The Representation of Atmospheric Blocking and the Associated Low-Frequency Variability in Two Seasonal Prediction Systems

PANOS J. ATHANASIADIS AND ALESSIO BELLUCCI
CMCC, Bologna, Italy

LEON HERMANS ON, ADAM A. SCAIFE, CRAIG MACLACHLAN, AND ALBERTO ARRIBAS
Met Office Hadley Centre, Exeter, United Kingdom

STEFANO MATERIA AND ANDREA BORRELLI
CMCC, Bologna, Italy

SILVIO GUALDI
CMCC, and INGV-CMCC, Bologna, Italy

(Manuscript received 15 April 2014, in final form 11 September 2014)

ABSTRACT
Primarily as a response to boundary forcings, certain components of the atmospheric intraseasonal variability are potentially predictable. Particularly referring to the extratropics, the current generation of seasonal forecasting systems is making advancements in predicting these components by realistically initializing many components of the climate system, using higher resolution and utilizing large ensemble sizes.

The operational seasonal prediction system of the Met Office (UKMO) and the corresponding system of the Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC) are analyzed in terms of their representation of different aspects of extratropical low-frequency variability. The UKMO system achieves unprecedented high scores in predicting the winter mean phase of the North Atlantic Oscillation (NAO; correlation 0.62) and the Pacific-North American pattern (PNA; correlation 0.82). The CMCC system, despite its smaller ensemble size and coarser resolution, also exhibits significant skill (0.42 for NAO, 0.51 for PNA). Low-frequency variability is underrepresented in both models, particularly in the eastern North Atlantic. Consequently, their intrinsic variability patterns (sectoral EOFs) are somewhat different from the observed patterns.

Regarding the representation of wintertime Northern Hemisphere blocking, after bias correction both systems exhibit a realistic climatology of blocking frequency. In this assessment, instantaneous blocking and large-scale persistent blocking events are identified using daily geopotential height fields at 500 hPa. The blocking signature on the circulation and the dependence of blocking frequency on the NAO are also quite realistic for both systems. Finally, the Met Office system exhibits significant skill in predicting the winter mean frequency of blocking that relates to the NAO.

1. Introduction
Extratropical teleconnections, the occurrence of blocking, and the associated modulation of the midlatitude westerly flow are important aspects of low-frequency variability. Part of the latter is attributed to the chaotic nature of the atmosphere and is inherently unpredictable. On the other hand, primarily as a response to boundary forcings, tropospheric low-frequency variability may include components that are potentially predictable. However, achieving the capacity to actually predict these components has been, to a certain degree,
elusive. Even from a theoretical viewpoint, particularly regarding the North Atlantic Oscillation (NAO), many studies have seen this variability pattern as unpredictable noise. Stephenson et al. (2000) show that the interannual variations in the wintertime NAO index have a broad spectrum that is close to being white noise, yet, as these authors note: “this pessimistic prospect for forecasting the NAO is not the whole story.” In fact, in a statistical assessment Johansson et al. (1998) found that the European climate exhibits some seasonal predictability, primarily associated with the NAO pattern. Folland et al. (2012) also demonstrated that hindcasts derived with statistical methods can have significant skill in predicting winter European climate. This predictability potential was supported also by the well-predicted 2010/11 strong negative NAO, which was a case study for Maidens et al. (2013).

The present study focuses on two aspects: first, to assess the representation of atmospheric blocking and low-frequency variability in two operational seasonal prediction systems [those of the Met Office (UKMO) and Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC); see next section for details] and second, to examine the associated predictability skill of given variability patterns and of wintertime blocking frequency in these systems. If these are essential it is because they relate directly to the predictability of the regional short-term climate, which is very important for many applications. As said, seasonal forecasting systems have had difficulty in making skillful predictions of the North Atlantic Oscillation (NAO) and other aspects of the European climate. However, it is encouraging that some newly developed operational systems are achieving higher levels of skill by realistically initializing most components of the climate system, using higher resolution and utilizing large ensemble sizes (Riddle et al. 2013; Scaife et al. 2014).

