratio better than the dynamical model does.

By combining the two hybrid Poisson and binomial models, the frequency of TC landfall over the Unites States is predicted. The synthesized hybrid model shows comparable or better prediction of TC landfall frequency relative to the dynamical model in terms of RMSE, correlation, and MSSS between predicted and observed TC landfall frequency and amplitude of interannual variation. We can conclude that the new hybrid model retains forecast skill up to lead month 5 with a correlation coefficient of 0.5 and a forecast error of 2.0 for TC landfall for the United States.

In this study, we used the results of retrospective forecasts by FLOR in which the initial state of the atmosphere and land components are not constrained by observations, while the oceanic component is. We hypothesize that if we initialized the atmospheric component, we might obtain better skill in predicting TC activity in the North Atlantic. Similar tests with FLOR show improved seasonal predictions of land surface conditions with atmospheric initializations (Jia et al. 2016). Recent studies show that the dynamical models have longer prediction skills of 2 weeks or more because of accurate simulation of intraseasonal oscillations such as the Madden–Julian oscillation (MJO; e.g., Xiang et al. 2015a,b; Nakano et al. 2015). Further, Murakami et al. (2015) showed that FLOR has the capability of simulating a strong MJO signal. If atmospheric initial conditions contain observed MJO phase and amplitude, the dynamical model may predict TC activity well at least for the shortest lead-month forecast of $L = 0$. Moreover, Murakami et al. (2015) showed better prediction of 1997/98 TC activity using the higher-resolution version of FLOR (i.e., HiFLOR) in addition to the better simulations of large-scale parameters than FLOR. In the future, we plan to construct a hybrid model using large-scale parameters simulated using HiFLOR to improve the skill in predicting TCs in the North Atlantic. The new hybrid model is computationally more expensive relative to a statistical model because the hybrid model requires simulated large-scale fields from a dynamical model. However, given the selected predictors, it is possible to obtain their forecasts for the areas of interest very easily trough the North American Multi-Model Ensemble project (NMME; Kirtman et al. 2014). Of particular interest is to apply the new hybrid model to the NMME predictions in order to see if we obtain higher skill by the multimodel ensembles.

Acknowledgments. This material is based in part upon work supported by the National Science Foundation under Grant AGS-1262099, Award NA14OAR4830101 to the Trustees of Princeton University (Gabriele Villarini). This report was prepared by Hiroyuki Murakami under Award NA14OAR4830101 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce. Gabriele Villarini also acknowledges financial support from the USACE Institute for Water Resources. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily

reflect the views of the National Oceanic and Atmospheric Administration, the U.S. Department of Commerce, or the U.S. Army Corps of Engineers. The authors thank Dr. Jan-Huey Chen, Dr. Lakshmi Krishnamurthy, and the review comments and suggestions by Dr. Phil Klotzbach and five anonymous reviewers.

REFERENCES

Ahrens, J. H., and U. Dieter, 1982: Computer generation of Poisson deviates from modified normal distributions. *ACM Trans. Math. Software*, **8**, 163–179, doi:10.1145/355993.355997.

Alessandri, A., A. Borrelli, S. Gualdi, E. Scoccimarro, and S. Masina, 2011: Tropical cyclone count forecasting using a dynamical seasonal prediction system: Sensitivity to improved ocean initialization. *J. Climate*, **24**, 2963–2982, doi:10.1175/2010JCLI3585.1.

Bove, M. C., J. B. Elsner, C. W. Landsea, X. Niu, and J. J. O'Brien, 1998: Effect of El Niño on U.S. landfalling hurricanes, revisited. *Bull. Amer. Meteor. Soc.*, **79**, 2477–2482, doi:10.1175/1520-0477(1998)079<2477:EOENOO>2.0.CO;2.

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

——, ——, P. J. Klotzbach, and C. W. Landsea, 2007: Seasonal tropical cyclone forecasts. *WMO Bull.*, **56**, 297–309.

Camp, J., M. Roberts, C. MacLachlan, E. Wallace, L. Hermanson, A. Brookshaw, A. Arribas, and A. A. Scaife, 2015: Seasonal forecasting of tropical storms using the Met Office GloSea5 seasonal forecast system. *Quart. J. Roy. Meteor. Soc.*, **141**, 2206–2219, doi:10.1002/qj.2516.

Chen, J.-H., and S.-J. Lin, 2011: The remarkable predictability of inter-annual variability of Atlantic hurricanes during the past decade. *Geophys. Res. Lett.*, **38**, L11804, doi:10.1029/2011GL047629.

