ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 36, JULY 2019, 753–765

• **Original Paper** •

Evaluation of the Forecast Performance for North Atlantic Oscillation Onset

Guokun DAI[1], Mu MU[1] and Zhina JIANG*[2]

[1]*Department of Atmospheric and Oceanic Sciences & Institute of Atmospheric Sciences, Fudan University, Shanghai* 200438*, China*

[2]*State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing* 100081*, China*

(Received 26 December 2018; revised 6 March 2019; accepted 12 April 2019)

ABSTRACT

By utilizing operational forecast products from TIGGE (The International Grand Global Ensemble) during 2006 to 2015, the forecasting performances of the European Centre for Medium-Range Weather Forecasts (ECMWF), National Centers for Environmental Prediction (NCEP), Japan Meteorology Agency (JMA) and China Meteorological Administration (CMA) for the onset of North Atlantic Oscillation (NAO) events are assessed against daily NCEP–NCAR reanalysis data. Twenty-two positive NAO (NAO+) and nine negative NAO (NAO−) events are identified during this time period. For these NAO events, control forecasts, one member of the ensemble that utilizes the currently most proper estimate of the analysis field and the best description of the model physics, are able to predict their onsets three to five days in advance. Moreover, the failure proportion for the prediction of NAO− onset is higher than that for NAO+ onset, which indicates that NAO− onset is harder to forecast. Among these four operational centers, ECMWF has performs best in predicting NAO onset, followed by NCEP, JMA, and then CMA.

The forecasting performance of the ensemble mean is also investigated. It is found that, compared with the control forecast, the ensemble mean does not improve the forecasting skill with respect to the onset time of NAO events. Therefore, a confident forecast of NAO onset can only be achieved three to five days in advance.

Key words: NAO onset, operational forecast, TIGGE dataset

Citation: Dai, G. K.,, M. Mu, and Z. N. Jiang, 2019: Evaluation of the forecast performance for North Atlantic Oscillation onset. *Adv. Atmos. Sci.*, **36**(7), 753–765, https://doi.org/10.1007/s00376-019-8277-9.

Article Highlights:

- Operational control forecasts can predict the onset of NAO events three to five days in advance.
- NAO− event onsets are more difficult to forecast compared with NAO+ event onsets.
- Ensemble mean forecasts cannot improve the skillful forecast time for NAO event onsets compared with control forecasts.

1. Introduction

The North Atlantic Oscillation (NAO) is a dominant atmospheric dipole mode in the Northern Hemisphere during wintertime, which is expressed as a seesaw sea level pressure (SLP) between the Icelandic low and the Azores high (Walker and Bliss, 1932; Feldstein, 2000; Woollings et al., 2008). Weather and climate in North America and Eurasia is greatly influenced by the NAO (Hurrell, 1995; Diao et al., 2015). During the positive phase of the NAO (NAO+), both the Icelandic low and the Azores high strengthen, which induces warming in North America and Europe. However, during the negative phase of the NAO (NAO−), both of these atmospheric centers weaken, which often causes cooling in North America and Europe (Marshall et al., 2001).

Many studies have investigated the NAO and its climatic effects from a monthly mean or seasonal mean perspective (Thompson and Wallace, 2000; Thompson et al., 2000; Zuo et al., 2016; Whan and Zwiers, 2017). They have found that extreme climatic events are usually associated with the NAO (Yiou and Nogaj, 2004; Scaife et al., 2008; Seager et al., 2010). Furthermore, Feldstein (2003) showed that the intrinsic time scale of the NAO is approximately two weeks. That is, the NAO can be investigated from a synoptic view. On this basis, we can investigate the different stages during its life cycle, such as its onset, development, and decay stages. From the perspective of its daily variability, many extreme cold events are related to NAO events. For example, Europe suffered from extreme cold during the winter of 2011/12, which was mainly caused by an atmospheric circulation transformation from NAO+ to eastern NAO− (Luo et al., 2014). In December 2013, a blizzard attacked the Middle East, which was due to the atmospheric circulation associated with an NAO+ decay stage (Luo et al., 2015). Therefore, skillful forecasts of NAO events may prevent a great amount of social loss.

* Corresponding author: Zhina JIANG
Email: jiangzn@cma.gov.cn

Utilizing daily forecast products, Johansson (2007) compared the forecast skill of the NAO from the European Centre for Medium-Range Weather Forecasts (ECMWF) and National Centers for Environmental Prediction (NCEP). Not only did this study calculate the anomaly correlation coefficient (ACC; Murphy and Epstein, 1989) of the 500 hPa geopotential field over the extratropical Northern Hemisphere (20°–80°N), but it also calculated the correlation of the NAO index (NAOI) between forecast products and observations. From their results, ECMWF offered improved skill for the NAO forecast compared to that of NCEP, regardless of the ACC or the correlation between the NAOI. Furthermore, Vitart (2014) assessed the NAO forecast skill of ECMWF from 2002 to 2011, again investigating it from the perspective of the NAOI correlation. The results showed that if a threshold of 0.60 was applied, the skillful forecast time of the NAO extended to 11 days in 2011 (Fig. 7 in his paper). However, both the correlation of the NAOI and the ACC for the geopotential field are descriptors of the NAO mode, and they cannot reflect the NAO intensity, let alone NAO events. In other words, previous studies have evaluated the forecast skill of the NAO mode rather than NAO events. Fortunately, Scaife et al. (2014) evaluated the seasonal forecast skill of NAO events, but without distinguishing between positive and negative phases. As we know, the features and mechanisms for the onset and development of NAO+ and NAO− events are quite different (Benedict et al., 2004; Luo et al., 2007; Barnes and Hartmann, 2010; Jiang et al., 2015). Benedict et al. (2004) pointed out that the formation of NAO+ is accompanied by anticyclonic wave breaking, and vice versa for NAO− events. Luo et al. (2007) used a theoretical model and found that NAO− events could be excited repeatedly, while downstream isolated dipole blockings could be triggered after NAO+ events decayed. All the above work inspired us to explore the forecasting performance of NAO+ and NAO− events separately.

Since atmospheric circulation is a complicated and chaotic system, a deterministic forecast is an estimation of the future atmosphere, which has large uncertainties (Lorenz, 1963). To improve the forecast skill, ensemble forecasts, which perturb initial conditions and model systems are applied. The ensemble mean of the perturbed forecasts may filter out the unpredictable part of the ensemble members and preserve the predictable part. Therefore, the ensemble mean is often more skillful than a deterministic forecast (Leith, 1974; Leutbecher and Palmer, 2008). Buizza and Leutbecher (2015) evaluated the ECMWF forecasting skill for the instantaneous grid-point 500 hPa geopotential height field in the Northern Hemisphere from 2012 to 2013. Based on their result, the deterministic forecast time was 17.0 days in advance for the Northern Hemisphere 500 hPa geopotential height, by utilizing the RMSE measurement. However, for the ensemble forecast, the forecast time was improved to 22.0 days by the measurement of the continuous ranked probability score. This inspired us to speculate that a longer skillful forecast time in advance might be achievable for NAO events if ensemble forecasts are applied.

Because NAO events are often related to extreme weather in the Northern Hemisphere during wintertime, it is necessary to investigate the forecasting performance for the onset of NAO events. Specifically: (1) How far in advance can NAO onset be predicted by different forecast centers? (2) Is there any difference in forecasting performance between the onset of positive and negative NAO events? (3) Can the ensemble mean forecast for the onset of NAO events achieve a better performance compared to the control forecast? The present paper seeks to answer these questions. In section 2, the data and methods used in the paper are described. The forecasting performances for the onset of NAO events based on control forecasts and ensemble mean forecasts are investigated in sections 3 and 4, respectively. Section 5 provides some discussion, and conclusions are drawn in section 6.

2. Data and methods

2.1. *Data*

The reanalysis data used in this paper are the daily SLP from the National Centers for Environmental Prediction and National Center for Atmosphere Research (NCEP–NCAR). The spatial resolution is $2.5° \times 2.5°$ and the time period is from 1 January 1958 to 31 December 2015 (Kalnay et al., 1996; Kistler et al., 2001).

The International Grand Global Ensemble (TIGGE) data are used to investigate operational weather forecasts from different centers. The program started in October 2006, which aimed to improve the forecasting skill for high impact weather events on a time scale from one day to two weeks. TIGGE data include daily weather forecasting products from 10 global centers (Park et al., 2008; Swinbank et al., 2016). In this paper, the daily operational forecasts of SLP from ECMWF, NCEP, the Japan Meteorology Agency (JMA) and the China Meteorological Administration (CMA) are used. The SLP fields were interpolated to a spatial resolution of $2.5° \times 2.5°$ before comparison. NCEP issues four forecasts daily, which start at 0000 UTC, 0600 UTC, 1200 UTC and 1800 UTC. Both ECMWF and CMA issue two forecasts, which start at 0000 UTC and 1200 UTC daily. However, JMA only issues one forecast, which starts at 1200 UTC. To obtain a fair comparison, a daily forecast starting from 1200 UTC is chosen. Moreover, the length of model integrations in ECMWF, NCEP, JMA and CMA are 15 days, 16 days, 9 days (11 days after 2013) and 10 days, respectively. For each operational center, there is a control forecast, which utilizes the currently most proper estimate of the analysis field and best description of the model physics, along with several ensemble perturbed members, which have perturbations on initial conditions and model systems (Park et al., 2008). The time period for comparison covers 1 November to 31 March from 2006 to 2015.