Climate models have been assessed in several studies in terms of their representation of the atmospheric blocking climatology (e.g., Masato et al. 2013; Anstey et al. 2013; Dunn-Sigouin and Son 2013). A sufficiently high resolution (in both the ocean and the atmosphere) has been found to be one of the key model characteristics for improving this representation (Scaife et al. 2011; Jung et al. 2012). In the ocean, the high resolution (0.25° or less) is necessary for better representing the mesoscale activity and, consequently, reducing the sea surface temperature (SST) model bias, for example at the Gulf Stream Extension, which is a crucial for cyclogenesis (Sanders 1986; Hoskins and Valdes 1990). In the atmosphere, high resolution (less than or about 0.5°) is important for better resolving the meridional eddy fluxes of heat and vorticity/momentum that are the essence of eddy–mean flow interaction and help maintain the low-frequency anomalies (e.g., Shotts 1983; Lau and Holopainen 1984; Lorenz and Hartmann 2003).

Mean bias correction has been found to improve the climatology of blocking (Scaife et al. 2010) by bringing the time-mean geopotential height meridional gradient to its correct location and magnitude. These corrections have a direct effect on the detected blocking. However, one should bear in mind that some of the blockings detected after bias correction may not actually occur in the raw model circulation and, therefore, the same is true for their climatic impacts. It should be mentioned, also, that Shukla and Mo (1983) and others have defined blockings as strong persistent anomalies, and in this view mean bias correction would not have an effect (since anomalies and variances are not affected).

Midlatitude blockings are also closely related to climate variability patterns. In the North Atlantic basin, the NAO and the variability of the jets and the storm track are all directly related to the occurrence of blockings (Masato et al. 2012; Davini et al. 2012; Athanasiadis et al. 2010; Woollings et al. 2008). In the view of all these, in order to better simulate the impacts of atmospheric blocking and natural climate variability it is fundamental to improve our climate models in their representation of not only the time-mean state but also the variance at different frequencies.

In this study we show that for both of the analyzed models, monthly variability of the eddy-driven jet and variability of the 500-hPa geopotential height are underestimated in the Atlantic sector, particularly over Greenland and at the eastern side of the basin. As a direct impact of these biases or errors, the principal EOFs in this sector do not exactly match the observed pattern, which corresponds to the NAO (see Fig. 10). Here it is argued that the predictive skill of the NAO and the associated blocking frequency (both affecting European climate) are limited by these deficiencies, albeit less so for the UKMO model.

2. Data and methods

In this study two seasonal forecasting systems are analyzed. One might argue that analyzing only two systems is clearly not the whole story and there is certainly space for broader multimodel comparisons. However, working with only these two systems allowed a more in-depth comparison of a representative low-resolution model and a high-resolution state-of-the-art model against observations. The choice of the Met Office system was based on its recently reported good predictive skill of the NAO (Scaife et al. 2014).

The winter hindcasts of the Met Office Global Seasonal Prediction System 5 (GloSea5, hereafter referred
to as UKMO) cover the period 1996–2011 with an ensemble of 24 members. Initialization dates centered on 1 November (25 October, 1 November, and 9 November) have been used, each contributing eight members to the ensemble. The winter hindcasts of the CMCC Seasonal Prediction System version 1.5 (hereafter referred to as CMCC) consist of nine members covering the period 1989–2006. These are all initialized on 1 November. To account for the uncertainty in the exact initial state, past atmospheric states (with a 12-h interval, thus going back to 28 October) are assigned as the initial condition on 1 November. All model data until the beginning of December (initial period) are discarded. Details of these two models can be found, in C. MacLachlan et al. (2014, unpublished manuscript), Materia et al. (2014), and Borrelli et al. (2012). Their main features are summarized in Table 1.

### a. Model and reanalysis data

Daily and monthly mean data have been used in this study covering the period December to March (DJFM). In particular, daily geopotential height fields at 500 hPa (Z500) and u wind at 250 hPa (U250) from the Interim European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-Interim; Dee et al. 2011) and from the above-mentioned hindcasts have been interpolated to a common regular 2.5° × 2.5° grid prior to further processing. In this analysis only the Northern Hemisphere is examined. In addition, monthly mean sea level pressure (MSLP) fields have been used for assessing the predictability of the mean circulation and its relation to blocking frequency. It should be noted that for the analysis of blocking only 5 (out of 8) members for each initialization date of the UKMO model were available at daily resolution.

