https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

Predictable Decadal Forcing of the North Atlantic Jet Stream by Sub-Polar North Atlantic Sea Surface Temperatures

Kristian Strommen[a], Tim Woollings[a], Paolo Davini[b], Paolo Ruggieri[c], and Isla R. Simpson[d]

[a]Department of Physics, University of Oxford
[b]Istituto di Scienze dell'Atmosfera e del Clima, Consiglio Nazionale delle Ricerche (CNR-ISAC), Torino
[c]Department of Physics and Astronomy, University of Bologna
[d]Climate and Global Dynamics Laboratory, National Centre for Atmospheric Research, Boulder CO

Correspondence: Kristian Strommen (kristian.strommen@physics.ox.ac.uk)

Abstract. It has been demonstrated that decadal variations in the North Atlantic Oscillation (NAO) can be predicted by current forecast models. While Atlantic Multidecadal Variability (AMV) in sea surface temperatures (SSTs) has been hypothesised as the source of this skill, the validity of this hypothesis and the pathways involved remain unclear. We show, using reanalysis and data from two forecast models, that the decadal predictability of the NAO can be entirely accounted for by the predictability of decadal variations in the speed of the North Atlantic eddy-driven jet, with no predictability of decadal variations in the jet latitude. The sub-polar North Atlantic (SPNA) is identified as the only potential source of an SST-based signal across the models and reanalysis, and the predictability of the jet speed is shown to be consistent with a forcing from the SPNA visible already within a single season. The pathway is argued to be tropospheric in nature, with the SPNA-induced heating extending up to the mid-troposphere, which alters the meridional temperature gradient around the climatological jet position. The link between SSTs and heatfluxes, which mediates the predictability, is shown to be underestimated in the forecast models by approximately a factor of two, with potential implications for the 'signal-to-noise paradox'. The relative roles of anthropogenic aerosol emissions and the AMOC at generating predictable SPNA variability are also discussed. The analysis is extensively supported by the novel use of a set of seasonal hindcasts spanning the 20th century and forced with prescribed SSTs.

1 Introduction

European winter weather is strongly influenced by the variability of North Atlantic eddy-driven jet, and it is therefore of high societal value to predict this variability as far in advance as possible. Recent studies have shown that, remarkably, retrospective ensemble forecasts ('hindcasts') are now able to skillfully predict some components of the low-frequency variability of the winter jet at lead times of up to 10 years (Smith et al., 2019; Athanasiadis et al., 2020). However, the exact source of the predictable signal and mechanisms involved remain unclear, making it uncertain to what extent one can rely on this skill to remain for genuine decadal forecasts of the future. The aim of this paper is to try to clarify these points.

In Simpson et al. (2018), by considering a jet index based on zonal winds at 700hPa, it was shown that the decadal variability of the jet is much stronger in March than in the boreal winter months December, January, February (DJF). This late winter jet variability was argued to arise from the internally generated component of Atlantic Multidecadal Variability (AMV) in

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

North Atlantic sea surface temperatures (SSTs), though the mechanisms were not elucidated; they also showed that models failed to capture the connection. In alignment with the enhanced decadal variability, the observed connection between the AMV and the jet was shown to be far greater in March than during DJF. Nevertheless, Smith et al. (2019) and Athanasiadis et al. (2020) showed that skillful decadal forecasts of the DJF-averaged North Atlantic Oscillation (NAO) could be achieved, and the latter suggested that the skill appeared to be driven by the AMV. Both Smith et al. (2019) and Athanasiadis et al. (2020) emphasised an apparent 'signal-to-noise paradox' in these forecasts, mimicking the phenomenon observed for seasonal winter NAO forecasts (Scaife and Smith, 2018). This 'paradox' effectively says that the real world appears to be much more predictable than the forecast model thinks it is, with the model underestimating the response to forcing or the response to predictable boundary conditions on seasonal-to-decadal timescales. A practical consequence of this behaviour is that models are likely underestimating the predictable component of decadal winter jet variability, suggesting that the results of Simpson et al. (2018) and Athanasiadis et al. (2020) are consistent with a hypothesis that the AMV is driving predictable decadal jet[1] variability from December through March. In fact, several studies have found that even the total decadal variability appears to be systematically underestimated in models (Kravtsov, 2017; Wang et al., 2017; Kim et al., 2018; Simpson et al., 2018; Bracegirdle, 2022).

Numerous studies have been conducted on the potential for air-sea coupling to generate links between the AMV and the jet, starting with the pioneering work of Bjerknes (1964). The decadal variability of the AMV itself has been hypothesised to be driven by a combination of the Atlantic Meridional Overturning Circulation (AMOC) (Bjerknes, 1964; Delworth et al., 1993; Kushnir, 1994), anthropogenic aerosols and other greenhouse gases (Booth et al., 2012; Bellomo et al., 2018; Robson et al., 2022), and stochastic atmospheric forcing (Hasselmann, 1976; Clement et al., 2015; O'Reilly et al., 2019). An excellent recent overview on these topics with more complete references can be found in Zhang et al. (2019). Different mechanisms have been put forward for how the AMV affects the jet, including the direct modulation of low-level baroclinicity and stationary waves by the North Atlantic SST anomalies (Kushnir, 1994; Msadek et al., 2011; Kushnir et al., 2002; Peings et al., 2016); forcing from the tropical Atlantic (Davini et al., 2015); and stratospheric pathways (Omrani et al., 2014). However, the response in climate models to imposed AMV anomalies appears inconsistent and model dependent (Ruggieri et al., 2021), and the period of highly reliable observational data is short, making it challenging to distinguish between different hypotheses.

One major source of uncertainty in and across many studies is that the decadal variability is an average over several different processes occurring on different timescales, due to (a) the continuous nature of air-sea coupling and (b) the fact that the AMV pattern itself evolves over time, with the anomalies in the sub-polar North Atlantic (SPNA) arising first before propagating towards the tropical Atlantic (Zhang et al., 2019; Wills et al., 2019). This makes it difficult to attribute causality between atmospheric versus oceanic forcing, and also to distinguish between the role played by particular regions in the Atlantic ocean, such as the sub-polar versus tropical North Atlantic. However, several studies have emphasised the importance of the sub-polar gyre in particular (Gastineau and Frankignoul, 2015; Woollings et al., 2015; Ortega et al., 2017; Wills et al., 2019), especially on longer timescales (Delworth et al., 2017).

In this paper, we make crucial use of two techniques to help address these challenges:

[1]Since the NAO is largely describing the variability of the jet, we will conflate these without comment for the remainder of the paper.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

1. The separation of the eddy-driven jet into two components, corresponding to the speed and latitude of the jet.

2. The use of two seasonal hindcast ensembles, named ASF20C and CSF20C, spanning the period 1900-2010. ASF20C is forced with prescribed, observed boundary conditions (Weisheimer et al., 2017), while CSF20C is fully coupled (Weisheimer et al., 2020) (more details in Section 2.1).

The first point is motivated by the fact that the variability and sensitivity of the latitude and speed of the jet are very different. The jet latitude exhibits multimodality and considerable variability on seasonal timescales, but shows no significant variability on decadal timescales beyond white noise (Woollings et al., 2010, 2014). The jet speed, on the other hand, is approximately Gaussian across all timescales and exhibits robust decadal variability (Woollings et al., 2014). Furthermore, Baker et al. (2017) showed that the latitude and speed respond differently to thermal forcing, and Woollings et al. (2015) showed that the processes responsible for latitudinal shifts in the jet clearly differ from those responsible for changes to the jet speed. This means that analysis based on single indices which amalgamate the latitude and speed (such as the NAO index) may struggle to identify robust links between the jet and SST anomalies.

To motivate the second point, we note that existing decadal forecasts only go back to 1954 at the earliest, leaving them with a relatively small effective sample size once any low-pass filtering or decadal averaging has taken place. The period 1954 to present also does not adequately sample the variability associated with the AMV and the AMOC. There is therefore great value in extending the analysis back to 1900. While taking decadal averages of an atmosphere-only seasonal hindcast obviously does not constitute an actual decadal *forecast*, it nevertheless turns out to be useful to think of it as being like a 'nudged' forecast, where both the atmospheric and oceanic state are being nudged back towards observations at the start of each winter (and moreover, for ASF20C, the SSTs are being forecasted perfectly). We will show that in fact the decadal variability reproduced by the seasonal hindcasts completely matches the predictable decadal variability of a genuine decadal forecast ensemble, justifying this perspective post hoc. This not only allows us to confidently use ASF20C/CSF20C to extend our analysis back to 1900, but also introduces two considerable benefits: the lack of coupling in ASF20C simplifies the question of causality, and the fact that the forecasts making up ASF20C/CSF20C only cover a single season simplifies the question of timescales.

We will show that on decadal timescales there is no predictability of the latitude of the jet, and that all the observed skill at predicting the winter NAO arises from the predictability of the *speed* of the jet. By comparing observations with the ASF20C/CSF20C seasonal hindcasts and the Decadal Prediction Large Ensemble (DPLE) decadal forecasts (Yeager et al., 2018), we argue that predictable forcing of the jet speed arises from SST anomalies in the SPNA. Furthermore, we argue that the predictable forcing occurring on decadal timescales arises as the accumulation of a smaller forcing taking place already on seasonal timescales, with no need to invoke mechanisms spanning multiple seasons as in several other studies. Finally, we argue that the response of the jet speed to SST anomalies in the SPNA can be understood simply as the adjustment of the jet to changes in the tropospheric meridional temperature gradient across the climatological jet core.

The structure of the paper is as follows. The data and methods are described in Section 2. In Section 3 we assess the predictability of the NAO, jet speed and jet latitude in the models, while Section 4 is devoted to identifying potential sources of skill from SSTs. Section 5 addresses the pathways, mechanisms and the 'signal-to-noise paradox', while Section 6 considers

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

the relative roles of aerosols, the AMOC and stochastic atmospheric forcing at generating predictable decadal variability in the SPNA. The results are finally summarised and discussed in Section 7.

2 Data and Methods

2.1 Data

2.1.1 ERA20C

We use the 20th century reanalysis dataset ERA20C, spanning 1890-2010, as our observational 'truth' (Poli et al., 2013). This reanalysis is constructed using a cycle of the Integrated Forecast System (IFS) forecast model. Due to difference in available observations between the beginning and end of the 20th century, the atmospheric component of ERA20C only assimilates surface pressure, in order to maintain coherence over the whole period. Both ocean and sea-ice boundary conditions come from the HadISST2.1.0.0 dataset (Rayner et al., 2003). Although it is known that the internal variability is underestimated in the early 20th century (Dell'Aquila et al., 2016), ERA20C constitutes a reasonable reference for the status of the North Atlantic climate.

2.1.2 ASF20C and CSF20C

The ASF20C model data considered comes from an atmosphere-only seasonal hindcast experiment covering the 20th century (Weisheimer et al., 2017). A 51 ensemble member seasonal forecast is initialised every 1st of November from 1901 to 2010 and allowed to run for 4 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 TL255, corresponding to roughly 80km grid spacing near the equator, with 91 levels in the vertical. The model is run in atmosphere-only mode with prescribed observed sea-surface temperatures (SSTs) with boundary conditions and initial conditions from ERA20C. Further details can be found in Weisheimer et al. (2017).