2.2. *NAOI and NAO events*

To identify an NAO event, the NAOI proposed by Li and Wang (2003) is used, which is defined as the difference in regionally normalized SLP zonally averaged over the Atlantic sector (i.e., longitudes from 80°W to 30°E between a line of midlatitude (35°N) and high latitude (65°N)]. This can be formulated as

$$\mathrm{NAOI} = \hat{P}_{35°\mathrm{N}} - \hat{P}_{65°\mathrm{N}} ,$$

where $\hat{P}$ represents the normalized SLP averaged from 80°W to 30°E. This index focuses on circulation in the North Atlantic sector, which provides a more faithful representation of the spatiotemporal variability of the NAO. Moreover, this index has the ability to recognize eastern NAO events, which performs better than the NAOI from the NOAA Climate Prediction Center (Luo et al., 2014).

Utilizing the method described in Li and Wang (2003) and the daily NCEP–NCAR reanalysis data, daily NAOI values from 1 January 1958 to 31 March 2015 are calculated. The climatic reference state is obtained by averaging daily SLP from 1958 to 2000, which is consistent with previous work (Li and Wang, 2003). To include more NAO cases, a time period from November to the following March is defined as wintertime.

According to the definition of an NAO event, if the NAOI is greater than 1.0 standard deviation for three or more consecutive days, an NAO event is identified. The period over which the NAOI is greater than 1.0 standard deviation is defined as the persistent episode and the first day of the persistent episode is considered as the onset day, while the last day of the event is called the decay day (Fig. 1). It is defined as an NAO+ (NAO−) event when the NAOI is positive (negative) during the persistent episode. This definition of NAO events is widely used in NAO-related studies (Luo et al., 2016; Song, 2016; Yao et al., 2016). Considering that there may be several intervals during NAO events, some rules should be made for selecting events. For two consecutive NAO events with the same phase, the second one will be omitted if the interval between the first decay and second onset is less than seven days. With these criteria, 22 NAO+ events and 9 NAO− events are selected for evaluation during the wintertime from 2006/07 to 2014/15 (Table 1). The duration is defined as the time interval

Table 1. The skillful forecast times and their corresponding forecast durations (numbers in brackets) of the control forecasts for the NAO events from 2006 to 2015.

	Date (duration)	ECMWF	NCEP	JMA	CMA
NAO+01	20061130 (16)	4 (≥6)	×	4 (≥5)	×
NAO+02	20061228 (23)	8 (≥7)	×	5 (≥4)	×
NAO+03	20071205 (4)	6 (3)	3 (5)	6 (≥3)	3 (4)
NAO+04	20080105 (3)	1 (5)	5 (3)	1 (3)	1 (4)
NAO+05	20080220 (11)	9 (≥6)	5 (6)	6 (≥3)	7 (≥3)
NAO+06	20080308 (5)	3 (4)	1 (5)	4 (≥5)	7 (≥3)
NAO+07	20081216 (3)	5 (9)	3 (4)	3 (≥5)	3 (≥7)
NAO+08	20090110 (3)	1 (3)	0 (-)	1 (3)	1 (3)
NAO+09	20091119 (5)	4 (4)	3 (5)	6 (≥3)	3 (≥7)
NAO+10	20110201 (7)	4 (≥11)	3 (9)	1 (≥8)	1 (≥8)
NAO+11	20111119 (3)	6 (6)	2 (3)	6 (≥3)	2 (3)
NAO+12	20120118 (3)	2 (3)	2 (3)	5 (≥3)	2 (3)
NAO+13	20120306 (5)	4 (6)	4 (4)	4 (≥5)	4 (≥6)
NAO+14	20131213 (16)	1 (≥14)	1 (≥15)	4 (≥5)	2 (4)
NAO+15	20140207 (3)	2 (3)	4 (3)	2 (3)	4 (3)
NAO+16	20140319 (5)	3 (3)	2 (4)	3 (3)	×
NAO+17	20141205 (3)	3 (3)	4 (4)	4 (≥4)	×
NAO+18	20141216 (7)	5 (3)	4 (7)	5 (≥3)	×
NAO+19	20141231 (17)	3 (≥12)	×	3 (≥6)	×
NAO+20	20150127 (3)	2 (3)	×	2 (3)	×
NAO+21	20150218 (22)	4 (≥10)	3 (≥12)	4 (≥5)	×
NAO+22	20150328 (4)	4 (4)	5 (4)	4 (≥5)	×
NAO−01	20070122 (5)	1 (4)	×	1 (4)	×
NAO−02	20081226 (10)	8 (≥6)	3 (≥13)	4 (≥5)	×
NAO−03	20090201 (3)	3 (≥12)	4 (≥11)	3 (≥6)	3 (≥7)
NAO−04	20090212 (3)	5 (≥10)	4 (≥12)	3 (≥6)	0 (-)
NAO−05	20100205 (33)	4 (5)	5 (≥11)	6 (≥3)	1 (≥9)
NAO−06	20101122 (16)	7 (5)	5 (≥11)	0.5*(≥3)	0.4 (≥5)
NAO−07	20110121 (3)	3 (3)	2 (7)	2 (3)	2 (3)
NAO−08	20121218 (5)	7 (≥7)	7 (5)	3 (≥4)	5 (≥4)
NAO−09	20130219 (6)	3 (≥10)	2 (≥13)	4 (≥5)	2 (≥8)

Notes: ≥6 means that the forecasted duration is longer than 6 days. However, we cannot get the exact number due to the data limitation; × represents a case from the respective center that is not investigated because the products for the case are partially missing. 0.5* indicates that for this event the center could forecast it five days in advance. However, the forecasted onset is one day earlier or later than the observation.

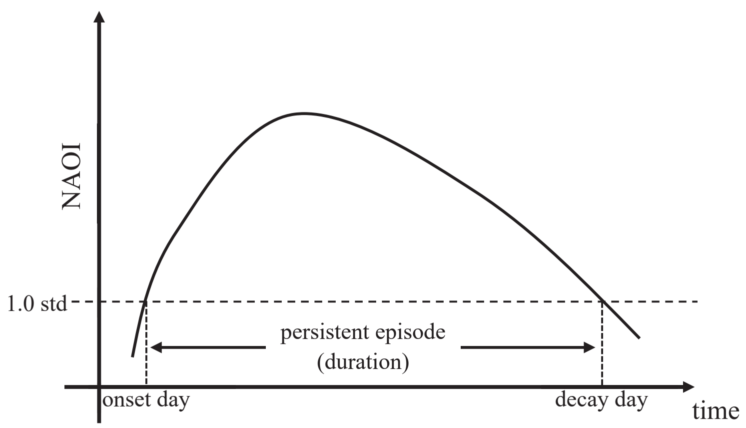

Fig. 1. Schematic diagram of an NAO+ event. An episode that has an NAOI larger than 1.0 standard deviation for at least three consecutive days is defined as a persistent episode or duration. The first (last) day on which the NAOI is larger than 1.0 standard deviation is called the onset (decay) day.

between the onset day and decay day. For these selected NAO events, their durations vary from 3 to 33 days. Although some events have durations as short as 3 days, others could be as long as nearly one month, they could all have impact on adjacent weather.

2.3. *Skillful forecast time*

The skillful forecast time is defined as the longest day that the NAO onset could be forecasted. Specifically, the forecast products from one day in advance are used to identify whether or not an NAO event occurs. The criterion for determination is whether the NAOI derived from the operational centers exceeds 1.0 standard deviation and persists for at least three consecutive days. The NAO onset can be forecasted one day in advance if the criterion is met, and the forecast made two days in advance is tested in the same way. Products with different forecast times are tested and the longest day that is able to meet the condition is defined as the skillful forecast time.

3. Evaluation of the control forecast performance

As already mentioned, the starting forecast time of 1200 UTC was chosen for these four operational centers because JMA only issues forecasts from 1200 UTC every day. However, the outputs from all operational centers are instantaneous fields at 0000, 0600, 1200 and 1800 UTC. Considering that the daily mean reanalysis SLP field is used in the definition of NAO events and our evaluation is made for daily mean SLP fields, operational forecast products at 0000 UTC, 0600 UTC, 1200 UTC and 1800 UTC are averaged as the daily mean SLP field.

3.1. *Evaluation of NAO+ onset*

In this subsection, we take the third case in Table 1 (an NAO+ event that occurred on 5 December 2007) as an example to illustrate how to define the skillful forecast time. The case was chosen because NCEP and CMA joined the TIGGE project in March 2007 and May 2007, respectively.

Figure 2 shows the SLP anomaly evolution during the onset of NAO+. We find that, before the onset of NAO+, there are several negative SLP anomalies at high latitudes (approximately 60°N) in the North Atlantic sector and the negative anomalies deepen with time. Moreover, positive anomalies move towards lower latitudes (approximately 30°N) in the North Atlantic sector and intensify after lag-4 day (Fig. 2d). Then, the negative-over-positive dipole SLP anomalies in the North Atlantic sector trigger the onset of NAO+.

