Clim Dyn
DOI 10.1007/s00382-017-3891-7

NMME-based hybrid prediction of Atlantic hurricane season activity

Daniel S. Harnos[1] · Jae-Kyung E. Schemm[1] · Hui Wang[1] · Christina A. Finan[1,2]

Received: 27 December 2016 / Accepted: 28 August 2017

Abstract A hybrid dynamical–statistical model is pursued for prediction of Atlantic seasonal hurricane activity driven by output of the North American Multimodel Ensemble (NMME). This is an updated version of a proven multiple linear regression method conditioned on forecast vertical wind shear from the Climate Forecast System and observed sea surface temperatures (SSTs). The method pursued for prediction utilizes August–October (ASO) Main Development Region (MDR; 10–20°N, 20–80°W) vertical wind shear and observed North Atlantic (NATL; 55–65°N, 30–60°W) SST averaged over the 3 months preceding the forecast in conjunction with the full hurricane climatology. NMME forecasts improve upon representations relative to individual members. The NMME multi-model mean better reproduces vertical wind shear distributions over the MDR and captures the observed relationships between SST and vertical wind shear with hurricane trend and interannual variability despite occasionally poor reproductions by individual members. Cross-validation reveals the multi-model average of the hybrid model outputs from the individual NMME members yields forecast errors 10–30% less than the individual members, while correlations with observed hurricane-related activity typically improve. The NMME methodology is shown to be competitive with official outlooks from Colorado State University and the National Oceanic and Atmospheric Administration over recent years.

Keywords Tropical cyclone · Hurricane · Typhoon · North American Multimodel Ensemble · Interannual variability · Seasonal prediction

1 Introduction

Hurricanes threaten coastal interests throughout the vicinity of the North Atlantic Ocean. Since 2005, the United States has averaged one hurricane-related event costing in excess of $1 billion per year.[1] Population shifts, increasing property values in coastal regions, and income growth have led to increased hurricane vulnerability (Pielke Jr. et al. 2008; Mendelsohn et al. 2012; Peduzzi et al. 2012; Willoughby 2012), placing further scrutiny on improved understanding and anticipation of hurricane activity. Furthermore, the last two decades have seen increased hurricane activity relative to historical records. Attribution of the recent trend appears to be a mix of internal variability and anthropogenic influences while direct attribution of each role in observational records remains difficult (e.g. Knutson et al. 2010 and references therein). Despite the recent hurricane trend increases, considerable interannual variability remains, as in 2005 (15

This paper is a contribution to the special collection on the North American Multi-Model Ensemble (NMME) seasonal prediction experiment. The special collection focuses on documenting the use of the NMME system database for research ranging from predictability studies, to multi-model prediction evaluation and diagnostics, to emerging applications of climate predictability for subseasonal to seasonal predictions. This special issue is coordinated by Annarita Mariotti (NOAA), Heather Archambault (NOAA), Jin Huang (NOAA), Ben Kirtman (University of Miami) and Gabriele Villarini (University of Iowa).

✉ Daniel S. Harnos
Daniel.Harnos@noaa.gov

[1] Climate Prediction Center, NCEP/NWS/NOAA, College Park, MD, USA

[2] Innovim LLC, Greenbelt, MD, USA

[1] https://www.ncdc.noaa.gov/billions/time-series.

Published online: 13 September 2017

hurricanes) to 2006 (5 hurricanes) or 2009 (3 hurricanes) to 2010 (12 hurricanes).

The National Oceanic and Atmospheric Administration (NOAA) has been operationally issuing Hurricane Season Outlooks (HSO) for the Atlantic basin since 1998 and for the Eastern North Pacific basin since 2005. These HSO, issued first in May and then revised in August, are intended to draw attention and increase awareness regarding the forthcoming hurricane seasons. Numerous other public and private entities also provide seasonal hurricane forecasts for the Atlantic, including the first by Gray (1984a, b) and Colorado State University through more recently Davis et al. (2015) and the University of Arizona. The foundational relationships used for many of these forecasts are lagged relations between the observed atmospheric and oceanic states preceding the hurricane season to project the forthcoming season. These approaches are often conditioned on the state of the El Niño-Southern Oscillation (ENSO) and/or Atlantic Multidecadal Oscillation (AMO; Schlesinger and Ramankutty 1994). ENSO has been shown on interannual time scales to alter the vertical wind shear distribution over the Atlantic (e.g. Gray 1984a; Goldenberg and Shapiro 1996; Bell and Chelliah 2006; Wang et al. 2014) with reduced (increased) shear supporting (inhibiting) hurricane activity. The AMO is linked to variations in the oceanic currents of the Atlantic that alter the sea surface temperature (SST) distributions on decadal time scales with warmer (cooler) SSTs tied to increased (decreased) hurricane activity (e.g. Landsea et al. 1999, Enfield et al. 2001; Goldenberg et al. 2001; Bell and Chelliah 2006). The AMO's influence on hurricane activity appears tied to excitation of the Atlantic Meridional Mode (Vimont and Kossin 2007; Kossin and Vimont 2007) that alters the SST and vertical wind shear patterns throughout the tropical Atlantic.

Alternatives exist to hurricane seasonal activity predictive methods conditioned on preseason observed atmospheric and oceanic states. One such practice is through dynamical prediction of hurricane activity by counting individual events in general circulation models (GCMs) as by Vitart and Stockdale (2001), Vitart et al. (2007), and Schemm and Long (2014). Other methods utilize a hybrid framework, whereby GCM forecasts are used to train statistical models for hurricane activity predictions (e.g. Wang et al. 2009a; Vecchi et al. 2011; Villarini and Vecchi 2013; Li et al. 2013; Davis et al. 2015; Choi et al. 2015; Kim et al. 2017). Hybrid approaches can leverage the prior constraint of historical conditions and observed hurricane activity as in a purely statistical model, while also considering potential changes in the climatic state over the lead period that methods conditioned solely on observations may miss. Conversely, hybrid models can suffer if predictability from the associated GCM is limited relative to the realized atmospheric or oceanic state.

One approach towards improving single GCM predictability is through the use of multi-model ensembles. These methods aid in the conveyance of forecast uncertainty (e.g. Palmer et al. 2000; Kirtman 2003; DeWitt 2005; Min et al. 2009), while multi-model means have been shown to enhance predictability relative to the individual GCMs (e.g. Krishnamurti et al. 1999; Wang et al. 2009b; Becker et al. 2014). Accordingly, predictive methods for hurricane activity utilizing a multi-model approach may be desirable due to improved uncertainty conveyance for seasonal activity in addition to potential greater accuracy.

Here we explore applicability of the hybrid model of Wang et al. (2009a) based upon an empirical linear regression model forced with vertical wind shear and SST relationships adapted to inputs from the North American Multi-model Ensemble (NMME; Kirtman et al. 2014) for prediction of North Atlantic hurricane activity. Through such an approach the performance of the original model based on the National Centers for Environmental Prediction (NCEP) Climate Forecast System (CFS) model can be compared to other NMME coupled GCMs (CGCMs) and the multi-model mean. These predictions are reliant upon the GCMs ability to accurately forecast vertical wind shear and SST months in advance, thereby also necessitating an investigation of NMME member predictive skill for these variables. The end result strives to use the hybrid model output for producing deterministic and probabilistic forecast guidance that is then aggregated with other forecast tools to produce the NOAA HSO. Section 2 overviews the datasets utilized and the development of the prediction methodology. Predictive skill for NMME vertical wind shear and SST is quantified in Sect. 3. Relationships between observed hurricane activity and vertical wind shear and SST are examined in Sect. 4. Section 5 evaluates the retrospective predictive skill of the NMME-based hybrid model, while Sect. 6 compares the NMME methodology to the official seasonal outlooks from Colorado State University (CSU) and NOAA. Final discussions and conclusions are presented in Sect. 7.

2 Data and methods

The CGCM hindcast (retrospective forecast) data is taken from the NMME Phase II Archive hosted by the National Center for Atmospheric Research (NCAR).[2] Wind forecasts are not a mandatory published NMME variable; therefore only models that voluntarily include this information are included in this study. These models, their ensemble methodologies, and associated references are described in Table 1. The CGCM data is output on the 1° NMME

[2] http://www.earthsystemgrid.org/search.html?Project=NMME.

Table 1 NMME CGCM members and associated information

CGCM name	Operating center	Ensemble size	Ensemble arrangement	References
CanCM3	Canadian Meteorological Centre	10	All first of the month 0000 UTC	Merryfield et al. (2013)
CanCM4	Canadian Meteorological Centre	10	All first of the month 0000 UTC	Merryfield et al. (2013)
CCSM4	National Center for Atmospheric Research	10	All first of the month 0000 UTC	Kirtman et al. (in preparation)
CFSv2	National Centers for Environmental Prediction	12	Four members (0000, 0600, 1200, 1800 UTC) every fifth day. Last three runs prior to the 8th day of initial month are used	Saha et al. (2014)

CanCM3 and CanCM4 are combined here into the 20 ensemble member CanCM34

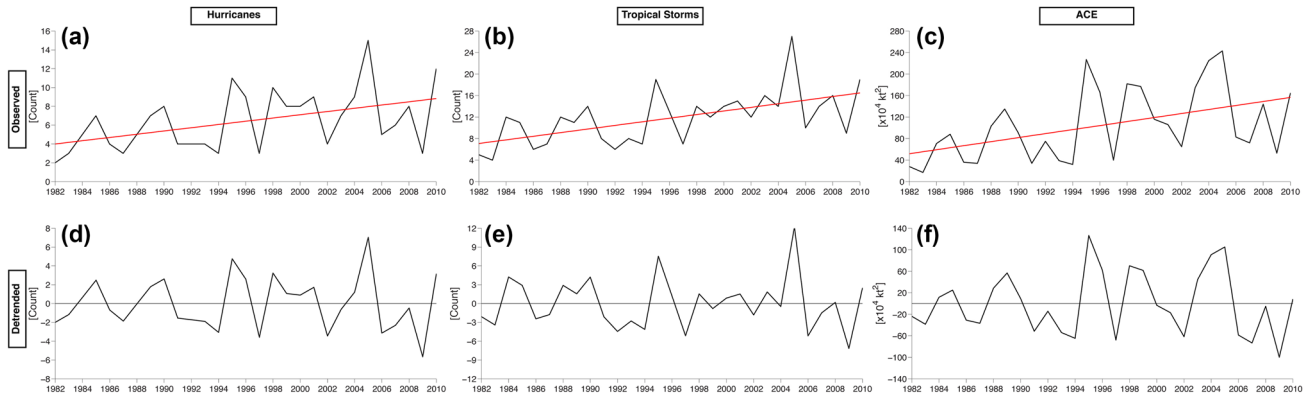


Fig. 1 Time series of the observed annual North Atlantic hurricanes (**a**), tropical storms (**b**), and ACE (**c**) during 1982–2010 (*black line*), with trend (*red line*) overlaid. **d**–**f** are as in **a**–**c** but with trend removed from observations

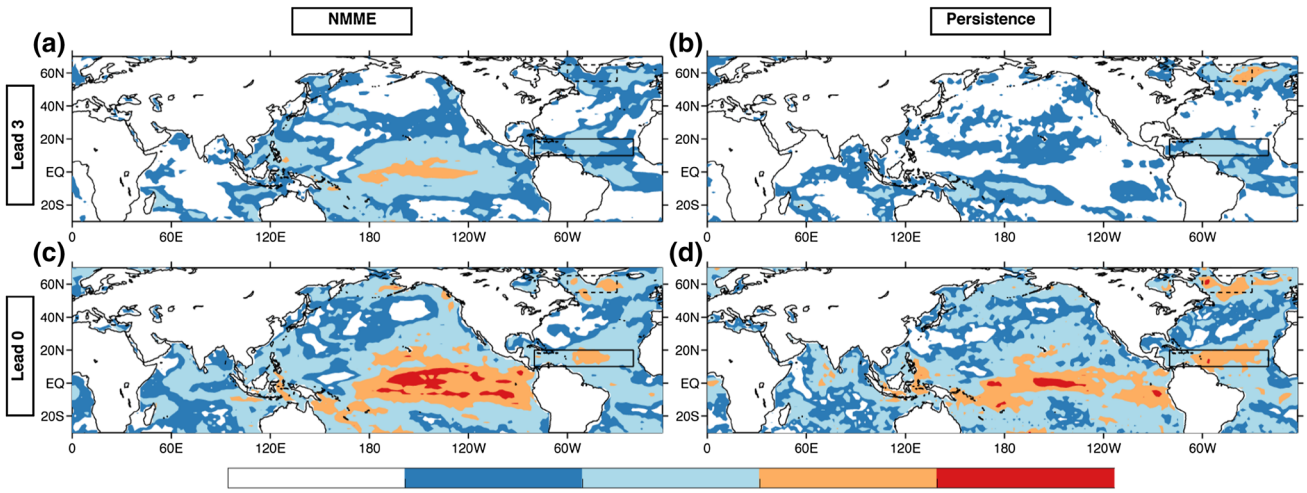
grid, with the exception of the CFSv2 on a 2.5° grid. In an effort to limit the influence of the highly similar CanCM3 and CanCM4 models (Merryfield et al. 2013) that would otherwise comprise half of the NMME average solution, both models' ten ensemble members are combined into a 20-member CanCM34.[3] All models utilize a common hindcast period of 1982–2010 with initial conditions taken from each month between April (3-month lead, hereafter referred as "lead 3") and July (0-month lead, referred as "lead 0"). Unless otherwise mentioned, all subsequent analyses and anomalies are taken relative to the 1982–2010 period.

