Geophysical Research Letters

RESEARCH LETTER
10.1029/2024GL108472

Understanding the Intermittency of the Wintertime North Atlantic Oscillation and East Atlantic Pattern Seasonal Forecast Skill in the Copernicus C3S Multi-Model Ensemble

L. H. Baker[1], L. C. Shaffrey[1], S. J. Johnson[2], and A. Weisheimer[3,4]

[1]Department of Meteorology, National Centre for Atmospheric Science (NCAS), University of Reading, Reading, UK, [2]European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK, [3]Research Department, European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK, [4]Department of Physics, National Centre for Atmospheric Science (NCAS), University of Oxford, Oxford, UK

Key Points:

- Seasonal forecast skill of the wintertime North Atlantic Oscillation (NAO) and East Atlantic Pattern (EA) is intermittent
- Well-forecast NAO/EA winters generally occur when there is substantial tropical forcing

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

L. H. Baker,
l.h.baker@reading.ac.uk

Citation:

Baker, L. H., Shaffrey, L. C., Johnson, S. J., & Weisheimer, A. (2024). Understanding the intermittency of the wintertime North Atlantic Oscillation and East Atlantic Pattern seasonal forecast skill in the Copernicus C3S multi-model ensemble. *Geophysical Research Letters*, *51*, e2024GL108472. https://doi.org/10.1029/2024GL108472

Received 7 FEB 2024
Accepted 26 JUN 2024

Author Contributions:

Conceptualization: L. H. Baker, L. C. Shaffrey
Formal analysis: L. H. Baker
Funding acquisition: L. C. Shaffrey
Investigation: L. H. Baker
Methodology: L. H. Baker
Supervision: L. C. Shaffrey
Visualization: L. H. Baker
Writing – original draft: L. H. Baker
Writing – review & editing: L. H. Baker, L. C. Shaffrey, S. J. Johnson, A. Weisheimer

Abstract

The wintertime North Atlantic Oscillation (NAO) and East Atlantic Pattern (EA) are the two leading modes of North Atlantic pressure variability and have a substantial impact on winter weather in Europe. The year-to-year contributions to multi-model seasonal forecast skill in the Copernicus C3S ensemble of seven prediction systems are assessed for the wintertime NAO and EA, and well-forecast and poorly-forecast years are identified. Years with high NAO predictability are associated with substantial tropical forcing, generally from the El Niño Southern Oscillation (ENSO), while poor forecasts of the NAO occur when ENSO forcing is weak. Well-forecast EA winters also generally occurred when there was substantial tropical forcing, although the relationship was less robust than for the NAO. These results support previous findings of the impacts of tropical forcing on the North Atlantic and show this is important from a multi-model seasonal forecasting perspective.

Plain Language Summary

The wintertime North Atlantic Oscillation (NAO) and East Atlantic Pattern (EA) are two important indicators of atmospheric variability in the North Atlantic. They can have a substantial impact on European winter weather. The ability of seasonal forecast models to forecast the NAO and EA varies from year to year. This intermittency of forecast skill is investigated in seven different seasonal forecast systems from the Copernicus C3S database, by focusing on the most well-forecast and poorly-forecast years. Years where the NAO is well-forecast are associated with substantial tropical forcing, generally from the El Niño Southern Oscillation (ENSO), while poor forecasts of the NAO occur when ENSO forcing is weak. Similar but weaker results hold for the EA. These results are valuable for increasing the usability of seasonal forecasts by identifying conditions under which forecasts are more likely to be skillful.

1. Introduction

The winter North Atlantic Oscillation (NAO) is the leading mode of pressure variability in the North Atlantic (Hurrell, 1995). The positive phase of the NAO is associated with a stronger North Atlantic jet, and typically with mild, wet winters across northern Europe. The negative phase of the NAO is associated with a weaker jet, more frequent occurrence of atmospheric blocking, and cold, dry winters in northern Europe. For example, the strongly negative NAO in winter 2009/10 was associated with extremely cold weather over northern Europe (Cohen et al., 2010), while the strongly positive NAO in winter 2019/20 was associated with very mild and wet weather in northern Europe (Hardiman et al., 2020). The second mode of North Atlantic pressure variability is the East Atlantic Pattern (EA); together with the NAO this modulates the position of the North Atlantic jet. For example, the strongly positive EA in the winter of 2013/14 deflected the North Atlantic jet south over northwestern Europe and resulted in a succession of storms and widespread flooding (Huntingford et al., 2014). Being able to skillfully forecast these two leading modes of wintertime variability for the season ahead could provide substantial socioeconomic benefits to Europe in terms of forecasting, for example, energy demand, flood responses, agriculture and finance.

It has been demonstrated that there is significant skill in winter NAO seasonal forecasts (e.g., Kang et al., 2014; Scaife et al., 2014; Stockdale et al., 2015). Baker, Shaffrey, Sutton, et al. (2018) showed that five of the seven seasonal forecast systems in the EUROSIP multi-model ensemble (MME) had significant skill in forecasting the winter NAO for a 20-year hindcast period (1992–2011), and a MME constructed from these five skillfull systems had a correlation skill of 0.73 for the winter NAO. A similar level of multi-model skill from three seasonal

10.1029/2024GL108472

forecast systems was also found by Athanasiadis et al. (2017) for the NAO. More recently, Lledó et al. (2020) performed an analysis of five seasonal hindcasts from the Copernicus C3S database. They focussed on the leading four modes of variability of 500 hPa geopotential height (Z500) in the North Atlantic, over the hindcast period 1993–2016. Lledó et al. (2020) showed the multi-model correlation skill for the wintertime NAO was statistically significant with a value between 0.3 and 0.4. This skill is lower than that for the EUROSIP systems, which were generally older versions of the C3S modeling systems with smaller ensemble sizes. For the winter EA, Lledó et al. (2020) found that only one of the C3S systems had significant skill, but Thornton et al. (2023) have since shown the C3S multimodel ensemble to have significant skill for forecasting the EA in late autumn/early winter for the 1993–2016 hindcast period. In contrast, Baker, Shaffrey, and Scaife (2018) found significant skill (correlation 0.5) for wintertime seasonal forecasts of a pressure-based index similar to the EA Pattern in the Met Office (MetO) GloSea5 system for winters 1992–2011.

