Jet latitude regimes and the predictability of the North Atlantic Oscillation

Kristian Strommen

Physics Department, University of Oxford, UK

Correspondence
K. Strommen, AOPP, Clarendon Laboratory, Oxford University, Parks Road, Oxford OX1 3PU, UK.
Email: kristian.strommen@physics.ox.ac.uk

Funding information
European Commission (Horizon 2020), Grant/Award Number: 64127; Thomas Phillips and Jocelyn Keene Stipendiary Junior Research Fellowship in Science of Climate

Abstract
In recent years, numerical weather prediction models have begun to show notable levels of skill at predicting the average winter North Atlantic Oscillation (NAO) when initialised one month ahead. At the same time, these model predictions exhibit unusually low signal-to-noise ratios, in what has been dubbed a 'signal-to-noise paradox'. We analyse both the skill and signal-to-noise ratio of the Integrated Forecast System (IFS), the European Centre for Medium-range Weather Forecasts (ECMWF) model, in an ensemble hindcast experiment. Specifically, we examine the contribution to both from the regime dynamics of the North Atlantic eddy-driven jet. This is done by constructing a statistical model which captures the predictability inherent to the trimodal jet latitude system, and fitting its parameters to reanalysis and IFS data. Predictability in this regime system is driven by interannual variations in the persistence of the jet latitude regimes, which determine the preferred state of the jet. We show that the IFS has skill at predicting such variations in persistence: because the position of the jet strongly influences the NAO, this automatically generates skill at predicting the NAO. We show that all of the skill the IFS has at predicting the winter NAO over the period 1980–2010 can be attributed to its skill at predicting regime persistence in this way. Similarly, the tendency of the IFS to underestimate regime persistence can account for the low signal-to-noise ratio, giving a possible explanation for the signal-to-noise paradox. Finally, we examine how external forcing drives variability in jet persistence, and highlight the role played by transient baroclinic eddy feedbacks to modulate regime persistence.

Keywords
eddy feedback, eddy-driven jet, jet latitude, NAO, persistence, seasonal prediction, signal-to-noise paradox, winter predictability
1 | INTRODUCTION

The leading mode of variability in the North Atlantic region is referred to as the North Atlantic Oscillation (NAO), which captures a frequently recurring dipole structure of pressure anomalies. During the winter season the NAO is to a large extent capturing the location of the storm track (Rogers, 1997) and therefore has a big influence on European winter climate. Despite the fact that, due to the chaotic nature of the atmosphere, any individual storm is not predictable for more than a few weeks ahead, evidence has begun to accumulate which suggests that the average winter conditions may be predictable (Smith et al., 2016). Furthermore, modern numerical weather prediction (NWP) models have begun to show notable and consistent levels of skill at capturing this predictable signal when initialized one month ahead (though it should be noted that comparable skill was already reported in Müller et al. (2005)). In a particularly noteworthy example, studies conducted by the UK Met Office system GloSea5 (Scaife et al., 2014a; Dunstone et al., 2016) showed that their seasonal prediction system displayed skillful ensemble hindcasts over the 35-year period 1980–2015. Subsequently, the multi-model study Athanasiadis et al. (2017) showed that comparable levels of skill were attainable by other models as well.

Two questions have naturally sprung out of all this work. Firstly, can one identify concrete external drivers that account for the majority of interannual NAO variability? A significant number of studies have attempted to address this question, both in terms of locating drivers of the observed NAO as well as studying which of these drivers are responsible for model skill. A good overview of these can be found in Smith et al. (2016). Some of the more consistently proposed drivers include Atlantic sea surface temperatures (Czaja and Frankignoul, 1999; Rodwell et al., 1999; Saunders and Qian, 2002; Dunstone et al., 2016), the El Niño Southern Oscillation (ENSO; Dong et al., 2000; Moron and Gourand, 2003; Toniazzo and Scaife, 2006; Brönnimann, 2007; Li and Lau, 2012; Drouard et al., 2013; Jiménez-Esteve and Domeisen, 2018), the stratospheric polar vortex (Baldwin and Dunkerton, 2001; Kidston et al., 2015; Dunstone et al., 2016), the quasi-biennial oscillation (QBO; Anstey et al., 2013; Scaife et al., 2014b; O’Reilly et al., 2019) and Arctic sea ice, particularly in the Kara region (Strong and Magnusdottir, 2011; Koenigk et al., 2016; Wang et al., 2017).

The second big question arises from the observation, first noted in Eade et al. (2014) for GloSea5 and subsequently corroborated by the multi-model study Baker et al. (2018), that the signal-to-noise ratio of model NAO ensemble predictions tends to be too low. In other words, the predictable component of the model appears to be a lot smaller than that of the real world. This phenomenon has been described as a “signal-to-noise paradox” in Dunstone et al. (2016); Scaife and Smith (2018) gives an overview. As a consequence, large ensembles are required in order for the ensemble mean to amplify the predictable signal sufficiently. Because a resolution of this “paradox” could lead to more skilful predictions, or equivalent skill with a smaller ensemble, it is of great interest to understand what is producing this low signal-to-noise ratio in models.

In Kumar et al. (2014) and Siegert et al. (2016), “linear regression style” statistical models aimed at representing NAO predictability were analysed, suggesting that the paradox may be due to the model not adequately representing a linear predictable signal in the magnitude of the NAO. This implies that one possible explanation is one or more overly weak teleconnections in the model. Both North Atlantic sea-surface temperatures (SSTs; Rodwell et al., 1999; Mehta et al., 2000; Gastineau et al., 2013; Scaife et al., 2013) and the QBO (O’Reilly et al., 2019) have been suggested as examples of such.

An alternative hypothesis was proposed in Strommen and Palmer (2019), based on the idea that the NAO is driven by regime dynamics. The possible influence of regimes on North Atlantic circulation has been studied since the 1980s, often by way of classifying distinct, frequently recurring and quasi-persistent geopotential height patterns, of which the two phases of the NAO are examples (Vautard, 1990; Michelangeli et al., 1995; Dawson et al., 2012). Studies have shown that both frequency of occurrence of individual regimes and the transition rates between different regimes can be influenced by external forcing (Charlton-Perez et al., 2018); it was suggested in Palmer (1999) that anthropogenic forcing may also influence regime behaviour in a similar manner. In Strommen and Palmer (2019), a statistical model was presented which captures the predictability inherent to such regime dynamics on seasonal time-scales. By envisaging the atmosphere as a two-state system, with transitions taking place on daily time-scales, it was shown that interannual variability in such a system is determined by changes in the preferred regime state each year and that this preference is in turn driven by variations in the persistence time-scales of the two regimes. Furthermore, if NWP models systematically underestimate the level of regime persistence relative to the real world, then the signal-to-noise paradox emerges as a natural feature of the system. That models struggle to replicate many aspects of regimes, including persistence, has been observed in e.g., Strommen et al. (2019), which highlighted the importance of horizontal resolution.

The purpose of the present paper is to build on the framework of Strommen and Palmer (2019) in three key ways. Firstly, the statistical model in that work was not fitted directly against NWP model data, instead relying
on a freely varying “regime fidelity” parameter to demonstrate that observed skill and signal-to-noise behaviour could be captured within a reasonable range of parameter values. We will show here, by explicitly fitting our statistical model, that the signal and signal-to-noise ratios of the European Centre for Medium-range Weather Forecasts (ECMWF) seasonal prediction system (the Integrated Forecast System, IFS) can indeed be fully accounted for by its skill at capturing interannual variations in regime behaviour alone. This is done by basing our statistical model not on geopotential height regimes, but on the trimodal structure of the North Atlantic eddy-driven jet (Woollings et al., 2010a). Because the NAO is to a large extent reflecting variations in the jet speed and latitude (Athanasiadis et al., 2010; Woollings et al., 2010a), one may \textit{a priori} expect to account for much of the model NAO skill in terms of its skill at predicting shifts in the jet speed and latitude. In fact, in Parker et al. (2019), simple linear regression techniques were used to show that the majority of the IFS’s skill at predicting the NAO could be attributed to its skill at predicting shifts in the jet latitude alone. However, this approach is only valid once the trimodality has been smoothed away by taking seasonal means, which renders the probability distribution Gaussian. Given the fundamentally trimodal structure of the jet latitude, we would argue that a more appropriate way to analyse the contribution from jet latitude variations is through regime dynamics. As explained above, in the framework of the regime model from Strommen and Palmer (2019), the predictable component is not the average seasonal jet latitude itself, but rather the average seasonal persistence time-scales of the jet latitude regimes (which in turn determine the preferred jet latitude regime). By truncating the trimodal structure to a two-state system, we show that the IFS has skill at predicting seasonal shifts in these persistence time-scales, and hence that the signal-to-noise paradox can be explained by NWP models having systematic biases in regime persistence, as proposed in Strommen and Palmer (2019). The success of such a two-state truncation is consistent with the hypothesis of Woollings et al. (2008), which also viewed the NAO via a two-state system using the jet.

Secondly, having in this way fitted a regime dynamics model to jet latitude data for both reanalysis and the IFS, we analyse external sources of interannual variability of regime persistence. We show that commonly cited predictors of the NAO, such as ENSO, Kara Sea ice and the stratospheric polar vortex all project onto regime persistence. Of particular note here is the role of ENSO, which, while long thought to influence the NAO, does not have any significant linear correlation with the NAO index, implying that any such influence is necessarily nonlinear. We demonstrate significant linear correlations of ENSO (via the Niño 3.4 index) with regime persistence time-scales, implying that regime dynamics may be a practical way to quantitatively measure the nonlinear influence of ENSO on the NAO.

Finally, the link between NAO variability and persistence time-scales of the eddy-driven jet immediately suggests that a critical role may be played by local dynamics (as opposed to remotely driven forcing), in the form of transient baroclinic eddy forcing. There is a vast literature demonstrating the role played by such forcings in maintaining jet anomalies (Hoskins et al., 1983; Shutts, 1983; Robinson, 1996; Lorenz and Hartmann, 2001; 2003; Zurita-Gotor et al., 2014), and it is therefore reasonable to expect these to influence the persistence time-scales of the jet regimes used in our statistical model. By defining a suitable metric based on the eddy momentum flux convergence, we show that this is indeed the case, and that this is in fact the dominant source of jet latitude variability. For this metric we find that individual IFS ensemble members cannot generally reproduce the observed forcing strength, with a large ensemble mean required to boost the signal; a similar inadequacy for GloSea5 was observed in Scaife et al. (2019). This suggests that the signal-to-noise paradox may be driven not just by weak teleconnections but also a weak eddy forcing in NWP models.

The structure of the paper is as follows. In Section 2 we describe the data used, provide basic details of the IFS model and make some basic comments and definitions of key metrics; this includes the “ratio of predictable components” (RPC) metric (Eade et al., 2014), which is used as a measure of the signal-to-noise ratio. In Section 3, in order to keep the paper fairly self-contained, we provide a basic description of the regime dynamics model constructed in Strommen and Palmer (2019). In Section 4, we discuss the relation between the jet and the NAO on daily and seasonal time-scales, including the influence of the trimodal jet latitude structure, for both reanalysis and the IFS. This analysis is used to motivate the explicit statistical model fitting carried out in the same section. In Section 5, we show the results of this fitting, and verify that the high skill and low signal-to-noise ratio (i.e., high RPC value) of the IFS can be explained using our statistical model. In Section 6 we analyse the influence of both external forcing and local eddy forcing on regime persistence time-scales, both for reanalysis and the IFS. Finally, in section 7 we make our concluding remarks.