Various approaches have been taken in previous studies for the detection of blocking in observational and model data, including, among others, the reversal of the potential temperature meridional gradient on the dynamical tropopause [potential vorticity (PV) = 2 surface] by Pelly and Hoskins (2003) and Tyrlis and Hoskins (2008), and the reversal of the meridional geopotential height gradient at 500 hPa as in Tibaldi and Molteni (1990), Scherrer et al. (2006), and Masato et al. (2013). In this study we use daily Z500 fields following three slightly different methods for the detection of midlatitude blockings. These are outlined below.

For the analysis of blocking, the mean bias has been subtracted from the daily Z500 fields. Given that for both models this bias shows only a weak dependence on lead time (the mean biases for the months of December, January, February, and March are not fundamentally different from each other) it is acceptable to remove monthly biases from the daily fields. Alternative, more elaborate bias correction methods can be employed, but this is not considered necessary for the type of work that is undertaken here.

### b. Blocking detection methods

Two different methodologies are used for the detection of blockings. In both mean bias correction is applied as suggested in Scaife et al. (2010). The first methodology follows the one-dimensional approach introduced by Tibaldi and Molteni (1990) with the difference that the central blocking latitude (CBL) is allowed to vary with longitude as in Pelly and Hoskins (2003). The CBL is supposed to follow the zone of maximum baroclinic activity, where the actual blocking of the westerly flow occurs. Thus, at each longitude the CBL is defined here as the latitude where the standard deviation of the synoptic frequency Z500 anomalies takes its maximum. These anomalies are computed for DJFM by applying a standard Eulerian 2–6.5-day bandpass filter to the ERA-Interim daily Z500 fields. One may argue that the CBL should be defined separately for each model and the observations; however, 1) the corresponding differences were found to be very small (not shown), and 2) when comparing blocking statistics it is preferred for these to refer to the same location. A historical period representative of both model hindcasts is used (1989–2011), and the resulting CBL profile differs very slightly from other similar definitions used in the literature (e.g., Barriopedro et al. 2010).

The second methodology for blocking detection is two-dimensional and is outlined in Scherrer et al. (2006).
According to this, blockings are detected separately at each grid point within the latitude zone 35°–75° N. Here, mainly for illustration purposes, this zone is expanded slightly to 30° N. However, it is known that anomalies that may be detected as “blockings” at low latitudes have different characteristics compared to the traditional mid- and high-latitude blockings (the former tend to coincide with the permanent subtropical anticyclones) and thus they are not regarded as such.

Both of the above-mentioned methodologies result in a binary daily index, which is referred to as instantaneous blocking (IB). As expected, along the grid points that define the CBL, the 1D and 2D blocking detection methods give comparable but not identical results (section 3). Continuing with the 2D methodology, to retain only the instantaneous blockings that have a large-scale character, a spatial threshold is applied. According to this, a detected instantaneous blocking at a given point and time is referred to as large-scale blocking (LSB) if instantaneous blocking is also detected in all grid points (seven in total) that fall within a centered zonal window of 15° longitude. A further selection according to temporal persistence leads to the subset of blocking events (BEs) that are required to last for at least 5 consecutive days. In particular, the criterion requires that on each of these 5 days there should be at least one grid point within a centered 5° lat × 10° lon window for which an LSB is detected. For the used regular grid, such a window includes 3 × 5 = 15 grid points. If not clearly stated otherwise, any reference to blocking implies blocking detection performed after mean bias correction has been applied.

