# Seasonal Forecasting of Wind and Waves in the North Atlantic Using a Grand Multimodel Ensemble

RAY BELL[a] AND BEN KIRTMAN

*Department of Atmospheric Sciences, Rosenstiel School of Marine and Atmospheric Science,
University of Miami, Miami, Florida*

ABSTRACT

This study assesses the skill of multimodel forecasts of 10-m wind speed, significant wave height, and mean wave period in the North Atlantic for the winter months. The 10-m winds from four North American multimodel ensemble models and three European Multimodel Seasonal-to-Interannual Prediction project (EUROSIP) models are used to force WAVEWATCH III experiments. Ten ensembles are used for each model. All three variables can be predicted using December initial conditions. The spatial maps of rank probability skill score are explained by the impact of the North Atlantic Oscillation (NAO) on the large-scale wind–wave relationship. Two winter case studies are investigated to understand the relationship between large-scale environmental conditions such as sea surface temperature, geopotential height at 500 hPa, and zonal wind at 200 hPa to the NAO and the wind–wave climate. The very strong negative NAO in 2008/09 was not well forecast by any of the ensembles while most models correctly predicted the sign of the event. This led to a poor forecast of the surface wind and waves. A Monte Carlo model combination analysis is applied to understand how many models are needed for a skillful multimodel forecast. While the grand multimodel ensemble provides robust skill, in some cases skill improves once some models are not included.

## 1. Introduction

Accurate seasonal forecasts of wind and waves in the North Atlantic will provide benefits to coastal land management (Tsimplis et al. 2005), marine vessel routing (Coll et al. 2013; Heij and Knapp 2015; Bell and Kirtman 2018), renewable energy (Clark et al. 2017; Torralba et al. 2017), and oil and gas activities (Colman et al. 2011).

The variability of the winter North Atlantic wind–wave climate is dominated by a small number of modes. Previous studies have shown the North Atlantic Oscillation (NAO) to have the dominant effect on wind–wave variability and, to a lesser degree, the east Atlantic pattern (Bacon and Carter 1991; Wolf and Woolf 2006; Dodet et al. 2010; Stopa and Cheung 2014;

Martínez-Asensio et al. 2016; Castelle et al. 2017; Zubiate et al. 2017). The NAO is the term given to the north–south oscillation of mass that results in variability of the latitude of the North Atlantic midlatitude storm track, which is most prominent in winter (Stephenson et al. 2003). It is traditionally measured using mean sea level pressure (MSLP) fluctuations between Iceland and the Azores. During the NAO's positive (negative) phase the storm track takes more of a northerly (southerly) track.

It is important to understand the dynamics of the NAO to understand sources of predictability. The self-maintaining nature of the North Atlantic storm track is discussed in Hoskins and Valdes (1990) and Ambaum and Novak (2014). However, we are more interested in how the storm track behaves overall during the winter months. Brayshaw et al. (2009) noted that midlatitude orography and the land–sea temperature contrast on the east coast of North America play a role in determining the southwest–northeast tilt of the upper-tropospheric jet. In addition, sea surface temperature (SST) gradients in the North Atlantic act to perturb the jet stream. Gradients in the

---

western North Atlantic, such as the Gulf Stream, strengthen the jet downstream, while gradients in the North Atlantic Current weaken the jet (Brayshaw et al. 2009; O'Reilly et al. 2017b).

In addition, the NAO is influenced by remote SSTs, such as those in the tropical Pacific. The mechanisms by which ENSO influences the NAO remain unclear and are a topic of debate. Some studies look at the stratospheric pathway, where during warm events (El Niños) the polar vortex weakens and the frequency of sudden stratospheric warming events increases (Butler et al. 2017). These events descend from the stratosphere, weakening the upper-level westerlies and the resultant NAO is negative (Brönnimann et al. 2007; Bell et al. 2009; Hansen et al. 2017). Kim et al. (2012) found the seasonal forecast model CFSv2 simulates this El Niño–negative NAO response better than the seasonal forecast model ECMWF System 4. Others argue for a tropospheric pathway via the role of quasi-stationary waves kicked off from diabatic heating in the tropical Pacific, which are enhanced by eddy–mean flow interactions in the extratropics (Graf and Zanchettin 2012; Jiménez-Esteve and Domeisen 2018). Butler et al. (2014) separate these two pathways and find the stratospheric pathway mostly influences the North Atlantic, whereas the tropospheric pathway mostly influences North America. It is also possible El Niño–Southern Oscillation (ENSO) SSTs can modulate North Atlantic north tropical SSTs (Alexander et al. 2002), which then in turn affect the NAO. Zhang et al. (2018) found that the ENSO–NAO relationship is highly nonlinear and can depend on the type of ENSO event. Finally, SSTs in the South Atlantic are also thought to play a role in the NAO via a barotropic response to upper-level convergence over the Caribbean (Robertson et al. 2000).

Despite seasonal forecasts often struggling in the extratropics due to a low signal-to-noise ratio (Eade et al. 2014), recent studies have found the NAO to be predictable on seasonal time scales (Scaife et al. 2014; Siegert et al. 2016; Yang et al. 2015; Dunstone et al. 2016; Jha et al. 2016; Athanasiadis et al. 2017; Wang et al. 2017). This is in part due to the drivers of seasonal NAO variability being predictable on seasonal time scales. For example, tropical Pacific SSTs are often at the crux of a skillful seasonal forecast (Shukla 1998) and its impact on the NAO is captured in the Global Seasonal Forecast System version 5 (GloSea5) model [see Fig. 2 in Scaife et al. (2014)], which has good skill at forecasting the NAO. Atlantic SST has also shown to be predictable on seasonal time scales (Rodwell and Folland 2002), such as in the North Atlantic subpolar gyre.

TABLE 1. Seasonal forecast models used in this study. Ten ensembles are used throughout as shown in the parentheses in the number of ensembles row.

| Model | CanCM3 | CanCM4 | CCSM4 | CFSv2 | GloSea5 | System 4 | System 5 |
|---|---|---|---|---|---|---|---|
| Center | CCCMA | CCCMA | UM | NCEP | Met Office | ECMWF | Météo-France |
| No. of ensembles | 10 | 10 | 10 | 24 (10) | 28 (10) | 15 (10) | 15 (10) |
| Ensemble perturbation | Assimilation runs | Assimilation runs | Prior 12-hourly | 5-day lagged start | 8-day lagged start | Atmospheric singular vectors | Atmospheric singular vectors |
| Atmosphere resolution | T63 (2.8°) | T63 (2.8°) | f19 (0.9° × 1.25°) | T126 | N216 (0.7°) | T255 | T255 |
| Ocean resolution | ~100 km | ~100 km | 1° | 0.5° | 0.25° | 1° | 1° |
| Model top | 1 hPa | 1 hPa | ~3 hPa | L64 (0.2 hPa) | L85 (85 km) | L91 (0.01 hPa) | L91 (0.01 hPa) |
| Reference | Merryfield et al. (2013) | Merryfield et al. (2013) | Infanti and Kirtman (2016) | Saha et al. (2014) | MacLachlan et al. (2015) | Molteni et al. (2011) | Météo-France (2016) |
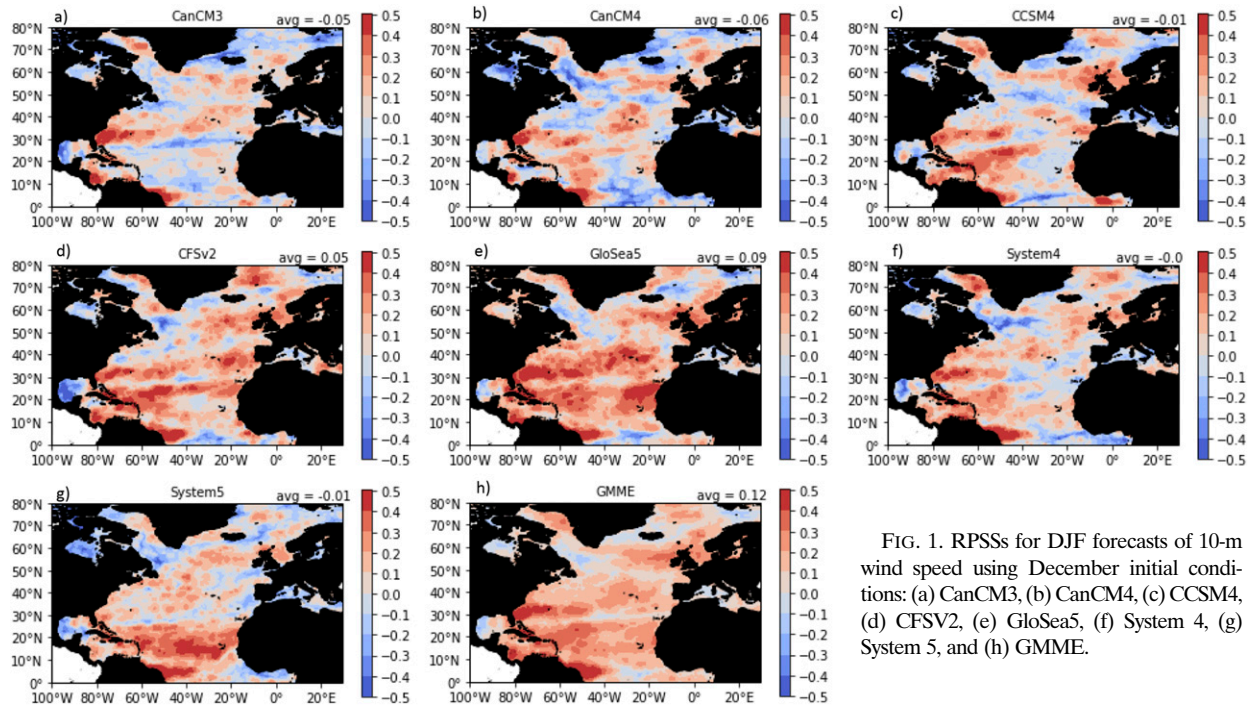
FIG. 1. RPSSs for DJF forecasts of 10-m wind speed using December initial conditions: (a) CanCM3, (b) CanCM4, (c) CCSM4, (d) CFSV2, (e) GloSea5, (f) System 4, (g) System 5, and (h) GMME.

While it is important to understand and predict the physical mechanisms forcing the NAO (Scaife et al. 2016), the role of ensembles and multimodel forecasts increases the skill of forecasting the NAO. Figure 3 in Scaife et al. (2014) shows that increasing the number of ensembles increases the correlation of the ensemble-mean forecast with the observed NAO. This skill limit is determined by the quotient of how well the ensemble members correlate with the observations over the correlation between pairs of ensembles members. Sardeshmukh et al. (2000) and Eade et al. (2014) also note how reliable predictions may be achieved using a large ensemble size to reduce noise. Multimodel forecasts have shown their benefit over a single model forecast in recent years (Palmer et al. 2004; Kirtman et al. 2014), especially in their superiority of probabilistic verification metrics. The skill increase in multimodel forecasts comes from error cancellation of model biases. Doblas-Reyes et al. (2003) found a multimodel forecast of the NAO was superior to individual models using European models from the DEMETER project (Palmer et al. 2004); however, NAO skill is marginal in these old-generation models. Athanasiadis et al. (2017) investigated the NAO and Arctic Oscillation (AO) using CFSv2, GloSea5, and CMCC seasonal forecasts. They found an increase in the correlation coefficient with the observed NAO in part coming from the increased ensemble size. Ehsan et al. (2017) found good skill in predicting winter precipitation in two subtropical regions using NMME and the ECMWF System 4 models.

In this study we make use of a grand multimodel ensemble (GMME) to investigate the skill of seasonal forecasts of wind and waves in the North Atlantic. The GMME is a comprehensive multimodel study using models in the North American Multimodel Ensemble (NMME) and models in the European Multimodel Seasonal-to-Interannual Prediction project (EUROSIP). EUROSIP is designed to provide multimodel forecasts from independent coupled seasonal forecasting systems (see https://www.ecmwf.int/en/forecasts/documentation-and-support/long-range/seasonal-forecast-documentation/eurosip-user-guide/multi-model for more information). While, previous studies have focused on seasonal forecasting of surface winds, to our knowledge, this is the first study looking at seasonal ocean wave forecasts using a multimodel ensemble.

This paper is structured as follows. The datasets, model experiments, and skill metrics are described in section 2. The spatial skill levels of the forecasts are shown in section 3. Section 4 shows the relationship between wind and waves and MSLP as well as quantifying how well the models forecast year-to-year variability. Section 5 shows the relationship of the NAO to large-scale environmental conditions,
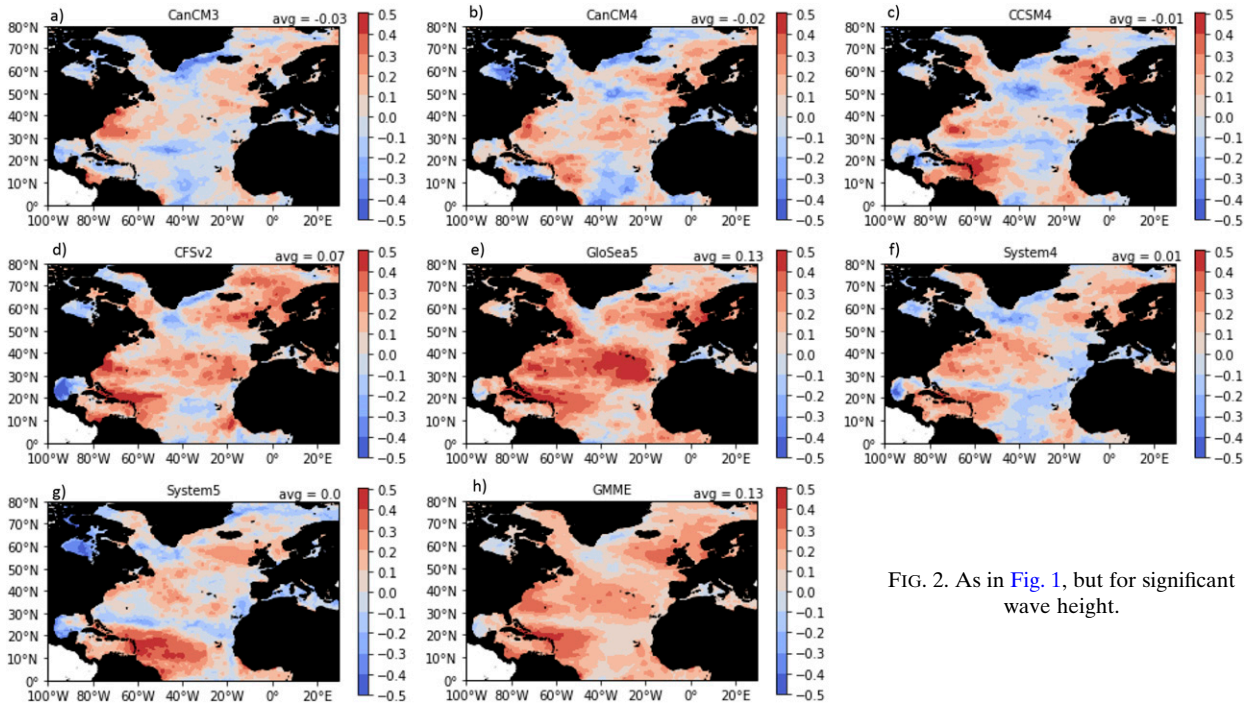
FIG. 2. As in Fig. 1, but for significant wave height.

and section 6 discusses its impact on local surface winds and waves. Examples of where the multimodel ensemble performed well and poorly are given in section 7. Section 8 shows forecast error growth with increasing lead time, and section 9 quantifies the skill improvement of the multimodel forecasts. Finally, the discussion is given in section 10, and the conclusions are presented in section 11.
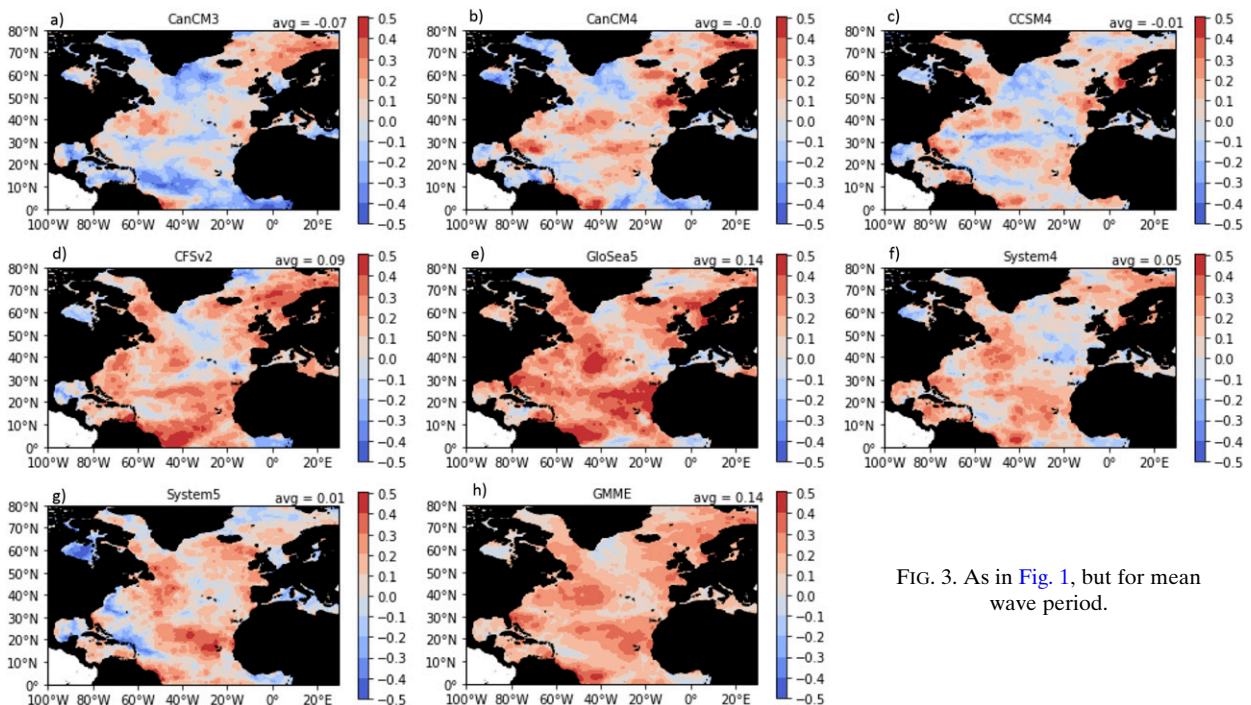


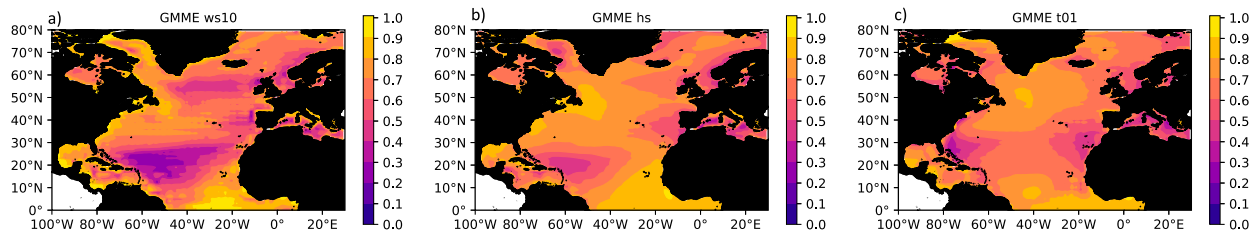FIG. 3. As in Fig. 1, but for mean wave period.

