Retrospective Forecasts of the Hurricane Season Using a Global Atmospheric Model Assuming Persistence of SST Anomalies

MING ZHAO

University Corporation for Atmospheric Research, and NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

ISAAC M. HELD AND GABRIEL A. VECCHI

NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

(Manuscript received 2 February 2010, in final form 9 April 2010)

ABSTRACT

Retrospective predictions of seasonal hurricane activity in the Atlantic and east Pacific are generated using an atmospheric model with 50-km horizontal resolution by simply persisting sea surface temperature (SST) anomalies from June through the hurricane season. Using an ensemble of 5 realizations for each year between 1982 and 2008, the correlations of the model mean predictions with observations of basin-wide hurricane frequency are 0.69 in the North Atlantic and 0.58 in the east Pacific. In the North Atlantic, a significant part of the degradation in skill as compared to a model forced with observed SSTs during the hurricane season (correlation of 0.78) can be explained by the change from June through the hurricane season in one parameter, the difference between the SST in the main development region and the tropical mean SST. In fact, simple linear regression models with this one predictor perform nearly as well as the full dynamical model for basin-wide hurricane frequency in both the east Pacific and the North Atlantic. The implication is that the quality of seasonal forecasts based on a coupled atmosphere–ocean model will depend in large part on the model's ability to predict the evolution of this difference between main development region SST and tropical mean SST.

1. Introduction

Predictions of the statistics of storm activity in the upcoming hurricane season have a long history (see Camargo et al. 2007b for a review). Until recently, these predictions have been based on statistical models, using the state of the atmosphere and oceans in some period before the start of the hurricane season as predictors (e.g., Gray 1984b; Elsner and Jagger 2006; Klotzbach and Gray 2009). In recent years, dynamical atmospheric and coupled atmosphere–ocean models have begun to be applied to this problem. These models can be used to predict indices such as vertical shear, low-level vorticity, midtropospheric relative humidity, or potential intensity, which can then be related statistically to storm genesis (e.g., Wang et al. 2009). Or, more ambitiously, the storms developed by the atmospheric models can be used directly (e.g., Vitart 2006; Vitart et al. 2007; Camargo and Barnston 2009; LaRow et al. 2010). The latter approach obviously becomes more appealing as these dynamical models move to finer horizontal resolution and their simulations of storm statistics become more credible. This paper takes the latter approach, using a global atmospheric model with 50-km horizontal resolution, but we also use the results to justify a simple statistical model.

The climatological statistics of tropical storms forming in our dynamical model, as well as the interannual variability in the model when run over observed sea surface temperatures (SSTs), have been described in Zhao et al. (2009, hereafter ZHLV). The quality of the simulation has encouraged us to use this model as a tool to explore issues related to predictions of seasonal storm statistics. We do not describe coupled model simulations here, but use the atmospheric model in isolation to estimate bounds on the likely quality of future coupled model predictions. The simulations in ZHLV are not predictions since they assume knowledge of the observed SSTs

Corresponding author address: Dr. Ming Zhao, NOAA/Geophysical Fluid Dynamics Laboratory, Princeton University, Forrestal Campus 201, Forrestal Rd., Princeton, NJ 08540-6649.
E-mail: ming.zhao@noaa.gov

DOI: 10.1175/2010MWR3366.1

throughout the hurricane season, but they do provide an estimate of an upper bound to the quality of forecasts with coupled systems based on this atmospheric model. Here we consider the opposite limit, constructing a lower bound on the quality of attainable predictions, by taking the SST anomalies observed prior to the hurricane season and assuming that these anomalies persist through the season. Any prediction that can improve on this persistence forecast for SST should improve on the skill of these predictions. A possible caveat arises from the possibility that the uncoupled system might not mimic certain aspects of the hurricane-scale fully coupled dynamics even if given correct large-scale SSTs. The high correlations, at least in the North Atlantic and east Pacific, in ZHLV and in analogous studies with other atmospheric models, such as LaRow et al. (2008), provide justification for this approach.

By comparing simulations forced by persisted SST anomalies with the more accurate simulations using observed SSTs, one can hope to understand which features in the evolving SST patterns result in the loss of skill. This would help efforts in improving coupled model predictions by focusing attention on those aspects of the SST field that matter for this application. It also motivates the construction of statistical models that focus on the indices that explain the loss in skill.

We limit the discussion that follows to the North Atlantic and eastern Pacific. As discussed in ZHLV, these are the basins for which the interannual variations in the model's storm frequency correlate best with observations. This is not necessarily because the simulations are of higher quality in these basins than elsewhere, but at least in part because signals associated with interannual variations in SST stand out from the noise most clearly in these basins. The prediction integrations are designed to capture the bulk of the storm season in these two basins.

2. Hindcasts of North Atlantic and east Pacific storm frequency

We refer to the model here as C180HIRAM2.1 [i.e., the High-Resolution Atmospheric Model (HIRAM); C180 refers to a model with a cubed-sphere dynamical core (Putman and Lin 2007) with 180×180 grid points on each face of the cube, resulting in grid sizes ranging from 43.5 to 61.6 km]. In addition to the cubed-sphere dynamical core and resolution, the model differs from the Geophysical Fluid Dynamics Laboratory (GFDL) Atmospheric Model version 2.0 (AM2) in its convection and cloud schemes, as described in ZHLV. The source of the observed SSTs is the Met Office Hadley Centre Sea Ice and SST model version 1.1 (HadISST 1.1; Rayner et al. 2003), while observed storm statistics are taken from the International Best Track Archive for Climate Stewardship (IBTrACS, available online at http://www.ncdc.noaa.gov/oa/ibtracs/; Kruk et al. 2010). In the experiments described in ZHLV, the atmospheric model is used to generate four realizations of the period 1981–2005, running continuously for the full period over observed SSTs. For this paper, we have continued these four integrations up to and including 2008. Below we refer to these observed SST experiments as the Atmospheric Model Intercomparison Project (AMIP) experiments.

