University of Reading

Seasonal predictability of the winter North Atlantic Oscillation from a jet stream perspective

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Parker, T., Woollings, T., Weisheimer, A., O'Reilly, C., Baker, L. ORCID: https://orcid.org/0000-0003-0738-9488 and Shaffrey, L. ORCID: https://orcid.org/0000-0003-2696-752X (2019) Seasonal predictability of the winter North Atlantic Oscillation from a jet stream perspective. Geophysical Research Letters, 46 (16). pp. 10159-10167. ISSN 1944-8007 doi: 10.1029/2019GL084402 Available at https://centaur.reading.ac.uk/88178/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1029/2019GL084402

Publisher: Wiley

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Geophysical Research Letters

RESEARCH LETTER
10.1029/2019GL084402

Seasonal Predictability of the Winter North Atlantic Oscillation From a Jet Stream Perspective

Tess Parker[1,2], **Tim Woollings**[1], **Antje Weisheimer**[1,3,4], **Chris O'Reilly**[1], **Laura Baker**[5], **and Len Shaffrey**[5]

[1]Department of Physics, University of Oxford, Oxford, UK, [2]Now at School of Earth, Atmosphere and Environment, Monash University, Clayton, Victoria, Australia, [3]European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK, [4]National Centre for Atmospheric Science, Department of Physics, University of Oxford, Oxford, UK, [5]National Centre for Atmospheric Science, University of Reading, Reading, UK

Key Points:
- Multilinear regressions of jet latitude and jet speed are used to investigate prediction skill for the winter NAO
- Data are sourced from ECMWF and UKMO operational seasonal forecast systems and a long seasonal hindcast experiment
- Prediction skill for the winter NAO is due to skill in predicting fast variations in jet latitude

Supporting Information:
- Supporting Information S1

Correspondence to:
T. Parker,
tess.parker@monash.edu

Citation:
Parker, T., Woollings, T., Weisheimer, A., O'Reilly, C. H., Baker, L., & Shaffrey, L. (2019). Seasonal predictability of the winter North Atlantic Oscillation from a jet stream perspective. *Geophysical Research Letters*, *46*, 10,159–10,167. https://doi.org/10.1029/2019GL084402

Received 2 JUL 2019
Accepted 9 AUG 2019
Accepted article online 19 AUG 2019
Published online 30 AUG 2019

©2019. The Authors.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Abstract The winter North Atlantic Oscillation (NAO) has varied on interannual and decadal timescales over the last century, associated with variations in the speed and latitude of the eddy-driven jet stream. This paper uses hindcasts from two operational seasonal forecast systems (the European Centre for Medium-range Weather Forecasts's seasonal forecast system, and the U.K. Met Office global seasonal forecast system) and a century-long atmosphere-only experiment (using the European Centre for Medium-range Weather Forecasts's Integrated Forecasting System model) to relate seasonal prediction skill in the NAO to these aspects of jet variability. This shows that the NAO skill realized so far arises from interannual variations in the jet, largely associated with its latitude rather than speed. There likely remains further potential for predictability on longer, decadal timescales. In the small sample of models analyzed here, improved representation of the structure of jet variability does not translate to enhanced seasonal forecast skill.

1. Introduction

The leading mode of variability in the large-scale circulation over the North Atlantic in winter is the North Atlantic Oscillation (NAO), which strongly impacts the weather and climate in the Euro-Atlantic sector. Seasonal hindcasts performed using current-generation dynamical forecast models have demonstrated encouraging skill in forecasting the winter NAO (Derome et al., 2005; Dunstone et al., 2016; Müller et al., 2005; Palmer et al., 2004; Rodwell & Doblas-Reyes, 2006; Scaife et al., 2014) and the related Arctic Oscillation (Riddle et al., 2013).

Interestingly, the level of skill itself appears to vary on decadal timescales (see, e.g., Müller et al., 2005; Shi et al., 2015). To investigate this variation, an ensemble of atmospheric seasonal hindcast experiments using the European Centre for Medium-range Weather Forecasts (ECMWF)'s Integrated Forecasting System model for the period 1900–2009 was created to explore the skill in predicting the winter NAO over a long period (Weisheimer et al., 2017). This study found interdecadal variations in predictive skill (with larger skill in recent decades than for decades in the mid-twentieth century), which appear to be physically based, as the skill covaries with the characteristics of the general circulation itself (see also O'Reilly et al., 2017).

While the NAO is most often defined in terms of variations in sea level pressure in the Northern Hemisphere, the underlying dynamics are related to variability of the Atlantic eddy-driven jet stream associated with eddy-mean flow interactions (Thompson et al., 2013). Specifically, the NAO reflects changes in both the speed and latitude of the jet stream. These two features are linearly independent in their interannual variability (Woollings et al., 2014) and likely sensitive to different dynamical drivers (Baker et al., 2018). Jet latitude and jet speed therefore provide two potentially independent pathways for predictable signals to emerge in the NAO. These may also occur on different timescales, since over the last century the NAO has been more strongly affected by the jet latitude on interannual timescales but by the speed on decadal timescales (Woollings et al., 2015). An additional complexity, however, is that the decadal variations in jet speed can modulate the amount of variability in latitude on shorter timescales, with decades where the jet was strong and steady being interspersed with decades where the jet was weak and variable (Woollings et al., 2017).