In evaluating the NMME system performance, multiple CGCMs are evaluated individually before eventual aggregation into the NMME mean. Hindcast vertical wind shear is taken as the difference of zonal wind (u) between 200 and 850 hPa (u_{200}–u_{850}). Wind shear is evaluated over the months of August–October (ASO) due to this period accounting for approximately 90% of historical hurricane activity, as in numerous prior studies (e.g. Bell and Chelliah 2006; Knutson et al. 2007), despite the official Atlantic hurricane season spanning 1 June through 30 November. ASO seasonal mean wind shear is obtained via averaging the three individual monthly mean values. Forecast u values at leads 3 and 0 during 2011–2016 are also utilized, with identical treatment to the hindcasts with the exception of the CFSv2 forecasts being taken across the first 8 days of the month instead of every fifth day (32 members).

Observed SST is taken from the NOAA Optimum Interpolation SST (OISST) Version 2 (Reynolds et al. 2002) on a 1° grid for use as a possible predictor. For subsequent comparison, reanalysis SST and u data are taken from the NCEP CFS Reanalysis (CFSR; Saha et al. 2010) on a 2.5° grid. ASO means are generated by averaging across those respective months.

Retrospective hurricane data are taken from the NOAA Hurricane Best Track Database (HURDAT; Landsea et al. 2004). Hurricane-related information is referred to as "Atlantic" throughout this document, but is intended to only describe events within the northern hemisphere portion of the basin. Predictands in the hybrid model include

[3] Hindcast MDR shear values at lead 0 between CanCM3 and CanCM4 were correlated at 0.95, otherwise stated ≥90% of the variance of the individual CGCMs could be retained by averaging the two together while limiting their combined influence on the NMME mean.



		CanCM34	CCSM4	CFSv2	NMME	Persistence
MDR	Lead 3	0.71	0.71	0.50	0.67	0.71
	Lead 0	0.85	0.91	0.62	0.83	0.90
NATL	Lead 3	0.79	0.10	0.47	0.64	0.85
	Lead 0	0.84	0.79	0.61	0.84	0.90

Fig. 2 Spatial distributions of anomaly correlations for 1982–2010 between observed and hindcasts initialized in April (lead 3) with ASO SSTs for NMME (**a**) and persisted CFSR SST anomaly values from April through ASO (**b**). **c** and **d** are as in **a** and **b** except using hindcasts initialized in and persisted from July (lead 0), respectively. Correlations are only shaded at ≥0.4 and when statistically significant above the 1% level using a two-sided Monte Carlo test. *Boxes* indicating the MDR (10–20°N, 20–80°W) and NATL (55–65°N, 30–60°W) are shown with *solid black* and *dashed black lines*, respectively. Correlations for the MDR and NATL are listed below the figure and include the NMME member CGCMs

number of hurricanes, number of tropical storms, number of major hurricanes, and accumulated cyclone energy (ACE) as a percentage of the 1981–2010 median (89.6×10^4 kt^2). Here results for major hurricanes are omitted for brevity. Figure 1 shows observed hurricane count, tropical storm count, and ACE for the Atlantic for 1982–2010 along with their linear trend, detrended activity, and anomalous activity. An increasing trend in each predictand is seen over time (Fig. 1a–c), and with relatively inactive (1982–1994) and active (1995–2010) eras. These periods are closely tied to the AMO (Goldenberg et al. 2001) and Atlantic Multidecadal Mode (Vimont and Kossin 2007; Patricola et al. 2014). Standard deviations of hurricane count, tropical storms, and ACE are 3.2, 5.0, and 67.4×10^4 kt^2, respectively, while the corresponding detrended time series (Fig. 1d–f) exhibit standard deviations of 2.8, 4.1, and 59.4×10^4 kt^2, respectively. Aforementioned values support interannual variability dominating the 1982–2010 period, accounting from 67 to 79% of the variance depending upon the predictand.

In developing the empirical regression model, anomalies of SST and wind shear are utilized. For cross-validation over the hindcast period (1982–2010), anomalies are generated by omitting the target year and building the climatology based on the remaining 28 years of hindcast data, while for the 2011–2016 forecasts the anomalies are taken relative to the climatology from the full 1982–2010 period. A multiple linear regression analysis is then performed on the anomalous SST and vertical wind shear values averaged over specified regions, defined in Sect. 4. These regression coefficients are then applied to the hindcast (or forecast) anomalies of each CGCM ensemble member for the target year to predict the interannual component of hurricane count and ACE. These interannual values are then added back to their respective long-term mean values over a climatological period, yielding the final forecast numbers. In defining climatology, the full 29-year period is utilized with average seasonal activity of 6.4 hurricanes, 12.0 tropical storms, and 104.4×10^4 kt^2 of ACE (117% of median).

3 CGCM performance for SST and vertical wind shear prediction

With accurate depiction of SST and wind shear being the foundation of the hybrid model, it is necessary to evaluate the representation of these fields by the NMME models first. Anomaly correlation (AC) scores of the hindcast ASO SST for the NMME models with the CFSR SST as observations at leads 3 and 0 are shown in Fig. 2a, c. Also included are the scores of SST persistence for initial conditions of lead

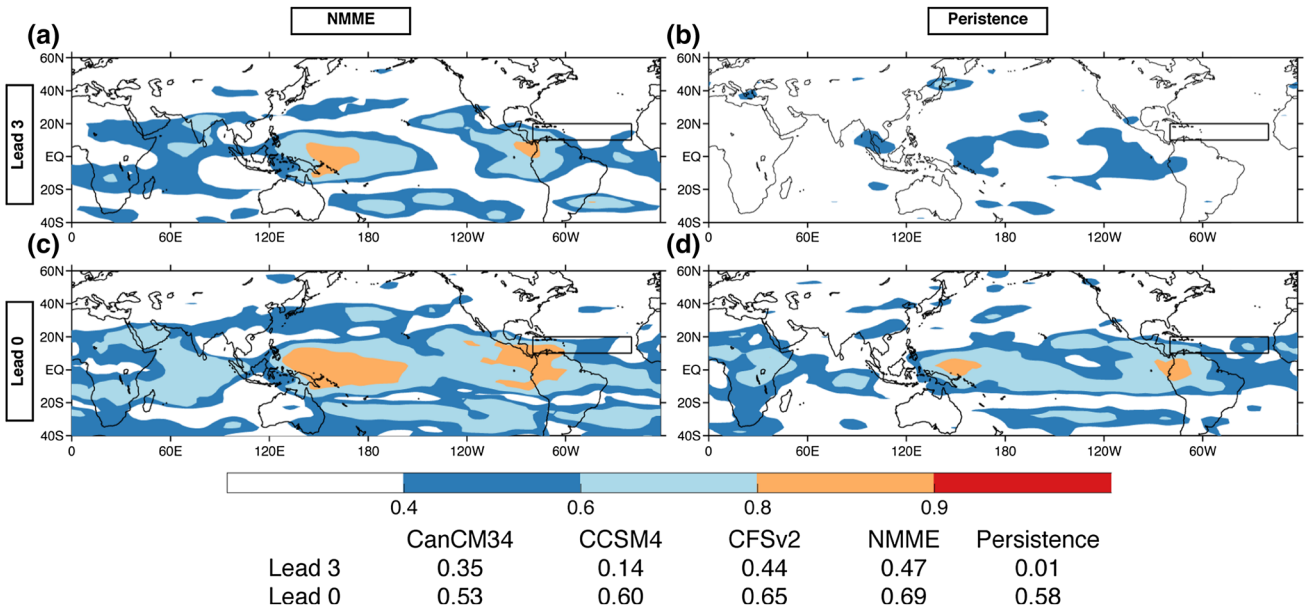


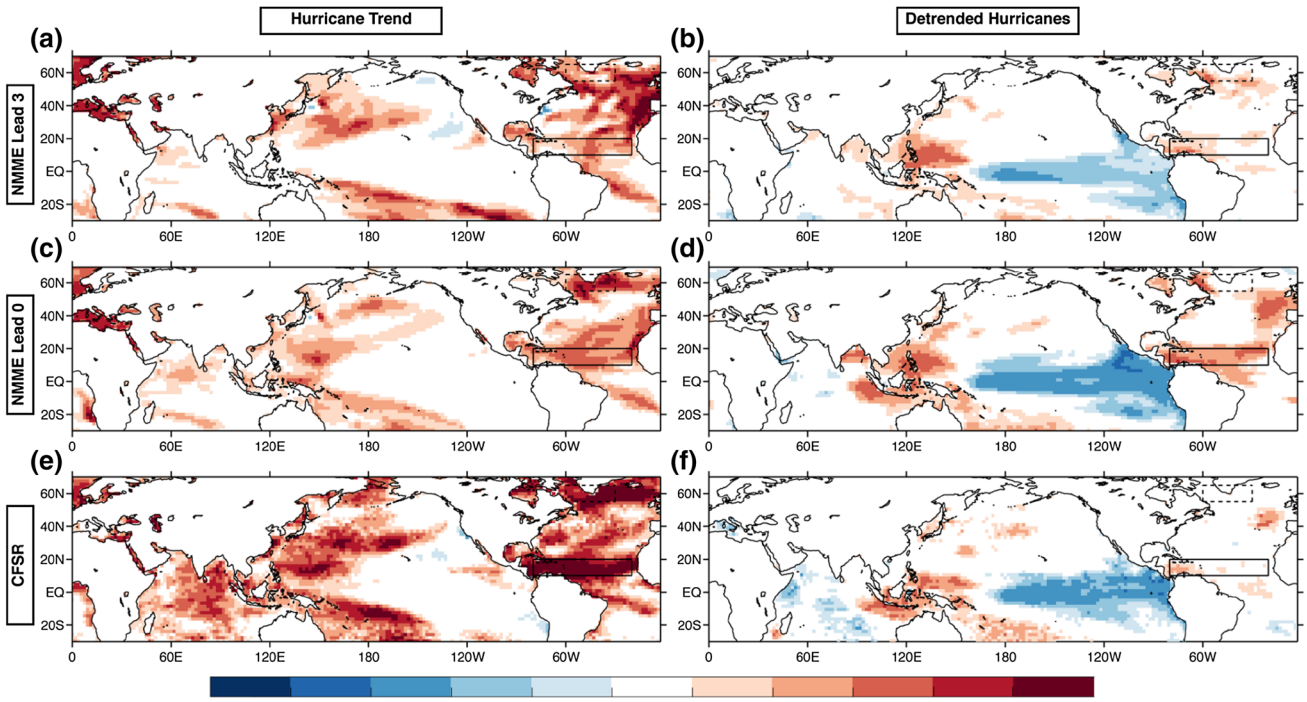
Fig. 3 As in Fig. 2, but for vertical wind shear. Correlations for the MDR are listed below the figure and include the NMME member CGCMs

3 (April) and lead 0 (July) respectively (Fig. 2b, d). ASO SSTs are reproduced best throughout the tropical Pacific, particularly extending east of the date line where El Niño events typically manifest. NMME SST forecasts typically outperform persistence, with notable exceptions south of Greenland and just north of the equator in the Atlantic. The former North Atlantic region (NATL; 55–65°N, 30–60°W), is associated with deep water formation within the thermohaline circulation (Koltermann et al. 1999), and is linked to hurricane activity through the AMO (Goldenberg et al. 2001). The latter region is the Atlantic's main development region (MDR; 10–20°N, 20–80°W), where the majority of the basin's hurricanes form. Across each of these regions at both leads the persisted SST anomalies prove superior to the NMME forecasts. Some of the poor performance in the NMME for these focal regions appears due to the CFSv2 at both leads while CCSM4 lead 3 forecasts in the NATL are notably poor with an AC score of 0.1. Marginal skill decreases are generally seen in both NMME forecast and persisted SSTs as a function of lead. Barnston and Tippett (2013) note a shift in Central Pacific SST bias characteristics post-1999 in CFSv2 forecasts that is attributed to a spurious signal from the CGCM being initialized with CFSR conditions. CCSM4 is also initialized with CFSR, with this bias after 1999 potentially impacting both CGCM forecasts as well as the CFSR being used as reanalysis. As will be shown later, forecast SSTs are not used as predictors in the hybrid model so that these issues would not affect the prediction results.