Baker, Shaffrey, Sutton, et al. (2018) noted that in some winters, all or most forecast systems successfully captured the sign and magnitude of the NAO, while in other winters the NAO was poorly forecast by all or most systems. Thus the forecast skill is intermittent, rather than being constant throughout the hindcast period. The main aim of the present study is to further our understanding of this intermittency of seasonal forecast skill for the wintertime (DJF) NAO and EA, with a particular focus on identifying key processes present in the most well-forecast and most poorly-forecast NAO and EA years. Understanding these processes is important in the context of identifying potential situations in which seasonal forecasts are expected to be more reliable, sometimes referred to as "windows of opportunity" (Mariotti et al., 2020). To do this, we consider the performance of the seasonal hindcasts in individual years. Case studies of individual years tend to focus on well-forecast, and often extreme years, for example, the negative NAO in winter 2009/10 (Fereday et al., 2012) or the positive NAO in winter 2019/20 (Hardiman et al., 2020). These aid our understanding of key processes that were active in these years, and highlight potential drivers of predictability of North Atlantic pressure variability. Less work has been done focusing on poorly-predicted years. However, these years could be useful to help to identify which processes are not represented well by the forecast models.

Previously identified drivers of NAO predictability include tropical sea-surface temperature anomalies such as El Niño Southern Oscillation (ENSO) (Hardiman et al., 2019; Ineson & Scaife, 2009), the Indian Ocean Dipole (Hardiman et al., 2020), and western Pacific SSTs (Huntingford et al., 2014). Stratospheric influences on the NAO include the quasi-biennial oscillation (Boer & Hamilton, 2008), sudden stratospheric warmings (SSWs) and strong stratospheric polar vortex events (Scaife et al., 2016). Extratropical surface processes may also play a role including sea-ice anomalies (Hall et al., 2017), Eurasian snow cover (Cohen & Fletcher, 2007) and Atlantic SSTs (Czaja & Frankignoul, 1999; Rodwell et al., 1999). The pathways by which these processes influence the NAO are complex, typically involving interactions between different climate drivers. For the EA, Maidens et al. (2021) demonstrated a link between rainfall anomalies in the tropical North Atlantic and the EA, while Thornton et al. (2023) showed a link between ENSO and the EA in late autumn/early winter. Other recent studies have shown links between ENSO and forecast skill in the North Atlantic through analysis of teleconnection pathways (Williams et al., 2023) and relaxation experiments (Knight et al., 2022). The studies hypothesized that the strength of these teleconnections in seasonal forecast models may be too weak, leading to the signal-to-noise paradox (Eade et al., 2014).

The present study investigates the processes related to predictability of the winter NAO/EA from a multi-model seasonal forecast perspective. We focus on the following objectives:

- Assess the overall skill of the Copernicus C3S multi-model seasonal forecasts of the DJF NAO and EA
- Assess the yearly contributions to NAO and EA forecast skill to identify the most well- and poorly-forecast years
- Identify the key processes in the years with the most well-forecast and poorly-forecast NAO and EA

The paper is structured as follows. In Section 2 the methods and data used in the study are described. The results are presented in Section 3. Finally, a discussion and conclusions are given in Section 4.

2. Data and Methodology

2.1. Data

Data from seven seasonal forecast systems in the C3S archive have been used, namely ECMWF, MetO, Meteo-France (MetFr), DWD, CMCC, NCEP and JMA. Details of the model versions and configuration of the DJF forecasts and hindcasts are given in Table S1 in Supporting Information S1. Hindcasts/forecasts initialized on or before 1 November are used, so the DJF forecasts have a lead-time of 2–4 months. Hindcasts from all systems cover the period 1993–2016 (here we refer to winter seasons by the year of the December; e.g., winter 1993/4 is referred to as 1993). These are combined with forecasts over the period 2017–2022 (where available), giving a total of 30 years. The forecast data consist of the real-time operational forecasts for the corresponding winter seasons, for example, the winter 2017/18 forecast was run operationally on 1 November 2017. The hindcasts are run using typically smaller ensemble sizes (e.g., for the ECMWF system there are 25 hindcast ensemble members compared to 51 in the forecasts); these are detailed in Table S1 in Supporting Information S1. For some systems, regular updates to the operational systems mean that the forecasts in different years are produced by different versions of the model. We make the assumption that these differences in model formulation are not likely to have a significant impact on our results.

Reanalysis data from ERA5 (Hersbach et al., 2020) are used as a reference to evaluate the forecast systems against. ERA5 provides an ensemble of reanalysis estimates of various atmospheric fields, derived by combining observations and model simulation. Here the ensemble mean, monthly-mean fields of mean sea-level pressure (MSLP) are used.

2.2. Definition of NAO and EA Indices

To assess the uncertainty in the definition of the NAO, two definitions are used here, as in Baker, Shaffrey, Sutton, et al. (2018). The first is the difference in MSLP between the Azores (37°N, 25°W) and Iceland (65°N, 22°W), following Scaife et al. (2014). The second, referred to as the box-based NAO, is the difference in MSLP averaged over a southern box (90°W–60°E, 20°N–55°N) and a northern box (90°W–60°E, 55°N–90°N) (Stephenson et al., 2006). The broader spatial scale of the box-based NAO accounts for the fact that there are often differences in the location of the NAO centers of action in models compared to observations (Walz et al., 2018).

The EA is characterized by the pressure variation centered in the mid-North Atlantic, to the west of the UK. To define the EA, the MSLP at (52.5°N, 27.5°W) was used, following Moore et al. (2011).

2.3. Evaluation Metrics

The anomaly correlation coefficient (ACC) is used to provide a measure of ensemble mean correlation skill (see Supporting Information S1 for full definition). ACC can range from −1 to 1, with 1 indicating perfect skill, and values less than zero indicating no skill.