Figure 3 illustrates the specific method for defining the forecast time for the onset of NAO+ events. From the NCEP–NCAR reanalysis data, the onset day for the event is 5 December 2007. From Fig. 3 it is apparent that, for the onset of NAO+, ECMWF can predict six days in advance. However, for NCEP, JMA and CMA, the skillful predictions of NAO+ onset are three, six and three days in advance, respectively.

Utilizing the same method, the skillful forecast time for 22 NAO+ events from 2006/07 to 2014/15 are identified with forecast products from the four centers (Table 1). It can be seen that, for the 22 NAO+ event onsets, almost all of them can be predicted in advance, except for one case from NCEP, which occurred on 10 January 2009. The skillful forecast time has a large range, from one to nine days. For the pre-

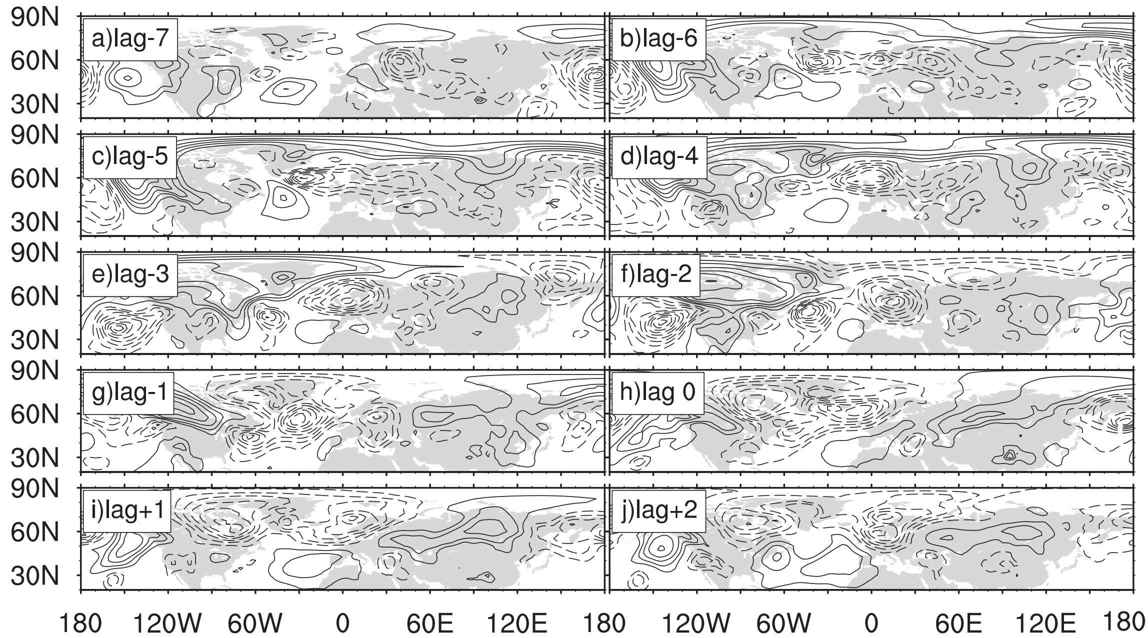

Fig. 2. SLP anomaly evolution for an NAO+ event. Lag 0 corresponds to the onset day of the NAO+ event, which is 5 December 2007. The contour interval is 5 hPa. Solid and dashed lines represent positive and negative values, respectively.

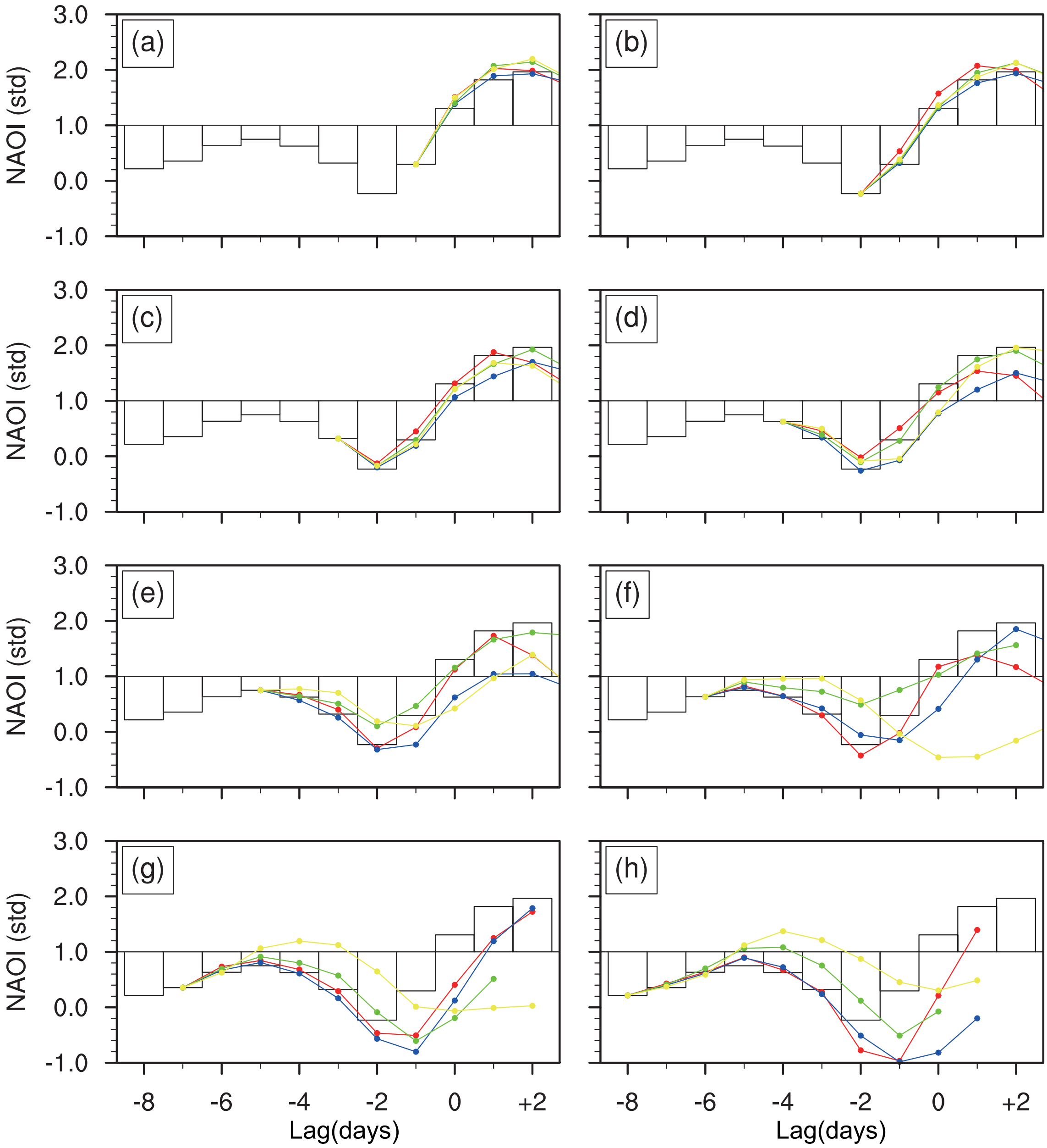

Fig. 3. Schematic illustrations of the NAO+ event onset forecast times, with the panels (a) to (h) representing forecast times of one to eight days in advance, respectively. The vertical axis is the normalized NAOI, and the histogram represents the reanalysis index. Red, blue, green and yellow lines represent the forecast indexes from ECMWF, NCEP, JMA and CMA, respectively. Lag 0 corresponds to the onset of the NAO+ event onset on 5 December 2007.

diction of NAO+ onset, the mean skillful forecast time of ECMWF is 3.82 days, which is the longest among the four centers. The mean skillful forecast time is 3.18 days for NCEP, 3.77 days for JMA and 3.08 days for CMA. From the above results, we can conclude that the mean skillful forecast time for the onset of NAO+ onset is approximately three to four days, which is much less than the eleven days derived from NAOI correlation by Vitart (2014). This difference may be caused by two main reasons. On the one hand, Vitart (2014) focused on the relationships of the NAOI annually, whilst our investigations focus only on NAO events during wintertime. On the other hand, whether or not the NAOI forecast exceeds 1.0 standard deviation on a particular day is crucial in our assessment. However, Vitart (2014) paid more attention to the NAOI correlation between forecasts and observations. Moreover, the skillful forecast time from ECMWF is longer than that from NCEP, which is consistent with the results of Johansson (2007). However, in most cases, the predicted NAOI is smaller than that derived from the reanalysis data, though both have close values.