Given the weak but significant correlation of 0.31 (p value <0.001) between the observed and predicted winter NAO found in the ensemble of long seasonal hindcasts by Weisheimer et al. (2017), the current work aims to examine possible sources of this predictive skill. In particular, this study seeks to determine whether the skill arises from skill in the prediction of the latitude of the eddy-driven North Atlantic jet, or from prediction of the jet speed, or a combination of the two. Furthermore, the current research aims to determine whether the experimental seasonal hindcasts exhibit the same relationship between the NAO and jet latitude and speed as is observed in a reanalysis data set. The analysis is also applied over a shorter hindcast period to two operational models: the fifth generation ECMWF and U.K. Met Office global seasonal forecast systems.

2. Data and Methodology

2.1. Data

Reanalysis data are taken from the ECMWF twentieth century reanalysis (ERA-20C; Poli et al., 2016). This covers the period from 1900–2010 and is generated by assimilating only surface pressure and marine wind observations. The behavior of the jet is similar in the ERA-20C and the National Oceanic and Atmospheric Administration Twentieth Century Reanalysis (20CR; Compo et al., 2011), the other existing long reanalysis data set (Woollings et al., 2017). For the shorter period covered by the operational model re-forecasts, reanalysis data are taken from the Interim ECMWF Reanalysis (ERA-Interim, herein ERAI; Dee et al., 2011).

A long atmospheric seasonal hindcast experiment for boreal winters from 1900–2009 (ASF-20C) was performed by Weisheimer et al. (2017) using the atmospheric component of the ECMWF's Integrated Forecasting System model. The hindcasts were carried out with an atmosphere-only model, prescribing observed sea surface temperatures (SSTs) as the lower boundary. Feedbacks from the atmosphere onto the SSTs are neglected, and the model assumes perfect atmospheric forcing from the SSTs. The global atmospheric seasonal forecast data set consists of 51 ensemble members, with forecasts initialized on 1 November of each year using ERA-20C, and run to the end of February of the following year. Boreal winter seasons of December to February are labeled according to the year in which the forecasts were initialized.

We also analyze hindcasts from the more recent, fifth generation of the ECMWF seasonal forecasting system (SEAS5; Johnson et al., 2018). SEAS5 includes updated atmospheric and ocean models and adds an interactive sea ice model. This is a slightly different model version from that used in the ASF-20C experiment, and the horizontal resolution is about doubled. Data from the 25-member ensemble initialized on 1 November, and extending 7 months into the future, are available for the years 1981–2016.

For comparison, hindcast data from the fifth generation U.K. Met Office operational global seasonal forecast system (GloSea5; MacLachlan et al., 2015) is also considered. This model has approximately 50 km atmospheric resolution and includes an ocean model with interactive sea-ice. The hindcasts cover the winter period for the years 1993–2015, and here consist of 21 ensemble members, with seven members initialized on each of 25 October, 1 November, and 9 November every year.

For both operational seasonal hindcasts, data from the common period of 1993–2009 are used so that the impact of the different models on the results is made from a basis in which as much data as possible remains fixed, making any attributions clearer.

2.2. Methods

The latitude and speed of the eddy-driven jet is calculated following Woollings et al. (2010). Here, the daily mean zonal wind at 850 hPa is zonally averaged over a longitudinal sector of 0–60°W over the North Atlantic, and neglecting winds poleward of 75° and equatorward of 15°. A 9-day running mean of the resulting field is then calculated, but the data are not low-pass filtered as in Woollings et al. (2010) since the seasonal forecast data do not cover the full calendar year. The jet speed is then defined as the maximum zonally-averaged westerly wind speed for each day, using the 9-day running mean centered on that day, and the jet latitude as the latitude at which this maximum is identified. Seasonal means (December to February) are calculated as simple averages for each year, without the Fourier filtering applied by Woollings et al. (2010) and thus no deseasonalizing of the daily data is performed here. As stated by the authors, results are not qualitatively different if the low-pass filtering is omitted or the winds at one isobaric level only are used.

Statistical significance of correlations is estimated using a t test based on an effective sample size:

$$N_{\text{eff}} = N\left(\frac{1-\rho_{1a}\rho_{1b}}{1+\rho_{1a}\rho_{1b}}\right), \tag{1}$$