Following Weisheimer et al. (2019), the relative contribution of an individual year i to the overall correlation skill is given by

$$\text{contribution}_i = \frac{\bar{x}_i\, y_i}{\sigma_{\bar{x}}\sigma_y} \tag{1}$$

where y_i and $\bar{x}_i$ are, respectively, the observed and ensemble mean anomalies in year i, and σ_y and $\sigma_{\bar{x}}$ are the standard deviations over time of the observation and ensemble mean, respectively. Anomalies are taken relative to the timeseries mean of the relevant quantities.

The ratio of predictable components (RPC, Eade et al., 2014) is a measure of over or underconfidence, and represents the ratio of the predictability of the real world, to the predictability of the model. It is defined fully in Supporting Information S1. Values of RPC significantly greater than one indicate that the system is underconfident.

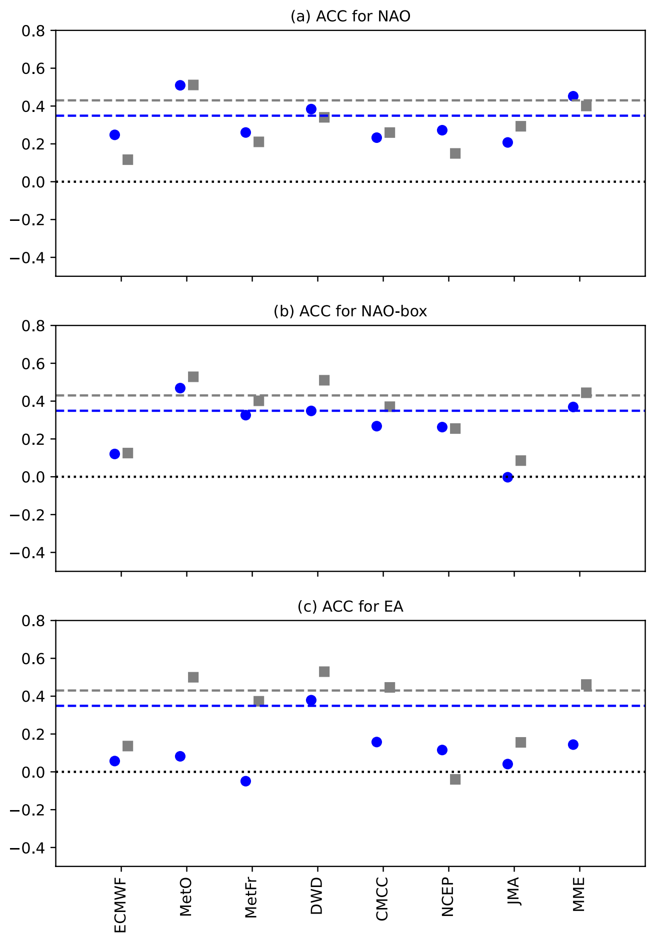

Figure 1. Summary of anomaly correlation coefficient (ACC) for each C3S forecast system evaluated against ERA5. The rows show quantities for the point-based North Atlantic Oscillation (NAO) (top), NAO-box (middle) and East Atlantic Pattern (EA) (bottom). ACC values are shown for the full hindcast/forecast period 1993–2022 (blue circles) and for the reduced hindcast period 1993–2011 (gray squares). Blue and gray dashed lines shows the value for significance at the 5% level for the 30-year and 19-year sample sizes, respectively.

3. Results

3.1. Evaluation of Winter NAO and EA Seasonal Forecast Skill in the C3S Systems

Figure 1 shows the ACC for the DJF NAO and EA indices for the period 1993–2022 (blue circles), for each system and for the MME. The MME was constructed by averaging over all seven systems, giving each system equal weighting. Only MetO and DWD show significant skill for the point-based NAO index (Figure 1a); the other systems have ACC values around 0.2–0.3. The ACC for the MME is 0.45 which is actually lower than the ACC for the highest single system (MetO). Skill for the box-based NAO index is similar (Figure 1b), but with much lower skill for the ECMWF and JMA systems than for the point-based NAO. For the EA (Figure 1c), the skill is very low (less than 0.2 in most systems), with only DWD showing significant skill. To give a broader perspective, Figure S1 in Supporting Information S1 shows maps of ACC for DJF MSLP. There is little skill in the North Atlantic region, apart from regions of significant skill around the NAO centers of action in MetO and DWD. In DWD there is also significant skill extending through the central North Atlantic, consistent with the significant skill for the EA in this model. Equivalent ACC maps for Z500 (not shown) show very similar patterns of skill to those for MSLP, the main differences being slightly better skill in the north-east of the domain in the MetO and DWD hindcasts.

It is of interest to compare these results with those found for the EUROSIP multimodel (Baker, Shaffrey, Sutton, et al., 2018), where the ACC ranged from −0.1 to 0.6 for individual EUROSIP systems. To give a more direct comparison between C3S and EUROSIP, the gray squares in Figure 1 show the ACC for the C3S systems for the common period with EUROSIP (1993–2011). For NAO and NAO-box, ACC values for individual systems range from 0.1 to 0.5. For the box-based NAO, in five out of seven systems the skill is slightly higher for the shorter period than for the full period, while for the point-based NAO the skill is similar in both periods. Examining the skill for the later part of the period (2012–2022) separately (gray diamonds in Figure S2 in Supporting Information S1) shows a substantial drop in skill for the box-based NAO compared to the earlier period, but not for the point-based NAO. For the EA (Figure 1c), the skill for the 1993–2011 period is much higher than for the full period, with four systems and the MME showing significant skill, compared with just one system for the full period. In contrast there is no skill for the 2012–2022 period in any system except NCEP (Figure S2 in Supporting Information S1).

Figure S3 in Supporting Information S1 shows scatter plots of the RPC against ACC for each system, for both the full hindcast period and the shorter period 1993–2011. In general, there is a correspondence between higher RPC and higher skill: systems with significant skill are generally also underconfident (RPC > 1). There is a marked difference in RPC for the EA between these two periods: for the full period there is just one underconfident system (RPC > 1), while in the shorter period four systems are underconfident.