We focus mainly on the relationship between the NAOI derived from the reanalysis data and the forecast products in previous works. However, the NAOI cannot represent all circulation features in the North Atlantic sector, although the

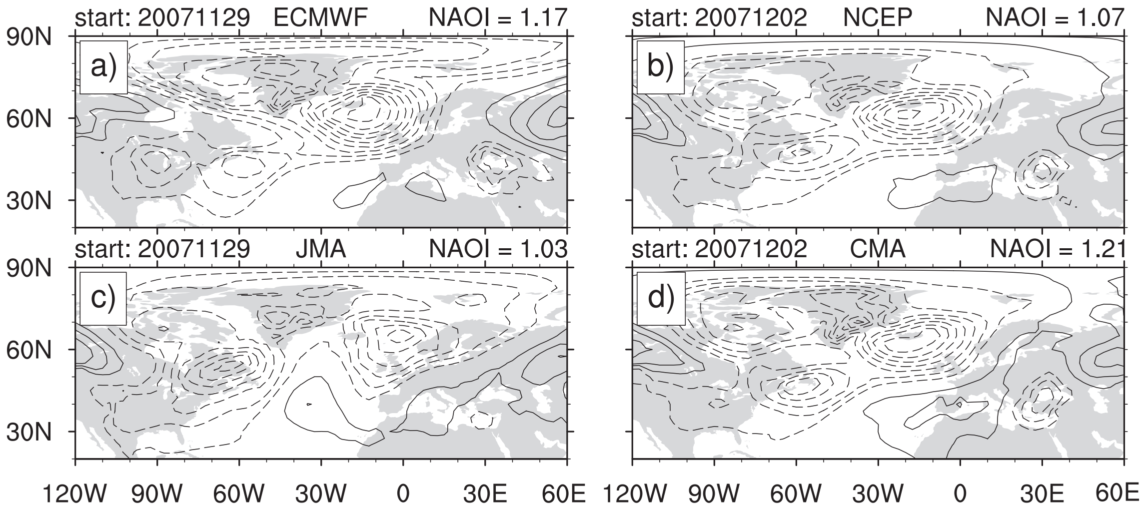

Fig. 4. SLP anomalies for the NAO+ event that occurred on 5 December 2007, as forecast by the different centers at their respective skillful forecast times: (a) ECMWF on 29 November 2007; (b) NCEP on 2 December 2007; (c) JMA on 29 November 2007; (d) CMA on 2 December 2007. Contour interval: 5 hPa.

index can be used to measure the phase and intensity of the NAO. To what extent, then, can circulation in the North Atlantic sector be predicted? To answer this, the predicted SLP anomalies for the onset of the NAO+ event on 5 December 2007 from the four centers are given in Fig. 4, with their own respective forecast times. The forecasted SLP anomalies show a multicenter structure of negative anomalies, which has close similarity with that of the reanalysis (Fig. 2). To quantify this similarity, the ACC of the SLP anomalies in the North Atlantic sector between the forecast and the reanalysis data is calculated. The ACC is 0.852 for ECMWF with a skillful lead time of six days, 0.954 for NCEP with a skillful lead time of three days, 0.780 for JMA with a skillful lead time of six days, and 0.931 for CMA with a skillful lead time of three days, all of which are much larger than 0.60. This indicates that the evaluation of NAO onset with the NAOI in our work is more rigorous.

Similar work was carried out for the onset of 22 NAO+ events (Fig. 5a). For most cases, the ACC exceeds 0.60, which can be viewed as skillful forecasts (Hollingsworth et al., 1980). However, there are two cases with low ACCs. We find that the ACC has a low value of 0.398 in ECMWF with a lead time of eight days for the second NAO+ case (onset on 28 December 2006, Fig. 6), and the ACC is 0.442 from CMA with a lead time of seven days for the sixth NAO+ case (onset on 8 March 2008, Fig. 7). For the second NAO+ case, there is a negative SLP anomaly mainly located over the mid to high latitudes between 70°W and 20°W. A positive SLP anomaly occupies western Europe and exhibits a west-negative and east-positive pattern in the North Atlantic sector (Fig. 6i). With one to six days in advance, the ECMWF can predict SLP anomalies with west-negative and east-positive patterns. Furthermore, not only does the ACC between the forecast products and the reanalysis data exceed 0.80, but the predicted NAOI also has a close value with that of the reanalysis data (Figs. 6a–f). However, with a lead time of seven

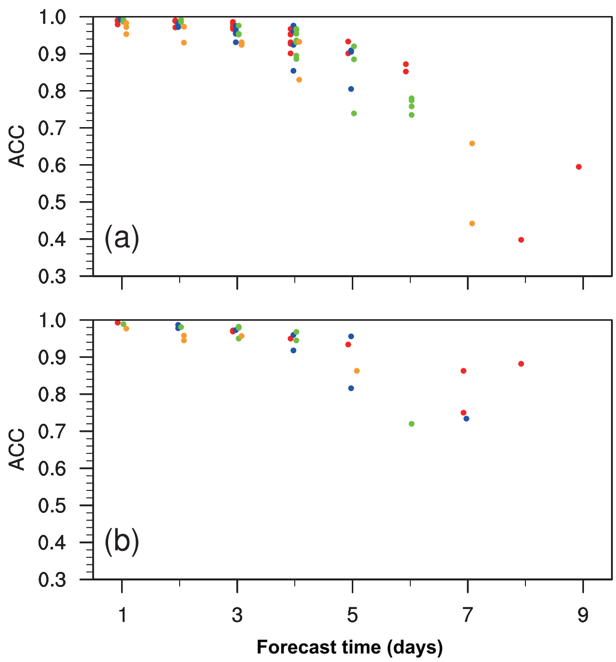

Fig. 5. The ACC of the SLP anomalies in the North Atlantic sector between the reanalysis and the products from four centers with their respective forecast times for (a) 22 NAO+ events and (b) 9 NAO− events. Red, blue, green and yellow dots represent the forecasts from ECMWF, NCEP, JMA and CMA, respectively.

to eight days, the forecasted negative SLP anomaly in southern Greenland is more intensive, and the forecasted positive anomaly is more northward than the observed one. Moreover, the negative and positive SLP anomalies from the forecast show a northwest–southeast tilt rather than the zonal tilt that is observed from the reanalysis (Figs. 6g and h). As a consequence of this pattern, the ACC has a low value of 0.64 for seven days (0.40 for eight days), and the forecasted NAOI

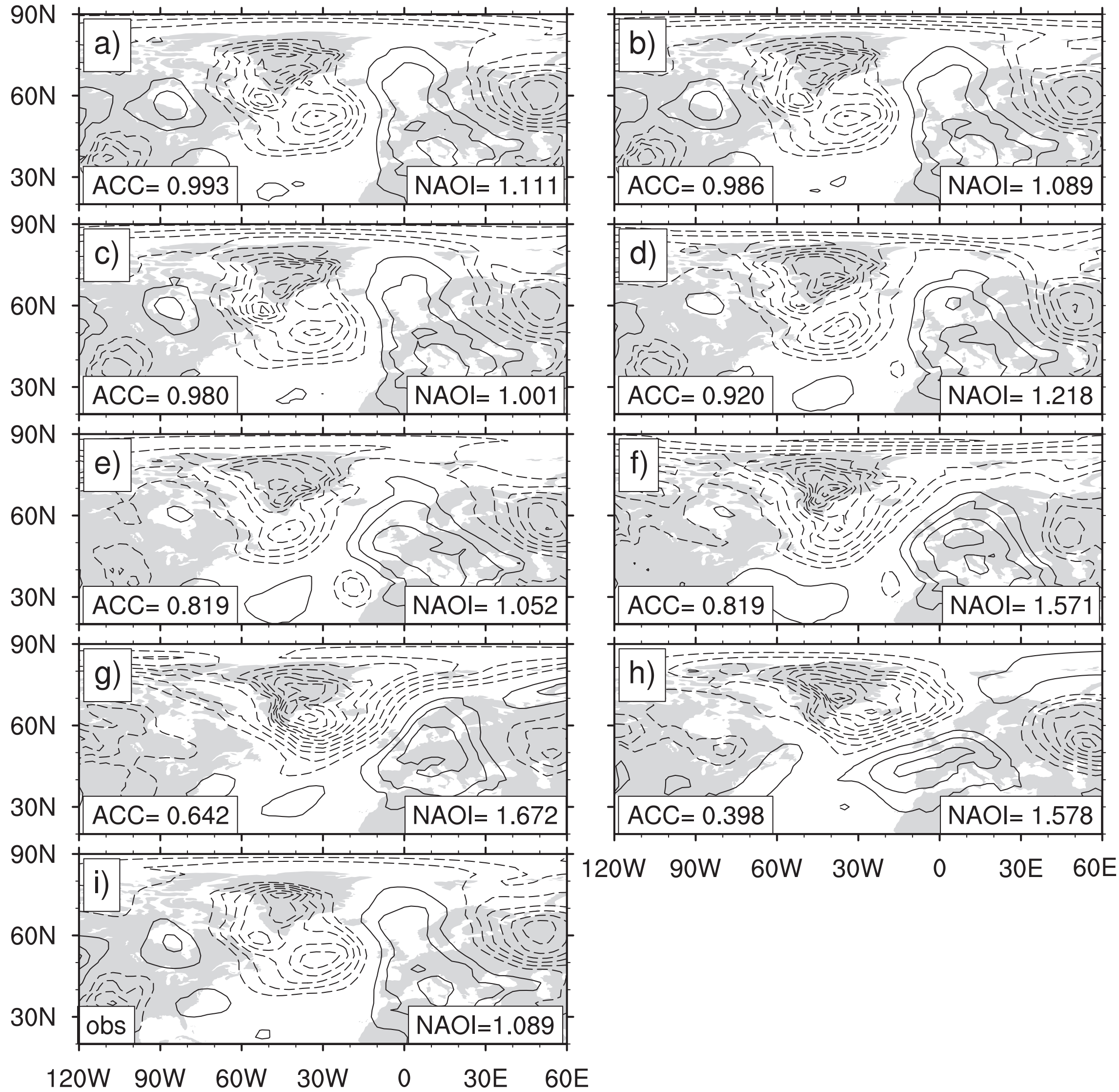

Fig. 6. The forecasted and observed SLP anomalies in the North Atlantic sector on 28 December 2006. Panels (a) to (h) represent the ECMWF forecast one to eight days in advance, respectively. Panel (i) represents the SLP anomaly derived from the reanalysis (contour interval: 5 hPa). Solid and dashed lines represent positive and negative values, respectively. The zero line is omitted.

has a much larger value than that of the reanalysis as well. Although the forecasted NAOI can trace the onset of the NAO+ event, the forecasted SLP anomaly is dissimilar from that of the reanalysis.