For our hindcast experiments, we compute the global SST anomaly field for the month of June in each year from 1982 to 2008 using the mean of the years 1982–2005 to define the anomaly. (The shorter period is used to define anomalies because the years 2006–08 were added at a late point in the execution of these experiments.) For this analysis, we start from 1982 because this is the first full year in which satellite retrievals of SST are regularly incorporated into the HadiSST product (Rayner et al. 2003), a potential source of inhomogeneity in SST observations between 1981 and 1982. We integrate from 1 June to the end of December for each year after adding this anomaly to the climatological mean SST for each month. The initial states of the atmosphere and the land are not estimated from observations, they are simply taken from one of the AMIP experiments. We run five realizations for each year by perturbing the initial state. The atmospheric state diverges within a few days among different realizations and gradually evolves into completely different solutions after a couple of weeks. We refer to these persisted June SST anomaly experiments as the forecast (FCST) experiments. For both the AMIP and the FCST experiments we analyze the basin-wide annual hurricane count for the months from July to December since the FCST experiments can only be considered retrospective forecasts for this period. The storm-tracking algorithm used is described in appendix B of ZHLV. Our focus on hurricane counts is motivated in part by the suggestion that counts of weaker tropical storms are more sensitive to the details of the tracking algorithm, as described in ZHLV.

Figure 1a shows the time series of annual (July–December) Atlantic hurricane counts from both the FCST experiments and the observations. For comparison, the results from the AMIP experiments are shown in Fig. 1b. For both the FCST and the AMIP experiments, the number of hurricane counts has been normalized by multiplying by a constant factor (1.17 for AMIP and 1.06 for FCST) so that each ensemble mean time-mean reproduces the observed time-mean count. Between the execution of the AMIP and FCST experiments, the dynamical core of the model was updated slightly to improve efficiency and stability. Preliminary

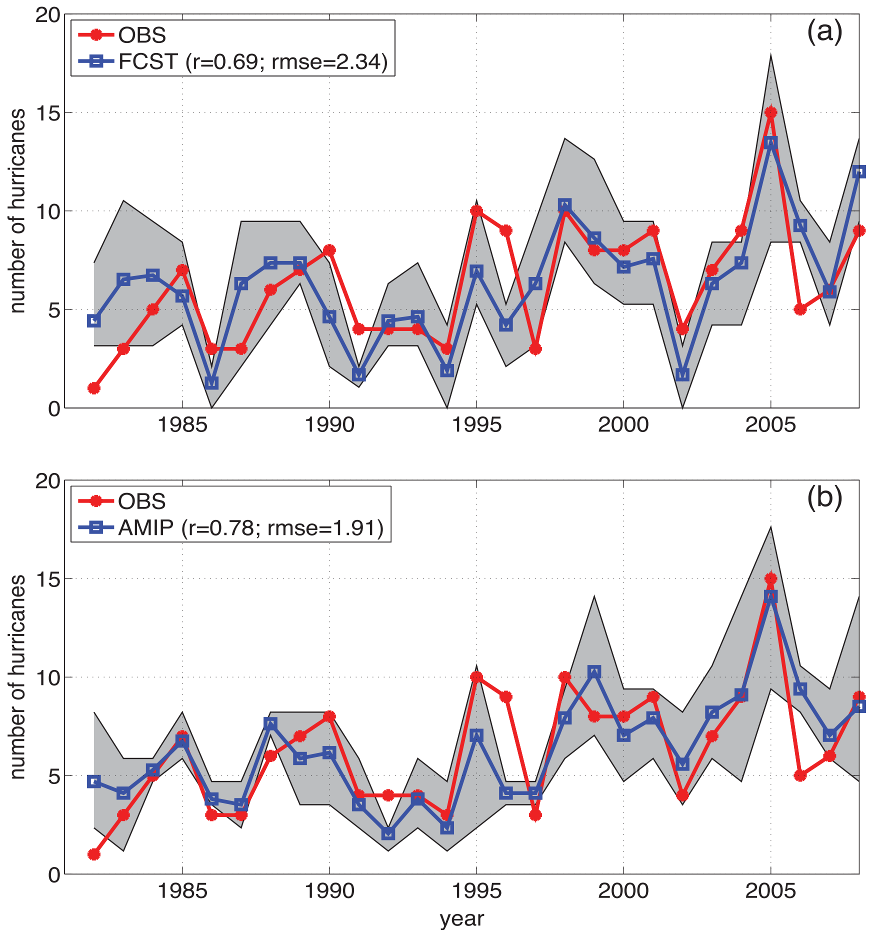

FIG. 1. (a) North Atlantic July–December hurricane counts for each year for the period of 1982–2008. IBTrACS observations (red circles), five-member ensemble mean from the FCST experiments (blue squares), and the maximum and minimum number for each year from the five-member integrations (shaded area). (b) As in (a), but for the AMIP experiments with four-member ensemble. Model time series are normalized as described in text.

tests, focusing on the Atlantic, suggested that this difference was of negligible importance for this study, but after all realizations were completed it became evident that there was a small difference in the total storm activity, which was more significant in the east Pacific. (The difference is less than 1 hurricane per year in the North Atlantic and between 1 and 2 in the east Pacific.) It is for this reason that we normalize the AMIP and FCST ensembles separately. The normalization will not influence the correlations between the model and the observations, but will reduce the model's RMSE slightly (while we have computed this bias correction with the full FCST ensemble, excluding the prediction year from this correction makes little difference).

Both the ensemble mean and the full range over the ensemble are shown in Fig. 1. The correlation between the ensemble mean hurricane counts and the observations is 0.78 for the AMIP experiments. This is slightly lower than the value in ZHLV (0.83), due to the addition of the years 2006–08. The reduction in correlation is entirely due to 2006, for which all members of the ensemble predict more hurricanes than were observed. From the AMIP to the FCST experiments we see a reduction in skill that can be measured by an increase in the RMS count error from 1.91 to 2.34 and the drop in the correlation from 0.78 to 0.69.

While one expects a few observations to lie outside of the range spanned by all 9 of these realizations, 2006 and 1996 seem to stand out as years for which this atmospheric model is least successful, with the former year being too active in the model and the latter too quiescent. In contrast, neither the AMIP nor the FCST ensemble spreads have difficulty in encompassing the peak activity in 2005.