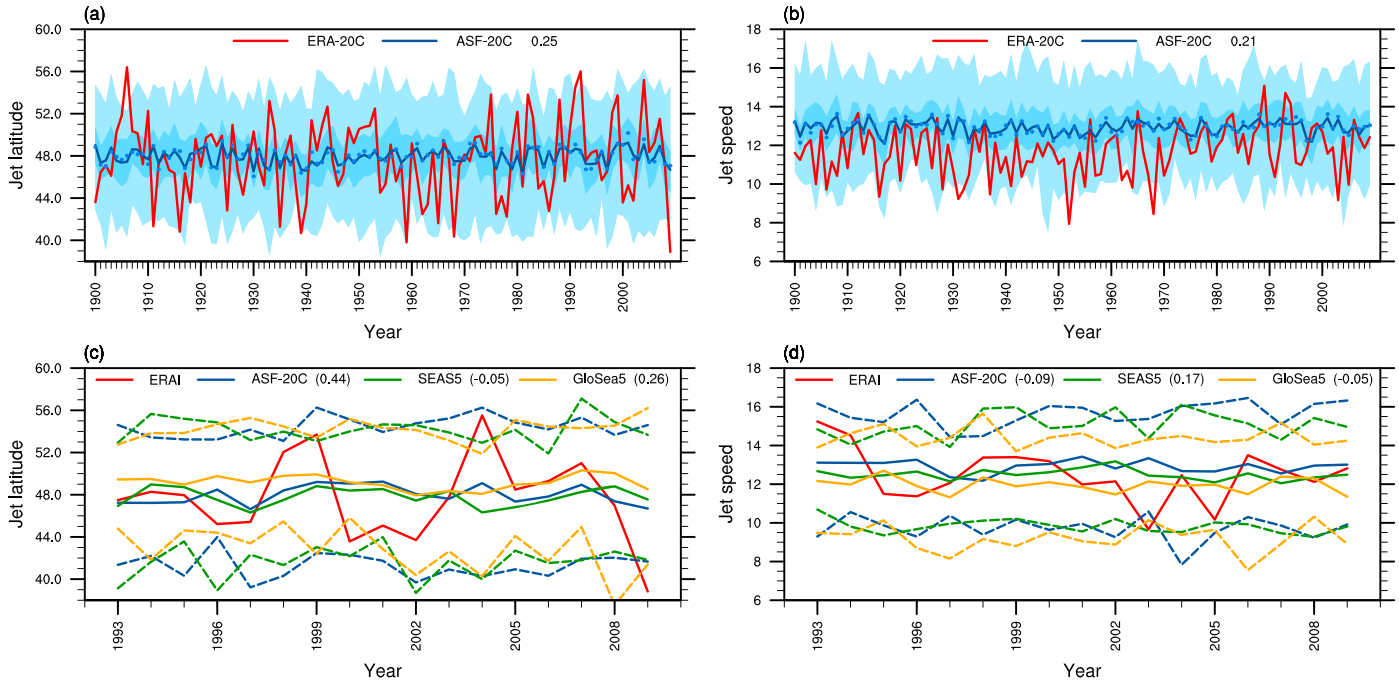

Figure 1. Seasonal mean DJF 1900–2009 for (a) jet latitude and (b) jet speed for ERA-20C (red line) and ASF-20C ensemble mean (blue line), median (blue dots), interquartile range (dark blue shading), and spread (light blue shading). Common period DJF 1993–2009 for (c) jet latitude and (d) jet speed for ERAI (red), ASF-20C (blue), SEAS5 (green), and GloSea5 (orange). Solid lines are ensemble means and dashed lines indicate ensemble spreads. Correlations given in legend (not statistically significant at 95% level in brackets). DJF = December–February; ECMWF = European Centre for Medium-range Weather Forecasts; ERA-20C = ECMWF twentieth century reanalysis; ASF-20C = atmospheric seasonal forecasts of the twentieth century; SEAS5 = fifth generation of the ECMWF seasonal forecasting system; ERAI = Interim ECMWF Reanalysis; GloSea5 = fifth generation U.K. Met Office operational global seasonal forecast system.

where N is the actual sample size and ρ_{1a} and ρ_{1b} are the lag-1 autocorrelation coefficients of the two time series (Wilks, 2011).

Here, the reference NAO index is defined as the principal component (PC) of the leading empirical orthogonal function (EOF) of 500-hPa geopoential height anomalies in the Atlantic sector (90°W to 30°E, 30–90°N) from the ERA-20C reanalysis data (Weisheimer et al., 2017). This index is in very good agreement with other estimates of the NAO, correlating with the PC-based sea level pressure index at 0.96 (Hurrell et al., 2003). The same method is applied to the ERAI data to obtain a reference NAO index over the shorter analysis period. The ASF-20C forecast NAO index is obtained by projecting the individual ensemble members onto the ERA-20C reference EOF (Weisheimer et al., 2017), and the same projection method is applied to the SEAS5 and GloSea5 ensembles with respect to the ERAI reference EOF.

3. Results

3.1. Seasonal Jet Analyses

The time series of seasonal mean jet latitude and speed for the various forecast ensembles versus the reanalyses are shown in Figure 1. The correlation skills of the ASF-20C ensemble mean jet latitude and speed with ERA-20C (Figures 1a and 1b) are both statistically significant, although very low (0.25 and 0.21, respectively). The verification data for jet latitude for all but 13 years and for speed for all but 10 years are within the model spread. In general, the ASF-20C hindcasts exhibit a positive bias in jet strength, with the median and mean jet speed greater than the reanalysis seasonal mean for the majority of the forecast period. The results in Woollings et al. (2017) suggest that the mean state bias in jet strength may impact jet variability in models, and this is investigated in section 3.3.

In order to assess the impact of different models on the results, we now compare the ASF-20C, SEAS5, and GloSea5 ensemble means with ERAI over the common period of December 1993 to February 2010, or 17 winter seasons (Figures 1c and 1d). In this way the different models are compared with the same reanalysis over the same period, and thus, no additional complexity is introduced as a result of, for example,

Table 1
Correlations (Statistically Significant at 95% Level in Bold) for Jet Latitude and Speed for the "Fast" and "Slow" Components, Between ASF-20C and ERA-20C

ASF-20C/ERA-20C	Latitude	Speed
	11-year running mean	
Fast component	**0.26**	0.18
Slow component	0.09	0.41
	31-year running mean	
Fast component	**0.34**	**0.26**
Slow component	−0.11	0.78

Note. Components show correlation between the observed NAO and the reconstructed model NAO when that regression element is omitted. NAO = North Atlantic Oscillation; ASF-20C = atmospheric seasonal forecasts of the twentieth century; ECMWF = European Centre for Medium-range Weather Forecasts; ERA-20C = ECMWF twentieth century reanalysis.

longer-term variability, or the sampling issue associated with the uncertainties related to the exact hindcast length. For the ASF-20C ensemble mean jet latitude the correlation with ERAI has increased compared with that for the 110-year experiment with ERA-20C. This suggests that the recent period was potentially more predictable than the century as a whole, as discussed in section 1. This higher correlation is no longer significant at the 95% level due to the shortness of the period used, but with a p value of 0.08 is close to significant. The other systems show weaker correlations and no significant skill over this period (see legends in Figure 1). The shortness of this period is clearly a limitation in assessing hindcast skill. For completeness, the time series and correlations for the full hindcast periods for SEAS5 and GloSea5 are shown in Figure S1 in the supporting information. Given that there is very little variability in the ensemble means in terms of absolute units, Figure S2 shows the observations and the ensemble means on separate axes.