In addition, we will sometimes make use of the CSF20C hindcast as well. This hindcast is identical to ASF20C except it is run with dynamic coupling between the atmosphere and ice/ocean, and is initialised using the coupled reanalysis CERA20C (Laloyaux et al., 2018). The configuration is described in Weisheimer et al. (2020) and is similar to the SEAS5 operational seasonal forecast at the European Centre for Medium-range Weather Forecasts (Johnson et al., 2019). The ocean model used is NEMO version 3.4 (Madec and the NEMO team, 2016), and the ice model is LIM2 (Fichefet and Maqueda, 1997): both are run at a 1 degree horizontal resolution.

2.1.3 Decadal Prediction Large Ensemble (DPLE)

DPLE is made up of a suite of 40-member ensemble forecasts, each initialised on November 1st and run for 10 years. Forecasts are initialised every year from 1954 to 2015. To be consistent with CSF20C/ERA20C we restrict to the overlapping period 1954-2010. When assessing the decadal forecast skill of DPLE, we always take averages over the entire 10-year period. For

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

example, suppose we have a timeseries J made up of DJF averages of some quantity in reanalysis, and we want to assess the capacity of DPLE to predict $\overline{J}$, where overline denotes the average across the 10-year period 1954-1964. Then the DPLE forecast of $\overline{J}$ is taken to be the ensemble mean over $\overline{x_k}(k = 1, \ldots, 40)$, where the x_k are the 10-year forecasts initialised on November 1st 1954, with k referring to ensemble member. By doing this for consecutive 10-year periods we obtain estimates of the decadal variability predicted by DPLE which we can correlate with the observed decadal variability.

Importantly, we do not perform any drift-correction of any kind in our analysis. The main goal of this paper is to understand how the atmosphere responds to decadally varying SSTs, and this can be assessed in DPLE irrespective of any drift taking place. Furthermore, it is not customary to de-drift seasonal forecasts, so no drift correction is done for ASF20C/CSF20C: not de-drifting DPLE therefore makes the analysis more directly comparable between the forecast products. For DPLE we will only ever consider two timescales: the response taking place in the first season or the 10-year mean across the whole forecast period. The drift taking place in the former is small, and the imprint of the drift in the latter is smoothed out by the large averaging-window. Finally, we note that Athanasiadis et al. (2020) found that robust decadal NAO forecast skill still exists in DPLE if drift-correction is carried out.

DPLE is run using CESM version 1.1, using the same model and component configuration as that used in the CESM-LE (Kay et al., 2015). The atmosphere component is version 5 of the Community Atmosphere Model (CAM5: Hurrell et al. (2013)), with a horizontal resolution of around 1 degree and 30 levels in the vertical. The ocean component is version 2 of the Parallel Ocean Program (Danabasoglu et al., 2012) and the sea ice model is version 4 of the Los Alamos National Laboratory (LANL) Community Ice Code (Hunke et al., 2010). Both are run at a 1 degree spatial resolution. Further details can be found in Yeager et al. (2018).

2.1.4 CMIP data

In order to assess the mechanisms, we will be examining intermodel spread in the mean state of the jet across a number of climate model simulations. We analyse data drawn from the 5th (CMIP5) and 6th (CMIP6) phases of the coupled model intercomparison project. We analyse the historical experiments for 31 single-member CMIP6 models (Eyring et al., 2016) detailed in Table A1 of Appendix A, and a total of 71 ensemble members from 28 distinct CMIP5 models (Taylor et al., 2012) detailed in Table A2. These historical experiments consist of coupled uninitialised climate runs forced with historical greenhouse gas and aerosol forcings over the 20th century, after a spin-up from a free-running pre-industrial control run.

To further boost the sample size, we also make use of model data produced as part of the PRIMAVERA project (Roberts et al., 2018b), detailed in Table A3. These coupled simulations all follow the HighResMIP protocol (Haarsma et al., 2016), and are therefore initialised in 1950, following a short 50-year spin-up. The simulations span the 65 years between 1950 and 2015, and use the same historical forcings as CMIP6. Six underlying models were used, each run at a number of different resolutions: CMCC-CM2 (Cherchi et al., 2019), CNRM-CM6 (Voldoire et al., 2019), EC-Earth3 (Haarsma et al., 2020), ECMWF-IFS (Roberts et al., 2018a), HadGEM3-GC31 (Williams et al., 2018), MPI-ESM1-2 (Gutjahr et al., 2019), and AWI-CM-1.0 (Sein et al., 2017).

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

Note that the model data is the same as that used in Dorrington et al. (2022), and the tables of this paper have been reproduced from there with permissions. In total, the climate model data used consists of more than 8400 DJF seasons generated from 76 different climate models.

In addition to the historical CMIP6 data, we make use of two piControl CMIP6 (500-year and 603-year long respectively) integrations from EC-Earth3 (Döscher et al., 2022). These simulations are atmosphere-ocean coupled runs with pre-industrial forcings. Such simulations are specifically interesting owing to their large internal variability induced by an internally-driven centennial oscillation of the AMOC (Meccia et al., 2022).

2.2 Methods

2.2.1 Metrics

The NAO index is computed as the leading EOF of daily DJF 500hPa geopotential height anomalies in the Euro-Atlantic sector 30-90N, 80W-40E. A seasonal cycle is estimated by taking the average daily NAO value, for each day in DJF, across all years available. This cycle is then removed, along with any linear trend. When computing the NAO index for ensemble forecast data, all the ensemble member data is used to compute the EOF, after which each members geopotential height field is projected onto the EOF pattern to obtain the individual NAO indices.

To compute indices of the jet speed and latitude, we follow the simplified methodology of Parker et al. (2019). Wind fields are first interpolated to a regular 1 degree grid. Daily DJF 850hPa zonal winds are then restricted to the region 15-75N, 60-20W and smoothed with a 5-day running mean. For any given day, the jet is said to be located at the latitude where the magnitude of the zonally averaged winds in this region peaks. The jet latitude on that day is precisely this latitude, while the jet speed is the magnitude of the maximum. As with the NAO, a seasonal cycle and linear trend is removed.

An 'SPNA' index measuring SST variability in the sub-polar North Atlantic is also used extensively. This is defined as the DJF-averaged SSTs, averaged over the region 49-57N, 50-25W. The motivation for this precise choice is given in the main text. The results are not sensitive to small shifts in this region.

2.2.2 Statistics

Our default stance on significance testing is to explicitly specify a statistical model representing the null hypothesis, and then generate confidence intervals by making 10,000 random draws using the model. Because the motivation behind the choice of each statistical model depends on the situation at hand, we introduce each such model in the main text as and when it is required.

In order to have a second set of confidence intervals independent of the authors own statistical models, we make use of the 'Fourier Phase Shuffling' method (Ebisuzaki, 1997). This method can be loosely described as follows. First compute the Fourier transform of the timeseries of interest. Secondly, for each Fourier mode, replace the computed phase by a randomly chosen one. Third, convert the resulting Fourier series back to a timeseries. The resulting randomly generated timeseries is guaranteed to have the same autocorrelation (at all lags) and degrees of freedom as the original timeseries. Confidence intervals generated

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

by drawing 10,000 such random 'phase shuffles' can therefore be viewed as testing a null hypothesis where the timeseries is random with a given autocorrelation structure. In particular, this null hypothesis is typically much stricter than one using an AR1 fit, which only specifies the autocorrelation at lag 1. This is especially the case for timeseries based on SSTs, which often carry memory well beyond a single season.

Finally, for all questions of statistical significance, in cases where ASF20C/CSF20C and DPLE seem to differ in some regard, we always choose to place higher credence in ASF20C/CSF20C, due to the fact that (a) their sample size is approximately twice as large ($N = 109$ vs $N = 56$); and (b) they have more ensemble members (51 vs 40).

3 Predictable jet variability in ASF20C and DPLE

3.1 Predictability of the jet speed and not the jet latitude

We first examine what low-frequency jet variability is skillfully reproduced by ASF20C. Figure 1 shows 10 and 30-year running means of the NAO, jet latitude and jet speed for ERA20C and the ASF20C ensemble mean. Note that the ASF20C ensemble mean standard deviation is considerably smaller than in ERA20C due to the 'signal-to-noise paradox': the timeseries have therefore been normalised to have mean 0 and standard deviation 1, to allow for an easier visual comparison. It can be seen that ASF20C skillfully reproduces decadal NAO variability across the entire period 1900-2010, with a correlation coefficient of $\sim 0.4 - 0.5$ depending on the choice of smoothing: using the raw seasonal data gives a correlation of 0.34. The 10/30-year correlations closely match those reported in Smith et al. (2019) and Athanasiadis et al. (2020) using genuine decadal forecasts, suggesting that the decadal forecast skill they reported can likely be extended all the way back to 1900. Figure 1 also clearly shows that ASF20C cannot skillfully reproduce decadal jet latitude variability, but *can* skillfully reproduce decadal jet speed variability. Furthermore, if we regress out the (decadally averaged) ensemble mean jet speed from the ensemble mean NAO and correlate the residual with the observed NAO, we obtain ≈ 0.1 using 10-year averages and ≈ -0.1 using 30-year averages. Therefore, all of the skill that ASF20C has at reproducing decadal NAO variability can be accounted for by the jet speed. Figure B1 in Appendix B shows that, similarly, the coupled hindcast CSF20C has significant skill at predicting decadal variations in the speed but not the latitude.

Figure 2 shows that the same conclusion is true for the DPLE forecasts: there is no apparent predictability of decadal shifts in the jet latitude, but high skill at predicting shifts in the jet speed. Note that taking 30-year means is much less sensible for DPLE, given its shorter coverage of 56 years, but these are included anyway for direct comparison with Figure 1. The relatively large negative correlation emerging for 30-year jet latitude variability is likely to just be noise.

How significant are the jet speed correlations reported in Figures 1 and 2? The interannual lag-1 correlation is approximately 0 for all three data sets, and we are therefore justified in assuming a null hypothesis of the DJF jet speed as being random draws from a fitted Gaussian distribution. The approximate Gaussianity of the jet speed distribution has been previously noted (Woollings et al., 2010; Parker et al., 2019). By simulating 10,000 randomly generated DJF jet speed timeseries for each data set, taking 10 or 30-year running means and then computing correlations, we build up a distribution of correlations that can be obtained by chance. Note that taking running means will automatically introduce considerable autocorrelation to the resulting

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

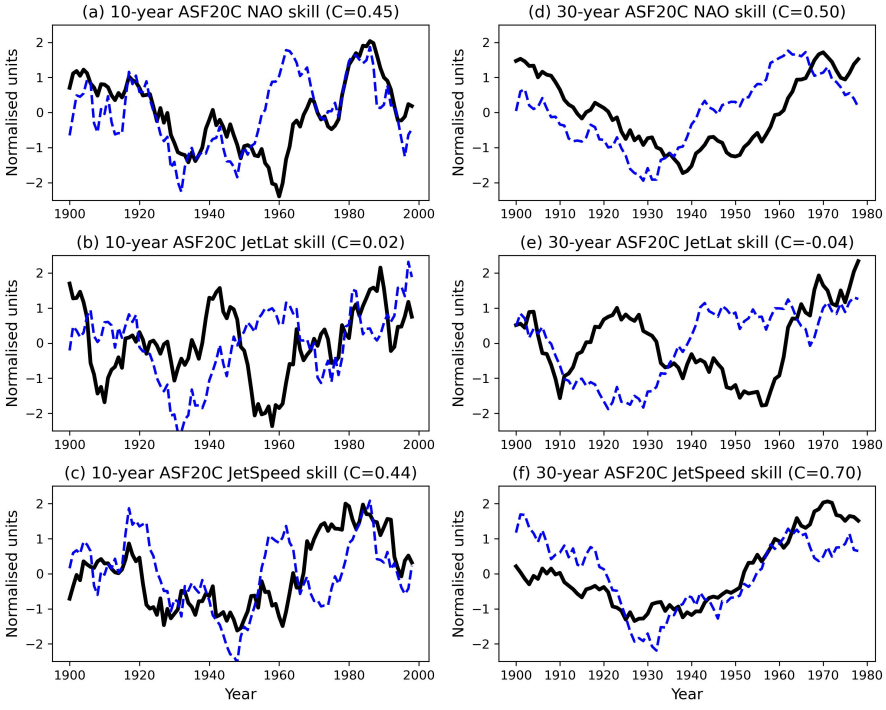

Figure 1. Timeseries of 10-year running DJF means of (a) the NAO, (b) the jet latitude, and (c) the jet speed. The same but with 30-year running means in (d), (e) and (f). The thick black curves are always ERA20C and the dashed blue curves are always the ASF20C ensemble mean. The timeseries have been normalised to have mean 0 and standard deviation 1, for visualisation purposes. The value C in each subplot is the correlation between the two timeseries.