As for the sixth NAO+ case, the reanalysis shows a negative-over-positive SLP anomaly, with a northeast–southwest tilt in the North Atlantic sector (Fig. 7i). With a lead time shorter than six days, the SLP anomaly forecasted by CMA shows a northeast–southwest tilted dipole, which is similar to that in the reanalysis, and the ACC between the forecast and the reanalysis exceeds 0.85 (Figs. 7a–f). When the lead time approaches seven days, the forecasted SLP anomalies exhibit a meridional dipole with no tilt (Fig. 7g). As a result, the forecasted ACC is as low as 0.44, and the NAOI has a value of 1.574, which is much larger than the value of 1.036 derived from the reanalysis. That is, although the forecasted NAOI can trace the NAO+ onset with a lead time of seven or eight days, the predicted circulation pattern is dissimilar to that of the reanalysis. This confirms the conclusion that, for NAO+ onset, skillful forecasts can only be made several days (approximately three to four days, on average) in advance.

3.2. *Evaluation of NAO− onset*

Similar to the evaluation for NAO+ events, we perform an assessment on the nine NAO− event onsets as well. The forecast times for the onset of these NAO− events are shown in Table 1. Among the cases, most of the onsets can be forecast several days in advance, ranging from one to eight days. However, for the 33 forecasts in total, three of them failed. The proportion of NAO− event onset forecasting failures is much higher than that of NAO+ events, which only has one

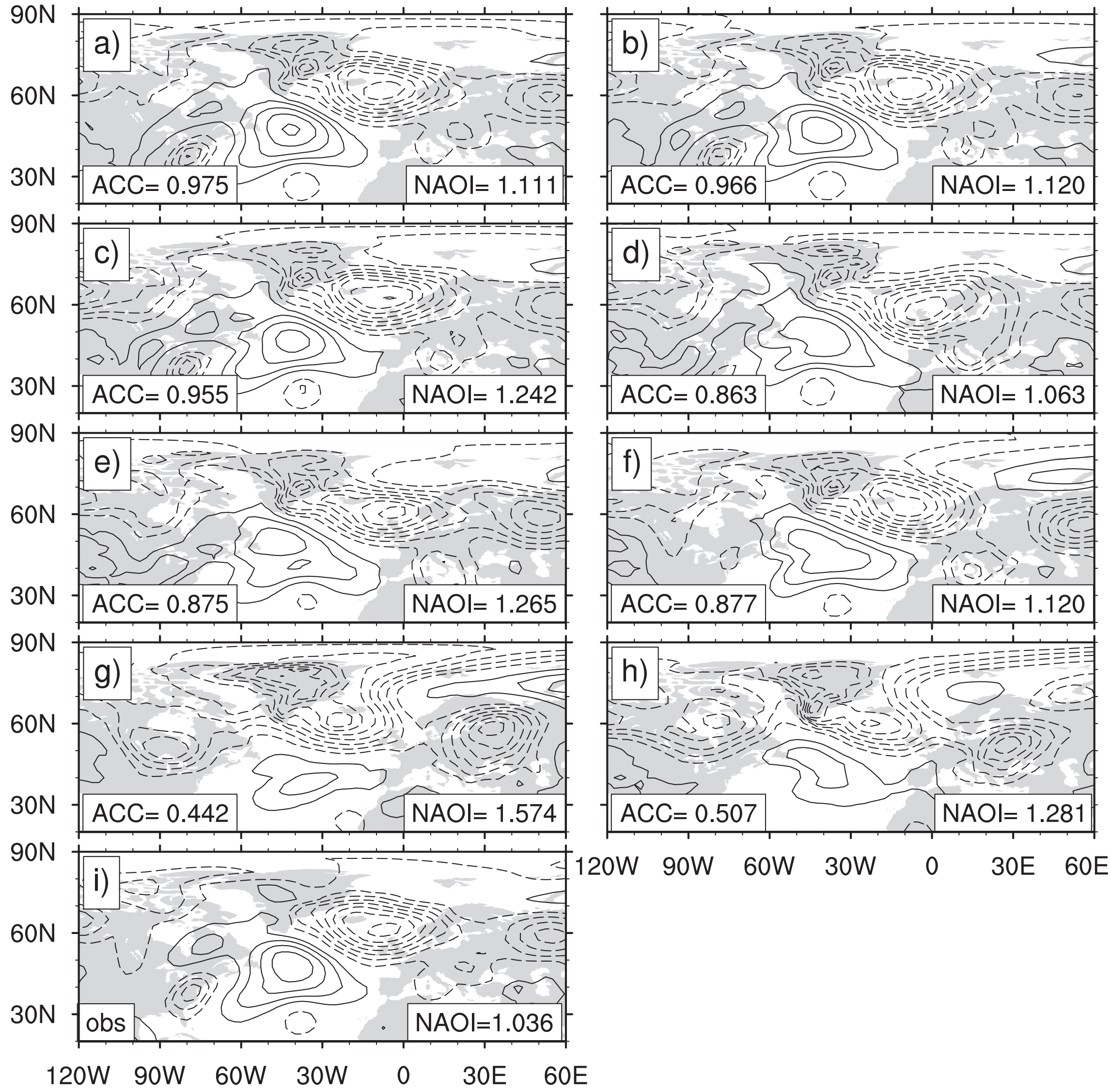

Fig. 7. As in Fig. 6 but for the case on 8 March 2008 from CMA.

failure among the 75 forecasts. This may indicate that forecasting NAO− onset is more difficult than forecasting NAO+ onset. On average, the skillful forecast time for NAO− onset is 4.56 days for ECMWF. ECMWF has the longest forecast time among the four centers, which is the same result as for NAO+ onset. The mean skillful forecast time for these NAO− events is 4.00 days, 3.25 days and 2.60 days for NCEP, JMA and CMA, respectively.

Similar to the NAO+ onset forecasts, most NAO− forecasts underestimate the intensity of the SLP field in meeting the criterion of NAO onset. However, by calculating the ACC of the SLP anomalies in the North Atlantic sector between the forecast products and the reanalysis data, we find that the ACC exceeds 0.7, even for a lead time of eight days (Fig. 5b). That is, the forecast is skillful for NAO− onset even if the lead time is longer than one week. This may be due to the small number of NAO− used in the assessment.

From the above results, we can conclude that the four chosen centers have the ability to forecast NAO onset several days in advance. However, the skillful forecast time is short (three to five days, on average). On the other hand, the proportion of failures for NAO− onset prediction is higher than that for NAO+, which indicates that NAO− events are harder to predict. This difference in forecasting performance may relate to their different physical mechanisms. Specifically, NAO− (NAO+) events being difficult (easy) to predict is likely related to the strong (weak) nonlinearity of NAO− (NAO+) events, behaving with weak (strong) energy dispersion (Luo et al., 2007). This strong (weak) nonlinearity and weak (strong) energy dispersion makes NAO− (NAO+) events difficult (easy) to predict.

4. Ensemble mean forecast performance

In the section above, we discussed the control forecast performance for the prediction of NAO onset. However, besides the control forecast, operational forecast centers also produce ensemble forecasts by perturbing initial conditions and model systems.

In operational forecasts, ECMWF and JMA both perturb initial conditions with the singular vector method. Each center produces 25 pairs of perturbations on the initial conditions. Together with the model stochastic process, ECMWF and JMA each produce 50 perturbed forecasts (the number of initial perturbation pairs in JMA reduced to 13, and the number of perturbed ensemble forecasts decreased to 26 after 2014). NCEP produces 10 pairs of initial perturbations with ensemble transform rescaling. Together with the model perturbations, there are 20 perturbed forecasts in the NCEP ensemble. However, for CMA, seven pairs of initial perturbations are generated with the breeding vector method. Without any perturbations in the model integration, CMA produces 14 perturbed forecasts in its ensemble. Previous investigations have proven that the ensemble mean can achieve a higher skill in operational forecasts (Leutbecher and Palmer, 2008). However, for the onset of NAO events, whether or not ensemble mean forecasts can improve the performance is unknown.

As already mentioned, there are 50 perturbed forecasts from ECMWF. Together with the control forecast, the ensemble mean is obtained by averaging 51 forecasts with the same weight. Similarly, the ensemble mean forecasts of JMA, NCEP and CMA are achieved in the same way.

