

Seasonal Forecasts of Tropical Cyclone Activity in a High-Atmospheric-Resolution Coupled Prediction System*

JULIA V. MANGANELLO,⁺ KEVIN I. HODGES,[#] BENJAMIN A. CASH,⁺ JAMES L. KINTER III,^{+,@}
ERIC L. ALTSHULER,⁺ MICHAEL J. FENNESSY,⁺ FREDERIC VITART,[&] FRANCO MOLteni,[&] AND
PETER TOWERS[&]

⁺ *Center for Ocean–Land–Atmosphere Studies, George Mason University, Fairfax, Virginia*

[#] *Department of Meteorology, University of Reading, Reading, United Kingdom*

[@] *George Mason University, Fairfax, Virginia*

[&] *European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom*

(Manuscript received 28 July 2015, in final form 23 November 2015)

ABSTRACT

Seasonal forecast skill of the basinwide and regional tropical cyclone (TC) activity in an experimental coupled prediction system based on the ECMWF System 4 is assessed. As part of a collaboration between the Center for Ocean–Land–Atmosphere Studies (COLA) and the ECMWF called Project Minerva, the system is integrated at the atmospheric horizontal spectral resolutions of T319, T639, and T1279. Seven-month hindcasts starting from 1 May for the years 1980–2011 are produced at all three resolutions with at least 15 ensemble members. The Minerva system demonstrates statistically significant skill for retrospective forecasts of TC frequency and accumulated cyclone energy (ACE) in the North Atlantic (NA), eastern North Pacific (EP), and western North Pacific. While the highest scores overall are achieved in the North Pacific, the skill in the NA appears to be limited by an overly strong influence of the tropical Pacific variability. Higher model resolution improves skill scores for the ACE and, to a lesser extent, the TC frequency, even though the influence of large-scale climate variations on these TC activity measures is largely independent of resolution changes. The biggest gain occurs in transition from T319 to T639. Significant skill in regional TC forecasts is achieved over broad areas of the Northern Hemisphere. The highest-resolution hindcasts exhibit additional locations with skill in the NA and EP, including land-adjacent areas. The feasibility of regional intensity forecasts is assessed. In the presence of the coupled model biases, the benefits of high resolution for seasonal TC forecasting may be underestimated.

1. Introduction

Skillful seasonal forecasting of tropical cyclone (TC) activity continues to be a scientific challenge with implications for society and economies. Seasonal prediction is predicated on the fact that signals of oceanic origin [sea surface temperature (SST), sea ice, etc.], changes in radiative forcing (solar, greenhouse gases,

aerosols, etc.), and land surface conditions (e.g., soil moisture) have long time scales (seasons and years) and may exhibit predictable evolution. This in turn can significantly influence the atmospheric circulation and provide a certain degree of predictability on the one-season to 1-yr time scales (e.g., Shukla et al. 2000). Predictability of the seasonal mean TC activity in particular has its origin in the high dependence of TC statistics on the atmospheric and oceanic conditions, such as SST and vertical wind shear (VWS) (Gray 1979). In the tropics, changes in the large-scale circulation are strongly related to changes in the SST distribution, such as El Niño–Southern Oscillation (ENSO), which is considered to be predictable on seasonal time scales (e.g., Kim et al. 2012; MacLachlan et al. 2015). From this perspective, seasonal TC activity may be considered a stochastic process modulated by the seasonal climatic conditions.

* Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JCLI-D-15-0531.s1>.

Corresponding author address: Julia V. Manganello, Center for Ocean–Land–Atmosphere Studies, George Mason University, 113 Research Hall, Mail Stop 2B3, 4400 University Dr., Fairfax, VA 22030.
E-mail: jvisneva@gmu.edu

The observed relationships between large-scale climate variability and TC statistics have led to the development of the first seasonal forecasts of TC activity, which are based on statistical methods (see review by [Camargo et al. 2010](#)). Such forecasts continue to be issued operationally for the North Atlantic (NA), eastern North Pacific (EP), western North Pacific (WP), and Australian regions by a number of governmental agencies, academic institutions, and private companies.

As a result of model development, global atmospheric general circulation models (AGCMs) are becoming increasingly skillful at explicitly simulating TCs, including their genesis and life cycle (e.g., [Goerss et al. 2004](#); [Halperin et al. 2013](#); [Met Office 2014](#); [Roberts et al. 2015](#)). Consequently, seasonal prediction systems based on such models have also become an attractive tool for TC forecasting. Using high-resolution AGCMs forced by predicted or persistent SST anomalies (SSTAs) is one example of this approach (e.g., [LaRow et al. 2010](#); [Zhao et al. 2010](#); [Chen and Lin 2011, 2013](#)). For instance, [Chen and Lin \(2011, 2013\)](#) have demonstrated high skill in retrospective seasonal forecasts of the NA TC frequency during 1990–2010 using a 25-km AGCM, with less success in the North Pacific. They conclude that the assumption of persistent SSTAs may be less applicable to the WP and not adequate for the 2011–13 seasons ([Chen and Lin 2014](#)). The authors also suggest that further improvement of the TC seasonal predictions is partly dependent on the improved model resolution.

Dynamical seasonal forecasts of TC activity based on coupled ocean–atmosphere general circulation models (CGCMs) have been investigated since the late 1990s and issued operationally at the European Centre for Medium-Range Weather Forecasts (ECMWF) beginning in 2001 ([Vitart and Stockdale 2001](#)). These earlier and many-present operational seasonal prediction systems employ relatively coarse-resolution AGCMs to predict the seasonal evolution of the large-scale climate. Although low-resolution models are capable of simulating TC-type vortices with some realistic features (see review by [Walsh 2008](#)), there are known deficiencies in a number of climatological characteristics such as genesis patterns, mean TC frequency, tracks, structure, and intensity distribution (e.g., [Manganello et al. 2012](#); [Camargo 2013](#); [Strachan et al. 2013](#); [Roberts et al. 2015](#)). Yet, interannual variability of the basinwide, seasonally aggregated TC activity metrics like TC frequency and accumulated cyclone energy (ACE)¹ can be

quite realistic in such models provided that the variability of the large-scale circulation is well simulated (e.g., [Vitart et al. 1997, 1999](#)). The forecast skill of these TC activity measures is therefore found to be competitive to statistical forecasts, particularly when a multimodel ensemble technique is applied, and is dependent on the skill of forecasts of the large-scale circulation ([Vitart and Stockdale 2001](#); [Vitart 2006](#); [Vitart et al. 2007](#)). While the consensus is that high model resolution is essential for the realistic simulation of TCs, there have nonetheless been very few studies that have directly examined this influence on the skill of the seasonal TC activity forecasts. Recent work by [Vecchi et al. \(2014\)](#) and [Camp et al. \(2015\)](#) has demonstrated that reasonably skillful predictions of regional, in addition to basinwide, seasonal TC activity can be achieved using high-resolution coupled climate models. A question arises whether these regional TC forecasts would also benefit from the systematic increase in the model resolution.

In this paper, we evaluate the performance of retrospective forecasts of the seasonal mean TC activity, and overall TC climatology, in an experimental high-resolution seasonal prediction system similar to the ECMWF System 4 (hereafter System 4; [Molteni et al. 2011](#)). As part of an international collaboration called Project Minerva ([Zhu et al. 2015](#)), the system is integrated at atmospheric horizontal resolutions ranging from T319 to T639 and T1279. The coarsest resolution (T319; ~62-km grid) is already higher than in most current operational seasonal prediction systems (e.g., [Molteni et al. 2011](#); [Saha et al. 2014](#)). The finest resolution (T1279; ~16-km grid) is presently used operationally at the ECMWF for (uncoupled) medium-range weather forecasts. We examine whether further increasing atmospheric resolution beyond the “TC permitting” range (20–100-km grid; e.g., [Zhao and Held 2012](#)) leads to an improved skill of the basinwide and regional TC activity hindcasts. The influence of the ensemble size on the forecast skill is also addressed. To evaluate the impact of the large-scale climate on the quality of these hindcasts, we assess the skill of the relevant basin-specific climatic conditions and their relationship with the TC activity hindcasts compared to observations. The potential influence of the coupled-model biases on these connections is discussed.

The paper is organized as follows. [Section 2](#) contains a description of the modeling system and numerical experiments, methodologies of identifying and tracking TCs, and observational data used in the study. The climatology of TC formation and tracks and intensity distributions are briefly described in [section 3](#). Analysis of the seasonal forecast skill of the basinwide and regional TC activity is presented in [sections 4 and 5](#), respectively. A summary of the results and some concluding remarks are given in [section 6](#).

¹ ACE is an integral measure of the TC activity and is computed by integrating the squared peak wind speed at each time interval along a track and accumulating over all tracks in a season (see [Bell et al. 2000](#)).

TABLE 1. TC identification criteria.

Criteria	Horizontal resolution of the Minerva forecasting system		
	T1279	T639	T319
1) Surface (10 m) wind speed threshold (m s^{-1} ; intensity threshold).	15.4 ^a	14.4 ^{a,b}	13.4 ^{a,b}
2) Difference in vorticity between 850 and 250 hPa (a warm core condition). ^c	Larger than zero for all resolutions.		
3) Vorticity max at each level (6 levels) between 850 and 250 hPa (a coherent vertical structure condition). ^c	Applied to all resolutions.		
4) Criteria 1–3 are achieved for four consecutive time steps (24 h).	Applied to all resolutions.		
5) Cyclogenesis (first identification) occurs between 0°–20°N over land and 0°–30°N over oceans.	Applied to all resolutions.		

^a Observed tropical storm threshold for 10-min MSW is used.

^b Surface wind speed threshold is further adjusted for model resolution based on Fig. 2 in Walsh et al. (2007). Values derived from a selection of Hurricane Research Division wind analysis are used.

^c Vorticity is truncated at spectral T319 horizontal resolution (N160) common to all resolutions.

2. Methodology

a. Modeling system and experimental setup

Project Minerva employs a coupled operational long-range prediction system based on the System 4 (Molteni et al. 2011). The two modeling systems have very similar configurations in terms of the ocean model, coupling, initialization, and ensemble perturbation generation methods. The ocean model is Nucleus for European Modelling of the Ocean (NEMO; Madec 2008), version 3.0, on the ORCA1 grid, which has a horizontal resolution of about 1° (with equatorial refinement of $1/3^\circ$) and 42 levels in the vertical. The ocean–atmosphere coupling is implemented from the start and occurs with a 3-h coupling frequency. The unperturbed initial conditions for the atmosphere come from the ECMWF interim reanalysis (ERA-Interim; Dee et al. 2011) and Ocean Reanalysis System 4 (ORA-S4) for the ocean. Ozone initial conditions are taken from the seasonally varying climatology. Stratospheric volcanic aerosols are included; time variation of greenhouse gases is specified as well. More details about System 4, its initialization, and its ensemble generation can be found in Molteni et al. (2011).

The main differences between System 4 and the Minerva forecasting system are in their component AGCMs. Both use the ECMWF Integrated Forecast System (IFS; ECMWF 2015), cycle 36r4 at spectral T255 horizontal resolution in System 4, and cycle 38r1 at three different spectral horizontal resolutions in Project Minerva. These resolutions are T319, T639, and T1279, corresponding approximately to 62-, 31-, and 16-km grid spacing, respectively. The ECMWF IFS is a spectral, semi-implicit, semi-Lagrangian hydrostatic model with 91 levels in the vertical and a model top in the mesosphere at 0.01 hPa.

Our study is based on a subset of Minerva integrations, which includes 7-month hindcasts started from 1 May initial conditions during 1980–2011 and consisting of 15 ensemble members for the T1279 and T639 configurations and 51 members for the T319. These experiments are respectively referred to as T1279, T639, and T319 hereafter. When comparing results, we use all 51 ensemble members of T319, unless otherwise noted. Upper-air data for all model configurations are converted to the common T319 resolution prior to the analysis. An evaluation of the modeling system's ability to represent the climatology is in the supplementary material (see Fig. S1).

b. Identification and tracking of tropical cyclones

Predicted storms are identified explicitly in the model data using an objective feature-tracking methodology. The initial TC identification and tracking is similar to that used in Bengtsson et al. (2007) and is based on the tracking algorithm of Hodges (1994, 1995, 1999). Vortices are detected in the Northern Hemisphere (NH) as maxima in the 6-hourly relative vorticity field averaged over 850-, 700-, and 600-hPa levels, with values greater than $5 \times 10^{-6} \text{ s}^{-1}$ (at a spectral horizontal resolution of T63). Vertical averaging of vorticity is found to produce more coherent tracks and to capture more of the life cycle of storms that may include an African easterly wave (AEW) precursor, compared to the tracking based on a single-level vorticity field (Serra et al. 2010). A posttracking lifetime filter of 2 days is employed. The TC identification criteria (see Table 1) are applied to the raw tracks to separate the simulated TCs from other synoptic systems. As a result, model storms tend to include both earlier and later stages of a life cycle than the observed

storms. Further details of the TC identification and tracking can be found in [Manganello et al. \(2012\)](#).

An additional filter is used to remove spurious regionally confined storms in the Caribbean Sea off the northern tip of South America, which are endemic at coarser resolutions (see [section 3](#)). This feature is also found in other versions of the IFS at low resolutions and appears to be a consequence of the insufficiently resolved sharp orography in the northern Venezuelan highlands (see [Manganello et al. 2012](#)). An application of this filter results in a significantly improved hindcast skill in the NA for the T319 model with no major changes for the T639 and T1279.

Our analysis is performed for the NH only, and the results are reported for the NA, EP, and WP basins, as the skill for the north Indian Ocean is found to be quite low. A storm is assigned to a particular basin if it reaches its peak intensity there. For instance, it may originate as an easterly wave over the Caribbean Sea and propagate into the EP (e.g., [Serra et al. 2010](#)). If it reaches its lifetime maximum intensity there it is classified as an EP TC. The EP TCs also include the central North Pacific storms.

