RESEARCH LETTER

Key Points:
- The winter NAO can be skilfully predicted months ahead
- The signal-to-noise ratio of the predictable signal is anomalously low
- Predictions of the risk of regional winter extremes are possible

Skillful long-range prediction of European and North American winters

A. A. Scaife[1], A. Arribas[1], E. Blockley[1], A. Brookshaw[1], R. T. Clark[1], N. Dunstone[1], R. Eade[1], D. Fereday[1], C. K. Folland[1,2], M. Gordon[1], L. Hermanson[1,3], J. R. Knight[1], D. J. Lea[1], C. MacLachlan[1], A. Maidens[1], M. Martin[1], A. K. Peterson[1], D. Smith[1], M. Vellinga[1], E. Wallace[1], J. Waters[1], and A. Williams[1]

[1]Met Office Hadley Centre, Exeter, UK, [2]Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden, [3]Willis Research Network

Abstract Until recently, long-range forecast systems showed only modest levels of skill in predicting surface winter climate around the Atlantic Basin and associated fluctuations in the North Atlantic Oscillation at seasonal lead times. Here we use a new forecast system to assess seasonal predictability of winter North Atlantic climate. We demonstrate that key aspects of European and North American winter climate and the surface North Atlantic Oscillation are highly predictable months ahead. We demonstrate high levels of prediction skill in retrospective forecasts of the surface North Atlantic Oscillation, winter storminess, near-surface temperature, and wind speed, all of which have high value for planning and adaptation to extreme winter conditions. Analysis of forecast ensembles suggests that while useful levels of seasonal forecast skill have now been achieved, key sources of predictability are still only partially represented and there is further untapped predictability.

1. Introduction

Despite recent advances in weather and climate forecasting, skillful predictions of year to year fluctuations in winter North Atlantic Oscillation [*Walker and Bliss*, 1932] and associated changes in weather at lead times of months have until recently been elusive [*Johansson*, 2007; *Kim et al.*, 2012; *Smith et al.*, 2012]. This is because climate models have shown little extratropical atmospheric circulation response to slowly varying components of the climate system such as the ocean [*Kushnir et al.*, 2006], which might otherwise provide long-range predictability. As a result, while many state-of-the-art seasonal forecast systems show significant predictability for tropical climate, only low forecast skill is generally found in the extratropics [*Arribas et al.*, 2011; *Kim et al.*, 2012]. This has led to the conclusion that little predictability may exist for key extratropical events such as extreme winters [*Jung et al.*, 2011]. However, climate models are imperfect and predictability could well be underrepresented. Indeed, past forecast systems have occasionally shown signs of skill in extratropical circulation [*Palmer et al.*, 2004; *Müller et al.*, 2005], and encouraging levels of skill for the Arctic Oscillation were recently reported by *Riddle et al.* [2013]. An improvement in long-range forecasting of the extratropics would generate enormous benefit to society as it would allow planning in highly populated regions of the Northern Hemisphere for the risk of severe winter weather including winter wind storms [*Renggli et al.*, 2011] and disruption to transport [*Palin et al.*, 2013] networks for example.

The single most important factor for year to year fluctuations in the seasonal climate around the Atlantic Basin is the state of the North Atlantic Oscillation (NAO) and its hemispheric equivalent, the Arctic Oscillation. Year to year variability in the NAO describes the state of the Atlantic jet stream and is directly related to near-surface winds and hence winter temperatures (through advection) across North America, Europe, and other regions around the Atlantic Basin. We present estimates of the predictability of the surface NAO and winter climate from the Met Office seasonal forecast system Global Seasonal forecast System 5 (GloSea5) which has high ocean resolution, a comprehensive representation of the stratosphere, and interactive sea ice physics, all of which mediate predictable teleconnections to the North Atlantic as shown below.