AC scores of vertical wind shear corresponding to lead 3 and lead 0 are shown in Fig. 3a, c. Vertical wind shear AC scores are relatively lower compared to those of SST (Fig. 2), with largest values focused within the equatorial Pacific where values exceed 0.8 east on Indonesia and south of Panama at both leads. Persisted wind shear values at both leads are generally inferior to the NMME predictions, apparent in both the extent of high AC values and their magnitudes. Of particular interest is the MDR where NMME yields AC scores of 0.47 at lead 3 and 0.69 at lead 0 while corresponding persistence values are 0.01 and 0.58. NMME AC scores across the MDR also exceed the comparable values of individual member CGCMs, demonstrating the improved ability of NMME to reproduce the observed wind shear distribution for the MDR relative to individual models. Relative to SST, wind shear predictive skill over the MDR is lower and decreases more rapidly as function of lead-time.

4 Atlantic hurricane activity relationships with SST and vertical wind shear

To assess how closely SST and vertical wind shear are associated relative to long-term trend and interannual variability in Atlantic hurricane activity, correlations are computed between the ASO environmental conditions and the time series of the hurricane linear trend (red line in Fig. 1a) and detrended interannual variability (Fig. 1d). Since hurricane counts are well correlated to tropical storm counts (0.91 overall; 0.90 for interannual variability) and ACE (0.90 overall; 0.87 for interannual variability), spatial correlations with environmental variables are only shown for



		CanCM34	CCSM4	CFSv2	NMME	CFSR
MDR T (D)	Lead 3	0.51 (0.19)	0.37 (0.29)	0.24 (0.23)	0.42 (0.25)	0.75 (0.25)
	Lead 0	0.47 (0.41)	0.62 (0.37)	0.23 (0.39)	0.49 (0.41)	
NATL T (D)	Lead 3	0.74 (0.18)	-0.02 (0.13)	0.26 (0.27)	0.47 (0.27)	0.75 (0.04)
	Lead 0	0.67 (0.20)	0.62 (0.19)	0.30 (0.24)	0.61 (0.24)	

Fig. 4 Spatial distributions of correlation of hindcast ASO SST for April initial conditions with long-term hurricane trend (*left column*) and detrended interannual variation (*right column*) for NMME at lead 3 (**a, b**), NMME at lead 0 (**c, d**), and CFSR observations during ASO (**e, f**). Correlations are only shaded at ≥|0.4| and when statistically significant above the 1% level using a two-sided Monte Carlo test. *Boxes* are the same as in Fig. 2. Correlations for the MDR and NATL with the hurricane trend (*red line* in Fig. 1a) and detrended time series (Fig. 1d) are listed below the figure without and with parentheses respectively, and include the NMME member CGCMs

hurricane activity in this section. Other predictands exhibit comparable results.

SST correlations with hurricane activity in terms of trend and interannual components are shown in Fig. 4. NMME exhibits broad regions of positive trend correlations across the Atlantic at both leads, indicative of the increasing trend in hurricane activity (Fig. 1a) being associated with warmer SSTs. Observed ASO SSTs (Fig. 4e) also exhibit these positive correlations, albeit at a greater magnitude than the hindcast projections with average values of 0.75 for the MDR and NATL. The NMME mean correlations with hurricane trend across the MDR and NATL are both below observations at each lead, with increased values at the shorter lead time. CanCM34 is the closest performing NMME member to observations for each lead and focal region, with the CFSv2 being the most disparate from the observed trend relationship but this cannot be wholly attributed to CFSR being used for initialization given the performance of CCSM4. Notable is CCSM4's exceptionally poor relationship with hurricane trend at lead 3 for the NATL, with a correlation of −0.02.

SST relationships with interannual hurricane variability (Fig. 4b, d, f) typically exhibit lesser correlations than those with hurricane trend. Positive correlations are again apparent across the MDR and NATL, linking increased SST values to enhanced hurricane activity. The NATL shows limited association with interannual hurricane activity, as observed correlations are only 0.04. Conversely the MDR exhibits an observed correlation of 0.25 with the detrended hurricane time series, identical to that seen in the NMME at lead 3. The NMME at lead 0 sees this value increase to 0.41. Greatest correlation magnitudes in NMME and observations with hurricane interannual variability are negative values spanning the Niño 3.4 and Niño 3 (5°S–5°N; 90–150°W) regions with peaks of ≤−0.5. These strong values are expected, given that anomalous warming (cooling) in the Central and Eastern Pacific associated with El Niño (La Niña) tends to reduce (enhance) Atlantic hurricane activity (Gray 1984a; Goldenberg and Shapiro 1996; Bell and Chelliah 2006; Wang et al. 2014).

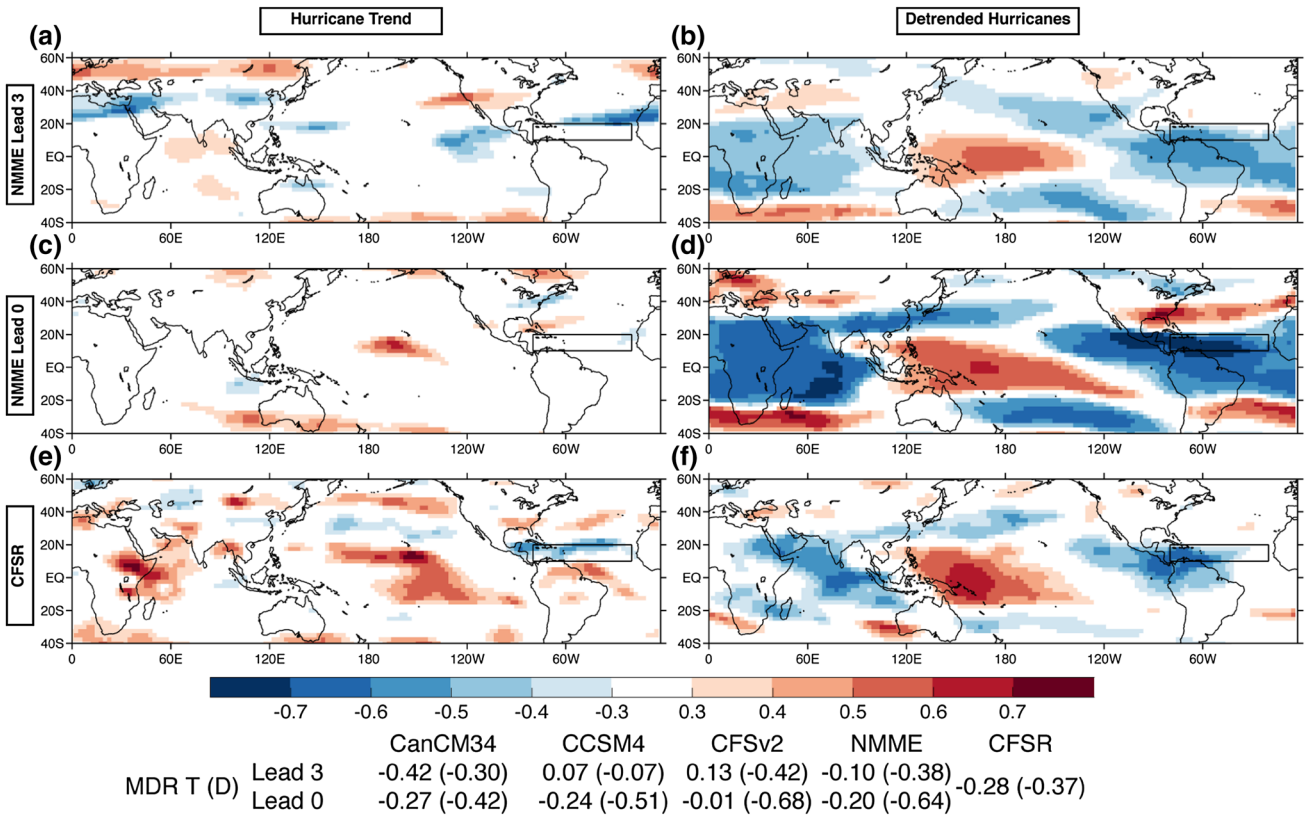


Fig. 5 As in Fig. 4, but for vertical wind shear. Correlations for the MDR with the hurricane trend (*red line* in Fig. 1a) and detrended time series (Fig. 1d) are listed below the figure without and with parentheses respectively, and include the NMME member CGCMs

Correlation of vertical wind shear anomalies with hurricane trend and detrended hurricane activity is shown in Fig. 5. Little relationship is seen between vertical wind shear and hurricane trend across the MDR in NMME at either lead (Fig. 5a, c), despite CFSR average values of −0.28 implying reduced vertical wind shear playing a role in the increased hurricane activity since 1982. NMME's poor performance appears due to CCSM4 and CFSv2 at lead 3 when both get the correlation's sign wrong (0.07 and 0.13 respectively), while CFSv2's relationship is negligible at lead 0 (−0.01). One potential explanation for the lead 3 shear correlations with long-term trend from CCSM4 and CFSv2 possessing the wrong sign may be due to the ENSO predictability barrier that is most pronounced in boreal Spring (e.g. Barnston et al. 2012) at longer leads. ENSO activity did shift over the cross-validation period, as ASO observations in 1982–1994 and 1995–2010 show each period experienced 31% of years with El Niño conditions while the former (latter) period saw 8% (38%) La Niña conditions.[4] Given the knowledge that El Niño (La Niña) is less (more) conducive for Atlantic hurricane activity (e.g. Bell and Chelliah 2006), the increased La Niña frequency by a factor of five late in the validation period and poor predictions of the future ENSO state may be responsible for the lead 3 shear correlations with trend of CCSM4 and CFSv2.

Vertical wind shear correlations with hurricane interannual variability (Fig. 5b, d, f) exhibit greater MDR averaged consistency as at each lead all member CGCMs, the NMME mean, and CFSR observations exhibit negative correlations. CFSR indicates the observed relationship across the MDR to be −0.37, which is similar to the NMME lead 3 value (−0.38) that nearly doubles at lead 0 (−0.64). One consistent pattern between NMME and CFSR correlations with detrended hurricane activity is the negative values across the western Indian Ocean and eastern Pacific Ocean, while the equatorial region extending east from the maritime continent to near the dateline possesses positive correlations. These patterns reassert Walker circulation enhancement during La Niña events being conducive for increased Atlantic hurricane activity.

Despite inconsistencies by the NMME members in reproducing wind shear, interannual variability predictions for the MDR still appear skillful. This ability of NMME to reproduce vertical wind shear across the MDR is shown in

[4] http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml.