As a final remark to the reader more keen to view regimes through the prism of Z500 fields, we note that the work in Madonna et al. (2017) allows one to essentially translate the jet perspective presented here to a geopotential height picture. We do not comment on this further.
2 | DATA AND METHODOLOGY

2.1 | Data

The IFS model data considered comes from a seasonal hindcast experiment covering the 20th Century (Weisheimer et al., 2017). A 51-ensemble member seasonal forecast is initialised in every 1 November from 1900 to 2010 and allowed to run for four months, thereby producing a December–January–February (DJF) prediction for every year in this period. The model used is version CY41R1 of the IFS. Its spectral resolution is T255, corresponding to roughly 80 km grid spacing near the Equator, with 91 levels in the vertical. The model is run in atmosphere-only mode with prescribed SSTs. Because the hindcasts cover the early part of the century where observations are limited, the reanalysis product ERA20C (Poli et al., 2013) is used to initialise the model atmosphere; this assimilates only observations of surface pressure. Both ocean and sea-ice boundary conditions come from the HadISST2.1.0.0 dataset (Rayner et al., 2003). Further details can be found in Weisheimer et al. (2017).

While this paper will focus on the recent period 1980–2010, where more comprehensive reanalysis datasets assimilating satellite data are available, we will use ERA20C as our primary source of observations for consistency. Therefore all wind and geopotential height data used is from ERA20C, with SSTs and sea-ice coming from HadISST2. The exception for this is for the QBO, where the ERA-Interim dataset (Dee et al., 2011) is used. This is due to the fact that ERA20C has a very weak QBO, owing to the fact that surface pressure data alone do not constrain the stratospheric circulation sufficiently. However, the QBO may have been partially forcing the surface pressure data (which does severely constrain the NAO), so it is important to use a more realistic representation of it to assess its potential impact.

The NAO is defined as the leading empirical orthogonal function (EOF) of detrended daily geopotential height anomalies at 500 hPa (Z500), when restricted to the North Atlantic region 90°W–30°E, 30°N–90°N. Anomalies are relative to a seasonal climatology computed across the full ERA20C time period 1900–2010. The daily NAO of an individual IFS ensemble member is obtained by projecting its geopotential height field onto the reference EOF from ERA20C. A winter NAO timeseries is obtained by taking the DJF mean of the daily timeseries. It should be noted that the principal component timeseries produced in this way is not equivalent to that obtained by taking the leading EOF of DJF Z500 means. This latter approach only captures variability across seasons, while the use of daily data will also capture variability within seasons. However, the EOFs themselves are almost identical in their spatial pattern, with both capturing the familiar NAO dipole.

The speed and latitude of the eddy-driven jet are computed as in Parker et al. (2019). Daily zonal means of zonal winds are restricted to the region 0°–60°W, 15°–75°N and 850 hPa, before being smoothed with a 9-day running mean. For each day of the DJF season, the speed of the jet is identified as the maximum wind speed attained in this domain and the latitude of the jet is the latitude at which this maximum is located. As noted in Parker et al. (2019), this procedure produces qualitatively similar results to more complex methods using winds at multiple levels or further filtering.

In Section 6.1 we consider various external predictors. For ENSO, we use the Niño 3.4 index, defined as the average across the region 5°S–5°N, 190°–240°E. Since studies show that the response in the North Atlantic typically lags ENSO by around one month (Jiménez-Esteve and Domeisen, 2018), we consider the mean of this index over the November–December–January (NDJ) period to capture the influence on the North Atlantic DJF season. A QBO timeseries is defined by taking the DJF mean of zonal equatorial (10°S–10°N) winds at 50 hPa, zonally averaged. The stratospheric polar vortex is measured using the November zonal mean winds at 50 hPa averaged between 60° and 80°N. Kara Sea ice is measured as the November sea-ice anomaly, relative to the period 1978–2016, in the box 30°–75°E, 67°–80°N. The data are detrended to remove the linear trend resulting from declining sea-ice. Note that, for these last two predictors, November means only were used to avoid confounding issues resulting from interactions and feedbacks between these and the NAO. Because of the slow time-scales of both ENSO and the QBO, as well as their spatial separation from the centre of NAO action, this was not deemed to be necessary in these cases.

2.2 | Methodology

We are interested in three key metrics assessing model skill, all closely related. We define Corr(EnsMean, Obs) to be the Pearson correlation coefficient between the ensemble mean of some forecast and observations over a given time period. We will refer to this informally as the actual predictability. On the other hand, we define Corr(EnsMean, Mem) to be the average correlation between the ensemble mean and individual members of the ensemble. This is sometimes referred to as the potential predictability. However, as the signal-to-noise paradox has demonstrated clearly, this metric is not necessarily a meaningful upper bound on skill, so we follow the convention of Strommen and Palmer (2019) and refer to this instead as the model predictability of the
system, as it measures the model’s ability to predict itself.

The third metric of interest is the “ratio of predictable components”, RPC, defined in Eade et al. (2014). One first defines the notion of the “predictable component” of a given system as the square root of the total fraction of variance that can be predicted in that system. When applied to observations we obtain PC(Obs), and when applied to a model we obtain PC(Mod). The RPC is then by definition

\[
RPC_{\text{true}} = \frac{PC(\text{Obs})}{PC(\text{Mod})}.
\]

In Eade et al. (2014) they provide an estimate of RPC amenable to computation via the following inequality:

\[
RPC_{\text{est}} = \frac{\text{PC(Obs)}}{\text{PC(Mod)}} \geq \frac{\sqrt{\text{Corr(EnsMean, Obs)}^2}}{\sqrt{\sigma_{\text{sig}}^2 / \sigma_{\text{tot}}^2}}.
\]

where \(\sigma_{\text{sig}}^2\) is the ensemble mean variance and \(\sigma_{\text{tot}}^2\) is the average variance of individual ensemble members. We define the RPC estimate to be the absolute value of this last term:

\[
RPC_{\text{est}} = \left| \frac{\text{Corr(EnsMean, Obs)}}{\sqrt{\sigma_{\text{sig}}^2 / \sigma_{\text{tot}}^2}} \right|.
\]

Note the absolute value is necessary because we are taking a positive square root. As noted in Strommen and Palmer (2019), if all ensemble members are assumed to share the same variance (which would in expectation hold given an infinite ensemble), then \(RPC_{\text{est}}\) is precisely equal to the ratio of actual predictability divided by model predictability. Hence this metric will be 1 if all ensemble members and observations behave as random draws from the same underlying distribution. Further discussion of this metric can be found in Strommen and Palmer (2019). While \(RPC_{\text{est}}\) is an estimate of the true RPC, we will without comment refer to this estimate simply as the RPC.

In terms of significance testing, 95% confidence intervals are, unless otherwise stated, generated by a Monte Carlo method of explicitly creating \(> 1,000\) estimates using the statistical model in question. This is the case for all metrics computed with the regime dynamics model constructed in this paper. For correlations between various observed time series (such as external predictors as in Section 6.1), Monte Carlo samples were generated by a bootstrap resampling method. In practice, intervals found in this way were not appreciably different from those obtained simply using the standard error of correlations, so for simplicity we have used these.

Finally, we will frequently be examining statistical properties of probabilities, including relationships between different time series of probabilities. To facilitate this, we typically work with log-odds rather than probabilities when manipulating or generating probabilities. From an assumed distribution of probabilities \(P\) (taking values between 0 and 1), a new, more Gaussian distribution \(\hat{P}\) (taking values between \(-\infty\) and \(+\infty\)) is obtained via the log-odds transformation

\[
P = \log \left( \frac{P}{1-P} \right).
\]

Specific values \(\hat{p}\) drawn from this latter distribution correspond to the probability \(p\) from \(P\) given by

\[
p = \frac{1}{1 + e^{-\hat{p}}}.
\]

Because the distributions of probabilities we work with are log-normal or reverse log-normal, the log-odds distribution will be approximately Gaussian and so linear regression techniques can be used meaningfully. Note that a distribution \(Y\) is called reverse log-normal if \(1-Y\) is log-normal. While we often and without comment treat log-odds and probabilities interchangeably in discussion, the distinction is made clear when this is important.

2.3 Non-stationarity of predictability and skill metrics for the period 1980–2010

It has been noted that the skill the IFS has at predicting the winter NAO is non-stationary over the Twentieth Century (O’Reilly et al., 2017). Because there appears to be variability not only in the role of external forcing but also local regime dynamics, it would be ideal to examine the framework presented in this paper across multiple periods of the Twentieth Century. However, because this would complicate the presentation of the paper considerably, we will focus on the recent period 1980–2010 on questions of predictability. Estimates of distributions or parameters that are independent of skill are computed using the full period 1900–2010, while those clearly dependent on skill are computed using the period 1980–2010.

In this modern period, using the definitions of Section 2.2, the IFS ensemble mean winter NAO has a correlation with the observed winter NAO of 0.49. The model predictability is 0.19 and the \(RPC_{\text{est}}\) over this period is 1.80. These are the numbers that we wish to explain.
3 | THE REGIME-BASED LINK BETWEEN THE JET AND THE NAO

3.1 | The distributions of jet latitude and jet speed

It is well known that the NAO, while often depicted using sea-level pressure/Z500 anomalies, can be viewed as capturing variations in both the speed and orientation of the eddy-driven jet (Thompson et al., 2003; Woollings et al., 2008). In this view, days in which the NAO is in its positive phase (NAO+ days) typically correspond to the jet being shifted northwards while days in the negative phase (NAO– days) correspond to a southward shift (e.g., Figure 1 of Woollings et al. (2010a)).

By taking as its starting point an analysis of variations in the jet, as opposed to variations in the NAO, Woollings et al. (2010b) found that the jet tends to exhibit regime-like behaviour in terms of its location. Concretely, they identified three preferred locations, corresponding to the trimodality of the probability density function (pdf) describing the daily latitudinal position of the jet. A comparison of this pdf for ERA20C and the IFS was done in Parker et al. (2019), which the reader should consult for a more comprehensive comparison. For completeness, we include a figure showing the pdfs here as well. To estimate a pdf from jet latitude data, we use a simple Gaussian mixture model assuming three peaks. As noted in Parker et al. (2019), the IFS also exhibits a trimodal jet, so this assumption on the mixture model is appropriate for the IFS as well. The estimated pdfs of both ERA20C and the IFS, using all 110 years of data, are plotted in Figure 1a. The three peaks are clearly visible and correspond closely to those in Parker et al. (2019). The shading indicates an estimate of uncertainty of the ERA20C probabilities, obtained by fitting the mixture model to individual 30-year chunks of the 110 years. While the locations of the peak are fairly stable, the relative sizes of each of the three modes vary more with time. Therefore, in line with Parker et al. (2019), we identify three distinct jet latitude regimes: Southern (15–40°N), Central (40–52°N) and Northern (52–70°N). The IFS appears to be biased towards spending proportionally more time in the Central regime at the expense of the Southern and Northern regimes. By contrast, the pdf of daily jet speeds, shown in Figure 1b, is visibly Gaussian. The IFS has a clear bias towards stronger mean jet speeds than ERA20C. Finally, one can check (e.g., by computing correlations) that the speed and latitude on both daily and seasonal time-scales are largely independent of each other.