FIG. 4. Signal-to-noise ratios of DJF forecasts using December initial conditions from the GMME: (a) 10-m wind speed, (b) significant wave height, and (c) mean wave period.

## 2. Methodology

### a. Seasonal forecast models

The seasonal forecast models used in this study are shown in Table 1. It is worth noting that CFSv2 and GloSea5 are lagged ensemble starts, whereas all other models initialize on the first of every month. Both CFSv2 and GloSea5 have high ocean resolutions: 0.5° and 0.25°, respectively. Whereas the other models have ocean resolutions of approximately 1°. A higher ocean resolution is advantageous as it provides more robust sea surface temperature gradients, which are important for accurately simulating synoptic atmospheric variability (Kirtman et al. 2012; Parfitt et al. 2017). The models

with a resolved stratosphere are GloSea5, System 4, and System 5. Ten ensembles are used for each model throughout this study to ensure a fair comparison and to remove any bias caused by the lagged starts of CFSv2 and GloSea5. All datasets have been interpolated to 1°. The data are averaged over the winter months of December–February (DJF) for analysis after forcing the wave model with high-frequency 6-hourly wind data. The time period of interest is from winter 1992/93 to winter 2009/10. While there are issues with using a small sample size (Siegert et al. 2016; O'Reilly et al. 2017a), it is the only time period with overlapping data for all seven models. The environmental variables of SST, MSLP, geopotential height at 500 hPa ($Z_{500}$), and zonal
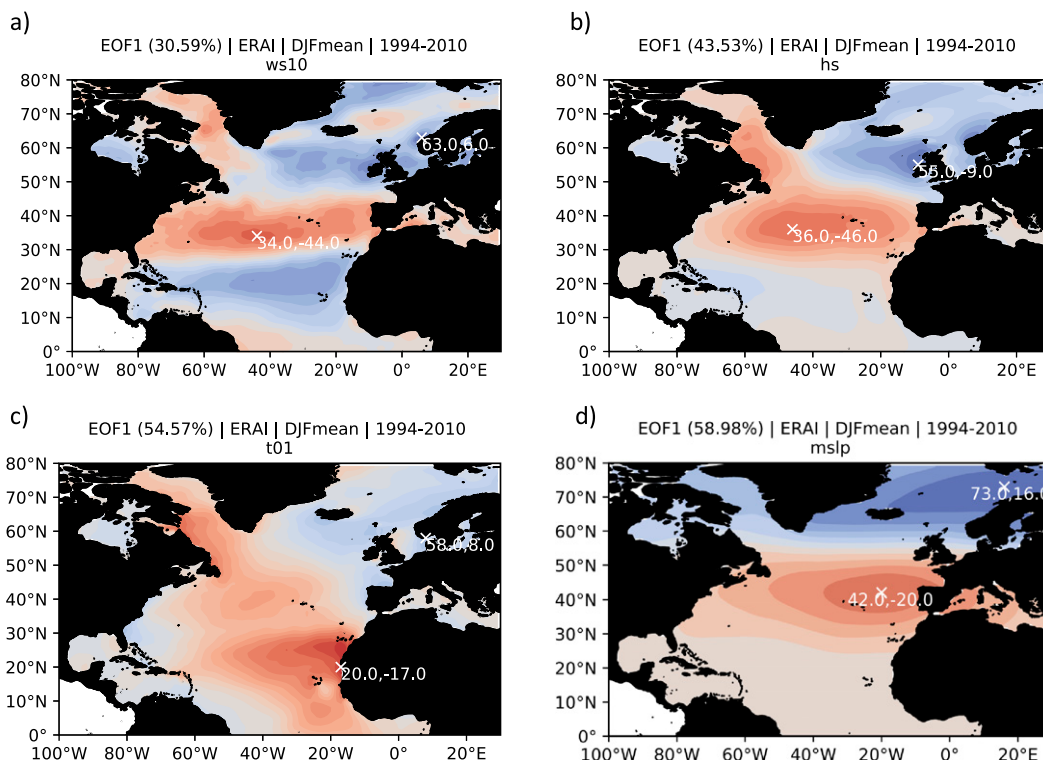


FIG. 5. The observed loading patterns of the first EOF for the DJF average: (a) 10-m wind speed, (b) significant wave height, (c) mean wave period, and (d) MSLP in the spatial domain 100°W–30°E, 0°–80°N, for the years 1994–2010 in the ERA-Interim dataset. The eastern Pacific and land points have been masked out of the calculation.
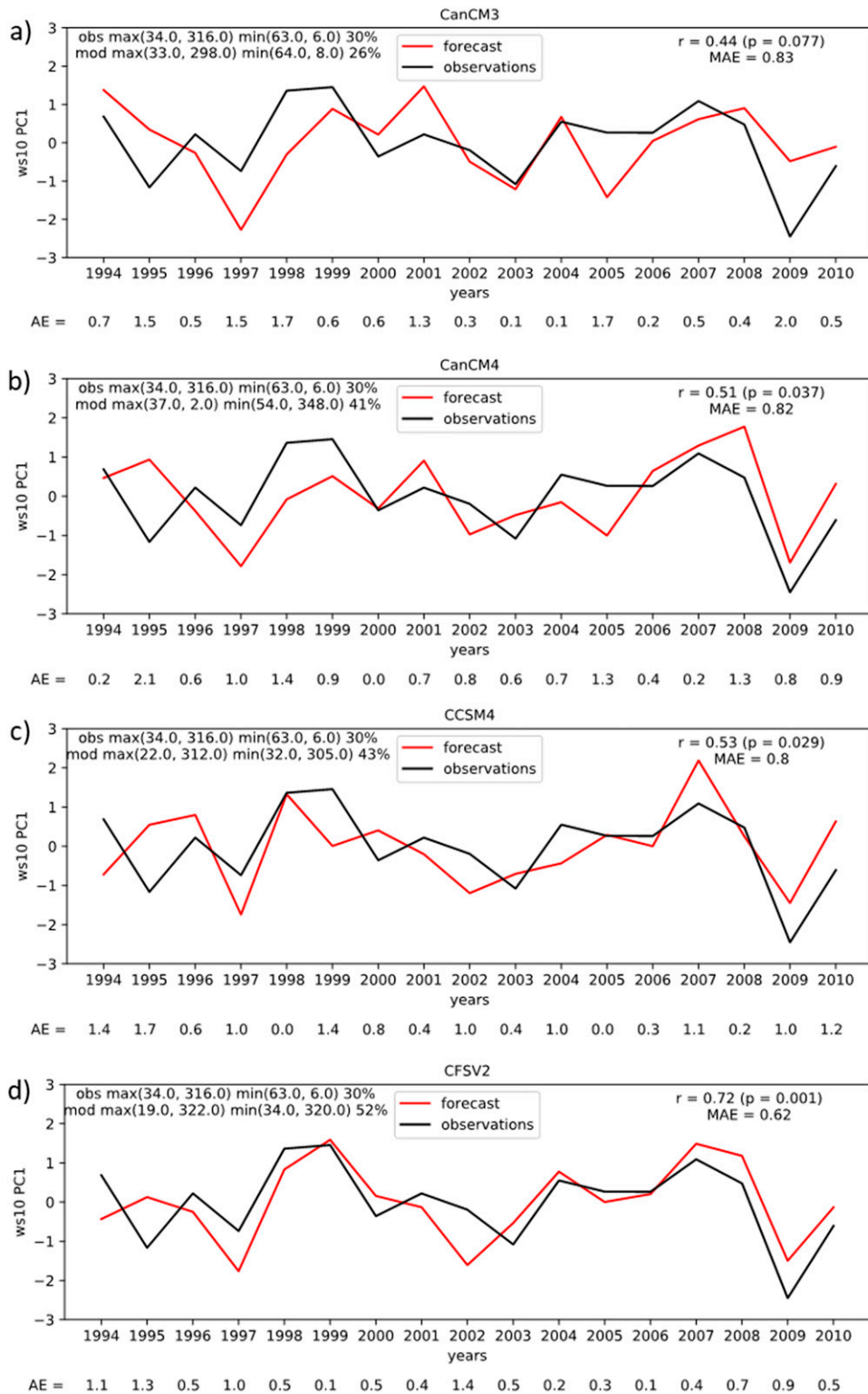
FIG. 6. The black line in all panels indicates the observed 10-m wind speed PC1 corresponding to the EOF in Fig. 5a. The red line is the model forecast using December initial conditions for (a) CanCM3, (b) CanCM4, (c) CCSM4, (d) CFSV2, (e) GloSea5, (f) System 4, (g) System 5, and (h) GMME. The correlation coefficient, *p* value, and mean absolute error are given in the top right-hand corner for each model. The absolute error for each year is given below the *x* axis for each model. The locations of the maximum and minimum of the EOF are shown in the top left along with the variance explained by the EOF.
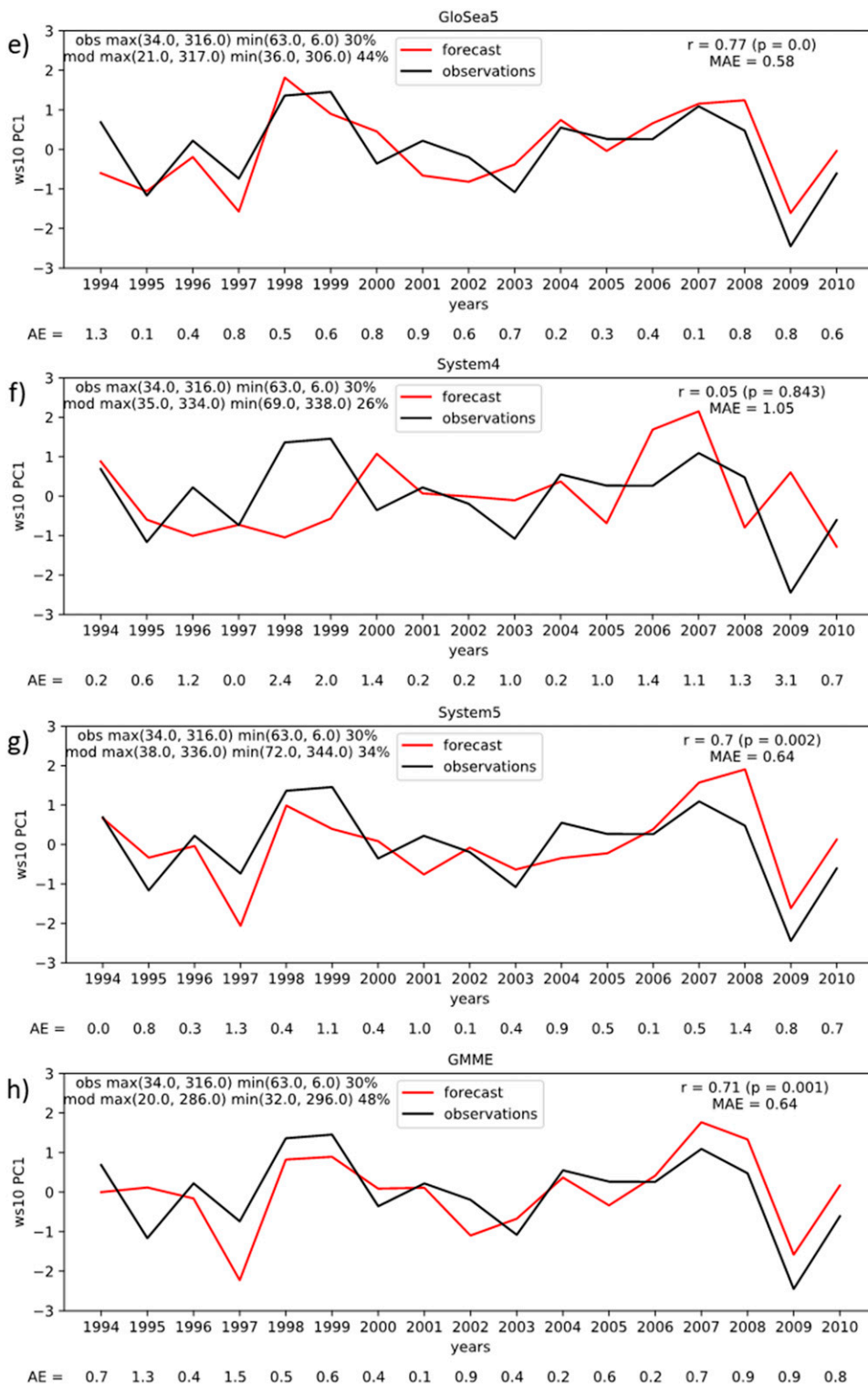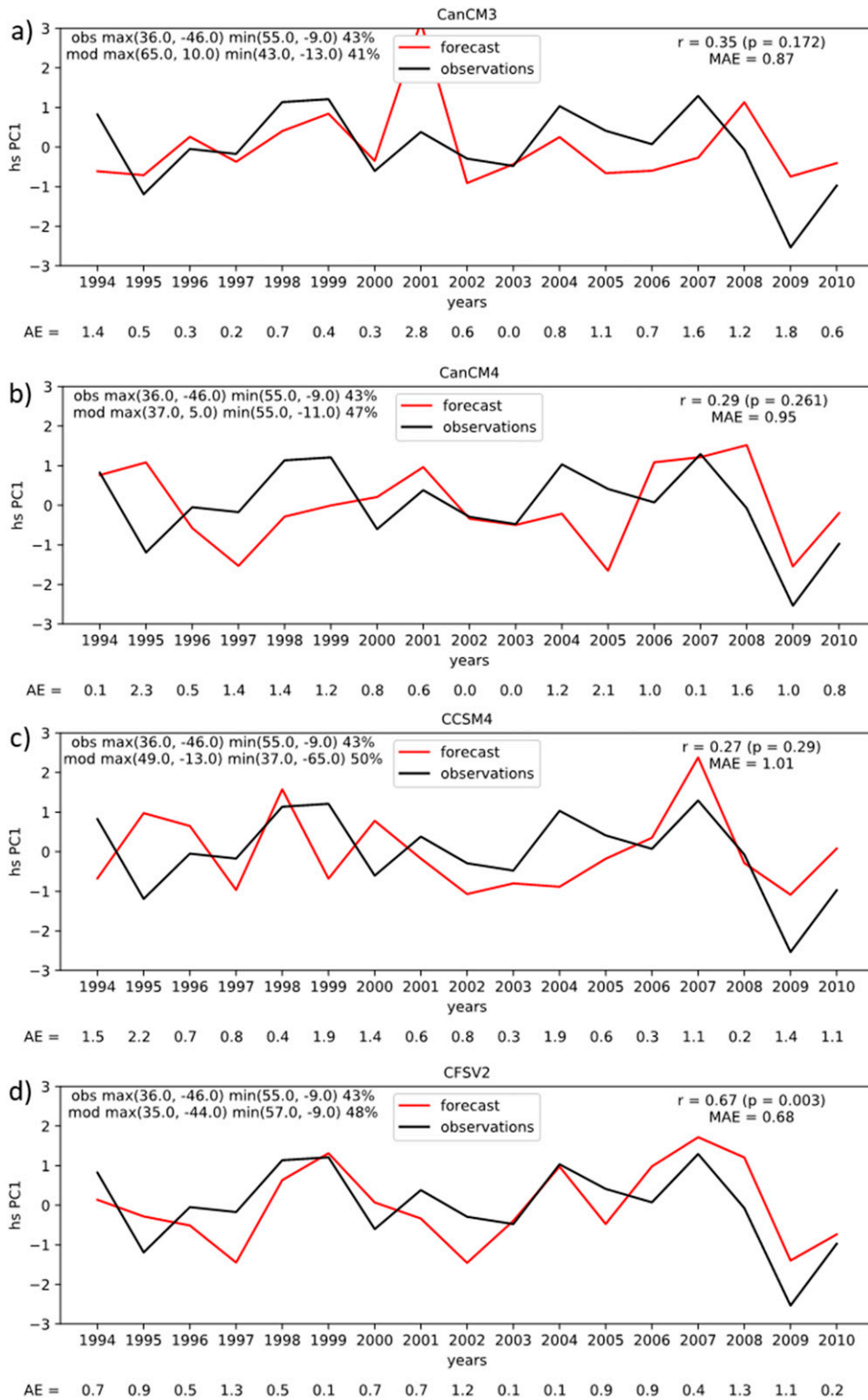
FIG. 6. (*Continued*)