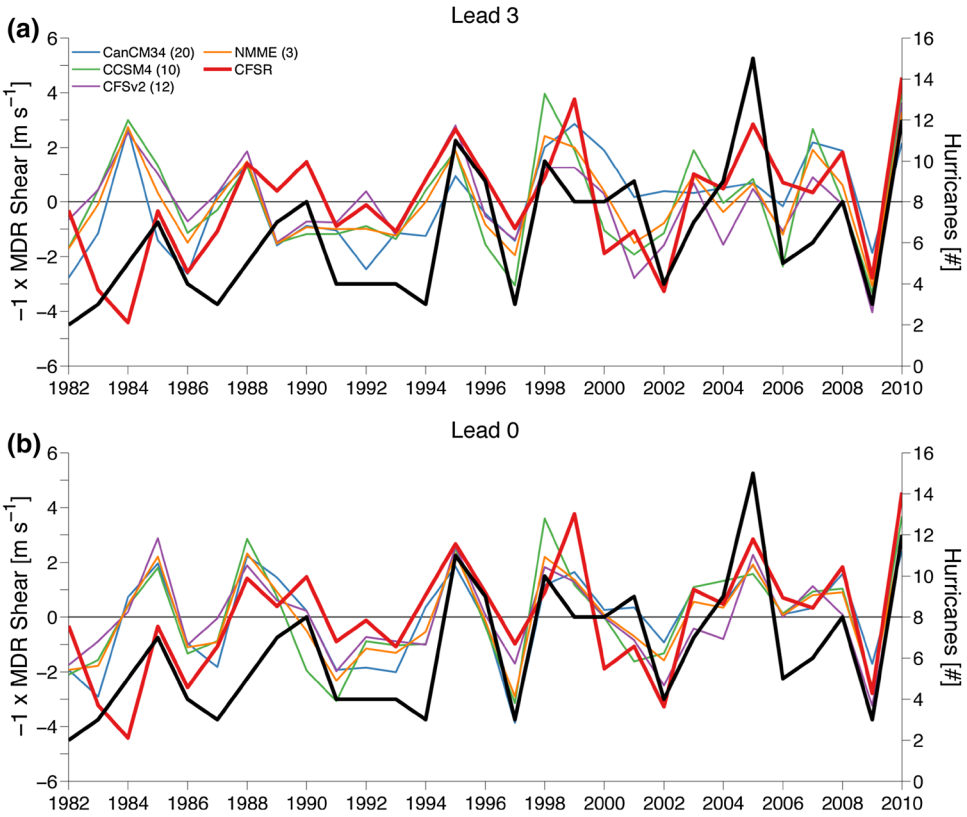


Fig. 6 Time series of the negative of ASO hindcast wind shear anomalies (1982–2010) for the MDR with lead 3 (**a**) and lead 0 (**b**) initial conditions compared to observed North Atlantic hurricane activity (*black line*)

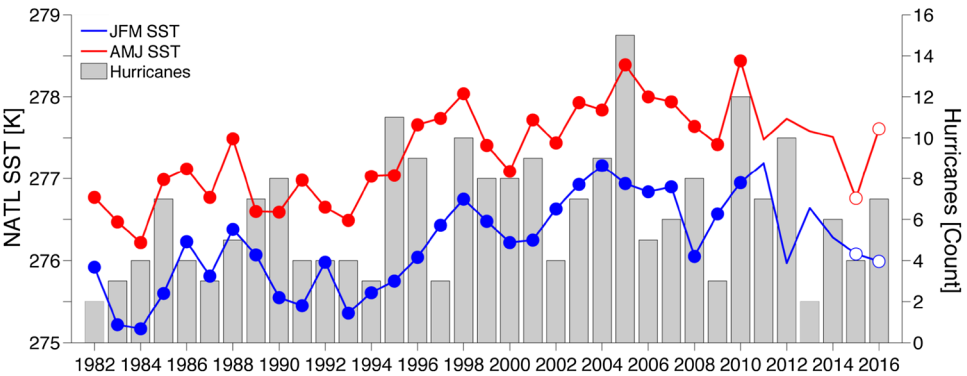


Fig. 7 Time series of observed preseason OISSTv2 SST within the NATL during JFM (*blue*) and AMJ (*red*) along with annual hurricane activity. *Filled circles* on the SST time series indicate the data was included in the hindcast period, while *open circles* are indicate usage during the 2015 and 2016 real-time forecast evaluation

Fig. 6 relative to observed CSFR shear. CGCM hindcast MDR shear anomalies tend to be of the same sign and have comparable magnitudes to observed CFSR values, with some notable exceptions (e.g. lead 3 in 1984). As expected, improvement is also generally seen from lead 3 to lead 0 with some exceptions (e.g. 1985) and NMME member projections also tend to cluster together, with few exceptions (e.g. CanCM34 at lead 3 for 1992). Observed hurricane activity is also shown in Fig. 6, with the inverse relationship between hurricanes (or tropical storms or ACE, not shown) and vertical wind shear readily apparent. The upward trend in NMME wind shear is not as clear as in CFSR, with this potentially responsible for the poor correlation between forecast wind shear and hurricane trend (Fig. 5a, c, e).

5 Prediction of Atlantic hurricane activity with NMME

In selecting predictors for the regression model, Wang et al. (2009a) serves as a starting point. That study revealed NATL SST and MDR shear were generally independent of one another, whereas SST in the Niño 3.4 region or MDR were highly correlated with MDR shear due to the ocean–atmosphere coupling associated with ENSO. Here, similar results are found with lead 0 NMME mean MDR SST and vertical wind shear correlated at a value of −0.53, while the corresponding Niño 3.4 SST and MDR vertical wind shear have correlations of 0.80 (SST in the tropical Pacific region from Wang et al. (2009a) exhibits a similar correlation of 0.79

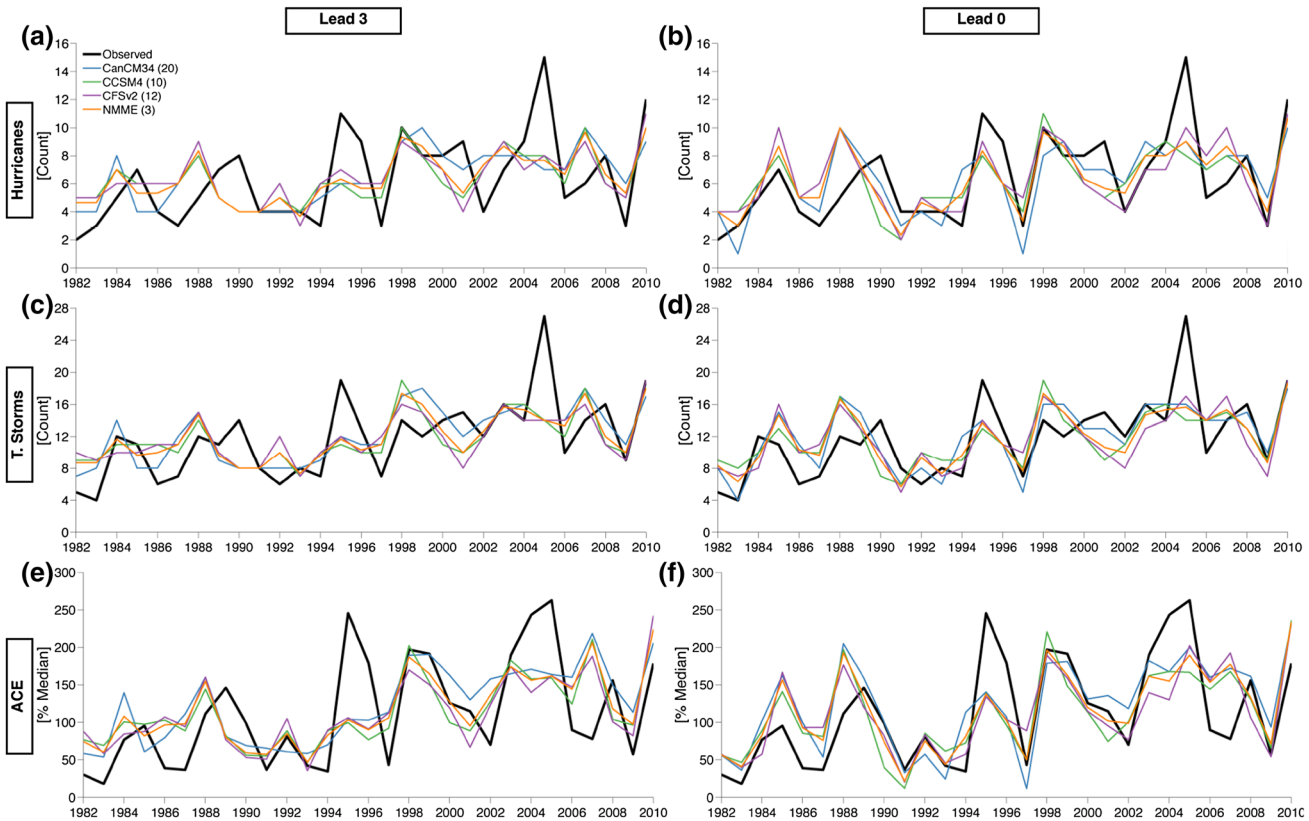


Fig. 8 Hindcasts of North Atlantic hurricane activity with predictors of observed preseason 3-month NATL SST and ASO hindcast MDR wind shear with April (**a**) and July (**b**) initial conditions. **c** and **d** are as in **a** and **b** but for a predictand of tropical storms, while **e** and **f** are for ACE

with NMME lead 0 MDR vertical wind shear). NATL SST is distinct from the tropical predictors with correlations <0.3 between NMME lead 0 MDR vertical wind shear and CFSR preseason NATL SST. Given the superior skill exhibited by persistence relative to NMME forecasts across the NATL, preseason SST observations may serve as a potential predictor. Figure 7 depicts January–March (JFM) and April–June (AMJ) averaged SSTs for the NATL along with hurricane activity over the associated year. The NATL region is closely associated with the AMO, with the change in AMO phase prior to the 1995 hurricane season apparent in the sudden increase in NATL SST values, and the subsequent period tied to increased hurricane activity (e.g. Trenberth and Shea 2006).

Given the lack of a codependence between NATL SST and MDR vertical wind shear, these regions are isolated for generating independent hurricane predictions. Anomaly correlations of ASO NATL SST for persisted values exceed those from NMME hindcasts (Fig. 2), while this region is highly correlated with the increasing trend in hurricane activity (Figs. 4, 7). Anomaly correlations of ASO MDR vertical wind shear meet or exceed persisted values for all NMME member CGCMs and the multi-model mean at leads 3 and 0 (Fig. 3), while these projections are highly correlated with interannual hurricane variability (Fig. 5).[5] As such, this methodology utilizes predictors of NATL observed SST averaged over the 3 months preceding the forecast month (i.e. January–March for lead 3 and April–June for lead 0) and CGCM hindcast MDR vertical wind shear during ASO.

The NMME member CGCMs are run through the 1982–2010 cross-validation period with the hybrid model to produce annual hurricane, tropical storm, and ACE predictions for the Atlantic. These individual model ensemble mean predictions are then averaged, without weighting to account for ensemble size, to produce the NMME mean predictions. Figure 8 shows the time series of predictions for each CGCM and the multi-model mean for hurricanes, tropical storms, and ACE. Underdispersive tendencies are apparent for every predictand at both leads, particularly

[5] The 1999 shift in bias characterstics from CFSR initialized models (CFSv2 and CCSM4) noted by Barnston and Tippett (2013) for Central Pacific SST was similarly evaluated for MDR vertical wind shear relative to the NCEP/NCAR reanalysis (Kalnay et al. 1996), and no distinctive shifts in bias characteristics were observed over 1982–2010 (not shown), particularly relative to CanCM3 and CanCM4, that are initialized with ERA-Interim (Dee et al. 2011).

Table 2 Correlations and RMSE (in parentheses) of hybrid model North Atlantic hindcast (1982–2010) predictions for the listed predictand with lead 3 and lead 0 initial conditions for NMME member CGCMs and NMME mean

Predictand	Lead	CanCM34	CCSM4	CFSv2	NMME	Trend
Hurricanes	Lead 3	0.44 (2.86)	0.46 (2.80)	0.46 (2.80)	0.51 (2.26)	0.43 (2.83)
	Lead 0	0.68 (2.30)	0.62 (2.47)	0.67 (2.36)	0.70 (1.69)	
Tropical storms	Lead 3	0.59 (4.05)	0.52 (4.21)	0.47 (4.36)	0.57 (3.11)	0.58 (3.97)
	Lead 0	0.70 (3.56)	0.57 (4.06)	0.64 (3.80)	0.66 (2.94)	
ACE	Lead 3	0.54 (55.99)	0.54 (55.89)	0.52 (56.94)	0.58 (46.83)	0.47 (57.87)
	Lead 0	0.72 (45.95)	0.70 (47.21)	0.67 (49.35)	0.73 (37.80)	

Hurricane and tropical storm predictions are a function of count, while ACE is a function of percentage of median. The last column denotes values with respect to observed linear trend as a predictor (i.e. the red lines in Fig. 1a–c)

for active hurricane seasons (i.e. 1995, 2005). Differences between the individual CGCMs and NMME mean in the predictions for individual seasons are all ≤3 hurricanes or tropical storms and ≤50% of median ACE at each lead, with typical differences of 1 hurricane, 2 tropical storms, and 10–20% median ACE.

Hindcast performance of the member CGCMs and multimodel mean predictions is quantified in Table 2 via correlation and root-mean-square error (RMSE) of predicted hurricane, tropical storm, and ACE activity respectively with observations. As expected from the greater anomaly correlations (Figs. 2, 3), improved correlations and RMSE are seen in each predictand at shorter forecast leads. Correlations of observed and predicted hurricane activity are greatest in the member CGCMs at lead 3 for the CCSM4 and CFSv2 (0.46) and lead 0 from the CanCM34 (0.68) while the NMME has correlations of 0.51 and 0.70 at leads 3 and 0 respectively, outperforming all CGCMs at both leads. The NMME yields the lowest hurricane RMSE values with reductions of 20–30% relative to the best performing member. Also as a point of comparison in Table 2 are correlation and RMSE of a prediction solely utilizing the long-term trend (i.e. red lines from Fig. 1a–c) as a form of persistence. While the NMME members generally exhibit comparable correlations and RMSE to the trend predictor at lead 3, the individual CGCMs exhibit superior predictions at lead 0. Most importantly, the NMME predictions surpass the correlations and RMSE of the trend predictions at both leads (with the exception of tropical storm predictions at lead 3), further indicating the utility of the interannual predictions from the multimodel ensemble.