c. EOF computation

One way to examine the representation of low-frequency variability in the models is to compare the empirical orthogonal functions (EOFs) computed from the model data to those computed from the ERA-Interim reanalysis. Particularly in the Atlantic sector, the variability of blocking is directly related to the dominant EOF, which has been used as an alternative definition for the NAO (Hurrell 1995). As representative of different levels in the troposphere, U250, Z500, and MSLP fields are used for this purpose. The respective monthly climatology of each model and the ERA-Interim reanalysis (for the corresponding data period) is first subtracted from the monthly mean fields. The hindcast periods are relatively short, and thus removing a long-term trend from the data was not considered necessary. The EOF analysis was performed in two different sectors, representing the North Atlantic (105° W–30° E) and the North Pacific (120° E–105° W). This choice is discussed in more detail in Athanasiadis et al. (2010). In calculating the eigenvectors and the associated principal component (PC) time series we have area-weighted the data to account for the uneven resolution of the spherical coordinate grid (Baldwin et al. 2009). Hemispheric patterns are produced by regressing the hemispheric fields onto the sectoral PCs. Within the respective sectors, these patterns define the corresponding EOFs and have the same units as the data. Hereafter we refer to these hemispheric patterns as the sectoral EOFs, or variability patterns.

3. Blocking climatologies

First we examined the capability of the two seasonal forecasting models to reproduce the observed Northern Hemisphere climatology of blocking frequency in one dimension—that is, as a function of longitude along the CBL (Fig. 1). The resulting profiles of wintertime blocking frequency are shown in Fig. 2. It should be noted that the blocking frequency varies strongly from one decade to another (see Håkkinen et al. 2011) and therefore, with respect to blocking and low-frequency variability, each model (UKMO and CMCC) should be compared with observations (ERA-Interim) in the period covered by the respective model hindcasts. At a first glance at Fig. 2, one would say that the agreement between the model profiles and the observational analysis from ERA-Interim is quite good, especially in the domain adjacent to the eastern North Atlantic (30° W–30° E) where most of the blocking activity is found. In contrast to the abovementioned domain, significant
departures of the opposite sign are found off the West Coast of North America around the 210°E meridian. Generally, comparing these model profiles with the respective ones produced without the use of mean bias correction (not shown), it was found that for both models the use of mean bias correction increases the blocking frequency in the neighborhood of the prime meridian (30°W–30°E) and decreases it around the 210°E meridian. In effect, this brings the blocking frequency of both models closer to the observed in the first domain, in agreement with Scaife et al. (2010), but overreduces the blocking frequency of the UKMO model in the second domain, leading to departures of the opposite sign.

To get a more detailed picture of the spatial distribution of blocking frequency, results from the two-dimensional analysis of blocking are shown in Fig. 3. This displays the climatology of instantaneous blocking frequency for each of the two models and the respective period of ERA-Interim reanalysis. First, it is seen that between the two ERA-Interim periods (1989–2006 and 1996–2011) there are small but noticeable differences in blocking frequency. Over Greenland, in particular, from

---

2 Referring to this domain, it should be noted that the CBL (−45°N) is located to the south of the Alaskan blocking area (Renwick and Wallace 1996).
the earlier to the later period there seems to be a relative increase of about 30%, which is consistent with the negative trend in the winter-mean NAO index from the 1990 to the present (http://www.cpc.ncep.noaa.gov). Again, both models seem to be performing well in capturing the main features of the observed spatial distribution of blocking frequency.

If one focuses on the large-scale persistent blockings, referred to as blocking events, the departures are more pronounced. Figure 4 shows the corresponding climatologies, where it can be seen that Greenland blocking is particularly underrepresented in the CMCC model. Given that instantaneous blocking is underrepresented in that area, persistent blocking is expected to be more strongly underrepresented because blocking frequency generally drops quasi-exponentially with duration (Barriopedro et al. 2010). In addition, it can be argued that high resolution is more important for the representation of persistent blocking, during which the eddy fluxes are fundamental for the maintenance of the blocking anomalies. Furthermore, Greenland is a huge orographic obstacle, and although it is not as sharp as
other mountain ranges, still it is less well represented at low resolutions, with all the effects that this may have on the circulation. Finally, as Scaife et al. (2010) comment, high resolution in the ocean is also crucial for the representation of blocking because it reduces mean-state biases through SST biases.