3.2 | Daily and seasonal links between the jet and the NAO

In order to create a statistical model, we need to establish a quantitative link between the jet and the NAO. Because we are ultimately interested in the DJF mean NAO, a first approach might be to relate the DJF means of the jet latitude and speed to the DJF mean NAO. This was the approach taken in Parker et al. (2019), where the DJF NAO was decomposed into a jet speed and jet latitude component using multiple linear regression. Indeed, collapsing daily data to seasonal means completely eliminates the trimodal structure, rendering the pdf of seasonal mean jet latitudes normal: this can be seen in Figure 1c,d, showing histograms of DJF means of jet latitude and speed respectively. By comparing these DJF decompositions of the NAO into a latitude and speed component for the IFS and ERA20C, Parker et al. (2019) concluded that the IFS skill was, in this framework, attributable to the jet latitude skill alone. However, the fundamentally trimodal structure of the daily jet latitude suggests that this linear approach may not be the most appropriate and comes at the cost of removing structure which may contain valuable information.

In fact, the failure of linear regression to account for the link between the jet and the NAO on daily time-scales can be seen clearly in Figure 2, which shows scatter-plots of daily jet latitudes for ERA20C in (a) and the IFS in (c). Similarly, daily jet speeds against daily NAO values are shown for ERA20C in (b) and the IFS in (d). While the relationship between the jet speed and the NAO appears to be linear in nature, the relationship between the jet latitude and the NAO is visibly nonlinear, both for ERA20C and the IFS. In particular, negative NAO values may occur both at very low and very high latitudes. For convenience, the points in (a) and (c) have been marked according to which of the three regimes they are in. The three regimes each appear to have their own distinct relationship with the NAO; given the trimodality of jet latitudes, such a trimodal influence is perhaps to be expected a priori. By comparison, Figure 3a–d shows the same scatter-plots but using DJF means instead of daily values. While the relationship between the jet latitude and the NAO now appears to be linear, the comparison with Figure 2 suggests that this is at least in part because the nonlinear structure has been smeared away by averaging. Note that this nonlinear relationship between the jet latitude and the NAO was also noted in Woollings et al. (2010a). Its potential implications for NAO predictability do not seem to have been commented on in the literature.

It is important to note that the relationship in Figure 2a, resembling a continuous parabola, does not in
FIGURE 1 Estimated probability distributions of (a) the daily jet latitude, (b) the daily jet speed, (c) the DJF mean jet latitude, and (d) the DJF mean jet speed. In all plots, ERA20C is shown in solid black and the IFS in dashed red. The full period 1900–2010 is used. In (a), the grey shading indicates a 95% confidence interval around the ERA20C frequencies, estimated by recomputing the pdf over multiple distinct periods of the 20th century.

and of itself imply regime behaviour. The work in Fyfe and Lorenz (2005) and Monahan and Fyfe (2006) shows that such a parabola can arise from purely Gaussian latitudinal variations in a jet of approximately constant shape. Indeed, if one Taylor expands the zonal wind strength as a function of latitudinal deviations, then the quadratic term becomes non-negligible if these deviations are large, as is the case here. If the NAO is also reflecting the strength of the zonal mean jet, then this may lead to a quadratic relationship between the jet latitude and the NAO. However, because jet latitude variations are known to exhibit regime behaviour, we argue that a quantitative relationship between the jet and the NAO requires the framework of regime dynamics to capture fully.

As noted in the Introduction, the predictable component across a given season is, in a regime system, the preferred regime state. A simple way to capture this is to simply count, in each DJF, the number of days spent in each regime. Figure 3 also shows a scatter-plot of the number of days spent in the Southern regime each season against the seasonal NAO mean, for ERA20C (e) and the IFS (f). There is a strong linear relationship, with a correlation of −0.6 for ERA20C and −0.67 for the IFS. The relationship is such that seasons where the jet spends a lot of time in the Southern regime correspond on average to NAO− seasons, consistent with the aforementioned finding in Woollings et al. (2010b) that NAO− seasons correspond on average to a southward shift in the jet. The correlations obtained for this regime-inspired metric are notably higher than those obtained by simply correlating DJF means of jet latitude (−0.6 versus 0.39), implying that representing the regime-like behaviour of the jet can add quantitatively practical information. We will show in the next section that this relationship can be inferred to good approximation from the daily jet latitude structure.

Similar scatter-plots of the days spent in the Central (Northern) regime against the winter NAO do not show as strong a signal, with a correlation of 0.45 (0.14). In terms of the jet latitude influence on the NAO, the dominant signature appears to be the proportion of time spent in the Southern regime versus the Central and Northern regimes. Hence, while the jet latitude itself is trimodal, when considered in terms of its influence on the NAO, a bimodal truncation, obtained by conceiving the Central and Northern regimes as one, appears to be appropriate. This is consistent with the analysis of Woollings et al. (2010b), which effectively showed that a two-state mixture model approximation of the NAO selects precisely this
FIGURE 2  Scatter plots of daily NAO indices against (a, c) jet latitude and (b, d) jet speed for (a, b) ERA20C and (c, d) IFS. In each plot, the value C in the title is the linear correlation between the two quantities. In (a, c), latitudes in the Southern (< 40°N), Central (40°N to 52°N) and Northern (>52°N) regimes are shown as blue crosses, black dots and red triangles respectively.

bimodal jet latitude truncation (cf. figure 14 in Woollings et al. 2010b). While further motivation for this truncation will be discussed in later sections, we note that it does appear to throw away some of the information in Figure 2. It is possible that a more delicate use of this information could improve our results further, but the bimodal approximation we consider will be sufficient for our purposes here.

4 | THE REGIME DYNAMICS OF THE TRIMODAL JET

In the previous section we argued that the influence of the jet latitude on the NAO should be understood via the regime behaviour visibly taking place on daily time-scales. We showed that while the jet latitude is trimodal, its influence on the NAO is reasonably well captured by a bimodal truncation. In this section we will use this to construct a statistical model aimed at capturing the predictable signal coming from this bimodal regime structure contained in the winter NAO index.

Note that the statistical model we construct will only use information from the jet latitude, ignoring the jet speed entirely. While Figures 3c,d show there is a strong linear relationship between the DJF mean jet speed and the DJF mean NAO, with correlations of around 0.6, the IFS does not have any skill at predicting average jet speeds. This was already a conclusion of Parker et al. (2019), but for our purposes it is sufficient to note that the correlation between the DJF jet speed of the IFS ensemble mean and the DJF jet speed of ERA20C is ≈ −0.02. Therefore we would not expect jet speeds to be making a contribution to skill.

4.1 | The Markov model

In this section we recall the main features of the statistical model of bimodal regime dynamics constructed in Strommen and Palmer (2019): we refer to this model informally as the “Markov model”. Readers should refer to Strommen and Palmer (2019) for a more complete discussion of this model.

We are going to imagine the jet as a bimodal system operating on daily time-scales. From this perspective, a DJF season corresponds to approximately 90 days, where, for each day, the jet resides in one of two states. The natural mathematical framework for generating 90 days of jet regime states is a two-state Markov chain. We will show in the next section that the autocorrelation of daily jet
FIGURE 3 Scatter plots of DJF mean jet (a, b) latitude, (c, d) speed and (e, f) number of days spent in the Southern jet regime against DJF mean NAO indices for (a, c, e) ERA20C and (b, d, f) IFS. In each plot, the value C in the title is the linear correlation between the two quantities. The full period 1900–2010 is used.

Latitudes are well captured with an AR(1) process, implying that the Markov property does hold in this case and that it is not necessary to consider memory beyond one day.

Consider a two-state Markov chain. Denote the two states by the deliberately ambiguous monikers A and B; they will be defined in the next section. The two states have persistence probabilities denoted by $p_A$ and $p_B$ (ranging between 0 and 1). If we imagine the chain operating on daily time-scales, this means concretely that if the jet is in the A (B) regime on a given day, it will remain in this regime the next day with probability $p_A$ ($p_B$) or transition to the other regime with probability $1 - p_A$ ($1 - p_B$).

To each state we now associate probability distributions which we denote by $X_A$ and $X_B$. These distributions represent the range of daily NAO index values that can be obtained when the jet is in states A and B respectively; they are assumed to be stationary. A single season of NAO values is now generated as follows. We assume that the probabilities $p_A$ and $p_B$ are fixed at the start of December and remain constant during the season. The jet will, on 1 December, be in one of the two states, after which it will carry out a 90-day random walk in the Markov chain. On each day when the jet is in state A, we draw a random number from the distribution $X_A$, which will be the NAO index value on that day. Similarly, on each day when the jet is in state B, the NAO index on that day is a random draw from $X_B$. At the end of the season, the 90 daily NAO values are then averaged to generate a DJF mean NAO.

If we let $\pi_A$ and $\pi_B$ denote the long-term average proportion of time spent in state A and B respectively, then one can show (Strommen and Palmer, 2019) that

\[
\pi_A = \frac{1 - p_B}{2 - p_A - p_B},
\]

and hence the preferred state is determined by the relative values of $p_A$ and $p_B$, with no preferred state if, and only if, $p_A = p_B$. Furthermore, if $Y = Y(p_A, p_B)$ represents the distribution of DJF NAO means obtained through this process, it was shown in Strommen and Palmer (2019) that $Y$ has expected value

\[
\mathbb{E}(Y) = \pi_A \mathbb{E}(X_A) + \pi_B \mathbb{E}(X_B) \tag{1}
\]

and variance

\[
\text{Var}(Y) = \frac{\pi_A \text{Var}(X_A) + \pi_B \text{Var}(X_B)}{90}, \tag{2}
\]

where 90 is the number of days in the season. Because the means $\mathbb{E}(X_A)$ and $\mathbb{E}(X_B)$ of $X_A$ and $X_B$ are constant.
by assumption, this implies that the only source of predictability in a season comes from the persistence probabilities \( p_A \) and \( p_B \), these determine the values of \( \pi_{A} \) and \( \pi_{B} \) (i.e. the preferred state of the jet) which, via the above equations, determine the expected NAO value. Predictability of interannual variability in jet persistence therefore corresponds to predictability of interannual NAO variability. This predictability would be induced by seasonal deviations in the persistence probabilities from their climatological means, and the ability of NWP models to detect these. Possible triggers of such deviations will be considered in Section 6.

A caveat to the above is that in seasons where \( p_A \) and \( p_B \) are close to each other, that is, seasons where the jet only has a very weak preference to one particular state, then 90 days may not be a sufficient number of days to witness the above theoretical formulae, which become exact only in the limit of infinitely many days. Such a situation, in which the position of the jet may be dominated by internal variability, would correspond to a winter with a low level of predictability. Conversely, years in which \( p_A \) and \( p_B \) are strongly separated would correspond to a winter with a high level of predictability. We will see that, overall, there appears to be a sufficient amount of predictability in interannual variations of persistence probabilities to fully account for observed NAO predictability.