It is of interest to determine the timescales on which the skill in ASF-20C is realized. An 11-year running mean of the jet latitude and speed is first used to determine that part of the time series that varies on approximately decadal timescales (the "slow" component). Subtracting these slow components from the full time series for jet latitude and speed yields time series that vary on shorter than decadal timescales (the "fast" component). Of these, only the fast component of the jet latitude shows a significant correlation; the fast component of jet speed and the slow components of both jet speed and latitude show no significant correlation between ERA-20C and the ASF-20C ensemble mean, although the correlation for the slow jet speed is relatively high (Table 1). Similarly, removing the 31-year running mean, which represents variations on multidecadal timescales, results in statistically significant correlations for the fast component of both the jet latitude and speed. The slowly varying components are, in contrast, not statistically significantly correlated with the observations, although the correlation for slow jet speed is relatively high (Table 1). In summary, for both of the time windows used here, any skill in predicting the jet indices comes from the fast component of the variance.

3.2. Link to the NAO

For the period 1900–2009, the ASF-20C ensemble mean anomaly correlation coefficient for the NAO index is 0.31 and is highly statistically significant (Weisheimer et al., 2017). While still low, this correlation is higher than those for the jet indices, suggesting that the model may be achieving skill in the NAO from both of the independent jet features. To test this, a multiple linear regression is performed to determine the relationship between the NAO and the jet latitude and speed. The intention is to attribute the predictable signal in the NAO to either jet latitude or speed, or both. As this is a simple linear approach, it ignores the non-stationarity in the NAO, which likely has a slightly different spatial pattern on different timescales (Woollings et al., 2015).

The regression coefficients for the ASF-20C ensemble mean NAO are 0.66 for jet latitude and 0.38 for speed, combining to explain 57% of the variability in the NAO. The corresponding coefficients for ERA-20C are 0.74 for jet latitude and 0.35 for speed, explaining 67% of the variability in the reanalysis NAO (Figure 2a). Although these values appear different, Figure 2c reveals that the observed regression values are in fact captured within the range of behavior in the model ensemble. There is clearly a considerable spread in the ensemble regression coefficients, even when using the full 110-year period.