As a general comment, we would like to emphasize that although mean bias correction appears to be a successful correction technique, it may not be really “fair.” That is because the observed blockings are all-important in determining the mean state, and thus, in reverse, correcting the model climatological field leads to an artificially good blocking climatology. This is even more so if one considers the eddy–mean flow interaction: fixing the mean state with bias correction certainly does not alter the variability (transient eddies) in the model, and therefore it leads to a dynamically inconsistent realm. The important thing to consider is that—given that what matters in a model is the climatic impacts of the blockings and not the blockings themselves—the mean bias correction cannot make a model better than it actually is. Nevertheless, the fact that the mean-bias correction leads to a realistic distribution of blocking frequency suggests that the daily variability in the models is reasonable.

These arguments originate from theoretical considerations. However, their relevance to reality can be seen

---

**Fig. 4.** As in Fig. 3, but only for the persistent, large-scale blockings referred to as blocking events (see text for details).
in Fig. 5, which shows composites of the full and the anomaly Z500 fields (without bias correction) on the occurrence of blocking within two indicative areas. Knowing that over southern Greenland for DJFM the UKMO model has a negative Z500 bias and the CMCC model a positive one (not shown), the effect of the bias removal prior to the detection of the blockings is evident in that area by the weaker meridional gradient in the UKMO composites and the stronger one in the CMCC composites (green contours). This indicates, as one would expect, that when a blocking is detected after bias correction an actual reversal of the gradient does not necessarily occur in the full model fields.

On the other hand, mean bias correction is justified by examining the respective composites of the Z500 anomalies (also in Fig. 5). The latter reveal the kind of anomaly fields that correspond to blocking. In effect, with the adopted two-dimensional blocking detection methodology, blockings correspond to a dipole pattern of Z500 anomalies with the positive anomaly on the poleward side (thus impeding the westerly flow at the middle of the dipole). This pattern is similar in the two models and the observations, indicating that the detected blocking modifies the raw model circulation in the same way that it modifies the observed circulation. The abovementioned dipole pattern can also be seen in Davini et al. (2012) (see Greenland blocking in their Fig. 5). In the next section we explore the representation of variability in the two models.

4. Variances and patterns of variability

Continuing the assessment of the two models regarding the representation of low-frequency variability, we present first the standard deviation of the monthly U250 fields. We have chosen this field because it expresses the variability of the jet streams. Vial and Osborn (2012) argue that the time-mean atmospheric circulation biases affect the frequency of blocking episodes, and the maximum event duration in the Euro-Atlantic region. Moreover, the variability of the jet streams has received renewed attention in relation to the dominant extratropical variability patterns. Athanasiadis et al. (2010) present a number of reasons for which the jet variability can be considered more fundamental than the traditional Z500 diagnostic. In any case, from a dynamical point of view it is obvious that a model should be able to simulate successfully the preferred jet stream positions and the associated transitions. Figures 6 and 7 serve to compare each of the two models with the respective observations in this aspect, although the latter shows only the Atlantic sector dominant EOF.

An important feature in the Atlantic sector, clearly seen in the bottom panels of Fig. 6, is that the jet variability exhibits at least two distinct maxima over the central and eastern North Atlantic. This is the case for the observed monthly anomalies. Similarly, Athanasiadis et al. (2010) show that daily teleconnectivity exhibits a tri-pole pattern in the North Atlantic domain (see also Woollings et al. 2010). However, the two models examined here are lacking this feature, indicating an underrepresentation of jet stream variability at the area where blocking is more frequent. Instead, they exhibit increased monthly variability off the coast of North America (Gulf of Mexico and Quebec). These biases are fully consistent with the respective dominant EOFs shown in Fig. 7. In these, comparing with the observed patterns, the two models exhibit weaker loadings at the central and eastern North Atlantic and stronger loadings at its western side.

As blocking is mostly related to the so-called intermediate frequency variability (periods of 7–15 days), we compare also the distributions of the standard deviation of the respective Z500 band-passed transients. These are shown in Fig. 8. Again, it can be seen that the CMCC model (more strongly than the UKMO model) underrepresents variability in the North Atlantic, while it has excessive variability in the North Pacific. This was found to hold also for the monthly Z500 variability (not shown). Figure 9 shows the dominant Z500 monthly EOF for the Atlantic sector. In it this can be seen that both models exhibit patterns that are slightly different from the observed ones (i.e., they are not dipoles). Such slight differences were found also in the representation of the PNA pattern3 (not shown).