A final important point here is that, in principle, the probabilities \( p_A \) and \( p_B \) may be completely independent of each other. It is therefore necessary to consider the impact of any remote or local forcing on the system through both probabilities separately, with the overall impact being a combination of both. This will be analysed further in later sections.

4.2 | How predictable are jet latitude dynamics?

Fitting the above framework to actual jet latitude data requires first and foremost the identification of the two states \( A \) and \( B \). The conclusion of Section 3.2 suggests that we identify \( A \) with the Southern regime and \( B \) with the combined Central and Northern states. However, because we are ultimately interested in the question of predictability, we also need to take into account what the IFS does and does not have skill at predicting. To this end, we considered both potential bimodal truncations. Truncation 1 is defined as the one described above: a “South” regime corresponding to the Southern regime and a “North” regime corresponding to a combined Central and Northern regime. Truncation 2 is defined by having a ”South” regime corresponding to the combined Southern and Central regimes and a “North” regime corresponding to the Northern regime. For each truncation, it is straightforward to estimate, for each regime, the seasonal persistence probabilities for both ERA20C and each IFS ensemble member. Supporting Figure S1 shows estimates, for each regime, of the IFS skill at predicting the probabilities of ERA20C, measured as the correlation between the ensemble mean persistence and the persistence of ERA20C; only data from the period 1980–2010 were used. While the skill at predicting persistence of the North regime is largely insensitive to the choice of truncation, with a correlation of \( \approx 0.3 \) either way, there is a dramatic difference for the South regime. For Truncation 1, the IFS has no notable skill at all, with a correlation of \( \approx 0.3 \). However, for Truncation 2 the IFS skill goes up to \( \approx 0.5 \). That is, while the truncation most relevant for determining the NAO is Truncation 1, the truncation actually predictable by the IFS appears to be Truncation 2.

While the Truncation 2 South correlation of 0.5 is statistically significant, the corresponding North correlation of 0.3 is not outside the 95% confidence interval of the no-skill null hypothesis. To bolster confidence in both these correlations, we therefore considered the same correlations over earlier 30-year chunks. Concretely, we computed analogous correlations over the thirty “chunks” of data 1950–1980, 1951–1981, ..., 1979–2009, 1980–2010. The mean across all these chunks are \( \approx 0.55 \) and \( \approx 0.32 \) for the South and North regimes respectively. In fact, these correlations are effectively stationary across each chunk, including for the very first period 1950–1980 which has no overlap with the modern period 1980–2010. The IFS therefore does appear to have robust levels of skill at predicting the persistence probabilities in Truncation 2, at least since 1950 onwards. Earlier periods begin to show evidence of the drop in NAO skill discussed in O’Reilly et al. (2017) and so were not included.

The fact that the IFS only has skill at predicting Truncation 2 is perhaps puzzling. However, there is evidence to suggest that the mechanisms driving jet variability are qualitatively different in the midlatitudes, where the Southern and Central regime occur, compared to the higher-latitude Northern regime closer to Greenland. In particular, Barnes and Hartmann (2010) showed that transient eddy feedbacks acted to reinforce the jet in the midlatitudes only, with northern jet anomalies being more closely associated with Rossby wave variability. If the IFS is simply able to broadly detect if jet variability is going to be dominated by baroclinic eddies or Rossby waves in a given season, then this may be expected to lead to skill at predicting persistence in the Truncation 2 context only, even if the IFS is not able to clearly distinguish between the Southern and Central regime. Furthermore, in White et al. (2019) it was shown that a large proportion of Northern regime events appear to be localised Greenland “tip-jet events”, a
surface-level phenomenon which appears to be less closely related to the equivalent barotropic phenomenon of the NAO; this further reinforces the idea that jet persistence is likely to be less relevant for the Northern regime. In Section 6.2, we will independently confirm that eddy forcing appears to be crucial for understanding persistence of the Southern and Central regimes, but less so for the Northern. Further analysis as to why IFS skill is limited to Truncation 2 will be left for future work.

Since it is Truncation 1 which is most relevant for determining the NAO, the lack of IFS skill for this truncation would at first sight appear to be a major obstruction towards explaining NAO skill using regime dynamics. However, distinguishing between a preference towards the Northern regime or the Southern and Central regimes can still be expected to contribute towards NAO skill, albeit to a lesser extent. We will see in Section 5 that this contribution is in fact sufficient to account for the entirety of observed NAO skill.

Finally, Figure 4 shows the autocorrelation of daily jet latitudes for both ERA20C and the IFS, along with the theoretical autocorrelation functions assuming the data is AR(1) (i.e., satisfies the Markov property). The AR(1) functions match the observed data closely, confirming that the daily jet latitudes satisfy the Markov property to a sufficient degree to justify the approach of Section 4.1. It is also clear that the IFS is in general under-persistent compared to ERA20C, with consistently reduced levels of autocorrelation.

4.3 | Determination of the Markov model distributions

The analysis in the previous section suggests that we need to effectively be working with both Truncation 1 and Truncation 2 in the Markov model in a manner we will now sketch out. Truncation 2 (Southern+Central and Northern) will be used for determining the seasonal persistence probabilities of ERA20C and the extent to which the IFS can predict these. These will allow us to build up a quantitative measure of where the jet is spending most of its time each season. For each day spent (by ERA20C or the IFS) in the combined Southern+Central regime, we simply assume that the jet will be in the Southern regime according to the proportions implied from the full, trimodal jet latitude pdf: that is, with probability \( P(\text{Jet in Southern} | \text{Jet in Southern or Central}) \). This sketch will be fleshed out fully in the next section, but we mention it now to better motivate the figures that follow.

Figure 5 shows the estimates of \( X_A \) and \( X_B \) in Truncation 1 for both ERA20C and the IFS. Hence in (a) is shown the distribution of possible daily NAO values when the jet is in the Southern regime, and in (b) the same for when the jet is in either of the Central and Northern regimes. In concordance with earlier studies, the Southern jet is strongly associated with the NAO–, with ERA20C (the IFS) having mean of \(-0.7 (-0.94)\). The two combined regimes have a weaker association with the NAO+; with ERA20C and the IFS both having a mean of 0.21. Hence, on average, regime dynamics would imply that, if the jet spends \( N \) days in the Southern regime, the DJF mean NAO for that winter would, for ERA20C and the IFS respectively, be

\[
\text{NAO}_{\text{ERA20C}}^{\text{Reg}} = -0.7N + (90 − N) \cdot 0.21 \approx -0.010 \cdot N + 0.21, \tag{3}
\]

\[
\text{NAO}_{\text{IFS}}^{\text{Reg}} = -0.9N + (90 − N) \cdot 0.21 \approx -0.012 \cdot N + 0.21. \tag{4}
\]

On the other hand, directly fitting a linear relationship to the data in Figure 3e,f gives

\[
\text{NAO}_{\text{ERA20C}}^{\text{Fit}} \approx -0.016 \cdot N + 0.32, \tag{5}
\]

\[
\text{NAO}_{\text{IFS}}^{\text{Fit}} \approx -0.020 \cdot N + 0.34. \tag{6}
\]

Because the standard error of the slope coefficients in (5) and (6) is only around 0.002, the mismatch between these two estimates seems robust for both ERA20C and the IFS. Indeed, it is unlikely that jet latitude dynamics alone can account for the full story of the jet and the NAO. One possible explanation for the mismatch is that, in seasons...
where there is a strong forcing of the jet towards the Southern regime, this forcing may also be modulating the speed of the jet, which we know influences the NAO independently of the latitude. Studies have also noted that the jet shows notable intra-seasonal structure which is not fully captured by the Markov model (Franzke et al., 2011; Novak et al., 2015). In any case, we will see that the agreement is good enough to allow regime dynamics alone to account for the skill of the IFS.

Figure 6 shows estimates of the year-to-year distributions of seasonal persistence probabilities for Truncation 2. Hence in (a) is shown the distribution for the Northern regime, and in (b) is shown the same for the combined Southern and Central regimes. As in Strommen and Palmer (2019), the histograms are generated by explicitly estimating a probability for every individual year. These were judged by eye to be well captured by reverse log-normal pdfs and fitted to the data using the standard Python package *scipy*. It can be seen that both regimes show high levels of average persistence, with peaks close to or exceeding 90%. The IFS shows a small but notable tendency towards under-persistence, particularly for the Northern regime. When similar distributions are fitted to each of the three jet latitude regimes individually (not shown), it is similarly found that the IFS is under-persistent for each. This is consistent with the reduced autocorrelation in Figure 4, suggesting that this tendency towards reduced persistence is a robust feature of the IFS. For further analysis of preferred transitions in the jet system, the reader may consult Franzke et al. (2011) and Novak et al. (2015). Because our bimodal system is determined by the two persistence probabilities alone, such preferred transitions in the three-state system will not play a role in this paper.

The above information will allow us to generate seasonal persistence probabilities for ERA20C and link jet latitude regimes to the NAO. The final ingredient is to statistically represent the skill the IFS has at predicting the persistence probabilities of ERA20C, which we now address.

4.4 Representing model skill and parameter fitting

Our statistical model of NAO predictability will proceed as follows. Each winter season we draw two persistence probabilities for ERA20C, corresponding to the two regimes of Truncation 2, using the distributions shown in Figure 6. These are then used to generate a 90-day random walk across the system, which is finally translated, via the the distributions $X_A$ and $X_B$ of Truncation 1, into a DJF NAO mean. To simulate an ensemble hindcast of the IFS, we will also, each season, generate a pair of persistence probabilities for every ensemble member individually. In order to represent the observed IFS skill at predicting the persistence in ERA20C, the probabilities drawn for the ensemble members need to be correlated with the probabilities drawn for ERA20C. As a preliminary step towards this, we need to examine the extent to which the persistence probabilities of the two regimes are independent of each other.

Figure 7 shows how the probabilities (shown as log-odds) of Truncation 2 (i.e., Southern+Central and

![Figure 5](image-url) Estimated histograms of the possible NAO values obtained when the jet is (a) in the combined Central+Northern regime and (b) the Southern regime, for ERA20C (black solid) and IFS (red dashed). This corresponds to Truncation 1 in the text. The values $M$ in the legend denote the mean of the distributions. The full period 1900–2010 is used.
FIGURE 6 Estimated reverse log-normal distributions of seasonal persistence probabilities for (a) the Northern regime and (b) the combined Southern+Central regime, for ERA20C (black solid) and the IFS (red dashed). This corresponds to Truncation 2 in the main text. The values of ‘max’ in the legend denote the peaks of the distributions. The full period 1900–2010 is used.