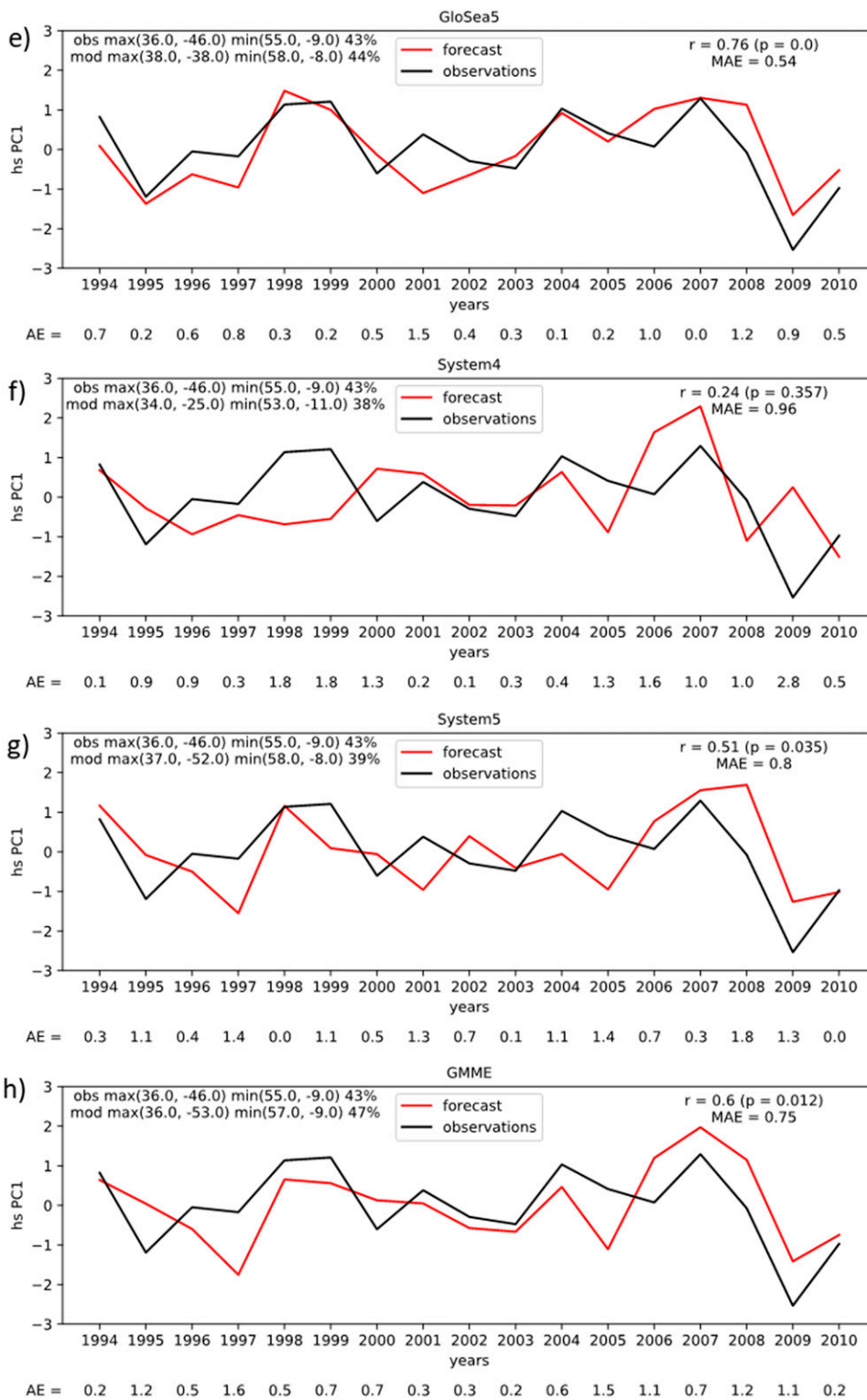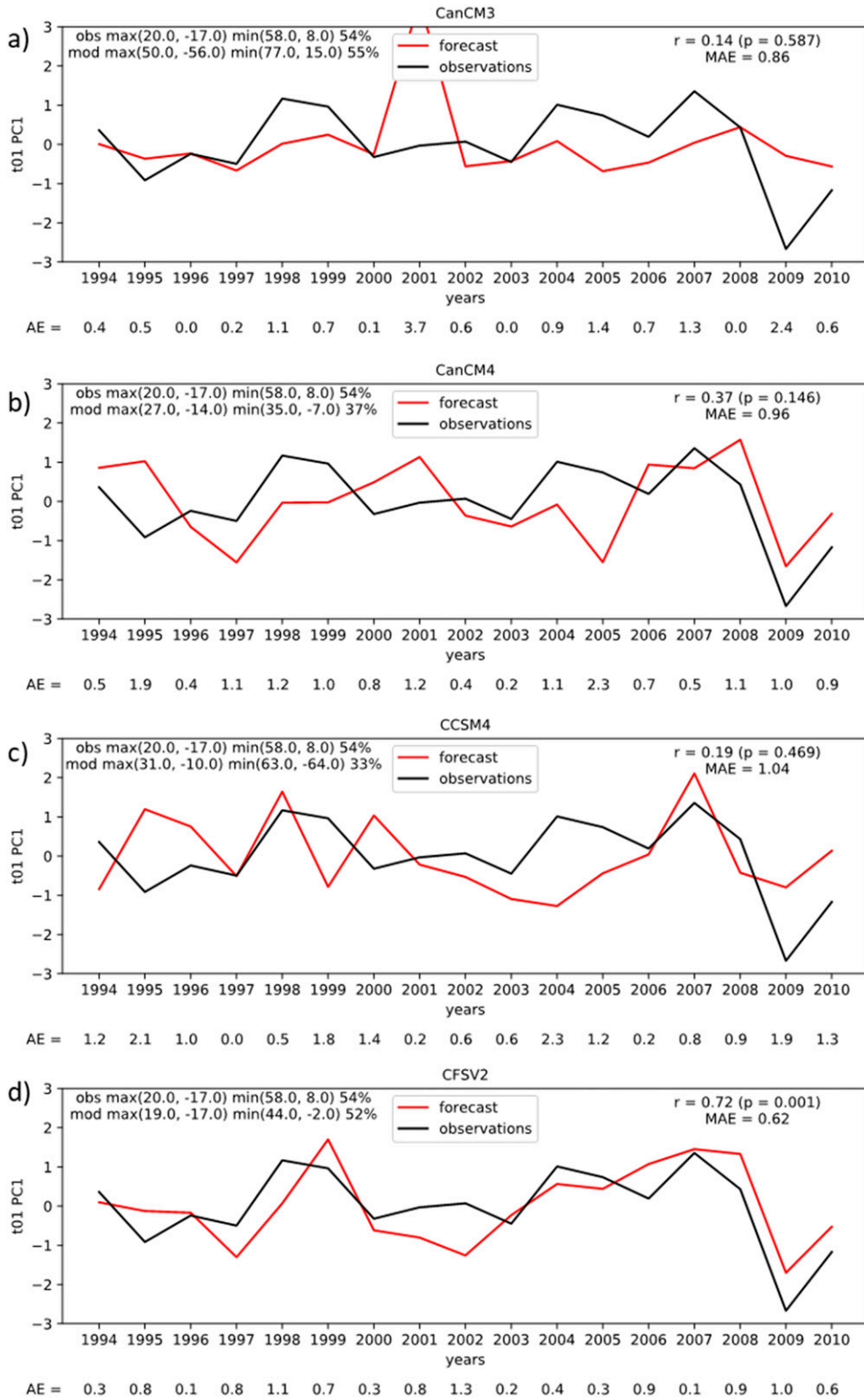FIG. 7. As in Fig. 6, but for significant wave height.

FIG. 7. (*Continued*)

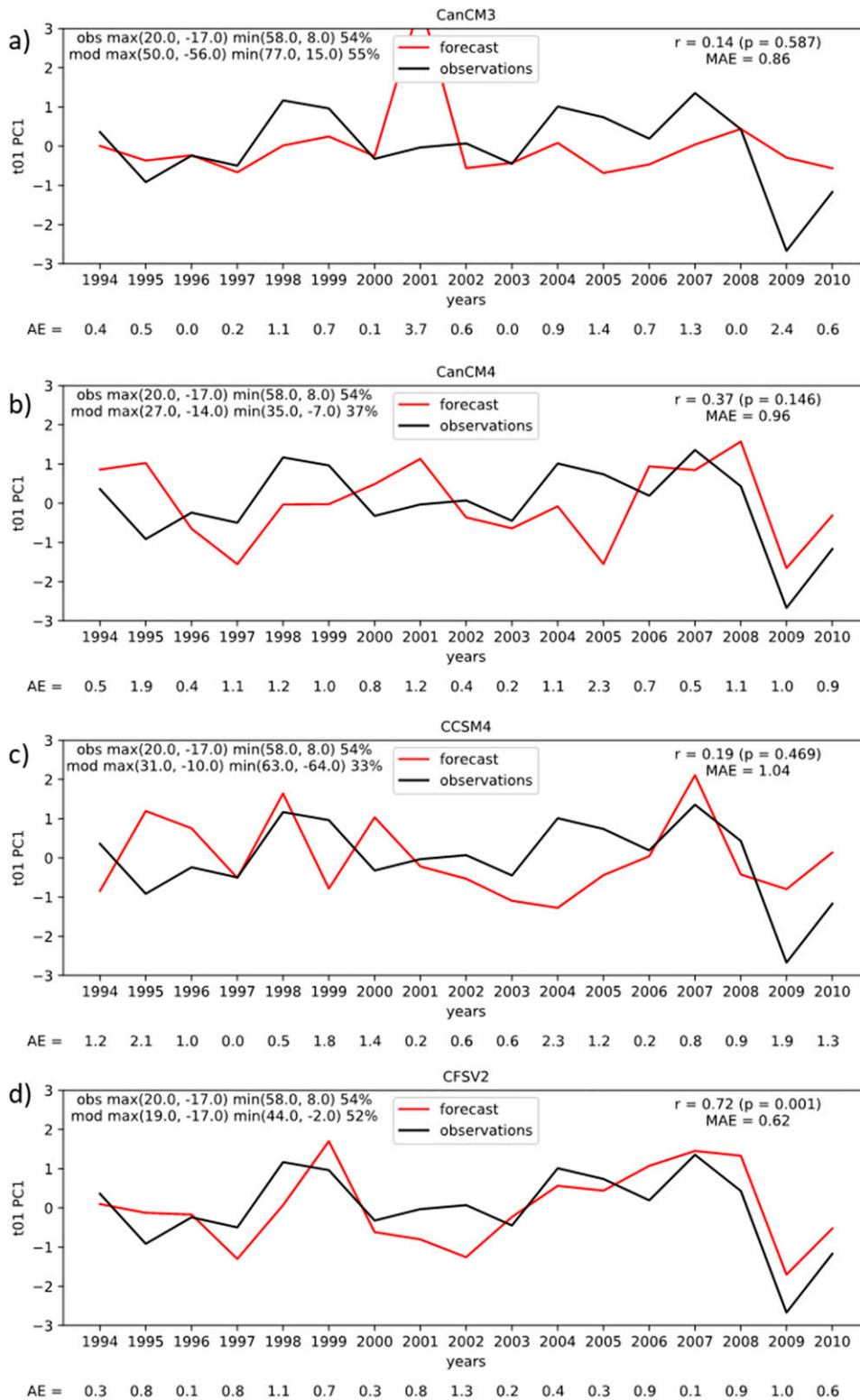FIG. 8. As in Fig. 6, but for mean wave period.

FIG. 8. (*Continued*)

wind at 200 hPa ($U_{200}$) are investigated in relation to the NAO. Unfortunately, data for sea ice, stratosphere diagnostics, and ocean heat content diagnostics were not available for all seven models and therefore are not included in this study.

### b. Ocean wave model

The 10-m winds from the seasonal forecast models are used to force WaveWatchIII experiments. The 4.18 version (Tolman et al. 2014) is used with Ardhuin et al. (2010) wind input and dissipation source terms. The experiments are integrated over a large Atlantic domain (100°W–30°E, 70°S–80°N) to ensure swell waves are captured. ETOPO1 (Amante and Eakins 2009) is used for the bathymetry, and islands smaller than 1° are included at obstruction points defined by Chawla and Tolman (2008). In this study the wave variables of significant wave height ($H_s = 4\sqrt{E}$), where $E$ is the total variance given as the sum over all frequency and direction bins: $\int_0^{2\pi} \int_0^{\infty} F(f, \sigma)\, df\, d\sigma$, and mean wave period ($T_{01} = T_{m0,1} = 2\pi\sigma^{-1}$), are investigated.

### c. Reanalysis data

Model forecast results are compared to reanalysis data, which are used as close to the observed values and have the advantage of spatial homogeneity. The European Centre for Medium-Range Weather Forecasts interim reanalysis (ERA-Interim; Dee et al. 2011) is used as previous studies have shown, compared to other reanalyses, that ERA-Interim has the smallest bias for 10-m winds (Stopa and Cheung 2014). The horizontal resolution of ERA-Interim is ~80 km; however, the data are interpolated to 1°. Even though ERA-Interim is coupled to a wave model, it uses a different wave model than that used in this study (Wave Assimilation Model; Bidlot et al. 2007). Therefore, the ERA-Interim winds are used as input to the WaveWatchIII model and compared to the forecasts. We have opted not to bias correct the winds, such as is done in Durrant et al. (2013, 2014), because of the complexity of a multimodel approach. Using the unbiased winds from all models and ERA-Interim allows for a fair comparison. The surface winds provide mechanic energy into the ocean, which generates wind waves (Simmonds and Keay 2002). The wind waves develop into swell waves once they propagate away from the energy source.

### d. Forecast skill assessment

The quality of the forecasts is assessed using the probabilistic verification metric of rank probability skill score (RPSS). The RPSS measures the probabilistic skill relative to a climatological forecast and a value greater

than zero is said to be better than using climatology as the predictor (Weigel et al. 2007). The systematic error is removed from the models by subtracting the observed climatology from the predictor climatology before the predictions and observations are split into terciles. This is discussed fully in Bell and Kirtman (2018).

## 3. Spatial probabilistic skill assessment

### a. 10-m wind speed

The ability of the models to predict the DJF 10-m wind speed using December initial conditions is shown in Fig. 1. RPSS values greater than zero, in red, show where the forecast models have skill. Figs. 1a–g show the skill for the individual models in alphabetical order as they are presented in Table 1. The models have varying levels of skill, and the patterns are slightly different for each model. The west subtropical region around 30°N is the most predictable region as all models tend to be able to forecast this region. Skill becomes noisier in the extratropics; however, some models such as CCSM4, CFSv2, and GloSea5 are skillful around the British Isles. CFSv2 and GloSea5 have high levels of skill across the entire North Atlantic, as shown by the average RPSS value given at the top right in each panel. RPSS is noisier in the other models, and the spatial average RPSS value is close to zero. Figure 1h shows the RPSS for the GMME. The spatial pattern is smooth, and the spatially averaged RPSS value is greater than any of the individual models.

### b. Significant wave height

The RPSS of forecast $H_s$ (Fig. 2) is mostly similar to the RPSS of the 10-m wind speed (Fig. 1), but there are spatial differences. Skill in the 10-m wind speed is related to skill in $H_s$ downstream of the winds such as in the southwest and northeast regions. This is due to wave growth along the direction of the winds and is observed in other basins (e.g., Bell and Kirtman 2018). Skill can be observed in three distinct regions: the western tropical North Atlantic, the subtropical North Atlantic, and the northeast North Atlantic. CFSv2 and GloSea5 have the highest levels of forecast skill compared to other models. The GMME has a similar level of skill to GloSea5 when the RPSS is averaged over all grid points in the spatial domain.

### c. Mean wave period

Figure 3 shows the RPSS of the forecast mean wave period $T_{01}$. Not only is $T_{01}$ related to the 10-m wind

speed but to the wind direction and fetch length as well. The three regions of predictability have a southwest–northeast tilt, which reflects the $T_{01}$ mean climate. CFSv2 and GloSea5 are the most skillful models across the entire basin but System 5 has large regional RPSS values in the middle of and to the northeast of the basin. The regional skill in forecasting 10-m wind speed in CCSM4 does not translate into a skillful $T_{01}$ forecast, most likely because of limited skill in the wind–wave generation regions (the west coast of Africa and the U.S. East Coast). The winds in these regions are important as errors in wind direction and fetch length here will lead to errors in $T_{01}$ downstream. Similar to Fig. 2, the GMME forecast of $T_{01}$ is as skillful as the best model, GloSea5.