2. Predictability of the North Atlantic Oscillation

The forecasts used here were produced using the Met Office Global Seasonal forecast System 5 (GloSea5). The climate model at the core of this forecast system is Hadley Centre Global Environmental Model version 3
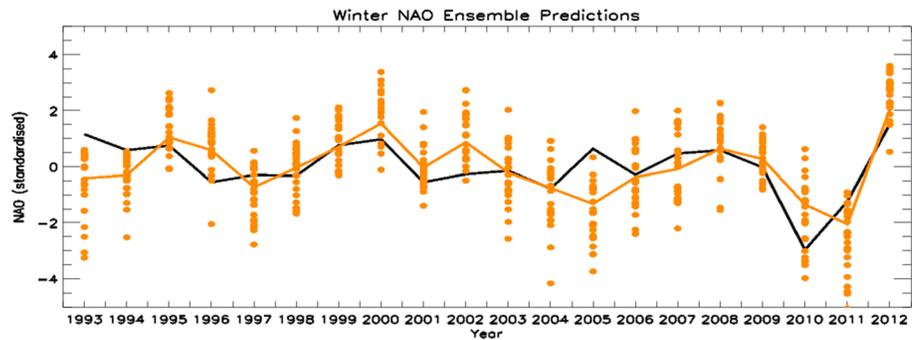
Figure 1. Predictability of the winter North Atlantic Oscillation. The NAO in observations (black line), ensemble mean forecasts (orange line), and individual ensemble members (orange dots) in winter (December to February (DJF)) hindcasts. The NAO is measured as the sea level pressure difference between Iceland and the Azores, but the skill is insensitive to the precise definition as large-scale patterns are frequently well captured (Figure S1). Observations, ensemble mean, and ensemble members are normalized by their respective standard deviations. Anomalies are for December to February, and forecasts were initialized from dates centered on 1 November. The correlation score of 0.62 is significant at the 99% level according to a *t* test and allowing for the small lagged autocorrelation in forecasts and observations.

with atmospheric resolution of 0.83° longitude by 0.55° latitude, 85 quasi-horizontal atmospheric levels, and an upper boundary at 85 km near the mesopause. The ocean resolution is 0.25° globally in both latitude and longitude with 75 quasi-horizontal levels. This resolution is necessary to reduce key biases in the ocean and atmosphere and give a realistic winter blocking climatology in the model [*Scaife et al.*, 2011]. A 24-member ensemble of forecasts was run for each winter in the period 1993 to 2012 with lagged start dates centered on 1 November (25 October, 1 November, and 9 November) and eight members initialized on each of the three start dates. Members from the same start date differ only by stochastic physics [*Arribas et al.*, 2011]. Initial atmospheric and land surface data were taken from ECMWF Re-Analysis (ERA)-Interim observational reanalyses, and initial conditions for the global ocean and sea ice concentration were from the Forecasting Ocean Assimilation Model (FOAM) system [*Blockley et al.*, 2013]. This configuration allows very skillful predictions of various slowly varying components of the climate system to be made for the coming winter (Table S1).