Northern) are related. In (a), (c) and (e) are shown scatter-plots of the Southern+Central probabilities versus the Northern probabilities for ERA20C, the IFS ensemble members and the IFS ensemble mean respectively. In (b), (d) and (f) are shown the same, but with data restricted to 1980–2010. The value $C$ in each denotes the correlation between the two quantities. The consistent negative sign of the correlation across all data suggests a link between the two probabilities. Part of this may be due to noise introduced from the fact that we always estimate probabilities over only 90 days, which may in some seasons be insufficient to obtain independent estimates of both. However, the fact that the correlation goes up significantly for the IFS ensemble mean, peaking at $\approx -0.6$ in the modern period, is taken as evidence that there is a robust signal here. Such a signal, if it existed, would naturally be amplified considerably by taking a large ensemble mean. As mentioned before, Barnes and Hartmann (2010) showed that eddy feedback is not playing a major role for the Northern regime, suggesting that increased or reduced persistence there may simply be reflecting changes in mid-latitude persistence.

To formalise both the connection between the two probabilities and the IFS skill at predicting both, we utilise the framework of the “signal+noise” method from Siegert et al. (2016) as follows. Note that the following will all be phrased in terms of log-odds of probabilities.

Let $\hat{p}_{\text{obs}}$ be the persistence probability of regime $A$ for ERA20C in a given winter season and $\hat{p}_{\text{mod}}$ the corresponding probability for a single IFS ensemble member. Similarly, let $\hat{q}_{\text{obs}}$ and $\hat{q}_{\text{mod}}$ be the probabilities for regime $B$. Then we will assume that these four probabilities satisfy the following decompositions:

\begin{align}
\hat{p}_{\text{obs}} &= s + \epsilon_A + \mu_{\text{obs},A}, \\
\hat{p}_{\text{mod}} &= \beta_A s + \eta_A + \mu_{\text{mod},A}, \\
\hat{q}_{\text{obs}} &= \lambda s + \epsilon_B + \mu_{\text{obs},B}, \\
\hat{q}_{\text{mod}} &= \beta_B \lambda s + \eta_B + \mu_{\text{mod},B}.
\end{align}

Here the variables $s$, $\epsilon_A$ and $\eta_A$ are normally distributed with

$$s \sim \mathcal{N}(0, \sigma_s), \quad \epsilon_A \sim \mathcal{N}(0, \sigma_{\epsilon_A}), \quad \eta_A \sim \mathcal{N}(0, \sigma_{\eta_A}),$$

with the analogous notation for the $B$ variables. These are seasonally varying parameters, with $s$ representing the predictable component (i.e., signal) of $\hat{p}_{\text{obs}}$, and $\epsilon_A$ the unpredictable component (i.e., noise). We assume that $\hat{q}_{\text{obs}}$ is determined, modulo a noise term $\epsilon_B$, by $s$, via the scaling constant $\lambda$. This will ensure a degree of correlation between the two probabilities, the extent of which is determined by values of the parameters. The probabilities of the model are assumed also to be split into a signal and noise component. For regime $A$, the signal is assumed to equal a scaling factor $\beta_A$ times the true signal $s$. Hence $\beta_A$ is capturing the model’s ability to capture the predictable signal. A value of $\beta_A < 1$ corresponds to the model damping the signal; $\beta_A = 0$ corresponds to no predictability, with $\beta_A = 1$ corresponding to “perfect” skill at forecasting $\hat{p}_{\text{obs}}$. Similarly, the model captures the predictable signal
Figure 7  Scatter plots of the (log-odds of) persistence probabilities of the combined Southern+Central regime against those of the Northern regime: (a) ERA20C (1900–2010), (b) ERA20C (1980–2010), (c) IFS ensemble members (1900–2010), (d) IFS ensemble members (1980–2010), (e) IFS ensemble mean (1900–2010) and (f) IFS ensemble mean (1980–2010). Note that the ensemble mean persistence probability is the mean over all the ensemble member probabilities.

\( \lambda s \) of \( q_{\text{obs}} \) according to a scaling factor \( \beta_B \). Finally, all the \( \mu \) terms simply represent the climatological means of the respective distributions.

Conceptually therefore, we assume that in each DJF season there is precisely one predictable component in the jet system, namely a signal in how persistent the Southern+Central regime is going to be in reality. This completely determines, modulo unpredictable chaos, how persistent the Northern regime is going to be as well. The model is then able to capture the signal in both regimes according to fixed scaling parameters. The ability of the model to make good predictions thus depends on the relative size of the signal \( s \) compared to \( \epsilon \) (i.e., the signal-to-noise ratio of \( p_{\text{obs}} \)), and the relative size of the scaled signal \( \beta s \) compared to \( \eta \) (i.e., the signal-to-noise ratio of \( p_{\text{mod}} \)); similarly for \( \tilde{q} \).

It is clear that this is by no means the only way to represent the system, nor is any good fit of the model parameters unique. However, we will show that this representation, along with a very simple parameter fit, captures the main features of the system. As discussed in Section 2.3, we will fit parameters using the 1980–2010 data only, to ensure that we are not unfairly biasing our results. Parameter fitting is carried out using a “moments estimator” approach (Siegert et al., 2016), in which theoretical formulate of key statistical quantities (means and covariances) for the system defined by Equations (7)–(10) are derived and compared to estimates from the data. The details are technical and can be found in Appendix A of the Supporting Information.

The parameters obtained are listed in Table 1. Using Equations (7)–(10) with these parameters, we can generate a large sample of artificial probabilities. Such a random sample is shown in Figure 8 and should be compared to Figure 7b,d,f. It can be seen that the overall structure is recreated, including the means, variances and correlations, suggesting that the parameter fitting was effective. The parameter values themselves are physically sensible as well. That \( \lambda \approx -1 \) is saying precisely that any forcing which acts to increase the persistence of the Southern+Central regime will also result in an exactly corresponding decrease in the persistence of the Northern regime and vice versa. Both \( \beta \) parameters are between 0 and 1, meaning that the model tends to dampen any forced increase in persistence, consistent with the conclusions of Strommen and Palmer (2019). The noise terms of the IFS
Table 1: Estimated parameters for the signal + noise decompositions of persistence probabilities

| Parameter | Estimate |
|-----------|----------|
| $\sigma_s$ | 0.46     |
| $\sigma_{s_A}$ | 0.55  |
| $\beta_A$ | 0.25     |
| $\sigma_{s_B}$ | 0.62  |
| $\sigma_{\epsilon_B}$ | 0.56  |
| $\sigma_{\eta_B}$ | 0.68  |
| $\beta_B$ | 0.37     |
| $\lambda$ | -0.95    |
| $\mu_{\text{obs}, A}$ | 2.72  |
| $\mu_{\text{obs}, B}$ | 2.11  |
| $\mu_{\text{mod}, A}$ | 2.70  |
| $\mu_{\text{mod}, B}$ | 1.81  |

Note: These values are for log-odds

are slightly larger in amplitude than those of ERA20C, but this discrepancy is small compared to the dampened signal.

As a basic sensitivity test, we verified that the estimated parameter values were virtually unchanged when omitting any one of the 30 years between 1980 and 2010 before doing the fitting. We similarly found little change if fitting was carried out using the longer period 1970–2010. The inclusion of earlier periods began to show evidence of a decrease in skill (manifesting itself as smaller $\beta$ parameters) and so is not considered further.

Finally, by exploiting the fact that the mean of $\eta_A$ is zero, we can estimate the time series of the persistence signal $s$ using the following relationship:

$$\overline{p}_{\text{mod}} - \mu_{\text{mod}, A} \approx \beta_{A s},$$

where the overline denotes the ensemble mean. Dividing the quantity on the left-hand side by our estimated value of $\beta_A$ produces the desired signal, a key ingredient in our hindcast simulations using the Markov model.

4.5 Simulating a seasonal hindcast

We now have all the ingredients necessary for simulating an IFS hindcast of the NAO during the period 1980–2010 using our Markov model. The extent of skill found in such a simulation will give an estimate of how much NAO skill can be accounted for purely through the regime dynamics of the daily jet latitude. The procedure for simulating a single winter season proceeds as follows.

1. Pick the estimated value of $s$ for the year in question, using Equation 11.
2. Draw random values of $\epsilon_A, \epsilon_B, \eta_A, \eta_B$.
3. Use Equations (7)–(10) with all the above selected values to generate persistence probabilities for both ERA20C and the IFS.
4. Initialize both ERA20C and the IFS to start in either the Northern or combined Southern+Central regime. Let both of them perform a 90-day random walk through the Markov chain specified by the persistence probabilities generated in step 3.
5. On days when either ERA20C or the IFS land in the Southern+Central regime, randomly specify it as having landed in the Southern or Central regime according to the relative proportions of the two regimes.
6. In this way, generate 90 days of residency in the full trimodal jet latitude structure.
7. Count the number of days spent in the Southern regime for both ERA20C and the IFS. Use Equations (3) and (4) to generate a DJF NAO mean value for both.

Repeat this procedure 30 times to generate a full 30-year hindcast simulation of the period 1980–2010.
every such simulation, both ERA20C and the IFS ensemble members will have differing persistence probabilities, due to the random component in step 2, but all will have the same underlying signal estimated at the end of the previous section. A schematic of the Markov model is shown in Figure 9 to provide visual assistance to the reader.

In step 4, we note that the conditional probability of the jet being in the Southern regime given that we know it is in the Southern+Central combined regime is \( \approx 33\% \) for both ERA20C and the IFS. Hence when either ERA20C or the IFS land in this combined regime, we specify it as having landed in the Southern regime with that probability.

We would like to emphasise that, while we have endeavoured to fully motivate our choices in the creation of this statistical model, the final results are fairly insensitive to many of the particulars. For example, Equation (3) can be used to generate NAO values for both ERA20C and the IFS with no real change. Furthermore, the use of Equations (3) and (4) themselves assume that the daily impact of the jet latitude on the NAO is always “average” in magnitude. If instead we draw random daily values from the estimated distribution to generate more variability in these daily impacts, the results are still broadly similar, because this variability is mostly averaged away when combining results from multiple hindcast simulations. We chose to use the average values to isolate the impact of jet latitude variations alone. As was the case in Strommen and Palmer (2019), most sensible ways of implementing the Markov model are likely to produce qualitatively similar results.

5 | ESTIMATING NAO SKILL WITH REGIME DYNAMICS

Using the methodology outlined in Section 4.5, we can now simulate a large number of hindcasts and estimate both the expected level of correlation and signal-to-noise ratios (i.e., RPC values). Concretely, we repeat the procedure outlined in the previous section 10,000 times to produce this number of estimates of our desired metrics. These are then used to generate distributions of correlations and RPC values.

Carrying this out gives the following estimates for a 51-member ensemble:

\[
\text{Corr(EnsMean, Obs)} = 0.51 \pm 0.20,
\]
\[
\text{Corr(EnsMean, Mem)} = 0.17 \pm 0.06,
\]
\[
\text{RPC}_{\text{est}} = 2.10 \pm 0.91,
\]

where the uncertainty estimates are two standard deviations around the mean. Recalling the actual values from Section 2.3 of 0.49, 0.19 and 1.80, we see that there is very good qualitative agreement, with all three metrics well within the range of sampling variability. In particular, all the IFS skill at predicting the NAO can be accounted for purely through its ability to predict jet latitude persistence in a regime dynamics system.