### d. Signal-to-noise ratio

The RPSS results should be looked at in context with the signal-to-noise ratio (StN ratio; e.g., Kumar 2009) to provide an understanding of how confident a forecast is. We use the methodology of Athanasiadis et al. (2017) and define signal as the standard deviation of the ensemble means and noise as the standard deviation of the ensemble members. This calculation yields a range between zero and one with zero indicating regions that have large noise and therefore large ensemble spread. We show the StN ratio for 10-m wind speed (Fig. 4a), $H_s$ (Fig. 4b), and $T_{01}$ (Fig. 4c) from the GMME. It is interesting to note the regions of low StN ratio tend to occur downstream of the large-scale circulation, such as to the east of the Caribbean and to the west of western Europe. Therefore, forecast skill in the center of the basin, where the StN ratio is large can be viewed with a high level of confidence.

## 4. Temporal deterministic skill assessment

The spatial pattern of variability of winter wind and waves in the North Atlantic is largely explained by the first empirical orthogonal function (EOF1) in the spatial domain 100°W–30°E, 0°–80°N. We show the spatial loading pattern of EOF1 for 10-wind speed (Fig. 5a), $H_s$ (Fig. 5b), $T_{01}$ (Fig. 5c), and MSLP (Fig. 5d); in addition, the variances explained by EOF1 are 31%, 44%, 55%, and 59%, respectively. The spatial patterns look markedly similar for all panels. For MSLP and $H_s$ there is a dipole in the northeast North Atlantic that resembles that of the NAO. From hereon we refer to MSLP EOF1 and the NAO. For 10-wind speed the pattern is more tripolar, as there is some variability in the tropics. Finally, EOF1 of $T_{01}$ shows a northeast–southwest tripole, along the

TABLE 2. MSLP PC1 model ranking. The AE model sum is the summation of the absolute error from all model forecasts.

| | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | Rank sum |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AE model sum | 4.6 | 9.4 | 5.9 | 8.7 | 7.1 | 5.9 | 5.3 | 4.3 | 2.9 | 1.9 | 4.5 | 6 | 5 | 5.4 | 9.1 | 10.4 | 4.5 | |
| CanCM3 | 1 | 5 | 3 | 8 | 7 | 2 | 7 | 7 | 2 | 8 | 3 | 8 | 2 | 4 | 6 | 7 | 5 | 85 |
| CanCM4 | 2 | 8 | 6 | 4 | 6 | 7 | 3 | 5 | 5 | 4 | 8 | 7 | 4 | 1 | 7 | 4 | 7 | 88 |
| CCSM4 | 8 | 1 | 1 | 6 | 1 | 5 | 6 | 2 | 7 | 6 | 7 | 3 | 3 | 8 | 1 | 2 | 8 | 75 |
| CFSv2 | 6 | 3 | 2 | 2 | 5 | 1 | 4 | 4 | 8 | 3 | 2 | 2 | 6 | 2 | 4 | 1 | 1 | 58 |
| GloSea5 | 7 | 6 | 7 | 3 | 3 | 3 | 1 | 1 | 4 | 1 | 4 | 6 | 8 | 5 | 3 | 6 | 3 | 74 |
| System 4 | 3 | 2 | 8 | 1 | 8 | 8 | 8 | 1 | 1 | 5 | 1 | 4 | 7 | 7 | 2 | 8 | 6 | 80 |
| System 5 | 5 | 7 | 4 | 7 | 2 | 6 | 2 | 8 | 6 | 7 | 6 | 1 | 1 | 3 | 8 | 3 | 4 | 80 |
| GMME | 4 | 4 | 5 | 5 | 4 | 4 | 5 | 3 | 3 | 2 | 5 | 5 | 5 | 6 | 5 | 5 | 2 | 72 |
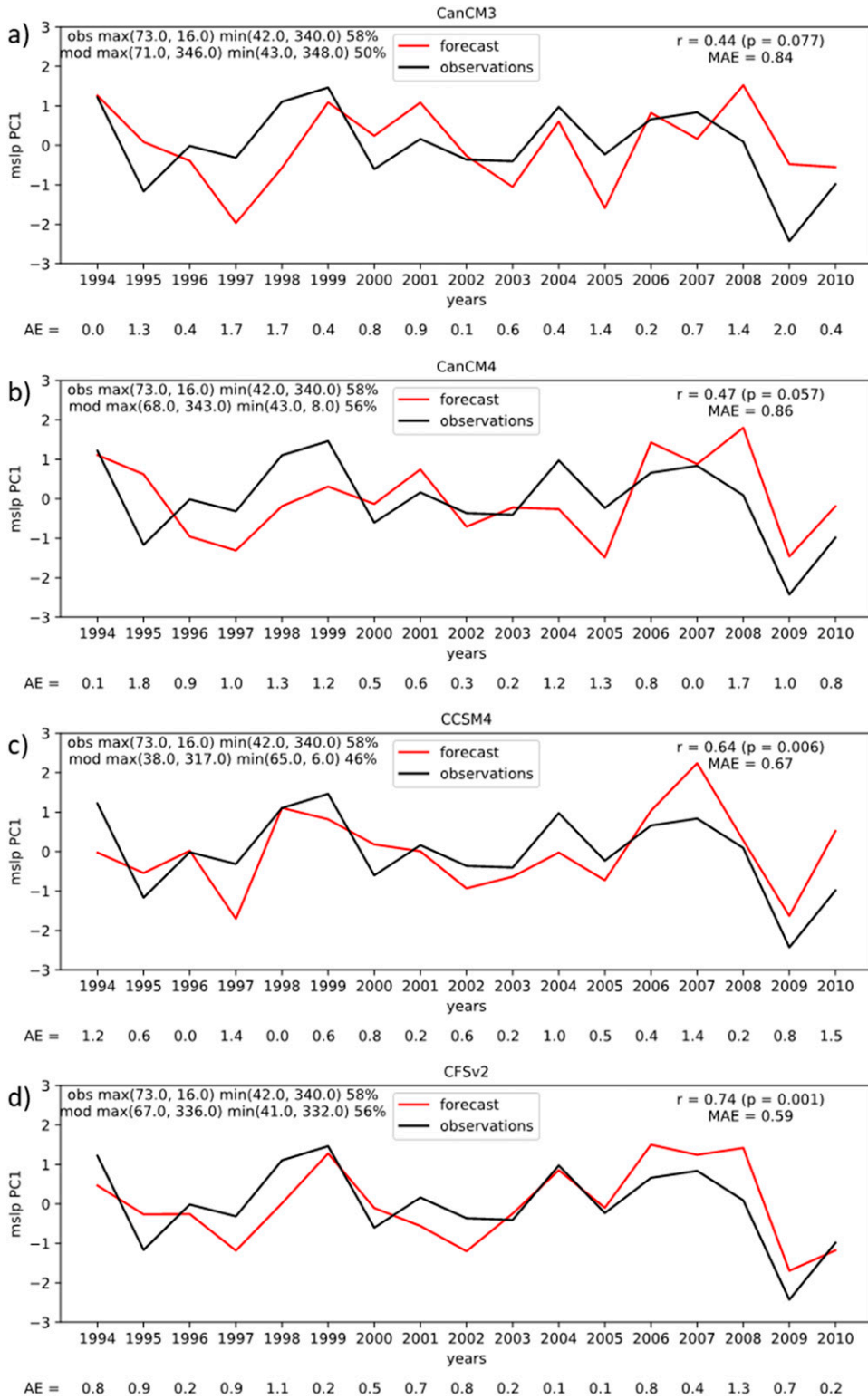
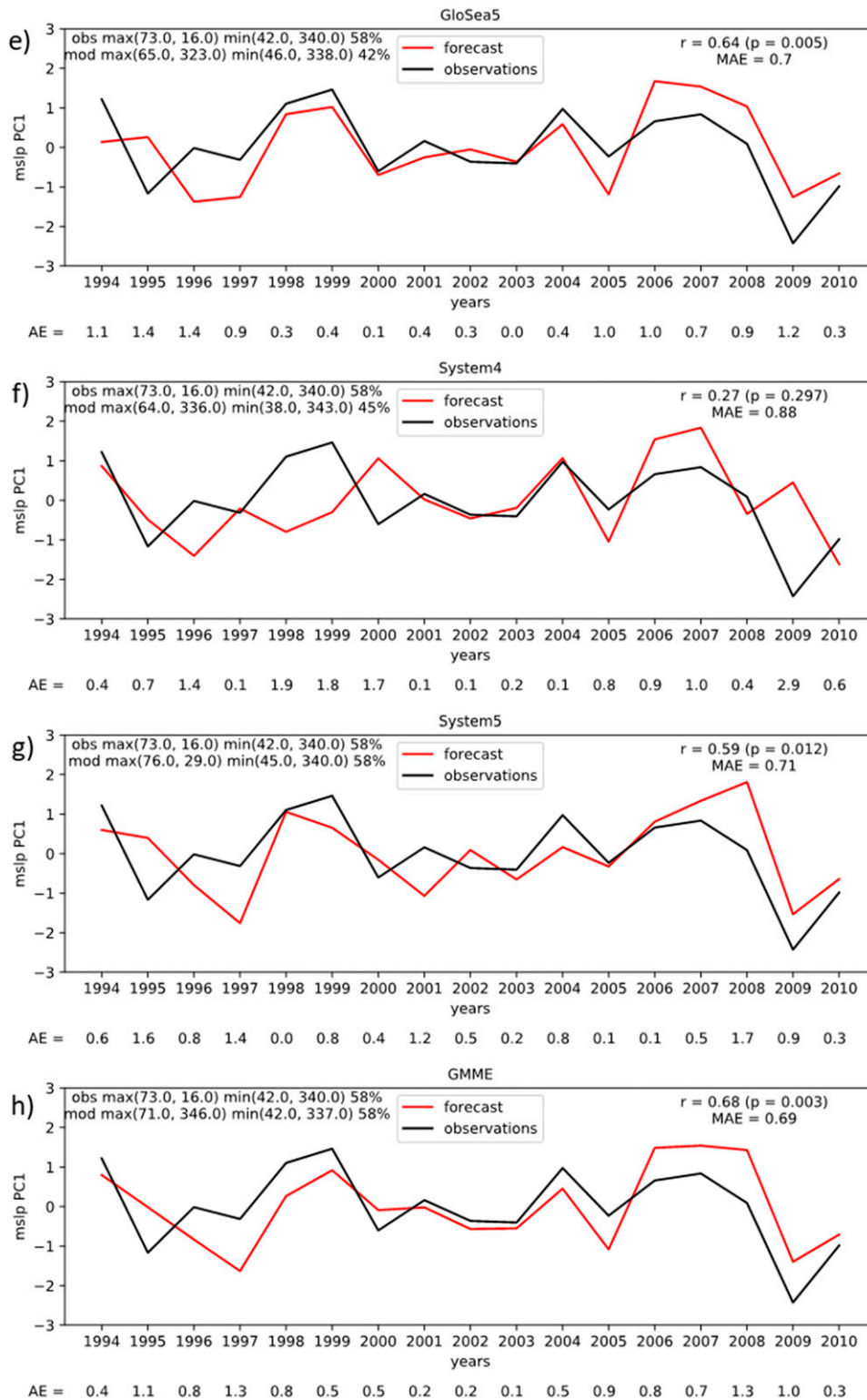FIG. 9. As in Fig. 6, but for mean sea level pressure.
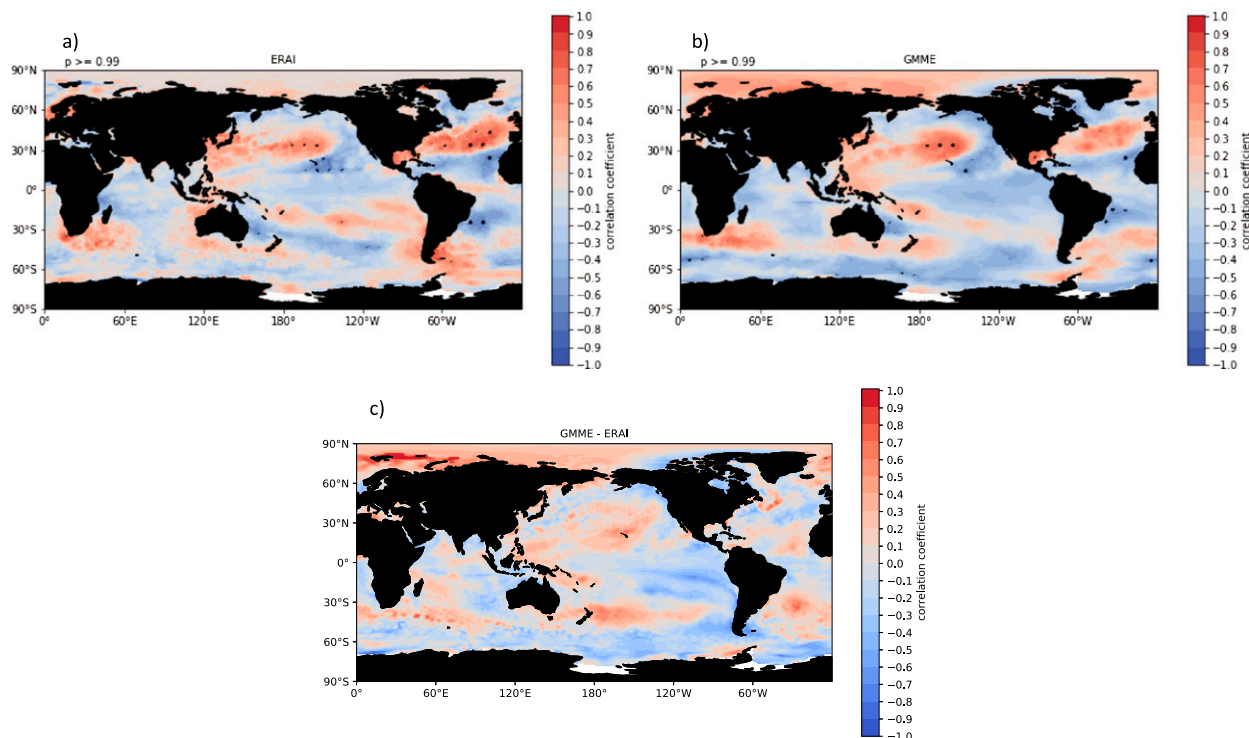
FIG. 9. (*Continued*)

FIG. 10. (a) Observed correlation coefficient of DJF SST and MSLP PC1 (see Fig. 9). (b). GMME mean forecast of DJF SST and MSLP PC1 correlation coefficients using December initial conditions. Stippling shows points significant at the 99% level calculated using a one-sided $t$ test for positive values and calculated using the survival function for negative values.

direction of the 10-m wind. There is a strong case to be made here that a skillful forecast of surface wind and waves largely depends on a skillful forecast of the NAO due to similar patterns of variability. For the temporal analysis we use the first principal component (PC1) related to EOF1 calculated using the methodology of Dawson (2016). The advantage of using a PC instead of a point-based analysis is that it will take into account structural differences between the models.

*a. 10-m wind speed time series*

The red lines in Figs. 6a–g are ensemble-mean forecast PC1s for the individual models using December initial conditions. Additional information about the model's PC1 is given in the top-left corner of the panels. Figure 6h is the GMME PC1 forecast. The correlation coefficient, $p$ value of the correlation coefficient, and mean absolute error are given in the top-right corner of all panels. The absolute error for each year is given below the panel. The PC1 is well predicted in CFSv2, GloSea5, and System 5, with correlation coefficient values of 0.72, 0.77, and 0.7, respectively. The GMME forecast has a correlation coefficient of 0.71 and is similar to the superior models. The other

models, excluding System 4, have correlation values between 0.44 and 0.53. System 4 has a low correlation coefficient value and is largely a results of two years when it failed to capture the sign of the forecast: 1998 and 2009.

*b. Significant wave-height time series*

The results of year-to-year forecasting of $H_s$ (Fig. 7) are similar to that of 10-m wind speed but with slightly lower skill. For example, the best models—CFSv2, GloSea5, and System 5—have correlation coefficient values between 0.51 and 0.76. This is not surprising as the observed PC1 results are very similar for $H_s$ and 10-m wind speed, such as the large negative event in 2009. For the GMME (Fig. 7h) the forecasts are generally good. However, the 1997 forecast predicted a negative pattern of EOF1, whereas a neutral pattern was observed.