3.2. Interannual Variation in NAO and EA Forecast Skill

For brevity, in the remaining sub-sections we focus on the results for the point-based NAO and EA indices and show equivalent results for the NAO-box index in Supporting Information S1.

Figure 2a shows timeseries of standardized anomalies of the NAO in ERA5 and the ensemble mean from each of the C3S forecast systems. Standardized values (anomalies relative to the hindcast period, divided by the standard deviation) are used to allow for a visual comparison between the hindcasts and ERA5, since the magnitudes of the hindcast ensemble mean anomalies are much smaller than the observed magnitude. For the NAO index, winter

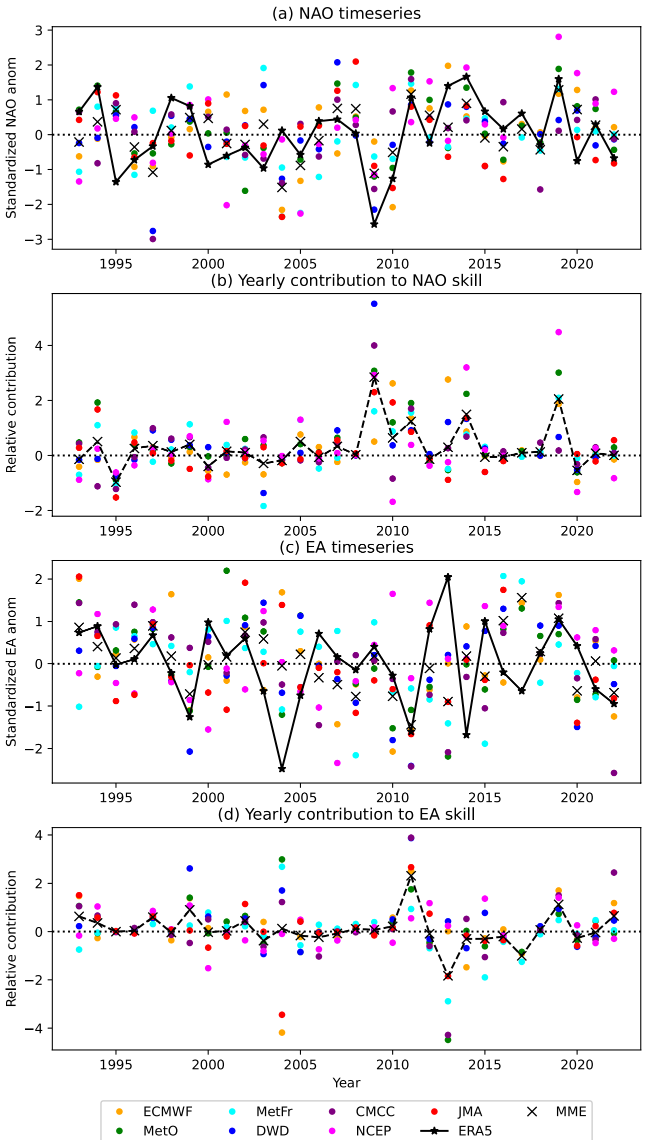

Figure 2. Timeseries of standardized ensemble mean hindcasts and forecasts of DJF (a) North Atlantic Oscillation (NAO) and (c) East Atlantic Pattern (EA) for each C3S forecast system (colors), the multi-model ensemble (black "+" symbols) and ERA5 (bold black lines). Relative contribution of each year to the anomaly correlation coefficient (ACC) skill for (b) NAO and (d) EA.

2009/10 stands out as a strongly negative NAO year which was well forecast by all models, in agreement with Lledó et al. (2020). 2011/12, 2014/15, and 2019/20 were positive NAO years that were well forecast by all seasonal forecast systems. In contrast the negative NAO in 1995/6 and 2020/21 were not forecast by any of the C3S systems. For the EA (Figure 2c), the positive EA in winters 1997/8 and 2019/20 were well forecast by all systems, as was the negative EA in 2011/12. In contrast the negative EA in 2004/5 and 2017/18 were not captured by the models, and neither was the positive EA in 2013/14.

To further evaluate the forecasts in individual winters, we decompose the contribution that each winter has to the ACC skill (see Section 2.3). Only a few years make a large positive contribution to the MME skill for NAO (Figure 2b), while the contribution from most years is close to zero. For the MME and most of the forecast systems, 2009 makes the biggest contribution to skill for the NAO. In 2011/12, 2014/15, and 2019/20 there is also a consistently positive contribution for all models. In contrast, 1995/6 and 2020/21 show a clear negative contribution to skill for all models. For the EA (Figure 2d), again it can be seen that only a few years make large positive or negative contributions to MME skill. 1997/8, 2011/12, and 2019/20 are all positive contributors to skill for all models. In contrast 2013/14 contributes negatively for most models and 2017/18 contributes negatively to skill in the three available forecasts. 2004/5 has an interesting split between models, with the contribution to skill being negative for two models and positive for the rest. Two years (2011/12 and 2019/20) had good forecasts of both the NAO and EA, while 2020/21 had poor forecasts of both the NAO and EA. The other well and poorly forecast years were identified for only one of these. This is because when the NAO or EA is close to 0, the contribution to skill will by definition always be small. This is the case in, for example, 2009/10, where the NAO was well-forecast: here the EA was close to 0 in both ERA5 and the hindcasts, so although there was good correspondence between observation and forecasts, the resulting contribution to skill was 0. This also applies to the near-neutral observed NAO in 2004/5.

The intermittency in skill identified here, and the generally agreement between seasonal forecast models about which winters are well and poorly forecast for each index, implies that particular years may be more predictable than others. In the next section we investigate potential sources of this predictability.

3.3. Drivers of Predictability in Years With Well and Poorly Forecast NAO/EA

In this section we aim to identify the key drivers of predictability of the NAO and EA. This is done by first assessing the key processes and potential drivers active in the well- and poorly-forecast NAO/EA years identified in the previous section. Based on these findings, we use contingency tables to assess the relationship between tropical forcing and NAO/EA predictability.