Hindcast tropical storm correlations are typically larger than those for hurricanes. Tropical storms exhibit the strongest correlations from CanCM34 (0.59 and 0.70 at leads 3 and 0, respectively), superior to the corresponding NMME values of 0.57 and 0.66. As with hurricanes, tropical storm RMSE values are again reduced for NMME relative to all individual members, here on the order of 20–35% improvement over the best member. Hindcast ACE projections exhibit correlations of a similar magnitude to tropical storms and typically larger than those for hurricanes. NMME outperforms all members with correlations of 0.58 at lead 3 and 0.73 at lead 0. Yet again, RMSE is reduced for NMME relative to the member GCMs, here by 15–25% over the best performing members across each lead.

Cross-validated ACE values exhibit the strongest correlations with observations at lead 3 from the NMME (0.58), and at lead 0 (0.73). RMSE for ACE predictions are smallest for the NMME mean, again 10–15% less than the best performing NMME member CGCM. In summary the NMME predictions are typically at or near the top in terms of correlation performance, but the NMME mean exhibits consistently reduced RMSE relative to the best performing CGCMs by around 10–15%.

Given errors in CGCM initialization and boundary conditions, an uncertainty component in the predictions is desirable to account for possible errors. One such approach for this has been to add and subtract the standard deviation of the error between hindcast predictions and observations, as in Davis et al. (2015). Such an approach is undesirable here however, as it rewards poor predictions that allow the error's standard deviation, and by extension the predicted range, to increase. For example, utilizing the mean hurricane activity for 1982–2010 (6.38 hurricanes) would produce an error standard deviation of 3.20 hurricanes for a range of 3–10 hurricanes after rounding. This envelope would capture 86% of the hurricanes during 1982–2010, surpassing the 84% mark that the Davis model captures over their full 1950–2013 period. Instead here the ensemble nature of the CGCMs is exploited, and the standard deviation of the predictions for each year from the individual CGCM member projections is added and subtracted to their ensemble mean prediction to yield the predictive envelope. Such a method is constrained by the CGCM projections rather than the quality of the forecast, permitting a more meaningful predictive range. For the NMME prediction, the standard deviations of the member CGCMs are averaged, without accounting for weighting by ensemble size, and then used in the same manner as the individual CGCMs to devise the uncertainty range.

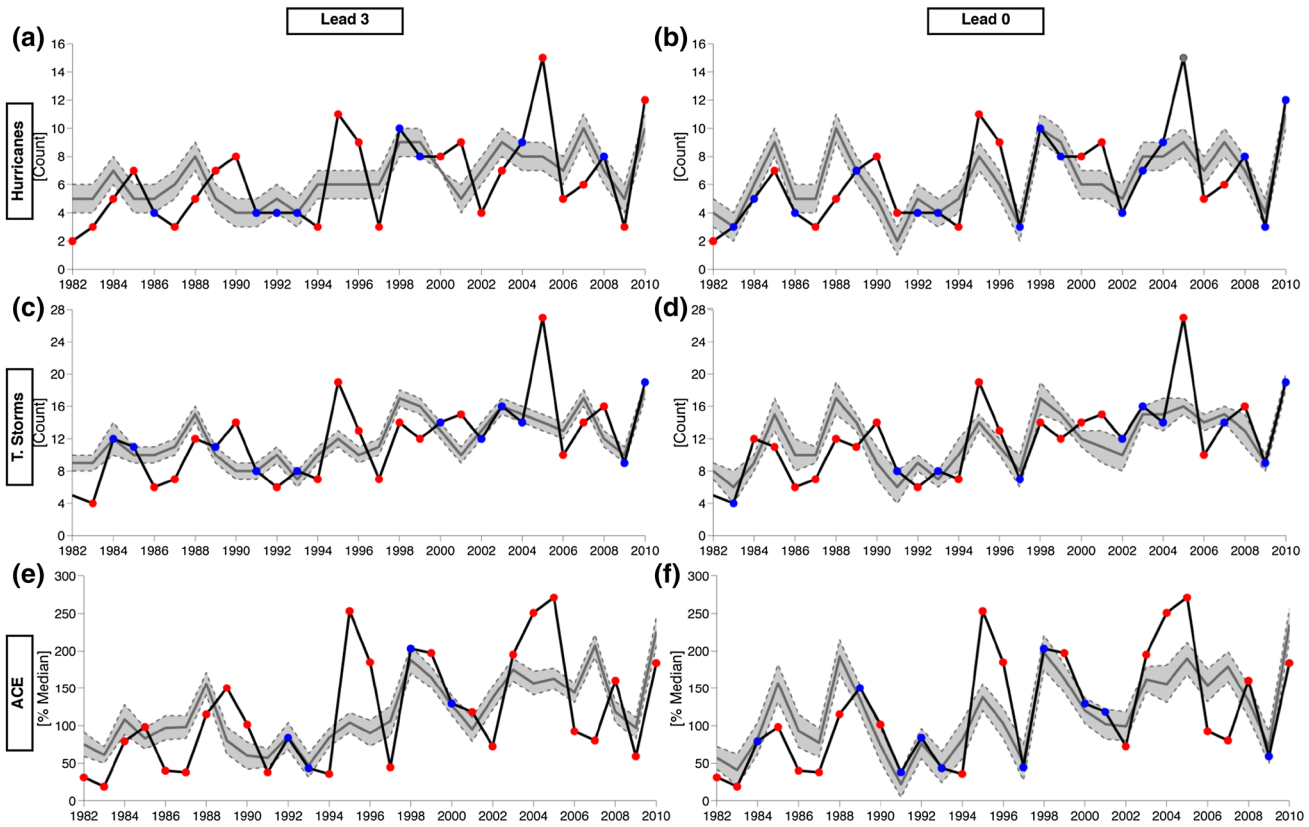


Fig. 9 Hindcast NMME mean North Atlantic hurricane predictions (*dark gray line*), and predictive envelope indicated by ±1 standard deviation averaged from member CGCMs (*light gray* region bounded by *dark gray dashed lines*) at lead 3 (**a**) and lead 0 (**b**) initial conditions. *Black line* indicates observed hurricane activity, while *blue* (*red*) *dots* denote that year's observed activity lying within (*outside*) the predictive envelope. *Panels* (**c**) and (**d**) are as in (**a**) and (**b**) but for a predictand of tropical storms, while (**e**) and (**f**) are for ACE

The NMME cross-validation period predictions incorporating uncertainty are shown in Fig. 9 for hurricanes, tropical storms and ACE. NMME hurricane projections (Fig. 9a, b) at lead 3 capture only 28% of observed seasons yet by lead 0 these values improve substantially, capturing 52% of seasonal activity. The predictive envelope does not only correctly project hurricane seasons near historical normals, as for example the 2010 season (12 hurricanes) and 1997 season (3 hurricanes) are both captured at lead 0. Tropical storm projections (Fig. 9c, d) see 38% (34%) of predictions within the envelope at lead 3 (lead 0). ACE (Fig. 9e, f) is within the NMME predicted range for 14% (34%) of the hindcast record at lead 3 (lead 0).

Some of the performance in Fig. 9 appears due to underdispersive tendencies in the NMME predictions. This is readily apparent if comparing the forecast spread of the NMME predictions from Fig. 9 to predictions using an observed long-term trend fit (i.e. red lines from Fig. 1a–c) and ±1 standard deviation of observed detrended variability (from Fig. 1d–f), with this method capturing 86% of hurricane and tropical storm predictions and 62% of ACE predictions (not shown). The issue with such a framework arises in that the average spread of the observed trend predictions is quite substantial, such as ACE on average being a range of 130% of median. The trend predictions, while a potential baseline to quantify skill, are overly broad such that value added to the users of these predictions is highly questionable. This is also apparent in that the forecast spread values of this framework are typically double those issued by Colorado State University (CSU) and NOAA in their seasonal outlooks (see the following section). A balance must be struck between the observed variance and that conveyed in forecast uncertainty, so the forecast remains of value to the end users. Subsequent models added to the NMME in the future may help increase overall diversity of predictions and broaden the resulting forecast uncertainty.

While uncertainty is conveyed in the preceding analysis, this method fails to leverage the full range of information contained in the multimodel ensemble, as it relies upon the mean and variance among the 42 ensemble members. Since NOAA's HSO conveys the odds for below-, near-, and above-normal seasonal activity it is desirable to similarly convey the odds of each tercile. Such probabilities can be developed by binning individual ensemble member forecasts from the

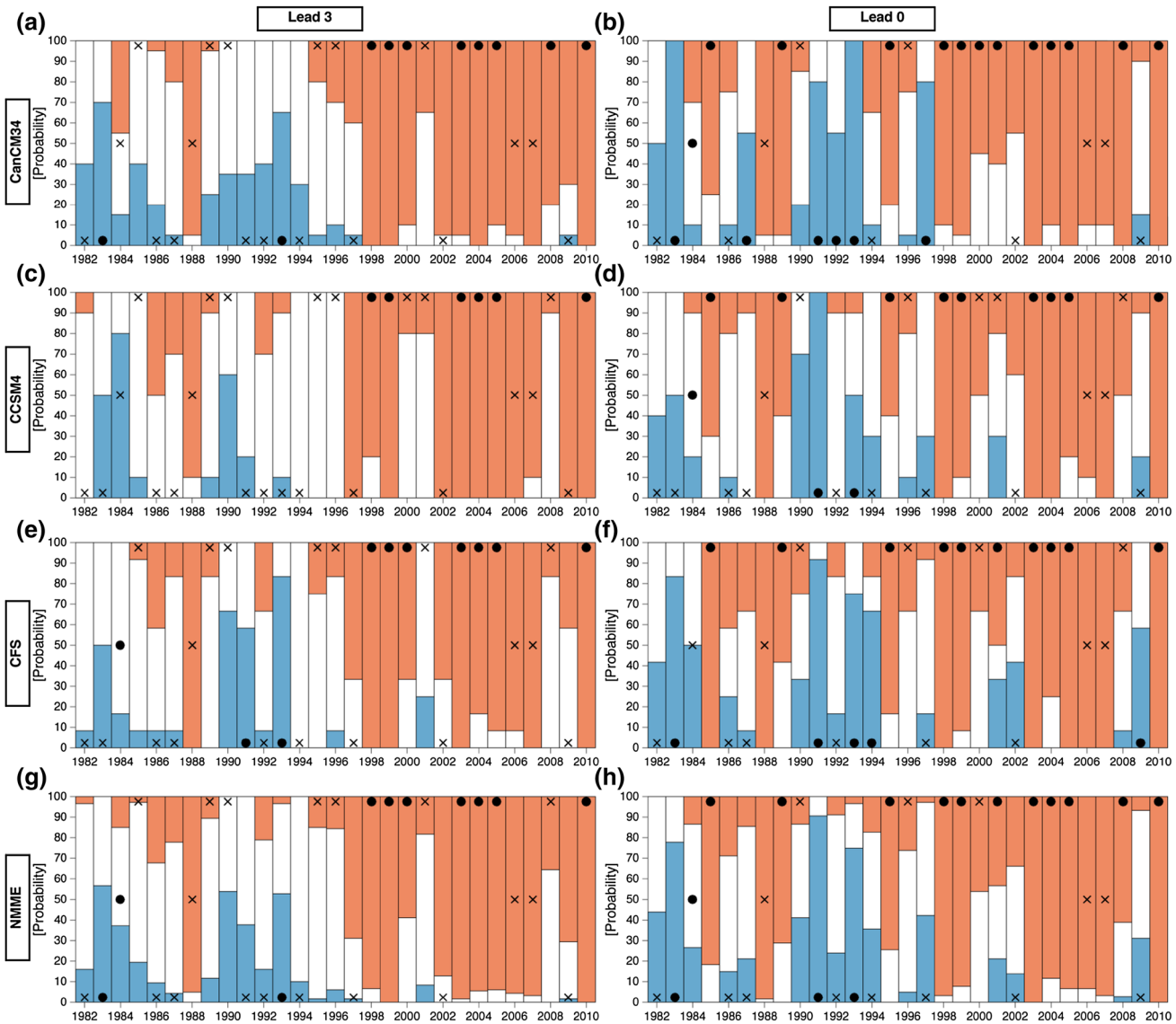


Fig. 10 Probabilistic forecasts of below-, near-, or above-normal seasonal hurricane activity in blue, white, and red respectively, over the hindcast period from CanCM34, CCSM4, CFSv2, and the NMME ensemble at lead 3 (*left column*) and lead 0 (*right column*). Markers are placed for each year near the *bottom, middle*, or *top* of the *column* to delineate whether that season verified as below-, near-, or above-normal respectively. *Circles* indicate the category with the greatest probability being forecast correctly, whereas an *x* indicates an incorrect forecast

hybrid methodology into their respectively classification relative to climatology. Below-normal activity is defined as: <5 hurricanes, 9 tropical storms, and 71.4×10^4 kt^2 of ACE, while above-normal activity is defined as: >7 hurricanes, 12 tropical storms, and 120×10^4 kt^2 of ACE, with near-normal constituting the range between the two classifications. Probabilities are generated for each member CGCM by counting the proportion of forecasts in each tercile category, while the multimodel probability is constructed by averaging the probabilities of each CGCM.