The TC activity is computed for the May–November (MJJASON) season, which encompasses the whole period of integration and also represents the bulk of the annual TC activity in the NA, EP, and WP basins. This period includes the range of deterministic atmospheric prediction (about 2 weeks) when the initialized atmospheric state can directly affect the prediction. However, we do not believe that the inclusion of the full month of May significantly affects the seasonal mean results since this is a month of very weak TC activity in all the basins reported.

c. Observational and reanalysis data

To compare the simulated TCs with those observed, we use data from the International Best Track Archive for Climate Stewardship (IBTrACS, version v02r01; [Knapp et al. 2010](#)). IBTrACS uses 10-min average wind speed at 10-m elevation for the maximum sustained wind (MSW) estimate, which closely corresponds to the model definition of MSW (see [Table 1](#)). We also use the same conversion coefficient between 1- and 10-min winds equal to 0.88 (see [Knapp et al. 2010](#)) to adjust TC thresholds. Thus, the “tropical storm” threshold of 17.5 m s^{-1} (34 kt; $1 \text{ kt} \approx 0.51 \text{ m s}^{-1}$) defined for the 1-min MSW becomes 15.4 m s^{-1} (30 kt) for the 10-min MSW. For the direct comparison with model-simulated tracks, IBTrACS data are processed by applying criteria 1 and 4 of [Table 1](#).

Surface fields and pressure-level analysis from the ERA-Interim for the period 1980–2011 are used to

TABLE 2. Climatological means of the TC frequency and the ACE for the MJJASON season of 1980–2011 for IBTrACS (OBS) and Minerva forecasts. Differences between the model results and observations that are statistically significant at the 95% confidence level using a two-sided Student’s *t* test are shown in boldface. Degrees of freedom are computed taking into account serial correlation in the time series using [Bretherton et al. \(1999\)](#) formula for effective sample size of order 1.

	NA	EP	WP
	TC frequency		
OBS	9.8	15.6	22.5
T1279	10.2	10.2	26.5
T639	9.9	10.7	28.7
T319	9.0	11.1	30.1
	ACE (10^4 kt^2)		
OBS	86.2	119.1	206.4
T1279	68.8	65.2	157.6
T639	49.1	49.9	136.1
T319	27.0	37.7	104.1

compute observational estimates of the atmospheric and SST-based indices in [section 4](#).

3. Predicted TC climatology

A brief review of the simulated TC climatology is given below with the purpose of demonstrating the overall level of skill, its dependence on model resolution, and its role as a potential aid in diagnosing the skill of seasonal forecasts in later sections.

In the NA the seasonal mean TC frequency is quite realistic in all Minerva hindcasts ([Table 2](#)), contrary to some recent studies ([Strachan et al. 2013](#); [MacLachlan et al. 2015](#); [Roberts et al. 2015](#)). There is a clear increase of the TC frequency with the resolution [as has been reported previously in, e.g., [Manganello et al. \(2012\)](#), [Strachan et al. \(2013\)](#), and [Roberts et al. \(2015\)](#)] where T639 and T1279 attain values very close to observations. This is largely a result of an enhanced eastern main development region (MDR; 7.5° – 22.5°N , 80° – 20°W) genesis in the higher-resolution models ([Figs. 1a–d](#)), which is also found in the above studies, and could be partly due to better tracking of the AEWs with the new tracking procedure ([section 2b](#)). In contrast, TC genesis in the western MDR and the Caribbean Sea is weaker than in observations (and the latter center is misplaced farther southeast from its observed location in T639 and T319, as mentioned in [section 2b](#)). These biases may be related to suppressed convective activity in the two regions (not shown). Although poor genesis in the Caribbean Sea and the Gulf of Mexico is common among recent models (e.g., [Strazzo et al. 2013](#)), the latter center is simulated quite well in all Minerva hindcasts. The overall distribution of tracks is realistic, albeit the track

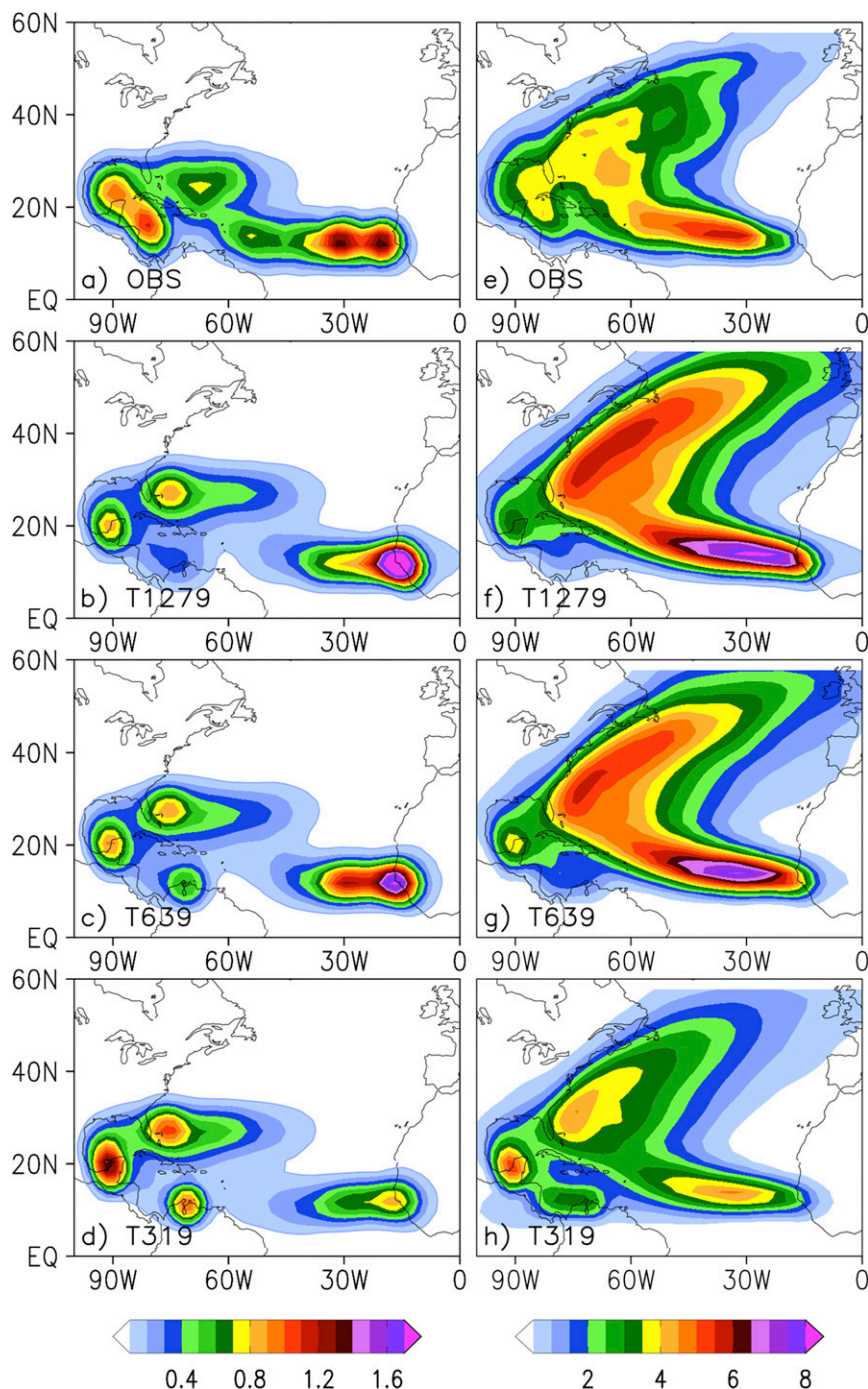


FIG. 1. NA (left) genesis and (right) track densities as number density per season per unit area equivalent to a 5° spherical cap for (a),(e) IBTrACS (OBS) and Minerva hindcasts at (b),(f) T1279; (c),(g) T639; and (d),(h) T319 resolutions based on MJJASON of 1980–2011.

density is too low in the MDR in T319 and somewhat overpredicted in T1279 and T639 both in the MDR and the western subtropical NA (Figs. 1e–h).

Compared to the NA, the seasonal mean TC frequency is systematically under- and overpredicted in

the EP and WP, respectively (Table 2). In these two basins, the climatological mean TC count also decreases with an increase in the model resolution. In the EP, this is mainly due to more storms originating from easterly waves in the Caribbean Sea at coarser

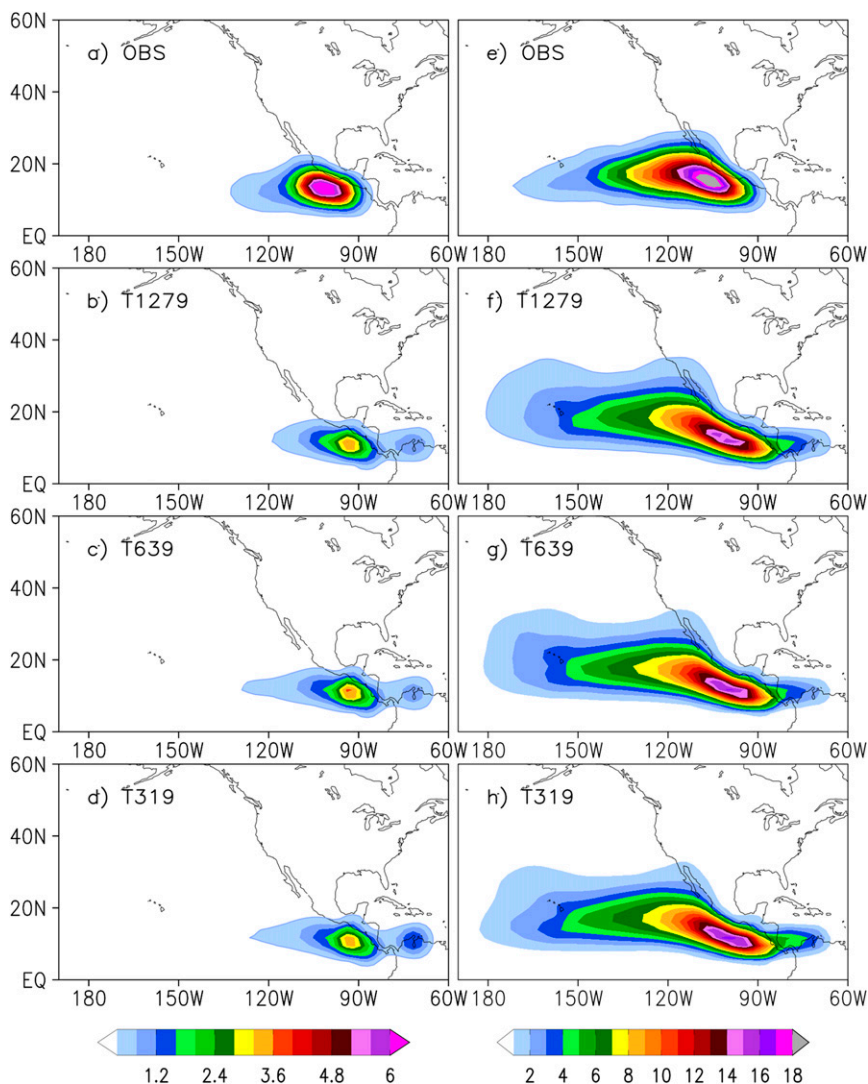


FIG. 2. As in Fig. 1, but for the EP.

resolutions (Fig. 2), likely related to spurious wave activity off the northern tip of South America (see section 2b). The main center of the EP cyclogenesis off the Mexican Pacific coast is equally underrepresented in all Minerva hindcasts. In the WP, the maximum concentration of cyclogenesis occurs in the Philippine Sea and the South China Sea at lower resolutions, as opposed to the southeastern part of the basin in the observations (Figs. 3a–d). There are also more central Pacific storms in Minerva hindcasts. These two biases appear to be related to the intertropical convergence zone (ITCZ) errors (see the supplementary material), likely as a result of a significant cold bias in the cold tongue region (see Zhu et al. 2015). Overall, the WP seasonal TC count, genesis, and track densities (Figs. 3e–h) are best represented at the highest resolution, although there

is still insufficient genesis close to the equator in the T1279 model.