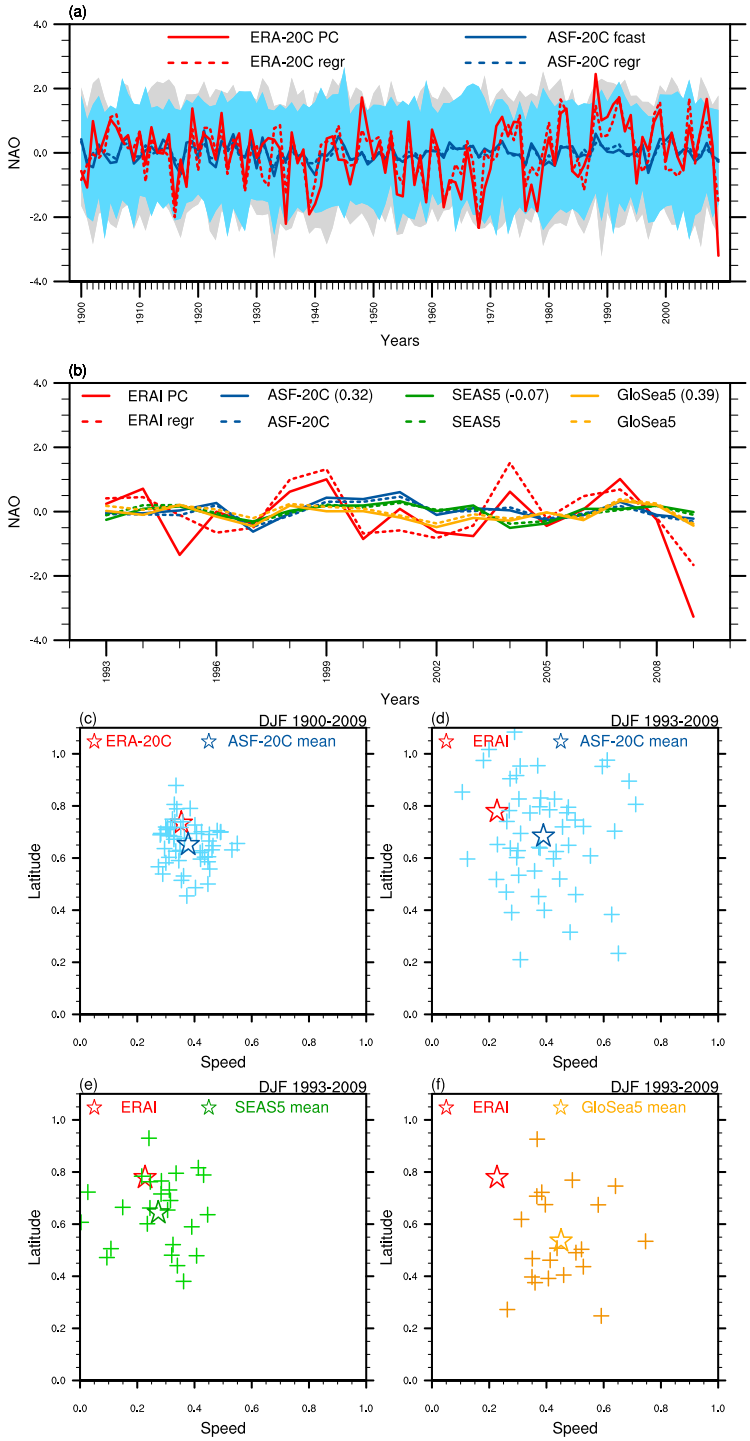

Figure 2. Seasonal NAO index for (a) DJF 1900–2009 for ERA-20C and ASF-20C and (b) DJF 1993–2009 for ERAI and the models. Solid lines are NAO index from PC analysis and dashed lines are reconstructed from regression coefficients. Correlations given in legend (not statistically significant at 95% level in brackets). For (a), gray shading shows original forecast ensemble spread, and blue shading reconstructed index spread. (c–f) Scatter plots of multiple linear regression coefficients of the NAO for jet latitude and speed for (c) ASF-20C and ERA-20C for DJF 1900–2009; and for DJF 1993–2009 with ERAI for (d) ASF-20C, (e) SEAS5, and (f) GloSea5. Red stars denote reanalysis coefficients, colored stars forecast ensemble means, and colored pluses ensemble members. NAO = North Atlantic Oscillation; DJF = December–February; ASF-20C = atmospheric seasonal forecasts of the twentieth century; ERAI = Interim ECMWF Reanalysis; GloSea5 = fifth generation U.K. Met Office operational global seasonal forecast system; ERA-20C = ECMWF twentieth century reanalysis; ECMWF = European Centre for Medium-range Weather Forecasts.

Table 2
Impact of Individual Multilinear Regression Terms on NAO Skill

ASF-20C (reconstructed)/ERA-20C	Latitude	Speed
11-year running mean		
Fast component	0.21 *(−0.18)*	0.43 (0.04)
Slow component	0.40 (0.01)	0.39 (0.00)
31-year running mean		
Fast component	0.24 *(−0.18)*	0.45 (0.03)
Slow component	0.44 (0.02)	0.42 (0.00)

Note. Components show correlation between the observed NAO and the reconstructed model NAO when that regression element is omitted. Figures in brackets give the change in correlation compared to that between observed ERA-20C and reconstructed ASF-20C NAO over the 11-year (0.39) and 31-year (0.42) segments of the time series. NAO = North Atlantic Oscillation; ASF-20C = atmospheric seasonal forecasts of the twentieth century; ECMWF = European Centre for Medium-range Weather Forecasts; ERA-20C = ECMWF twentieth century reanalysis.

For the period 1993–2009, the correlation skill for the NAO index with ERAI is 0.32 for ASF-20C, −0.07 for SEAS5 and 0.39 for GloSea5, but none of these are statistically significant (Figure 2b). Focusing on the shorter, common period, gives an even larger spread between ensemble members, though some interesting features are apparent. First, the ASF-20C ensemble remains consistent with the observations (Figure 2d). The SEAS5 ensemble is also consistent with the observations (Figure 2e), even though there is noticeable weakening of the link between jet speed and the NAO, when compared with ASF-20C. The GloSea5 ensemble appears to show a discrepancy from the observations in that the model NAO is too strongly related to jet speed and too weakly to latitude.

To diagnostically attribute skill in the NAO to the jet features, we use the multiple linear regression approach for the ASF-20C data. For both the 11-year and 31-year running means, jet latitude and speed on both fast and slow timescales are used as the four predictors for the NAO. This gives an estimated time series for the NAO reconstructed only from information of the jet indices. These individual components are removed one by one to determine their contribution to the NAO skill. The full correlation between the observed ERA-20C NAO, and the reconstructed ASF-20C ensemble mean NAO calculated using multiple linear regression, is 0.39 for the 11-year running mean time window, and 0.42 for the 31-year running mean. Table 2 shows the correlations obtained after the component for each of the four predictors is removed from the reconstructed ASF-20C ensemble mean NAO. The clear result is that only the faster components of jet latitude appreciably reduce the correlation skill of the calculated NAO (by approximately 0.18; these changes in skill are shown in italics). The effect on the correlation skill is negligible for the other components, and for faster components of speed the skill actually increases slightly when those coefficients are neglected.

From this simple linear approach using multiple linear regression, it is apparent that in ASF-20C at least, the only source of predictability in the NAO that can be significantly attributed to the jet is that due to fast variations in jet latitude.

3.3. Distributions of Jet Speed and Jet Latitude

We now investigate the models' representation of jet variability in more detail. Probability density functions (PDFs) of jet speed and latitude are shown in Figure 3. For each model ensemble the separate PDFs of latitude and speed are shown, as well as the joint 2-D distribution. The kernel method of Silverman (1981) has been used to estimate the PDFs (e.g., Woollings et al., 2010). The trimodal distribution of jet latitude (Woollings et al., 2010) is reasonably well represented in the models, although the southerly jet regime appears lacking in the GloSea5 forecast ensemble.