timeseries. Table 1 summarises the result of this significance testing. This shows that for ASF20C, whether considering 10 or 30 year smoothing, the decadal jet speed correlations are only significant to within $\rho < 0.1$ and not $\rho < 0.05$: the same is true for CSF20C (see Figure B1). For DPLE, the conclusion is the same except that the correlations using 30-year smoothing appear to be significant also with $\rho < 0.05$. However, as cautioned already, the small effective sample size here means that this level of confidence is probably not robust. It goes without saying that the jet latitude correlations are never significant.

Our results here contrast with those of Smith et al. (2019) and Athanasiadis et al. (2020), which both report statistically significant decadal NAO forecasts with $\rho < 0.05$. The fact that their significance tests differ from ours and consider the NAO rather than the jet speed will at least partly explain this. The smoothing they use is also very close, but not identical to, the 10-year running means we use. Table 1 shows that the 10-year ASF20C and DPLE jet speed correlations sit neatly between the bounds of the 90 and 95% confidence interval, and it is therefore plausible that small shifts in the null hypothesis used could shift ρ to below 0.05.

For this article, we will assume that ASF20C and DPLE really do have significant skill at predicting decadal variations in the jet speed. We believe this assumption is strengthened by the analysis we will conduct later, which shows a clear source and

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

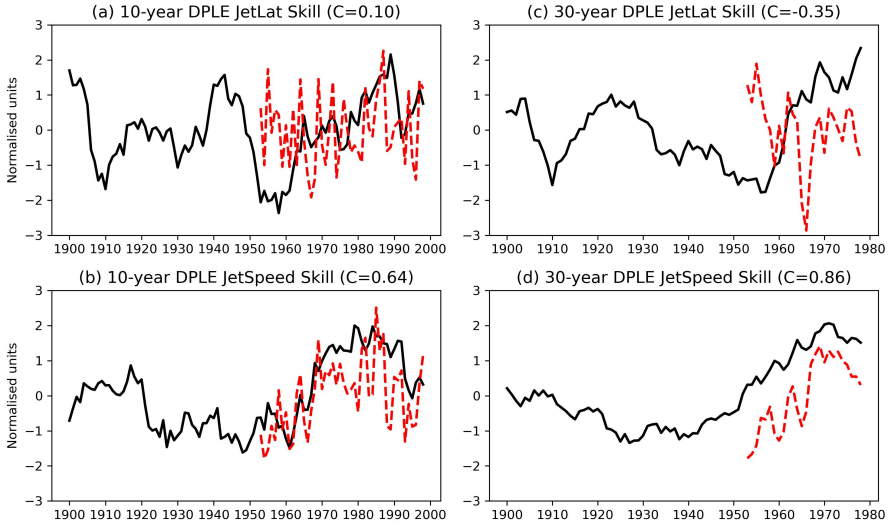

Figure 2. Timeseries of 10-year running DJF means of (a) the jet speed, and (b) the jet latitude. The same but with 30-year running means in (c) and (d). The thick black curves are always ERA20C and the dashed red curves are always the DPLE ensemble mean. The timeseries have been normalised to have mean 0 and standard deviation 1, for visualisation purposes. The value C in each subplot is the correlation between the two timeseries.

JetSpeed correlations	10 year smoothing	30 year smoothing
Corr(ERA20C, ASF20C)	0.44	0.70
90% confidence interval	±0.42	±0.70
95% confidence interval	±0.48	±0.80
Corr(ERA20C, DPLE)	0.64	0.86
90% confidence interval	±0.62	±0.80
95% confidence interval	±0.71	±0.85

Table 1. DJF jet speed correlations between different data sets at different levels of smoothing. Confidence intervals assume a null hypothesis of the interannual jet speed timeseries being uncorrelated Gaussian noise.

pathway for this skill, thereby justifying a high prior for such skill *post hoc*. The 95% significance found in Smith et al. (2019) and Athanasiadis et al. (2020), the 90% significance we find here, and the consistency of our analysis across two independent forecast models can then be used to justify updating this prior to give strong overall confidence in the skill.

3.2 Treating low-frequency variability in ASF20C and DPLE as equivalent

The similar behaviour of ASF20C and DPLE suggests that the extra information provided to ASF20C (the correct climate state every November 1st and the correct SSTs at all times) is not adding any extra jet speed skill beyond what is expected from

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

an actual decadal forecast. We interpret this as strong evidence for the assertion that the decadal timescale signals responsible for predictable shifts in the jet speed are fully represented in ASF20C. We will therefore consider ASF20C and DPLE as 'equivalent', in the sense that we can study the sources of skill and mechanisms using both data sets in a directly comparable manner. The same conclusion is drawn for CSF20C. Possible alternative explanations for the similar low-frequency variability are discussed in the Discussion and Conclusions.

The assertion that ASF20C captures the decadal forecast signals has two important consequences. Firstly, since each of the 109 ASF20C winter forecasts only explicitly knows about the initial conditions and boundary forcings of the season in question, the sources of predictable decadal jet speed forcing must be present in their entirety within a single winter season. In particular, the predictable atmospheric signals, and the mechanisms involved, do not require multi-year lags as some studies have suggested they do (Peings and Magnusdottir, 2014b; Kwon et al., 2020). The timeseries of initial conditions or boundary forcing may of course contain multi-annual memory, but any given initialised ASF20C atmospheric forecast has no direct knowledge of this. Secondly, since the ASF20C forecasts are uncoupled, the forcing of the jet does not essentially depend upon atmosphere-ocean coupling, in the sense that no within-season-feedbacks are necessary. We will make use of these two points repeatedly to simplify the analysis and reasoning.

We emphasise straight away that while surface coupling appears to not be required to reproduce the low frequency jet speed variability, missing or deficient coupling may play a role in why the ensemble mean signal is so small (the 'signal-to-noise paradox'). This point is returned to in the discussion of Section 7.

4 Sources of predictability from sea surface temperatures

4.1 Why we restrict attention to sea surface temperatures only

The troposphere broadly speaking has a decorrelation timescale of about 2 weeks, and this is also true for the daily jet speed index (not shown). This strongly suggests that the skill in both ASF20C and DPLE cannot be explained by the persistence of atmospheric initial conditions. It is possible in principle that the skill comes from the stratosphere (Omrani et al., 2014), which has a decorrelation timescale of several months. However, the initial conditions of ASF20C (CSF20C) come from ERA20C (CERA20C), which only assimilates surface variables. Because these do not strongly constrain the stratosphere, the stratospheric variability in ERA20C is highly biased, and there is consequently limited scope for skill from the stratospheric initial conditions in ASF20C. The bias in stratospheric initial conditions of ASF20C and the visible impact of these on forecast skill has been examined previously (O'Reilly et al., 2019). We therefore choose to not consider potential stratospheric sources of skill in our analysis.

Given this restriction, the skill is necessarily generated by a slow timescale forcing from the boundary conditions, i.e. the SSTs, ice and anthropogenic emissions such as carbon dioxide and aerosols, and natural forcings such as volcanic aerosol and solar variability. The impact of ice on decadal jet variability has been found to be small compared to SSTs (Peings and Magnusdottir, 2016), leading us to exclude ice from our analysis. As for anthropogenic emissions, the most obvious way for these to alter the jet is *indirectly*, by altering the radiative forcing and hence surface temperatures, either globally (as in the

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

case of carbon) or locally (e.g. in the case of sulphate aerosols). Such a forcing on the jet would manifest itself as an apparent forcing from surface temperatures. While this could potentially include forcing from temperatures over land, the decorrelation timescale over land is much faster than for SSTs, making a decadal timescale forcing seem less likely. We therefore ultimately restrict our attention to searching for sources of skill in the SSTs. The role of aerosols are discussed again in Section 6.1, including the possibility of a more direct impact of aerosols on the jet by altering the local radiative fluxes in the troposphere itself (as opposed to at the surface).

4.2 Sub-polar North Atlantic SSTs as a common signal across observations and forecasts

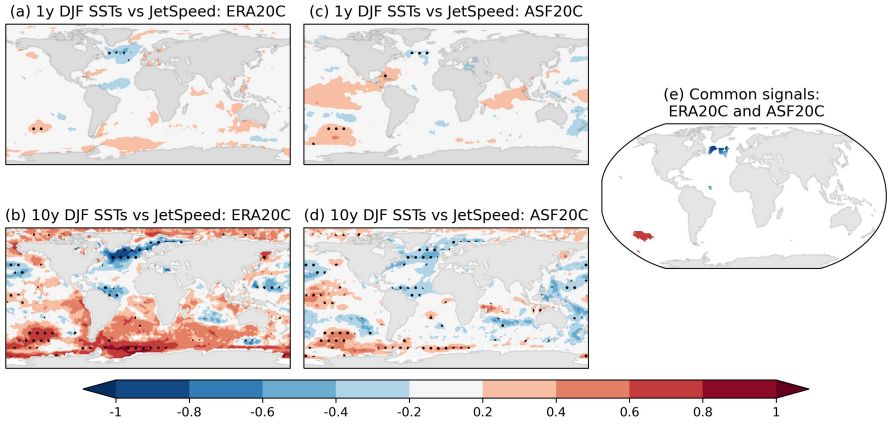

Figure 3. Correlations between the DJF JetSpeed timeseries and DJF SSTs at each gridpoint. In (a) for ERA20C using seasonal data; (b) for ERA20C with a 10-year running mean applied; (c) for ASF20C using seasonal data; (d) for the ASF20C ensemble mean with a 10-year running mean applied. In (e) are shown the correlations from (b) that are common across all subplots and significant in each subplot by itself. Stipling indicates significance using a two-tailed t-test. The period 1900-2010 is used.

In order to locate potential sources of skill in SSTs, we compute correlations of the winter jet speed against winter SSTs at every gridpoint. Because of our conclusion that these sources must be visible already on seasonal timescales (see Section 3.2), we do this using both the raw DJF timeseries as well as timeseries obtained by applying a 10-year running mean. Sources of skill are then assumed to correspond to regions of significant correlations that are common to ERA20C, ASF20C, CSF20C and DPLE across both interannual and decadal timescales. While it is of course possible that the locations of signals differs somewhat across the datasets (e.g. due to mean state biases in the forecast models), the simplest - and more physically sound - possibility is that the signals are identical across the datasets, so we consider this possibility first.