The climatological mean ACE is significantly lower than the observed in all three basins and at all resolutions, except for the NA at T1279 (Table 2). This is largely a consequence of the model's low skill in simulating the most intense storms in terms of the 10-m wind speed even at the highest T1279 resolution. The simulated frequency distributions of the lifetime maximum 10-m wind speed do not reproduce secondary peaks (not shown), which is in contrast with our previous analyses of the earlier IFS cycle at T1279 forced by the observed SST and sea ice (Manganello et al. 2012; Manganello et al. 2014a). (As a side note, the northwestern bias in the WP genesis in Minerva precludes the development of the most intense typhoons, which primarily form in

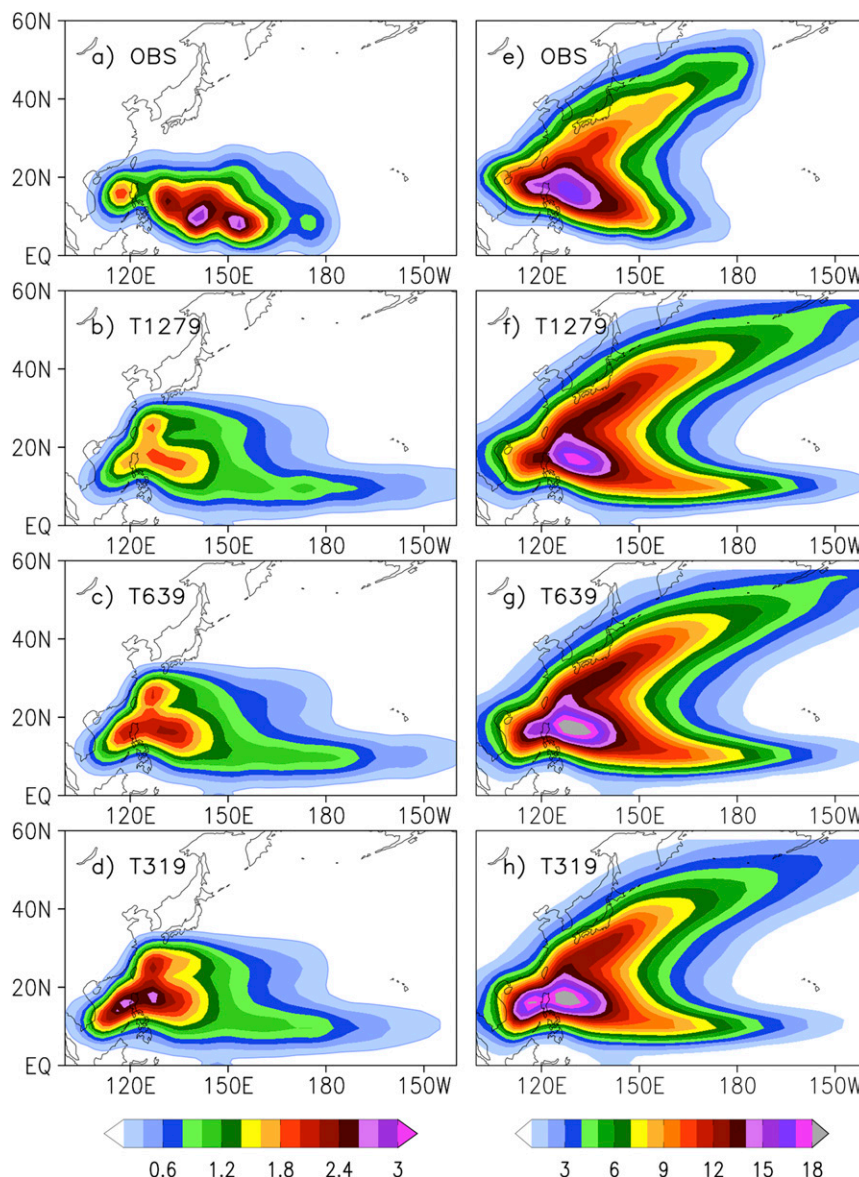


FIG. 3. As in Fig. 1, but for the WP.

the southeastern part of the domain.) Overall, the T1279 intensity distributions are realistic up to wind speeds corresponding to Saffir–Simpson category 3 hurricanes in the NA and category 2 storms in the EP and WP. The strongest model TCs have peak wind speeds of 55.9 m s^{-1} in the NA, 52.9 m s^{-1} in the EP, and 59.6 m s^{-1} in the WP, which are all equivalent to the category 4 storms. If, on the other hand, the TC intensity is assessed using the lifetime minimum sea level pressure (SLP), the resultant T1279 distributions show much better correspondence to observations, particularly in the NA (Fig. S2 in the supplemental material). (For the T639 and T319 models, the intensity distributions continue to

be too narrow and skewed toward the lowest intensities.) These results are similar to the findings of Chen and Lin (2013) and suggest a possibility of storm category forecasting with the T1279 Minerva system using an SLP-based classification.

4. Seasonal forecasts of the basinwide TC activity

Over the period 1980–2011, Minerva demonstrates somewhat modest but significant skill in predicting the interannual variability of the seasonal mean TC frequency in all three basins (except for the NA in T319; Table 3). The highest correlations are 0.51 in the NA

TABLE 3. Linear correlation coefficients between the ensemble mean predicted and observed (IBTrACS) TC frequency and ACE for MJJASON of 1980–2011. Correlation coefficient values for the detrended time series are given in parentheses. Boldface values indicate that correlation coefficients are statistically significant at the 95% confidence level using a one-sided Student's *t* test and taking into account serial correlation in the time series using Bretherton et al. (1999) formula for effective sample size of order 2.

	NA	EP	WP
	TC frequency		
T1279	0.51 (0.48)	0.58 (0.45)	0.44
T639	0.48 (0.57)	0.52 (0.38)	0.46
T319	0.28 (0.34)	0.56 (0.42)	0.52
	ACE		
T1279	0.61 (0.59)	0.70 (0.64)	0.72
T639	0.64 (0.69)	0.71 (0.65)	0.76
T319	0.48 (0.50)	0.72 (0.66)	0.68

(T1279), 0.58 in the EP (T1279), and 0.52 in the WP (T319). The correlations for the ACE, which could potentially capture more of the climate influence and be less sensitive to the details of the TC identification, show much higher values reaching 0.64 in the NA (T639), 0.72 in the EP (T319), and 0.76 in the WP (T639). (The sensitivity of the forecast skill to atmospheric resolution is addressed later in section 4c.) Although a multimodel ensemble (MME) approach can be quite successful in improving the prediction of TCs (Vitart 2006; Vitart et al. 2007), it does not lead to more skillful forecasts in our study. The Minerva MME-based scores do not generally exceed the best individual model's score (not shown). This could be an indication that merely changing the atmospheric horizontal resolution based on a single model may not produce a sufficiently diverse ensemble where model biases cancel each other.

To get more insight into the above results, we recomputed correlations for the individual subperiods: 1980–89, 1990–99, and 2000–11 (Table 4). Minerva hindcasts exhibit significant variations in the level of skill from one decade to another. Correlations can reach values as high as 0.83 for the TC frequency (2000–11; T639) and 0.92 for the ACE (1990s; T1279) in the NA; 0.83 for the TC frequency and 0.86 for the ACE (both 1990s; T319) in the EP; and 0.75 for the TC frequency (2000–11; T319) and 0.87 for the ACE (1990s; T639) in the WP. But more importantly, in the 1980s correlations for both measures of the TC activity for all models and basins are low and mainly insignificant, and particularly so in the NA. This may be related to the fact that before 1989, ORA-S4 (see section 2a) uses ERA-40 fluxes as opposed to ERA-Interim, where the latter are found to improve the mean state and interannual variability of ocean fields, especially in the NA (Molteni et al. 2011). The influence of certain climate factors (e.g., West

TABLE 4. Linear correlation coefficients between the ensemble mean predicted and observed (IBTrACS) TC frequency and ACE for MJJASON of 1980–89 (P1), 1990–99 (P2), and 2000–11 (P3). Boldface values indicate that correlation coefficients are statistically significant at the 95% confidence level using a one-sided Student's *t* test and taking into account serial correlation in the time series using Bretherton et al. (1999) formula for effective sample size of order 2.

	NA			EP			WP		
	P1	P2	P3	P1	P2	P3	P1	P2	P3
	TC frequency								
T1279	−0.19	0.66	0.73	0.39	0.73	0.62	0.40	0.49	0.56
T639	−0.19	0.77	0.83	0.46	0.75	0.12	0.22	0.64	0.70
T319	−0.25	0.40	0.67	0.27	0.83	0.46	0.31	0.72	0.75
	ACE								
T1279	0.11	0.92	0.56	0.53	0.81	0.29	0.61	0.81	0.69
T639	0.34	0.90	0.62	0.46	0.81	0.24	0.50	0.87	0.84
T319	0.13	0.82	0.54	0.48	0.86	0.23	0.47	0.81	0.70

African precipitation, North Atlantic Oscillation) on the seasonal NA TC activity is known to vary depending on whether the background climate conditions are more or less favorable for cyclogenesis and development (Fink et al. 2010; Caron et al. 2015). It is possible that the Minerva forecasting system captures some interactions but not others or does not reproduce their timing, which may also contribute to the low skill in the NA during the inactive period of the 1980s. The analysis of the predictability of these influences has not been done, which is beyond the scope of the current paper. It is noteworthy that because of the poor skill in the NA in the 1980s, correlations for the rest of the period 1990–2011 improve substantially compared to the full 32-yr record and reach 0.73, 0.73, and 0.63 for the TC frequency and 0.74, 0.78, and 0.66 for the ACE for the T1279, T639, and T319 model, respectively (cf. Table 3). The forecast skill in the EP and WP improves as well but overall to a lesser extent (not shown).

Accuracy, or a small difference between the ensemble mean forecast and observation, which is another measure of skill, is shown by the root-mean-square error (RMSE; Table 5). RMSE is computed after model data are calibrated using historical data: simulated TC frequency and ACE for each ensemble member are scaled by the ratio of the observed and predicted ensemble-mean values for the period 1980–2011, without cross validation (Table 2). Such calibration removes systematic bias in the simulated ensemble-mean quantities. RMSE values for the TC frequency and ACE are rather large, especially in the NA and EP, compared to their climatological means (Table 2) and other forecasting models and methods (e.g., Vitart 2006; Zhao et al. 2010; Vecchi et al. 2011). In the EP, this is related to the fact that predicted year-to-year variations are rather weak,

TABLE 5. RMSE between the calibrated ensemble mean predicted and observed (IBTrACS) TC frequency and ACE for MJJASON of 1980–2011 (see text for more detail). RMSE values for the detrended time series are given in the parentheses.

	NA	EP	WP
	TC frequency		
T1279	3.5 (2.9)	3.6 (3.6)	3.4
T639	3.6 (2.7)	3.8 (3.7)	3.4
T319	3.9 (3.1)	3.7 (3.6)	3.3
	ACE (10^4 kt ²)		
T1279	42.4 (38.2)	43.2 (42.7)	45.2
T639	41.6 (35.2)	43.8 (42.9)	43.2
T319	47.3 (41.5)	45.0 (43.8)	47.9

although multiyear variability is captured quite well (see [section 4b](#)). In the NA, the opposite is true, and high RMSE values are related to the absence of a positive trend in these two metrics over the studied period, which is a distinct model bias further discussed in [section 4a](#). In all basins, RMSE appears to be larger for the lowest-resolution model (T319), and more so for the ACE than the TC frequency.

The quality of forecasts can also be assessed using a skill score based on the RMSE and called the root-mean-square skill score (RMSSS). It measures the relative improvement of the forecast over some benchmark (usually low skilled) forecast like climatology. RMSSS is defined as one minus the ratio of the RMSE of the forecasts to the RMSE of the “forecasts” of climatology ([WMO 2002](#)). Minerva hindcasts of the TC frequency and the ACE show positive RMSSS values indicating potential improvement over a climatological hindcast ([Table 6](#)). The scores are rather modest for the TC frequency but show overall higher values for the ACE.

The skill of ensemble forecasts is also assessed using probabilistic diagnostics, which incorporate information about the ensemble distribution. One such measure is statistical reliability (or consistency). It estimates the degree to which forecast probabilities match the observed frequencies and can be represented by the ratio of the ensemble spread (averaged over all forecast years) to the RMSE (SPRvERR; e.g., [Buizza et al. 2005](#)). In a perfectly reliable ensemble forecast, the forecast uncertainty is fully accounted for, and the SPRvERR is equal to one. Minerva retrospective forecasts of the TC frequency can be considered as highly reliable (after calibration), except in the NA ([Table 7](#)). The hindcasts of the ACE are more underdispersed (or overconfident), but less so in the WP at the highest resolutions. The SPRvERR for the ACE also exhibits sensitivity to model resolution where the T319 hindcasts are less reliable than the T639 and T1279. This is due

TABLE 6. RMSSS for the TC frequency and the ACE for MJJASON of 1980–2011 (see text for more detail). RMSSS values for the detrended time series are given in the parentheses.

	NA	EP	WP
	TC frequency		
T1279	0.14 (0.12)	0.19 (0.11)	0.10
T639	0.12 (0.18)	0.14 (0.07)	0.11
T319	0.04 (0.06)	0.16 (0.09)	0.14
	ACE		
T1279	0.19 (0.19)	0.24 (0.19)	0.28
T639	0.21 (0.25)	0.23 (0.19)	0.31
T319	0.10 (0.12)	0.21 (0.17)	0.24

both to relatively larger RMSE and narrower ensemble spread (not shown).

a. North Atlantic

The retrospective skill in the NA TC frequency and the ACE is further illustrated in [Fig. 4](#) using the T1279 results. In this basin, Minerva captures interannual variations quite well (particularly after 1990) but fails to reproduce multidecadal-scale changes: generally low activity before about 1994 and much higher activity thereafter (e.g., [Goldenberg et al. 2001](#)). This bias is a feature of the Minerva system and is also present in the T319 and T639 models for both measures of the TC activity (not shown). It may be partly responsible for the relatively low skill in this basin. Indeed, correlation scores computed for the detrended time series are generally higher ([Table 3](#)). The RMSE drops significantly and becomes comparable to the other studies cited above ([Table 5](#)). RMSSS and reliability also improve where SPRvERR exceeds 0.9 for the TC frequency ([Tables 6 and 7](#)).

The seasonal NA TC activity is modulated by local changes in the SST, SLP, and VWS, among other factors (e.g., [Landsea 2000](#)), and remote climate variations such as ENSO (e.g., [Camargo et al. 2010](#) and references therein). Relative SST index (SST-REL) defined as the difference between SST in the NA MDR and that in the

TABLE 7. The SPRvERR for the TC frequency and the ACE for MJJASON of 1980–2011. SPRvERR values for the detrended time series are given in parentheses.