*c. Mean wave period time series*

Most of the models that did not accurately forecast the time series of 10-m wind speed (and therefore $H_s$) did not forecast the time series of $T_{01}$ (Figs. 8a–c,f). The 10-m wind speed alone does not explain the results of the $T_{01}$ variability. It is the large-scale wind climate that
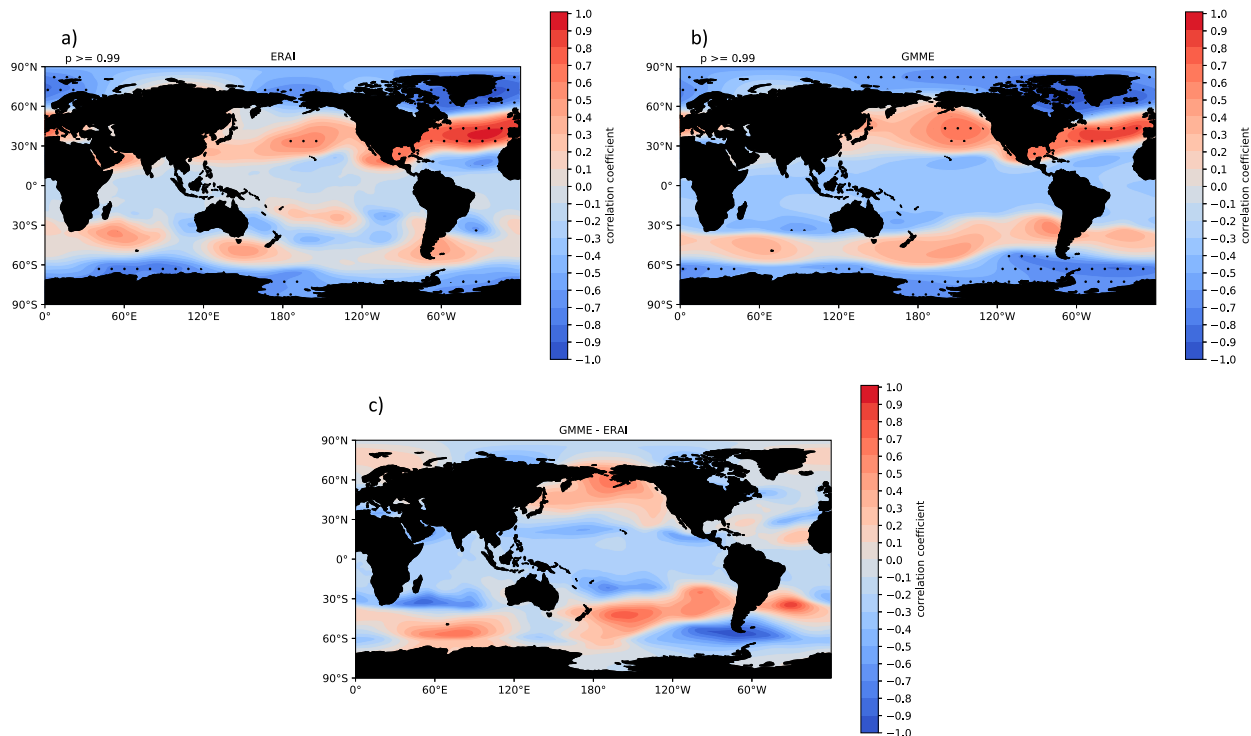
FIG. 11. As in Fig. 10, but for geopotential height at 500 hPa.

controls the wind direction and fetch length. Therefore, local and remote regions require accurate wind forecasts. The models that do accurately forecast the time series of 10-m wind speed also capture the variability of $T_{01}$ to a strong degree, such as GloSea5, which has a correlation coefficient of 0.84. For the models that have poor forecasts for the 10-m wind speed the resultant $T_{01}$ is worse. In these cases (CanCM3, CCSM4, and System
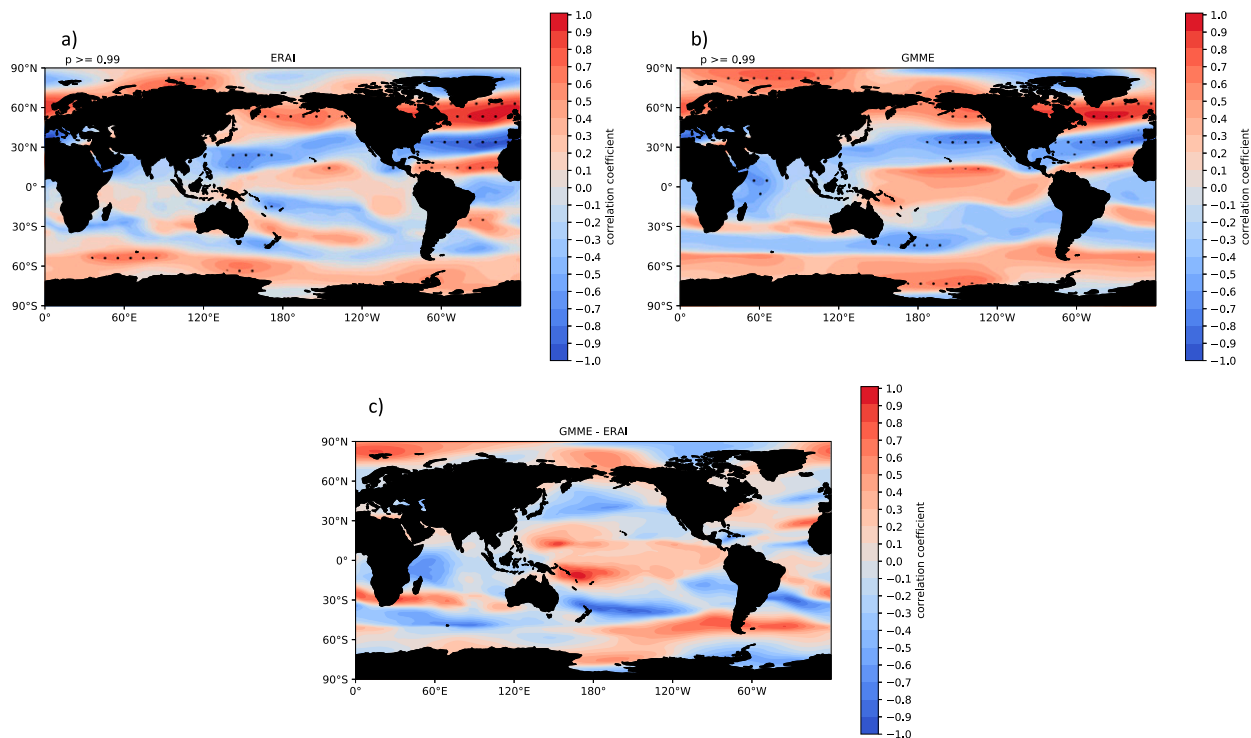


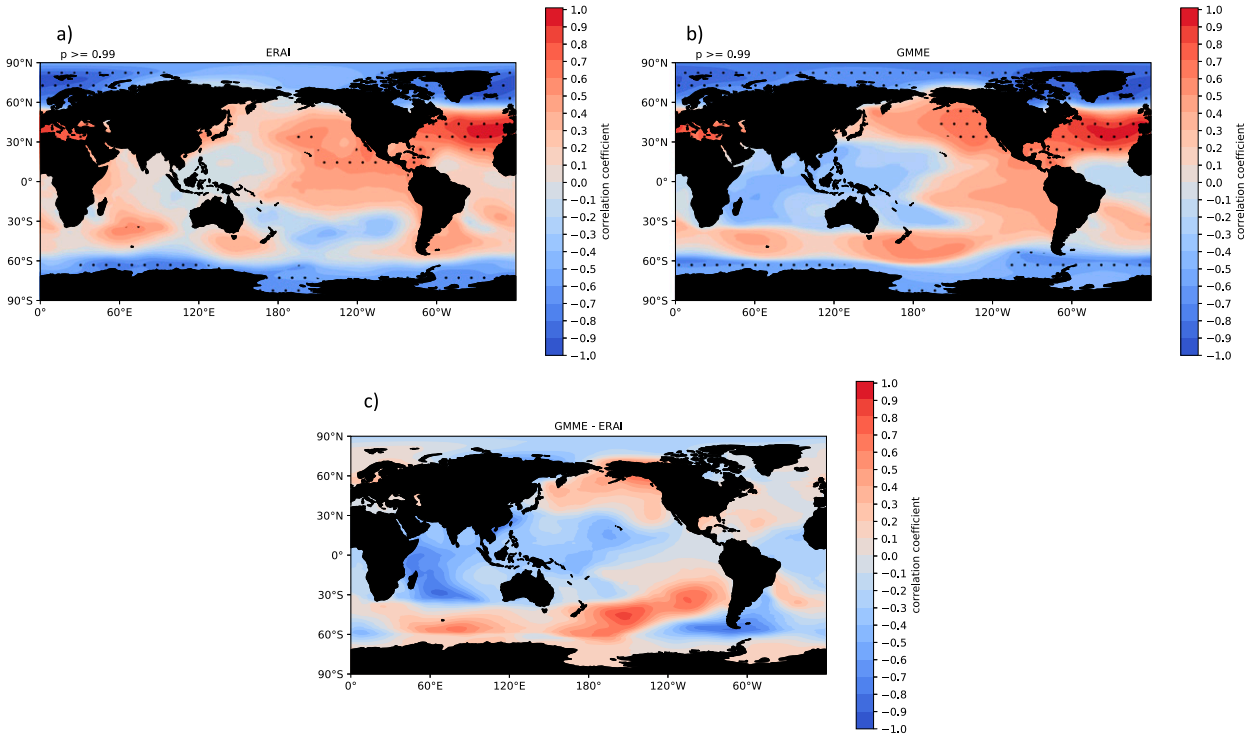FIG. 12. As in Fig. 10, but for zonal wind at 200 hPa.

FIG. 13. As in Fig. 10, but for MSLP.

4) there is error propagation arising from differences in wind speed, wind direction, and fetch length.

### d. NAO time series

The NAO is well predicted in CFSv2 and GloSea5, with correlation coefficient values of 0.74 and 0.64, respectively. CCSM4 is shown to predict the NAO well with a correlation coefficient value of 0.64; however, wave forecast skill in CCSM4 is limited to the northeast North Atlantic, where the NAO is most dominant. The GMME forecast has a correlation coefficient of 0.69 and is only behind CFSv2. The results are similar to
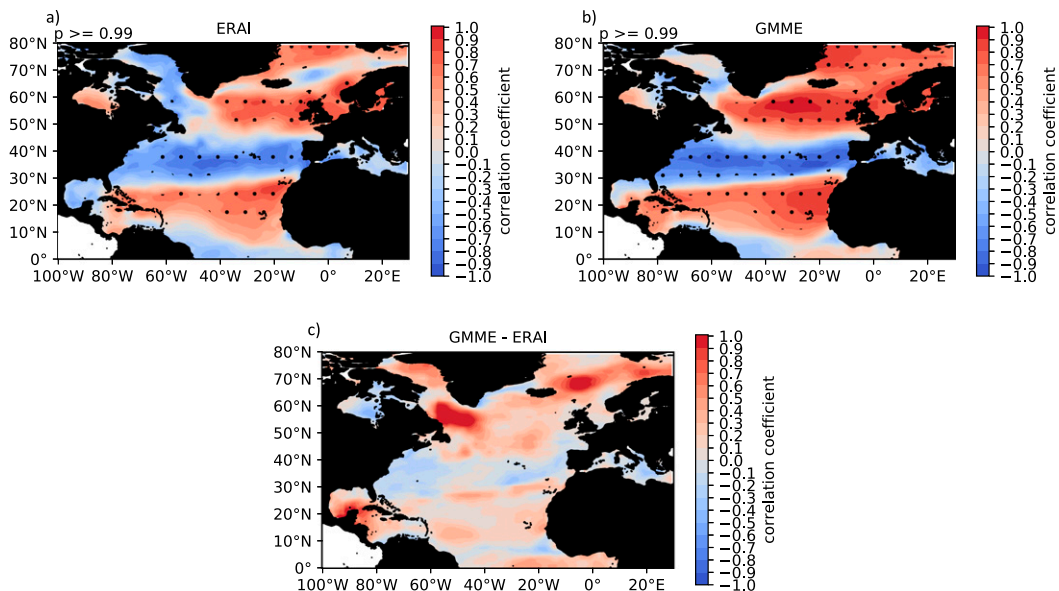


FIG. 14. As in Fig. 10, but for wind speed at 10-m height in the North Atlantic.
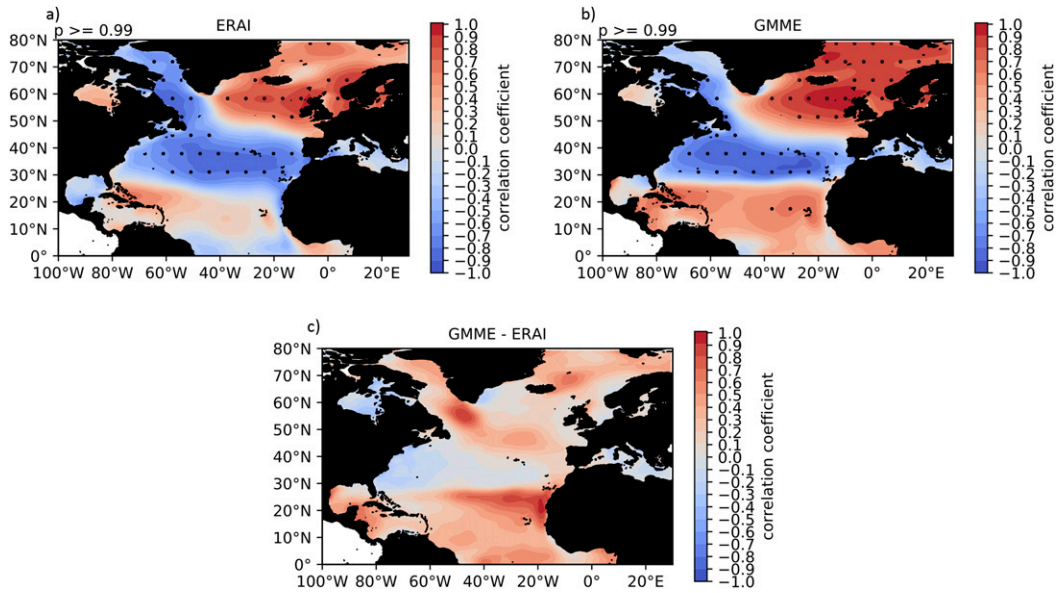
FIG. 15. As in Fig. 10, but for significant wave height in the North Atlantic.

those of the wind and wave variables discussed above. This demonstrates there is a strong relationship between the NAO, surface winds, and surface waves.

Table 2 shows the ranking of all eight models in terms of yearly absolute error. The last column is the sum of the ranks. Table 2 highlights the importance of using multimodel forecasts, as the "best" model varies from year to year. CFSv2 has the best ranking out of all the models, followed by GMME, GloSea5, and CCSM4. System 4 and System 5 have

similar rankings, followed by CanCM3 and CanCM4. The top row in Table 2 gives the sum of the models' absolute error and defines years that are more predictable than others. This is discussed further in section 7.

## 5. NAO and large-scale environmental conditions

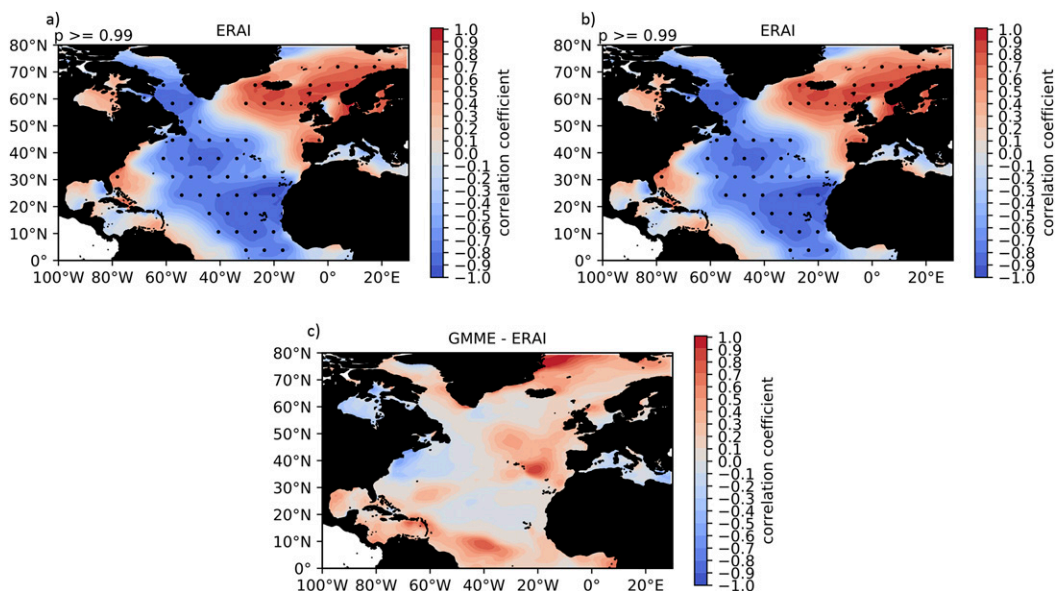The NAO is influenced by a combination of teleconnections from large-scale environmental conditions