Similarly, Fig. 10 displays the dominant North Atlantic EOFs of monthly MSLP, which is an alternative way to define a pattern representing the NAO. Although both models do a good job in reproducing the main character of the observed pattern, small but significant differences can be seen between the CMCC model and the corresponding observed pattern (centers are slightly misplaced and the loadings differ). In contrast, the UKMO model does not suffer from these discrepancies. In accordance to the abovementioned excess of monthly variance found in the North Pacific, the Atlantic PC 1 for both models has a stronger projection in the North

3 The dominant Z500 monthly EOFs for the Pacific–North American sector (150°–290°E) were compared. The main difference is that for both models the corresponding pattern was found to exhibit stronger variability in the northern part of the North Pacific compared to the observed pattern, which is consistent with the biases in Z500 monthly variance. The percentage of the total variance explained by EOF 1 was also slightly higher in the models.
Pacific compared to the observations, indicating a stronger covariance with this sector.

The misrepresentation of teleconnection patterns, to a lesser or greater degree, is one of the remaining major deficiencies of current climate models. In the North Atlantic, in particular, where a number of different dynamical processes play simultaneously their own role in determining the low-frequency variability (Swanson...
(2002; Folland et al. 2012), it is not surprising that the models have some problems simulating this variability, particularly aloft.

5. Predictive skill at seasonal time scales

We have already discussed some of the shortcomings of the examined models in realistically representing the low-frequency variability of the extratropical wintertime circulation. So far, without considering the temporal evolution of the features, we have examined only the time-mean distributions and the respective biases, looking, for example, at the climatology of blocking frequency and the distribution of variances and covariability. Yet, ultimately, the quality of a forecasting system is most dominantly represented by its predictive skill. In this regard, this assessment continues with the examination of some measures of predictive skill related to the low-frequency variability in the extratropics.

Recalling that the strongest model deficiencies were found in the eastern North Atlantic basin, it is imperative to examine the models’ skill in predicting the phase of the NAO. For comparison and for completeness we also examine the predictive skill for the PNA, which is
the dominant variability pattern in the North Pacific sector. At this point it should be noted that for a global circulation model to exhibit some predictive skill beyond the weather forecast time scale (1–2 weeks) it is fundamental to have an ensemble of forecasts so that the signal-to-noise ratio is sufficiently enhanced. Scaife et al. (2014) discuss the rationale underlying this claim and show how the predictive skill depends on the ensemble size. As a measure of predictive skill we use the correlation coefficient between the observed and the forecast winter [December–February (DJF)] mean indices of two variability patterns: the PNA and the NAO indices as defined as in Wallace and Gutzler (1981) using, correspondingly, the monthly standardized anomalies of Z500 and MSLP at the respective centers of action.

Figure 11a shows the abovementioned time series for the UKMO model and ERA-Interim, quoting also the corresponding correlation coefficients. For the NAO this was originally shown in Scaife et al. (2014). For comparison, Fig. 11b shows the same but for the CMCC model and the respective period of ERA-Interim. All
the quoted correlations passed a $t$ test of statistical significance at the 0.95 level against the null hypothesis of nonpositive correlation. To account for autocorrelation, the effective ensemble size was calculated separately for each model according to Bretherton et al. (1999). In fact, for the UKMO case, the correlations passed a similar test even at the 0.99 level. Compared to CMCC, the UKMO ensemble mean exhibits a higher predictive skill for both the NAO and the PNA. It should be noticed that for UKMO the corresponding predictive skill for the Z500 Atlantic EOF 1 (which can be considered as an alternative way to define NAO; see Fig. 7c) was found to be as high as 0.74. The abovementioned values are Pearson’s linear correlations; however, Spearman’s rank correlations were not found to be much different. The CMCC model also exhibits significant skill for both patterns. Considering its lower resolution, smaller ensemble size (Table 1), and its greater biases seen in the previous section, it is not so surprising that the skill of the CMCC model is surpassed by the UKMO model. In both systems the skill for the PNA (0.82 for UKMO and 0.51 for CMCC) exceeds the skill for the NAO (0.62 for UKMO and 0.42 for CMCC) in agreement with Johansson (2007), who reports higher skill for the PNA.
pattern versus the NAO. Perhaps this difference in predictability points to different dynamical processes dominating the low-frequency variability in the corresponding sectors.