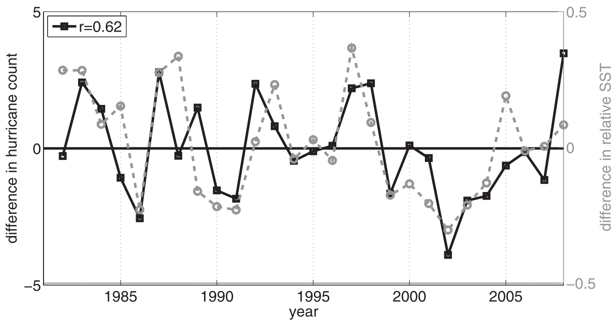

FIG. 2. Time series of the difference in yearly (July–December) Atlantic hurricane counts between the FCST and the AMIP ensemble means shown in Fig. 1 (solid line). The time series of the difference between yearly Atlantic June and ASO relative SST anomalies (dashed line; see text for the definition of Atlantic relative SST).

To understand the reduction in skill from the AMIP to the FCST experiments, we characterize the difference in SSTs between the two sets of experiments by an index based on results of Swanson (2008), Vecchi et al. (2008), Knutson et al. (2008), and ZHLV. The index is defined as the difference between the SSTs averaged over the North Atlantic main development region (MDR; 10°–25°N, 80°–20°W) and the SSTs averaged over the entire tropics (30°S–30°N). This index is referred to below as the Atlantic relative SST, or Φ_{NA}. ZHLV (see their Fig. 16) shows that Φ_{NA} anomaly for the August–September–October (ASO) hurricane season explains much of the interannual variance in hurricane counts in the AMIP experiments as well as the spread in global warming simulations using the twenty-first-century SST anomalies projected by different global coupled models. Figure 2 shows that the difference between the value of Φ_{NA} anomaly for June and the value averaged over ASO explains much of the difference in ensemble mean Atlantic hurricane counts between the FCST and the AMIP experiments. The correlation between the difference in hurricane counts and the difference in Φ_{NA} anomaly is 0.62 [significant at $p = 0.01$, estimating the effective degrees of freedom (Wilks 2006, p. 144) from the autocorrelation of 0.3 for Φ_{NA} at a 1-yr lag]. This correlation indicates that a significant fraction of the deterioration in skill from AMIP to FCST can be thought of as due to the errors made in the value of the Φ_{NA} anomaly by using persistence of the June SST anomaly. This figure also provides evidence that the small difference in the dynamical core used in the AMIP and FCST runs is not responsible for the bulk of the reduction in correlation, since this reduction can be explained in large part by the Φ_{NA} index.

One can check whether this correlation between the difference in Φ_{NA} anomaly and the difference in storm counts in Fig. 2 is as large as could be expected given the noise in the model. As discussed in ZHLV, the standard deviation in Atlantic hurricane count across an ensemble with identical SSTs in this model is roughly $0.5 + 0.2N$, where N is the mean hurricane count for a given year. Using this estimate, assuming Gaussian statistics, one can estimate the expected value of the correlation between the two curves in Fig. 2 if the difference in Φ_{NA} anomaly is the only relevant difference in the SST fields. The expected correlation is 0.78 with a standard deviation of 0.07, significantly larger than the value of 0.62 obtained above. The implication is that, while the change in Φ_{NA} anomaly is responsible for a large part of the signal, the model's Atlantic hurricane activity is also responding to other aspects of the change in the SST anomaly field over the prediction period.

A hint as to the source of this information in the SST field is provided by the correlation map between the basin-averaged North Atlantic hurricane count difference between the FCST and AMIP experiments and the local change in relative SST anomaly from June to the ASO mean. Figure 3a shows that besides the large values in the Atlantic MDR itself, there are also significant correlations in the eastern and western equatorial Pacific. We can first regress the difference in Atlantic hurricane count to the difference in Φ_{NA} anomaly to obtain a linear relationship and then use it to obtain a residual count difference, which is not explained by the difference in Φ_{NA} anomaly. Figure 3b shows a correlation map between the residual count difference and local change of relative SST. Evidently the effect of eastern and western equatorial Pacific SST cannot be completely captured by the change in the Φ_{NA} index. Indeed, in a multiple linear regression calculation, the correlation coefficient increases from 0.62 to 0.72 if we take into account both the Atlantic MDR relative SST and the eastern Pacific relative SST (see Fig. 3). The correlation further increases to 0.79 (a value as large as could be expected given the noise level of the model and the size of the ensembles) when the western equatorial Pacific region (see Fig. 3) is also included in the multiregression model.

The relative SSTs in the two identified Pacific regions display significant correlation to El Niño–Southern Oscillation (ENSO) indices; in particular, the correlation with the Niño-3.4 index is 0.54 for the eastern Pacific region and −0.78 for the western Pacific region. ENSO has long been recognized to be an important factor in impacting the hurricane activities in the North Atlantic basin (e.g., Gray 1984a; Goldenberg and Shapiro 1996; Camargo et al. 2007a). Since our AMIP experiments use observed SSTs, the ENSO state is already accounted for through the boundary forcing. The correlation between

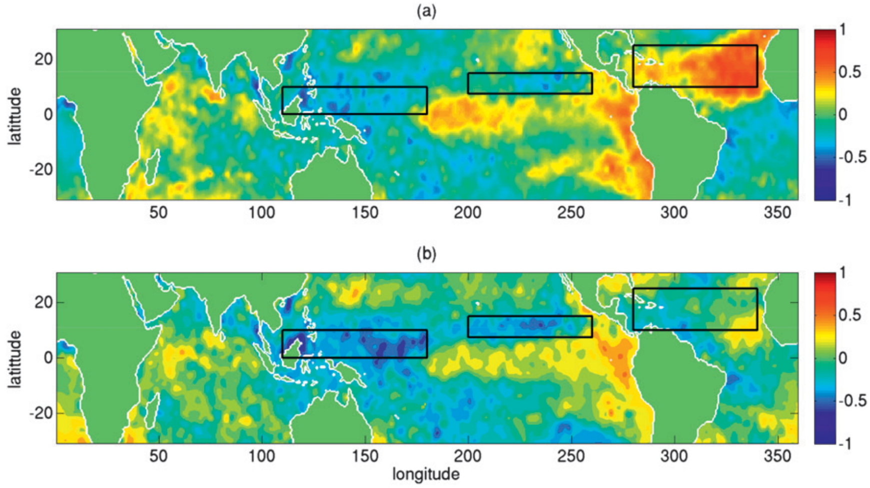

FIG. 3. (a) Correlation map between the local change in relative SST anomaly from June to the ASO mean and the basin-averaged North Atlantic hurricane count difference between the FCST and AMIP experiments. (b) As in (a), but the total North Atlantic hurricane count difference is replaced with a residual count difference that is not explained by the difference in Atlantic MDR relative SST anomaly (see text for the details). The three boxes show the North Atlantic MDR (10°–25°N, 80°–20°W), the east Pacific region (7.5°–15°N, 160°–100°W), and the west Pacific region (0°–10°N, 110°E–180°), which explain most of the difference in North Atlantic hurricane count between the FCST and AMIP experiments.