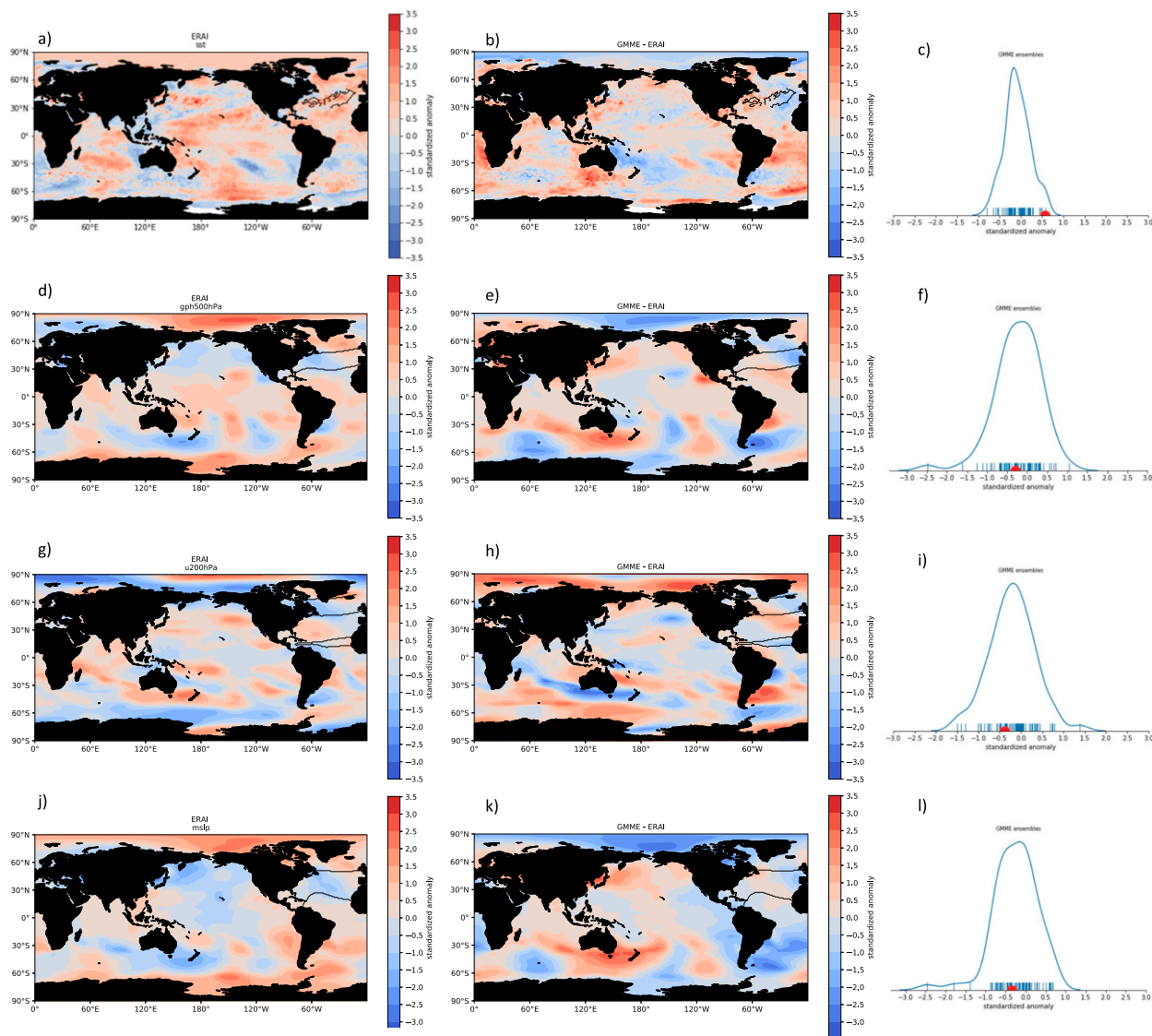FIG. 16. As in Fig. 10, but for mean wave period in the North Atlantic.

FIG. 17. (left) Observed standardized anomalies for DJF 2002/03 of (a) SST, (d) geopotential height at 500 hPa, (g) zonal wind at 200 hPa, and (j) MSLP. The black contour shows regions that have a correlation coefficient with MSLP PC1 greater than 0.5. (b),(e),(h),(k) As in the left panels, but using the GMME mean. (c),(f),(i),(l) Kernel density estimates of values averaged over the black contour in the observed plot. The red dot is the observed value, the short blue lines along the $x$ axis are the data for the individual ensemble forecasts, and the blue line is the kernel density estimate.

as well as internal processes. The results in this section investigate the relationship between large-scale environmental conditions and the NAO. The results ensure that the forecasts are predicting the NAO for the right reasons. Variables considered include SST, $Z_{500}$, $U_{200}$, and MSLP.

*a. Sea surface temperature*

The observed correlation coefficient of MSLP PC1 (the black line in Fig. 9) with SST is presented in Fig. 10a. Significant points at the 99th percentile value are shown via stippling. There is a strong dipole in the

North Atlantic, as discussed in Brayshaw et al. (2009), that has a large influence on the NAO. There are however dipoles outside of the North Atlantic that are correlated with the NAO. Examples include the regions in the subtropical North Pacific and the South Atlantic. Figure 10b shows the same relationship using the GMME forecast and Fig. 10c is GMME minus the observations. The GMME captures the large-scale spatial relationship between SST and the NAO; however, there are some regional differences. For example, the GMME has a larger correlation with eastern tropical Pacific SSTs, whereas this relationship
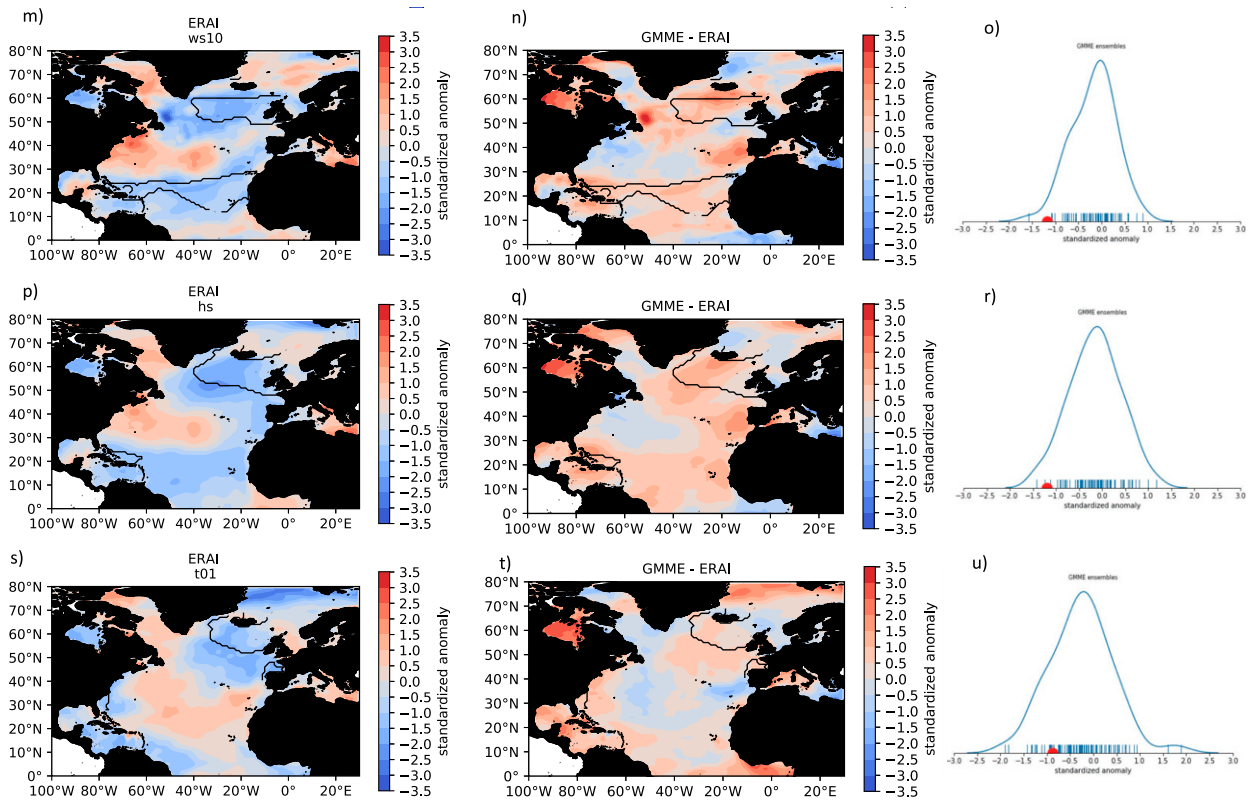
FIG. 17. As in (a)–(l), but for (m)–(o) wind speed at 10-m height, (p)–(r) significant wave height, and (s)–(u) mean wave period.

is more noisy in ERA-Interim. The GMME forecast correlation pattern in the North Atlantic is slightly weaker than observed, and few points are statistically significant.

### b. Geopotential height at 500 hPa

The observed correlation of MSLP PC1 with $Z_{500}$ is presented in Fig. 11a. There is a clear NAO pattern in the North Atlantic and to a lesser degree a signal in the North Pacific. Points in the Southern Ocean that are significant are likely to be related to the southern annular mode, which shows variability over a wide range of time scales (Simmonds and King 2004). The GMME forecast correlation of MSLP PC1 with $Z_{500}$ is very good in the North Atlantic (Fig. 11b). Outside of the North Atlantic, the relationship is stronger in GMME than observed (Fig. 11c), as the GMME forecasts a broad hemispheric oscillation in both hemispheres.

### c. Zonal wind at 200 hPa

The $U_{200}$–NAO correlation coefficient is a noisier field than the previous environmental variables (Fig. 12). The observed pattern in the North Atlantic shows the largest signal in the NAO region but there are regions of alternating correlation coefficient signs that extend south throughout the Atlantic and east to central Europe

(Fig. 12a). There is also a tripolar pattern in the North Pacific. The GMME forecasts of the $U_{200}$–NAO correlation coefficient are much smoother than observed and the banding is too zonal (Fig. 12b). The pattern in the North Pacific is also too strong in the GMME forecasts.

### d. Mean sea level pressure

The observed correlation coefficient of DJF MSLP with North Atlantic MSLP PC1 (Fig. 13a) shows an NAO pattern that is expected, given the construction of MSLP PC1. There is also a strong relationship with mean sea level pressure in the central to northeast Pacific. The region with the largest correlation coefficient is shifted farther north in the GMME forecast compared to the observations (Fig. 13c). In addition, the negative correlation coefficient in the western Pacific is much stronger in the GMME forecast. However, the differences are small in the regions with known physical relationships to the NAO: the North Atlantic and tropical eastern Pacific.

## 6. NAO and winds and waves in the North Atlantic

The impact of the NAO on the winter wind and wave climate in the North Atlantic is presented below for 10-m wind speed, $H_s$, and $T_{01}$.
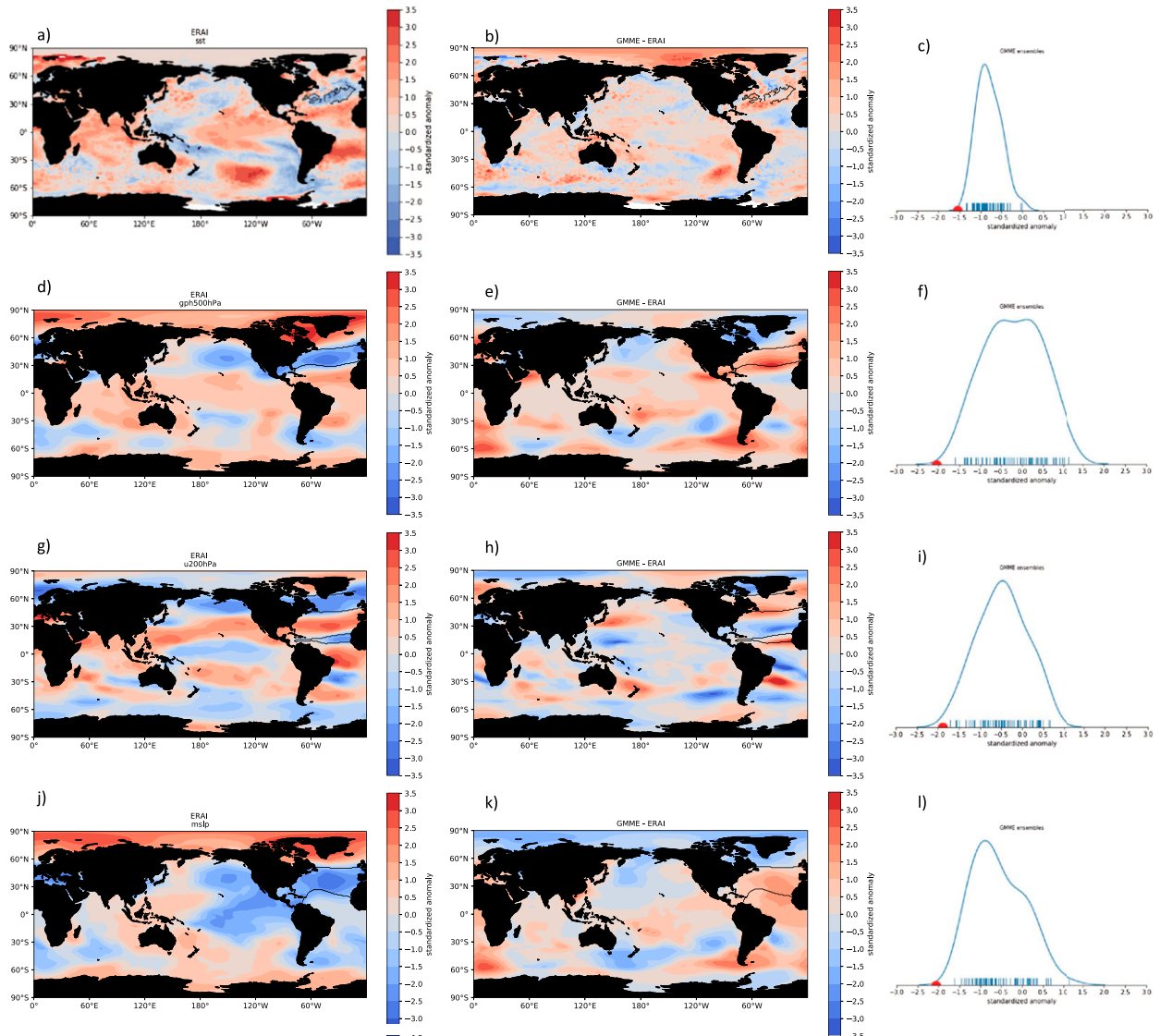
FIG. 18. As in Fig. 17, but for DJF 2008/09.

### a. 10-m wind speed

The NAO's impact on 10-m wind speed is a tripolar pattern (Fig. 14a), as discussed in previous studies (e.g., Fig. 5a in Zubiate et al. 2017). It is encouraging that the GMME forecast captures the large-scale tripolar pattern. Some regional differences exist such as the GMME forecast showing a correlation of the opposite sign in the Gulf of Mexico compared to the observations. The GMME in addition has a much stronger correlation coefficient in the extratropics (Fig. 14c). It is worth noting that the spatial pattern of the NAO–10-m wind speed correlation is strongly related to the RPSS spatial pattern (Fig. 1 c), highlighting the importance of the NAO in predicting winter 10-m wind speed.

### b. Significant wave height

The NAO influences $H_s$ in a northeast–southwest pattern that takes into account wave growth along the wind direction (Fig. 15a). The GMME correctly predicts this large-scale pattern (Fig. 15b) but the regional differences, such as stronger correlation in the tropics and extratropics, arise due to the wind forcing (cf. Fig. 15c with Fig. 14c).

### c. Mean wave period

The NAO tripolar pattern for $T_{01}$ has an east–west pattern along the major axis of the jet stream (Fig. 16a). During NAO positive (negative) years the jet stream is stronger (weaker) and farther north (south).
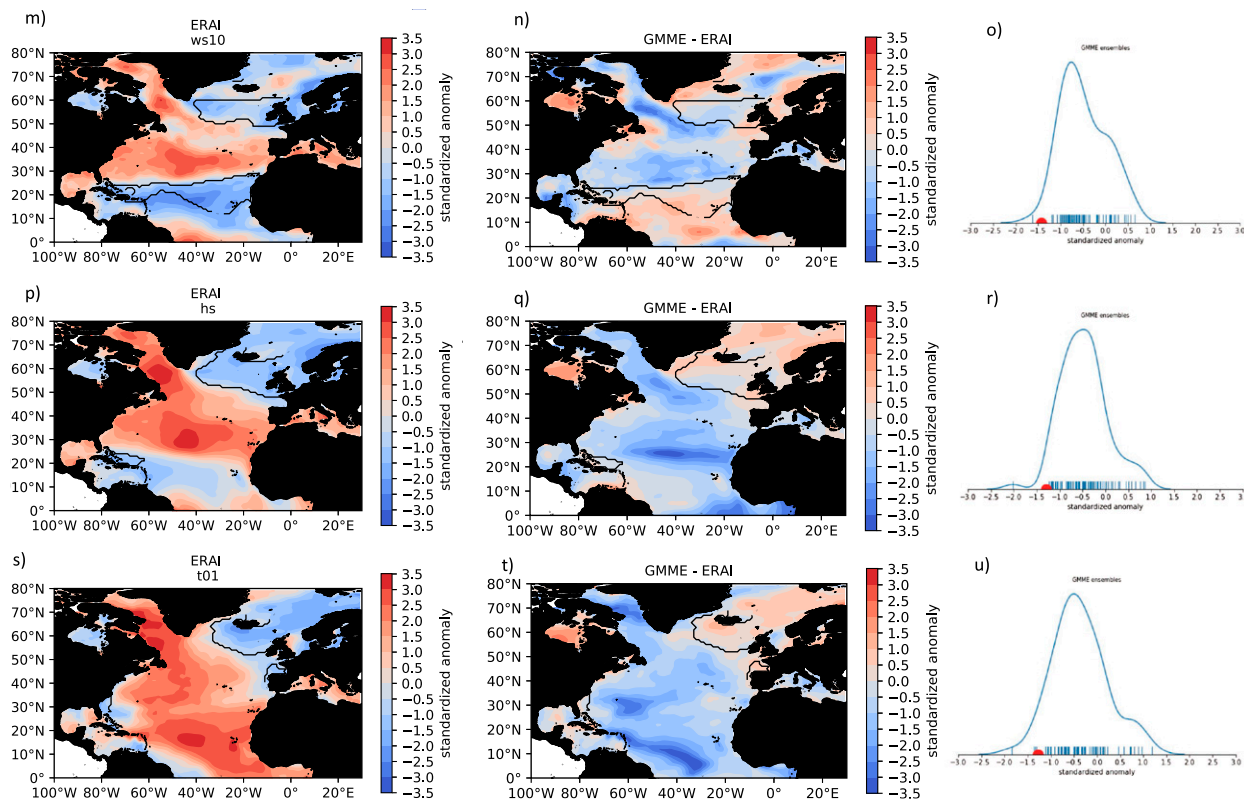
FIG. 18. (*Continued*)

This results in longer- (shorter) period waves in the northeast part of the basin. The differences in the GMME forecast and the observations are small in the central part of the basin, the negative correlation area (Fig. 16c). Similar to 10-m wind speed and $H_s$, the relationship is too strong in the tropics and extratropics.