Figure 10a shows how the actual predictability and model predictability vary as a function of the ensemble size: the thick lines show estimates using the Markov model and the stipled lines show estimates obtained by a Monte Carlo resampling of the 51 member IFS ensemble. It can be seen that the actual predictability begins to asymptote as the ensemble size grows. If we set the \( \beta \) parameters to 1 (i.e., maximise model skill), the expected ensemble mean correlation with 51 ensemble members is \( \approx 0.62 \pm 0.18 \); increasing the ensemble size to 1,000 does not lead to any increase, implying that the asymptotic limit seen in Figure 10a is the maximum expected correlation that can be achieved purely from jet latitude dynamics. Consequently, the proportion of NAO variance that can be accounted for by jet latitude persistence alone is around 0.62\(^2 \approx 0.38\). Note also that while the Markov model was fitted to the full 51 member ensemble, its scaling to smaller ensemble sizes closely matches that of the IFS resamples. Figure 10b similarly shows how RPC scales with ensemble size. This metric is more sensitive to ensemble size and does not begin to approach the theoretical limit of 0.51/0.16 \( \approx 3.2 \) until the ensemble size approaches 1,000. Increasing the \( \beta \) parameters to 1 and using 1,000 ensemble members brings the RPC down to \( \approx 1.05 \), with the actual predictability at 0.62 (as mentioned) and the model predictability at 0.60. This implies that the low signal-to-noise ratio here is not due to the IFS having noisier regime dynamics \( (\sigma_\epsilon > \sigma_r) \), but is almost entirely accounted for by its damping of persistence \( (\beta < 1) \).

Figure 11 shows a typical hindcast simulation generated with the Markov model. In (a) is shown the

\[
\text{Corr(EnsMean, Obs)} = 0.51 \pm 0.20,
\]
\[
\text{Corr(EnsMean, Mem)} = 0.17 \pm 0.06,
\]
\[
\text{RPC}_{\text{est}} = 2.10 \pm 0.91,
\]
raw hindcast, with the simulated ERA20C (black), IFS ensemble mean (red) and grey dots representing individual ensemble members. In (b) is shown the same simulated ERA20C and IFS ensemble mean but now with both time series normalized to have standard deviation 1. This should be compared to figures 1 and 2 of Dunstone et al. (2016) or figures 8 and 9 in Strommen and Palmer (2019). The ensemble mean correlation and RPC in this case were 0.59 and 2.43 respectively. The ensemble variance is almost identical to that of ERA20C, which is a general feature of the Markov model. Note also the visible presence of negative skew in the ensemble. In fact, this negative skew is a natural feature of the Markov model, owing to the higher climatological mean persistence of the Southern + Central regime compared to the Northern regime. It is well known that the winter NAO is significantly negatively skewed, implying that this skew can be explained naturally through regime dynamics.

While the correlations and RPC values, including their scaling with ensemble size, can also be reproduced with more linear techniques, as in Siegert et al. (2016) for the UK Met Office model, the Markov model has some key additional achievements. Firstly, it fully respects the trimodal structure of the jet, without relying on collapsing this structure by taking DJF means. Secondly, it straightforwardly produces a negatively skewed NAO, which (to our knowledge) cannot be achieved through linear techniques alone without artificially inserting the skew or explicitly imposing some form of state-dependence (e.g., through the inclusion of multiplicative noise; Sura et al., 2005). Finally, by highlighting the role of daily time-scale jet persistence, the Markov model points towards a concrete physical mechanism driving predictability, namely transient baroclinic eddy forcing. This is discussed further in the next section.

6 | EXTERNAL AND INTERNAL DRIVERS OF REGIME VARIABILITY

Having demonstrated that the IFS skill at predicting the winter NAO can be attributed to its ability at predicting interannual variations in the extent of jet persistence, we now examine potential drivers of such interannual variability. First we examine external sources of forcing (teleconnections), before examining local mechanisms in the form of eddy forcing.

Because a comprehensive and in-depth study of predictors of regime dynamics would likely make the present paper excessively long, we keep the discussion brief. We aim to examine this more in future work.

6.1 | External forcing

As discussed in the Introduction, some of the most commonly proposed sources of external NAO forcing are sea-ice in the Kara Sea region, ENSO (Niño 3.4) and the stratosphere, both in the form of the stratospheric polar vortex and the QBO. We will examine the influence of precisely these factors on seasonal persistence probabilities in both the Southern+Central and Northern regimes. We remind the reader that details of how these predictors are defined was included in Section 2.1.
Note that, while in our implementation of the Markov model the entirety of the predictable signal comes from the Southern+Central persistence, this may well be an oversimplification. As mentioned already in Section 4.2, it was shown in Barnes and Hartmann (2010) that, while persistence driven by eddy feedback was a key mechanism in the midlatitudes, variability further north is driven more by Rossby waves. This implies that the influence of external forcing may be different, and hence partially independent, for the two regimes.

Figure 12 shows correlations between the various predictors and persistence probabilities of ERA20C, the IFS ensemble mean and the average across individual IFS ensemble members; for the Northern and the Southern+Central regimes. In the North, the significant predictors appear to be Kara Sea ice and ENSO. The overall largest influence in ERA20C comes from Kara Sea ice, with a correlation of around 0.45, in agreement with previous studies (Wang et al., 2017). More striking is the significant linear correlation, of around 0.4, between ENSO and the Northern persistence probabilities here, which is far from obvious. Indeed, a major obstacle towards a quantitative estimate of the influence of ENSO on the NAO is that, while studies suggest there is a relationship between the two, there is virtually zero linear correlation between the Niño 3.4 time series and the NAO index (e.g. Wang et al., 2017). The significant correlation here suggests that jet latitude regime dynamics may offer a route towards such quantitative estimates of what proportion of NAO variance is explained by ENSO. However, while ENSO is influencing the persistence of ERA20C, the IFS does not appear to capture this at all, with neither the ensemble mean nor individual members showing any significant correlation. We remind the reader that, because the hindcasts considered used prescribed SSTs, the Niño 3.4 index is identical for both ERA20C and all IFS ensemble members. Therefore this seems to be a failure in the propagation of the ENSO signal sufficiently far downstream. The IFS similarly fails to capture the strong link with Kara Sea ice. Neither ERA20C nor the IFS show any notable link between the polar vortex or the QBO with Northern persistence. The eddy forcing (EF) component (rightmost entry of the plots) will be discussed next section.

In the Southern+Central regime, the story is a little different. All the predictors seem to contribute similar levels of predictability, with ERA20C having a correlation of around 0.3 with Kara Sea ice, ENSO and both stratospheric indices. Unlike in the Northern case, the IFS ensemble mean has comparable correlations with all of these except ENSO. This implies that, while ENSO may be a driver of variability in ERA20C, it is playing little or no role in the NAO predictability of the IFS. Instead, Kara Sea ice and the stratosphere seem to be playing more important roles, with the ensemble mean correctly replicating the relationship seen for ERA20C. As before, the eddy forcing term is discussed next section.

This supports the conclusions of earlier studies, namely that Kara Sea ice and the stratosphere are playing a significant role in driving NAO predictability. Our framework suggests that these may act primarily by influencing the extent of jet persistence, particularly in the midlatitudes. There is evidence that ENSO is playing a significant role in regulating persistence in ERA20C, especially (and curiously) in the Northern regime, but this is not replicated by the IFS.
While the size of the errorbars suggests that caution is required, it is interesting to note that the correlations in the two regimes are not perfectly inverse to each other. While the influence of Kara Sea ice and ENSO do naturally anti-correlate across the two (e.g., increased Kara Sea ice increases Northern persistence and simultaneously decreases Southern+Central persistence), this is not the case for the two stratospheric indices. Instead, the stratosphere seems only to modulate persistence in the Southern+Central regime, with no notable effect in the north. This again lends support to the idea that the two regimes are, to an extent, modulated independently of each other. We observe that this provides a potential quantitative way to estimate nonlinear influences of teleconnections on the NAO, using the fact that the DJF mean NAO is a nonlinear function of both persistence probabilities.

Finally, we briefly address the question of dependence amongst the different predictors. This is to some extent mitigated by the choice of which time period our predictors are considered over. In particular, our ENSO index, when restricted to November months, has almost zero correlation with both November Kara Sea ice, the November polar vortex or the DJF mean QBO index. As such, these November-only predictors are independent of the November-December-January ENSO index used. Studies have also suggested a link between Kara Sea ice and the stratospheric circulation (Cohen et al., 2013; Kim et al., 2014) and indeed we find positive correlations between November sea-ice and the November polar vortex and DJF mean QBO. However, if we subtract away the sea-ice signal from the stratospheric indices using linear regression, their correlations with persistence probabilities do not appreciably change. Therefore the predictors considered appear to be broadly independent, insofar as their impact on jet persistence is considered.

6.2 | Local eddy forcing

There is by now an extensive literature on the role of transient baroclinic eddy feedback in maintaining jet persistence (e.g., Hoskins et al. (1983); Shutts (1983); Robinson (1996); Lorenz and Hartmann (2001); Lorenz and Hartmann (2003); Zurita-Gotor et al. (2014); Novak et al. (2015) and references therein) as well as for the NAO more directly (Barnes and Hartmann, 2010). It is therefore natural to consider the extent to which eddy forcing is modulating the persistence probabilities central to the Markov model. To do so, we compute the daily eddy momentum flux convergence of 250 hPa ERA20C winds:

$$E_{250} := \frac{\partial (-u''_{250} v''_{250})}{\partial y}$$

where $u''_{250}$ ($v''_{250}$) is the zonal (meridional), 2–6 day bandpass filtered wind field at 250 hPa. Positive values of this quantity correspond to regions where the transient eddies are accelerating the westerlies (Hoskins et al., 1983). Lorenz and Hartmann (2003) showed that the eddies generally act to reinforce the jet through a positive feedback but, as noted in Woollings et al. (2010b), this feedback is not limited to the three jet regimes and is present at all latitudes. In particular, we verified that when restricting to days when the jet is classified as being...
in the Southern+Central regime (that is, at latitudes below 52°N), one finds a notable overlap between the zonal wind anomalies and the eddy forcing field, as expected (Figure S2).

This is used to motivate the following metric, aimed at measuring the extent to which the eddies are reinforcing the jet in the Southern+Central regime. For a given winter season, consisting of 90 days, we take the dot product of the wind anomalies and eddy flux convergence, restricted to the Southern+Central region only. The mean across all 90 daily dot products is then defined as the strength of the eddy forcing in the Southern+Central regime in this winter:

$$\text{EF}_{\text{South}} = \left\langle E_{250} \cdot u'_{850} \right\rangle,$$

where $u'_{850}$ is the 850 hPa daily wind anomaly, $\left\langle \cdot \right\rangle$ denotes the DJF mean and all vectors have been restricted to the latitudes 15–52°N. The daily dot product will be positive on days when the eddies are strongly reinforcing the jet. Figure S2 implies that this will, on average, be the case and hence this metric would be expected to correlate with measures of seasonal jet persistence. While an analogous metric $\text{EF}_{\text{North}}$ for the Northern regime could be considered, the results of Barnes and Hartmann (2010) suggest that eddy feedback is not playing a major role in the higher latitudes. We verified that $\text{EF}_{\text{North}}$ was on average an order of magnitude smaller than $\text{EF}_{\text{South}}$, corroborating this reduced influence of eddies. For these reasons, we will consider only the $\text{EF}_{\text{South}}$ as a possible predictor of persistence.