Figure 3 shows gridpoint correlations for ERA20C and ASF20C, at interannual (top row) and 10-year (bottom row) timescales. The significance test used at each gridpoint (shown with stipling) is that the correlation is statistically significant ($\rho < 0.05$) using a simple two-tailed t-test; for ASF20C we further stipulate that gridpoint correlations must also be significant in ERA20C, and that the sign of the correlation is the same as that of ERA20C. This ensures that only genuinely common signals are

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

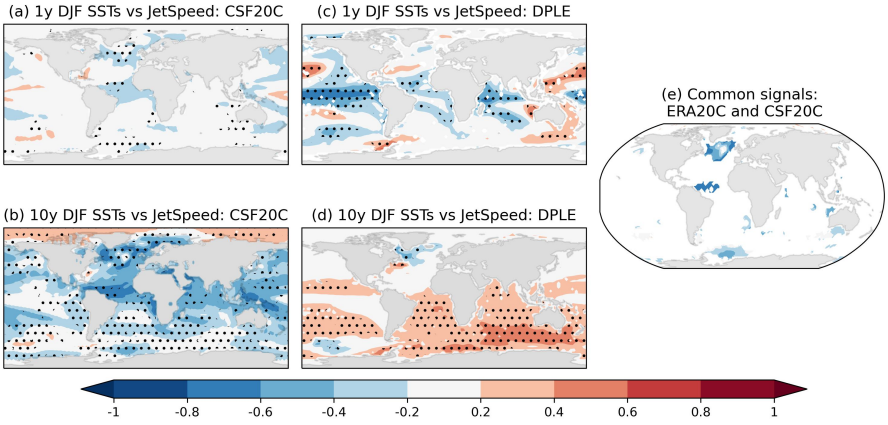

Figure 4. Correlations between the DJF JetSpeed timeseries and DJF SSTs at each gridpoint. In (a) for CSF20C using seasonal data; (b) for the CSF20C ensemble mean with a 10-year running mean applied; (c) for DPLE using seasonal data; (d) for the DPLE ensemble mean with a 10-year running mean applied. In (e) are shown the correlations from (b) that are common across subplots (a) and (b), as well as subplots (a) and (b) of Figure 3, and significant in each of these four subplots by themselves. Stipling indicates significance using a two-tailed t-test. The period 1900-2010 is used.

290 highlighted. Note that the two-tailed t-test is not a sufficiently robust test of significance here, especially when using 10-year smoothing, and is only used as a 'first pass' test: more robust tests will then be applied to any noteworthy regions. Figure 3(e) shows only those 10-year correlations from (b) that are significant and common across all panels (a), (b), (c) and (d). This immediately suggests that there are only two potential signals common to ERA20C and ASF20C on interannual-to-decadal timescales: the SPNA, and an isolated patch in the south Pacific.

295 Figure 4 shows the gridpoint correlations for CSF20C and DPLE, with panel (e) showing 10-year correlations common across ERA20C and CSF20C. The significance testing is similar, with correlations in CSF20C only deemed significant if they are significant in and of themselves and of the same sign as in ERA20C. Note that the south Pacific patch has now essentially vanished. Comparing Figure 3(e) with Figure 4(e) therefore implies that the SPNA is the only location where significant correlations can be found on interannual-to-decadal timescales for ERA20C, ASF20C and CSF20C simultaneously. Inspection
300 of the signals in DPLE shows that the SPNA is also clearly highlighted on 10-year timescales (Figure 4(d)). On interannual timescales (Figure 4(c)), the signal from the SPNA is evidently weaker and mostly confined to the eastern side. However, because the sample size of ERA20C, ASF20C and CSF20C is almost twice that of DPLE, we consider it justified to weight these datasets higher than DPLE, and suggest that the weaker interannual DPLE signal is simply due to the reduced sample size.

305 To summarise the above discussion, the SPNA is the only plausible common region of interannual-to-decadal correlations between SSTs and the jet speed across ERA20C, ASF20C, CSF20C and DPLE. Crucially, the coupled DPLE forecasts have skill at predicting the decadal SST variability in the SPNA region (Yeager et al., 2018; Yeager, 2020). We therefore define

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

an SPNA timeseries as the DJF averaged SSTs across the domain 49-57N, 50-25W. This box was chosen because it encloses almost precisely the negative correlations highlighted in Figure 3(e). The remainder of the analysis we carry out is not sensitive to small shifts in the definition of this box. The signal in the tropical Atlantic visible in some of the subplots of Figures 3 and 4 will be discussed again in Section 7.

It is worth mentioning the phenomenon of false discovery rates, which would a priori be expected to be high when considering correlations across thousands of gridpoints with high spatial autocorrelation. However, based on the results and discussion of previous sections, we are now assuming that (a) there is decadal forecast skill which needs to be explained and (b) the source of this skill is most likely coming from SSTs. A rejection of any and all gridpoint correlations as 'false discoveries' would be directly counter to this assumption, and is therefore not done. The discovery of a region of correlations common across 4 data sets on the other hand can be seen as providing additional evidence towards our assumption.

4.3 Significance of the SPNA-JetSpeed link and the question of causality

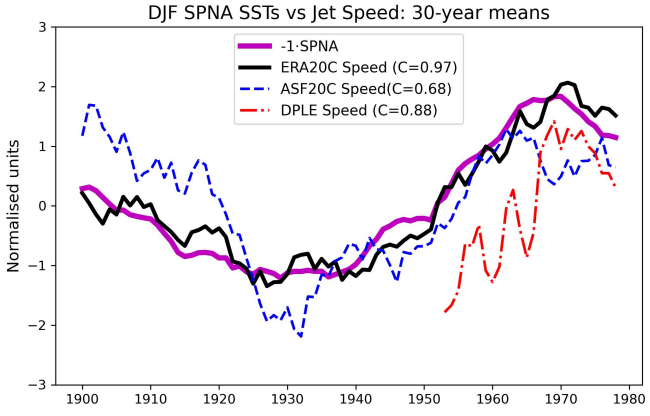

Figure 5. Timeseries of 30-year running means of ERA20C SPNA SSTs (thick purple: the sign has been flipped for visual convenience), ERA20C jet speeds (thick black), ASF20C ensemble mean jet speeds (dashed blue) and DPLE ensemble mean jet speeds (dashed red). The values of C in the legend indicate the correlation between the jet speed timeseries and the relevant SPNA timeseries (i.e., SSTs are drawn from ERA20C in the case of the ERA20C and ASF20C jet speeds, and from DPLE in the case of DPLE jet speeds). The timeseries have been normalised to have mean 0 and standard deviation 1.

Table 2 summarises the correlations between SPNA SSTs and jet speed for the different datasets across different timescales. Figure 5 visualises the 30-year running mean timeseries: the sign of the SST index has been flipped for visual convenience. Two 95% confidence intervals for ERA20C/ASF20C are reported in Table 2, based on two different null hypotheses. The first, labelled 'Simple Model', models the interannual SPNA timeseries with an AR1 process (with a lag of 1 year), and the interannual jet speed as a normal distribution as before: these are fitted to the dataset in question and 10,000 randomly drawn timeseries are generated. Correlations are then computed using the raw timeseries as well as timeseries obtained by 10 and

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

SPNA vs JetSpeed	Raw	10 year smoothing	30 year smoothing
ERA20C	**-0.34**	**-0.80**	**-0.95**
ASF20C	-0.19	-0.35	-0.67
CSF20C	**-0.27**	**-0.57**	-0.70
95% CI - Simple Model	±0.19	±0.51	±0.78
95% CI - Phase Shuffle	±0.27	±0.76	±0.88
DPLE	-0.13	-0.28	**-0.88**
95% CI - Simple Model	±0.29	±0.72	±0.90
95% CI - Phase Shuffle	±0.35	±0.90	±0.85

Table 2. Correlations between DJF averaged SPNA SSTs and DJF averaged jet speed at different levels of smoothing. The null hypotheses used to generate the two confidence intervals are explained in the text. Correlations significant using the 'Simple Model' null hypothesis are highlighted in bold.

325 30-year smoothing, in order to produce the distribution of null hypothesis correlations. The second, labelled 'Phase Shuffle', uses the more severe Fourier Phase Shuffle method discussed in Section 2.2. Note that the confidence intervals for ERA20C, ASF20C and CSF20C are almost identical, so we use their average intervals as a common confidence interval for all three. Note also that the confidence intervals differ a lot for DPLE due to its sample size being around half that of the other three datasets.

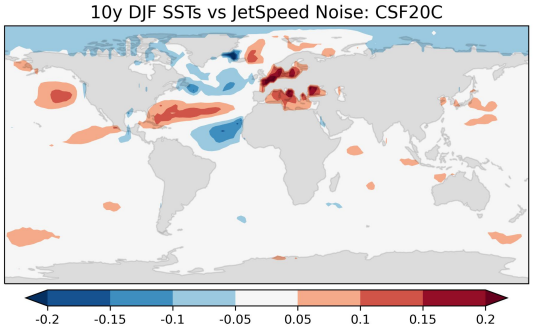

Figure 6. Correlations between the 'stochastic noise' component of the 10-year averaged DJF CSF20C jet speed and 10-year averaged DJF SSTs at each gridpoint. The period 1900-2010 is used. Note the different colorscale compared to Figures 3 and 4.

330 The correlation between SPNA and the jet speed is highly significant in ERA20C for all timescales considered here, against both null hypotheses considered, suggesting a robust physical link between these quantities. The correlations obtained by ASF20C and DPLE, while considerable, are not statistically significant with $\rho < 0.05$, though the interannual correlations in ASF20C become significant against a weaker test of $\rho < 0.1$ (not shown). The correlations in CSF20C are significant against the 'Simple Model' for interannual and 10-year timescales, but not against the 'Phase Shuffle' test. The DPLE correlations

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

are significant at 30-year smoothing, but, as previously noted, this amount of smoothing is unlikely to be sensible here[2]. Of course, the lack of consistent statistical significance does not mean the physical link in ERA20C is not simulated by the forecast models, only that these correlations would not be sufficient to establish such a link when viewed in isolation. However, the presence of jet speed skill in the forecasts, the clear significance of the correlations in ERA20C, and the narrowing down of potential signals to the SPNA can all be used to justify updating the prior probability that such a link exists. This is further strengthened by the way CMIP models independently corroborate the existence of such a link, discussed in Section 5.

There is an obvious question concerning the causality of such an SPNA-JetSpeed link, especially since SSTs in the SPNA are known to be particularly sensitive to forcing from the jet (Visbeck et al., 2003; Barrier et al., 2014; Ma et al., 2020). Are the correlations in Table 2 in fact purely indicative of stochastic atmospheric forcing on the ocean? In response to this we offer three points which suggest that at least a considerable portion of the correlations are due to a genuine forcing from the ocean to the atmosphere. The related question of the role of stochastic forcing in driving SPNA SST variability is considered in Section 6.3.

Firstly, we remind the reader that ASF20C is uncoupled and uses prescribed boundary forcing, so correlations on interannual timescales in ASF20C necessarily imply causality from the SSTs to the atmosphere and *not* vice versa. While this is not true for CSF20C and DPLE, which are coupled, the fact that the correlations they attain are entirely comparable to those of ASF20C suggests that forcing from the SSTs to the atmosphere constitutes a pragmatic common explanation for the correlations found in all three forecast models.