Niño-3.4 SST index and the AMIP-simulated hurricane count is -0.4 for the 1982–2008 period, slightly lower than that for the observations (-0.5). This level of correlation is however much lower than that between Φ_{NA} and the hurricane count for both the model and the observations.

If one uses the difference in the Niño-3.4 index instead of Φ_{NA} to explain the deterioration in skill from AMIP to FCST, the correlation with the difference in simulated hurricane counts is only 0.27. Using both Φ_{NA} and Niño-3.4 index in a multiple regression increases the correlation from 0.62 for Φ_{NA} alone to only 0.65. This is in large part due to the high correlation from June to the ASO mean in Niño-3.4 (>0.8), since the "spring barrier" has already been passed by June. Nevertheless, as we will discuss in section 3, the Φ_{NA} index alone does not capture the full effect of ENSO on hurricane variability. Furthermore, the persisted June SST anomaly also tends to underestimate the ENSO signal, since ENSO events tend to build over boreal summer and fall (Harrison and Larkin 1998). We leave a more systematic exploration of how the changes in SST anomalies over the forecast interval affect the skill of the FCST ensemble for future work.

Figure 4 shows similar plots as in Fig. 1, but for the east Pacific. As in Fig. 1, both the FCST and the AMIP hurricane counts have been separately normalized by multiplying by constant factors (1.39 for AMIP and 1.11 for FCST) to correct the model bias. Because the simulated storms in the east Pacific tend to be of smaller scale than those in the North Atlantic, they appear to be more sensitive to the minor changes in the dynamical core referred to above. The AMIP normalization factor is larger than that used in ZHLV (1.25) for the entire season (April–December) since the AMIP experiments tend to underproduce east Pacific hurricanes for July–September while slightly overproducing in May and June, as shown in Fig. 5 in ZHLV. Overall, the correlation between the ensemble mean east Pacific hurricane counts and the observations is 0.58 for the FCST experiments with a RMS count error of about 3 hurricanes. This indicates a relatively small drop of correlation compared to the AMIP experiments (correlation = 0.65), with the RMS count error being slightly smaller than in AMIP.

Similar to the procedure used in the North Atlantic, we can define a relative SST index for the east Pacific (Φ_{EP}). We choose the east Pacific MDR region as (7.5°–15°N, 160°–80°W). As discussed below, Φ_{EP} is itself a useful predictor of storm count in the east Pacific. However, the persisted June value of Φ_{EP} anomaly is very strongly correlated with the ASO mean (correlation >0.9), so there is little advantage in predicting the evolution of this index over the forecast period, in distinction to the Atlantic case. Consistently, the differences in the simulations when using observed and persisted anomalies are well within the model's noise level.

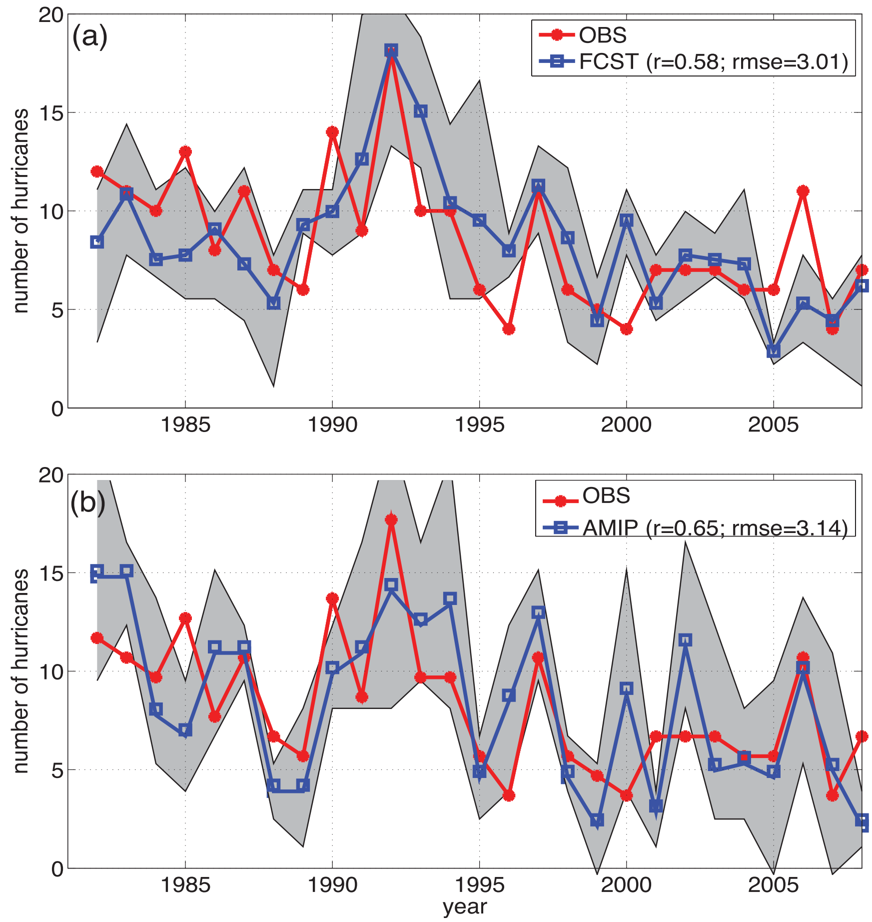

FIG. 4. As in Fig. 1, but for the east Pacific.