3.3.1. Potential Drivers of Predictability in the Most Well- and Poorly-Forecast Years

In order to understand more about potential drivers of predictability, we focus on the well-forecast and poorly-forecast winters identified in the previous section based on relative contribution to skill (namely 1997/8, 2009/10, 2011/12, 2014/15, and 2019/20; and 1995/6, 2004/5, 2013/14, 2017/18, and 2020/21, respectively). Table 1a gives a summary of key features from existing literature relating to potential drivers of predictability in each of these years.

Table 1
(a) Table Summarising the Potential Drivers of the Most Well-Forecast and Poorly-Forecast Years (See Text for Details). (b) and (c): Contingency Tables for Years With Strong/Neutral ENSO, and Good/Poor NAO and EA Forecasts, Respectively

Year	Index	Potential drivers	Tropical forcing?	SSW?	Strong SPV?
(a)					
Well-forecast years					
1997	EA+	Strong El Niño	Y	N	N
2009	NAO−	High Eurasian snow cover extent; weakening of the polar vortex; two SSWs; strong El Niño and easterly QBO (Cohen et al., 2010; Fereday et al., 2012)	Y	Y	N
2011	NAO+, EA−	Persistent La Niña since the previous winter	Y	N	N
2014	NAO+	Moderate El Niño strengthening throughout winter. Atypical atmospheric response (Peng et al., 2018). Cold North Atlantic SSTs throughout winter (Duchez et al., 2016)	Y	N	N
2019	NAO+, EA+	Neutral ENSO but very strong positive Indian Ocean Dipole. Strong Stratospheric Polar Vortex (SPV) (Hardiman et al., 2020)	Y	N	Y
Poorly-forecast years					
1995	NAO−	Cold polar stratospheric temperatures (Manney et al., 1996) and resulting low stratospheric ozone levels (Müller et al., 1997). Weak La Niña	N	N	Y
2004	EA−	No strong signal from tropics or other common NH teleconnections (Santos et al., 2007). Very cold Arctic stratosphere temperatures (El Amraoui et al., 2008)	N	N	Y
2013	EA+	Neutral ENSO but heavy rainfall over the west Pacific, Indonesia and the eastern Indian Ocean; strong westerly QBO (Huntingford et al., 2014). Weakened Aleutian Low; strong polar vortex	Y	N	Y
2017	EA−	Weak La Niña, SSW in February (Knight et al., 2021)	N	Y	N
2020	NAO−, EA+	Moderate La Niña. Strong SSW in early January persisting to early February (Lee, 2021). Low autumn Barents-Kara Arctic sea ice (Lu et al., 2021)	Y	Y	N

	Poor NAO	Good NAO
(b)		
Strong ENSO	2	6
Neutral ENSO	6	1

	Poor EA	Good EA
(c)		
Strong ENSO	2	5
Neutral ENSO	4	3

Note. DJF El Niño Southern Oscillation (ENSO) Magnitude Was Computed From the ESRL/NOAA Niño 3.4 Index Based On HadISST by Calculating the Absolute Value of Anomalies Relative to the Full Available Period (Winters 1870–2022). Strong/neutral ENSO years are defined as the upper/lower tercile DJF ENSO magnitude, respectively. Good/poor forecast years are defined as those in the upper/lower tercile MME contribution to skill for NAO/EA.

Based on Table 1a, a common feature of the five well-forecast years is that they all occurred when the Tropical Pacific was in a strong or relatively strong El Niño or La Niña state, apart from 2019 (in which there was a strongly positive Indian Ocean Dipole (IOD)). In contrast, the five poorly-forecast winters occurred when ENSO was in a neutral state, except for 2020/21 (moderate La Niña). Furthermore, all five poorly-forecast years had strong stratospheric anomalies: in winters 1995/6, 2004/5, and 2013/14 the polar vortex was strong; and in both 2017/18 and 2020/21 SSWs occurred in late winter. Only one of the well-forecast years (2009/10) had a SSW occurring, and only one had a strong stratospheric polar vortex (2019/20).

3.3.2. The Relationship Between Tropical Forcing and Winter NAO and EA Predictability

To further investigate the relationship between tropical forcing and the predictability of the winter NAO and EA, we focus on the link with ENSO, but also discuss other potential sources of predictability from the tropics. Tables 1b and 1c show contingency tables of ENSO strength versus the MME contribution to skill of each year for the NAO and the EA, respectively. ENSO strength is defined as the absolute value of the DJF Niño 3.4 anomaly (i.e., El Niño and La Niña years are considered together). “Strong” and “neutral” ENSO years are defined as the top and bottom terciles (10 years) in terms of ENSO magnitude. “Good” and “poor” forecast years are defined as the top and bottom terciles of the MME contribution to NAO/EA skill.

Table 1b shows a clear correspondence between well-forecast NAO winters occurring when the ENSO signal is strong, and poorly-forecast NAO winters occurring when ENSO is neutral ENSO. For the EA (Table 1c), the results show the same general pattern but the results are weaker than for the NAO.

The relationship between strong and weak ENSO years, and good and poor NAO and EA forecasts (respectively), can be further shown by calculating the skill of subsets of years. For the NAO, the ACC for the MME drops to 0.34 when strong ENSO years are excluded and increases to 0.62 when weak ENSO years are excluded. For the EA the corresponding ACC values are 0.08 and 0.28, respectively.

4. Discussion and Conclusions

The Copernicus C3S multi-model seasonal forecast ensemble has been used to investigate the intermittency of seasonal forecast skill for the wintertime NAO and EA, and to understand more about processes driving the predictability of these modes of North Atlantic variability.

For the winter NAO, the overall ACC skill for the C3S forecast systems was similar to that for the previously studied EUROSIP forecast systems. The box-based NAO and the EA showed a general drop in skill between the earlier part (1993–2011) and the later part (2012–2022) of the period studied; this drop in skill was not seen for the point-based NAO. Some variation in skill between relatively short hindcast periods is expected, and is due to the influence of individual years on the overall hindcast skill. It was also found that a small number of years make large positive or negative contributions to the skill, and these years are generally consistent between models.