Forecast probabilities of hurricane activity from each CGCM and the NMME mean are shown in Fig. 10 along with the observed classification of each season. At lead 3 (Fig. 10a, c, e), 10 seasons have their hurricane activity forecast in the proper classification by CanCM34, CFSv2, and NMME while CCSM4 only classifies 6 seasons properly. The NMME probabilities are nonzero for every hurricane forecast category that verifies with the exception of the below normal 2002 season. Unsurprisingly, forecasts improve at lead 0 for each of the CGCMs and NMME. CanCM34 correctly forecasts 19 of the 29 seasonal hurricane activity, while CFSv2 and NMME accurately capture 15 seasons, and CCSM4 gets 12 right. NMME at lead 0 outperforms the member CGCMs in that it fails to have

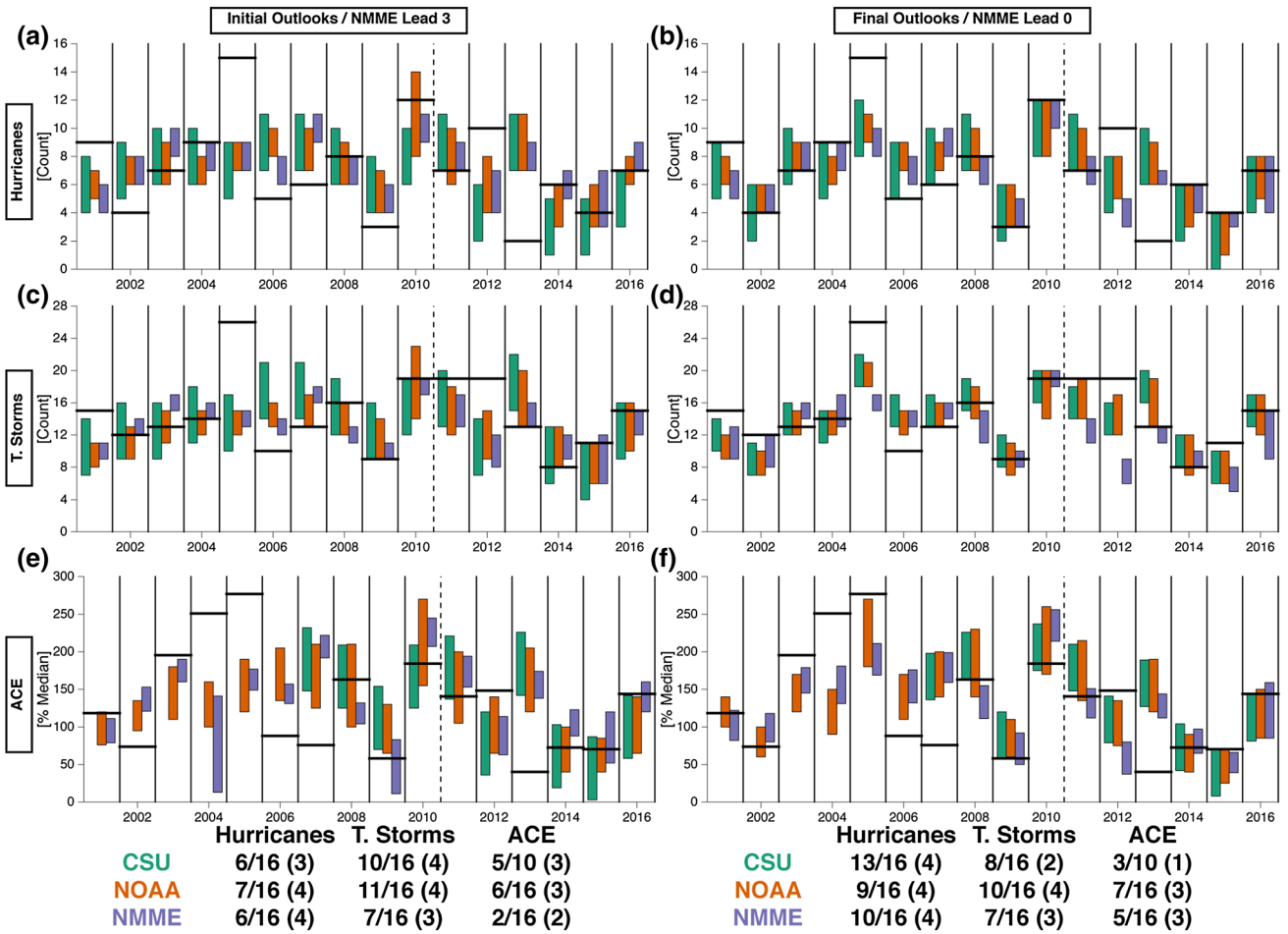


Fig. 11 Forecast ranges from CSU (*green*), NOAA (*orange*), and the hybrid NMME method (*purple*) since 2001 for listed predictands for initial outlooks (issuance in April for CSU and NMME, May for NOAA) in the *left column* and final outlooks (issuance in July for NMME, August for CSU and NOAA) in the *right column*. *Horizontal black lines* indicate observed activity. The *vertical black dashed line* indicates the transition from hindcast to real-time analyses from the NMME methodology in 2011. Proportion of predictands forecast correctly is detailed below each column, with values since 2011 in parentheses

any seasonal forecast misses with a 0% forecast probability for the category that verifies. At both leads, and across all models, above-normal hurricane activity is forecast most accurately while near- and below-normal seasons exhibit the poorest performance, which appears tied to the NATL predictor and frequent above-normal seasons observed since 1995. Given the overconfidence apparent in the probabilities at each lead in excess of 80% that fail to verify, reliability could be improved from this methodology. A more advanced methodology incorporating historical errors could improve reliability and overall probabilistic skill (e.g. Unger et al. 2009).

6 Real-time NMME North Atlantic hurricane forecast performance

With the hybrid model trained over the 29-year hindcast period from 1982 to 2010, the regression coefficients that are developed from that period can be evaluated over the 2011–2016 "real-time" period. While ideally the sample sizes between the training (hindcast) and testing (real-time) populations would be comparable, there are several arguments against balancing the two. First, given the limited availability of climate records and the forecasts being verified once per year, shrinking the training period inevitably yields a decline in forecast skill due to reduced sampling. Second, the NMME hybrid forecasts are used operationally in their current format and altering the sampling for the purposes of this publication creates discontinuity between the implemented version at NOAA and the version herein.

Third, there is ample precedent for disproportionate training and testing datasets in climate research, particularly relating to seasonal tropical cyclone forecasts with reported real-time results typically being for only a single season (e.g. Wang et al. 2009a; Kim and Webster 2010; Vecchi et al. 2011; Li et al. 2013; Villarini and Vecchi 2013). Given these limitations the real-time results herein should be taken cautiously due to limited sampling, but demonstrate utility of the hybrid methodology in an operational setting. The question arises as to how the hybrid methodology performs as a comparison to the official outlooks from CSU and NOAA. Figure 11 depicts the predicted ranges for the initial and final hurricane outlooks from CSU, NOAA, and the NMME forecast for 2001–2016. NMME results for 2001–2010 are from the cross-validation, while real-time NMME forecast data for 2011–2014 is processed through the hybrid model retrospectively despite the hybrid methodology then being unavailable to NOAA forecasters. Some caution should be taken in interpreting Fig. 11, given that the NOAA and NMME outlooks are not mutually exclusive for 2015–2016, and the CSU and NOAA outlooks potentially leverage information regarding early season tropical activity that developed prior to the outlook issuance while the NMME forecast has no such benefit. NOAA and NMME forecasts have specified ranges, while CSU outlooks are deterministic with a single value forecast. CSU outlooks are converted into forecast ranges using ± 1 standard deviation of the cross-validated standard error from 1982 to 2010.[6]

In the initial outlooks (Fig. 11a, c, e; April for CSU, May for NOAA, and lead 3 for NMME) the NMME forecasts are within 1 correct forecast difference from CSU and NOAA, with the exception of tropical storms and ACE during the hindcast period of 2001–2010. Some successes of the NMME at the longer lead are apparent in the real-time period, with the NMME outlooks being the only of the three forecasts to correctly capture the observed activity in both 2015 and 2016 for all three predictands while in use at NOAA. NMME has also forecast more seasons correctly at the longer lead (4) than CSU (3) over the real-time period, despite having a smaller forecast range to work with in each of the six seasons. For the 2001–2016 final outlooks (Fig. 11 b, d, f) results are more mixed, with CSU having the best hurricane outlooks but the lowest proportion of ACE forecasts to verify, with NOAA performing the best for tropical storm and ACE predictions but being inferior to NMME hurricane forecasts over this period. During the real-time period from 2011 onward all methods predict four of six hurricane seasons correctly, while NOAA (4) outperforms NMME (3) and CSU (2) in accurate tropical storm forecasts, and ACE forecasts are best from NOAA and NMME (3) while CSU only captures the 2014 season correctly. On a whole, the NMME method is competitive with CSU and NOAA outlooks, particularly over the real-time period where it matches or surpasses CSU performance at the shortest forecast lead. These results are further compelling for the utility of the NMME hybrid forecast, particularly given that: the NMME outlook is available prior to that of CSU and NOAA, the NMME forecast lacks any insight into antecedent tropical activity during the year that the other outlooks can take into account, and the NOAA outlook being an aggregation of a wealth of tools (including the NMME guidance). The initial operational use of the NMME hybrid methodology at CPC for the 2015 season is explored in greater detail in "Appendix 1".

7 Conclusions

Here a hybrid dynamical–statistical model for North Atlantic hurricane activity is developed utilizing NMME inputs. The model is unique for its hybrid nature bridging statistical relationships with the dynamical component to predict the future atmospheric state rather than relying upon lagged relationships. The dynamical component is further enhanced through the NMME representation that incorporates multiple CGCMs with their inherent strengths and weaknesses for reproducing the observed atmospheric and oceanic state to provide a broader perspective and increase uncertainty conveyance. Deterministic and probabilistic forecasts of hurricane season activity is yielded by these methods, and used operationally to supplement as a guidance tool in producing the NOAA HSO.

A predictive method is developed; utilizing ASO forecast MDR vertical wind shear and observed NATL preseason SST and the full retrospective climatology. With this methodology, the multi-model mean prediction based upon the NMME member CGCMs yields typical reductions of 10–35% in RMSE dependent upon the predictands, relative to the best performing member CGCM. NMME mean correlations with observed activity surpass the majority of member CGCMs and typically all members. The statistical envelope of the NMME projections for hurricane, tropical storm, and ACE activity exhibits considerable skill in reproducing observations for both inactive and active seasons. Initial probabilistic forecast guidance also proves skillful over the hindcast period, but appear underdispersive and overconfident and could be subsequently improved by incorporating more advanced methods (e.g. Unger et al. 2009). NMME forecasts are competitive with CSU and NOAA outlooks in recent years, despite being available prior to the release of those products and lacking the knowledge of any antecedent tropical activity prior to the official forecast release.

[6] As per Sect. 3 of http://webcms.colostate.edu/tropical/media/sites/111/2017/04/2017-04.pdf.

Difficulties remain for extremely active (e.g. 2005) and inactive seasons (e.g. 2009), however subseasonal contributions such as MJO impacts (e.g. Klotzbach 2014; Klotzbach and Oliver 2015) and extratropical impacts via Rossby wave breaking (e.g. Zhang et al. 2016) may complicate hurricane variability at the seasonal scale. The proposed subseasonal NMME project (Pegion 2015) can introduce opportunities to refine the NMME hybrid model approach at these timescales and potentially yield improved predictions.