	NA	EP	WP
	TC frequency		
T1279	0.77 (0.90)	1.1 (1.0)	1.0
T639	0.74 (0.95)	1.0 (1.0)	1.1
T319	0.79 (0.96)	1.1 (1.1)	1.1
	ACE		
T1279	0.75 (0.79)	0.83 (0.82)	0.92
T639	0.77 (0.88)	0.80 (0.80)	0.93
T319	0.67 (0.74)	0.71 (0.71)	0.73

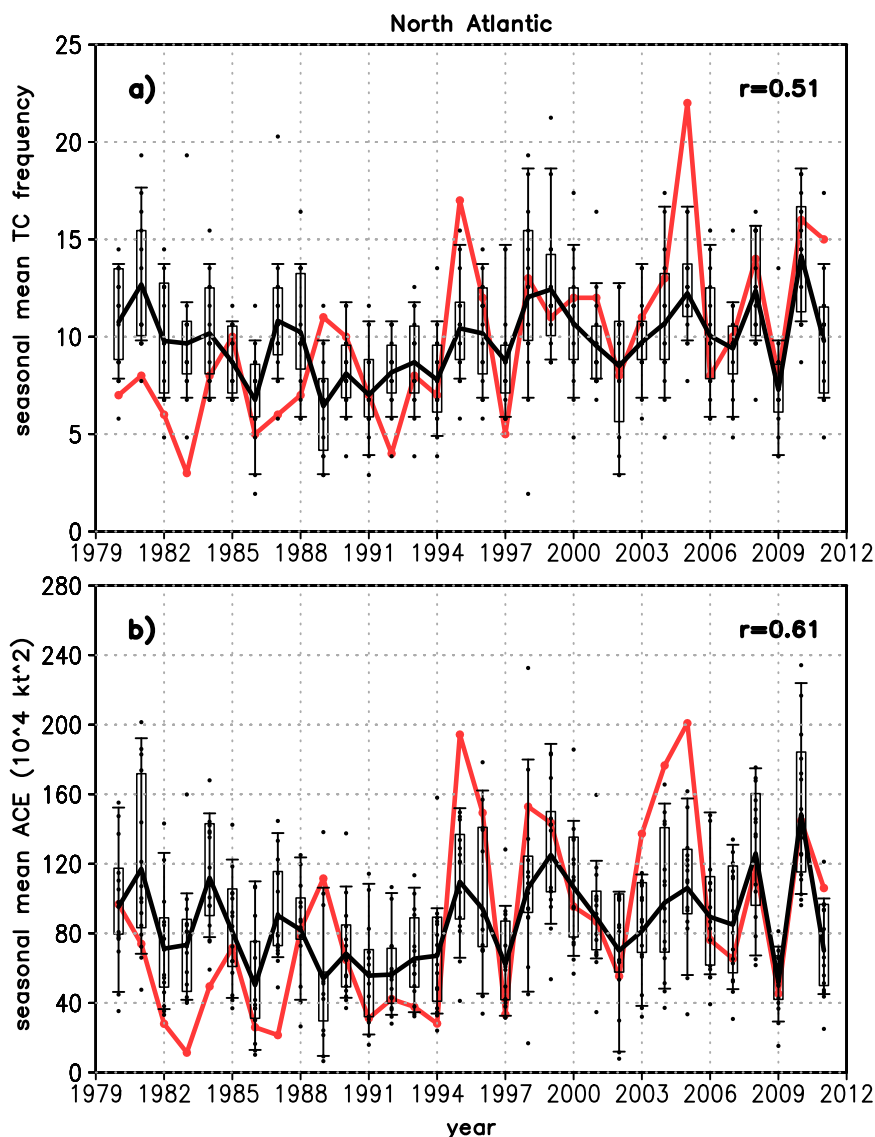


FIG. 4. Retrospective forecasts of the NA MJJASON (a) TC frequency and (b) ACE for 1980–2011. In both panels, red lines show the observed time series and black lines show the calibrated ensemble-mean forecasts using T1279 (see text for more details). Black dots denote calibrated forecasts from the individual ensemble members. Box-and-whisker plots delineate the 25th–75th and 10th–90th percentile ranges, respectively. Correlation coefficients between the observed time series and ensemble-mean forecasts are shown in the top-right corner of each panel.

global tropics (30°S – 30°N) has also been used to skillfully predict seasonal NA hurricane activity (e.g., [Zhao et al. 2010](#); [Vecchi et al. 2011](#)). Minerva hindcasts of the NA TC frequency and the ACE display realistic correlations with several large-scale climatic indices, with the exception of the VWS averaged over the MDR (VWS-MDR), which by itself could explain as much as 75%–80% of the variability in both metrics ([Table 8](#)). Correlations with the SST-REL [and SST averaged over the tropical Pacific (SST-PAC)] are also stronger. The

seasonal hindcast skill of the SST-based indices is very high in Minerva where all correlations exceed 0.8 [[Table 9](#); see also [Zhu et al. \(2015\)](#) for more detail on ENSO skill], whereas it is lower for VWS-MDR and SLP averaged over the basin's respective MDR (SLP-MDR; [Table 9](#)). VWS-MDR is known to respond to local as well as remote SST changes, where it has a tendency to diminish as a result of local warming ([Knaff 1997](#)) and increase because of the warming of the tropical Pacific and the tropical Indian Ocean ([Goldenberg and Shapiro](#)

TABLE 8. Linear correlation coefficients between the large-scale climatic indices and the TC frequency (top row) and the ACE (bottom row) for MJJASON of 1980–2011. Model correlations are based on the ensemble-mean values. Observed correlations (OBS) are between the IBTrACS TC frequency and the ACE and the climatic indices based on the ERA-Interim data. The indices are 1) SST-MDR [7.5°–22.5°N, 80°–20°W in the NA; 7.5°–15°N, 160°–80°W in the EP (Zhao et al. 2010); and 10°–20°N, 130°–160°E in the WP]; 2) SST-PAC (30°S–30°N, 120°E–80°W, where NA SSTs are masked out); 3) SST-REL (see Zhao et al. 2010, Vecchi et al. 2011); 4) the Niño-3.4 index (5°S–5°N, 120°–170°W); 5) EIO SSTAs averaged over 10°S–22.5°N, 75°–100°E (Zhan et al. 2011); 6) SSTG defined as the difference between the SST averaged over 40°–20°S, 160°E–170°W and 0°–16°N, 125°–165°E (Zhan et al. 2013; SSTG here is averaged over May–July); 7) VWS-MDR; and 8) SLP-MDR. Boldface values indicate that correlation coefficients are statistically significant at the 95% confidence level using a one-sided Student's *t* test and taking into account serial correlation in the time series using Bretherton et al. (1999) formula for effective sample size of order 2. Model correlations that are significantly different from their observed values using Fisher's *Z* statistic are marked with an asterisk.

	North Atlantic				Eastern North Pacific				Western North Pacific			
	OBS	T1279	T639	T319	OBS	T1279	T639	T319	OBS	T1279	T639	T319
SST-REL	0.65	0.87*	0.86*	0.75	0.63	0.58	0.53	0.72	—	—	—	—
	0.67	0.75	0.82	0.86	0.58	0.63	0.54	0.64				
SST-MDR	0.61	0.65	0.65	0.57	—	—	—	—	—	—	—	—
	0.62	0.50	0.54	0.60								
SST-PAC	−0.18	−0.39	−0.36	−0.26	—	—	—	—	—	—	—	—
	−0.16	−0.53	−0.55	−0.51								
Niño-3.4	−0.42	−0.52	−0.40	−0.20	0.43	0.39	0.13	0.37	0.27	0.30	0.30	0.33
	−0.48	−0.64	−0.61	−0.56	0.32	0.48	0.33	0.39	0.73	0.68	0.61	0.60
EIO SSTA	—	—	—	—	—	—	—	—	−0.54	−0.64	−0.38	−0.65
									−0.28	−0.38	−0.26	−0.46
SSTG	—	—	—	—	—	—	—	—	−0.60	−0.51	−0.52	−0.59
									−0.65	−0.63	−0.61	−0.68
VWS-MDR	−0.65	−0.90*	−0.87*	−0.74	−0.48	−0.62	−0.50	−0.74	−0.42	−0.44	−0.17	−0.62
	−0.70	−0.86	−0.90*	−0.90*	−0.43	−0.64	−0.55	−0.64	−0.33	−0.39	−0.09	−0.52
SLP-MDR	−0.68	−0.72	−0.76	−0.68	−0.64	−0.63	−0.41	−0.59	−0.58	−0.74	−0.77	−0.84*
	−0.56	−0.67	−0.73	−0.79	−0.48	−0.69	−0.53	−0.57	−0.54	−0.74	−0.75	−0.81

1996; Latif et al. 2007). Therefore, the net response of the VWS-MDR in the model would depend on the realism of the atmospheric teleconnections within the global tropics. Multiple regression analysis of the VWS-MDR onto the SST averaged over the basin's respective MDR (SST-MDR) and SST-PAC (see Table S1 in the supplemental material) implies a stronger control of the tropical Pacific over the NA VWS-MDR variability and in turn the variability of the NA TC frequency and the ACE in Minerva hindcasts. Geographically, the largest differences occur over the eastern MDR (not shown), which is one of the main cyclogenesis regions. The tropical Pacific also has a stronger influence over the simulated SLP variability in the tropical and subtropical NA (not shown). Further analysis of this issue is beyond the scope of the present study. We only note that these VWS errors are not present in the T1279 results from Project Athena (Kinter et al. 2013; the hindcast and AMIP-style integrations with an earlier cycle of the IFS forced by the observed records of SST and sea ice, not shown), which strongly points to Minerva coupled model biases as the source.

Errors in the representation of tropical heating and/or atmospheric teleconnections may also affect the realism of longer-time-scale variability, such as multidecadal

trends. In contrast to ERA-Interim, Minerva hindcasts of the VWS-MDR, SLP-MDR, and SST-REL have insignificant trends, similar to the NA TC frequency and the ACE (Table S2 in the supplemental material; see also Manganello et al. 2014b). Since simulated trends in the SST-MDR are quite realistic, particularly for T1279 (Table S2), these errors could in part be because of a stronger influence of the tropical Pacific variability that acts to offset multidecadal changes intrinsic to the tropical NA. The increase in the NA TC activity over the past few decades has also been linked with the downward trends in TC outflow temperature associated with a cooling tropical tropopause layer (Emanuel et al. 2013), or rather with upper tropospheric (UT; 300–150 hPa) temperature trends according to Vecchi et al. (2013). Temperature trends in Minerva hindcasts over the NA MDR tend to be positive in the UT and insignificant at the 100- and 50-hPa levels (not shown), which could further limit the realism of the TC activity trends in this basin.

b. North Pacific

In contrast to the NA, multiyear variability is well captured in the EP, while year-to-year variations are rather weak (Fig. 5). Detrending the TC frequency and the ACE often leads to lower hindcast skill in this basin

TABLE 9. Correlation coefficients between the simulated and observed large-scale climatic indices in Table 8. Boldface indicates that values are statistically significant at the 95% confidence level using a one-sided Student's *t* test and taking into account serial correlation in the time series using the Bretherton et al. (1999) formula for effective sample size of order 2.

	North Atlantic			Eastern North Pacific			Western North Pacific		
	T1279	T639	T319	T1279	T639	T319	T1279	T639	T319
SST-REL	0.84	0.83	0.83	0.93	0.90	0.92	—	—	—
SST-MDR	0.86	0.86	0.86	—	—	—	—	—	—
SST-PAC	0.83	0.81	0.82	—	—	—	—	—	—
Niño-3.4	0.88	0.87	0.87	0.88	0.87	0.87	0.88	0.87	0.87
EIO SSTA	—	—	—	—	—	—	0.81	0.77	0.76
SSTG	—	—	—	—	—	—	0.85	0.81	0.83
VWS-MDR	0.71	0.70	0.73	0.84	0.81	0.83	0.52	0.45	0.55
SLP-MDR	0.54	0.55	0.51	0.82	0.83	0.86	0.76	0.79	0.80

(Tables 3 and 5–7). The historical forecasts of the WP TC activity are overall quite skillful both on interannual and decadal time scales (Fig. 6). There are only a few years markedly outside the 10th–90th percentile range in this basin (1984 and 1990 for the TC frequency; and 1987, 1991, 1999, and 2011 for the ACE).

The seasonal EP TC activity is also influenced by the variability in the large-scale environmental variables, such as SST-REL, VWS, and SLP, which is partly related to ENSO (e.g., Camargo et al. 2010; Zhao et al. 2010). The observed correlations of the EP TC frequency and the ACE with the corresponding climatic indices are quite well reproduced in Minerva, with the exception of perhaps EP VWS-MDR, which has somewhat lower correlations (Table 8). In addition, the hindcast skill of these EP indices is quite high and exceeds the skill of the respective NA indices (Table 9), although the variance of the EP VWS-MDR is significantly lower than in the reanalysis (not shown). It is noteworthy that the EP and NA TC activity hindcasts vary out of phase, which is quite similar to the observations (Wang and Lee 2009): correlations for the full period of 1980–2011 are -0.55 , -0.56 , -0.32 , and -0.50 for the TC frequency and -0.61 , -0.62 , -0.43 , and -0.55 for the ACE from IBTrACS and T1279, T639, and T319 models, respectively. (All the above correlations are statistically significant, except for the TC frequency in the T639 model.)

Interannual variations of the WP TC activity are largely determined by changes in the location and strength of the monsoon trough (e.g., Chen et al. 1998, and references therein). ENSO has a dominant influence on the interannual variability of the monsoon trough as well as VWS and thermodynamic conditions in the region (e.g., Camargo et al. 2010, and references therein; Wu et al. 2012), which leads to mostly southeast-to-northwest shifts in the TC genesis reflected in the low correlation between the TC frequency and the Niño-3.4 index. On the other hand, the influence of ENSO on the TC intensity and lifetime is much larger (Wang and

Chan 2002; Camargo and Sobel 2005), resulting in a high correlation between the ACE and the Niño-3.4 index. These contrasting effects of the ENSO influence are well reproduced in Minerva, including the correlations with the WP VWS-MDR and, to a lesser extent, the WP SLP-MDR (Table 8). Zhan et al. (2011) have demonstrated that SSTAs in the east Indian Ocean (EIO) are an additional factor that modulate seasonal WP TC activity. In contrast with ENSO, EIO SSTAs significantly affect the basinwide TC frequency but have weaker influence on the intensity and subsequently ACE. Minerva hindcasts also seem to capture these differences (Table 8), which is in contrast with the results of Chen and Lin (2013), who found no significant statistical relation between EIO indices and model-predicted storms in this basin. Recently, Zhan et al. (2013) discovered that the spring SST gradient (SSTG) between the southwestern Pacific and the western Pacific warm pool is significantly anticorrelated with the WP TC frequency and an integral measure of the WP TC activity during the typhoon season. Minerva reproduces these observed correlations as well (Table 8), suggesting that both the atmospheric response to the evolving SSTA, ocean–atmosphere interactions and regional teleconnections in the model have some fidelity in this basin. The hindcast skill of the WP MDR indices, the EIO SSTAs, and the SSTG in Minerva is also fairly high (Table 9).