Defining the three regimes based on the PDFs in Figure 3 (and in agreement with Masato et al., 2016), the northerly (N) and southerly (S) regimes are defined for jet latitudes above 51°N and below 39°N, respectively. For all three forecast ensembles, the strongest wind speeds occur when the jet is in the central (C) regime, in agreement with Woollings et al. (2010) and Woollings et al. (2017).

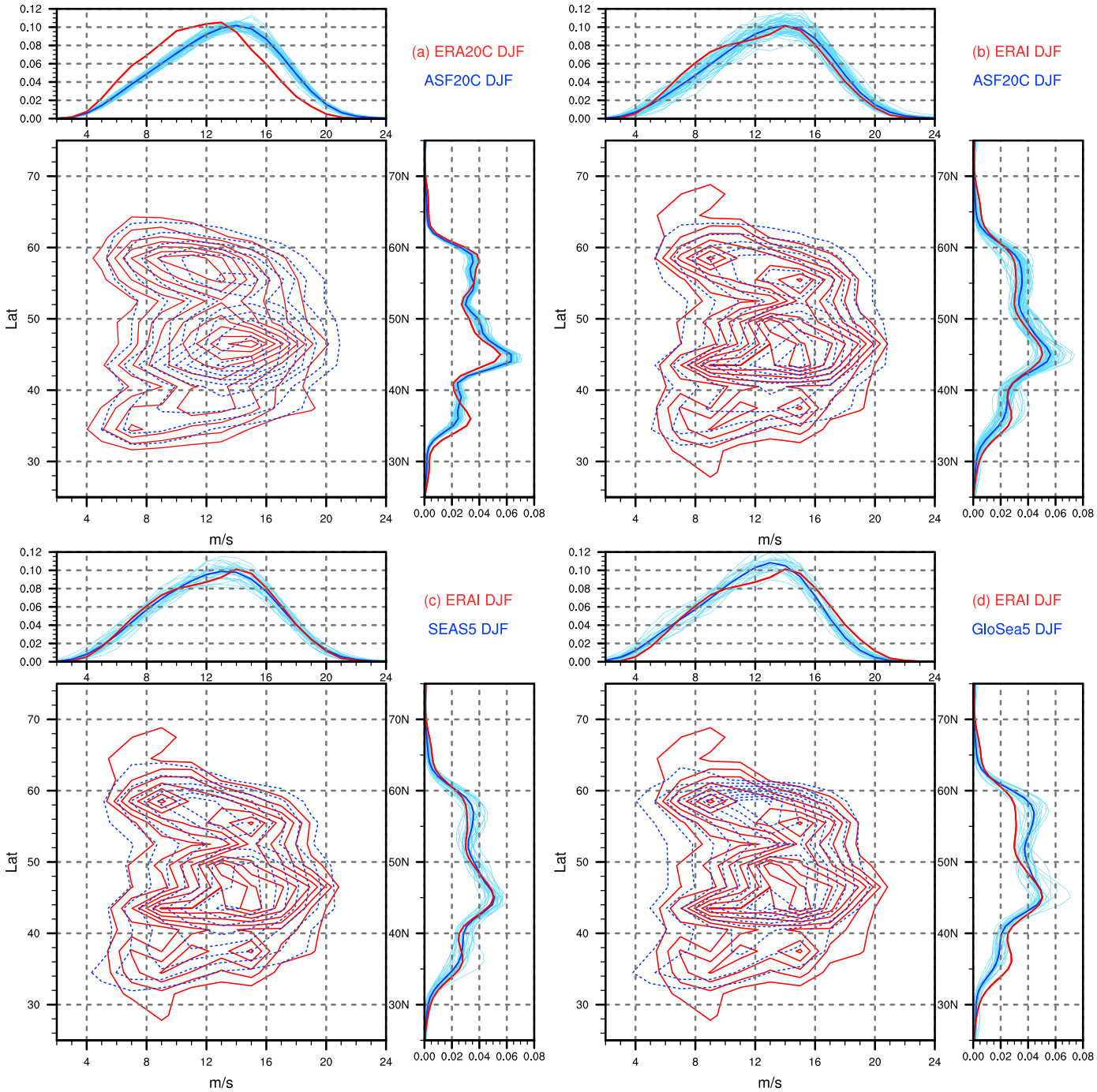

Figure 3. Distributions of jet latitude and speed. Two-dimensional distribution shows speed on the *x* axis (m/s) and latitude on the *y* axis, with data binned every 2 m/s and 3° of latitude. Contours are drawn at intervals of four occurrences. Corresponding one-dimensional distributions for speed at the top and latitude to the right. Reanalysis is shown in red, ensemble mean in blue, and individual ensemble members in light blue. For DJF 1900–2009 for (a) ERA-20C and ASF-20C; and DJF 1993–2009 for ERAI and (b) ASF-20C, (c) SEAS5, and (d) GloSea5. ECMWF = European Centre for Medium-range Weather Forecasts; ERA-20C = ECMWF twentieth century reanalysis; ASF-20C = atmospheric seasonal forecasts of the twentieth century; SEAS5 = fifth generation of the ECMWF seasonal forecasting system; ERAI = Interim ECMWF Reanalysis; GloSea5 = fifth generation U.K. Met Office operational global seasonal forecast system; DJF = December–February.