Figure 1 shows the skill of predicting the year to year fluctuations in the winter surface NAO (difference in sea level pressure between Iceland and Azores) at lead times of 1 to 4 months, well beyond weather forecast time scales. The resulting correlation coefficient between the ensemble average of 24 forecast members per winter and the observed surface NAO is 0.62 in GloSea5. This is statistically significant at the 99% level of confidence (using a *t* test and allowing for the small lagged autocorrelation in model and observations). It confirms potential predictability hinted at in statistical studies [*Folland et al.*, 2012; *Cohen and Jones*, 2011] and atmospheric simulations with prescribed ocean conditions [*Rodwell et al.*, 1999; *Mehta et al.*, 2000; *Bretherton and Battisti*, 2000] and supports recent results for the Arctic Oscillation [*Riddle et al.*, 2013], using a seasonal forecast system based on first physical principles. The value achieved here greatly exceeds persistence forecast skill (0.15) and suggests that useful levels of seasonal forecast skill for the surface NAO can be achieved in operational dynamical forecast systems. Our result is also insensitive to the details of the model or the hindcast. For example, a repeat hindcast using a new dynamical core [*Walters et al.*, 2013] resulted in a similar correlation score of 0.6, as did removal of individual strong and predictable NAO winters such as 2009/2010 [*Fereday et al.*, 2012] or the poorly predicted winter of 2004/2005 (Figure S1 in the supporting information). None of these changes reduces the significance of the correlation below the 95% level. Note also that the forecast skill in our system arises largely from interannual variability rather than trends or low-frequency variability, as differences in the surface NAO from 1 year to the next are skillfully predicted with a correlation coefficient of 0.46 which is also significant at the 95% level, particularly for years which project strongly on to the NAO (Figure S1). As a further check we also calculated probabilistic skill and reliability scores. Relative operating characteristic scores [*World Meteorological Organization (WMO)*, 1992] for lower tercile winter temperatures are 0.70 for the Northern Hemisphere (20–90 N), 0.77 for North America (50–165 W, 10–85 N), and 0.65 for Europe (30 W–40E, 30–80 N), with high levels of reliability for all of these regions.
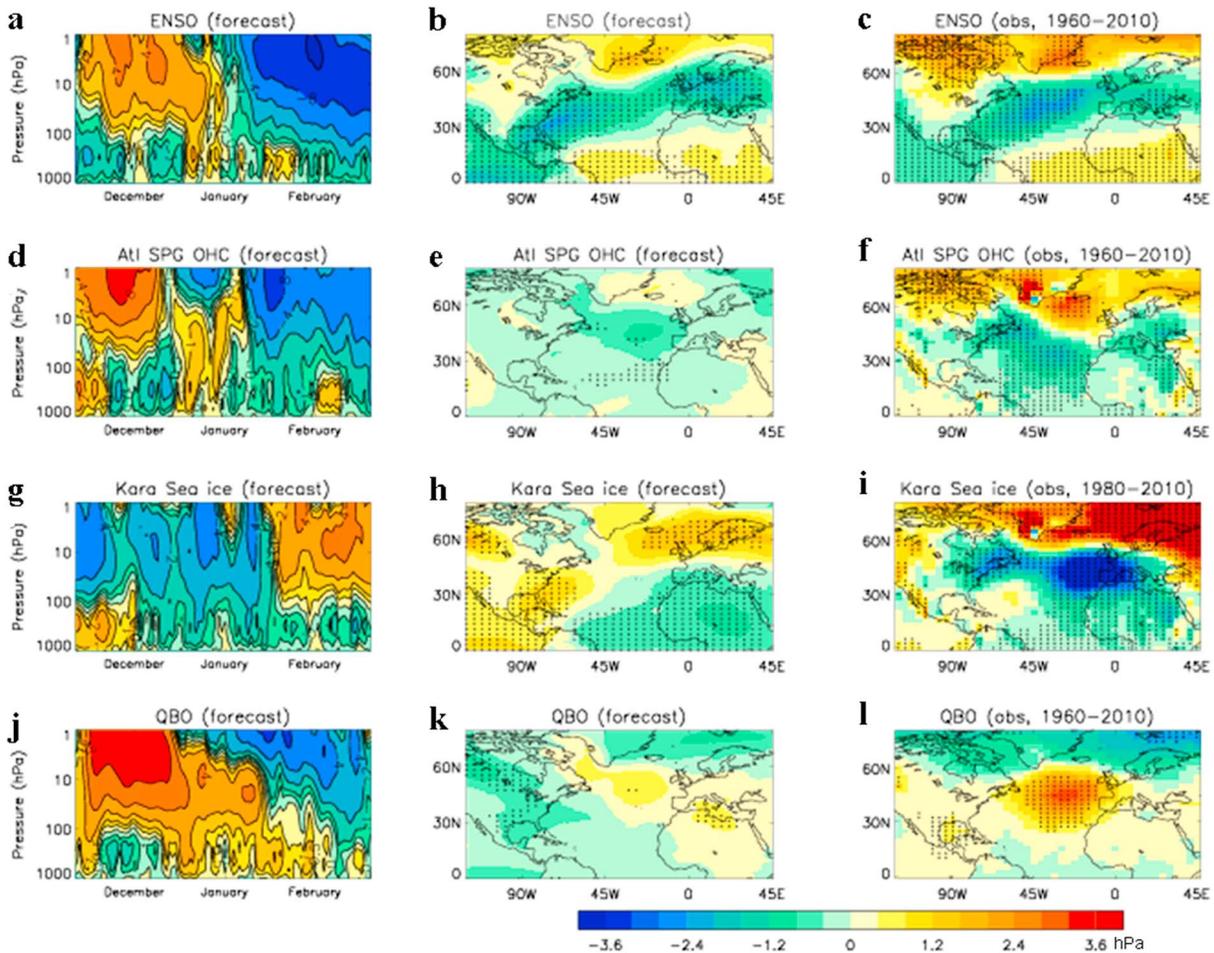
Figure 2. Predictable teleconnections to the North Atlantic. (a, d, g, and j) Winter (DJF) jet stream winds (m s^{-1}) and (b, e, h, and k) sea level pressure (hPa) signals in forecasts and (c, f, i, and l) observations, from the El Niño–Southern Oscillation (ENSO) shown in Figures 2a–2c, North Atlantic Ocean heat content in the subpolar gyre shown in Figures 2d–2f, sea ice area in the Kara Sea shown in Figures 2g–2i, and the quasi-biennial oscillation shown in Figures 2i–2k. Resulting winter differences are shown between upper and lower terciles of each factor measured in November. All factors are highly predictable (Table S1). Jet stream winds are composite daily zonal wind anomalies at 60 N for the Atlantic sector (90°W–60°E). Pattern correlations between predicted and observed sea level pressure patterns were 0.83, 0.14, 0.44, and 0.45, respectively, and significance at the 90% level is shown by hatching.