Secondly, positive SST anomalies in the SPNA are associated with positive heatflux anomalies, i.e. a flux from the ocean to the atmosphere: this is shown and discussed in Section 5.4 (see Figure 12). If the SPNA was just passively responding to the jet, then the sign of the heatflux would be expected to be negative, since a weaker jet should transfer less heat into the ocean. The positive sign is therefore a clear indication of forcing from the ocean.

Thirdly, quantitative estimates suggest the stochastic atmospheric forcing is weak on decadal timescales. One such estimate can be done as follows. Let x_k denote the 10-year averaged jet speed timeseries for each of the k CSF20C ensemble members ($k = 1, \ldots, 50$). We can decompose x_k as $x_k = s + \epsilon_k$, where s is the ensemble mean and ϵ is the residual for each member. Here the ϵ_k can be viewed as different instances of stochastic atmospheric variability, with s representing a forced signal (e.g. from SSTs). We then subtract the ensemble mean from each x_k and concatenate the resulting ϵ_k's back to back for all k. By correlating this timeseries with the CSF20C SSTs at each gridpoint (where ensemble members are again concatenated back to back in the same order), we get an estimate of the link between the noise terms and the SSTs, shown in Figure 6. It can be seen that the while the stochastic components of the 10-year jet speed in CSF20C exhibit a tripole pattern consistent with atmospheric forcing (Seager et al., 2000), the overall association between the stochastic component and SSTs appears weak, with correlations of around -0.06 (note the colorscale). It is unclear how such a weak link can be responsible for the high correlations found in the ensemble mean. A different estimate of the stochastic forcing using the method of Ma et al. (2020) was also carried out, wherein the stochastic forcing is diagnosed using the spread of ensemble members around the ensemble

[2]The fact that the 95% CI is smaller for the 30-year smoothing than the 10-year smoothing is likely due to the overly small effective sample size when using 30-year smoothing for DPLE.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

mean each winter. The result is shown in Figure B3, and yields the same qualitative conclusion that the stochastic forcing is notably weaker than the ensemble mean signal[3].

To conclude, we consider it likely that the SPNA is a source of predictable jet speed forcing which is common across all four datasets.

4.4 Decadal timescale forcing as the accumulation of seasonal timescale forcing

In Section 3.2, we argued that the presence of skill in ASF20C implies that predictable jet speed forcing is taking place already within a single season. Having now argued that this forcing is coming from SPNA SSTs, it remains to test whether the forcing taking place on seasonal timescales suffices to explain the bulk of the forcing taking place on decadal timescales.

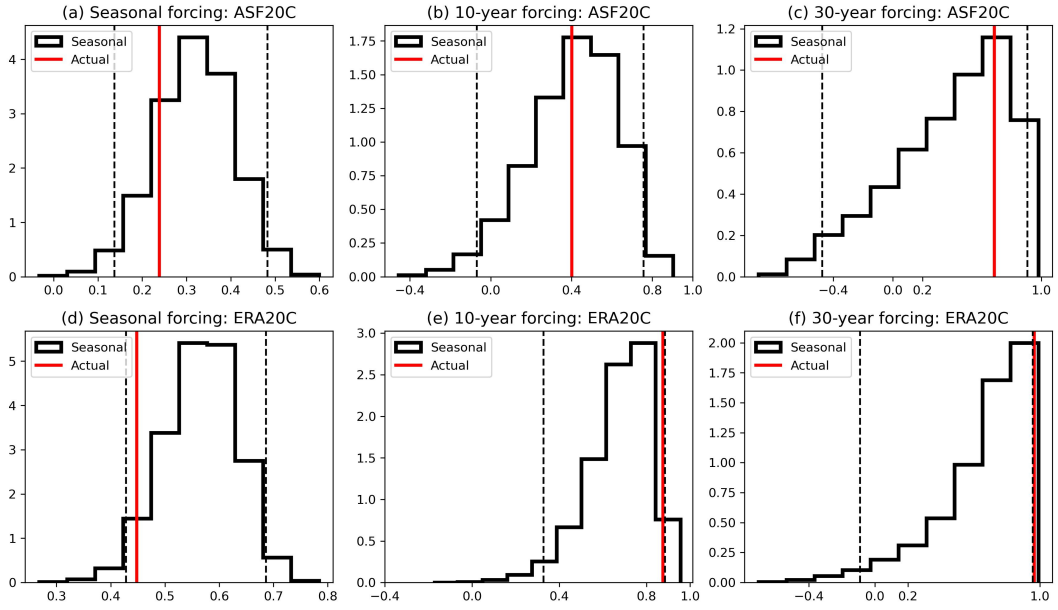

Figure 7. Histograms of the SPNA-JetSpeed correlations expected by chance from the seasonal timescale link, (a) and (d); taking 10-year averages of the seasonal link, (b) and (e); taking 30-year averages of the seasonal link, (c) and (f). See main text for details. The top row uses data fitted to ASF20C and the bottom to ERA20C. In each subplot the red line indicates the actually observed correlation in that dataset, and the dashed lines indicate the 95% confidence interval.

[3]The correlations in this second method are on the face of it larger than those obtained from the first method. However, this is at least in part due to the sample size in the second method being around half that of the first ($N = 50$ vs $N = 99$), and the fact that it only considers forcing within individual 10 year periods. The stochastic forcing across several consecutive decades, adjusted for sample size, would therefore be expected to be smaller, consistent with the estimates obtained using the first method.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

To test this, we model the SPNA-JetSpeed system as follows. For a given dataset, we perform a linear regression of the SPNA against JetSpeed to get

$$\text{JetSpeed} = a \cdot \text{SPNA} + b + \epsilon \tag{1}$$

for some constants a, b and a noise term ϵ which is normally distributed with mean 0. We then generate a random, synthetic timeseries of DJF SPNA by modelling it as an AR1 (with a lag of 1 year) as before: a corresponding JetSpeed timeseries is then obtained using the linear relationship just derived. We then take 10 and 30-year running means of these synthetic DJF timeseries and compute the correlations between them. By repeating this procedure 10,000 times we can assess what decadal timescale correlations are expected from taking running means of the interannual link.

The result of this for ASF20C is summarised in Figures 7(a), (b) and (c). The synthetic seasonal timescale correlations (Figure 7(a)) capture the observed seasonal correlation, essentially by construction. After taking running means, the distribution of synthetic correlations becomes notably skewed. For both 10 and 30-year running means (Figure 7(b) and (c)), the observed correlation is still comfortably within the 95% confidence interval of the synthetic correlations, and falls within the most likely bin in both cases. This implies that the decadal timescale skill of ASF20C is completely explained by its interannual skill, as expected. The same conclusion can be seen to hold for CSF20C by comparing Table 2 with Figure 7; we also verified the analysis using the shorter DPLE (not shown).

The same analysis for ERA20C is shown in Figures 7(d), (e) and (f). The fact the observed correlation on 10 and 30-year timescales straddles or even exceeds the 95% CI indicates that these correlations are higher than what might be expected purely from the seasonal timescale forcing. In other words, our simple model of the SPNA-JetSpeed link seems to just about fail for ERA20C. One obvious reason why this model may be inappropriate for ERA20C is the existence of amplifying feedbacks due to atmosphere-ocean coupling. These could introduce additional forcing terms or non-linear dependencies in (1) and thereby generate higher correlations.

Nevertheless, the linear seasonal timescale forcing alone predicts that 10-year correlations of around 0.8 and 30-year correlations of around 0.9 are very likely (Figure 7(e) and (f)). It is thus clear that while ERA20C shows some indication of additional forcing, the vast majority of the decadal signals can be explained using the seasonal signals.

4.5 Relationship with Atlantic Multidecadal Variability

We have proposed that decadal timescale predictability of the NAO arises from a forced response to SSTs in the SPNA. It is natural to ask how this relates to previous studies suggesting a forced response from the AMV, with the NAO response typically lagging the AMV by several years (Zhang et al., 2019). We suggest that this apparent discrepancy is a result of two factors.

Firstly, while the AMV pattern on interannual timescales is the famous 'horseshoe' pattern, the pattern largely collapses to a signal confined to the SPNA after performing decadal timescale smoothing (Delworth et al., 2017; Simpson et al., 2018). This suggests that as far as decadal atmospheric variability is concerned, the sub-polar North Atlantic SSTs are more relevant than the 'horseshoe' pattern. The importance of this region has been highlighted previously (Gastineau and Frankignoul, 2015; Woollings et al., 2015; Ortega et al., 2017; Delworth et al., 2017; Wills et al., 2019).

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

Secondly, anomalies in the SPNA (the northern part of the 'horseshoe') are observed to lead anomalies in the tropical Atlantic (the southern part of the 'horseshoe') by around 2 years (Zhang, 2007). It is likely that this propagation of anomalies takes place both via subsurface ocean dynamics and coupling with the atmosphere (Zhang et al., 2019). This means that performing analysis based on the standard AMV 'horseshoe' pattern is likely mixing together several different processes happening on different timescales. Since the SPNA anomalies arise first, restricting attention to these might be expected to give a clearer picture and better highlight causal pathways from the ocean to the atmosphere. This argument has also recently been made by Wills et al. (2019). Note that this second point is presumably related to the first. The 'horseshoe' pattern includes the effects of the interannual timescale processes associated with the propagation of SST anomalies along the 'horseshoe': taking decadal timescale averaging would be expected to remove these and highlight processes operating at much slower timescales, such as those associated with deep convection in the Labrador sea. This argument was also made in Delworth et al. (2017).

We therefore suggest that the apparent existence of an NAO response lagging the AMV by several years is, in fact, largely reflecting the existence of an essentially instantaneous atmospheric response to sub-polar North Atlantic SSTs, with the multiyear lag being an artefact of using an AMV index which averages across several different processes. A final point here is that the reverse impact of the NAO on the AMV has also been noted to be essentially lagged (Ma et al., 2020), additionally complicating analysis based on the AMV index.

5 Physical pathways and the signal-to-noise paradox

5.1 What affects the speed of the jet in climate models?

In order to shed light on the pathways involved, it is helpful to first understand the related question of what affects the climatological speed of the eddy-driven jet in climate models. To this end, we show in Figure 8 correlations between the multimodel climatological DJF jet speeds and climatological DJF temperatures (SSTs and air temperatures at 500hPa) at every gridpoint for 76 coupled CMIP5 and CMIP6 climate models (see Section 2). The period used to compute the climatologies is 1980-2015; linear trends are removed from the jet speeds and gridpoint timeseries before computing correlations to account for global warming.