## 7. Case studies

The advantage of a multimodel ensemble forecast is explored by looking at individual winter forecasts. Spatial anomalies and probability distributions are investigated for the environmental patterns discussed in section 5. These case studies are selected in terms of the best and worst forecasts using the sum of the models' absolute error for each year (Table 2). Winter 2002/03 is chosen as a good forecast as all the models predicted the NAO. All model forecasts, apart from CanCM3, had an absolute error of 0.2 or less for this year for the NAO. Winter 2008/09 is chosen as a poor forecast as all models fail to capture the magnitude of the negative NAO event. The model sum absolute error for this winter is 10.4, of which CanCM3 and System 4 made large contributions to.

### a. Winter 2002/03: A good multimodel ensemble forecast

The NAO during winter 2002/03 was close to zero. It is therefore surprising that this year was well forecast as models tend to struggle to predict near-normal conditions compared to above- or below-average conditions (van den Dool and Toth 1991). The SST in the North Atlantic was slightly warmer than average (Fig. 17a). The GMME forecast is closer to average in the North Atlantic, in particular, in the region where the correlation coefficient is greater than 0.5 for the MSLP PC1 and SST (the black contour in Fig. 17b and see Fig. 10b). The ensembles are averaged over this region and shown as a probability density function in Fig. 17c. The red circle is the observed value. The GMME $Z_{500}$ forecast as shown in Fig. 17f is good, as most ensembles predict near-normal conditions, although two ensembles forecast large negative anomalies. Weaker $U_{200}$ in the region with large correlation with the NAO is correctly forecast in the GMME (Fig. 17h). The difference between the GMME forecast and observed values for MSLP is small in the North Atlantic (Fig. 17k). While the large-scale environmental conditions are well forecast for the neutral NAO winter, its impact on the wind and wave climate is

not as well captured. Most ensembles predict near-normal conditions for the wind and waves (Figs. 17o,r,u); however, below-average conditions were observed.

### b. Winter 2008/09: A poor multimodel ensemble forecast

The NAO during winter 2008/09 is the strongest negative NAO in the forecast period. There were large SST anomalies in the North Atlantic in the region, which is strongly related to the NAO Fig. 18a, which explains the large negative NAO. The GMME did well in predicting this pattern, but it struggled to forecast the magnitude of anomalies, as shown in the probability density function (Fig. 18c). The $Z_{500}$ shows a clear negative NAO pattern (Fig. 18d). The GMME forecasts the Northern Hemisphere response (Fig. 18e), however, it does not capture the magnitude of the blocking event over the Labrador Sea, as well as the magnitude of the low in the central North Atlantic. Similarly, the GMME forecasts of $U_{200}$ and MSLP are of the right sign but are too weak (Figs. 18g–l). The wind and wave patterns are also too weak in the GMME forecasts although there are a few ensembles that do correctly capture the observed pattern (Figs. 18m–u).

## 8. Forecast error growth with lead time

Figure 19 shows the RPSS averaged over points in the North Atlantic where the correlation with MSLP PC1 is larger than 0.5 (see black contour in Fig. 17) for different lead-time forecasts. For lead-time zero (December initial conditions) all parameters are skillfully predicted with RPSS values greater than 0.1. The parameter that is most predictable is SST with an RPSS of 0.4, followed by MSLP. For November initial conditions predicting DJF, a lead time of 1 month, the RPSS drops for all parameters but the forecasts remain mostly still more accurate than when using a climatological forecast. The predictability of $T_{01}$ drops notably in the November initial conditions. Finally, using a lead time of 2 months, the October initial conditions forecasts show negligible skill in the North Atlantic for almost all parameters. The SST forecast skill drops markedly to being much worse than a climatological forecast. However, there is still a small predictable signal in the MSLP.

## 9. How many models are needed for a skillful multimodel forecast?

We show the RPSSs of all possible model combinations averaged over the region defined in Fig. 19 for 10-m wind speed (Fig. 20a), $H_s$ (Fig. 20b), $T_{01}$ (Fig. 20c), and MSLP (Fig. 20d). The maximum number of model
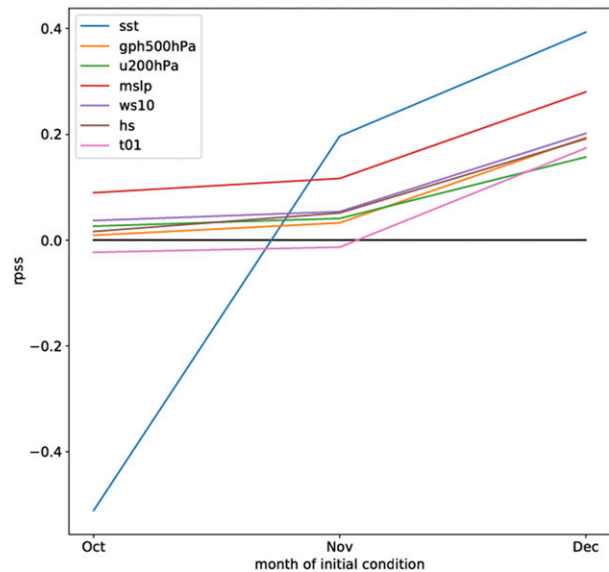


FIG. 19. RPSSs (averaged over the same points as in Fig. 17) for DJF forecasts using October–December initial conditions with the GMME mean. The variables, in order in the legend, are SST (blue), geopotential height at 500 hPa (orange), zonal velocity at 200 hPa (green), MSLP (red), wind speed at 10-m (violet), significant wave height (brown), and mean wave period (pink). The black line is the zero line, which is the RPSS for a climatological forecast.

combinations with seven models is 35 using two- and three-model combinations and the results are shown in the box-and-whisker plot. For all parameters the median skill for two models, denoted by the orange line, is double that of the median skill of the single models. The RPSS for the seven-model combination is marginally less than that of the best model (GloSea5) for 10-m wind speed (0.2 and 0.21). However, removing CanCM3, the only model with RPSS less than zero, gives an RPSS greater than any individual model, 0.215. The best forecast is given by combining CCSM4, CFSv2, and GloSea5. The relationship is similar for $H_s$ and $T_{01}$. The best forecast for $H_s$ is obtained by combining CFSv2 and GloSea5, whereas the best forecast for $T_{01}$ is obtained by combining CanCM4, CFSv2, and GloSea5. The relationship for MSLP is different. The seven-model combination is better than any of the individual models. A combination of CCSM4, CFSv2, GloSea5, and System 5 provides the best forecast.

## 10. Discussion

Seasonal forecasting of DJF winds and waves in the North Atlantic are shown to be skillful using December initial conditions with a combination of four NMME models and three EUROSIP models. The large multimodel ensemble dataset provides a comprehensive
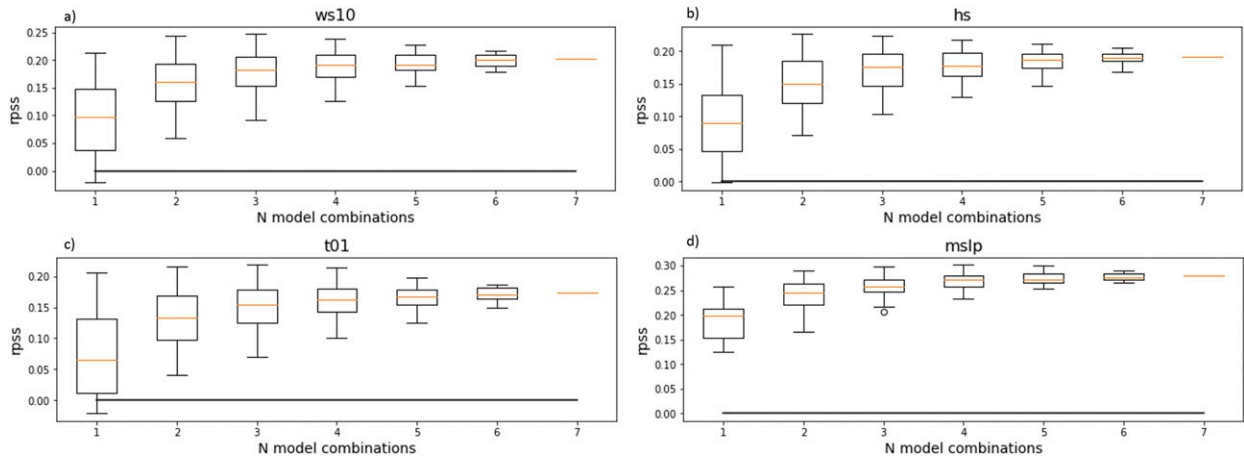
FIG. 20. RPSSs (averaged over the same points as in Fig. 17) for DJF forecasts of (a) 10-m wind speed, (b) significant wave height, (c) mean wave period, and (d) MSLP using December initial conditions with all possible multimodel combinations. The box-and-whiskers plots display the multimodel combinations with the box outlining the bottom to top quartiles and the orange line as the median. The whiskers extend to the first datum less (greater) than the first quantile minus (plus) 1.5 multiplied by the interquartile range. The black line in all panels is the zero line, which is the RPSS for a climatological forecast.

assessment of the uncertainty associated with seasonal prediction. There are two standout models in the assessment of individual models—CFSv2 and GloSea5; both of which have a high level of skill at predicting winds and waves in the North Atlantic as well as the year-to-year variability. The correlation coefficient of the ensemble mean forecast with the observed NAO is 0.74 for CFSv2 and 0.64 for GloSea5. Kim et al. (2012) found CFSv2 captures the expected winter atmospheric response to ENSO, moreso in the Pacific, but to a degree in the North Atlantic. In addition, they showed CFSv2 has good predictability skill for SST in the North Atlantic. GloSea5's ability to forecast the NAO at various time scales is given in Scaife et al. (2014), Athanasiadis et al. (2017), and Clark et al. (2017). Scaife et al. (2014) determine GloSea5's skillful forecast is due to the high-resolution ocean model (0.25°) as well as a resolved stratosphere. Further modeling studies may break down the role of a high-resolution ocean model versus a well-resolved stratosphere and their importance on the NAO.

The benefit of a multimodel ensemble forecast is shown in the spatial maps of RPSS (Figs. 1–3) and the Monte Carlo model combination analysis (Fig. 20). The results are in line with other studies that found multimodel ensemble forecasts are superior to individual model forecasts (Palmer et al. 2004; Kirtman et al. 2014 and references therein). The results in this case however showed one could be more selective in the model combinations. For example, selecting the three or four best models often gave a slightly improved forecast compared to the GMME. However, the best model combination is dependent on the variable of interest. Using all models available still provides robust skill. It is impossible to know a priori which models will yield the highest skill in combination for the operational forecasting. One approach to address this, which has been explored in O'Connor et al. (2017), is to use machine learning to weight models based on their spatial and temporal biases. In light of this, using all models is still recommended.

Correlation maps of the NAO with large-scale environmental conditions for observations and the seasonal forecasts help us to understand the model processes when predicting the NAO and its impact on surface winds and waves. However, mathematically the calculation merely yields an association, and the results must be interpreted by considering dynamical plausibility. For SST (Fig. 10), the dipole in the North Atlantic is expected although there is still debate regarding the role of the ocean versus atmospheric processes in forcing the NAO (Czaja et al. 2013). The SST dipole in the North Pacific is likely associated with the Arctic Oscillation (AO; Thompson and Wallace 1998). This pattern emerges as the AO and NAO are highly correlated and the physical mechanisms associated with annual modes for explaining the existence of the NAO are not well understood (Ambaum et al. 2001). One could undertake an idealized atmospheric general circulation model experiment with SST forcing in the North Pacific and climatology elsewhere to understand this further. For $Z_{500}$ (Fig. 11), the pattern in the North Atlantic–European sector is well understood for its impact on the geostrophic winds.

The case study of winter 2002/03, when the GMME correctly predicted near-normal NAO conditions, was a central Pacific El Niño year (McPhaden 2004). The $Z_{500}$ winter anomaly shows a wave train emanating from east of Hawaii (see Fig. 17d) into the North Atlantic. However, this did not project onto the NAO pattern and produced more of an east–west dipole in the North Atlantic. An average over the region related to the NAO therefore gives near-normal conditions. The model forecasts did not capture this wave train, and therefore the spatial structure is different. The poor NAO forecast during winter 2008/09 is well studied in Harada et al. (2010) and Ayarzagüena et al. (2011). A major stratospheric warming took place, but the external factors were unfavorable so the event was not well forecast. In this particular case it is likely that the atmosphere influenced the North Atlantic ocean beneath the storm track. It is worth mentioning that the majority of models correctly predicted the sign of the forecast. It was only CanCM3 and System 4 that predicted more near-normal conditions. The poor NAO forecast translated to a poor forecast of surface winds and waves in the North Atlantic with almost all ensembles not capturing the magnitude of the event.

Forecast skill, defined as RPSS of winds and waves in the North Atlantic, decreases to near zero when increasing the lead time to 2 months ahead of DJF. SST forecast skill drops markedly but MSLP remains predictable at a lead time of 2 months despite other local environmental conditions being difficult to forecast. There is potential here to increase the lead time for the NAO forecast and explore regions of predictability as was done in Dunstone et al. (2016).

## 11. Conclusions

Wintertime wind and wave variability is largely controlled by the NAO. Therefore, any skill in forecasting the wind and waves in the North Atlantic starts with assessing the forecasting skill of the NAO. The use of a grand multimodel ensemble (GMME) provides a large ensemble size that is advantageous for forecasting regions that are inherently noisy such as the extratropics. The GMME provides skillful forecasts of winds and waves in the North Atlantic during the winter months although it is limited to a 1-month lead time. The most predictable regions are around the British Isles, the Azores, and to the east of the Caribbean. The GMME also predicts the year-to-year winter variability of winds and waves to a high degree. An assessment of the forecasting large-scale environment conditions and their relationship to the NAO shows that the

models capture the local response in the North Atlantic but tend to have a larger correlation in the tropical Pacific than expected. The forecast response of the NAO to surface winds and waves is well captured and explains the spatial pattern of the rank probability skill score: regions that have the strongest correlation with the NAO are more predictable. The NAO does not explain the entire pattern though and it is likely other factors are at play. These include other modes of variability such as the eastern Atlantic pattern as well as remote wind forcing controlling the fetch length. However, these are outside the scope of this study and will be investigated in a future manuscript. A look at two winter case studies shows how sea surface temperatures (SSTs) affect the atmospheric variability, which in turn affects ocean winds and waves. The NAO of 2002/03 was well forecast despite it being near normal as near-normal conditions are hard to forecast. The tropical Pacific SSTs were well forecast, but the local SSTs were not. This has the impact of reducing the skill of the winds and waves in the North Atlantic. The NAO forecast during winter 2008/09 was poor despite the models capturing the sign of the event. Error was propagated in the forecast from the poor NAO forecast to a poor surface wind forecast and a poor ocean wave forecast. A multimodel ensemble forecast provides robust skill and is often better than any individual model for probabilistic skill assessments. The optimum forecast can be produced by comparing all possible model combinations, and the best skill was found by combining three or four models, depending on the parameter of interest.

### REFERENCES

Alexander, M., I. Bladé, M. Newman, J. R. Lanzante, N.-C. Lau, and J. D. Scott, 2002: The atmospheric bridge: The influence of ENSO teleconnections on air–sea interaction over the global oceans. *J. Climate*, **15**, 2205–2231, https://doi.org/10.1175/1520-0442(2002)015<2205:TABTIO>2.0.CO;2.

Amante, C., and B. Eakins, 2009: ETOPO1 1 Arc-Minute Global Relief Model: Procedures, data sources and analysis. NOAA Tech. Memo. NESDIS NGDC-24, 19 pp., https://www.ngdc.noaa.gov/mgg/global/relief/ETOPO1/docs/ETOPO1.pdf.

Ambaum, M. H. P., and L. Novak, 2014: A nonlinear oscillator describing storm track variability. Quart. J. Roy. Meteor. Soc., 140, 2680–2684, https://doi.org/10.1002/qj.2352.

——, B. J. Hoskins, and D. B. Stephenson, 2001: Arctic Oscillation or North Atlantic Oscillation? J. Climate, 14, 3495–3507, https://doi.org/10.1175/1520-0442(2001)014<3495:AOONAO>2.0.CO;2.