3. Sources of Predictability

While we cannot assess causality without additional experiments, which are beyond the scope of this study, digging deeper into the forecasts reveals several potential sources of predictable signals (Table S1 in the supporting information). One source of predictability originates in the tropical Pacific. Previous studies have shown that the El Niño–Southern Oscillation can drive interannual variations in the NAO [*Brönnimann et al.*, 2007] and hence Atlantic and European winter climate via the stratosphere [*Bell et al.*, 2009]. Figures 2b and 2c confirm that this teleconnection to the tropical Pacific is active in our experiments, with forecasts initialized in El Niño/La Niña conditions in November tending to be followed by negative/positive NAO conditions in winter. Established mechanisms [*Bell et al.*, 2009] operate in the forecasts, with deep easterly anomalies occurring in the extratropical jet stream after descending from the stratosphere in midwinter (Figure 2a).

Previous studies also identify precursors to the NAO in North Atlantic Ocean temperatures [*Rodwell et al.*, 1999; *Frankignoul*, 1985; *Rodwell and Folland*, 2002]. By selecting forecasts in years with a warm or cold north Atlantic subpolar gyre in November, we can examine the resulting winter signal in the atmospheric circulation. Forecasts starting from cold/warm North Atlantic states also result in winter predictions with more positive/negative NAO (Figures 2d–2f), although pattern correlation is low in this case. Note that although the ENSO
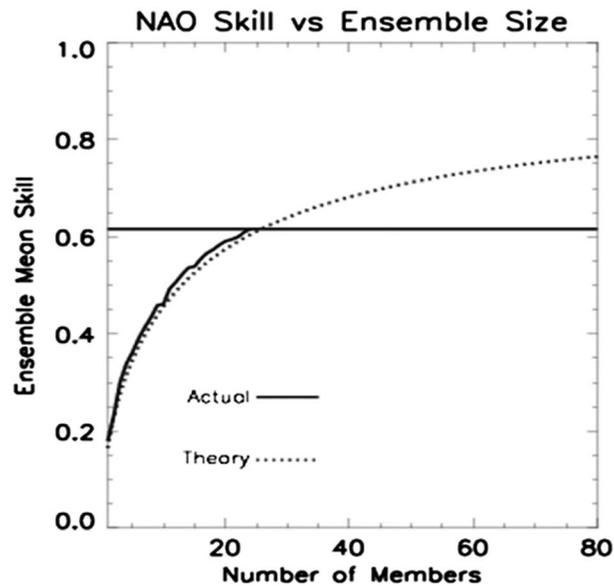
NAO Skill vs Ensemble Size