Through a similar method as used in section 3, the skillful forecast time of the ensemble mean is assessed for the onset of 22 NAO+ events. From the results (Table 2), we can see that most can be predicted one to seven days in advance. The average skillful forecast time for NAO+ onset is 3.86 days for ECMWF, 3.35 days for NCEP, 3.27 days for JMA, and 2.77 days for CMA. We find that the skillful forecast time of the ensemble mean is not significantly improved compared with the control forecast. Moreover, the skillful forecast times derived by the ensemble mean of the JMA and CMA are somewhat shorter than that derived by the control forecast. For example (i.e., case 5 on 20 February 2008), the skillful forecast time derived by the control forecast for NAO+ onset is nine days in advance for ECMWF, six days in advance for JMA, and seven days in advance for CMA. However, from the perspective of the ensemble mean, the skillful forecast time is seven days for ECMWF, four days for JMA, and five days for CMA. It seems that the skillful forecast time for NAO+ onset derived by the ensemble mean is shorter than that by the control forecast. For the above case, this may be due to the fact that the ensemble mean weakens the strength of the NAO mode, so that it is too weak to satisfy the NAO event criteria. The limited ensemble members and limited cases presented here is another possible reason. Moreover, for case 8 of the NAO+ events, which occurred on 10 January 2009, neither the control forecast nor the ensemble mean forecast from NCEP could predict its onset in advance.

To further verify the circulation pattern forecasted by the ensemble mean, the ACC of the SLP field between the reanalysis and the ensemble mean forecast in the North Atlantic sector are calculated at NAO onset. From Fig. 8a, it is obvious that ensemble mean forecasts have higher ACCs than that of control forecasts in most cases, especially for cases

Table 2. As in Table 1 but for the ensemble mean forecasts.

	ECMWF	NCEP	JMA	CMA
NAO+01	4 (⩾6)	×	4 (⩾5)	×
NAO+02	6 (⩾9)	×	4 (⩾5)	×
NAO+03	5 (3)	3 (5)	4 (⩾4)	3 (4)
NAO+04	1 (5)	5 (4)	1 (4)	1 (4)
NAO+05	7 (4)	5 (4)	4 (⩾5)	5 (⩾4)
NAO+06	3 (4)	1 (5)	2 (4)	7 (⩾3)
NAO+07	5 (6)	3 (4)	2 (3)	3 (⩾7)
NAO+08	1 (3)	0 (-)	1 (3)	1 (3)
NAO+09	5 (5)	6 (3)	4 (⩾4)	3 (4)
NAO+10	4 (⩾9)	3 (7)	1 (⩾8)	1 (⩾8)
NAO+11	6 (4)	2 (3)	6 (⩾3)	3 (4)
NAO+12	2 (3)	2 (3)	3 (3)	2 (3)
NAO+13	4 (7)	3 (5)	4 (⩾5)	1 (6)
NAO+14	1 (⩾14)	1 (⩾14)	1 (⩾8)	2 (5)
NAO+15	3 (3)	5 (3)	1 (3)	4 (3)
NAO+16	3 (3)	2 (3)	5 (⩾3)	×
NAO+17	3 (4)	3 (4)	2 (3)	×
NAO+18	6 (4)	4 (4)	5 (⩾3)	×
NAO+19	3 (⩾12)	×	3 (⩾6)	×
NAO+20	3 (3)	×	4 (⩾4)	×
NAO+21	4 (⩾11)	3 (⩾12)	5 (⩾4)	×
NAO+22	6 (5)	6 (7)	6 (3)	×
NAO−01	1 (3)	×	1 (3)	×
NAO−02	3 (⩾10)	3 (⩾9)	3 (⩾6)	×
NAO−03	4 (5)	4 (⩾11)	2 (3)	3 (⩾6)
NAO−04	5 (⩾10)	0 (-)	3 (⩾6)	0 (-)
NAO−05	6 (⩾9)	5 (⩾10)	6 (⩾3)	1 (⩾9)
NAO−06	8 (⩾7)	1 (⩾14)	6 (⩾3)	0.5 (⩾6)
NAO−07	3 (3)	2 (3)	3 (3)	2 (3)
NAO−08	6 (5)	7 (5)	3 (⩾6)	5 (⩾4)
NAO−09	3 (5)	1 (12)	1 (⩾5)	2 (⩾8)

when the ACC in the control forecast is less than 0.60. This illustrates that the ensemble mean could achieve a more accurate circulation pattern in the North Atlantic sector, although it cannot improve the skillful forecast time of NAO+ onset compared with the control forecast.

Similar work was carried out for the onset of nine NAO− onsets. The ensemble mean skillful forecast times from the four centers are shown in Table 2. From the results, we can see that the ensemble mean forecasts can predict NAO− onset one to eight days in advance for most cases. On average, the ensemble mean skillful forecast time is 4.33 days for ECMWF, 3.29 days for NCEP, 3.11 days for JMA, and 2.60 days for CMA.

Similarly, the skillful forecast time for NAO− onset derived by the ensemble mean does not offer any improvement. Moreover, the skillful forecast time derived by the ensemble mean is somewhat shortened compared with that of the control forecast, except for CMA. For example, the control forecast from ECMWF can predict the second NAO− onset case (which occurred on 26 December 2008) eight days in advance, but the ensemble mean skillful forecast time is only three days in advance. For the fourth NAO− onset case (which occurred on 12 February 2009), the control forecast from NCEP can predict its onset four days in advance. How-

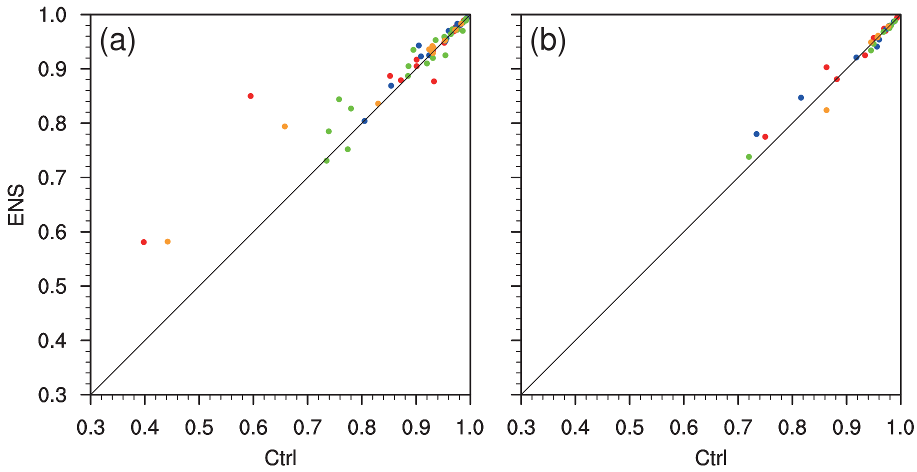

Fig. 8. Comparison of ACC values derived by the control forecasts and ensemble mean forecasts for the onset of (a) NAO+ events and (b) NAO− events.

ever, the ensemble mean forecast could not recognize this NAO− event.

However, not all of the ensemble means shorten the valid forecast time of NAO− onset. For example, in the control forecast, JMA could not predict the sixth NAO− onset case (which occurred on 22 November 2010), but its ensemble mean resulted in a skillful forecast time of six days in terms of the prediction of NAO− onset. Similarly, NCEP could predict this case eight days in advance from its ensemble mean forecast, which is longer than that from its control forecast time (seven days). This may indicate that whether or not the ensemble mean forecast achieves a longer skillful forecast time for NAO onset depends on the case and the model.

Inspired by the results for NAO+ events, we also investigated the circulation pattern in the North Atlantic sector for NAO− events forecasted by ensemble mean at onset time. From Fig. 8b, we can also see that the ensemble mean only improves the SLP forecast slightly, because the control forecast has done a good job of predicting the SLP (Fig. 5b).

5. Discussion

Besides the above evaluation of the operational forecast performance for NAO onset, there are still many aspects that need to be discussed.

5.1. *Forecast performance of NAO duration*

It is widely known that different durations of NAO events can generate different influences on weather over local and adjacent regions. Therefore, the forecast performance for NAO event duration should be further investigated. Similar to the definition of NAO duration in reanalysis data, the NAO duration from the TIGGE dataset is also defined as the time interval between the onset day and decay day, but with the products starting from the corresponding skillful forecast time in advance. For example, for the NAO+03 event, which occurred on 5 December 2007, ECMWF can forecast the onset six days in advance (i.e., the start time is 1200 UTC 29 November 2007). With the forecast starting at 1200 UTC 29 November 2007, the corresponding onset day is 5 December 2007 and the decay day is 7 December 2007. Thus, the duration of this NAO event from ECMWF is three days. JMA can also forecast the NAO onset with the start time of 1200 UTC 29 November 2007. However, due to its forecast length limit, it cannot forecast the event decay with six days in advance of the onset. Thus, the duration is marked as ⩾3 days. But for NCEP and CMA, the onset forecast assessment shows that their skillful forecast time for this event is three days in advance, so that their duration forecasts are assessed with the products starting from 1200 UTC 2 December 2007. The corresponding forecast durations are five and four days for NCEP and CMA, respectively. With this method, NAO duration forecast assessments are shown in brackets in Table 1.