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

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

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

Delworth, T. L., and M. E. Mann, 2000: Observed and simulated multidecadal variability in the Northern Hemisphere. *Climate Dyn.*, **16**, 661–676, doi:10.1007/s003820000075.

——, and Coauthors, 2006: GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *J. Climate*, **19**, 643–674, doi:10.1175/JCLI3629.1.

——, and Coauthors, 2012: Simulated climate and climate change in the GFDL CM2.5 high-resolution coupled climate model. *J. Climate*, **25**, 2755–2781, doi:10.1175/JCLI-D-11-00316.1.

——, F. Zeng, A. Rosati, G. A. Vecchi, and A. T. Witternberg, 2015: A link between the hiatus in global warming and North American drought. *J. Climate*, **28**, 3834–3845, doi:10.1175/JCLI-D-14-00616.1.

Deser, C., M. A. Alexander, S.-P. Xie, and A. S. Phillips, 2010: Sea surface temperature variability: Patterns and mechanisms. *Annu. Rev. Mar. Sci.*, **2**, 115–143, doi:10.1146/annurev-marine-120408-151453.

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

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

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

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

——, and ——, 2013: Frequency models. *Hurricane Climatology: A Modern Statistical Guide Using R*, Oxford University Press, 161–193.

——, K.-B. Liu, and B. Kocher, 2000: Spatial variations in major U.S. hurricane activity: Statistics and a physical mechanism. *J. Climate*, **13**, 2293–2305, doi:10.1175/1520-0442(2000)013<2293:SVIMUS>2.0.CO;2.

——, J. Murnane, and T. H. Jagger, 2006: Forecasting U.S. hurricanes 6 months in advance. *Geophys. Res. Lett.*, **33**, L10704, doi:10.1029/2006GL025693.

Folland, C. K., J. Knight, H. W. Linderholm, D. Fereday, S. Ineson, and J. W. Hurrell, 2009: The summer North Atlantic Oscillation: Past, present, and future. *J. Climate*, **22**, 1082–1103, doi:10.1175/2008JCLI2459.1.

Gnanadesikan, A., and Coauthors, 2006: GFDL's CM2 global coupled climate models. Part II: The baseline ocean simulation. *J. Climate*, **19**, 675–697, doi:10.1175/JCLI3630.1.

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

——, 1984b: Atlantic seasonal hurricane frequency. Part II: Forecasting its variability. *Mon. Wea. Rev.*, **112**, 1669–1683, doi:10.1175/1520-0493(1984)112<1669:ASHFPI>2.0.CO;2.

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

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

——, ——, ——, and ——, 1994: Predicting Atlantic seasonal hurricane activity by 1 June. *Wea. Forecasting*, **9**, 103–115, doi:10.1175/1520-0434(1994)009<0103:PABSTC>2.0.CO;2.

Hastie, T., R. Tibshirani, and J. Friedman, 2009: The wrong and right way to do cross-validation. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer-Verlag, 245–247.

Holland, G. J., 2007: Misuse of landfall as a proxy for Atlantic tropical cyclone activity. *Eos, Trans. Amer. Geophys. Union*, **88**, 349–356, doi:10.1029/2007EO360001.

Hurrell, J., and NCAR Research Staff, 2015: The climate data guide: Hurrell North Atlantic Oscillation (NAO) index (station-based). NCAR Climate Data Guide, accessed 20 April 2016. [Available online at https://climatedataguide.ucar.edu/climate-data/hurrell-north-atlantic-oscillation-nao-index-station-based.]

Jagger, T. H., and J. B. Elsner, 2010: A consensus model for seasonal hurricane prediction. *J. Climate*, **23**, 6090–6099, doi:10.1175/2010JCLI3686.1.

Jia, L., and Coauthors, 2015: Improved seasonal prediction of temperature and precipitation over land in a high-resolution

GFDL climate model. *J. Climate*, **28**, 2044–2062, doi:10.1175/JCLI-D-14-00112.1.

——, and Coauthors, 2016: The roles of radiative forcing, sea surface temperatures, and atmospheric and land initial conditions in U.S. summer warming episodes. *J. Climate*, doi:10.1175/JCLI-D-15-0471.1, in press.

Jones, P. D., T. Jónsson, and D. Wheeler, 1997: Extension to the North Atlantic oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland. *Int. J. Climatol.*, **17**, 1433–1450, doi:10.1002/(SICI)1097-0088(19971115)17:13<1433::AID-JOC203>3.0.CO;2-P.