Figure 3. Ensemble size and prediction skill. The increase of correlation skill score of ensemble mean forecasts of the NAO with increasing ensemble size, ranging from one member to 24 members (solid curve). The dotted curve shows a simple theoretical fit [*Murphy et al.*, 1990] to the increase in skill with ensemble size. The horizontal line shows the full 24-member skill realized in the forecast.

signal is of reasonable strength, in many of the cases of predictability in Figure 2, forecasts show evidence of the same mechanisms and patterns operating as in the real world but with a weaker signal; we return to this later.

Our third teleconnection to the NAO arises from the initialization of Arctic sea ice, particularly in the Kara Sea to the north of Europe. Interannual variability of sea ice and hence surface temperature is large here and has previously been connected to the generation of large-scale circulation anomalies [*Cohen and Jones*, 2011; *Yang and Christensen*, 2012]. Figures 2h and 2i show the association between sea ice anomalies in this region in November and the subsequent winter circulation in forecasts and observations. As identified in other studies [*Yang and Christensen*, 2012] low/high sea ice concentrations in the Kara Sea in November precede negative/positive NAO anomalies, with anomalous pressure gradients over northernmost Europe and the East Atlantic.

Our final teleconnection to the NAO arises from the quasi-biennial oscillation (QBO) in the tropical lower stratosphere. Interannual variability between westerly and easterly phases of the QBO has long been known to influence the troposphere in the Atlantic sector [*Ebdon*, 1975] in the sense shown in Figure 2, with westerly QBO being associated with a stronger extratropical jet, particularly in early winter [*Pascoe et al.*, 2005]. Figures 2j–2l again show a similar but weaker forecast signal.

4. Anomalous Signal-to-Noise Ratio

Despite the reproduction of known teleconnection patterns, it is clear from Figure 2 that the amplitude of signals in the forecasts is smaller than in observations. Similarly, while the ensemble mean signal in these forecasts correlates well with the observed NAO (corr = 0.62), the signal-to-noise ratio defined as the ensemble *mean* standard deviation divided by the total ensemble *member* standard deviation [*Kumar*, 2009] is low ($s = 0.2$). Despite this, the variability of the NAO from individual forecast members agrees well with observed variability and is around 8 hPa, so it is only the ensemble mean signal and not the variability of ensemble members that is too small. This presents something of a puzzle because for a perfect forecast system the expected signal-to-noise ratio and the correlation are directly related. Indeed, given the ensemble mean forecast correlation of 0.62, we would expect a signal-to-noise ratio much higher than found here [*Kumar*, 2009] (Figure 2). The answer lies in the weak signals in the forecast system (Figure 2) which result in the correlation between individual forecast members and the observations being several times higher than correlations between pairs of forecasts, a result similar to that found in atmosphere only experiments [*Mehta et al.*, 2000]. In summary, individual forecast members contain weaker predictable signals than the observations.