3. A simple statistical model for basin-wide hurricane frequency

The relative SST anomalies have a large controlling effect on the interannual variation of basin-wide hurricane activity as can be seen by a simple linear regression. Figure 5 shows scatterplots of yearly (July–December) hurricane counts versus yearly ASO Φ_{NA} and Φ_{EP} anomalies for the period of 1982–2008 from the combined AMIP and FCST experiments and the observations. In the North Atlantic, Φ_{NA} anomaly explains well over 50% of the variance in hurricane counts for both the models and the observations (correlations are 0.76 for the observations and 0.83 for the model ensemble mean; separation between AMIP and FCST shows little difference in both correlation and slope). The linear regression coefficient gives 8.8 (7.2–10.4) and 9.6 (6.3–13) (with the quantities in parentheses being the 95% confidence intervals), respectively, hurricanes per year per degree change of Φ_{NA} for the model and the observations. These values are slightly larger than that in ZHLV (7.8 hurricanes per year per kelvin), where results of both the AMIP and climate change experiments were considered together in the regression. Figure 5b shows a similar plot for the east Pacific, where there is also a clear positive correlation although with a lower coefficient (0.71 for the model and 0.67 for the observations). The linear regression coefficient gives respectively 10.9 (8–13.9) and 9.8 (5.4–14.10) hurricanes per year per degree change of Φ_{EP} for the model and the observations.

Overall, the model reproduces well the observed relationship between the basin's hurricane frequency and the basin relative SSTs for both the North Atlantic and the east Pacific. Interestingly, both ocean basins present similar dependence of hurricane frequency on their corresponding relative SSTs (roughly 1 hurricane per year for each 0.1 degree change of relative SSTs) despite their different climatological values. The large correlation between relative SST anomaly and the model hurricane frequency indicates that a simple linear regression model with Φ_{NA} (Φ_{EP}) anomaly as a single predictor can be useful to explain and predict the modeled North Atlantic (east Pacific) hurricane count.

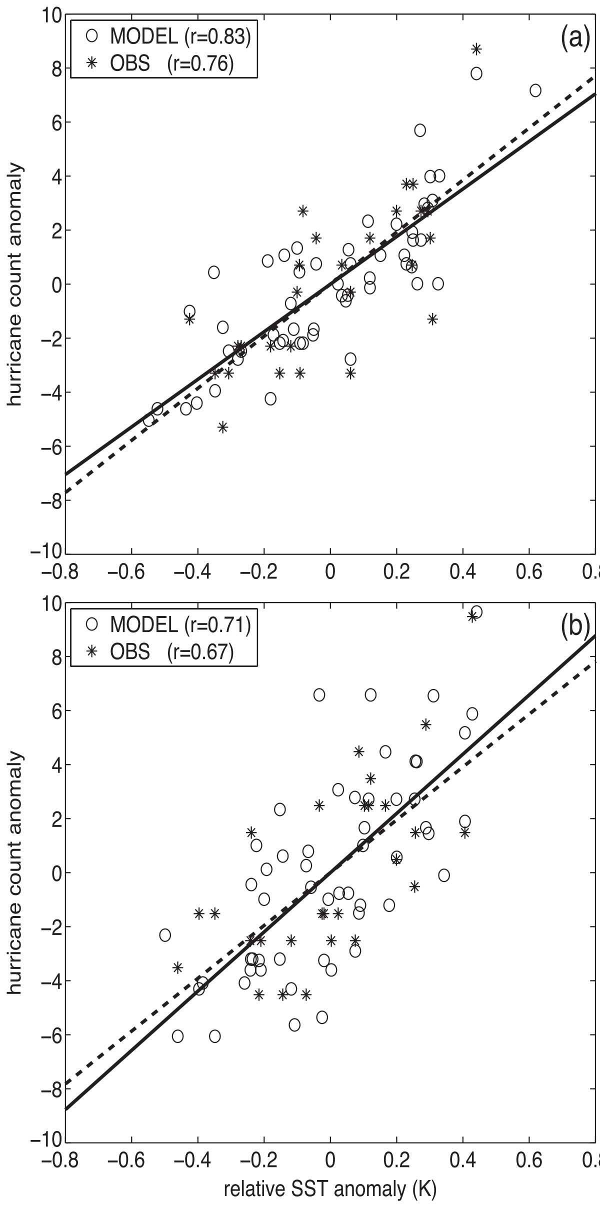

FIG. 5. (a) Scatterplot of the North Atlantic annual (July–December) hurricane count anomalies vs its ASO season MDR relative SST anomalies for each year for the period of 1982–2008. AMIP and FCST experiments (normalized) (circles) and observations (asterisks). Linear regression for the AMIP and FCST (solid lines) and observation (dashed lines). Correlations are shown in the legend. (b) As in (a), but for the east Pacific.

We first test a simple linear regression model obtained from Fig. 5 in reproducing both the North Atlantic and the east Pacific hurricane frequency. Since our dynamical model is not tuned to optimize the interannual variations in hurricanes, we choose regression slope based on the dynamical model results, that is 8.8 hurricanes per year per kelvin for the North Atlantic and 10.9 hurricanes per year per kelvin for the east Pacific. As expected, the linear model using the observed ASO relative SST anomaly produces time series of hurricane count about as well as the AMIP ensemble mean, with correlation = 0.76 and RMSE = 2 hurricanes for the North Atlantic and correlation = 0.67 and RMSE = 2.5 for the east Pacific.

Similar to the dynamical model using persisted June SST anomalies, we can use the June Φ_{NA} and Φ_{EP} anomalies in the linear regression model to predict the storm frequency in each hurricane season. Figure 6 shows how this simple linear prediction model compares with the dynamical model prediction (FCST). For the North Atlantic, the linear model produces a slightly lower skill than the FCST. The correlation is reduced from 0.69 to 0.55 and the RMS count error increases from 2.34 to 2.64 hurricanes. Much of the skill loss is due to several years prior to 1988. For the east Pacific, the linear prediction model performs slightly better than the FCST ensemble mean with slightly larger correlation (0.62) and smaller RMS count error (2.67 hurricanes).