Ardhuin, F., and Coauthors, 2010: Semiempirical dissipation source functions for ocean waves. Part I: Definition, calibration, and validation. J. Phys. Oceanogr., 40, 1917–1941, https://doi.org/10.1175/2010JPO4324.1.

Athanasiadis, P. J., and Coauthors, 2017: A multisystem view of wintertime NAO seasonal predictions. J. Climate, 30, 1461–1475, https://doi.org/10.1175/JCLI-D-16-0153.1.

Ayarzagüena, B., U. Langematz, and E. Serrano, 2011: Tropospheric forcing of the stratosphere: A comparative study of the two different major stratospheric warmings in 2009 and 2010. J. Geophys. Res., 116, D18114, https://doi.org/10.1029/2010JD015023.

Bacon, S., and D. J. T. Carter, 1991: Wave climate changes in the North Atlantic and North Sea. Int. J. Climatol., 11, 545–558, https://doi.org/10.1002/joc.3370110507.

Bell, C. J., L. J. Gray, A. J. Charlton-Perez, M. M. Joshi, and A. A. Scaife, 2009: Stratospheric communication of El Niño teleconnections to European winter. J. Climate, 22, 4083–4096, https://doi.org/10.1175/2009JCLI2717.1.

Bell, R., and B. Kirtman, 2018: Seasonal forecasting of winds, waves and currents in the North Pacific. J. Oper. Oceanogr., 11, 11–26, https://doi.org/10.1080/1755876X.2018.1438342.

Bidlot, J.-R., P. Janssen, and S. Abdalla, 2007: Impact of the revised formulation for ocean wave dissipation on the ECMWF operational wave model. ECMWF Tech. Memo. 509, 27 pp., https://www.ecmwf.int/sites/default/files/elibrary/2007/8228-revised-formulation-ocean-wave-dissipation-and-its-model-impact.pdf.

Brayshaw, D. J., B. Hoskins, and M. Blackburn, 2009: The basic ingredients of the North Atlantic storm track. Part I: Land–sea contrast and orography. J. Atmos. Sci., 66, 2539–2558, https://doi.org/10.1175/2009JAS3078.1.

Brönnimann, S., E. Xoplaki, C. Casty, A. Pauling, and J. Luterbacher, 2007: ENSO influence on Europe during the last centuries. Climate Dyn., 28, 181–197, https://doi.org/10.1007/s00382-006-0175-z.

Butler, A. H., L. M. Polvani, and C. Deser, 2014: Separating the stratospheric and tropospheric pathways of El Niño–Southern Oscillation teleconnections. Environ. Res. Lett., 9, 024014, https://doi.org/10.1088/1748-9326/9/2/024014.

——, J. P. Sjoberg, D. J. Seidel, and K. H. Rosenlof, 2017: A sudden stratospheric warming compendium. Earth Syst. Sci. Data, 9, 63–76, https://doi.org/10.5194/essd-9-63-2017.

Castelle, B., G. Dodet, G. Masselink, and T. Scott, 2017: A new climate index controlling winter wave activity along the Atlantic coast of Europe: The West Europe pressure anomaly. Geophys. Res. Lett., 44, 1384–1392, https://doi.org/10.1002/2016GL072379.

Chawla, A., and H. L. Tolman, 2008: Obstruction grids for spectral wave models. Ocean Modell., 22, 12–25, https://doi.org/10.1016/j.ocemod.2008.01.003.

Clark, R. T., P. E. Bett, H. E. Thornton, and A. A. Scaife, 2017: Skilful seasonal predictions for the European energy industry. Environ. Res. Lett., 12, 024002, https://doi.org/10.1088/1748-9326/aa57ab.

Coll, J., D. K. Woolf, S. W. Gibb, and P. G. Challenor, 2013: Sensitivity of ferry services to the Western Isles of Scotland to changes in wave and wind climate. J. Appl. Meteor. Climatol., 52, 1069–1084, https://doi.org/10.1175/JAMC-D-12-0138.1.

Colman, A. W., E. J. Palin, M. G. Sanderson, R. T. Harrison, and I. M. Leggett, 2011: The potential for seasonal forecasting of winter wave heights in the northern North Sea. Wea. Forecasting, 26, 1067–1074, https://doi.org/10.1175/WAF-D-11-00017.1.

Czaja, A., A. W. Robertson, and T. Huck, 2013: The role of Atlantic Ocean–atmosphere coupling in affecting North Atlantic oscillation variability. The North Atlantic Oscillation: Climatic Significance and Environmental Impact, Geophys. Monogr., Vol. 134, Amer. Geophys. Union, 1–35.

Dawson, A., 2016: Eofs: A library for EOF analysis of meteorological, oceanographic, and climate data. J. Open Res. Software, 4, e14, https://doi.org/10.5334/jors.122.

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, https://doi.org/10.1002/qj.828.

Doblas-Reyes, F. J., V. Pavan, and D. B. Stephenson, 2003: The skill of multi-model seasonal forecasts of the wintertime North Atlantic Oscillation. Climate Dyn., 21, 501–514, https://doi.org/10.1007/s00382-003-0350-4.

Dodet, G., X. Bertin, and R. Taborda, 2010: Wave climate variability in the north-east Atlantic Ocean over the last six decades. Ocean Modell., 31, 120–131, https://doi.org/10.1016/j.ocemod.2009.10.010.

Dunstone, N., D. Smith, A. Scaife, L. Hermanson, R. Eade, N. Robinson, M. Andrews, and J. Knight, 2016: Skilful predictions of the winter North Atlantic Oscillation one year ahead. Nat. Geosci., 9, 809–814, https://doi.org/10.1038/ngeo2824.

Durrant, T. H., D. J. M. Greenslade, and I. Simmonds, 2013: The effect of statistical wind corrections on global wave forecasts. Ocean Modell., 70, 116–131, https://doi.org/10.1016/j.ocemod.2012.10.006.

——, ——, ——, and F. Woodcock, 2014: Correcting marine surface winds simulated in atmospheric models using spatially and temporally varying linear regression. Wea. Forecasting, 29, 305–330, https://doi.org/10.1175/WAF-D-12-00101.1.

Eade, R., D. Smith, A. Scaife, E. Wallace, N. Dunstone, L. Hermanson, and N. Robinson, 2014: Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? Geophys. Res. Lett., 41, 5620–5628, https://doi.org/10.1002/2014GL061146.

Ehsan, M. A., and Coauthors, 2017: Skill and predictability in multimodel ensemble forecasts for Northern Hemisphere regions with dominant winter precipitation. Climate Dyn., 48, 3309–3324, https://doi.org/10.1007/s00382-016-3267-4.

Graf, H.-F., and D. Zanchettin, 2012: Central Pacific El Niño, the "subtropical bridge," and Eurasian climate. J. Geophys. Res., 117, D11102, https://doi.org/10.1029/2011JD016493.

Hansen, F., R. J. Greatbatch, G. Gollan, T. Jung, and A. Weisheimer, 2017: Remote control of North Atlantic Oscillation predictability via the stratosphere. Quart. J. Roy. Meteor. Soc., 143, 706–719, https://doi.org/10.1002/qj.2958.

Harada, Y., A. Goto, H. Hasegawa, N. Fujikawa, H. Naoe, and T. Hirooka, 2010: A major stratospheric sudden warming event in January 2009. J. Atmos. Sci., 67, 2052–2069, https://doi.org/10.1175/2009JAS3320.1.

Heij, C., and S. Knapp, 2015: Effects of wind strength and wave height on ship incident risk: Regional trends and seasonality. *Transp. Res.*, **37D**, 29–39, https://doi.org/10.1016/j.trd.2015.04.016.

Hoskins, B. J., and P. J. Valdes, 1990: On the existence of storm-tracks. *J. Atmos. Sci.*, **47**, 1854–1864, https://doi.org/10.1175/1520-0469(1990)047<1854:OTEOST>2.0.CO;2.

Hoyer, S., and J. J. Hamman, 2017: Xarray: N-D labeled arrays and datasets in Python. *J. Open Res. Software*, **5** (1), 10, https://doi.org/10.5334/jors.148.

Infanti, J. M., and B. P. Kirtman, 2016: Prediction and predictability of land and atmosphere initialized CCSM4 climate forecasts over North America. *J. Geophys. Res. Atmos.*, **121**, 12 690–12 701, https://doi.org/10.1002/2016JD024932.

Jha, B., A. Kumar, and Z.-Z. Hu, 2016: An update on the estimate of predictability of seasonal mean atmospheric variability using North American Multi-Model Ensemble. *Climate Dyn.*, https://doi.org/10.1007/s00382-016-3217-1, in press.

Jiménez-Esteve, B., and D. I. V. Domeisen, 2018: The tropospheric pathway of the ENSO–North Atlantic teleconnection. *J. Climate*, **31**, 4563–4584, https://doi.org/10.1175/JCLI-D-17-0716.1.

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, https://doi.org/10.1007/s00382-012-1364-6.

Kirtman, B. P., and Coauthors, 2012: Impact of ocean model resolution on CCSM climate simulations. *Climate Dyn.*, **39**, 1303–1328, https://doi.org/10.1007/s00382-012-1500-3.

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

Kumar, A., 2009: Finite samples and uncertainty estimates for skill measures for seasonal prediction. *Mon. Wea. Rev.*, **137**, 2622–2631, https://doi.org/10.1175/2009MWR2814.1.

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, https://doi.org/10.1002/qj.2396.

Martínez-Asensio, A., M. N. Tsimplis, M. Marcos, X. Feng, D. Gomis, G. Jordà, and S. A. Josey, 2016: Response of the North Atlantic wave climate to atmospheric modes of variability. *Int. J. Climatol.*, **36**, 1210–1225, https://doi.org/10.1002/joc.4415.

McPhaden, M. J., 2004: Evolution of the 2002/03 El Niño. *Bull. Amer. Meteor. Soc.*, **85**, 677–696, https://doi.org/10.1175/BAMS-85-5-677.

Merryfield, W. J., and Coauthors, 2013: The Canadian Seasonal to Interannual Prediction System. Part I: Models and initialization. *Mon. Wea. Rev.*, **141**, 2910–2945, https://doi.org/10.1175/MWR-D-12-00216.1.

Météo-France, 2016: Météo-France seasonal forecast ARPEGE system5 for EUROSIP: Technical description. Météo-France Tech. Rep., 12 pp. + appendixes, https://www.umr-cnrm.fr/IMG/pdf/system5-technical.pdf.

Molteni, F., and Coauthors, 2011: The new ECMWF seasonal forecast system (System 4). ECMWF Tech. Rep. 656, 49 pp., https://www.ecmwf.int/sites/default/files/elibrary/2011/11209-new-ecmwf-seasonal-forecast-system-system-4.pdf.

O'Connor, A., R. Bell, B. P. Kirtman, and J. Gorman, 2017: Long-range forecasting using COMPASS machine learning. *Seventh Int. Workshop on Climate Informatics*, Boulder, CO, NCAR, 57–60, https://dx.doi.org/10.5065/D6222SH7.

O'Reilly, C. H., J. Heatley, D. MacLeod, A. Weisheimer, T. N. Palmer, N. Schaller, and T. Woollings, 2017a: Variability in seasonal forecast skill of Northern Hemisphere winters over the twentieth century. *Geophys. Res. Lett.*, **44**, 5729–5738, https://doi.org/10.1002/2017GL073736.

——, S. Minobe, A. Kuwano-Yoshida, and T. Woollings, 2017b: The Gulf Stream influence on wintertime North Atlantic jet variability. *Quart. J. Roy. Meteor. Soc.*, **143**, 173–183, https://doi.org/10.1002/qj.2907.

Palmer, T. N., and Coauthors, 2004: Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER). *Bull. Amer. Meteor. Soc.*, **85**, 853–872, https://doi.org/10.1175/BAMS-85-6-853.

Parfitt, R., A. Czaja, and Y.-O. Kwon, 2017: The impact of SST resolution change in the ERA-Interim reanalysis on wintertime Gulf Stream frontal air–sea interaction. *Geophys. Res. Lett.*, **44**, 3246–3254, https://doi.org/10.1002/2017GL073028.

Robertson, A. W., C. R. Mechoso, and Y.-J. Kim, 2000: The influence of Atlantic sea surface temperature anomalies on the North Atlantic Oscillation. *J. Climate*, **13**, 122–138, https://doi.org/10.1175/1520-0442(2000)013<0122:TIOASS>2.0.CO;2.

Rodwell, M. J., and C. K. Folland, 2002: Atlantic air–sea interaction and seasonal predictability. *Quart. J. Roy. Meteor. Soc.*, **128**, 1413–1443, https://doi.org/10.1002/qj.200212858302.

Saha, S., and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, https://doi.org/10.1175/JCLI-D-12-00823.1.

Sardeshmukh, P. D., G. P. Compo, and C. Penland, 2000: Changes of probability associated with El Niño. *J. Climate*, **13**, 4268–4286, https://doi.org/10.1175/1520-0442(2000)013<4268:COPAWE>2.0.CO;2.

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

——, and Coauthors, 2016: Seasonal winter forecasts and the stratosphere. *Atmos. Sci. Lett.*, **17**, 51–56, https://doi.org/10.1002/asl.598.

Shukla, J., 1998: Predictability in the midst of chaos: A scientific basis for climate forecasting. *Science*, **282**, 728–731, https://doi.org/10.1126/science.282.5389.728.

Siegert, S., D. B. Stephenson, P. G. Sansom, A. A. Scaife, R. Eade, and A. Arribas, 2016: A Bayesian framework for verification and recalibration of ensemble forecasts: How uncertain is NAO predictability? *J. Climate*, **29**, 995–1012, https://doi.org/10.1175/JCLI-D-15-0196.1.

Simmonds, I., and K. Keay, 2002: Surface fluxes of momentum and mechanical energy over the North Pacific and North Atlantic Oceans. *Meteor. Atmos. Phys.*, **80**, 1–18, https://doi.org/10.1007/s007030200009.

——, and J. King, 2004: Global and hemispheric climate variations affecting the Southern Ocean. *Antarct. Sci.*, **16**, 401–413, https://doi.org/10.1017/S0954102004002226.

Stephenson, D. B., H. Wanner, S. Brönnimann, and J. Luterbacher, 2003: The history of scientific research on the North Atlantic Oscillation. *The North Atlantic Oscillation: Climatic Significance and Environmental Impact, Geophys. Monogr.*, Amer. Geophys. Union, Vol. 134, 37–50.

Stopa, J. E., and K. F. Cheung, 2014: Intercomparison of wind and wave data from the ECMWF Reanalysis Interim and the NCEP Climate Forecast System Reanalysis. *Ocean Modell.*, **75**, 65–83, https://doi.org/10.1016/j.ocemod.2013.12.006.

Thompson, D. W. J., and J. M. Wallace, 1998: The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.*, **25**, 1297–1300, https://doi.org/10.1029/98GL00950.

Tolman, H. L., and Coauthors, 2014: User manual and system documentation of WAVEWATCH III version 4.18. Marine Modeling and Analysis Branch Contribution 316, 282 pp. + appendixes, http://polar.ncep.noaa.gov/waves/wavewatch/manual.v4.18.pdf.

Torralba, V., F. J. Doblas-Reyes, D. MacLeod, I. Christel, and M. Davis, 2017: Seasonal climate prediction: A new source of information for the management of wind energy resources. *J. Appl. Meteor. Climatol.*, **56**, 1231–1247, https://doi.org/10.1175/JAMC-D-16-0204.1.

Tsimplis, M. N., and Coauthors, 2005: Towards a vulnerability assessment of the UK and northern European coasts: The role of regional climate variability. *Philos. Trans. Roy. Soc. London*, **363A**, 1329–1358, https://doi.org/10.1098/rsta.2005.1571.

van den Dool, H. M., and Z. Toth, 1991: Why do forecasts for "near normal" often fail? *Wea. Forecasting*, **6**, 76–85, https://doi.org/10.1175/1520-0434(1991)006<0076:WDFFNO>2.0.CO;2.

Wang, L., M. Ting, and P. J. Kushner, 2017: A robust empirical seasonal prediction of winter NAO and surface climate. *Sci. Rep.*, **7**, 279, https://doi.org/10.1038/s41598-017-00353-y.

Weigel, A. P., M. A. Liniger, and C. Appenzeller, 2007: The discrete Brier and ranked probability skill scores. *Mon. Wea. Rev.*, **135**, 118–124, https://doi.org/10.1175/MWR3280.1.

Wolf, J., and D. K. Woolf, 2006: Waves and climate change in the north-east Atlantic. *Geophys. Res. Lett.*, **33**, L06604, https://doi.org/10.1029/2005GL025113.

Yang, X., and Coauthors, 2015: Seasonal predictability of extratropical storm tracks in GFDL's high-resolution climate prediction model. *J. Climate*, **28**, 3592–3611, https://doi.org/10.1175/JCLI-D-14-00517.1.

Zhang, W., Z. Wang, M. F. Stuecker, A. G. Turner, F.-F. Jin, and X. Geng, 2018: Impact of ENSO longitudinal position on teleconnections to the NAO. *Climate Dyn.*, https://doi.org/10.1007/s00382-018-4135-1, in press.

Zubiate, L., F. McDermott, C. Sweeney, and M. O'Malley, 2017: Spatial variability in winter NAO–wind speed relationships in western Europe linked to concomitant states of the East Atlantic and Scandinavian patterns. *Quart. J. Roy. Meteor. Soc.*, **143**, 552–562, https://doi.org/10.1002/qj.2943.