In conclusion, we note that, based on the above analysis, we do not find any systematic differences associated with resolution of either the prediction skill of the climatic indices relevant to the TC activity or their relationship with the basinwide TC activity measures such as TC frequency or the ACE in any of the basins examined.

c. Sensitivity of the correlation skill to the ensemble size and the atmospheric resolution

Here, we examine the influence of the ensemble size on the correlation skill of the TC frequency and the ACE hindcasts to identify the “optimum” size of the

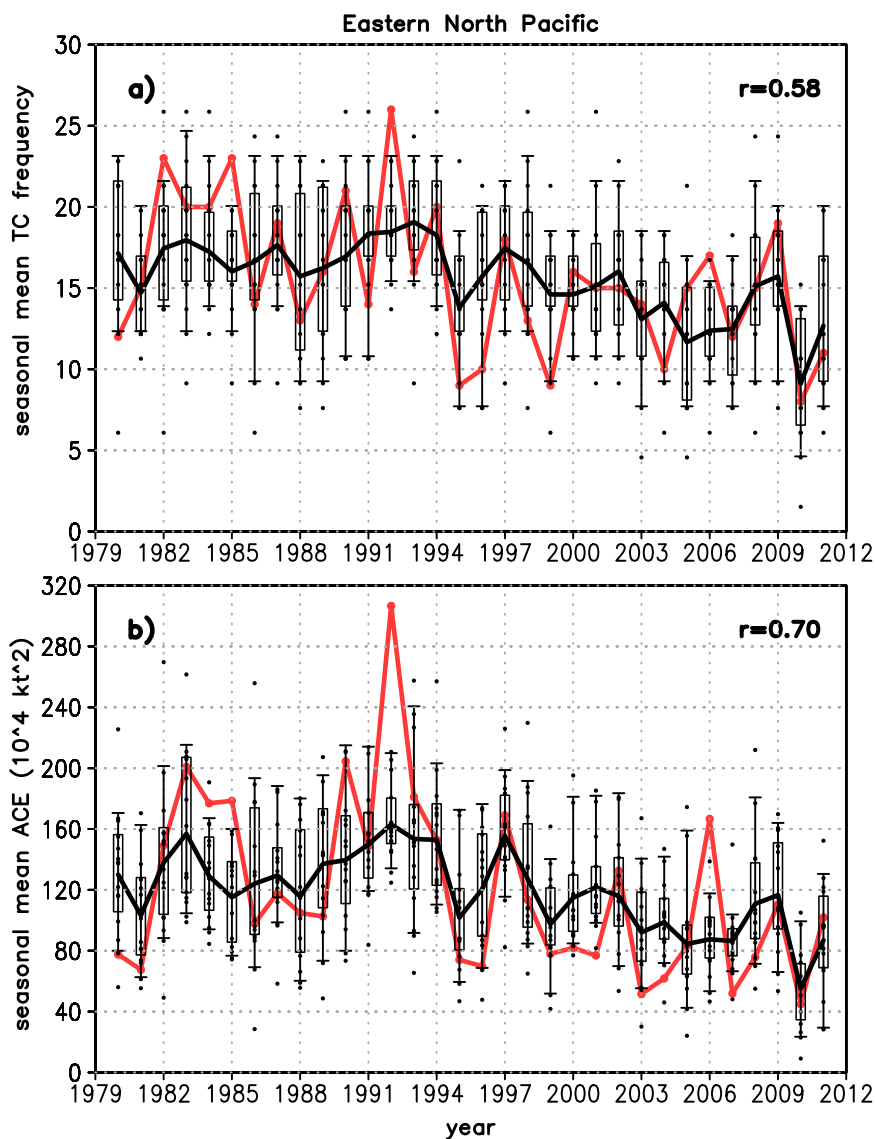


FIG. 5. As in Fig. 4, but for the EP.

ensemble and to rigorously evaluate the sensitivity of the correlation skill to the atmospheric horizontal resolution using identical ensemble sizes and taking into account uncertainty due to sampling. Figure 7 shows the mean values of the correlation coefficients as a function of the ensemble size for the TC frequency and the ACE. Also shown are the error bars that span the range within plus or minus one standard deviation of the mean. The associated uncertainty is partly due to sampling different combinations of the ensemble members and also due to randomly choosing a specific combination at each time in the record, as there is no connection between particular ensemble members from one season to another.

Figure 7 demonstrates that the fastest growth in skill occurs when the ensemble size increases from 1 to 5

members. The change in skill is still marked when the size of the ensemble is further increased to 15 members but occurs at a slower rate. Beyond that, the scores increase much more slowly than the extrapolation of the 15-member ensemble results would suggest (not shown) and saturate when the number of ensemble members reaches about 25–30. Coincidentally, a similar size of the ensemble (20–30 members) is recommended in operational hurricane forecasting to provide adequate estimates of the uncertainty (Gall et al. 2013).

Taking into account the uncertainty estimates and comparing ensembles of the same size, our results suggest that the degree to which atmospheric horizontal resolution affects correlation skill differs among the

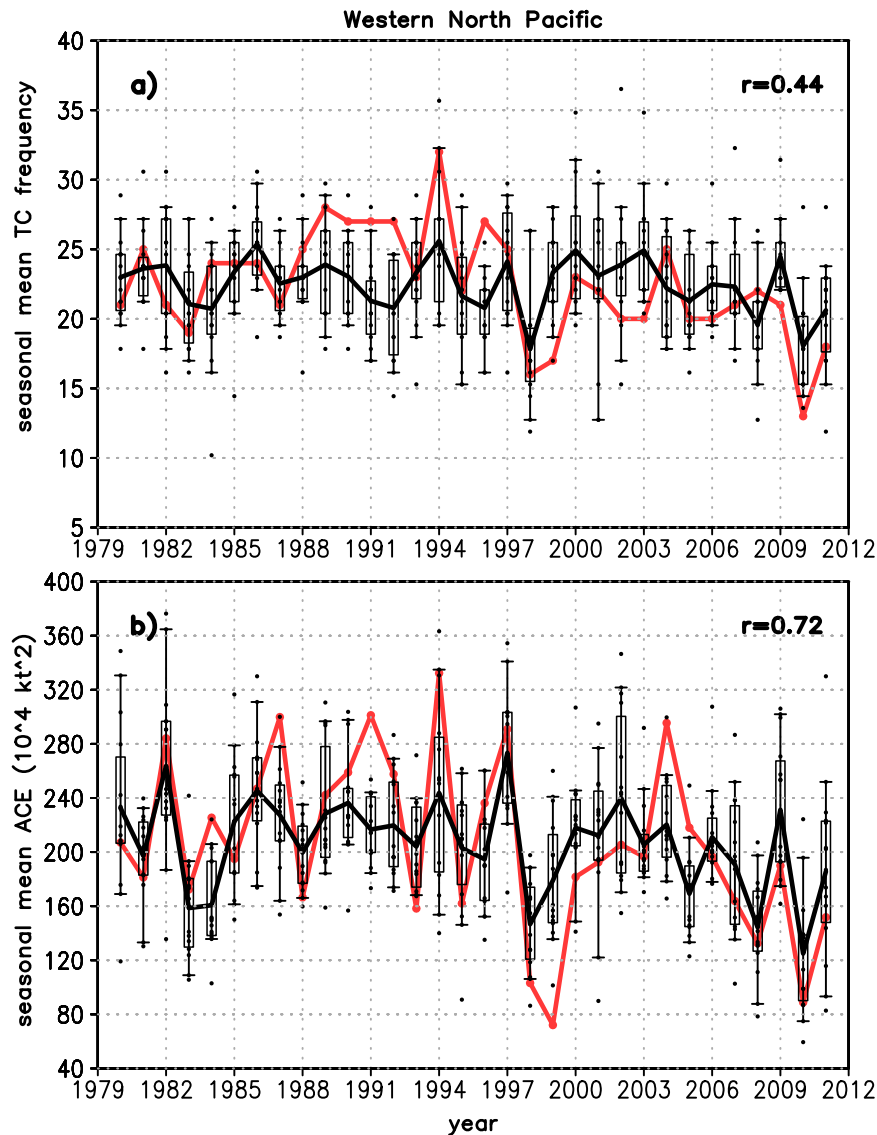


FIG. 6. As in Fig. 4, but for the WP.

basins. This influence also depends on the measure of the TC activity and is more evident for the ACE (where an increase in scores at higher resolutions is present in all three basins) than the TC frequency. In the NA, the impact of the resolution is most pronounced, and the T1279 and T639 scores are considerably higher than the T319 both for the TC frequency and the ACE (Figs. 7a,d), although the differences between the highest two resolutions are rather small. In the EP, the increase in correlation appears statistically significant only for T1279 in the case of the TC frequency and T639 for the ACE (Figs. 7b,e). The WP is the only basin where the correlation skill of the TC frequency hindcasts is not sensitive to model resolution (Fig. 7c). The ACE skill scores, however, do increase with the resolution but are

somewhat higher for the T639 model compared to T1279 (Fig. 7f).

5. Seasonal forecasts of the regional TC activity

Skillful seasonal forecasts of TC activity on subbasin or regional scales could significantly enhance the utility of such predictions (Vecchi and Villarini 2014). High-resolution dynamical models are regarded as one of the tools best suited for this purpose (e.g., Vecchi et al. 2014). It is therefore of interest to evaluate the performance of Minerva hindcasts in this respect.

Retrospective skill of regional TC forecasts is assessed here by means of Spearman rank correlation between the seasonal mean observed and predicted (all

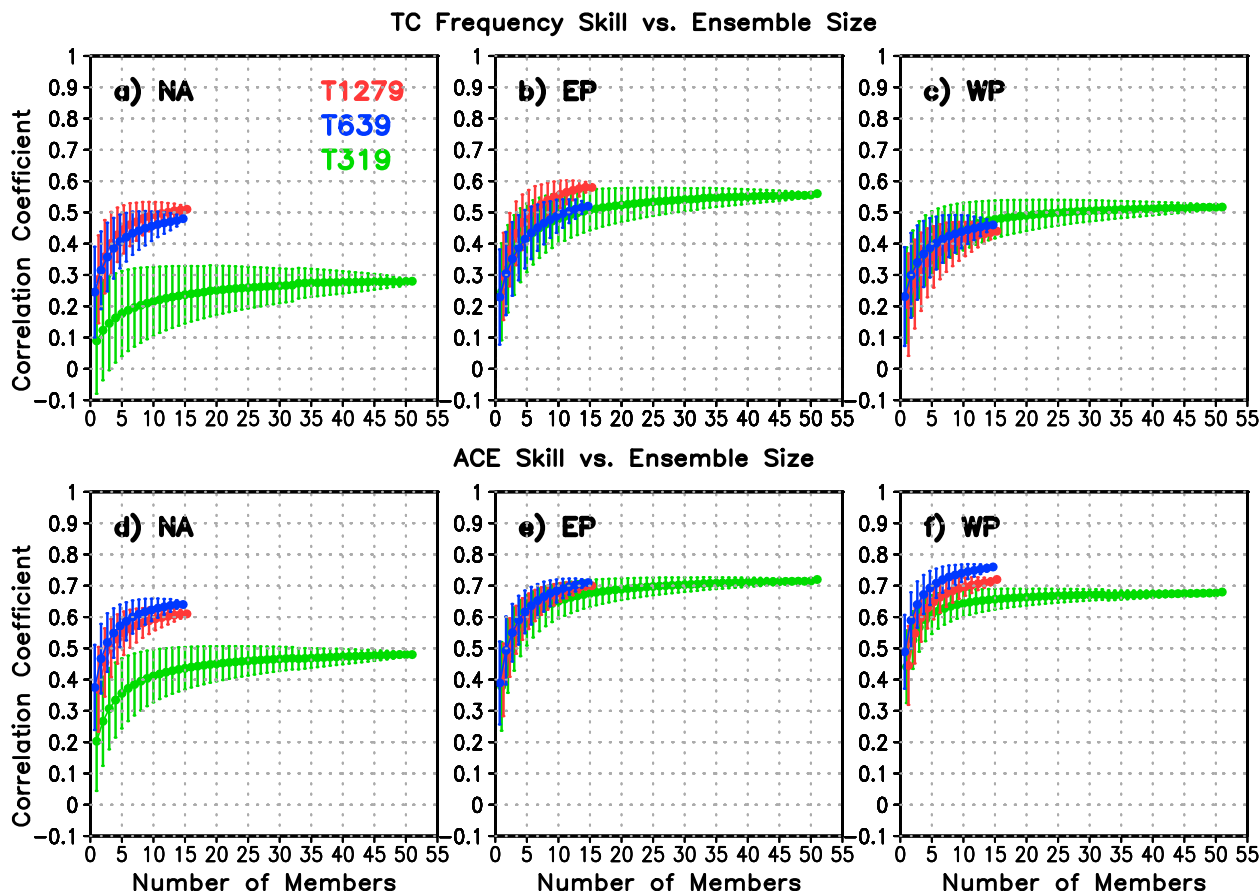


FIG. 7. Retrospective linear correlations between the MJJASON observed (IBTrACS) and predicted TC frequencies for 1980–2011 as a function of the ensemble size. Results are shown for the (a) NA, (b) EP, and (c) WP for T1279 (red), T639 (blue), and T319 (green). (d)–(f) As in (a)–(c), but for the ACE. Dots indicate mean values of the correlation coefficients, and error bars mark the range from plus to minus one standard deviation from the mean (see text for more details).

ensemble members) track densities (Fig. 8) similar to Vecchi et al. (2014, their Fig. 11). [We define track densities as number densities per season per unit area equivalent to a 5° spherical cap, which differs from the TC density definition used in Vecchi et al. (2014)]. Using this metric, significant skill in the NA is achieved mainly in the MDR and midlatitudes (Figs. 8a–c). The former region is important as many intense hurricanes with frequent landfall along the U.S. and Canadian coasts form and develop there (e.g., Kossin et al. 2010). Significant correlations are also present over the western Caribbean Sea (T639, Fig. 8b), part of the Gulf of Mexico (T1279, Fig. 8a), and along the U.S. mid-Atlantic seaboard (T1279, Fig. 8a). It is curious that only the T1279 model shows a fairly large area of significant skill in the immediate vicinity of the U.S. Atlantic coast, whereas in the coarser-resolution models this region is shifted to the northeast and is primarily over open waters (Figs. 8b,c). We suspect that this improvement is partly due to a more realistic relationship

of the track density variations with ENSO in this model, as discussed below.