It should be emphasised immediately that it is not possible to ascertain clear causality from such correlations, given the significant coupling between the jet, SSTs and air temperatures. Nevertheless, some useful points can be inferred. Figure 8(a) shows the typical tripole pattern in the North Atlantic, with the northern pole corresponding to the sub-polar North Atlantic and the southern pole corresponding to the tropical Atlantic. This is consistent with both a driving of the jet by the SSTs (Seager et al., 2000; Deser et al., 2007; Baker et al., 2019) and vice versa (Deser et al., 2010). Figure 8(b) shows that, once removed from the surface, the dominant signal is the meridional temperature gradient between the jet core and the poleward flank of the jet. By computing an index based on the difference in 500hPa DJF temperatures across the two peaks in Figure 8(b), we found that this meridional gradient accounts for more than 60% of the intermodel spread (correlation $C \approx 0.8$; not shown). Notably, at 500hPa there is no longer any sign of a strong influence from (or on) the tropics. The apparent signal from the Arctic may be related to the generation of planetary waves there, but we do not discuss this further.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

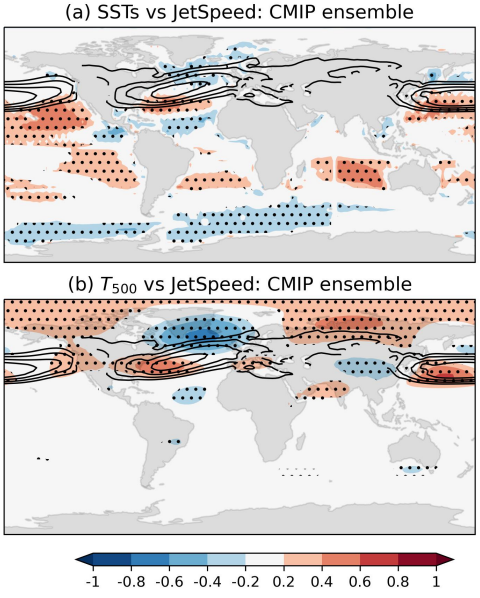

Figure 8. Filled contours: correlations between the spread in CMIP models' climatological DJF jet speeds and (a) the model spread in DJF SSTs at every gridpoint; (b) the model spread in DJF 500hPa air temperatures at every gridpoint. Line contours: the multimodel mean climatological zonal winds at 850hPa (contours drawn at 4, 6, 8 and 10 m/s). The climatologies are computed over the period 1980-2015. See main text for details of the models used.

Changes to meridional temperature gradients in the midlatitudes necessarily alter the eddy-driven jet, both directly and immediately via thermal wind balance and more indirectly by changing the baroclinicity. We interpret Figure 8(b) as saying that the meridional gradient most relevant for determining the speed of the eddy-driven jet is the local one between the jet core and the atmosphere situated directly above the SPNA. In particular, the full equator-to-pole gradient does not explain any of the intermodel jet speed spread. By contrast, in the southern hemisphere, Ceppi and Hartmann (2013) showed that the width of the Hadley Cell could explain intermodel spread in the southern hemisphere eddy-driven jet speed, suggesting a clear role for the tropics: they further reported that no such relationship could be found in the northern hemisphere. Our result here offers some clarification as to why, by highlighting the localised gradient as critical. It seems plausible that this is because the subtropical and eddy-driven jets are less strongly coupled in the North Atlantic compared to in the North Pacific or Southern Hemisphere.

5.2 The tropospheric pathway from sub-polar North Atlantic SSTs to the jet speed

Returning now to the question of how SPNA SSTs can modulate the jet speed, the results of the previous section suggest a simple explanation in keeping with the analysis of Woollings et al. (2015): anomalously cold SSTs in the SPNA cools the atmosphere aloft, thereby strengthening the local meridional temperature gradient around the jet, causing an intensification of the eddies and thereby an increased jet speed. Similarly for warm anomalies. To strengthen the case for this pathway, Figure

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

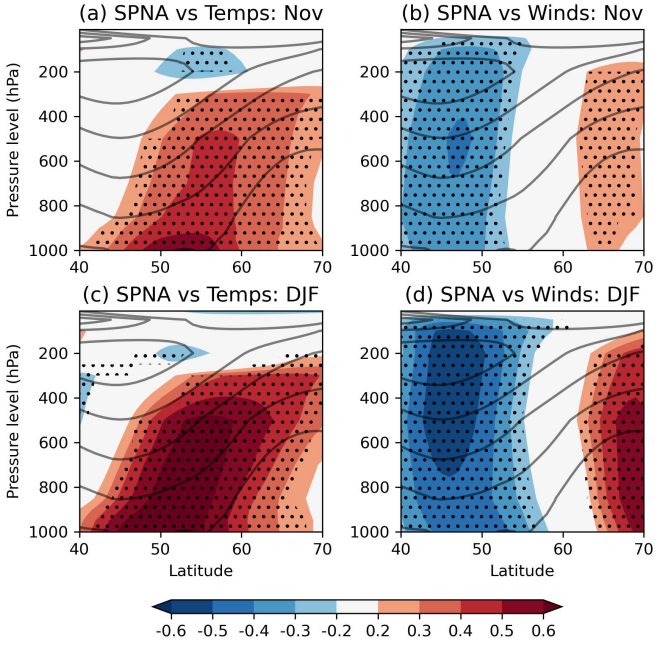

Figure 9. Correlations in ERA20C between SPNA SSTs and (left column) zonally averaged air temperatures at different pressure levels; (right column) zonally averaged zonal winds at different pressure levels. November averages are used in (a) and (b), while DJF averages are used in (c) and (d). The period 1900-2010 is used. The climatological zonal winds are shown in grey contours.

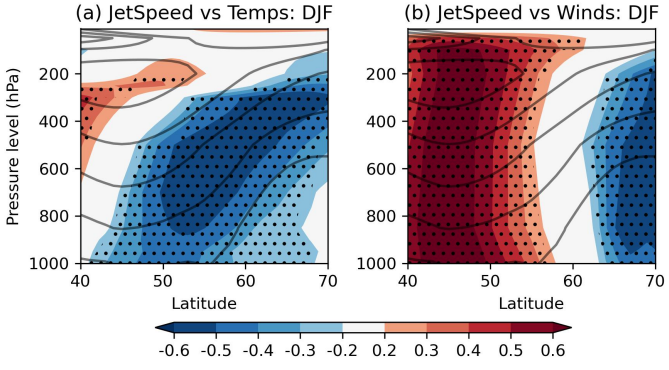

Figure 10. Correlations in ERA20C between jet speeds and (a) zonally averaged air temperatures at different pressure levels; (b) zonally averaged zonal winds at different pressure levels. DJF averages are used in both. The period 1900-2010 is used. The climatological zonal winds are shown in grey contours.

9 shows the vertical extent of the anomalies associated with SPNA SSTs in ERA20C, by correlating November SPNA SSTs with zonally averaged (a) temperatures and (b) zonal winds at different pressure levels: the averaging is done over longitudes

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

60W-20E. Figures (c) and (d) show the same but using DJF means. Figures 9(a) and (c) suggest that the heating does not remain confined to the surface, but extends up to around 300hPa, making the robust impact on the jet seen in Figure 9(b) plausible. For completeness, the reverse link from the jet speed to temperatures and winds is shown in Figure 10. Figure 10(a) shows that the heating pattern associated with the jet is maximal in the mid-troposphere. The heating associated with the SPNA (Figure 9(a) and (c)) projects well onto the jet pattern, but is maximal at the surface.

The structures are seen to be relatively barotropic in nature, especially in DJF. Detailed analysis of the tropospheric response to an SST anomaly in the North Atlantic, such as in Deser et al. (2007), show that the response generally proceeds in two stages. To begin with the induced anomaly is baroclinic in nature and localised to the heat source. In the second stage the baroclinic anomalies are removed by eddy heat and momentum transport, leading to a relatively barotropic structure with a more hemispheric scope. Figure 9 is therefore fully consistent with this tropospheric pathway having taken place.

To further test the importance of the tropospheric heating anomalies over the SPNA, we compute a rough estimate of the ERA20C jet speed anomalies expected from geostrophic wind balance as follows. We define a region labeled 'South' by 35-43N, 75-40W, which approximately covers the region in Figure 8(b) where the jet core and positive correlations overlap. As our 'North' region we use the same box used to define the SPNA region. We then assume a constant layer-thickness between 1000 and 300hPa, where 300hPa is the upper limit of the heating associated with the SPNA, as seen in Figure 9(c). Let $\hat{T}_N$ and $\hat{T}_S$ denote the layer-averaged DJF air temperature between 1000 and 300hPa in the North and South regions. By regressing our SPNA SST timeseries against $\hat{T}_N$, we obtain the SPNA-driven component which we denote by $\hat{T}_N(SPNA)$. We let $[\hat{T}_S]$ denote the average of $\hat{T}_S$ across all years 1900-2010. The timeseries of the difference $\hat{T}_N(SPNA) - [\hat{T}_S]$ thereby roughly measures the layer-averaged meridional temperature gradient across the jet core purely associated with SPNA SST variability. The discussion of the previous section suggests this gradient is the one most relevant for determining the jet speed. Geostrophic balance with a constant layer-thickness relates this gradient to the zonal winds U_{jet} in the jet core according to

$$U_{jet} = -\frac{R}{f} \cdot \log(1000/300) \cdot \frac{d}{dy}(\hat{T}_N(SPNA) - [\hat{T}_S]), \tag{2}$$

where $R \approx 287.05$ is the gas constant, $f \approx 0.0001$ is the midlatitude Coriolis, and $dy \approx 2553000$ metres. Here we have used that 1 degree latitude is roughly 111km. The U_{jet} timeseries is compared to the jet speed timeseries for ERA20C in Figure 11 using 30-year averages. They not only correlate extremely well ($C = 0.97$), but also share virtually identical magnitudes. In other words, the decadal jet speed variability of ERA20C can be fully accounted for by tropospheric temperature anomalies associated with the SPNA.

As always, inferring causality is challenging from diagnostics such as these. However, we note that if we rather correlate November SPNA anomalies against DJF temperatures/winds, we get a similar picture, albeit weaker in magnitude, especially for the winds (Figure B2). This suggests that the vertical heating is not purely coming from the adjustment of the jet itself, but also from processes induced by the SST anomalies. The fact that the heating associated with the SPNA peaks at the surface (Figure 9(a) and (c)), is also suggestive of SST driving. Further arguments as to why the correlations seen between the SPNA and the jet speed reflect forcing from the SSTs were already given in Section 4.3. Finally, many studies using idealised models and GCMs have demonstrated that SST anomalies can causally affect the jet (Palmer and Zhaobo, 1985; Deser et al., 2007;

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

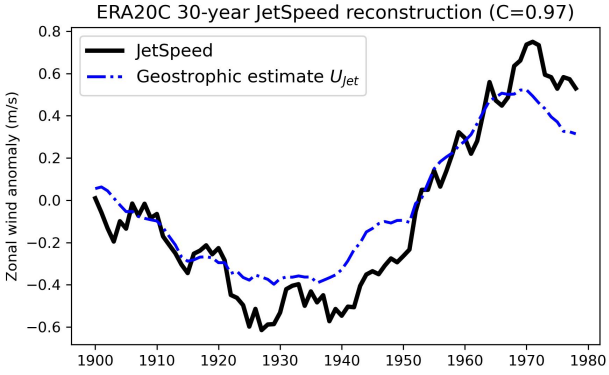

Figure 11. Black curve: 30-year averaged ERA20C DJF jet speed. Blue dashed curve: 30-year averages of the ERA20C U_{jet} timeseries, which roughly estimates the zonal wind variability associated with SPNA-induced changes to the meridional temperature gradient assuming geostrophic balance (see main text for details). The value C denotes the correlation coefficient between the two timeseries.

Hassanzadeh and Kuang, 2016; Baker et al., 2017, 2019). In particular, Baker et al. (2017) showed, using a dry model, that imposed heating anomalies modulate the speed of the eddy-driven jet by altering the meridional gradient around the jet core, making the causal pathway we propose here a plausible explanation for the observed decadal forecast skill.