Despite the comparable overall skill for the simple linear regression model, both the ENSO years and the very active year of 2005 provides evidence that the dynamical model is able to recover more information from the SST field than the simple relative SST index. Figure 7 shows a comparison of the North Atlantic yearly (July–December) hurricane counts averaged for the El Niño years (1982, 1987, 1991, 1992, 1994, 1997, and 2002) and the La Niña years (1988, 1998, 1999, 2000, 2001, and 2007) from the AMIP, the FCST, and the linear regression model using both the observed ASO and the persisted June relative SST anomaly. Evidently, both the AMIP and the FCST experiments produce a better contrast between the La Niña and El Niño years while their corresponding linear model predictions on average significantly underestimate the difference. An examination of the variation of Φ_{NA} with ENSO indicates that it is consistent with the suppression–enhancement of hurricane activity with El Niño–La Niña (not shown). Nevertheless, Fig. 7 also suggests clearly that this Φ_{NA} index alone does not capture the full effect of ENSO on storm variability.

Figure 8 shows a comparison of 2005 July–December hurricane counts from the AMIP and FCST ensembles and the linear regression model based on the observed ASO and persisted June relative SST anomalies. Both AMIP and FCST experiments produce more hurricanes on average than the linear model, although there is substantial spread among individual runs. The linear model significantly underestimates the number of hurricanes for 2005. A check for the spatial distribution of relative SST for 2005 shows a relative cooling of the east Pacific

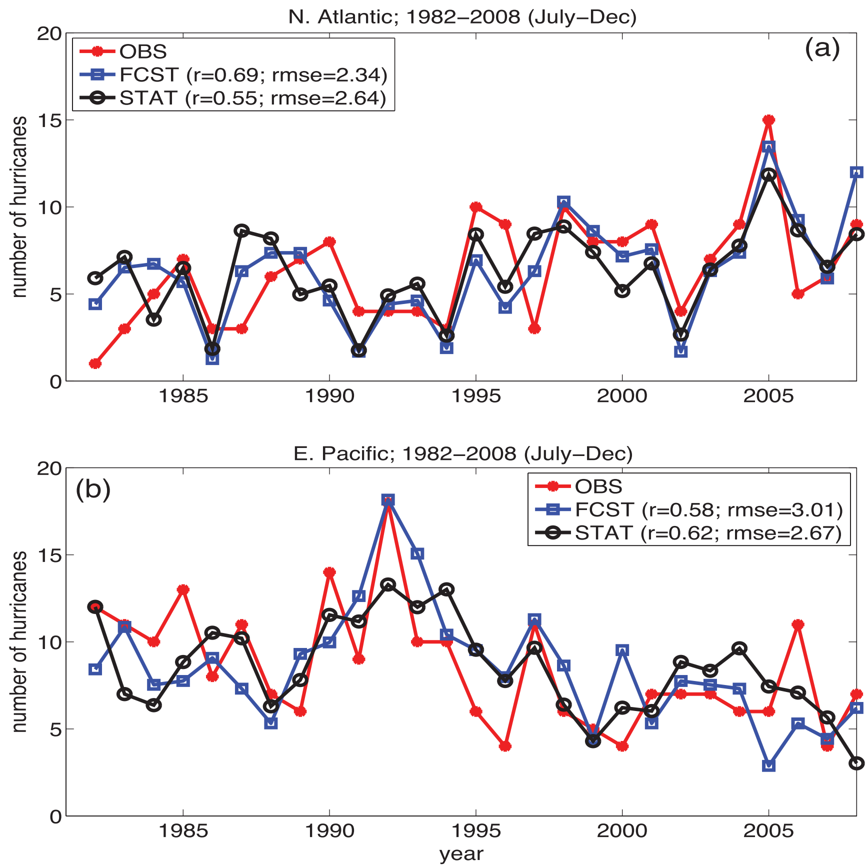

FIG. 6. A comparison of the linear prediction model (black circles) using June relative SST anomaly with the FCST ensemble mean (blue squares) and the observations (red circles) for (top) North Atlantic and (bottom) east Pacific. Correlations and RMS count errors are shown in the legend.

besides the strong relative warming in the North Atlantic MDR. This is consistent with further increased Atlantic hurricane activity as indicated by the correlation map in Fig. 3. This effect of remote region SST is presumably better captured by the dynamical model than the simple Φ_{NA} index.

It is interesting that this simple linear prediction model, using the persisted relative SST index, performs better for 2005 than the linear model using the actual ASO relative SST index. This is likely due to a cancellation of errors. The linear model underpredicts hurricane count in 2005 when using the observed hurricane season index, but the June value of this index happens to be greater than the ASO value, so that the linear model using the persisted index performs better. Given the large number of storms in this season, it is possible that the evolution of this relative SST index from June through the hurricane season was influenced by the hurricanes themselves, creating some ambiguity in how to interpret prescribed SST models.

To investigate the extent to which we can further extend the lead time for the linear prediction model using persisted SST anomalies, we use Φ_{NA} and Φ_{EP} anomalies for each month prior to the ASO hurricane season for the North Atlantic and east Pacific. Figure 9 shows both the correlation and the RMS count error for the hindcast period of 1982–2008 for a series of leading month predictions. For the North Atlantic, the model maintains significant skill back to April with correlation roughly 0.4 and RMS count error about 3 hurricanes. From April to May, the correlation increases to 0.5. The linear model skill increases continuously with reduced lead time with a strong improvement from June to July. In the east Pacific, there is a sharp improvement of model skill from April to May with correlation increasing from 0.4 to 0.57 and RMS count error decreasing from 3.5 to 2.8 hurricanes. However, after May, the skill of the linear model tends to level off in contrast to the steady improvement in the North Atlantic. These persistence forecasts provide a baseline against which to compare coupled model results.

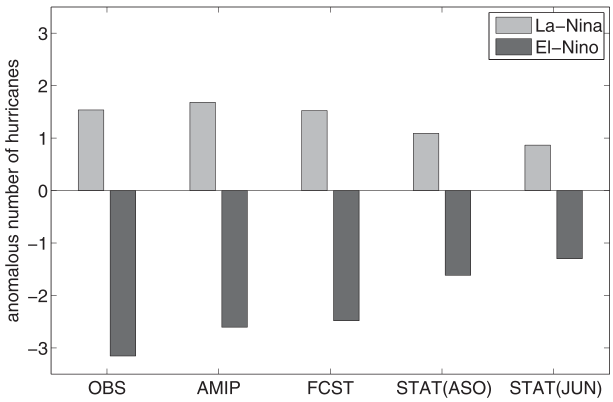

FIG. 7. A comparison of the North Atlantic yearly (July–December) hurricane counts averaged for the El Niño years and the La Niña years from the AMIP, the FCST, and the linear regression model using both the observed ASO [STAT(ASO)] and the persisted June [STAT(JUN)] relative SST anomaly.