The distributions are generally reproduced well by all systems, especially in comparison to older models (e.g., Hannachi et al., 2013). The ASF-20C ensemble overestimates the speed of the jet stream by a few meters per second (Figure 1b). It also shows a narrower distribution of jet latitude than the reanalysis, with a greater than observed occurrence of the C regime and a lesser occurrence of the N and S regimes (Figure 3a). At several latitudes, the reanalysis distribution lies outside the ensemble spread. An analysis of the frequency of events lasting 5 days or more for each of the regimes reveals that the number of long-lived events is underrepresented in the model for the N and S regimes, particularly for events lasting more than 20 days (not shown). For the shorter common period (Figure 3b) the ensemble mean distribution for ASF-20C exhibits similar behavior as for the full 110-year hindcast ensemble, although with a weaker discrepancy.

Turning to the other systems, the SEAS5 ensemble (Figure 3c) shows very similar distributions to the ERAI reanalysis, with the verification data lying mostly within the ensemble spread. In contrast, the GloSea5 ensemble (Figure 3d) exhibits biases in the shape of the jet latitude distribution, with the verification data lying outside the ensemble spread at many latitudes.

For the ASF-20C ensemble, a simple histogram of the standard deviation of jet latitude in each ensemble member (not shown) reveals that the observed standard deviation for both seasonal mean and daily values (~3.7° and 8.5°, respectively) lies outside the model distribution (~2.6–3.4° and 7.4–8°, respectively), indicating that the variability in jet latitude is underestimated in the hindcasts. This is consistent with the relationship identified by Woollings et al. (2017), indicating that the bias toward a stronger jet in the model could contribute to an underestimate of the jet latitude variability.

4. Discussion

This paper has examined the contributions from interannual and decadal variations in jet latitude and jet speed to the winter NAO skill in model hindcasts. This has been enabled particularly by the unprecedented size of the seasonal hindcast experiment ASF-20C, which is based on the ECMWF's atmospheric model and covers the boreal winter seasons from 1900–2009 using 51 ensemble members. Operational model hindcasts have also been investigated, but the short hindcast periods of these has limited the conclusions that can be drawn.

The ASF-20C seasonal hindcast experiment has demonstrated significant skill in capturing the fast (i.e., interannual) variations of jet latitude and, to a lesser extent, jet speed. The skill in prediction of the NAO in this experiment has been found to be dominated by the skill in predicting the fast variations in jet latitude. Further improvement in seasonal prediction of the NAO seems possible in this model, as the jet in the experiment is shown to be biased strong, at least over the longer period, and to underestimate the variability in latitude which is the major source of NAO skill.

In contrast, the slow, decadal variations in both latitude and speed are not captured, despite the pronounced multidecadal variability in jet speed in the reanalysis. This suggests that different processes are responsible for NAO variability on the decadal timescales, and hence that there may remain untapped sources of NAO skill, particularly on the longer timescale. The flip side of this, however, is that forecasting skill on the seasonal timescale in a model may not automatically translate to skill on longer timescales.

The models in general have a reasonable simulation of the Atlantic jet and its variability in latitude and speed, albeit with some discrepancies in individual cases. However, there is not a clear link between a model's simulation of the jet mean state characteristics and its skill in a prediction sense. For example, the ECMWF SEAS5 model was found to have the best representation of the jet latitude and speed distributions, but yet was not found to have significant skill in predicting either of these quantities or the NAO. This suggests that the major problem in this case lies not in the internal jet stream dynamics but in the pathways by which predictable drivers affect the jet stream.

It is interesting that, despite higher skill in predicting the NAO than other systems, the U.K. Met Office seasonal prediction system, GloSea5, appears to misrepresent the close link between the NAO and the jet indices. This could be important, since the jet latitude dominates the observed NAO, and has also been identified here as the clearest source of NAO skill. This discrepancy could contribute to the signal-to-noise issue in GloSea5, as predictable signals in jet latitude may not translate to as large a signal in terms of NAO index in the model as in observations.

Acknowledgments
This work was partly supported by NERC Grants NE/L01047X/1 (T. P. and T. W.) and NE/M005887/1 (C. O. R. and T. W.). The seasonal NAO indices and the jet speed and latitude indices used in this study are available in figshare (https://doi.org/10.26180/5bf747920c17f). The authors are grateful to two anonymous reviewers whose insightful comments helped to improve the presentation of our results.

References

Baker, L., Shaffrey, L., Sutton, R. T., Weisheimer, A., & Scaife, A. A. (2018). An intercomparison of skill and overconfidence/underconfidence of the wintertime North Atlantic Oscillation in multimodel seasonal forecasts. *Geophysical Research Letters*, *45*, 7808–7817. https://doi.org/10.1029/2018GL078838

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., et al. (2011). The twentieth century re-analysis project. *Quarterly Journal of the Royal Meteorological Society*, *137*, 1–28. https://doi.org/10.1002/qj.776

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*, 553–597. https://doi.org/10.1002/qj.828

Derome, J., Lin, H., & Brunet, G. (2005). Seasonal forecasting with a simple general circulation model. *Journal of Climate*, *18*, 597–609. https://doi.org/10.1175/JCLI-3289.1

Dunstone, N., Smith, D., Scaife, A. A., Hermanson, L., Eadie, R., Robinson, N., et al. (2016). Skilful predictions of the winter North Atlantic Oscillation one year ahead. *Nature Geoscience*, *9*, 809–814. https://doi.org/10.1038/ngeo2824

Hannachi, A., Barnes, E. A., & Woollings, T. (2013). Behaviour of the winter North Atlantic eddy-driven jet stream in the CMIP3 integrations. *Climate Dynamics*, *41*, 995–1007. https://doi.org/10.1007/s00382-012-1560-4