It was found that in general, winters in which the NAO was well forecast occurred when ENSO was strong, while poor NAO forecasts occurred when ENSO was neutral. These results are consistent with those of O'Reilly et al. (2020) which showed that low NAO forecast skill in the mid-twentieth Century coincided with low ENSO variability in this period, while more recent periods had more ENSO variability and higher NAO forecast skill. We note that the two strongest El Niño winters in the period studied (winters 1997/98 and 2015/16) were not associated with strongly negative NAO or with large positive contributions to NAO forecast skill. It has been previously shown that very strong El Niño events impact the North Atlantic differently from less strong events (King et al., 2023), and in particular do not project onto the NAO in the same way. In both these years, the observed NAO was close to zero and therefore the contribution to skill was small, even though most systems forecast the correct NAO sign. Another interesting year to consider is 2014/15, in which there was a moderate El Niño but a well-forecast positive NAO (i.e., not the expected response). Xie and Zhang (2017) concluded that the atypical atmospheric response to the El Niño state was due to internal variability. The positive NAO may therefore have been related to the persistent cold North Atlantic SSTs in this case (Duchez et al., 2016). There were also some years in the period studied where ENSO was not active and other tropical drivers had a notable impact on the North Atlantic. These include 2019, in which the IOD was unusually strong, while ENSO was weak, leading to an impact from the IOD on the North Atlantic (Hardiman et al., 2020); and 2013, in which there was heavy rainfall in the west Pacific (Huntingford et al., 2014). The importance of ENSO and other tropical drivers on NAO variability has been noted in other previous studies, typically focusing on individual models or individual years, for example, Ineson and Scaife (2009), Jiménez-Esteve and Domeisen (2018), and Scaife et al. (2017). In the present study we have investigated this relationship from a multi-model perspective.

The results are useful when considering the usability of seasonal forecasts in the context of windows of opportunity. By identifying conditions under which the state is more predictable, it is possible to highlight forecasts that may have more skill or be more reliable. The results suggest that if there is a strong El Niño or La Niña state, then a user might have more confidence in the forecasts for the upcoming winter season NAO and EA. This is particularly true for moderately strong El Niño events, the response of which projects strongly onto the NAO. Care should be taken in interpreting NAO forecasts in very strong El Niño years, in which the expected impacts on the North Atlantic are different (see, e.g., Toniazzo & Scaife, 2006) and may not lead to the same enhanced predictability seen under more moderate El Niño conditions. Enhanced predictability in years when ENSO is active has previously been shown for marine heatwaves (Jacox et al., 2022), monsoon rainfall (Dunstone et al., 2020) and for multi-year predictions for the Pacific region (Liu et al., 2023), highlighting potential windows of opportunity in active ENSO years in other regions and on longer timescales. Some ways in which windows of opportunity could be utilized in a practical forecasting context are discussed in, for example, Dunstone et al. (2023) and Mariotti et al. (2020).

Another aspect found in the most poorly-forecast years was large perturbations in the stratosphere in the absence of strong tropical forcing: either the occurrence of SSWs or of very strong stratospheric polar vortex events. This could suggest that the models fail to represent either stratospheric perturbations that are not driven by tropical forcing, and/or that they fail to represent the propagation of the signal from the stratosphere to the lower troposphere in the North Atlantic region. Future work could investigate this further and establish the cause of these deficiencies, with the hope of improving the representation of these processes in future seasonal forecast models.

Data Availability Statement

The seasonal hindcast and forecast data used in this study is freely available from Copernicus Climate Change Service, Climate Data Store (2018). The ERA5 reanalysis data (Hersbach et al., 2020) used in this study is freely available from Hersbach et al. (2023). The HadISST NINO3.4 data used in this study is freely available from https://climatedataguide.ucar.edu/climate-data/nino-sst-indices-nino-12-3-34-4-oni-and-tni.

Acknowledgments
LHB was funded by the Natural Environment Research Council (NERC), Grant NE/V001787/1. LCS and AW were funded by CANARI (NE/W004984/1). We thank the ECMWF Copernicus for making the C3S seasonal hindcast data via the Climate Data Store. We also thank two anonymous reviewers for their helpful comments.

References

Athanasiadis, P. J., Bellucci, A., Scaife, A. A., Hermanson, L., Materia, S., Sanna, A., et al. (2017). A multi-system view of wintertime NAO seasonal predictions. *Journal of Climate*, *30*(4), 1461–1475. https://doi.org/10.1175/JCLI-D-16-0153.1

Baker, L., Shaffrey, L., & Scaife, A. (2018). Improved seasonal prediction of UK regional precipitation using atmospheric circulation. *International Journal of Climatology*, *38*(S1), e437–e453. https://doi.org/10.1002/joc.5382

Baker, L., Shaffrey, L., Sutton, R., Weisheimer, A., & Scaife, A. (2018). An intercomparison of skill and overconfidence/underconfidence of the wintertime North Atlantic Oscillation in multimodel seasonal forecasts. *Geophysical Research Letters*, *45*(15), 7808–7817. https://doi.org/10.1029/2018GL078838

Boer, G., & Hamilton, K. (2008). QBO influence on extratropical predictive skill. *Climate Dynamics*, *31*(7), 987–1000. https://doi.org/10.1007/s00382-008-0379-5

Cohen, J., & Fletcher, C. (2007). Northern hemisphere winter surface temperature predictions based on land–atmosphere fall anomalies. *Journal of Climate*, *20*(16), 4118–4132. https://doi.org/10.1175/JCLI4241.1

Cohen, J., Foster, J., Barlow, M., Saito, K., & Jones, J. (2010). Winter 2009–2010: A case study of an extreme Arctic oscillation event. *Geophysical Research Letters*, *37*(17), L17707. https://doi.org/10.1029/2010GL044256

Copernicus Climate Change Service, Climate Data Store. (2018). *Seasonal forecast monthly statistics on single levels*. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.68dd14c3