Predictability among NMME members varies widely for both SST and vertical wind shear. SST anomaly correlations are generally superior to those of vertical wind shear. SST, and particularly ENSO, predictability appears to be a strong focus for the NMME member CGCMs (e.g. Kirtman et al. 2014; Merryfield et al. 2013; Saha et al. 2014) while lesser attention is seemingly paid toward the atmospheric circulation. This prioritization is manifested to some extent, in zonal wind not being deemed a mandatory output variable for NMME. Future attention towards the reproduction of the circulation characteristics would serve to improve the hybrid model, in addition to adding insight into other NMME variables that are deemed mandatory such as temperature and precipitation. The NMME hindcast archive continues to grow with models incorporating zonal wind information, such as the NASA Goddard Earth Observing System Model Version 5 (GEOS-5; Vernieres et al. 2012) and NCAR Community Earth System Model (CESM) Version 1 slated for inclusion. Inclusion of additional models will act to extend the perspective of the NMME multi-model mean and potentially lead to further improvements in predictability and hybrid model performance.

Acknowledgements This project was funded by NOAA's High Impact Weather Prediction Project (HIWPP) under the NMME extension. Feedback from two anonymous reviewers as well as internal reviews from Drs. Emily Becker and Peitao Peng greatly improved the quality of this manuscript. Additional helpful discussions with Gerry Bell, Lindsey Long, and Michelle L'Heureux helped drive this project. We acknowledge the agencies that support the NMME-Phase II system, and we thank the climate modeling groups (listed in Table 1) for producing and making available their model output. NOAA National Centers for Environmental Prediction, NOAA Climate Test Bed and NOAA Climate Program Office jointly provide coordinating support and led development of the NMME-Phase II system. NMME hindcast data was retrieved from the NCAR Earth System Grid repository that is supported financially by DOE, NASA, NOAA, and NSF. Maintenance, support, and development of the Earth System Grid repository is provided by CPC, IRI, and NCAR personnel.

Appendix 1: Initial performance of the hybrid NMME forecast for CPC operations

The NMME hybrid model was incorporated as a tool for the Atlantic Hurricane Seasonal Outlook issued by the Climate Prediction Center preceding the 2015 season. This year proved to be an interesting gauge, given building El Niño conditions early in the year, culminating with El Niño declaration by NOAA during March.[7] Observed JFM SST anomalies for the NATL (Fig. 7) were −0.14 °C relative to 1982–2014, the 16th lowest ranked value relative to the 1982–2014 period. Lead 3 forecasts for ASO wind shear in 2015 (Fig. 12) projected above normal wind shear across the MDR throughout the NMME members (Fig. 12a–c), with the strongest shear on the periphery of the MDR with the exception of CanCM34. CanCM34 forecasted the strongest MDR shear (4.91 m/s) with enhanced shear across much of the MDR, followed by the CCSM4 (2.06 m/s) and CFSv2 (0.46 m/s) that instead portrayed positive shear anomalies for the southwestern MDR and negative shear anomalies in the eastern MDR. The NMME mean shear forecast (Fig. 12d) for the MDR lies in between the CanCM34 and other members at 2.45 m/s, with modest positive anomalies throughout the MDR.

Climate conditions further suggested reduced 2015 hurricane activity at lead 0 during 2015. Observed NATL SST for AMJ (Fig. 7) was −0.58 °C relative to 1982–2014 for the 7th lowest ranked year. By early July, many ENSO predictions were indicating the potential for a major El Niño event to be established by ASO,[8] implying potential severe reduction in hurricane activity. NMME forecasts of lead 0 ASO vertical wind shear are shown in Fig. 13. Apparent are the extreme shear forecasts for the tropical Atlantic by CanCM34 (Fig. 13a), a value exceeding four standard deviations above normal and far larger than any shear magnitude forecast at this lead in the hindcast period. These strong shear projections were present in both CanCM3 and CanCM4 as they forecast +4 and +5 standard deviation events relative to their individual climatologies. While the extent of the CanCM34 positive shear anomalies extend across the entirety of the MDR, the greatest values generally are found in the southern Caribbean Sea. The remaining NMME members (Fig. 13b, c) are less extreme in their ASO shear projections, with CCSM4 exhibiting above normal shear from the northwest to southeast across the MDR with a mean magnitude of 3.40 m/s, while the CFSv2 forecasts near normal MDR shear (0.02 m/s) with negative values off the African coastline through 40°W. Strong positive shear anomalies are seemingly associated with the developing El Niño by CFSv2, however they are predominantly constrained to the East Pacific (peaking >18 m/s) and south of the MDR across the Atlantic basin. Averaging the three members yields the NMME mean in Fig. 13d with 4.00 m/s of shear across the

[7] http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/enso_disc_mar2015/ensodisc.pdf.

[8] http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/enso_disc_jul2015/ensodisc.pdf.

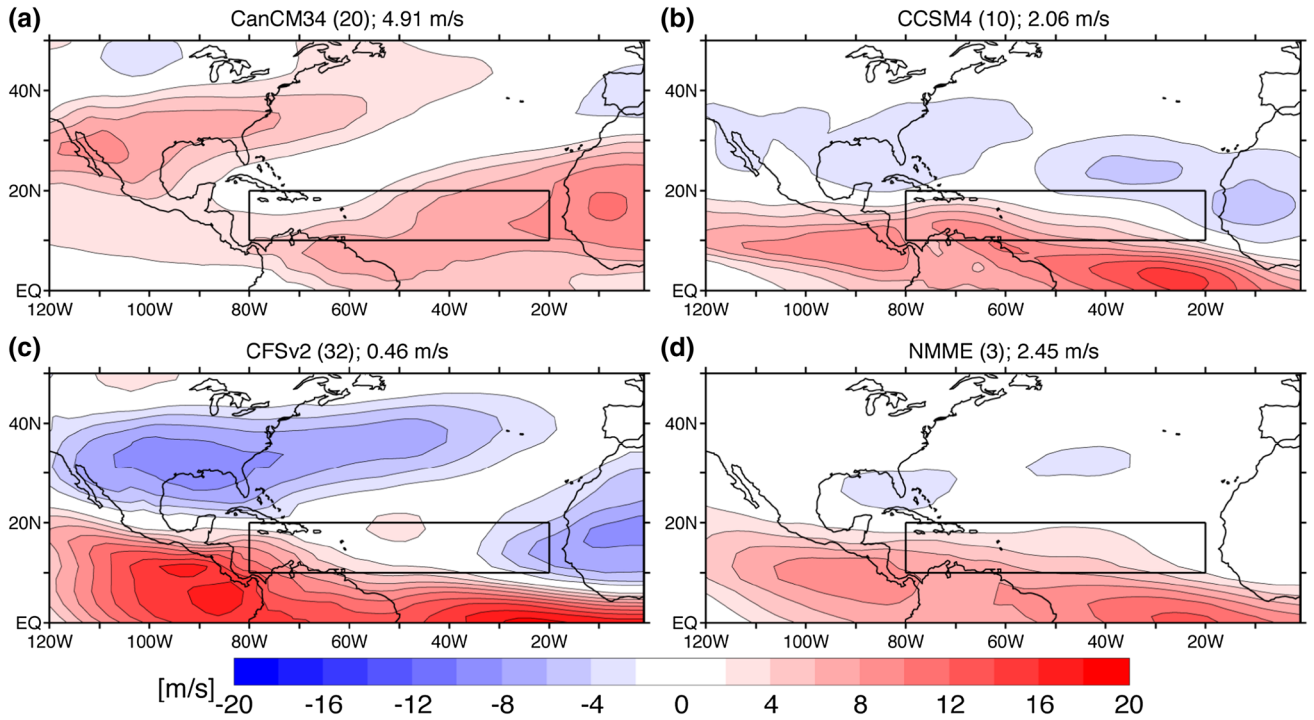


Fig. 12 Forecast anomalous vertical wind shear for ASO 2015 at lead 3 (April initial conditions) relative to 1982–2010 for CanCM34 (**a**), CCSM4 (**b**), CFSv2 (**c**), and multi-model mean (**d**). Contour intervals are 2 m/s, with the zero contour omitted. MDR averages are listed in the title of *each panel*

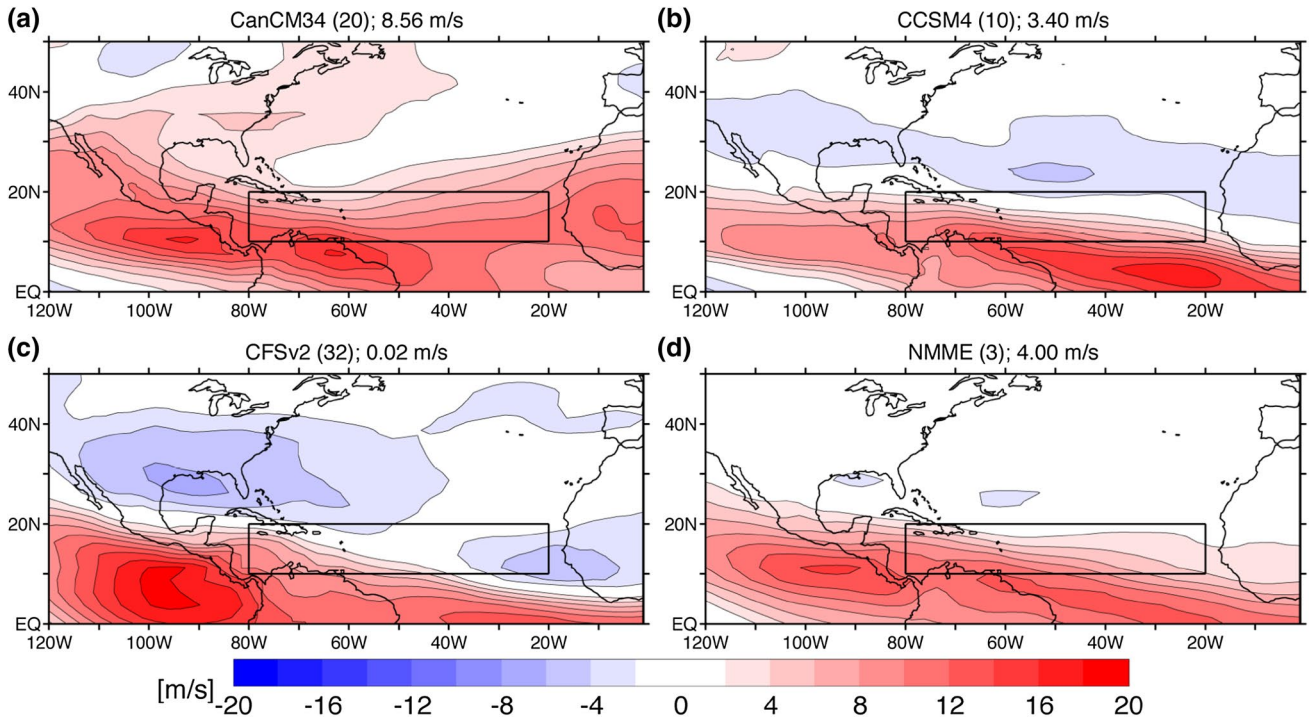


Fig. 13 As in Fig. 12, but at lead 0 (July initial conditions)

Table 3 Expected value and range (in parentheses) of predictions for 2015 North Atlantic hurricane, tropical storm, and ACE (as a percent of median) activity from the NMME models and ensemble mean

Predictand	Lead	CanCM34	CCSM4	CFSv2	NMME
Hurricanes	Lead 3	**3 (0–6)**	**5 (4–6)**	6 (5–7)	**4 (3–6)**
	Lead 0	1 (0–2)	**3 (2–4)**	**6 (4–7)**	**3 (2–5)**
Tropical storms	Lead 3	5 (0–10)	9 (8–10)	**11 (9–13)**	**8 (6–11)**
	Lead 0	2 (0–4)	7 (6–9)	**11 (9–13)**	7 (5–9)
ACE	Lead 3	**42 (0–98)**	**80 (64–96)**	102 (75–128)	**75 (42–108)**
	Lead 0	6 (0–23)	38 (23–48)	103 (74–132)	**49 (26–72)**

Bold values indicate observed activity fell within the forecasted range

MDR with this number strongly influenced by the CanCM34 projections both in magnitude and spatial distribution with positive anomalies extending across the full entirety of the MDR despite the weakness in the eastern Atlantic from the CCSM4 and CFSv2.