Despite the high skill in predicting extratropical winter climate, low signal-to-noise ratios mean that large forecast ensembles are still needed to achieve a given skill. This is illustrated by systematically sampling subsets of forecasts from the full ensemble of 24 members (Figure 3). Ensemble mean prediction skill for the NAO increases with the number of forecast members and is still increasing, albeit more slowly, as the full size of our ensemble is approached. Just this scenario has been previously examined from a statistical
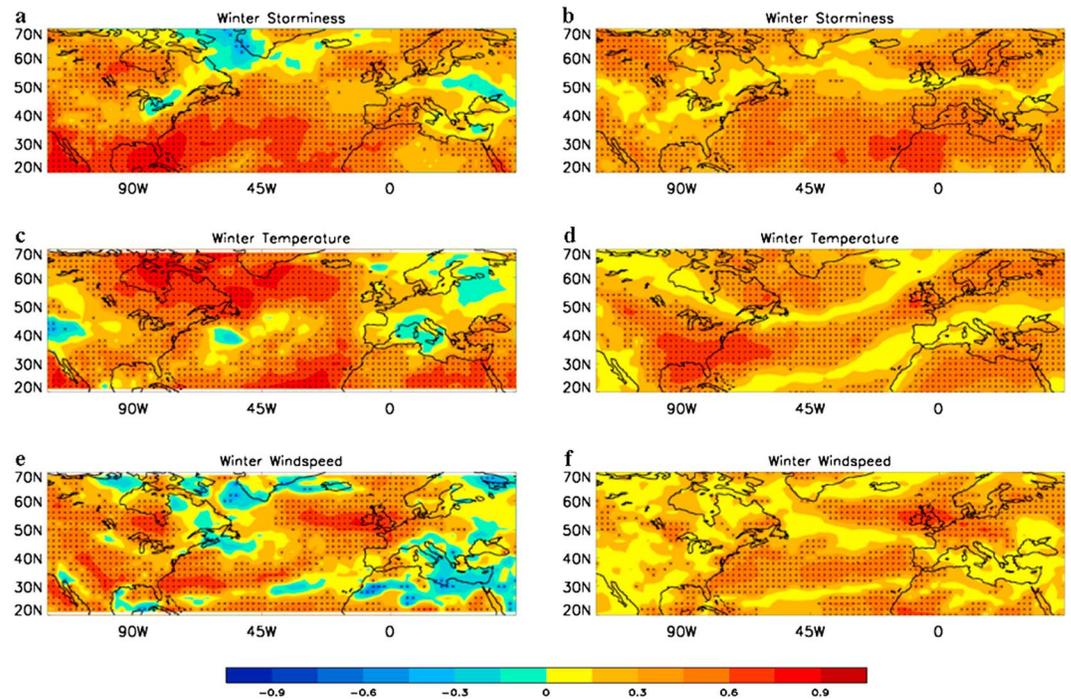
Figure 4. Forecast skill of surface winter weather conditions. Correlation score for (a and b) the frequency of winter storms (measured by tenth percentile daily sea level pressure minima), (c and d) winter mean temperature, and (e and f) winter mean wind speed (mean of daily values at 10 m altitude). Observed storminess, temperature, and wind speed are from the ERA-Interim reanalysis [*Dee et al.*, 2011] and are correlated with forecast storminess, temperature, and wind speed on the left and with forecast NAO on the right. Hatching indicates values above 90% statistical significance according to a student *t* test and allowing for autocorrelation. Scores on the right show the modulus of the correlation.

viewpoint [*Murphy*, 1990], and the skill limit of an infinite-sized ensemble depends only on the average correlation between pairs of forecast members and the average correlation between forecast members and observations. This limit exceeds 0.8 for the NAO in our system. Along with improvements in the modeled signal strength, increased ensemble size could therefore lead to further increases in seasonal forecast skill for the extratropics.