Inspecting the time series in Fig. 11 one may say that the strong El Niño of 1998 seems to enhance the predictive skill for PNA in both models. A similar influence may exist also for the NAO, yet quantifying its contribution to the predictive skill in a statistically robust way does not seem feasible due to the shortness of the time period covered by the model hindcasts containing very few strong El Niño and La Niña events. In any case, as indicated by Hoerling et al. (1997) and other studies, such an influence is not expected to be linear.

The recurrent persistent blocking events are associated with long-lasting circulation anomalies that fall in the intermediate/low-frequency range, well beyond the synoptic time scale of baroclinic eddies typically considered as 2–6 days. Therefore, one would expect that areas with high blocking frequency will also tend to exhibit high variance in these intermediate frequencies (referred to as LF). Of course, transient anomalies in this frequency band are not necessarily blockings; they can also be of the opposite polarity. Shukla and Mo (1983) define blockings as positive Z$_{500}$ anomalies with periods beyond 7 days, and so do Dole and Gordon (1983). However, in both of these studies it was found that the strongest and most persistent anomalies are
more frequently positive than negative. Also, Vial and Osborn (2012) find that the amplitude of LF variability is positively related to the blocking frequency and persistence in the Euro-Atlantic sector.

After these considerations it should be clear that biases in the LF variability are most likely associated with biases in the blocking frequency (at least in the Euro-Atlantic sector) and the same should be true in regard to the predictability of LF variability. In particular, given the strong observed relationship between high-latitude North Atlantic blocking and the NAO [e.g., in Scherrer et al. (2006), but also see in Pavan et al. (2000) their composites of blocking frequency on sectoral EOFs] one would expect that for a given model, a significant predictive skill in the NAO would “translate” also to a good predictive skill in the Euro-Atlantic blocking frequency and vice versa. Before we examine the latter predictive skill it is informative to examine the dependence of blocking frequency on the NAO. Figure 12 shows composite differences of IB (instantaneous blocking) and BE (blocking event) frequency onto the NAO index for ERA-Interim (1989–2011) and the two examined models. It is clear that in both models these signatures are realistic, although the respective magnitudes for IB are a bit weaker in the models compared to the observations. Consequently, if a model is skillful in predicting the winter-mean NAO it should also have some skill in predicting the winter-mean blocking frequency in the respective areas.

FIG. 10. As in Fig. 7, but for MSLP. Contour interval is 0.5 hPa.
Figure 13 shows the correlation coefficient between the observed and the modeled wintertime cumulative frequency of blockings. These blockings (IB/BE) have been summed up for all grid points within a two-dimensional window (10° lat × 20° lon) centered around the grid point of reference. In the Euro-Atlantic sector, and only for the UKMO model, it can be seen that the seasonal predictability of BE frequency in some key areas (where blocking activity is strong and related to the NAO variability) is comparable to the predictive skill for the NAO pattern itself (0.60). It should be noted that both of the respective time series (not shown) exhibit a slight positive trend [documented also in Davini et al. (2012)], which explains part of this correlation.

Similarly, the high correlations found in the mid-latitude North Pacific are consistent with the high predictive skill found for the PNA. Although not obvious, the high correlations at about 40°N are influenced by the frequent blockings occurring farther south (white contours) as the abovementioned summation window extends 5° north and south from the southern border of each grid box. In fact, after examining the respective blocking frequency composites on the PNA index (not shown), it is in exactly this area of the North Pacific where blocking exhibits its strongest dependence onto the PNA index. In other words, as for the NAO, the high predictive skill found for the blocking frequency in the North Pacific appears to be directly related to the high predictive skill found for the PNA pattern itself.

---

4 The blocking detection was performed separately for each ensemble member. Then, the corresponding monthly sums (blocked days) were averaged across the ensemble. Finally, the seasonal (DJFM) sum was taken.