NMME forecasts for the 2015 season in Table 3 for hurricane, tropical storm, and ACE activity were included as part of guidance towards development of the 2015 NOAA HSO. The 2015 season yielded 4 hurricanes, 11 tropical storms, and an ACE value of 59.3×10^4 kt^2 (64.2% of median). NMME predicted ranges capture the observed hurricane activity (Table 3) at both leads for hurricanes, with the best performing member being CCSM4, which captured the observed activity within its forecast range at both leads. Tropical storm activity was also reasonably well captured by NMME, with forecast ranges only failing to match observations for lead 0, but outperformed by the CFSv2 which saw observed activity fall within its predicted ranges at each lead. The failure by the NMME at lead 0 is largely attributable to the CanCM34 which projected an average of two tropical storms for the season, despite three tropical storms developing prior to the HSO update release in August 2015. NMME best captured ACE during 2015, with forecast ranges aligning with observations for both leads. Also noteworthy with ACE forecasts is that the CFSv2 failed to reproduce the observed ACE range for any lead, underscoring the original hybrid method conditioned solely on the CFS would have failed in its ACE projections for 2015. The 2015 hurricane season was somewhat unusual in that despite the near-normal tropical storm activity there were reduced hurricane and ACE activity. Nevertheless, the NMME prediction was able to reproduce the observed distributions for five of the six analyzed forecast combinations of forecast lead and predictands. This underscores the ability of the multimodel mean to aggregate diverse forecasts into an improved product that is not restricted to any singular model with its individual strengths, weaknesses, and biases.

References

Barnston AG, Tippett MK (2013) Predictions of Nino 3.4 SST in CFSv1 and CFSv2: a diagnostic comparison. Clim Dyn 41:1615–1633. doi:10.1007/s00382-013-1845-2

Barnston AG, Tippett MK, L'Heureux ML, Li S, Dewitt DG (2012) Skill of real-time seasonal ENSO model predictions during 2002-11: Is our capability increasing? Bull Amer Meteor Soc 93:631–651. doi:10.1175/BAMS-D-11-00111.1

Becker EJ, van den Dool H, Zhang Q (2014) Predictability and forecast skill in NMME. J Clim 27:5891–5906. doi:10.1175/JCLI-D-13-00597.1

Bell GD, Chelliah M (2006) Leading tropical modes associated with interannual and multidecadal fluctuations in North Atlantic hurricane activity. J Clim 19:590–612. doi:10.1175/JCLI3659.1

Choi W, Ho C-H, Kim J, Kim H-S, Feng S, Kang K (2015) A track-pattern based seasonal prediction of tropical cyclone activities over the North Atlantic. J Clim 29:481–494. doi:10.1175/JCLI-D-15-0407.1

Davis K, Zheng X, Ritchie EA (2015) A new statistical model for predicting seasonal North Atlantic hurricane activity. Weather Forecasting 30:730–741. doi:10.1175/WAF-D-14-00156.1

Dee D et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J Roy Meteor Soc 137:553–597. doi:10.1002/qj.828

DeWitt DG (2005) Retrospective forecasts of interannual sea surface temperature anomalies from 1982 to present using a directly coupled atmosphere-ocean general circulation model. Mon Weather Rev 133:2972–2995. doi:10.1175/MWR3016.1

Enfield DB, Mestas-Nuñez AM, Trimble PJ (2001) The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US. Geophys Res Lett 28:2077–2080. doi:10.1029/2000GL012745

Goldenberg SB, Shapiro LJ (1996) Physical mechanisms for the association of El Niño and West African rainfall with Atlantic major hurricane activity. J Clim 9:1169–1187. doi:10.1175/1520-0442(1996)009<1169:PMFTAO>2.0.CO;2

Goldenberg SB, Landsea CW, Mestas-Nuñez AM, Gray WM (2001) The recent increase in Atlantic hurricane activity: cause and implications. Science 293:474–479. doi:10.1126/science.1060040

Gray WM (1984a) Atlantic seasonal hurricane frequency. Part I: El Niño and 30-mb quasi-biennial oscillation influences. Mon Weather Rev 112:1649–1668. doi:10.1175/1520-0492(1984)112<1649:ASHFPI>2.0.CO;2

Gray WM (1984b) Atlantic seasonal hurricane frequency. Part II: forecasting its variability. Mon Weather Rev 112:1649–1668. doi:10.1175/1520-0493(1984)112<1669:ASHFPI>2.0.CO;2

Kalnay E et al (1996) The NCEP/NCAR 40-year reanalysis project. Bull Am Meteor Soc 77:437–471. doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Kim HM, Webster PJ (2010) Extended-range seasonal hurricane forecasts for the North Atlantic with a hybrid dynamical-statistical model. Geophys Res Lett 37:L21705. doi:10.1029/2010GL044792

Kim O, Kim HM, Lee MI (2017) Dynamical-statistical seasonal prediction for western North Pacific typhoons based on APCC multi-models. Clim Dynamics 48:71–88

Kirtman BP (2003) The COLA anomaly coupled model: Ensemble ENSO prediction. Mon Weather Rev 131:2324–2341. doi:10.1175/1520-0493(2003)131<2324:TCACME>2.0.CO;2

Kirtman BP et al (2014) The North American multimodel ensemble: phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. Bull Am Meteor Soc 95:585–601. doi:10.1175/BAMS-D-12-00050.1

Klotzbach PJ (2014) The Madden–Julian oscillation's impacts on worldwide tropical cyclone activity. J Clim 27:2317–2330. doi:10.1175/JCLI-D-13-00483.1

Klotzbach PJ, Oliver E. C. J. (2015) Modulation of Atlantic basin tropical cyclone activity by the Madden–Julian Oscillation (MJO) from 1905 to 2011. J Clim 28:204–217. doi:10.1175/JCLI-D-14-00509.1

Knutson TR, Sirutis JJ, Garner ST, Held IM, Tuleya RE (2007) Simulation of the recent multidecadal increase of Atlantic hurricane activity using an 18-km-grid regional model. Bull Am Meteor Soc 88:1549–1565. doi:10.1175/BAMS-88-10-1549

Knutson TR, McBride JL, Chan J, Emanuel K, Holland G, Landsea C, Held I, Kossin JP, Srivastava AK, Sugi M (2010) Tropical cyclones and climate change. Nat Geosci 3:157–163. doi:10.1038/ngeo779

Koltermann KP, Sokov AV, Terschnekov VP, Dobroliubov SA, Lorbacher K, Sy A (1999) Decadal changes in the thermohaline circulation of the North Atlantic. Deep Sea Res Part II 46:109–138. doi:10.1016/S0967-0645(98)00115-5

Kossin JP, Vimont DJ (2007) A more general framework for understanding Atlantic hurricane variability and trends. Bull Am Meteor Soc 88:1767–1781. doi:10.1175/BAMS-88-11-1767

Krishnamurti TN, Kishtawal CM, LaRow TE, Bachiochi DR, Zhang Z, Williford CE, Gadgil S, Surendran S (1999) Improved weather and seasonal climate forecasts from multimodel superensemble. Science 285:1548–1550. doi:10.1126/science.285.5433.1548

Landsea CW, Pielke RA Jr, Maestas-Nunez AM, Knaff JA (1999) Atlantic basin hurricanes. Clim Change 42:89–129

Landsea CW et al (2004) The Atlantic hurricane database re-analysis project: documentation for the 1851–1910 alterations and additions to the HURDAT database. In: Murnane RJ, Liu K-B (eds) Hurricanes and typhoons: past, present and future. Columbia University Press, New York, pp 177–221

Li X, Yang S, Wang H, Jia X, Kumar A (2013) A dynamical-statistical forecast model for the annual frequency of western Pacific tropical cyclones based on the NCEP Climate Forecast System version 2. J Geophys Res Atmos 118:12061–12074. doi:10.1002/2013JD020708

Mendelsohn R, Emanuel K, Chonabayashi S, Bakkensen L (2012) The impact of climate change on global tropical cyclone damage. Nat Clim Change 2:205–209. doi:10.1038/nclimate1357

Merryfield WJ et al (2013) The Canadian seasonal to interannual prediction system. Part I: models and initialization. Mon Weather Rev 141:2910–2945. doi:10.1175/MWR-D-12-00216.1

Min Y-M, Kryjov VN, Park C-K (2009) A probabilistic multimodel ensemble approach to seasonal prediction. Weather Forecasting 24:812–828. doi:10.1175/2008WAF2222140.1

Palmer TN, Brankovic C, Richardson DS (2000) A probability and decision-model analysis of PROVOST seasonal multimodel ensemble integrations. Q J Roy Meteor Soc 126:2013–2034. doi:10.1002/qj.49712656703

Patricola CM, Saravanan R, Chang P (2014) The impact of the El Niño-Southern Oscillation and Atlantic meridional mode on seasonal Atlantic tropical cyclone activity. J Clim 27:5311–5328. doi:10.1175/JCLI-D-13-00687.1

Peduzzi P, Chatenoux B, Dao H, De Bono A, Herold C, Kossin J, Mouton F, Nordbeck O (2012) Global trends in tropical cyclone risk. Nat Clim Change 2:289–294. doi:10.1038/nclimate1410

Pegion K (2015) Development of a subseasonal North American multi-model ensemble prediction system. AGU Fall Meeting, San Francisco

Pielke RAJ Jr, Gratz CW, Landsea D, Collins MA, Saunders, Musulin R (2008) Normalized hurricane damage in the United States: 1900–2005. Nat Haz Rev 9:29–42. doi:10.1061/(ASCE)/1527-6988(2008)9:1(29)

Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W (2002) An improved in situ and satellite SST analysis for climate. J Clim 15:1609–1625. doi:10.1175/1520-0442(2002)015<1609:AIISASS>2.0.CO;2

Saha S et al (2010) The NCEP climate forecast system reanalysis. Bull Am Meteor Soc 91:1015–1057. doi:10.1175/2010BAMS3001.1

Saha S et al (2014) The NCEP Climate Forecast System version 2. J Clim 27:2185–2208. doi:10.1175/JCLI-D-12-00823.1

Schemm J-KE, Long L (2014) CPC dynamic hurricane season prediction system upgrade with the NCEP CFSv2. 39th Climate Diagnostics Prediction Workshop, St. Louis, MO, National Oceanic and Atmospheric Association. http://www.cpc.ncep.noaa.gov/products/outreach/proceedings/cdw39_proceedings/D ay_3/Session_7/Long.pdf

Schlesinger ME, Ramankutty N (1994) An oscillation in the global climate system of period 65–70 years. Nature 367:723–726. doi:10.1038/367723a0

Trenberth KE, Shea DJ (2006) Atlantic hurricane and natural variability in 2005. Geophys Res Lett 33:L12704. doi:10.1029/2006GL026894

Unger DA, van den Dool H, O'Lenic E, Collins D (2009) Ensemble regression. Mon Weather Rev 137:2365–2379. doi:10.1175/2008MWR2605.1

Vecchi GA, Zhao M, Wang H, Villarini G, Rosati A, Kumar A, Held IM, Gudgel R (2011) Statistical–dynamical predictions of seasonal North Atlantic hurricane activity. Mon Weather Rev 139:1070–1082. doi:10.1175/2010MWR3499.1

Vernieres G, Rienecker MM, Kovach R, Keppenne CL (2012) The GEOS-iODAS: description and evaluation. Technical report series on global modeling and data assimilation 30, 73 p, NASA Goddard Space Flight Center, Greenbelt, MD, USA. NASA/TM-2012-104606/VOL30

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

Vimont DJ, Kossin JP (2007) The Atlantic meridional mode and hurricane activity. Geophys Res Lett 34:L07709. doi:10.1029/2007GL029683

Vitart F, 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

Vitart F, Stockdale TN (2001) Seasonal forecasting of tropical storms using coupled GCM integrations. Mon Weather Rev 129:2521–2537. doi:10.1175/1520-0493(2001)129<2521:SFOTSU>2.0.CO;2

Wang H, Schemm J.-K. E., Kumar A, Wang W, Long L, Chelliah M, Bell GD, Peng P (2009a) A statistical forecast model for Atlantic seasonal hurricane activity based on the NCEP dynamical season forecast. J Clim 22:4481–4500. doi:10.1175/2009JCLI2753.1

Wang B et al (2009b) Advance and prospectus of seasonal prediction: assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980–2004). Clim Dyn 33:93–117. doi:10.1007/s00382-008-0460-0

Wang H et al (2014) How well do global climate models simulate the variability of Atlantic tropical cyclones associated with ENSO? J Clim 27:5673–5692. doi:10.1175/JCLI-D-13-00625.1

Willoughby HE (2012) Distributions and trends of death and destruction from hurricanes in the United States, 1900–2008. Nat Haz Rev 13:57–64. doi:10.1061/(ASCE)NH.1527-6996.00000046

Zhang G, Wang Z, Dunkerton TJ, Peng MS, Magnusdottir G (2016) Extratropical impacts on Atlantic tropical cyclone activity. J Atmos Sci 73:1401–1418. doi:10.1175/JAS-D-15-0154.1