In the EP, significant correlations are found in the central part of the basin (Figs. 8d–f). There are additional broad regions of skill in the vicinity of the Central American coast and near Hawaii, only in the T1279 hindcasts (Fig. 8d). On the other hand, correlations near the Pacific coast of North America are insignificant or even negative. Since the TC activity in this part of the domain is strongly modulated by the Madden–Julian oscillation (MJO; Camargo et al. 2008), this result is consistent with the limited skill of current forecasting systems in predicting the MJO (about 25 days in Minerva). Overall, the fraction of the EP TC region exhibiting skill appears to increase with the model resolution, particularly in transition from T639 to T1279. The WP demonstrates the highest correlations among all the basins, which are largely confined to the southeast (Figs. 8g–i) and are hence indicative of the strong footprint of ENSO (see below). In contrast with the EP,

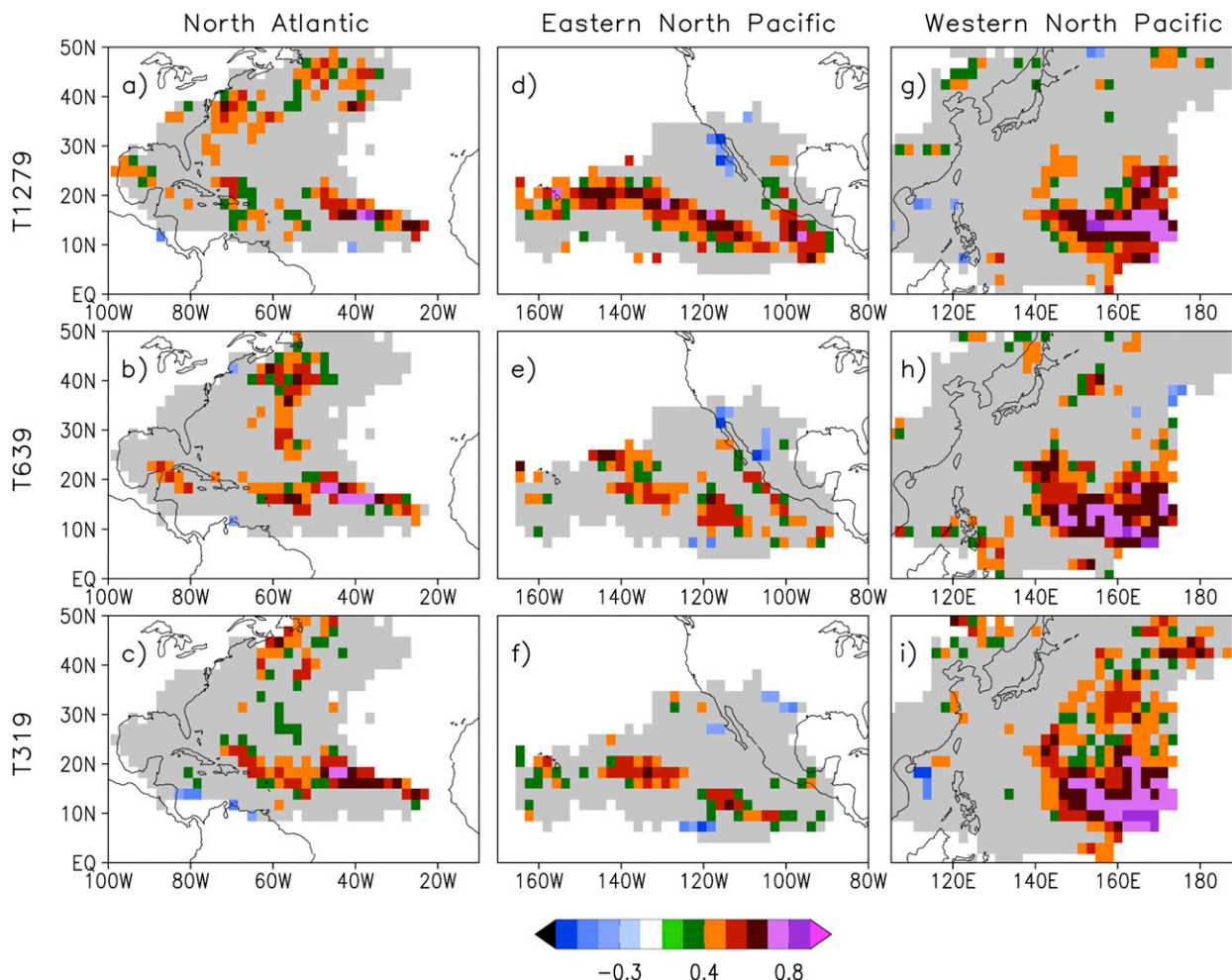


FIG. 8. Retrospective rank correlation between the MJJASON observed (IBTrACS) and predicted TC track densities as number densities per season per unit area equivalent to a 5° spherical cap for 1990–2011. Tracks from all ensemble members are used to compute seasonal mean predicted track densities. Results shown are for the (a)–(c) NA, (d)–(f) EP, and (g)–(i) WP for the (top) T1279, (middle) T639, and (bottom) T319 Minerva models. Color shading denotes values statistically significant at a two-sided $p = 0.1$ level. Gray shading indicates the regions where the observed track density is nonzero for at least 25% of the years.

higher resolution does not lead to improved skill here. Regrettably, there is practically no skill over the land-adjacent areas in this basin.

The skill of track density forecasts depends on the overall quality of the track climatologies, the skill of the basinwide TC frequency forecasts (Mei et al. 2014), and the fidelity of the relationship between track density variations and the large-scale modes of climate variability that are potentially predictable on seasonal time scales (Vecchi et al. 2014, and references therein). The latter is examined here in detail with respect to ENSO. In the observations, El Niño (La Niña) is associated with below-normal (above-normal) track density almost everywhere in the NA, but mostly in the MDR and the Gulf of Mexico (Fig. 9a; see also Mei et al. 2014). While Minerva hindcasts show similar correlation patterns,

there are noteworthy differences (Figs. 9b–d). The area of significant correlations is much larger in the model. Correlations are also too strong in the eastern MDR, more so at coarser resolutions. This is consistent with our analysis of the VWS-MDR (and SLP-MDR) variability in section 4a. Results with the T1279 model are overall the most realistic, taking into account the Gulf of Mexico, Caribbean Sea, and the central North Atlantic (Fig. 9b). At coarser resolutions, there is an increasingly strong relationship between track density along the eastern seaboard and ENSO, not found in the observations (Figs. 9c,d).

The observed EP track density exhibits a westward shift in response to ENSO (Fig. 9e; Camargo et al. 2008). While this change is generally captured in Minerva, the bulk of the track density increase occurs along the

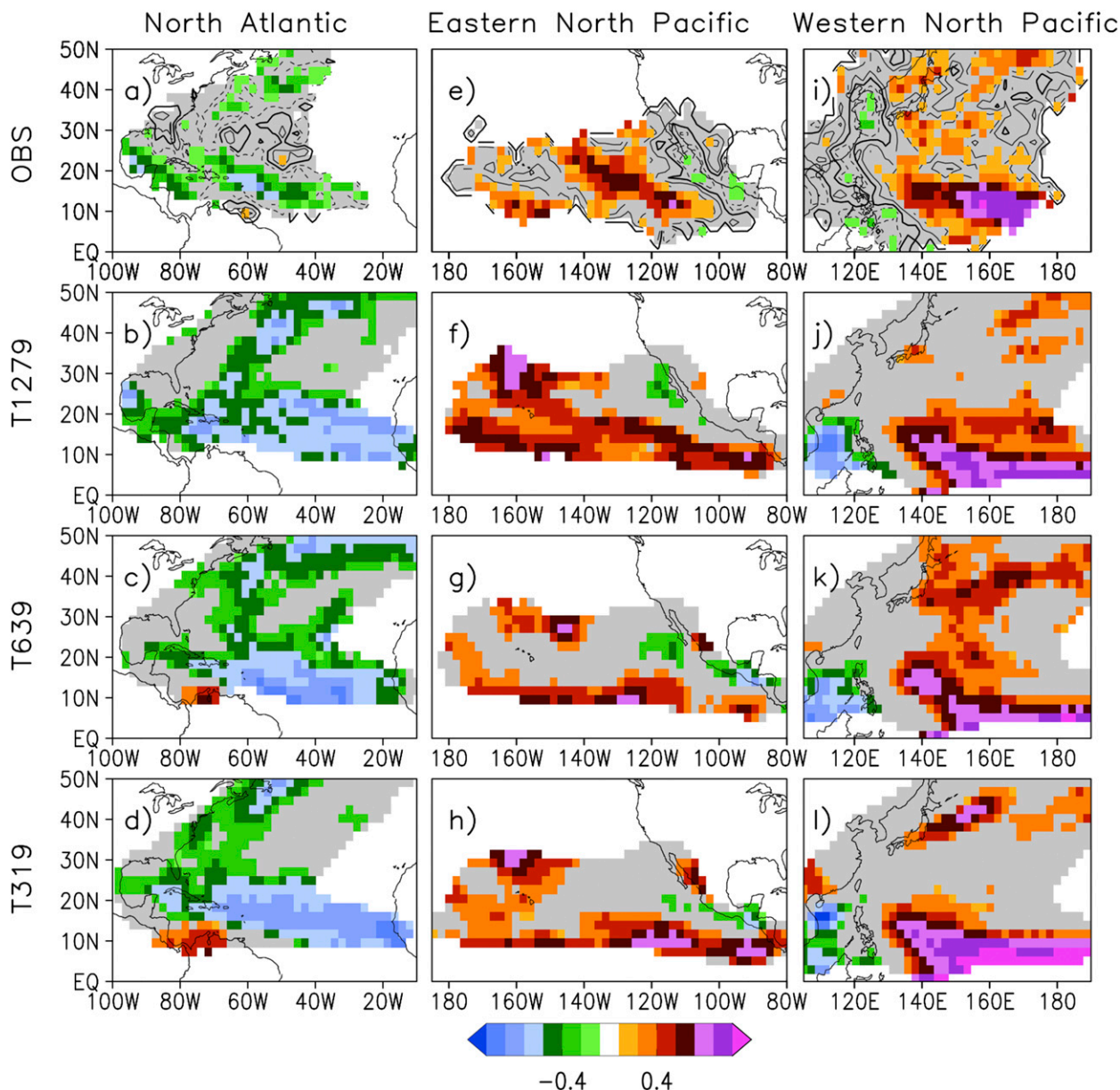


FIG. 9. Rank correlation between the MJJASON Niño-3.4 index and TC track densities as number densities per season per unit area equivalent to a 5° spherical cap for 1980–2011. Tracks from all ensemble members are used to compute seasonal mean predicted track densities. Results shown are for the (left) NA, (center) EP, and (right) WP for the (a),(e),(i) IBTrACS data; (b),(f),(j) T1279; (c),(g),(k) T639; and (d),(h),(l) T319 Minerva models. In (a), (e), and (i), the rank correlation is masked at the $p = 0.2$ level; nonsignificant values are shown by contours. Gray shading denotes the regions where the observed track density is nonzero for at least 25% of the years. In the remaining panels, the rank correlation is masked at the $p = 0.1$ level, and gray shading indicates the regions where the predicted track density is above a specified threshold for at least 25% of the years.

southern margin (less so in the T1279 model) and in the western part of the domain (Figs. 9f–h). Predictability in the central EP reflected in Figs. 8d–f could partly stem from high correlations of track density with the EP SST-REL over this region (not shown). ENSO also has a strong impact on the WP track density, with a displacement to the southeast and more northeastward

recurving tracks in El Niño years, and a shift to the northwest and largely straight tracks with landfalls in Southeast Asia during La Niña years (Fig. 9i; Camargo et al. 2007). All Minerva hindcasts reproduce these changes quite well, except that negative correlations in the South China Sea are too strong in the model (Figs. 9j–l). Positive correlations extending into the