5. Discussion: Implications for Regional Prediction

The NAO governs many aspects of European and North American winter weather, and predictability of the NAO therefore leads to similarly skillful predictions of surface winter climate (Figure 4). For example, the risk of damaging winter wind storms is highly relevant to the insurance sector [*Renggli et al.*, 2011], and this quantity can be predicted with high levels of skill across northern Europe and large areas of North America (Figure 4a). Similarly, winter temperatures have impacts on energy pricing and can disrupt transport networks [*Palin et al.*, 2013] but show predictability across large areas around the Atlantic Basin in our seasonal forecasts (Figure 4c). Finally, also related to atmospheric circulation, skillful prediction of near-surface winter wind speeds is demonstrated, again across large areas of Europe and North America (Figure 4e). This quantity is increasingly important as it governs year to year variations in the supply of wind-generated renewable energy. While Figures 4a, 4c, and 4e show large areas of significant correlation skill between the forecasts and observed historical conditions, there are patches of low skill for some fields in regions known to be affected by the NAO, such as temperature in northern Europe, which may arise due to imperfect model teleconnections. It is therefore interesting to ask how well the forecast NAO alone would serve as a proxy for regional prediction. Using only the forecast NAO (Figures 4b, 4d, and 4f) suggests that much of the skill in our forecasts arises from the prediction of the NAO alone. For example, the small regions where storminess is poorly predicted in Figure 4a coincide with regions where the NAO influence is weak (Figure 4b). Furthermore, while skill in North America arises from ENSO and the NAO, for regions such as

Europe where the NAO dominates, then simply using the forecast NAO may actually improve regional predictions (Figure 4d). Using either methodology, and assuming that the recent 20 year period is representative of coming years, predictions from this system could allow plans to be made months ahead for the risk of key weather-related impacts on society.

References

Arribas, A., et al. (2011), The GloSea4 ensemble prediction system for seasonal forecasting, *Mon. Weather Rev.*, *139*, 1891–1910.

Bell, C. J., L. J. Gray, A. J. Charlton-Perez, M. Joshi, and A. A. Scaife (2009), Stratospheric communication of El Niño teleconnections to European Winter, *J. Clim.*, *22*, 4083–4096.

Blockley, E. W., et al. (2013), Recent development of the Met Office operational ocean forecasting system: An overview and assessment of the new Global FOAM forecasts, *Geosci. Model Dev. Discuss.*, *6*, 6219–6278.

Bretherton, C. S., and D. S. Battisti (2000), An interpretation of the results from atmospheric general circulation models forced by the time history of the observed sea surface temperature distribution, *Geophys. Res. Lett.*, *27*, 767–770.

Brönnimann, S., E. Xoplaki, C. Casty, A. Pauling, and J. Luterbacher (2007), ENSO influence on Europe during the last centuries, *Clim. Dyn.*, *28*, 181–197.

Cohen, J., and J. Jones (2011), A new index for more accurate winter predictions. *Geophys. Res. Lett.*, *38*, L21701, doi:10.1029/2011GL049626.

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

Ebdon, R. A. (1975), The quasi-biennial oscillation and its association with tropospheric circulation patterns, *Met. Mag.*, *104*, 282–297.

Fereday, D., A. Maidens, A. Arribas, A. A. Scaife, and J. R. Knight (2012), Seasonal forecasts of Northern Hemisphere winter 2009/10, *Env. Res. Lett.*, *7*, doi:10.1088/1748-9326/7/3/034031.

Folland, C. K., A. A. Scaife, J. Lindesay, and D. Stephenson (2012), How predictable is European winter climate a season ahead?, *Int. J. Clim.*, doi:10.1002/joc.2314.

Frankignoul, C. (1985), Sea surface temperature anomalies, planetary waves and air-sea feed-back in the middle latitudes, *Rev. Geophys.*, *23*, 357–390.

Johansson, A. (2007), Prediction skill of the NAO and PNA from daily to seasonal time scales, *J. Clim.*, *20*, 1957–1975.

Jung, T., F. Vitart, L. Ferranti and J.-J. Morcrette (2011), Origin and predictability of the extreme negative NAO winter of 2009/10, *Geophys. Res. Lett.*, *38*, L07701, doi:10.1029/2011GL046786.