5.3 Why is the jet latitude not predictable?

Several studies have argued that an AMV-driven NAO response involves a shift in both the strength and location of the storm track and/or baroclinic region (Msadek et al., 2011; Peings and Magnusdottir, 2014a; Frankignoul et al., 2015; Peings et al., 2016; Ortega et al., 2017). Such systematic latitudinal shifts in the storm track/baroclinicity, and hence the eddies, would be expected to manifest as predictable latitudinal shifts in the eddy-driven jet itself. However, as discussed in Section 3, there is no decadal predictability of the latitude of the jet in ASF20C, CSF20C or DPLE. This suggests that on decadal timescales, the eddies are by and large located at their climatological position, and that the mechanism driving predictable jet variability does not essentially depend on latitudinal shifts in the storm track. Rather, such predictable changes appear to be more clearly related to the direct constraint of thermal wind balance along with an *intensification* of the storm track. This is fully consistent with Woollings et al. (2015), who highlighted the contrasting nature of forced shifts to the latitude and speed of the jet.

Of course, it cannot be ruled out that SST anomalies in the SPNA, or indeed elsewhere, force shifts in the latitude of the jet in the real world that are simply not captured by the forecast models we considered. However, until predictability has been established it does not seem possible to reject the null hypothesis that decadal variability in the jet latitude amounts to chaotic and unpredictable internal variability. Indeed, the inconsistent latitudinal response in multimodel studies such as Ruggieri et al. (2021) lends concrete evidence for such inherent chaos. Further evidence for this is given in Figure B4, which repeats the procedure of Section 4 to search for potential sources of jet latitude skill in ASF20C. This shows that there are no regions

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

of significant SST-jet latitude correlations common to both ERA20C and ASF20C, consistent with the lack of jet latitude predictability. We will return to this point in the discussion and conclusions.

5.4 The signal-to-noise paradox and turbulent heat transfer

As discussed in the Introduction, there is a 'signal-to-noise paradox' in the forecast models, which suggests that the models are responding too weakly to sources of predictable forcing. Since we have argued that the predictable forcing originates in the SPNA, it is natural to consider why the response to SST anomalies in this region may be too weak in the forecast models. As previously noted, the response of the atmosphere to an imposed SST anomaly proceeds in two stages, with the first stage dominated by the immediate localised response, and the second stage dominated by the eddies and eddy-meanflow interactions. There is already evidence to suggest that the second stage contributes to the 'paradox', due to a deficient eddy feedback in the forecast models (Hardiman et al., 2022). Here we would like to raise the possibility that biases in the first stage may also be contributing.

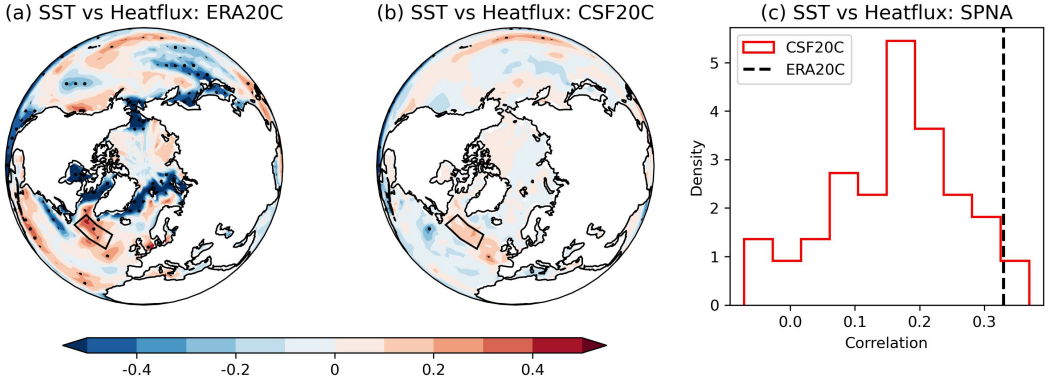

Figure 12. In (a): correlations between DJF SPNA SSTs and DJF heatfluxes at every gridpoint, using ERA20C (1900-2010). In (b): the same but using the CSF20C ensemble mean. Stipling in (a) and (b) denotes statistical significance ($\rho < 0.05$ with a two-tailed t-test). In (c): the distribution of correlations obtained by CSF20C by correlating area-averaged SPNA SSTs with area-averaged SPNA heatfluxes using all 51 ensemble members (red histogram); the black dashed line is the correlation in ERA20C.

The immediate atmospheric response to an SST anomaly is a result of forcing exerted by heatfluxes, which we understand here as the sum of the sensible and latent fluxes. To assess this heatflux forcing, we plot, in Figure 12(a) and (b), correlations between DJF averaged heatfluxes and SSTs at every gridpoint in ERA20C and the coupled CSF20C ensemble: qualitatively identical results are obtained if regression coefficients are used instead. Note that no additional decadal filtering is performed here. It can be seen that in ERA20C, there is a clear positive correlation in the SPNA: a positive SST anomaly is associated with a positive (i.e., upwards) heatflux. This would be expected to result in a heating of the atmosphere aloft, consistent with Figure 9(a). Note that this result is not inconsistent with that of Gulev et al. (2013), who report a negative correlation between an 'AMV index' and heatfluxes on interannual timescales. The index used in Gulev et al. (2013) is defined as the average SSTs

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

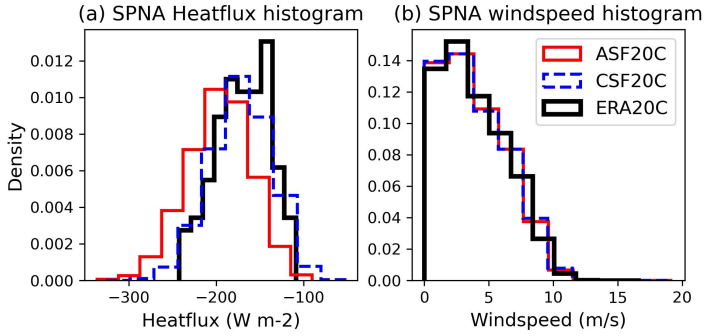

Figure 13. Histograms of DJF area-averaged SPNA (a) heatfluxes and (b) windspeeds, for ERA20C (black), ASF20C (red) and CSF20C (blue). The period 1900-2010 is used.

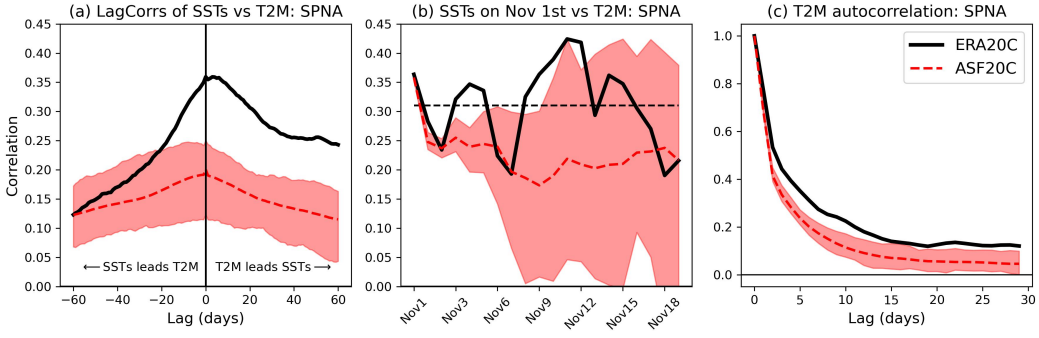

Figure 14. In (a): daily lag correlations between area-averaged SPNA SSTs and area-averaged SPNA two-meter temperatures (T2M). In (b): correlations between area-averaged SPNA SSTs on November 1st and between area-averaged SPNA T2M on subsequent days. In (c): daily autocorrelation of area-averaged SPNA T2M. In each case the thick black line is ERA20C, the dashed red line is the ASF20C ensemble mean, and the red shading shows the full range of the ASF20C ensemble spread. Only days in DJF are used, covering the period 1900-2010.

across a domain notably further south than our SPNA region, and inspection of Figure 12(a) shows that the correlations south of the SPNA are indeed largely negative.

In CSF20C, while the sign of the correlation in the SPNA is positive, its magnitude is about half the size (around 0.33 in ERA20C and 0.16 in CSF20C on average). The distribution of correlations between area-averaged SPNA SSTs and area-averaged SPNA heatfluxes obtained by CSF20C ensemble members is shown in Figure 12(c). This demonstrates that the correlation obtained in ERA20C is at the upper end of the ensemble spread, and in fact falls outside the 98th percentile. Given the positive sign of the correlations, we conclude that the communication from the ocean to the atmosphere is weaker in CSF20C than in ERA20C.

These correlations are even smaller in ASF20C (not shown), though it is not clear how these correlations should be interpreted, since the lack of coupling means the heatflux variability is completely different in nature compared to CSF20C and

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

ERA20C[4]. However, given that ASF20C and CSF20C share the same atmospheric component, it is reasonable to suppose that the communication from the ocean to the atmosphere is similarly biased in ASF20C. This is also supported by the fact that both ASF20C and CSF20C share a similar bias in the overall heatfluxes emerging from the SPNA: see Figure 13(a). Since ASF20C shares the same SSTs as ERA20C, these biases are not simply due to an overly cold SPNA. Nor is it due to excessively low windspeeds in this region: see Figure 13(b).

To further understand the impact of this bias, we show in Figure 14(a) the lagged correlations between daily SSTs and 2-metre temperature (T2M) in the SPNA region for ERA20C and ASF20C. A bias in these lagged correlations would be expected in ASF20C when T2M leads the SSTs, due to the lack of coupling. However, a substantial bias is also evident when the SSTs lead T2M, even (and especially) at lag 0, with none of the 51 ensemble members able to replicate the behaviour of ERA20C. This implies that in ASF20C, there is a substantial underestimation of how tightly linked SSTs and near-surface temperatures are, and Figure 12 suggests that this is at least in part due to an underestimation of the associated heatfluxes. The lag correlations in Figure 14(a) are computed using all days in DJF, so this underestimation could in principle be taking place late in the winter season. However, Figure 14(b) shows that the same bias is visible even when only correlating SSTs on November 1st (the initialisation day) with the T2M on subsequent days. We conclude that the SST-T2M link is likely too weak in ASF20C at all times. In particular, the weak link seems unlikely to be purely a result of deficiencies in the atmospheric adjustments due to the eddies, as these would be expected to manifest more clearly only after some time has passed. Instead, the weak link seems more consistent with biases in the surface temperature forcing. An inadequate surface temperature forcing from the slowly varying SSTs could also explain the reduced T2M autocorrelation in ASF20C, shown in Figure 14(c).

We have argued that SPNA SSTs force the jet by altering the meridional temperature gradient around the jet core, not just at the surface but at depth (Figure 9(a)). In ASF20C the atmospheric temperature response to SSTs is already substantially weaker than observed just 2 metres above sea-level. We therefore cautiously propose that the 'signal-to-noise paradox' in decadal NAO forecasts is at least partially due to deficiencies in the one-way communication from the ocean to the atmosphere in models. Note that biases in extratropical surface coupling have previously been suggested as contributing to the 'paradox' (Scaife and Smith, 2018), but the bias we highlight here is an uncoupled phenomenon.