If the NAO event duration derived from forecast products has the same length as that derived from NCEP–NCAR reanalysis data, it is called an accurate prediction of NAO event duration. Similarly, an underestimated (overestimated) duration prediction means that the predicted duration by the model is shorter (longer) than that derived from reanalysis data. For the control forecasts, 46 out of 75 NAO+ forecasts have an accurate prediction of duration. However, there are 18 NAO+ events whose durations are overestimated and 10 NAO+ events whose are underestimated. For the 33 NAO− events, 15 of them have accurate duration predictions, while there are 13 that are overestimated and four that are underestimated (Table 1). The results show that about half of the forecasts perform well in NAO duration prediction. As for the NAO events with durations longer than two weeks (5 NAO+ events and 2 NAO− events), the duration forecasts have good consistency with NCEP–NCAR reanalysis. However, the total duration cannot be predicted in these cases due to the limit of forecast length. But, for the NAO events with durations shorter than one week, numerical models tend to overestimate their durations. Overall, the forecasting performance for NAO+ duration is better than that of NAO−, which also indicates that NAO− events are harder to predict than NAO+

events. As for the ensemble mean forecasts of NAO duration, the results are similar to those of the control forecasts (Table 2).

5.2. *Performance sensitivity to sample number*

The mean skillful forecast time for the above is acquired by averaging all NAO cases for each operational center. However, there are some missing cases in NCEP and CMA. Therefore, in this subsection, only 13 NAO+ events and 7 NAO− events that are in common across the four operational centers are compared. Comparing Table 4 with Table 3, it can be seen that their mean skillful forecast times have changed and the sequence for the four centers has adjusted somewhat. For example, if all the cases are taken into consideration, the average skillful forecast time for NAO+ onset is 3.82 days for ECMWF, which is the longest among the four centers (Table 3). If only common cases are taken into consideration, JMA performs best for forecasting NAO+ onset (Table 4). However, the conclusion that the mean skillful forecast time for NAO onset is three to five days is still robust. Moreover, it can also be found that there are three failures among the 28 NAO− onset forecasts, but only one failure among the 52 NAO+ predictions (Table 4). This further confirms that NAO− onset is harder to forecast than NAO+ onset.

In addition, since the TIGGE project started in 2006, this study was limited by the available data length, thus possibly limiting the number of cases to establish statistical significance. To overcome this limitation, NAO events with a criterion of 0.6 standard deviations are also investigated. With this criterion, there are 33 NAO+ and 20 NAO− events during wintertime 2006/07 to 2014/15. From the results (Table 5), it can be seen that the skillful forecast time for NAO onset is three to four days on average, which is similar to the results using the criterion of 1.0 standard deviation. Furthermore, the results for these cases also show that ensemble mean forecasts make little contribution to improving the skillful forecast time for NAO onset. It is also found that NAO− onset is harder to forecast than NAO+ onset, since there are seven failures among the 73 NAO− forecasts and no failures among the 109 NAO+ forecasts.

Table 3. Comparison of the NAO event onset forecasts among the four centers.

		ECMWF	NCEP	JMA	CMA
NAO+ success ratio	ctrl	22/22	17/18	22/22	13/13
	ens-mean	22/22	17/18	22/22	13/13
NAO+ skillful forecast time	ctrl	3.82	3.18	3.77	3.08
	ens-mean	3.86	3.35	3.27	2.77
NAO− success ratio	ctrl	9/9	8/8	8/9	5/7
	ens-mean	9/9	7/8	9/9	5/7
NAO− skillful forecast time	ctrl	4.56	4.00	3.25	2.60
	ens-mean	4.33	3.29	3.11	2.60

Table 4. As in Table 3 but for the 13 NAO+ events and 7 NAO− events common across the four centers.

		ECMWF	NCEP	JMA	CMA
NAO+ success ratio	ctrl	13/13	12/13	13/13	13/13
	ens-mean	13/13	12/13	13/13	13/13
NAO+ skillful forecast time	ctrl	3.69	3.00	3.77	3.08
	ens-mean	3.62	3.25	2.62	2.77
NAO− success ratio	ctrl	7/7	7/7	6/7	5/7
	ens-mean	7/7	6/7	7/7	5/7
NAO− skillful forecast time	ctrl	4.57	4.14	3.50	2.60
	ens-mean	5.00	3.33	3.43	2.60

Table 5. As in Table 3 but for the NAO events with a threshold of 0.6 standard deviations.

		ECMWF	NCEP	JMA	CMA
NAO+ success ratio	ctrl	30/30	25/25	33/33	21/21
	ens-mean	30/30	24/25	33/33	18/20
NAO+ skillful forecast time	ctrl	3.87	3.28	3.45	2.43
	ens-mean	4.03	3.29	3.24	2.55
NAO− success ratio	ctrl	19/20	17/18	17/20	13/15
	ens-mean	19/20	17/18	17/20	13/15
NAO− skillful forecast time	ctrl	3.47	3.47	4.00	3.23
	ens-mean	3.63	3.23	3.24	3.08

5.3. *Limitation of the NAOI for ensemble forecasting*

In this study, the same NAOI was used for the control and ensemble mean forecasts, so that their forecast performances could be compared directly. However, as shown above, the NAOI derived from the ensemble mean usually has a small amplitude due to the averaging of ensemble members. This leads to the meridional dipole mode not being strong enough to meet the criteria of an NAO event. As a result, the skillful forecast time for NAO onset derived by the ensemble mean is usually short. It is possible that another NAOI index that considers each ensemble member might compensate for this shortcoming. However, this needs further investigation in future work.

6. Summary

Utilizing operational forecast products from ECMWF, NCEP, JMA and CMA from 2006 to 2015, we assessed the forecast performance for NAO onset. Twenty-two NAO+ events and nine NAO− events were selected during this period with daily NCEP–NCAR SLP reanalysis data based on the NAOI defined by Li and Wang (2003).

For the onset of NAO+ events, control forecasts from four centers can predict them several days in advance, ranging from one to nine days. On average, the skillful forecast time of NAO+ onset derived by the control forecast is 3.82 days for ECMWF, which is the longest among the four centers, and then 3.77 days, 3.18 days and 3.08 days for JMA, NCEP and CMA, respectively. The NAOI derived by the control forecast is close in value to that derived by the reanalysis data on NAO onset day, but with a weaker NAO mode. However, the ACC of the SLP anomalies in the North Atlantic sector between the control forecast and reanalysis exceeds 0.60 with a lead time of six days.

A similar evaluation was carried out for the onset of NAO− events. The four centers can predict the onset of NAO− events with a lead time of one to eight days for most cases. However, the failure proportion for NAO− onset prediction is higher than that for NAO+, which indicates that NAO− onset is harder to forecast. On average, the skillful forecast time derived by the control forecast is 4.56, 4.00, 3.25 and 2.60 days for ECMWF, NCEP, JMA and CMA, respectively. Similarly, the NAO mode derived by the control forecast with skillful forecast time on NAO onset day is also somewhat weaker than that from the reanalysis data. However, the ACC of the SLP anomalies in the North Atlantic sector between the control forecast and reanalysis exceeds 0.60, even when the lead time extends to eight days.

In addition, the ensemble mean forecast performance regarding NAO onset was also evaluated. The skillful forecast time derived by the ensemble mean is shorter than that derived by the control forecast, for both NAO+ and NAO− onset. This is because the NAO mode derived by the ensemble mean is weaker than that derived by the control forecast, which induces the NAO mode derived by the ensemble mean to be insufficiently strong to meet the criteria of an NAO event with the same lead time. However, the ensemble mean can forecast a more accurate circulation pattern in the North Atlantic sector than the control forecast with the same lead time. Moreover, for ensemble mean forecasts, the failure proportion of NAO− event forecasts is much larger than that of NAO+ event forecasts, which also suggests that the onset of NAO− events is harder to forecast than NAO+ events.

In the following, we try to present our answers to the three questions posed in the introduction.

(1) For the individual onset of an NAO event, operational weather forecasts can predict it three to five days in advance. On average, ECMWF has a skillful forecast time of 3.82 days for NAO+ onset and 4.56 days for NAO− onset derived by the control forecast, which is the longest among the four centers, followed by NCEP, JMA, and then CMA.

(2) From our investigation, the failure proportion for the prediction of NAO− onset is higher than that of NAO+ onset, regardless of whether the control forecast or ensemble forecast is used. This may indicate that the onset of NAO− events is comparatively harder to forecast.

(3) When compared with the control forecast, the ensemble forecasts do not improve the skillful forecast time for the onset of NAO events. However, the ensemble mean can produce a more accurate circulation pattern in the North Atlantic sector compared with the control forecast with the same lead time.

In this work, we have mainly focused on evaluating the forecasting performance with respect to NAO onset, as well as discussing that of the NAO duration. However, during the decay stage of an NAO+ event, a blocking in its downstream sector can be triggered (Luo et al., 2015). Moreover, an NAO− event can sometimes be transformed from an NAO+ event. The above events are possible sources of extreme weather in Eurasia (Luo et al., 2014). But how far in advance can we forecast the decay stage of an NAO event? Can this transformation be predicted in advance? These questions should be investigated in future work.