Kim, H.-M., P. Webster, and J. Curry (2012), Seasonal prediction skill of ECMWF System 4 and NCEP CFSv2 retrospective forecast for the Northern Hemisphere winter, *Clim. Dyn.*, *23*, doi:10.1007/s00382-012-1364-6.

Kumar, A. (2009), Finite samples and uncertainty estimates for skill measures for seasonal prediction, *Mon. Weather Rev.*, *137*, 2622–2631.

Kushnir, Y., W. A. Robinson, P. Chang, and A. W. Robertson (2006), The physical basis for predicting Atlantic sector seasonal-to-interannual climate variability, *J. Clim.*, *19*, 5949–5970.

Mehta, V., M. J. Suarez, J. V. Manganello, and T. L. Delworth (2000), Oceanic influence on the North Atlantic Oscillation and associated northern hemisphere climate variations: 1959–1993, *Geophys. Res. Lett.*, *27*, 121–124.

Müller, W. A., C. Appenzeller, and C. Schär (2005), Probabilistic seasonal prediction of the winter North Atlantic Oscillation and its impact on near surface temperature, *Clim. Dyn.*, *24*, 213–226.

Murphy, J. M. (1990), Assessment of the practical utility of extended range ensemble forecasts, *Q. J. R. Meteorol. Soc.*, *116*, 89–125.

Palin, E., et al. (2013), Future projections of temperature-related climate change impacts on the railway network of Great Britain, *Clim. Change*, doi:10.1007/s10584-013-0810-8.

Palmer, T. N., et al. (2004), Development of a European multimodel ensemble system for seasonal to interannual prediction (DEMETER), *Bull. Am. Meteorol. Soc.*, *85*, 853–872.

Pascoe, C. L., L. J. Gray, and A. A. Scaife (2005), A GCM study of the influence of equatorial winds on the timing of sudden stratospheric warmings. *Geophys. Res. Lett.*, *33*, L06825, doi:10.1029/2004JD004941.

Renggli, D., G. C. Leckebusch, U. Ulbrich, S. N. Gleixner, and E. Faust (2011), The skill of seasonal ensemble prediction systems to forecast wintertime windstorm frequency over the north Atlantic and Europe, *Mon. Weather Rev.*, *139*, 3052–3068.

Riddle, E. E., A. H. Butler, J. C. Furtado, J. L. Cohen, and A. Kumar (2013), CFSv2 ensemble prediction of the wintertime Arctic Oscillation, *Clim. Dyn.*, *41*, 1099–1116.

Rodwell, M. J., D. P. Rowell, and C. K. Folland (1999), Oceanic forcing of the wintertime North Atlantic Oscillation and European climate, *Nature*, *398*, 320–333.

Rodwell, M. J., and C. K. Folland (2002), Atlantic air-sea interaction and seasonal predictability, *Q. J. R. Meteorol. Soc.*, *128*(583), 1413–1443, doi:10.1002/qj.200212858302.

Scaife, A. A., et al. (2011), Improved Atlantic blocking in a climate model. *Geophys. Res. Lett.*, *38*, L23703, doi:10.1029/2011GL049573.

Walters, D. N., et al. (2013), The Met Office Unified Model Global Atmosphere 4.0 and JULES Global Land 4.0 configurations, *Geosci. Model Dev. Discuss.*, *6*, 2813–2881.

Smith, D., Scaife, A. A., and Kirtman, B. (2012), What is the current state of scientific knowledge with regard to seasonal and decadal forecasting? *Env. Res. Lett.*, *7*, 015602, doi:10.1088/1748-9326/7/1/015602.

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

WMO (1992), Standardized Verification System (SVS) for long-range forecasts (LRF). WMO Manual on the Global Data-Processing and Forecasting Systems. World Meteorological Organisation, Geneva, Switzerland.

Yang, S., and J. H. Christensen (2012), Arctic sea ice reduction and European cold winters in CMIP5 climate change experiments. *Geophys. Res. Lett.*, *39*, L20707, doi:10.1029/2012GL053338.