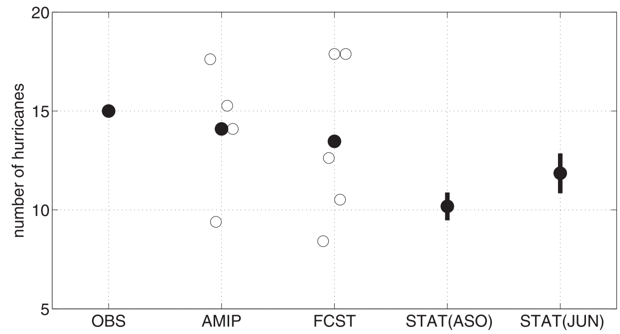

FIG. 8. A comparison of North Atlantic July–December hurricane counts from the AMIP, FCST, and a linear regression model using observed ASO [STAT(ASO)] and persisted June [STAT(JUN)] relative SST anomaly for the year 2005. Hurricane counts for AMIP and FCST are from Fig. 1: individual runs (open symbols) and ensemble mean (filled symbols). Error bars for the linear model shows 95% confidence level for the regression slope.

4. Conclusions

We have documented the level of skill of a global atmospheric model in predicting the hurricane count in the upcoming hurricane season with hindcast experiments (for 1982–2008) in which the SST anomaly is persisted from June. Using 5 realizations with a 50-km resolution model that simulates a fairly accurate global hurricane climatology, we find that the correlation of the ensemble mean (July–December) hurricane count to observations is 0.69 in the North Atlantic and 0.58 in the east Pacific. These correlations provide an estimate of the lower bound on the forecast skill, skill which should be improved by the ability to predict SST anomalies better than the persistence forecasts. Using this model, experiments with observed SSTs suggest that the correlation can be improved to about 0.78 in the North Atlantic with perfect SST prediction, but also that there is little improvement expected in the east Pacific. This distinction is interesting given the anticorrelation on interannual time scales between the two basins (Wang and Lee 2009), and the ability of the model to simulate this correlation, as described in ZHLV.

Despite the simplicity of this approach of persisting June SST anomaly, the level of skill for our dynamical model is generally comparable with those reported in the literature (e.g., Vitart 2006; Vitart et al. 2007; Camargo and Barnston 2009; LaRow et al. 2010). It should be noted, however, that there are many differences between these reports, such as the verification periods, the lead time, the metrics for evaluation, as well as the model calibrations. Furthermore, differences in model resolution, whether or not a multimodel ensemble is used, whether the forecast system is fully coupled with the ocean, and different choice of SSTs in an atmospheric model all play some role in variations of the forecast skill (Camargo et al. 2007b). A systematic evaluation of these models under controlled conditions would be a useful step to providing a formal assessment of the key controls to forecast skill among these models.

In our model, much of the difference in the North Atlantic between the ensemble with persisted SST and that with observed SSTs can be explained by the evolution in one SST index: the relative change of the SSTs in the main development region as compared to the tropical mean SST. While we do not claim that this is the optimal SST index for this purpose, it provides a useful first-order description of differences in the modeled North Atlantic and east Pacific hurricane frequencies. We have not searched for the best index systematically, due in part to the shortness of the time period of the simulations, but have tested this particular index because of arguments for its value in the literature (e.g., Swanson 2008; Vecchi et al. 2008; ZHLV). It is our

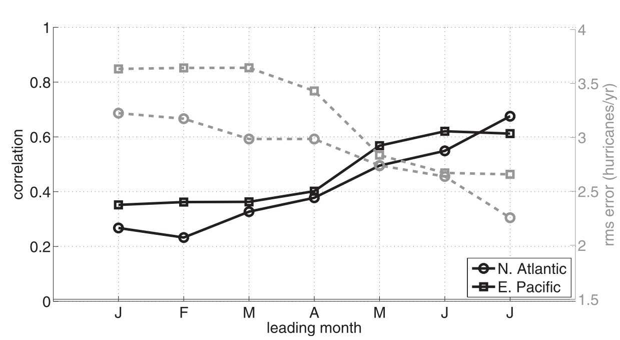

FIG. 9. Correlations (solid) and RMS count errors (dashed) for the linear prediction model using anomalous relative SSTs from each month prior to the ASO hurricane season for the North Atlantic (circles) and the east Pacific (squares).

estimate that, despite the noise level in the model (the spread among ensemble members), there is room for an improved index by taking into account more information from the Pacific or using different statistical model formulations (Villarini et al. 2010).

Even if it does not represent the full effect of the evolution of SST anomalies in our model, we suggest that this is a very useful and simple index to use when evaluating the value of SST forecasts for improving prediction of the seasonal hurricane count. The analogous index in the east Pacific has more persistence from June through the heart of the hurricane season, helping to explain why the forecasts by the dynamical model do not improve significantly when we use observed as opposed to persisted SST anomalies.

We have also demonstrated that the simplest linear model, with hurricane count increasing at roughly 1 storm per 0.1 degree increase in this relative SST index, performs about as well as the full dynamical model for the observed SST and nearly as well for the persisted June SST anomaly experiments. The linear prediction model with persisted SST anomalies generates significant skill back to April in the North Atlantic. Its skill improves continuously with reduced lead time. In contrast, for the east Pacific, the linear prediction model shows a sharp improvement of skill from April to May, with the skill tending to level off after May. While this simplest linear model is clearly imperfect, its skill in predicting basin-wide hurricane frequency appears to be as good as or better than more complex statistical models (e.g., Camargo et al. 2007b). However, to considerably improve the lead time for the hurricane seasonal forecast beyond the short time-scale persistence forecasts described here, a hybrid statistical-dynamical model (e.g., Wang et al. 2009; Vecchi et al. 2010, manuscript submitted to *Mon. Wea. Rev.*) or a dynamical coupled system (e.g., Vitart 2006; Vitart et al. 2007) would be needed.