The correlations between $\text{EF}_{\text{South}}$ and the persistence probabilities in the two regimes is shown in Figure 12, confirming our intuitive expectations. For both ERA20C and the IFS ensemble mean, the correlation in the case of the Southern+Central regime is around 0.62, significantly higher than that of any of the external predictors considered. While all these external predictors do enjoy non-zero correlations with the eddy forcing, if we subtract the influence of all these from the eddy forcing (by performing a multiple linear regression) the correlation between the forcing the Southern+Central persistence probabilities still remains high at around 0.45. Note that this is still larger than that of all the external predictors, implying that, while some of the eddy forcing is externally forced, a considerable amount of it may be either locally driven or unforced and essentially chaotic.

In the Northern regime, the eddy forcing has a smaller, negative correlation with seasonal persistence of around −0.3. This is consistent with the analysis of Section 4.4, which suggested that changes in the persistence characteristics of the Northern regime are merely a response to changes in the south.

7 | CONCLUSIONS

By truncating the trimodal jet into a bimodal system, we constructed a statistical model (the Markov model) based on capturing regime dynamics as in Strommen and Palmer (2019). In this system, the position of the jet is modelled using a two-state Markov chain which is driven by seasonal persistence probabilities, with the two states representing the Northern regime and the combined Southern+Central jet regimes. These persistence probabilities, assumed constant during a given winter but which otherwise vary from year to year, determine which of the two regimes the jet is likely to spend more time in. Because the average NAO conditions differ for the two regimes, with the more southern regime being associated with negative NAO values and the more northern regime associated with positive values, such interannual variations in the preferred jet regime translates into interannual variations in the NAO. Hence any knowledge of the expected persistence time-scales of the two regimes yields predictive power for the winter NAO.

We then showed that the IFS has robust skill at predicting these seasonal persistence probabilities over the modern period 1980–2010. This justified fitting the Markov model parameters using reanalysis and IFS model data, the latter in the form of a 51-member ensemble hindcast. The fitted model allowed us to estimate what the expected IFS ensemble mean correlation with the true NAO index would be if predictability were driven entirely by regime dynamics. Since this estimate, of 0.51 ± 0.20, is close to the observed correlation of 0.49, we concluded that all of the skill of the IFS may be coming purely from such regime dynamics. Because the Markov model does not include any contribution from the jet speed, this conclusion is consistent with that of Parker et al. (2019), who used linear regression techniques to attribute the IFS skill to the jet latitude. However, while linear techniques first requires one to collapse the trimodal structure of the jet by averaging, the Markov model respects the nonlinear structure inherent to the jet in a novel manner. A consequence of this is that the Markov model naturally produces a negatively skewed NAO, in accordance with observations. By highlighting the role of weak model persistence, this model also suggests a possible explanation of the “signal-to-noise paradox”. Thinking of regimes as potential wells, the weak persistence of the IFS may be a result of these potential wells being too shallow. As a result, any external forcing of the jet which attempts to push the jet preferentially towards either regime would not elicit as strong a response in the model as it ought to; the jet of the model is simply ejected from the relevant regime too swiftly.
While our statistical model captures predictability associated with transitions in the jet latitude, we omitted any contribution from the jet speed, which is likely overly simplistic given that studies suggest external drivers like ENSO can influence jet speeds (Drouard et al., 2013). In addition, earlier studies suggest that there is further intra-seasonal structure in jet variability, such as preferred transitions (Franzke et al., 2011; Novak et al., 2015), which are only indirectly represented in the Markov model. There is also evidence that thermal processes can play a crucial role in driving jet transitions (Novak et al., 2015), particularly for the southern regime (Madonna et al., 2019). As this thermal influence is qualitatively different from the eddy-driven forcing typically associated with higher levels of jet persistence, it is possible that this is a source of skill not represented in the Markov model. Representing the full trimodal nature of the link between the jet and the NAO on daily time-scales is also something to explore further, though some care will be required for this.

Finally, we considered both local and external drivers of variability in jet regime persistence. We found that Kara Sea ice, ENSO, the QBO and the stratospheric polar vortex, commonly cited as important drivers of NAO variability, all correlate with the estimated persistence probabilities from reanalysis. This suggests that the role of these teleconnections may be, at least in part, to modulate jet persistence rather than the NAO more directly. In the IFS, several of these external teleconnections were weaker than in reanalysis, with some being virtually non-existent. In particular, the IFS does not replicate the observed connection between ENSO and jet persistence, suggesting that ENSO may have very limited influence on skill. Locally we showed that a measure of the strength of the transient baroclinic eddy forcing also projects strongly onto the two persistence probabilities, particularly that of the southern regime. It appears that a substantial part of this eddy forcing is independent of the external drivers we considered, implying that much of the observed NAO predictability may not be easily attributable to remote teleconnections and may rather come from more complex local dynamics. This suggests that eddy feedbacks play a crucial role in modulating NAO predictability, and models that fail to produce realistic levels of feedback may struggle to achieve high NAO skill even with accurate teleconnections. In Scaife et al. (2019), by comparing seasonal predictions at varying horizontal resolutions, it was found that the strength of the eddy forcing on the NAO was systematically too weak, with no appreciable change until the resolution approached $\approx 15$ km grid-boxes. The importance of horizontal resolution may be necessary to eliminate the signal-to-noise paradox.

We end our discussion with two questions. Firstly, the results of this paper were based on the IFS in particular. In Baker et al. (2018) it was found that models can exhibit a broad range of ensemble mean correlations and RPC values. Can this variation across models be accounted for by their differing abilities at capturing regime persistence? Secondly, in the Markov model, predictability is explicitly induced by year-to-year variations in jet persistence, which is strongly modulated by local eddy forcing. Should the role of more remote teleconnections be understood, not through their direct influence on the NAO, but by their indirect influence on jet persistence via eddy forcing?

**ACKNOWLEDGEMENTS**

This work benefited from conversations with Tim Woollings, Tim Palmer, Fenwick Cooper, Marie Drouard and Peter Watson. Much of the work was carried out with funding from the European Commission under grant agreement 641727 of the Horizon 2020 research programme. The author was also funded by a Junior Research Fellowship at Jesus College, Oxford, during the writing of the manuscript. Processed IFS and ERA20C wind data were generously shared with me by Tess Parker, processed QBO data by Chris O’Reilly, and processed Kara Sea ice data by Fenwick Cooper. Finally, comments from two anonymous reviewers were of great help for clarifying several points.

**ORCID**

Kristian Strommen © https://orcid.org/0000-0001-5411-8109

**REFERENCES**

Anstey, J.A., Davini, P., Gray, L.J., Woollings, T.J., Butchart, N., Cagnazzo, C., Christiansen, B., Hardiman, S.C., Osprey, S.M. and Yang, S. (2013) Multi-model analysis of Northern Hemisphere winter blocking: model biases and the role of resolution. *Journal of Geophysical Research: Atmospheres*, 118, 3956–3971.

AThanasiadis, P.J., Wallace, J.M. and Wettstein, J.I. (2010) Patterns of wintertime jet stream variability and their relation to the storm tracks. *Journal of the Atmospheric Sciences*, 67, 1361–1381.

Athanasiadis, P.J., Bellucci, A., Scaife, A.A., Hermanson, L., Materia, S., Sanna, A., Borrelli, A., MacLachlan, C. and Gualdi, S. (2017) A multisystem view of wintertime NAO seasonal predictions. *Journal of Climate*, 30, 1461–1475.

Baker, L.H., Shaffrey, L.C., Sutton, R.T., Weisheimer, A. and 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

Baldwin, M.P. and Dunkerton, T.J. (2001) Stratospheric harbingers of anomalous weather regimes. *Science*, 294, 581–584. https://doi.org/10.1126/science.1063315
Barnes, E.A. and Hartmann, D.L. (2010) Dynamical feedbacks and the persistence of the NAO. *Journal of the Atmospheric Sciences*, 67, 851–865.

Brönnimann, S. (2007) Impact of El Niño–Southern Oscillation on European climate. *Reviews of Geophysics*, 45. https://doi.org/10.1029/2006RG000199

Charlton-Perez, A.J., Ferranti, L. and Lee, R.W. (2018) The influence of the stratospheric state on North Atlantic weather regimes. *Quarterly Journal of the Royal Meteorological Society*, 144, 1140–1151. https://doi.org/10.1002/qj.3280

Cohen, J., Jones, J., Furtado, J.C. and Tziperman, E. (2013) Warm Arctic, cold continents: a common pattern related to Arctic sea ice melt, snow advance, and extreme winter weather. *Oceanography*, 26. https://doi.org/10.5670/oceanog.2013.70

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

Dawson, A., Palmer, T.N. and Corti, S. (2012) Simulating regime structures in weather and climate prediction models. *Geophysical Research Letters*, 39(21). https://doi.org/10.1029/2012GL053284

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balsomeda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragnani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Holm, E.V., Isaksen, L., Källberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597.

Dong, B.W., Sutton, R.T., Jewson, S.P., O’Neill, A. and Slingo, J.M. (2000) Predictable winter climate in the North Atlantic sector during the 1997–1999 ENSO cycle. *Geophysical Research Letters*, 27, 985–988. https://doi.org/10.1029/1999GL010994

Drouard, M., Rivière, G. and Arbogast, P. (2013) The North Atlantic Oscillation response to large-scale atmospheric anomalies in the northeastern Pacific. *Journal of the Atmospheric Sciences*, 70, 2854–2874.

Dunstone, N., Smith, D., Scaife, A.A., Hermanson, L., Eade, R., Robinson, N., Andrews, M. and Knight, J. (2016) Skillful predictions of the winter North Atlantic Oscillation one year ahead. *Nature Geoscience*, 9, 809–814.

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

Franzke, C., Woollings, T. and Martius, O. (2011) Persistent circulation and preferred regime transitions in the North Atlantic. *Journal of the Atmospheric Sciences*, 68, 2809–2825. https://doi.org/10.1175/JAS-D-11-046.1

Fyfe, J.C. and Lorenz, D.J. (2005) Characterizing midlatitude jet variability: Lessons from a simple GCM. *Journal of Climate*, 18, 3400–3404.

Gastineau, G., D’Andrea, F. and Frankignoul, C. (2013) Atmospheric response to the North Atlantic Ocean variability on seasonal to decadal time scales. *Climate Dynamics*, 40, 2311–2330. https://doi.org/10.1007/s00382-012-1333-0

Hoskins, B.J., James, I.N. and White, G.H. (1983) The shape, propagation and mean-flow interaction of large-scale weather systems. *Journal of the Atmospheric Sciences*, 40, 1595–1612.

Jiménez-Esteve, B. and Domeisen, D.I. (2018) The tropospheric pathway of the ENSO–North Atlantic teleconnection. *Journal of Climate*, 31, 4563–4584.