Jones, S. C., and Coauthors, 2003: The extratropical transition of tropical cyclones: Forecast challenges, current understanding, and future directions. *Wea. Forecasting*, **18**, 1052–1092, doi:10.1175/1520-0434(2003)018<1052:TETOTC>2.0.CO;2.

Kim, H.-S., C.-H. Ho, and P.-S. Chu, 2011: Pattern classification of typhoon tracks using the fuzzy *c*-means clustering method. *J. Climate*, **24**, 488–508, doi:10.1175/2010JCLI3751.1.

——, ——, J.-H. Kim, and P.-S. Chu, 2012: Track-pattern–based model for seasonal prediction of tropical cyclone activity in the western North Pacific. *J. Climate*, **25**, 4660–4678, doi:10.1175/JCLI-D-11-00236.1.

Kim, H.-M., K. Edmund, M. Chang, and M. Zhang, 2015: Statistical–dynamical seasonal forecast for tropical cyclones affecting New York State. *Wea. Forecasting*, **30**, 295–307, doi:10.1175/WAF-D-14-00089.1.

Kirtman, B. P., and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1 Seasonal-to-interannual prediction; Phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc.*, **95**, 585–601, doi:10.1175/BAMS-D-12-00050.1.

Klotzbach, P. J., 2008: Refinements to Atlantic basin seasonal hurricane prediction from 1 December. *J. Geophys. Res.*, **113**, D17109, doi:10.1029/2008JD010047.

——, and W. M. Gray, 2003: Forecasting September Atlantic basin tropical cyclone activity. *Wea. Forecasting*, **18**, 1109–1128, doi:10.1175/1520-0434(2003)018<1109:FSABTC>2.0.CO;2.

——, and ——, 2004: Updated 6–11-month prediction of Atlantic basin seasonal hurricane activity. *Wea. Forecasting*, **19**, 917–934, doi:10.1175/1520-0434(2004)019<0917:UMPOAB>2.0.CO;2.

——, and ——, 2009: Twenty-five years of Atlantic basin seasonal hurricane forecasts (1984–2008). *Geophys. Res. Lett.*, **36**, L09711, doi:10.1029/2009GL037580.

——, and ——, 2012: Qualitative discussion of Atlantic basin seasonal hurricane activity for 2012. Department of Atmospheric Science, Colorado State University, 27 pp. [Available online at http://hurricane.atmos.colostate.edu/forecasts/2011/dec2011/dec2011.pdf.]

——, and ——, 2015: Extended range forecast of Atlantic seasonal hurricane activity and landfall strike probability for 2015. Department of Atmospheric Science, Colorado State University, 38 pp. [Available online at http://tropical.atmos.colostate.edu/forecasts/2015/june2015/jun2015.pdf.]

Knapp, K. R., M. C. Kruk, D. H. Levinson, H. J. Diamond, and C. J. Neuman, 2010: The international best track archive for climate stewardship (IBTrACS): Unifying tropical cyclone best track data. *Bull. Amer. Meteor. Soc.*, **91**, 363–376, doi:10.1175/2009BAMS2755.1.

Knutson, T. R., and Coauthors, 2013: Dynamical downscaling projections of twenty-first-century Atlantic hurricane activity: CMIP3 and CMIP5 model-based scenarios. *J. Climate*, **26**, 6591–6617, doi:10.1175/JCLI-D-12-00539.1.

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

——, K. A. Emanuel, and G. A. Vecchi, 2014: The poleward migration of the location of tropical cyclone maximum intensity. *Nature*, **509**, 349–352, doi:10.1038/nature13278.

——, T. R. Karl, T. R. Knutson, K. A. Emanuel, K. E. Kunkel, and J. J. O'Brien, 2015: Reply to 'Comment on ''Monitoring and understanding trends in extreme storms: State of knowledge.''' *Bull. Amer. Meteor. Soc.*, **96**, 1177–1179, doi:10.1175/BAMS-D-14-00261.1.

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, doi:10.1175/MWR-D-12-00254.1.

LaRow, T. E., Y.-K. Lim, D. W. Shin, E. P. Chassgnet, and S. Cocke, 2008: Atlantic basin seasonal hurricane simulations. *J. Climate*, **21**, 3191–3206, doi:10.1175/2007JCLI2036.1.