Czaja, A., & Frankignoul, C. (1999). Influence of the North Atlantic SST on the atmospheric circulation. *Geophysical Research Letters*, *26*(19), 2969–2972. https://doi.org/10.1029/1999GL900613

Duchez, A., Frajka-Williams, E., Josey, S. A., Evans, D. G., Grist, J. P., Marsh, R., et al. (2016). Drivers of exceptionally cold North Atlantic Ocean temperatures and their link to the 2015 European heat wave. *Environmental Research Letters*, *11*(7), 074004. https://doi.org/10.1088/1748-9326/11/7/074004

Dunstone, N., Smith, D., Yeager, S., Danabasoglu, G., Monerie, P.-A., Hermanson, L., et al. (2020). Skilful interannual climate prediction from two large initialised model ensembles. *Environmental Research Letters*, *15*(9), 094083. https://doi.org/10.1088/1748-9326/ab9f7d

Dunstone, N., Smith, D. M., Hardiman, S. C., Davies, P., Ineson, S., Jain, S., et al. (2023). Windows of opportunity for predicting seasonal climate extremes highlighted by the Pakistan floods of 2022. *Nature Communications*, *14*(1), 6544. https://doi.org/10.1038/s41467-023-42377-1

Eade, R., Smith, D., Scaife, A., Wallace, E., Dunstone, N., Hermanson, L., & Robinson, N. (2014). Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? *Geophysical Research Letters*, *41*(15), 5620–5628. https://doi.org/10.1002/2014GL061146

El Amraoui, L., Semane, N., Peuch, V.-H., & Santee, M. L. (2008). Investigation of dynamical processes in the polar stratospheric vortex during the unusually cold winter 2004/2005. *Geophysical Research Letters*, *35*(3), L03803. https://doi.org/10.1029/2007GL031251

Fereday, D. R., Maidens, A., Arribas, A., Scaife, A. A., & Knight, J. R. (2012). Seasonal forecasts of northern hemisphere winter 2009/10. *Environmental Research Letters*, *7*(3), 034031. https://doi.org/10.1088/1748-9326/7/3/034031

Hall, R. J., Scaife, A. A., Hanna, E., Jones, J. M., & Erdélyi, R. (2017). Simple statistical probabilistic forecasts of the winter NAO. *Weather and Forecasting*, *32*(4), 1585–1601. https://doi.org/10.1175/WAF-CD-16-0124.1

Hardiman, S. C., Dunstone, N. J., Scaife, A. A., Smith, D. M., Ineson, S., Lim, J., & Fereday, D. (2019). The impact of strong El Niño and La Niña events on the North Atlantic. *Geophysical Research Letters*, *46*(5), 2874–2883. https://doi.org/10.1029/2018GL081776

Hardiman, S. C., Dunstone, N. J., Scaife, A. A., Smith, D. M., Knight, J. R., Davies, P., et al. (2020). Predictability of European winter 2019/20: Indian Ocean Dipole impacts on the NAO. *Atmospheric Science Letters*, *21*(12), e1005. https://doi.org/10.1002/asl.1005

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al. (2023). *ERA5 monthly averaged data on pressure levels from 1940 to present*. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.6860a573

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Huntingford, C., Marsh, T., Scaife, A. A., Kendon, E. J., Hannaford, J., Kay, A. L., et al. (2014). Potential influences on the United Kingdom's floods of winter 2013/14. *Nature Climate Change*, *4*(9), 769–777. https://doi.org/10.1038/nclimate2314

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

Ineson, S., & Scaife, A. (2009). The role of the stratosphere in the European climate response to El Niño. *Nature Geoscience*, *2*(1), 32–36. https://doi.org/10.1038/ngeo381

Jacox, M. G., Alexander, M. A., Amaya, D., Becker, E., Bograd, S. J., Brodie, S., et al. (2022). Global seasonal forecasts of marine heatwaves. *Nature*, *604*(7906), 486–490. https://doi.org/10.1038/s41586-022-04573-9

Jiménez-Esteve, B., & Domeisen, D. I. V. (2018). The tropospheric pathway of the ENSO–North Atlantic teleconnection. *Journal of Climate*, *31*(11), 4563–4584. https://doi.org/10.1175/JCLI-D-17-0716.1

Kang, D., Lee, M.-I., Im, J., Kim, D., Kim, H.-M., Kang, H.-S., et al. (2014). Prediction of the Arctic Oscillation in boreal winter by dynamical seasonal forecasting systems. *Geophysical Research Letters*, *41*(10), 3577–3585. https://doi.org/10.1002/2014GL060011

King, M. P., Keenlyside, N., & Li, C. (2023). ENSO teleconnections in terms of non-NAO and NAO atmospheric variability. *Climate Dynamics*, *61*(5), 2717–2733. https://doi.org/10.1007/s00382-023-06697-8

Knight, J., Scaife, A., Bett, P. E., Collier, T., Dunstone, N., Gordon, M., et al. (2021). Predictability of European winters 2017/2018 and 2018/2019: Contrasting influences from the tropics and stratosphere. *Atmospheric Science Letters*, *22*(1), e1009. https://doi.org/10.1002/asl.1009

Knight, J., Scaife, A., & Maidens, A. (2022). An extratropical contribution to the signal-to-noise paradox in seasonal climate prediction. *Geophysical Research Letters*, *49*(23), e2022GL100471. https://doi.org/10.1029/2022GL100471

Lee, S. H. (2021). The January 2021 sudden stratospheric warming. *Weather*, *76*(4), 135–136. https://doi.org/10.1002/wea.3966

Liu, Y., Donat, M. G., England, M. H., Alexander, L. V., Hirsch, A. L., & Delgado-Torres, C. (2023). Enhanced multi-year predictability after El Niño and La Niña events. *Nature Communications*, *14*(1), 6387. https://doi.org/10.1038/s41467-023-42113-9

Lledó, L., Cionni, I., Torralba, V., Bretonnière, P.-A., & Samsó, M. (2020). Seasonal prediction of Euro-Atlantic teleconnections from multiple systems. *Environmental Research Letters*, *15*(7), 074009. https://doi.org/10.1088/1748-9326/ab87d2