There are hints that the dynamical model is extracting more useful information from the SST field than just the temperature of the main development region relative to the tropical mean. For example, the dynamical model with both observed and persisted June SST produces a larger contrast of North Atlantic hurricane activity between the La Niña and El Niño years, while the corresponding linear models using the index of relative SST as a single predictor significantly underestimate the difference. Further hints come from the dynamical model's ability to better simulate the very active 2005 season. One also hopes that dynamic models will eventually provide useful skill for intrabasin spatial distributions and intensities, beyond the simple basin wide hurricane counts discussed here. But the first goal in development of dynamical prediction of phenomena is to be competitive with the best statistical forecasts. By capturing the underlying dynamics, rather than putting in by hand the rules that are empirically found to connect hurricanes with climate, we have more confidence that the model results will be transferable to novel situations where our familiar rules might break down. Conversely, our confidence in the intrinsic relevance of predictors used in statistical forecasts is increased by dynamical simulations for which the predictors work just as well as for observations.

Our results strengthen the notion that SSTs alone provide most of the skill in seasonal hurricane forecasts. On interannual time scales, the predictable components of the tropical atmospheric environment directly affecting storm genesis, such as changes in vertical shear, atmospheric stability, and midtropospheric moisture, are likely to be very closely tied to these SSTs (e.g., Vecchi and Soden 2007; Latif et al. 2007; Camargo et al. 2007a; Tang and Neelin 2004; Garner et al. 2009). Furthermore, since climate anomalies in remote oceans, such as those occurring in the Pacific during the ENSO years, can affect both the tropical mean and the tropical Atlantic SST (e.g., Klein et al. 1999; Kushnir et al. 2006), seasonal climate forecasts and observations should be improved throughout the tropics to improve seasonal hurricane forecasts in the Atlantic and east Pacific.

Acknowledgments. This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract DE-AC02-05CH11231. The authors thank A. Wittenberg, S. Garner, T. Knutson, and G. Villarini for valuable discussions. The comments and suggestions from Suzana Camargo, Chris Landsea, and one anonymous reviewer helped improve the manuscript. Ming Zhao was supported in part by NSF Grant ATM-0612551 and in part under Award NA17RJ2612 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce. The findings are those of the authors and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration or the U.S. Department of Commerce.

REFERENCES

Camargo, S., and A. Barnston, 2009: Experimental seasonal dynamical forecasts of tropical cyclone activity at IRI. *Wea. Forecasting,* **24,** 472–491.

——, K. Emanuel, and A. Sobel, 2007a: Use of a genesis potential index to diagnose ENSO effects on tropical cyclone genesis. *J. Climate,* **20,** 4819–4834.

Camargo, S. J., A. G. Barnston, P. Klotzbach, and C. W. Landsea, 2007b: Seasonal tropical cyclone forecasts. *WMO Bull.,* **56,** 297–309.

Elsner, J., and T. Jagger, 2006: Prediction models for annual U.S. hurricane counts. *J. Climate,* **19,** 2935–2952.

Garner, S., I. Held, T. Knutson, and J. Sirutis, 2009: The roles of wind shear and thermodynamic stability in past and projected changes of Atlantic tropical cyclone activity. *J. Climate,* **22,** 4723–4734.

Goldenberg, S., and L. Shapiro, 1996: Physical mechanisms for the association of El Niño and West African rainfall with Atlantic major hurricane activity. *J. Climate,* **9,** 1169–1187.

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.

——, 1984b: Atlantic seasonal hurricane frequency. Part II: Forecasting its variability. *Mon. Wea. Rev.,* **112,** 1669–1683.

Harrison, D., and N. Larkin, 1998: El Niño–Southern Oscillation sea surface temperature and wind anomalies, 1946–1993. *Rev. Geophys.,* **36,** 353–399.

Klein, S., B. Soden, and N.-C. Lau, 1999: Remote sea surface temperature variations during ENSO: Evidence for a tropical atmospheric bridge. *J. Climate,* **12,** 917–932.

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

Knutson, T., J. Sirutis, S. Garner, G. Vecchi, and I. Held, 2008: Simulated reduction in Atlantic hurricane frequency under twenty-first-century warming conditions. *Nat. Geosci.,* **1,** 359–364, doi:10.1038/ngeo202.

Kruk, M., K. Knapp, D. Levinson, and J. Kossin, 2010: A technique for combining global tropical cyclone best-track data. *J. Atmos. Oceanic Technol.,* **27,** 680–692.

Kushnir, Y., W. Robinson, P. Chang, and A. Robertson, 2006: The physical basis for predicting Atlantic sector seasonal-to-interannual climate variability. *J. Climate,* **19,** 5949–5970.

LaRow, T., Y.-K. Lim, D. Shin, E. Chassignet, and S. Cocke, 2008: Atlantic basin seasonal hurricane simulations. *J. Climate,* **21,** 3191–3206.

——, 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.

Putman, W. M., and S.-J. Lin, 2007: Finite-volume transport on various cubed-sphere grid. *J. Comput. Phys.,* **227,** 55–78.

Rayner, R., D. Parker, E. Horton, C. Folland, L. Alexander, and D. Rowel, 2003: Global analyses 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.

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

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

Vecchi, G., and B. Soden, 2007: Effect of remote sea surface temperature change on tropical cyclone potential intensity. *Nature,* **450,** 1066–1071, doi:10.1038/nature06423.

——, K. Swanson, and B. Soden, 2008: Whither hurricane activity? *Nature,* **322,** 687–689.

Villarini, G., G. Vecchi, and J. Smith, 2010: Modeling the dependence of tropical storm counts in the North Atlantic basin on climate indices. *Mon. Wea. Rev.,* **138,** 2681–2705.

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

——, 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, C., and S.-K. Lee, 2009: Co-variability of tropical cyclones in the North Atlantic and the eastern North Pacific. *Geophys. Res. Lett.,* **36,** L24702, doi:10.1029/2009GL041469.

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.

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

Zhao, M., I. M. Held, S.-J. Lin, and G. A. Vecchi, 2009: Simulations of global hurricane climatology, interannual variability, and response to global warming using a 50 km resolution GCM. *J. Climate,* **22,** 6653–6678.