——, L. Stefanova, D.-W. Shin, and S. Cocke, 2010: Seasonal Atlantic tropical cyclone hindcasting/forecasting using two sea surface temperature datasets. *Geophys. Res. Lett.*, **37**, L02804, doi:10.1029/2009GL041459.

Latif, M., N. Keenlyside, and J. Bader, 2007: Tropical sea surface temperature, vertical wind shear, and hurricane development. *Geophys. Res. Lett.*, **34**, L01710, doi:10.1029/2006GL027969.

Lehmiller, G. S., T. B. Kimberlain, and J. B. Elsner, 1997: Seasonal prediction models for North Atlantic basin hurricane location. *Mon. Wea. Rev.*, **125**, 1780–1791, doi:10.1175/1520-0493(1997)125<1780:SPMFNA>2.0.CO;2.

Li, X., S. Yang, H. Wang, X. Jia, and A. Kumar, 2013: A dynamical-statistical forecast model for the annual frequency of western Pacific tropical cyclones based on the NCEP Climate Forecast System version 2. *J. Geophys. Res. Atmos.*, **118**, 12 061–12 074, doi:10.1002/2013JD020708.

Murakami, H., R. Mizuta, and E. Shindo, 2012: Future changes in tropical cyclone activity projected by multi-physics and multi-SST ensemble experiments using the 60-km-mesh MRI-AGCM. *Climate Dyn.*, **39**, 2569–2584, doi:10.1007/s00382-011-1223-x.

——, and Coauthors, 2015: Simulation and prediction of category 4 and 5 hurricanes in the high-resolution GFDL HiFLOR coupled climate model. *J. Climate*, **28**, 9058–9079, doi:10.1175/JCLI-D-15-0216.1.

Nakano, M., M. Sawada, T. Nasuno, and M. Satoh, 2015: Intraseasonal variability and tropical cyclogenesis in the western North Pacific simulated by a global nonhydrostatic atmospheric model. *Geophys. Res. Lett.*, **42**, 565–571, doi:10.1002/2014GL062479.

Pielke, R. A., J. Gratz, C. W. Landsea, D. Collins, M. A. Saunders, and R. Musulin, 2008: Normalized hurricane damage in the United States: 1990–2005. *Nat. Hazards Rev.*, **9**, 29–42, doi:10.1061/(ASCE)1527-6988(2008)9:1(29).

Portis, D. H., J. E. Walsh, M. El Hamly, and P. J. Lamb, 2001: Seasonality of the North Atlantic Oscillation. *J. Climate*, **14**, 2069–2078, doi:10.1175/1520-0442(2001)014<2069:SOTNAO>2.0.CO;2.

Ramsay, H. A., and A. H. Sobel, 2011: Effects of relative and absolute sea surface temperature on tropical cyclone potential intensity using a single-column model. *J. Climate*, **24**, 183–193, doi:10.1175/2010JCLI3690.1.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, and D. P. Rowell, 2003: Global analysis of sea surface temperature, sea ice, and night marine air temperature

since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, doi:10.1029/2002JD002670.

Saunders, M. A., and A. S. Lea, 2005: Seasonal prediction of hurricane activity reaching the coast of the United States. *Nature*, **434**, 1005–1008, doi:10.1038/nature03454.

Scaife, A. A., J. R. Knight, G. K. Vallis, and C. K. Folland, 2005: A stratospheric influence on the winter NAO and North Atlantic surface climate. *Geophys. Res. Lett.*, **32**, L18715, doi:10.1029/2005GL023226.

——, and Coauthors, 2014: Skillful long-range prediction of European and North American winters. *Geophys. Res. Lett.*, **41**, 2514–2519, doi:10.1002/2014GL059637.

Servain, J., I. Wainer, J. P. McCreary, and A. Dessier, 1999: Relationship between the equatorial and meridional modes of climatic variability in the tropical Atlantic. *Geophys. Res. Lett.*, **26**, 485–488, doi:10.1029/1999GL900014.

Smith, A. B., and R. W. Katz, 2013: US billion-dollar weather and climate disasters: Data sources, trends, accuracy and biases. *Nat. Hazards*, **67**, 387–410, doi:10.1007/s11069-013-0566-5.

Swanson, K. L., 2008: Nonlocality of Atlantic tropical cyclone intensities. *Geochem. Geophys. Geosyst.*, **9**, Q04V01, doi:10.1029/2007GC001844.

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

——, K. L. Swanson, and B. J. Soden, 2008: Whither hurricane activity? *Science*, **322**, 687–689, doi:10.1126/science.1164396.