Kidston, J., Scaife, A.A., Hardiman, S.C., Mitchell, D.M., Butchart, N., Baldwin, M.P. and Gray, L.J. (2015) Stratospheric influence on tropospheric jet streams, storm tracks and surface weather. *Nature Geoscience*, 8, 433–440. https://doi.org/10.1038/ngeo2424

Kim, B.-M., Son, S.-W., Min, S.-K., Jeong, J.-H., Kim, S.-J., Zhang, X., Shim, T. and Yoon, J.-H. (2014) Weakening of the stratospheric polar vortex by Arctic sea-ice loss. *Nature Communications*, 5(4646). https://doi.org/10.1038/ncomms5646

Koenigk, T., Caian, M., Nikulin, G. and Schimanke, S. (2016) Regional Arctic sea ice variations as predictor for winter climate conditions. *Climate Dynamics*, 46, 317–337. https://doi.org/10.1007/s00382-015-2586-1

Kumar, A., Peng, P. and Chen, M. (2014) Is there a relationship between potential and actual skill?. *Monthly Weather Review*, 142, 2220–2227. https://doi.org/10.1175/MWR-D-13-00287.1

Li, Y. and Lau, N.C. (2012) Impact of ENSO on the atmospheric variability over the North Atlantic in late winter – role of transient eddies. *Journal of Climate*, 25, 320–342.

Lorenz, D.J. and Hartmann, D.L. (2001) Eddy–zonal flow feedback in the Southern Hemisphere. *Journal of the Atmospheric Sciences*, 58, 3312–3327.

Lorenz, D.J. and Hartmann, D.L. (2003) Eddy–zonal flow feedback in the Northern Hemisphere winter. *Journal of Climate*, 16, 1212–1240.

Madonna, E., Li, C., Grams, C.M. and Woollings, T. (2017) The link between eddy-driven jet variability and weather regimes in the North Atlantic–European sector. *Quarterly Journal of the Royal Meteorological Society*, 143, 2960–2972. https://doi.org/10.1002/qj.3155

Madonna, E., Li, C. and Wettstein, J.J. (2019) Suppressed eddy driving during southward excursions of the North Atlantic jet on synoptic to seasonal time-scales. *Atmospheric Science Letters*, 20, e937. https://doi.org/10.1002/asl.937

Mehta, V.M., Suarez, M.J., Manganello, J.V. and Delworth, T.L. (2000) Oceanic influence on the North Atlantic Oscillation and associated Northern Hemisphere climate variations: 1959–1993. *Geophysical Research Letters*, 27, 121–124.

Michelangeli, P.-A., Vautard, R. and Legras, B. (1995) Weather regimes: recurrence and quasi-stationarity. *Journal of the Atmospheric Sciences*, 52, 1237–1256.

Monahan, A.H. and Fyfe, J.C. (2006) On the nature of zonal jet EOFs. *Journal of Climate*, 19, 6409–6424.

Moron, V. and Gourand, I. (2003) Seasonal modulation of the El Niño–Southern Oscillation relationship with sea level pressure anomalies over the North Atlantic in October–March 1873–1996. *International Journal of Climatology*, 23, 143–155. https://doi.org/10.1002/joc.868

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

Novak, L., Ambaum, M.H. and Tailleux, R. (2015) The life cycle of the North Atlantic storm track. *Journal of the Atmospheric Sciences*, 72, 821–833.
O'Reilly, C.H., Heatley, J., MacLeod, D., Weisheimer, A., Palmer, T.N., Schaller, N. and 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

O'Reilly, C.H., Weisheimer, A., Woollings, T., Gray, L.I. and MacLeod, D. (2019) The importance of stratospheric initial conditions for winter North Atlantic Oscillation predictability and implications for the signal-to-noise paradox. *Quarterly Journal of the Royal Meteorological Society*, 145, 131–146. https://doi.org/10.1002/qj.3413

Palmer, T.N. (1999) A nonlinear dynamical perspective on climate prediction. *Journal of Climate*, 12, 575–591.

Parker, T., Woollings, T., Weisheimer, A., O'Reilly, C., Baker, L. and Shaffrey, L. (2019) Seasonal predictability of the winter North Atlantic Oscillation from a jet stream perspective. *Geophysical Research Letters*, 46(16). https://doi.org/10.1029/2019GL084402

Poli, P., Hersbach, H., Tan, D.G.H., Dee, D.P., Thépaut, J.-N., Simmons, A., Peubey, C., Laloyaux, P., Komor, T., Berrisford, P., Dragnani, R., Trémolet, Y., Hölm, E.V., Bonavita, M., Isaksen, L. and Fisher, M. (2013). The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th Century assimilating surface observations only (ERA-20C). ERA report 14, ECMWF, Reading, UK.

Rayner, N.A., Parker, D.E., Horton, E.B., Folland, C.K., Alexander, L.V., Rowell, D.P., Kent, E.C. and Kaplan, A. (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres*, 108(D14). https://doi.org/10.1029/2002JD002670

Robinson, W.A. (1996) Does eddy feedback sustain variability in the zonal index?. *Journal of the Atmospheric Sciences*, 53, 3556–3569.

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

Rogers, J.C. (1997) North Atlantic storm track variability and its association to the North Atlantic oscillation and climate variability of Northern Europe. *Journal of Climate*, 10, 1635–1647.

Saunders, M.A. and Qian, B. (2002) Seasonal predictability of the winter NAO from North Atlantic sea surface temperatures. *Geophysical Research Letters*, 29(22). https://doi.org/10.1029/2002GL014952

Scaife, A.A. and Smith, D.M. (2018) A signal-to-noise paradox in climate science. *npj Climate and Atmospheric Science*, 1(28). https://doi.org/10.1038/s41612-018-0038-4

Scaife, A.A., Ineson, S., Knight, J.R., Gray, L., Kodera, K. and Smith, D.M. (2013) A mechanism for lagged North Atlantic climate response to solar variability. *Geophysical Research Letters*, 40, 434–439. https://doi.org/10.1002/grl.50099

Scaife, A.A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R.T., Dunstone, N., Eade, R., Fereday, D., Folland, C.K., Gordon, M., Hermanson, L., Knight, J.R., Lea, D.J., MacLachlan, C., Maidens, A., Martin, M., Peterson, A.K., Smith, D.M., Vellinga, M., Wallace, E., Waters, J. and Williams, A. (2014a) Skilful long-range prediction of European and North American winters. *Geophysical Research Letters*, 41, 2514–2519.

Scaife, A.A., Athanassiadou, M., Andrews, M., Arribas, A., Baldwin, M., Dunstone, N., Knight, J., MacLachlan, C., Manzini, E., Müller, W.A., Pohlmann, H., Smith, D.M., Stockdale, T. and Williams, A. (2014b) Predictability of the quasi-biennial oscillation and its northern winter teleconnection on seasonal to decadal timescales. *Geophysical Research Letters*, 41, 1752–1758. https://doi.org/10.1002/2013GL059160

Scaife, A.A., Camp, J., Comer, R., Davis, P., Dunstone, N., Gordon, M., MacLachlan, C., Martin, N., Nie, Y., Ren, H.L., Roberts, M., Robinson, W., Smith, D.M. and Vidale, P.L. (2019) Does increased atmospheric resolution improve seasonal climate predictions?. *Atmospheric Science Letters*, 20(e922). https://doi.org/10.1002/asl.922

Shotts, G.J. (1983) The propagation of eddies in diffusent jetstreams: Eddy vorticity forcing of ‘blocking’ flow fields. *Quarterly Journal of the Royal Meteorological Society*, 109, 737–761. https://doi.org/10.1002/qj.49710946204

Siegent, S., Stephenson, D.B., Sansom, P.G., Scaife, A.A., Eade, R. and Arribas, A. (2016) A Bayesian framework for verification and recalibration of ensemble forecasts: How uncertain is NAO predictability?. *Journal of Climate*, 29, 995–1012.

Smith, D.M., Scaife, A.A., Eade, R. and Knight, J.R. (2016) Seasonal to decadal prediction of the winter North Atlantic Oscillation: Emerging capability and future prospects. *Quarterly Journal of the Royal Meteorological Society*, 142, 611–617.

Strommen, K. and Palmer, T.N. (2019) Signal and noise in regime systems: A hypothesis on the predictability of the North Atlantic Oscillation. *Quarterly Journal of the Royal Meteorological Society*, 145, 147–163. https://doi.org/10.1002/qj.3414

Strommen, K., Mavilia, I., Corti, S., Matsueda, M., Davini, P., von Hardenberg, J., Vidale, P.L. and Mizuta, R. (2019) The sensitivity of Euro-Atlantic regimes to model horizontal resolution. *Geophysical Research Letters*, 46, 7810–7818. https://doi.org/10.1029/2019GL082843

Strong, C. and Magnusdottir, G. (2011) Dependence of NAO variability on coupling with sea ice. *Climate Dynamics*, 36, 1681–1689. https://doi.org/10.1007/s00382-010-0752-z

Sura, P., Newman, M., Penland, C. and Sardeshmukh, P. (2005) Multiplicative noise and non-Gaussianity: A paradigm for atmospheric regimes?. *Journal of the Atmospheric Sciences*, 62, 1391–1409.

Thompson, D.W.J., Lee, S. and Baldwin, M.P. (2003). Atmospheric processes governing the Northern Hemisphere Annular Mode/North Atlantic Oscillation, pp. 81–112 in The North Atlantic Oscillation: Climatic Significance and Environmental Impact. Hurrell, J.W., Kushnir, Y., Ottersen, G., Visbeck, M. (eds). American Geophysical Union, Washington DC.

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

Vautard, R. (1990) Multiple weather regimes over the North Atlantic: Analysis of precursors and successors. *Monthly Weather Review*, 118, 2056–2081.

Wang, L., Ting, M. and Kushner, P.J. (2017) A robust empirical seasonal prediction of winter NAO and surface climate. *Scientific Reports*, 7(279). https://doi.org/10.1038/s41598-017-00353-y

Weisheimer, A., Schaller, N., O'Reilly, C., MacLeod, D.A. and Palmer, T.N. (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
White, R.H., Hilgenbrink, C. and Sheshadri, A. (2019) The importance of Greenland in setting the northern preferred position of the North Atlantic eddy-driven jet. *Geophysical Research Letters*, 46, 14126–14134. https://doi.org/10.1029/2019GL084780

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

Woollings, T., Charlton-Perez, A., Ineson, S., Marshall, A.G. and Masato, G. (2010a) Associations between stratospheric variability and tropospheric blocking. *Journal of Geophysical Research: Atmospheres*, 115(D6). https://doi.org/10.1029/2009JD012742

Woollings, T., Hannachi, A. and Hoskins, B.J. (2010b) Variability of the North Atlantic eddy-driven jet stream. *Quarterly Journal of the Royal Meteorological Society*, 136, 856–868.

Zurita-Gotor, P., Blanco-Fuentes, J. and Gerber, E.P. (2014) The impact of baroclinic eddy feedback on the persistence of jet variability in the two-layer model. *Journal of the Atmospheric Sciences*, 71, 410–429.

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Strommen K. Jet latitude regimes and the predictability of the North Atlantic Oscillation. *Q J R Meteorol Soc*. 2020;146:2368–2391. https://doi.org/10.1002/qj.3796