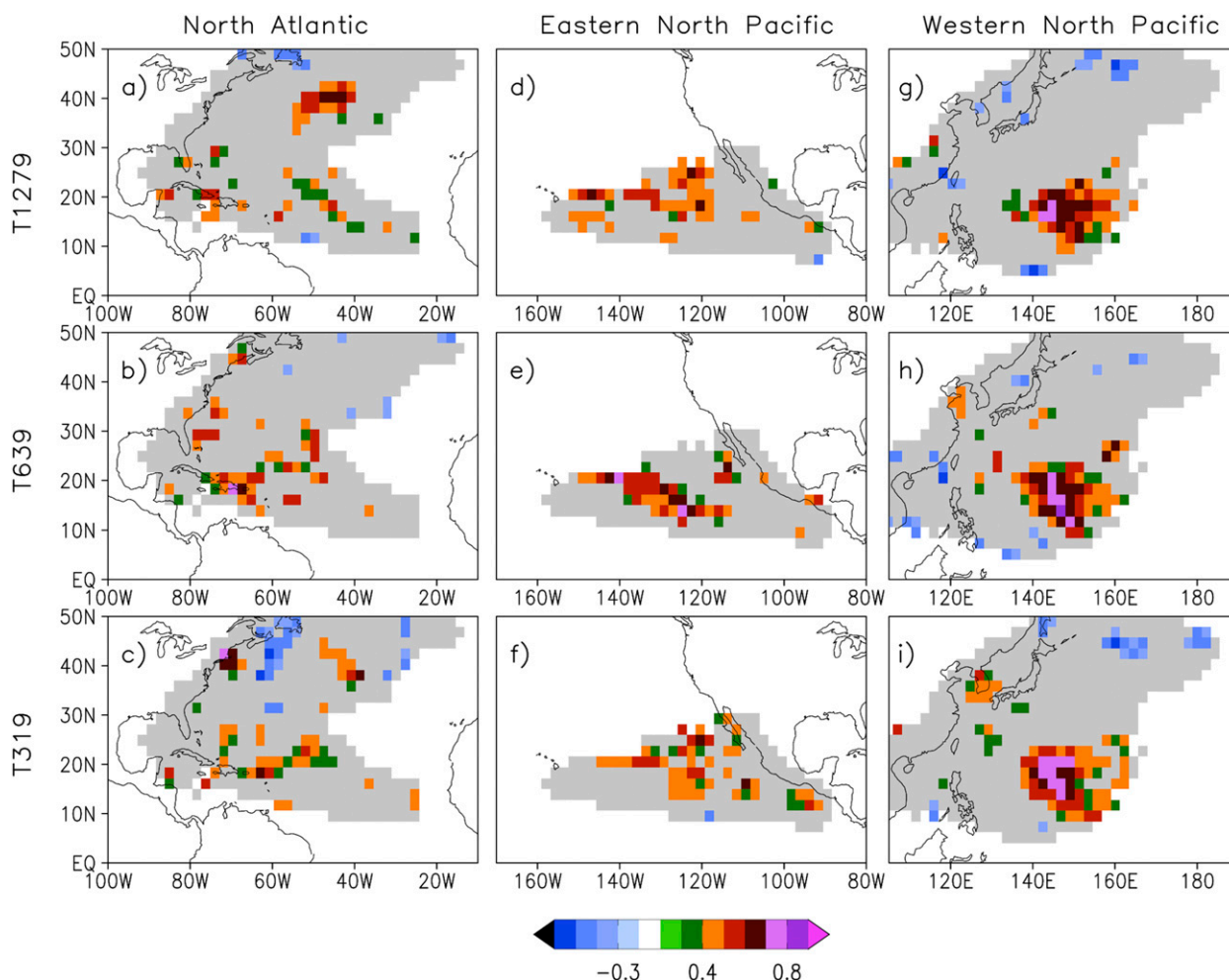


FIG. 10. Retrospective rank correlation between the MJJASON observed (IBTrACS) and predicted TC intensities as measured by 10-m wind speed and averaged over the area equivalent to a 5° spherical cap for 1990–2011. Tracks from all ensemble members are used to compute seasonal mean predicted intensities. Results shown are for the (a)–(c) NA, (d)–(f) EP, and (g)–(i) WP for the (top) T1279, (middle) T639, and (bottom) T319 Minerva models. Color shading denotes values statistically significant at a two-sided $p = 0.1$ level. Gray shading indicates the regions where the observed track density is nonzero for at least 25% of the years.

central Pacific are likely due to many simulated WP storms originating there (see Fig. 3), although tracking simulated storms through the earlier stages of the life cycle could also contribute to this result.

The ability to combine subbasin information with the seasonal TC intensity forecasts would enhance the value of regional TC activity predictions even further. This is a more challenging goal, since the outcome would be highly dependent on the skill of the regional track density (or a similar metric) forecasts discussed above. In addition, changes in storm intensity are driven by the ambient environmental conditions along the tracks (which can be modulated by broader-scale regional or remote modes of variability) as well as internal processes (e.g., Wang and Wu 2004). For these reasons, the following evaluation of the regional TC

intensity hindcasts should be regarded as mostly exploratory.

Figure 10 shows the retrospective skill of regional TC intensity forecasts assessed by means of Spearman rank correlation between the seasonal mean observed and predicted (all ensemble members) TC intensities expressed in terms of the 10-m wind speed and averaged over the area equivalent to a 5° spherical cap. In the NA, the area of significant skill is indeed rather limited and covers parts of the central tropical Atlantic, eastern Caribbean Sea, and off the southeastern coast of the United States (T1279 and T639 models; Figs. 10a–c). Over these regions, we also find significant correlations of the mean intensity and the Niño-3.4 index both in observations and hindcasts (not shown). In the EP, contiguous areas of significant skill are larger and mainly

found in the central and western parts of the basin (Figs. 10d–f). The WP also shows large areas of skill in the southeastern part of the domain (Figs. 10g–i). In the North Pacific, the locations of skillful intensity hindcasts appear to be directly linked to the regions of skillful track density hindcasts and could therefore be partly influenced by ENSO (see also Camargo et al. 2007; Camargo et al. 2008).

The above results show promise for regional intensity forecasting. A better understanding of the link between the regional track and intensity changes and other climate phenomena in addition to ENSO is needed. We do not find a clear dependence of the regional intensity hindcast skill (as defined above) on the model resolution.

6. Summary and conclusions

This study evaluates the skill of retrospective forecasts of the seasonal, basinwide, and regional tropical cyclone (TC) activity in Minerva, an experimental coupled prediction system integrated at high atmospheric resolutions ranging from 62 to 16 km. In the NA, the climatology of TC counts is quite realistic, largely as a result of the strong MDR genesis. At the highest resolution, the frequency of intense storms is still inadequate in all basins, except when intensity is assessed based on the minimum SLP, particularly in the NA. The EP and WP exhibit the familiar under- and overprediction of TCs, respectively. The northwest bias in the WP cyclogenesis, the prevalence of central Pacific storms, and relatively low intensities (compared to the NA) are strongly suggestive of the influence of the cold SST bias in the eastern equatorial Pacific.

Minerva demonstrates statistically significant skill in hindcasts of the seasonal mean TC frequency and ACE in all three basins, particularly for the period of 1990–2011, possibly as a result of improved ocean initialization. While the skill scores tend to be lower for the TC frequency, the ACE hindcasts are less reliable (underdispersed) except in the WP at the highest resolutions. Our analysis also suggests that the NA skill scores could be potentially much higher if it were not for an overly strong influence of the tropical Pacific variability on the NA MDR climate and the interannual and multidecadal time scales. On the other hand, the effects of aerosols and ozone, which are considered to be important drivers of the tropical NA climate variability (e.g., Evan et al. 2009, 2011; Emanuel et al. 2013), are not adequately represented in Minerva. The EP skill scores appear to be limited by weak interannual variability in the EP MDR. While our focus has been on the large-scale conditions, the EP cyclogenesis is also driven by wind surges,

AEWs, topographic effects, and ITCZ breakdowns (Camargo et al. 2010, and references therein). It is not clear to what degree these processes are resolved in Minerva and whether there is any predictability of the statistics of these events on seasonal time scales. Some of the highest scores are achieved in the WP, where Minerva demonstrates skill in simulating atmospheric response to SSTA, ocean–atmosphere interactions, and tropical teleconnections.

Higher atmospheric horizontal resolution improves skill scores for the ACE and, to a lesser extent, the TC frequency (where this effect is more basin specific), even though the influence of large-scale climate variations on these TC activity measures is largely independent of resolution changes in Minerva. The biggest gain occurs in transition from T319 to T639, while the differences between the T639 and T1279 models are generally not significant. This may indicate that the highest two resolutions are still too coarse to permit a qualitative improvement in TC genesis, while fine enough to better simulate intensification and intensity changes. Other possibilities include a suppression of sensitivity to atmospheric resolution by the convective parameterization or the relatively coarse ocean resolution.

Over broad areas of the NA and the North Pacific, Minerva exhibits significant skill in regional TC forecasts measured by retrospective rank correlations between the observed and predicted track densities. While most locations with skill are common to all resolutions, there are additional regions in the NA and EP (including land-adjacent areas) where significant correlations are achieved mostly by the T1279 model. It follows that, contrary to our resolution sensitivity analysis of the basinwide TC occurrence, there are advantages of further model refinement from 31 to 16 km for regional TC activity forecasting. Some of this improvement appears to stem from a more realistic relationship of the track density variations with ENSO at the highest resolution. Taking regional TC activity forecasting a step further, we assessed the feasibility of regional TC intensity forecasts. As expected, the areal coverage of significant skill in these forecasts is more limited, particularly in the NA, and does not show a clear dependence on the model resolution in contrast with the ACE forecasts. It is possible that intensity variations become more realistic at higher resolutions but do not occur in the same locations as in the observations.

A major source of predictability on seasonal time scales is the state of ENSO and the skill of models to predict ENSO is crucial for skillful prediction of TCs (e.g., Landsea 2000; Vitart and Stockdale 2001). Based on analyses by Zhu et al. (2015) and by us, we see that skillful prediction of the timing, phase, and magnitude of

the ENSO-related SSTAs may not be sufficient for this purpose. Rather, quality forecasts of the total SST distribution in the tropics, which determines the location and strength of the heating sources and therefore the response of the large-scale circulation, would be preferred. [The SSTA skill in Minerva is also relatively low in the western equatorial Pacific, which is a key region for generation of tropical teleconnections. A similar error in System 4 has been connected to a bias in near-equatorial winds in the western and central Pacific—a dominant factor in driving the coupled SST bias (Molteni et al. 2011).] The systematic model biases have been shown to strongly reduce the skill of the basinwide and regional TC activity forecasts in high-resolution CGCMs (Vecchi et al. 2014). In light of this, benefits of high resolution for seasonal TC activity forecasting, as explored in the current work, may be underestimated.

Acknowledgments. Funding of COLA for this study is provided by Grants from NSF (AGS-1338427), NOAA (NA09OAR4310058 and NA14OAR4310160), and NASA (NNX14AM19G). Computing resources on the Yellowstone supercomputer provided by the National Center for Atmospheric Research are also gratefully acknowledged.

REFERENCES

- Bell, G. D., and Coauthors, 2000: Climate assessment for 1999. *Bull. Amer. Meteor. Soc.*, **81**, 1328, doi:10.1175/1520-0477(2000)081<1328:CAF>2.3.CO;2.
- Bengtsson, L., K. Hodges, and M. Esch, 2007: Tropical cyclones in a T159 resolution global climate model: Comparison with observations and re-analyses. *Tellus*, **59A**, 396–416, doi:10.1111/j.1600-0870.2007.00236.x.
- Bretherton, C. S., M. Widmann, V. P. Dymnikov, J. M. Wallace, and I. Bladé, 1999: The effective number of spatial degrees of freedom of a time-varying field. *J. Climate*, **12**, 1990–2009, doi:10.1175/1520-0442(1999)012<1990:TENOSD>2.0.CO;2.
- Buizza, R., P. L. Houtekamer, G. Pellerin, Z. Toth, Y. Zhu, and M. Wei, 2005: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Mon. Wea. Rev.*, **133**, 1076–1097, doi:10.1175/MWR2905.1.
- Camargo, S. J., 2013: Tropical cyclones in high-resolution climate models. *U.S. CLIVAR Variations*, Vol. 11, No. 3, U.S. CLIVAR, Washington, DC, 4–11. [Available online at http://www.usclivar.org/sites/default/files/USCLIVAR_VARIATIONS_11_3_Fall2013.pdf.]
- , and A. H. Sobel, 2005: Western North Pacific tropical cyclone intensity and ENSO. *J. Climate*, **18**, 2996–3006, doi:10.1175/JCLI3457.1.
- , A. W. Robertson, S. J. Gaffney, P. Smyth, and M. Ghil, 2007: Cluster analysis of typhoon tracks. Part II: Large-scale circulation and ENSO. *J. Climate*, **20**, 3654–3676, doi:10.1175/JCLI4203.1.
- , —, A. G. Barnston, and M. Ghil, 2008: Clustering of eastern North Pacific tropical cyclone tracks: ENSO and MJO effects. *Geochim. Geophys. Geosyst.*, **9**, Q06V05, doi:10.1029/2007GC001861.
- , A. H. Sobel, A. G. Barnston, and P. J. Klotzbach, 2010: The influence of natural climate variability, and seasonal forecasts of tropical cyclone activity. *Global Perspectives on Tropical Cyclones, from Science to Mitigation*, 2nd ed. J. C. L. Chan and J. D. Kepert, Eds., World Scientific Series on Earth System Science in Asia, Vol. 4, World Scientific, 325–360.
- 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.
- Caron, L.-P., M. Boudreault, and C. L. Bruyere, 2015: Changes in large-scale controls of Atlantic tropical cyclone activity with the phases of the Atlantic multidecadal oscillation. *Climate Dyn.*, **44**, 1801–1821, doi:10.1007/s00382-014-2186-5.
- 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.
- , and —, 2014: New challenges and expectations of dynamical seasonal prediction of tropical cyclones. *31st Conf. on Hurricanes and Tropical Meteorology*, San Diego, CA, Amer. Meteor. Soc., 1A2. [Available online at <https://ams.confex.com/ams/31Hurr/webprogram/Paper244760.html>.]
- Chen, T.-C., S.-P. Weng, N. Yamazaki, and S. Kiehne, 1998: Interannual variation in the tropical cyclone formation over the western North Pacific. *Mon. Wea. Rev.*, **126**, 1080–1090, doi:10.1175/1520-0493(1998)126<1080:IVITTC>2.0.CO;2.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi:10.1002/qj.828.
- ECMWF, 2015: CY38R1 official IFS documentation. Accessed 4 January 2016. [Available online at <https://software.ecmwf.int/wiki/display/IFS/CY38R1+Official+IFS+Documentation>.]
- Emanuel, K. A., S. Solomon, D. Folini, S. Davis, and C. Cagnazzo, 2013: Influence of tropical tropopause layer cooling on Atlantic Hurricane activity. *J. Climate*, **26**, 2288–2301, doi:10.1175/JCLI-D-12-00242.1.
- Evan, A. T., D. J. Vimont, A. K. Heidinger, J. P. Kossin, and R. Bennartz, 2009: The role of aerosols in the evolution of tropical North Atlantic Ocean temperature anomalies. *Science*, **324**, 778–781, doi:10.1126/science.1167404.
- , G. R. Foltz, D. Zhang, and D. J. Vimont, 2011: Influence of African dust on ocean-atmosphere variability in the tropical Atlantic. *Nat. Geosci.*, **4**, 762–765, doi:10.1038/ngeo1276.
- Fink, A. H., J. M. Schrage, and S. Kotthaus, 2010: On the potential causes of the nonstationary correlations between West African precipitation and Atlantic hurricane activity. *J. Climate*, **23**, 5437–5456, doi:10.1175/2010JCLI3356.1.
- Gall, R., J. Franklin, F. Marks, E. N. Rappaport, and F. Toepfer, 2013: The Hurricane Forecast Improvement Project. *Bull. Amer. Meteor. Soc.*, **94**, 329–343, doi:10.1175/BAMS-D-12-00071.1.
- Goerss, J. S., C. R. Sampson, and J. M. Gross, 2004: A history of western North Pacific tropical cyclone track forecast skill. *Wea. Forecasting*, **19**, 633–638, doi:10.1175/1520-0434(2004)019<0633:AHOWNP>2.0.CO;2.
- Goldenberg, S. B., and L. J. Shapiro, 1996: Physical mechanisms for the association of El Niño and West African rainfall with