5 As mentioned in section 2b the anomalies detected at low latitudes are not actually considered as blockings.
FIG. 12. Composite differences of the monthly cumulative frequency of (left) instantaneous blocking and (right) blocking events onto the monthly NAO index, for (top) UKMO, (middle) ERA-Interim (1989–2011), and (bottom) CMCC. Positive values are in red and negative values in blue. The zero contour is omitted. Contour interval is 2 days per month for IB (left) and 1 day per month for BE (right). For the composites onto the standardized NAO index, the $\pm 0.5$ thresholds are used.
6. Summary and discussion

The ensemble hindcasts from two seasonal prediction systems (UKMO and CMCC) have been examined in terms of their representation of the wintertime Northern Hemisphere atmospheric blocking and the associated extratropical low-frequency variability. With regard to the respective dominant variability patterns, the NAO and the PNA, statistically significant predictive skill was found for both models, although this was higher for the UKMO model, which achieved correlations as high as 0.62 and 0.82 respectively for the winter-mean indices.

The seasonal averaging, as well as the ensemble averaging, is crucial for enhancing the signal-to-noise ratio. As discussed in Scaife et al. (2014), it should be noted that although the ensemble-mean forecast exhibits a strong correlation with the observations, its variance is considerably reduced due to the ensemble averaging. In most of the cases examined, the correlations between the individual ensemble members and the observations cover a wide range (from near zero or negative to values—in some cases—even higher than the correlation of the ensemble mean). Regarding the predictive skill for the NAO and the PNA pattern, it is worth mentioning that the ensemble-mean correlations for individual months were found to increase with lead time; that is, they seem to get higher as one moves from December to March. This peculiar, counterintuitive result may indicate the delayed emergence of a signal that arises from the initial state, possibly related to the ENSO teleconnection peaking in late winter as suggested, for example, in Manzini et al. (2006), or more directly via ocean–atmosphere interactions in the

FIG. 13. Correlation coefficient maps between the forecasted and the observed winter-mean (DJFM) cumulative frequencies of (left) instantaneous blocking and (right) blocking events, for (top) UKMO and (bottom) CMCC. Blocking occurrences have been integrated within $10^\circ$ lat $\times$ 20$^\circ$ lon window centered at each grid point. A semi-opaque gray mask is applied to correlations that do not pass the significance testing (see text). The white contours correspond to the climatological blocking frequency levels of 8 for IB (Fig. 3) and 1.5 for BE (Fig. 4).
North Atlantic (Czaja and Frankignoul 1999; Marshall et al. 2001). On the other hand, it could also be partly related to the initialization shock (decaying with lead time) having a negative impact on the predictive skill, particularly in the first month of the forecast.

A certain degree of uncertainty remains regarding the robustness of these statistics (high positive correlations interpreted as predictive skill) subject to different historical periods; to address this, better and longer sets of hindcasts are needed. Regarding the representation of low-frequency variability and the fidelity of the associated patterns (referring to the dominant EOFs), the CMCC model—which may be considered to represent low-resolution models with moderate ensemble sizes—was found to suffer from an underrepresentation of variance in the eastern North Atlantic basin more than the UKMO model.

As regards the predictive skill for the seasonal blocking frequency, the UKMO model exhibited considerable skill in the areas where the blocking frequency covaries more strongly with the NAO and the PNA pattern. These results call for more research to determine what drives this predictability and to what extent European and North American climate predictions benefit from it. For instance, Buehler et al. (2011) show that the occurrence of extreme cold or dry spells is dependent on North Atlantic blocking frequency, and therefore predicting blocking frequency may be of great value. Of course, blocking variability is not orthogonal to the NAO but directly related to it. Nevertheless, it is clear that winter variations in blocking frequency show significant predictability. Scaife et al. (2014) have already examined the origins of the NAO predictability, yet future dedicated studies may shed more light on the physical mechanisms that give rise to extratropical long-term predictability, referring also to the PNA pattern and the predictability of blocking frequency.

Acknowledgments. The CMCC authors were supported by the project GEMINA funded by the Italian Ministry of Education, University and Research and Ministry of Environment, Land and Sea. The Met Office authors were supported by the Joint DECC/Defra Met Office Hadley Centre Climate Programme (GA01101).

REFERENCES