——, M. Zhao, H. Wang, G. Villarini, A. Rosati, A. Kumar, I. M. Held, and R. Gudgel, 2011: Statistical–dynamical predictions of seasonal North Atlantic hurricane activity. *Mon. Wea. Rev.*, **139**, 1070–1082, doi:10.1175/2010MWR3499.1.

——, and Coauthors, 2013: Multiyear predictions of North Atlantic hurricane frequency: Promise and limitations. *J. Climate*, **26**, 7994–8016, doi:10.1175/JCLI-D-12-00464.1.

——, and Coauthors, 2014: On the seasonal forecasting of regional tropical cyclone activity. *J. Climate*, **27**, 7994–8016, doi:10.1175/JCLI-D-14-00158.1.

Villarini, G., and G. A. Vecchi, 2012: North Atlantic power dissipation index (PDI) and accumulated cyclone energy (ACE): Statistical modelling and sensitivity to sea surface temperature. *J. Climate*, **25**, 625–637, doi:10.1175/JCLI-D-11-00146.1.

——, and ——, 2013: Multiseason lead forecast of the North Atlantic power dissipation index (PDI) and accumulated cyclone energy (ACE). *J. Climate*, **26**, 3631–3643, doi:10.1175/JCLI-D-12-00448.1.

——, ——, and J. A. Smith, 2010: Modeling of the dependence of tropical storm counts in the North Atlantic basin on climate indices. *Mon. Wea. Rev.*, **138**, 2681–2705, doi:10.1175/2010MWR3315.1.

——, ——, T. R. Knutson, M. Zhao, and J. A. Smith, 2011: North Atlantic tropical storm frequency response to anthropogenic forcing: Projections and sources of uncertainty. *J. Climate*, **24**, 3224–3238, doi:10.1175/2011JCLI3853.1.

——, ——, and J. A. Smith, 2012: U.S. landfalling and North Atlantic hurricanes: Statistical modelling of their frequencies and ratios. *Mon. Wea. Rev.*, **140**, 44–65, doi:10.1175/MWR-D-11-00063.1.

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

Vitart, F., 2006: Seasonal forecasting of tropical storm frequency using a multi-model ensemble. *Quart. J. Roy. Meteor. Soc.*, **132**, 647–666, doi:10.1256/qj.05.65.

——, and T. N. Stockdale, 2001: Seasonal forecasting of tropical storms using coupled GCM integrations. *Mon. Wea. Rev.*, **129**, 2521–2537, doi:10.1175/1520-0493(2001)129<2521:SFOTSU>2.0.CO;2.

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

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

Wittenberg, A. T., A. Rosati, N.-C. Lau, and J. J. Ploshay, 2006: GFDL's CM2 global coupled climate models. Part III: Tropical pacific climate and ENSO. *J. Climate*, **19**, 698–722, doi:10.1175/JCLI3631.1.

Xiang, B., S.-J. Lin, M. Zhao, S. Zhang, G. Vecchi, T. Li, X. Jiang, L. Harris, and J.-H. Chen, 2015a: Beyond weather time scale prediction for Hurricane Sandy and Super Typhoon Haiyan in a global climate model. *Mon. Wea. Rev.*, **143**, 524–535, doi: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 prediction skill in a GFDL coupled model. *J. Climate*, **28**, 5351–5364, doi:10.1175/JCLI-D-15-0102.1.

Xie, S.-P., and J. A. Carton, 2004: Tropical Atlantic variability: Patterns, mechanisms, and impacts. *Earth's Climate: The Ocean–Atmosphere Interaction*, *Geophys. Monogr.*, Vol. 147, Amer. Geophy. Union, 121–142.

Yan, T., L. J. Pietrafesa, D. A. Dickey, P. T. Gayes, and S. Bao, 2015: Seasonal prediction of landfalling hurricanes along eastern seaboard of the United States. *Int. J. Climatol.*, **35**, 2647–2653, doi:10.1002/joc.4163.

Zhang, S., and A. Rosati, 2010: An inflated ensemble filter for ocean data assimilation with a biased coupled GCM. *Mon. Wea. Rev.*, **138**, 3905–3931, doi:10.1175/2010MWR3326.1.

Zhao, M., I. M. Held, and G. A. Vecchi, 2010: Forecasts of the hurricane season using a global atmospheric model assuming persistence of SST anomalies. *Mon. Wea. Rev.*, **138**, 3858–3868, doi:10.1175/2010MWR3366.1.