Lu, Q., Rao, J., Liang, Z., Guo, D., Luo, J., Liu, S., et al. (2021). The sudden stratospheric warming in January 2021. *Environmental Research Letters*, *16*(8), 084029. https://doi.org/10.1088/1748-9326/ac12f4

Maidens, A., Knight, J. R., & Scaife, A. A. (2021). Tropical and stratospheric influences on winter atmospheric circulation patterns in the North Atlantic sector. *Environmental Research Letters*, *16*(2), 024035. https://doi.org/10.1088/1748-9326/abd8aa

Manney, G. L., Santee, M. L., Froidevaux, L., Waters, J. W., & Zurek, R. W. (1996). Polar vortex conditions during the 1995–96 Arctic Winter: Meteorology and MLS ozone. *Geophysical Research Letters*, *23*(22), 3203–3206. https://doi.org/10.1029/96GL02453

Mariotti, A., Baggett, C., Barnes, E. A., Becker, E., Butler, A., Collins, D. C., et al. (2020). Windows of opportunity for skillful forecasts subseasonal to seasonal and beyond. *Bulletin of the American Meteorological Society*, *101*(5), E608–E625. https://doi.org/10.1175/BAMS-D-18-0326.1

Moore, G. W. K., Pickart, R. S., & Renfrew, I. A. (2011). Complexities in the climate of the subpolar North Atlantic: A case study from the winter of 2007. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 757–767. https://doi.org/10.1002/qj.778

Müller, R., Crutzen, P. J., Groo, J.-U., Bürhl, C., Russell, J. M., Gernandt, H., et al. (1997). Severe chemical ozone loss in the Arctic during the winter of 1995–96. *Nature*, *389*(6652), 709–712. https://doi.org/10.1038/39564

O'Reilly, C. H., Weisheimer, A., MacLeod, D., Befort, D. J., & Palmer, T. (2020). Assessing the robustness of multidecadal variability in Northern Hemisphere wintertime seasonal forecast skill. *Quarterly Journal of the Royal Meteorological Society*, *146*(733), 4055–4066. https://doi.org/10.1002/qj.3890

Peng, P., Kumar, A., & Hu, Z.-Z. (2018). What drove the Pacific and North America climate anomalies in winter 2014/15? *Climate Dynamics*, *51*(7), 2667–2679. https://doi.org/10.1007/s00382-017-4035-9

Rodwell, M. J., Rowell, D. P., & Folland, C. K. (1999). Oceanic forcing of the wintertime North Atlantic oscillation and European climate. *Nature*, *398*(6725), 320–323. https://doi.org/10.1038/18648

Santos, J., Corte-real, J., & Leite, S. (2007). Atmospheric large-scale dynamics during the 2004/2005 winter drought in Portugal. *International Journal of Climatology*, *27*(5), 571–586. https://doi.org/10.1002/joc.1425

Scaife, A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R., Dunstone, N., et al. (2014). Skillful long-range prediction of European and North American winters. *Geophysical Research Letters*, *41*(7), 2514–2519. https://doi.org/10.1002/2014GL059637

Scaife, A., Comer, R. E., Dunstone, N. J., Knight, J. R., Smith, D. M., MacLachlan, C., et al. (2017). Tropical rainfall, Rossby waves and regional winter climate predictions. *Quarterly Journal of the Royal Meteorological Society*, *143*(702), 1–11. https://doi.org/10.1002/qj.2910

Scaife, A., Karpechko, A. Y., Baldwin, M., Brookshaw, A., Butler, A., Eade, R., et al. (2016). Seasonal winter forecasts and the stratosphere. *Atmospheric Science Letters*, *17*(1), 51–56. https://doi.org/10.1002/asl.598

Stephenson, D. B., Pavan, V., Collins, M., Junge, M. M., Quadrelli, R., & Participating CMIP2 Modelling Groups. (2006). North Atlantic Oscillation response to transient greenhouse gas forcing and the impact on European winter climate: A CMIP2 multi-model assessment. *Climate Dynamics*, *27*(4), 401–420. https://doi.org/10.1007/s00382-006-0140-x

Stockdale, T. N., Molteni, F., & Ferranti, L. (2015). Atmospheric initial conditions and the predictability of the Arctic Oscillation. *Geophysical Research Letters*, *42*(4), 1173–1179. https://doi.org/10.1002/2014GL062681

Thornton, H. E., Smith, D. M., Scaife, A. A., & Dunstone, N. J. (2023). Seasonal predictability of the East Atlantic pattern in late autumn and early winter. *Geophysical Research Letters*, *50*(1), e2022GL100712. https://doi.org/10.1029/2022GL100712

Toniazzo, T., & Scaife, A. A. (2006). The influence of ENSO on winter North Atlantic climate. *Geophysical Research Letters*, *33*(24), L24704. https://doi.org/10.1029/2006GL027881

Walz, M. A., Donat, M. G., & Leckebusch, G. C. (2018). Large-scale drivers and seasonal predictability of extreme wind speeds over the North Atlantic and Europe. *Journal of Geophysical Research: Atmospheres*, *123*(20), 11518–11535. https://doi.org/10.1029/2017JD027958

Weisheimer, A., Decremer, D., MacLeod, D., O'Reilly, C., Stockdale, T. N., Johnson, S., & Palmer, T. N. (2019). How confident are predictability estimates of the winter North Atlantic Oscillation? *Quarterly Journal of the Royal Meteorological Society*, *145*(S1), 140–159. https://doi.org/10.1002/qj.3446

Williams, N. C., Scaife, A. A., & Screen, J. A. (2023). Underpredicted ENSO teleconnections in seasonal forecasts. *Geophysical Research Letters*, *50*(5), e2022GL101689. https://doi.org/10.1029/2022GL101689

Xie, J., & Zhang, M. (2017). Role of internal atmospheric variability in the 2015 extreme winter climate over the North American continent. *Geophysical Research Letters*, *44*(5), 2464–2471. https://doi.org/10.1002/2017GL072772