Hurrell, J. W., Kushnir, Y., Ottersen, G., & Visbeck, M. (2003). An overview of the North Atlantic Oscillation. In J. W. Hurrell, Y. Kushnir, G. Ottersen, & M. Visbeck (Eds.), *The North Atlantic Oscillation: Climate Significance and Environmental Impacts*, Geophysical Monograph Series (pp. 134). Washington, DC: American Geophysical Union. https://doi.org/10.1029/134GM01

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., et al. (2018). SEAS5: The new ECMWF seasonal forecast system. *Geoscientific Model Development Discussions*, *12*, 1087–1117. https://doi.org/10.5194/gmd-2018-228

MacLachlan, C., Arribas, A., Peterson, K. A., Maidens, A., Fereday, D., Scaife, A. A., et al. (2015). Global Seasonal forecast system version 5 (GloSea5): A high resolution seasonal forecast system. *Quarterly Journal of the Royal Meteorological Society*, *141*, 1072–1084. https://doi.org/10.1002/qj.2396

Masato, G., Woollings, T., Williams, K. D., Hoskins, B., & Lee, R. W. (2016). A regime analysis of Atlantic winter jet variability applied to evaluate HadGEM3-GC2. *Quarterly Journal of the Royal Meteorological Society*, *142*, 3162–3170. https://doi.org/10.1002/qj.2897

Müller, W., Appenzeller, C., & Schär, C. (2005). Probabilistic seasonal prediction of the winter North Atlantic Oscillation and its impact on near surface temperature. *Climate Dynamics*, *24*, 213–226. https://doi.org/10.1007/s00382-004-0492-z

O'Reilly, C., Heatley, J., MacLeod, D., Weisheimer, A., Palmer, T., Schaller, N., & Woollings, T. (2017). Variability in seasonal forecast skill of Northern Hemisphere winters over the twentieth century. *Geophysical Research Letters*, *44*, 5729–5738. https://doi.org/10.1002/2017GL073736

Palmer, T. N., Alessandri, A., Andersen, U., Cantelaube, P., Davey, M., Délécluse, P., & Déqué, M. (2004). Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bulletin of the American Meteorological Society*, *85*, 853–872. https://doi.org/10.1175/BAMS-85-6-853

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., et al. (2016). ERA-20C: An atmospheric reanalysis of the twentieth century. *Journal of Climate*, *29*(11), 4083–4097. https://doi.org/10.1175/JCLI-D-15-0556.1

Riddle, E. E., Butler, A. H., Furtado, J. C., Cohen, J. L., & Kumar, A. (2013). CFSv2 ensemble prediction of the wintertime Arctic Oscillation. *Climate Dynamics*, *41*, 1099–1116. https://doi.org/10.1007/s00382-013-1850-5

Rodwell, M. J., & Doblas-Reyes, F. J. (2006). Medium-range, monthly, and seasonal prediction for Europe and the use of forecast information. *Journal of Climate*, *19*(23), 6025–6046. https://doi.org/10.1175/JCLI3944.1

Scaife, A. A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R. T., Dunstone, N., et al. (2014). Skilful long-range prediction of European and North American winters. *Geophysical Research Letters*, *41*, 2514–2519. https://doi.org/10.1002/2014GL059637

Shi, W., Schaller, N., MacLeod, D., Palmer, T. N., & Weisheimer, A. (2015). Impact of hindcast length on estimates of seasonal climate predictability. *Geophysical Research Letters*, *42*, 1554–1559. https://doi.org/10.1002/2014GL062829

Silverman, B. W. (1981). Using kernel density estimates to investigate multimodality. *Journal of the Royal Statistical Society. Series B*, *43*, 97–99.

Thompson, D. W., Lee, S., & Baldwin, M. P. (2013). Atmospheric processes governing the Northern Hemisphere Annular Mode/North Atlantic Oscillation. In J. W. Hurrell, Y. Kushnir, G. Ottersen, & M. Visbeck (Eds.), *The North Atlantic Oscillation: Climatic significance and environmental Impact*, Geophysical Monograph Series (pp. 134). Washington, DC: American Geophysical Union. https://doi.org/10.1029/134GM05

Weisheimer, A., Schaller, N., O'Reilly, C., MacLeod, D. A., & Palmer, T. (2017). Atmospheric seasonal forecasts of the twentieth century: Multi-decadal variability in predictive skill of the winter North Atlantic Oscillation (NAO) and their potential value for extreme event attribution. *Quarterly Journal of the Royal Meteorological Society*, *143*, 917–926. https://doi.org/10.1002/qj.2976

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences*. Oxford: Elsevier Academic Press.

Woollings, T., Barnes, E., Hoskins, B., Kwon, Y.-O., Lee, R., Li, C., et al. (2017). Daily to decadal modulation of jet variability. *Journal of Climate*, *31*, 1297–1314. https://doi.org/10.1175/JCLI-D-17-0286.1

Woollings, T., Czuchnicki, C., & Franzke, C. (2014). Twentieth century North Atlantic jet variability. *Quarterly Journal of the Royal Meteorological Society*, *140*, 783–791. https://doi.org/10.1002/qj.2197

Woollings, T., Franzke, C., Hodson, D., Dong, B., Barnes, E., Raible, C., & Pinto, J. (2015). Contrasting interannual and multidecadal NAO variability. *Climate Dynamics*, *45*, 539–556. https://doi.org/10.1007/s00382-014-2237-y

Woollings, T., Hannachi, A., & Hoskins, B. (2010). Variability of the North Atlantic eddy-driven jet stream. *Quarterly Journal of the Royal Meteorological Society*, *136*, 856–868. https://doi.org/10.1002/qj.625