- Atlantic major hurricane activity. *J. Climate*, **9**, 1169–1187, doi:[10.1175/1520-0442\(1996\)009<1169:PMFTAO>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<1169:PMFTAO>2.0.CO;2).
- , C. W. Landsea, A. M. Mestas-Núñez, and W. M. Gray, 2001: The recent increase in Atlantic hurricane activity: Causes and implications. *Science*, **293**, 474–479, doi:[10.1126/science.1060040](https://doi.org/10.1126/science.1060040).
- Gray, W. M., 1979: Hurricanes: Their formation, structure and likely role in the tropical circulation. *Meteorology over the Tropical Oceans*, D. B. Shaw, Ed., Royal Meteorological Society, 155–218.
- Halperin, D. J., H. E. Fuelberg, R. E. Hart, J. H. Cossuth, P. Sura, and R. J. Pasch, 2013: An evaluation of tropical cyclone genesis forecasts from global numerical models. *Wea. Forecasting*, **28**, 1423–1445, doi:[10.1175/WAF-D-13-00008.1](https://doi.org/10.1175/WAF-D-13-00008.1).
- Hodges, K. I., 1994: A general method for tracking analysis and its application to meteorological data. *Mon. Wea. Rev.*, **122**, 2573–2586, doi:[10.1175/1520-0493\(1994\)122<2573:AGMFTA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1994)122<2573:AGMFTA>2.0.CO;2).
- , 1995: Feature tracking on the unit sphere. *Mon. Wea. Rev.*, **123**, 3458–3465, doi:[10.1175/1520-0493\(1995\)123<3458:FTOTUS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1995)123<3458:FTOTUS>2.0.CO;2).
- , 1999: Adaptive constraints for feature tracking. *Mon. Wea. Rev.*, **127**, 1362–1373, doi:[10.1175/1520-0493\(1999\)127<1362:ACFFT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<1362:ACFFT>2.0.CO;2).
- Kim, H.-M., P. J. Webster, and J. A. Curry, 2012: Seasonal prediction skill of ECMWF System 4 and NCEP CFSv2 retrospective forecast for the Northern Hemisphere winter. *Climate Dyn.*, **39**, 2957–2973, doi:[10.1007/s00382-012-1364-6](https://doi.org/10.1007/s00382-012-1364-6).
- Kinter, J. L., III, and Coauthors, 2013: Revolutionizing climate modeling with Project Athena: A multi-institutional, international collaboration. *Bull. Amer. Meteor. Soc.*, **94**, 231–245, doi:[10.1175/BAMS-D-11-00043.1](https://doi.org/10.1175/BAMS-D-11-00043.1).
- Knaff, J. A., 1997: Implications of summertime sea level pressure anomalies in the tropical Atlantic region. *J. Climate*, **10**, 789–804, doi:[10.1175/1520-0442\(1997\)010<0789:IOSSLP>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<0789:IOSSLP>2.0.CO;2).
- Knapp, K. R., M. C. Kruk, D. H. Levinson, H. J. Diamond, and C. J. Neumann, 2010: The International Best Track Archive for Climate Stewardship (IBTrACS). *Bull. Amer. Meteor. Soc.*, **91**, 363–376, doi:[10.1175/2009BAMS2755.1](https://doi.org/10.1175/2009BAMS2755.1).
- 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](https://doi.org/10.1175/2010JCLI3497.1).
- Landsea, C. W., 2000: El Niño–Southern Oscillation and the seasonal predictability of tropical cyclones. *El Niño and the Southern Oscillation: Multiscale Variability and Global and Regional Impacts*, H. F. Diaz and V. Markgraf, Eds., Cambridge University Press, 149–181.
- LaRow, T. E., 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](https://doi.org/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](https://doi.org/10.1029/2006GL027969).
- MacLachlan, C., and Coauthors, 2015: Global Seasonal forecast system version 5 (GloSea5): A high-resolution seasonal forecast system. *Quart. J. Roy. Meteor. Soc.*, **141**, 1072–1084, doi:[10.1002/qj.2396](https://doi.org/10.1002/qj.2396).
- Madec, G., 2008: NEMO reference manual, ocean dynamics component: NEMO-OPA. Version 3.0, Note du Pole de modélisation de l'Institut Pierre-Simon Laplace 27, 217 pp. [Available online at <http://www.nemo-ocean.eu/About-NEMO/Reference-manuals>.]
- Manganello, J. V., and Coauthors, 2012: Tropical cyclone climatology in a 10-km global atmospheric GCM: Toward weather-resolving climate modeling. *J. Climate*, **25**, 3867–3893, doi:[10.1175/JCLI-D-11-00346.1](https://doi.org/10.1175/JCLI-D-11-00346.1).
- , and Coauthors, 2014a: Future changes in the western North Pacific tropical cyclone activity projected by a multidecadal simulation with a 16-km global atmospheric GCM. *J. Climate*, **27**, 7622–7646, doi:[10.1175/JCLI-D-13-00678.1](https://doi.org/10.1175/JCLI-D-13-00678.1).
- , K. I. Hodges, and the Minerva Project Team, 2014b: Seasonal forecasts of the tropical cyclone activity in an ECMWF coupled operational prediction system. *Int. Conf. on Sub-seasonal to Seasonal Prediction*, College Park, MD, NOAA. [Available online at http://www.wmo.int/pages/prog/arep/wrwp/new/documents/07_Manganello.pdf.]
- Mei, W., S.-P. Xie, and M. Zhao, 2014: Variability of tropical cyclone track density in the North Atlantic: Observations and high-resolution simulations. *J. Climate*, **27**, 4797–4814, doi:[10.1175/JCLI-D-13-00587.1](https://doi.org/10.1175/JCLI-D-13-00587.1).
- Met Office, 2014: ENDGame: A new dynamical core for seamless atmospheric prediction. Met Office Doc., 27 pp. [Available online at http://www.metoffice.gov.uk/media/pdf/s/h/ENDGameGOVSci_v2.0.pdf.]
- Molteni, F., and Coauthors, 2011: The new ECMWF seasonal forecast system (System 4). ECMWF Tech. Memo, 656, 49 pp.
- Roberts, M. J., and Coauthors, 2015: Tropical cyclones in the UPSCALE ensemble of high-resolution global climate models. *J. Climate*, **28**, 574–596, doi:[10.1175/JCLI-D-14-00131.1](https://doi.org/10.1175/JCLI-D-14-00131.1).
- Saha, S., and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, doi:[10.1175/JCLI-D-12-00823.1](https://doi.org/10.1175/JCLI-D-12-00823.1).
- Serra, Y. L., G. N. Kiladis, and K. I. Hodges, 2010: Tracking and mean structure of easterly waves over the Intra-Americas Sea. *J. Climate*, **23**, 4823–4840, doi:[10.1175/2010JCLI3223.1](https://doi.org/10.1175/2010JCLI3223.1).
- Shukla, J., and Coauthors, 2000: Dynamical seasonal prediction. *Bull. Amer. Meteor. Soc.*, **81**, 2593–2606, doi:[10.1175/1520-0477\(2000\)081<2593:DSP>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<2593:DSP>2.3.CO;2).
- Strachan, J., P. L. Vidale, K. Hodges, M. Roberts, and M.-E. Demory, 2013: Investigating global tropical cyclone activity with a hierarchy of AGCMs: The role of model resolution. *J. Climate*, **26**, 133–152, doi:[10.1175/JCLI-D-12-00012.1](https://doi.org/10.1175/JCLI-D-12-00012.1).
- Strazzo, S., J. B. Elsner, T. LaRow, D. J. Halperin, and M. Zhao, 2013: Observed versus GCM-generated local tropical cyclone frequency: Comparisons using a spatial lattice. *J. Climate*, **26**, 8257–8268, doi:[10.1175/JCLI-D-12-00808.1](https://doi.org/10.1175/JCLI-D-12-00808.1).
- Vecchi, G. A., and G. Villarini, 2014: Next season's hurricanes. *Science*, **343**, 618–619, doi:[10.1126/science.1247759](https://doi.org/10.1126/science.1247759).
- , 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](https://doi.org/10.1175/2010MWR3499.1).
- , S. Fueglistaler, I. M. Held, T. R. Knutson, and M. Zhao, 2013: Impacts of atmospheric temperature trends on tropical cyclone activity. *J. Climate*, **26**, 3877–3891, doi:[10.1175/JCLI-D-12-00503.1](https://doi.org/10.1175/JCLI-D-12-00503.1).
- , 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](https://doi.org/10.1175/JCLI-D-14-00158.1).
- 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](https://doi.org/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](https://doi.org/10.1175/1520-0493(2001)129<2521:SFOTSU>2.0.CO;2).

- , J. L. Anderson, and W. F. Stern, 1997: Simulation of the interannual variability of tropical storm frequency in an ensemble of GCM integrations. *J. Climate*, **10**, 745–760, doi:[10.1175/1520-0442\(1997\)010<0745:SOIVOT>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<0745:SOIVOT>2.0.CO;2).
- , —, and —, 1999: Impact of large-scale circulation on tropical storm frequency, intensity, and location, simulated by an ensemble of GCM integrations. *J. Climate*, **12**, 3237–3254, doi:[10.1175/1520-0442\(1999\)012<3237:IOLSCO>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<3237:IOLSCO>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](https://doi.org/10.1029/2007GL030740).
- Walsh, K. J. E., 2008: The ability of climate models to generate tropical cyclones: Implications for prediction. *Climate Change Research Progress*, L. Peretz, Ed., Nova Publishers, 313–329.
- , M. Fiorino, C. W. Landsea, and K. L. McInnes, 2007: Objectively determined resolution-dependent threshold criteria for the detection of tropical cyclones in climate models and reanalyses. *J. Climate*, **20**, 2307–2314, doi:[10.1175/JCLI4074.1](https://doi.org/10.1175/JCLI4074.1).
- Wang, B., and J. C. L. Chan, 2002: How strong ENSO events affect tropical storm activity over the western North Pacific. *J. Climate*, **15**, 1643–1658, doi:[10.1175/1520-0442\(2002\)015<1643:HSEEAT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1643:HSEEAT>2.0.CO;2).
- 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](https://doi.org/10.1029/2009GL041469).
- Wang, Y., and C.-C. Wu, 2004: Current understanding of tropical cyclone structure and intensity changes—A review. *Meteor. Atmos. Phys.*, **87**, 257–278, doi:[10.1007/s00703-003-0055-6](https://doi.org/10.1007/s00703-003-0055-6).
- WMO, 2002: Standardised verification system (SVS) for long-range forecasts (LRF). Attachment II-9 to the manual on the GDPS, WMO 485, Vol. I, World Meteorological Organization, 21 pp.
- Wu, L., Z. Wen, R. Huang, and R. Wu, 2012: Possible linkage between the monsoon trough variability and the tropical cyclone activity over the western North Pacific. *Mon. Wea. Rev.*, **140**, 140–150, doi:[10.1175/MWR-D-11-00078.1](https://doi.org/10.1175/MWR-D-11-00078.1).
- Zhan, R., Y. Wang, and X. Lei, 2011: Contributions of ENSO and east Indian Ocean SSTA to the interannual variability of northwest Pacific tropical cyclone frequency. *J. Climate*, **24**, 509–521, doi:[10.1175/2010JCLI3808.1](https://doi.org/10.1175/2010JCLI3808.1).
- , —, and M. Wen, 2013: The SST gradient between the southwestern Pacific and the western Pacific warm pool: A new factor controlling the northwestern Pacific tropical cyclone genesis frequency. *J. Climate*, **26**, 2408–2415, doi:[10.1175/JCLI-D-12-00798.1](https://doi.org/10.1175/JCLI-D-12-00798.1).
- Zhao, M., and I. M. Held, 2012: TC-permitting GCM simulations of hurricane frequency response to sea surface temperature anomalies projected for the late-twenty-first century. *J. Climate*, **25**, 2995–3009, doi:[10.1175/JCLI-D-11-00313.1](https://doi.org/10.1175/JCLI-D-11-00313.1).
- , —, and G. A. Vecchi, 2010: Retrospective 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](https://doi.org/10.1175/2010MWR3366.1).
- Zhu, J., and Coauthors, 2015: ENSO prediction in Project Minerva: Sensitivity to atmospheric horizontal resolution and ensemble size. *J. Climate*, **28**, 2080–2095, doi:[10.1175/JCLI-D-14-00302.1](https://doi.org/10.1175/JCLI-D-14-00302.1).