Acknowledgements. The authors are grateful to the two anonymous reviewers for their helpful comments on this paper. This study was supported by the National Natural Science Foundation of China (Grant Nos. 41230420 and 41775001).

REFERENCES

Barnes, E. A., and D. L. Hartmann, 2010: Dynamical feedbacks and the persistence of the NAO. *J. Atmos. Sci.*, **67**(3), 851–865, https://doi.org/10.1175/2009JAS3193.1.

Benedict, J. J., S. Lee, and S. B. Feldstein, 2004: Synoptic view of the North Atlantic Oscillation. *J. Atmos. Sci.*, **61**(2), 121–144, https://doi.org/10.1175/1520-0469(2004)061<0121:SVOTNA>2.0.CO;2.

Buizza, R., and M. Leutbecher, 2015: The forecast skill horizon. *Quart. J. Roy. Meteor. Soc.*, **141**(693), 3366–3382, https://doi.org/10.1002/qj.2619.

Diao, Y. N., S. P. Xie, and D. H. Luo, 2015: Asymmetry of winter European surface air temperature extremes and the North Atlantic Oscillation. *J. Climate*, **28**(2), 517–530, https://doi.org/10.1175/JCLI-D-13-00642.1.

Feldstein, S. B., 2000: Is interannual zonal mean flow variability simply climate noise? *J. Climate*, **13**(13), 2356–2362, https://doi.org/10.1175/1520-0442(2000)013<2356:IIZMFV>2.0.CO;2.

Feldstein, S. B., 2003: The dynamics of NAO teleconnection pattern growth and decay. *Quart. J. Roy. Meteor. Soc.*, **129**(589), 901–924, https://doi.org/10.1256/qj.02.76.

Hollingsworth, A., K. Arpe, M. Tiedtke, M. Capaldo, and H. Savijärvi, 1980: The performance of a medium-range forecast model in winter–impact of physical parameterizations. *Mon. Wea. Rev.*, **108**(11), 1736–1773, https://doi.org/10.1175/1520-0493(1980)108<1736:TPOAMR>2.0.CO;2.

Hurrell, J. W., 1995: Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, **269**(5224), 676–679, https://doi.org/10.1126/science.269.5224.676.

Jiang, Z. N., X. Wang, and D. H. Wang, 2015: Exploring the phase-strength asymmetry of the North Atlantic Oscillation using conditional nonlinear optimal perturbation. *Adv. Atmos. Sci.*, **32**(5), 671–679, https://doi.org/10.1007/s00376-014-4094-3.

Johansson, Å., 2007: Prediction skill of the NAO and PNA from daily to seasonal time scales. *J. Climate*, **20**(10), 1957–1975, https://doi.org/10.1175/JCLI4072.1.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**(3), 437–472, https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50–year reanalysis: Monthly means CD–ROM and documentation. *Bull. Amer. Meteor. Soc.*, **82**(2), 247–268, https://doi.org/10.1175/1520-0477(2001)082<0247:TNNYRM>2.3.CO;2.

Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102**(6), 409–418, https://doi.org/10.1175/ 1520-0493(1974)102<0409:TSOMCF>2.0.CO;2.

Leutbecher, M., and T. N. Palmer, 2008: Ensemble forecasting. *J. Comput. Phys.*, **227**(7), 3515–3539, https://doi.org/10.1016/j.jcp.2007.02.014.

Li, J. P., and J. X. L. Wang, 2003: A new North Atlantic Oscillation index and its variability. *Adv. Atmos. Sci.*, **20**(5), 661–676, https://doi.org/10.1007/BF02915394.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

Luo, D. H., A. R. Lupo, and H. Wan, 2007: Dynamics of eddy-driven low-frequency dipole modes. Part I: A simple model of North Atlantic Oscillation. *J. Atmos. Sci.*, **64**(1), 3–28, https://doi.org/10.1175/JAS3818.1.

Luo, D. H., Y. Yao, and S. B. Feldstein, 2014: Regime transition of the North Atlantic Oscillation and the extreme cold event over Europe in January-February 2012. *Mon. Wea. Rev.*, **142**(12), 4735–4757, https://doi.org/10.1175/MWR-D-13-00234.1.

Luo, D. H., Y. Yao, A. G. Dai, and S. B. Feldstein, 2015: The positive North Atlantic Oscillation with downstream blocking and Middle East snowstorms: The large-scale environment. *J. Climate*, **28**(16), 6398–6418, https://doi.org/10.1175/JCLI-D-15-0184.1.

Luo, D. H., Y. Q. Xiao, Y. N. Diao, A. G. Dai, C. L. Franzke, and I. Simmonds, 2016: Impact of Ural blocking on winter warm Arctic–cold Eurasian anomalies. Part II: The link to the North Atlantic Oscillation. *J. Climate*, **29**(11), 3949–3971, https://doi.org/10.1175/JCLI-D-15-0612.1.

Marshall, J., and Coauthors, 2001: North Atlantic climate variability: Phenomena, impacts and mechanisms. *International Journal of Climatology*, **21**(15), 1863–1898, https://doi.org/10.1002/joc.693.

Murphy, A. H., and E. S. Epstein, 1989: Skill scores and correlation coefficients in model verification. *Mon. Wea. Rev.*, **117**(3), 572–582, https://doi.org/10.1175/1520-0493(1989)117<0572:SSACCI>2.0.CO;2.

Park, Y. Y., R. Buizza, and M. Leutbecher, 2008: TIGGE: Preliminary results on comparing and combining ensembles. *Quart. J. Roy. Meteor. Soc.*, **134**(637), 2029–2050, https://doi.org/10.1002/qj.334.

Scaife, A. A., C. K. Folland, L. V. Alexander, A. Moberg, and J. R. Knight, 2008: European climate extremes and the North Atlantic Oscillation. *J. Climate*, **21**(1), 72–83, https://doi.org/10.1175/2007JCLI1631.1.

Scaife, A. A., and Coauthors, 2014: Skillful long-range prediction of European and North American winters. *Geophys. Res. Lett.*, **41**(7), 2514–2519, https://doi.org/10.1002/2014GL059637.

Seager, R., Y. Kushnir, J. Nakamura, M. Ting, and N. Naik, 2010: Northern Hemisphere winter snow anomalies: ENSO, NAO and the winter of 2009/10. *Geophys. Res. Lett.*, **37**(14), L14703, https://doi.org/10.1029/2010GL043830.

Song, J., 2016: Understanding anomalous eddy vorticity forcing in North Atlantic Oscillation events. *J. Atmos. Sci.*, **73**(8), 2985–3007, https://doi.org/10.1175/JAS-D-15-0253.1.

Swinbank, R., and Coauthors, 2016: The TIGGE project and its achievements. *Bull. Amer. Meteor. Soc.*, **97**(1), 49–67, https://doi.org/10.1175/BAMS-D-13-00191.1.

Thompson, D. W. J., and J. M. Wallace, 2000: Annular modes in the extratropical circulation. Part I: Month-to-month variability. *J. Climate*, **13**(5), 1000–1016, https://doi.org/10.1175/1520-0442(2000)013<1000:AMITEC>2.0.CO;2.

Thompson, D. W. J., J. M. Wallace, and G. C. Hegerl, 2000: Annular modes in the extratropical circulation. Part II: Trends. *J. Climate*, **13**(5), 1018–1036, https://doi.org/10.1175/1520-0442(2000)013<1018:AMITEC>2.0.CO;2.

Vitart, F., 2014: Evolution of ECMWF sub-seasonal forecast skill scores. *Quart. J. Roy. Meteor. Soc.*, **140**(683), 1889–1899, https://doi.org/10.1002/qj.2256.

Walker, G. T., and E. W. Bliss, 1932: World weather V. *Mem. Roy. Meteor. Soc.*, **4**, 53–84.

Whan, K., and Zwiers, F., 2017: The impact of ENSO and the NAO on extreme winter precipitation in North America in observations and regional climate models. *Climate Dyn.*, **48**(5–6), 1401–1411, https://doi.org/10.1007/s00382-016-3148-x.

Woollings, T., B. Hoskins., M. Blackburn, and P. Berrisford, 2008: A new Rossby wave-breaking interpretation of the North Atlantic Oscillation. *J. Atmos. Sci.*, **65**(2), 609–626, https://doi.org/10.1175/2007JAS2347.1.

Yao, Y., D. H. Luo, A. G. Dai, and S. B. Feldstein, 2016: The positive North Atlantic oscillation with downstream blocking and middle East Snowstorms: Impacts of the North Atlantic Jet. *J. Climate*, **29**(5), 1853–1876, https://doi.org/10.1175/JCLI-D-15-0350.1.

Yiou, P., and M. Nogaj, 2004: Extreme climatic events and weather regimes over the North Atlantic: When and where? *Geophys. Res. Lett.*, **31**(7), L07202, https://doi.org/10.1029/2003GL019119.

Zuo, J. Q., H. L. Ren, W. J. Li, and L. Wang, 2016: Interdecadal variations in the relationship between the winter North Atlantic oscillation and temperature in south-Central China. *J. Climate*, **29**(20), 7477–7493, https://doi.org/10.1175/JCLI-D-15-0873.1.