It must be emphasised that the analysis we have shown here does not conclusively point to a bias in the heatflux response to a given SST anomaly. ERA20C is constructed using shortrange forecasts by a model closely related to ASF20C, with the data assimilation acting to adjust the initial conditions of these shortrange forecasts in order to produce a closer match between forecasts and observations with respect to the assimilated variables. What has been demonstrated is therefore that the effect of this data assimilation is to produce surface temperatures and heatfluxes more closely correlated with the observed SSTs. We have suggested that the forecast model has a bias in its heatflux response, and that the data assimilation is adjusting the fluxes and surface temperatures to counteract this. However, we cannot rule out that the appearance of higher correlations are due to adjustments entirely unrelated to the heatfluxes, such as wind speed or humidity adjustments in other regions. Much more careful analysis of the data assimilation process would be required to clarify this further, which goes beyond the scope of this

[4]Strictly speaking ERA20C is also not coupled. However, given that it is made up of consecutive 1-day forecasts with data assimilation, it can be assumed to capture all the effects of air-sea coupling, which are small on sub-daily timescales.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

paper. Nevertheless, two points are worth making here. Firstly, there does not seem to be any reason for the data assimilation to lead to higher SST-heatflux correlations *a priori*, suggesting that there probably is an associated model bias here. Secondly, there is in fact an obvious potential source of model biases in the heatflux response, which we now discuss.

The heatfluxes in models such as the IFS (which ASF20C/CSF20C is based on) are determined by parameterizations using several simplifications, including the Charnock formula for the surface roughness length (ECMWF, 2015). Greater roughness lengths are, broadly speaking, associated with increased friction and hence greater heat transfer. Experiments based on a GCM using the Charnock formula have found that varying the roughness length in a localised ocean region can have substantial impacts on the fluxes and hence surface temperatures (Kirk-Davidoff and Keith, 2008). Over the ocean, roughness is strongly influenced by unresolved processes such as waves and sea spray. In ASF20C these effects are represented by a globally fixed 'Charnock constant': in CSF20C this constant varies and is provided by the wave model. In either case there will undoubtedly be deficiencies, since the relationship between turbulent heat transfer and roughness length depends sensitively on the exact type of roughness and there is no general theory encapsulating all experiments: see e.g. the introduction to Wagner and Shishkina (2015). Note that the heatflux parameterization in the IFS is such that the heatflux on one timestep depends on the conditions of the previous timestep, and so a bias in the magnitude of the roughness length may also be expected to introduce biases in autocorrelation, such as those seen in Figure 14(c).

We therefore speculate that (a) the weak link between SSTs and heatfluxes (and consequently air temperatures) in the forecast models is partially related to unrealistic ocean surface roughness lengths in the SPNA, and (b) that this weak link may contribute to the 'signal-to-noise paradox' in decadal NAO forecasts. Further analysis would be required to fully justify such a hypothesis, which goes beyond the scope of the present paper. The relationship of our speculation here with the work of Zhang et al. (2021), who argued that the 'paradox' is somewhat mitigated in GCM simulations with an eddying ocean, is not clear to the authors at present[5].

6 Drivers of decadal SST variability

We have highlighted SSTs in the SPNA as the main source of decadal predictability of the jet. Understanding what is driving the variability of these SSTs, both in the real world and the forecast models, is essential for understanding how reliable forecast skill is likely to be in the future. Here we briefly examine three aspects: the role of aerosols, the AMOC, and stochastic atmospheric forcing.

6.1 The role of sulphate aerosols

The question of how aerosols affect North Atlantic SSTs is still actively debated, and could include the role of both direct radiative adjustments local to the region, as well as more indirect impacts from upstream emissions (Booth et al., 2012; Robson et al., 2022). Here we will only consider the hypothesis put forth by Robson et al. (2022), due to it being based on a large multimodel ensemble, unlike most other studies. In their study, they argued that sulphate aerosol emissions over the eastern

[5]In other words, we have no idea if/how resolving the eddies changes the roughness lengths generated by the wave model.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

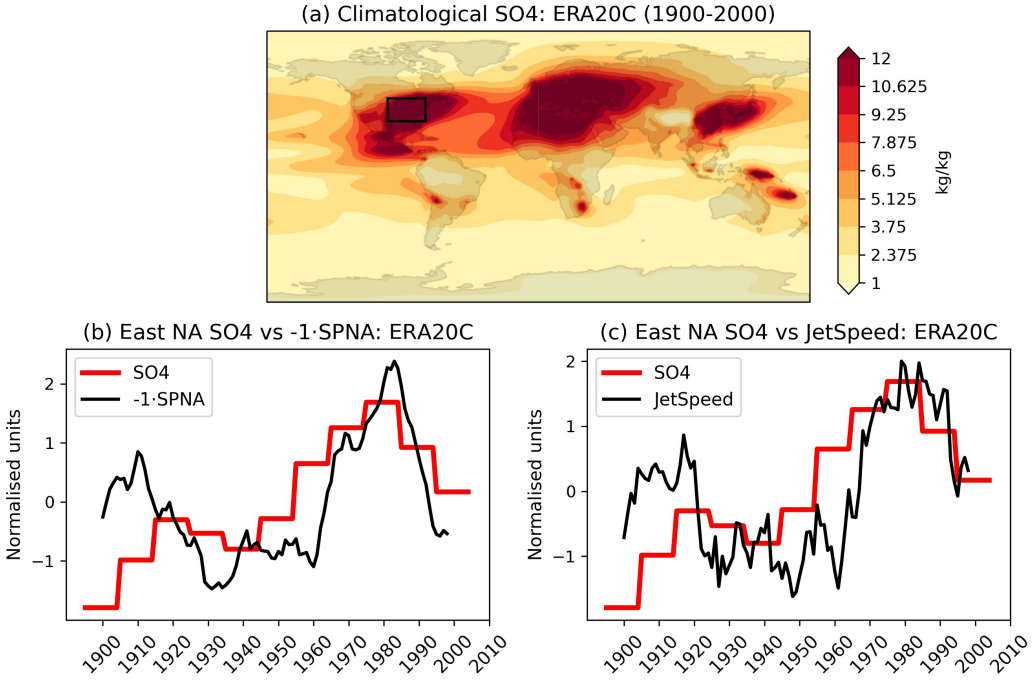

Figure 15. In (a): the climatological SO4 emissions of ERA20C, which is the same as those in ASF20C and CSF20C. In (b), timeseries of the 10-year averaged ERA20C SPNA SSTs and area-averaged East North American SO4. In (c), timeseries of the 10-year averaged ERA20C jet speed and area-averaged East North American SO4. The sign of the SPNA timeseries has been flipped for visual convenience. The domain of the East North American region has been marked with a black box.

part of North America led to a cooling of surface temperatures there, and that this cool air was then advected over the North Atlantic, thereby cooling the SSTs and ultimately modulating the AMOC. The possible direct impacts of emissions over the SPNA region itself will not be considered, since the emissions are generally very low in this region.

ERA20C, ASF20C and CSF20C all share the same prescribed aerosol emissions, namely those used in CMIP5 (Lamarque et al., 2010). The climatological SO4 emissions are shown in Figure 15(a), showing a clear concentration of emissions over the eastern North American region. To study the impact of these emissions on the SPNA and jet speed, we define an SO4 timeseries by averaging the emissions over the box 30-45N, 260-285E. This box is visualised in Figure 15(a) and approximately covers the eastern North American continent. This timeseries is compared to the 10-year averaged SPNA SSTs and ERA20C jet speed in Figures 15(b) and (c). The SO4 timeseries matches both the SSTs and jet speed in the period 1940-2010, with a correlation of around 0.8. However, in the earlier period 1900-1940, the SO4 emissions are anticorrelated with the SSTs and jet speed, with a correlation of around -0.3. Note that the CMIP5 aerosol forcing data is constant across each 10-year period, explaining the step-like quality of the SO4 timeseries.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

We conclude that while aerosol emissions over North America may have played a role in driving decadal SPNA variability, they are unlikely to explain the entire 20th century variability, especially in the early period when emissions are still very low. This is consistent with earlier work by Zhang et al. (2013). Note that the forecast models are able to capture the strong jet in this early period, despite the low aerosol emissions (Figure 1). If the aerosols are driving some of the SPNA variability, then this would be expected to lead to predictable shifts in the jet speed, as discussed in the present paper. It should also be noted that an aerosol-induced cooling of North America would strengthen the land-sea contrast, which would also be expected to drive a stronger jet (Portal et al., 2022). It is unclear if such a change in land-sea contrast alone can generate decadal timescale forecast skill.

6.2 The role of the AMOC

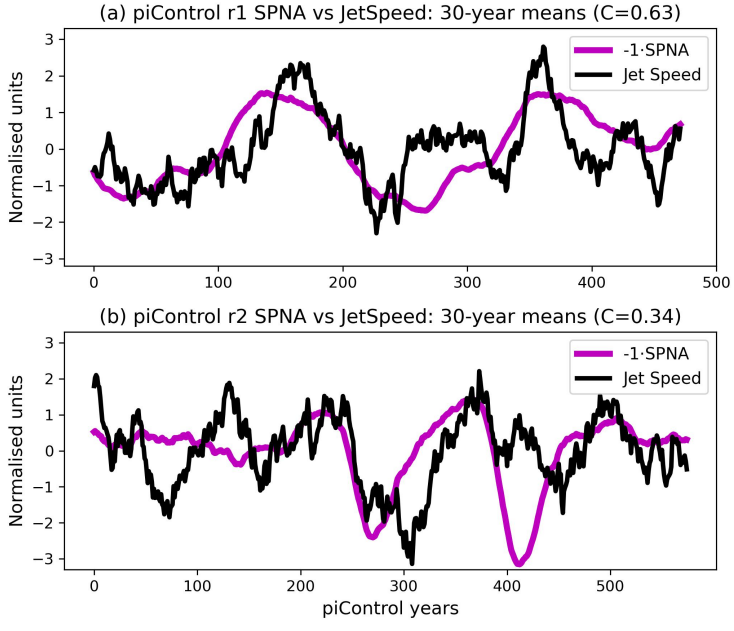

Figure 16. Timeseries of 30-year averages of the DJF SPNA SSTs (purple) and DJF jet speeds (black) in (a) the first EC-Earth3 pre-industrial control ensemble member, and (b) the same for the second ensemble member. The sign of the SPNA timeseries has been flipped for visual convenience. The value C in the titles indicates the correlation between the two timeseries. The timeseries have been normalised to have mean 0 and standard deviation 1.

To understand if the AMOC alone can drive SPNA-induced jet speed variability, we examine the two EC-Earth3 pre-industrial control simulations, which we refer to as piControl r1 and piControl r2. These simulations are particularly useful for assessing AMOC forcing, since they are known to exhibit large and striking centennial timescale AMOC oscillations (Meccia et al., 2022). The lack of any anthropogenic emissions means the associated variability in the SPNA will be entirely due to the AMOC.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

Figure 16 shows 30-year running DJF means of SPNA SSTs vs JetSpeed for the two ensemble members. The centennial AMOC oscillations are clearly visible in the SPNA timeseries. The correlations obtained between the SPNA and JetSpeed are -0.63 for the r1 member and -0.34 for the r2 member. A 95% CI using the 'Phase Shuffle' method is approximately ± 0.5, implying that the r1 correlation is highly significant while the r2 is not.

We cannot directly compare the correlations from these free-running coupled simulations to the ensemble mean correlations of ASF20C. However, we note that if we concatenate back-to-back 5 ASF20C ensemble members, we get a timeseries of length $5 \cdot 109 = 545$, approximately the same length as the piControl members. By randomly drawing 5 ASF20C members, concatenating them, and computing the correlations between the resulting SPNA and JetSpeed timeseries, we can compare ASF20C and piControl in a like-for-like manner. If we carry out 1000 random such draws, we find that the expected ASF20C correlation is around -0.2 with a 95% CI of approximately [-0.51, 0.18]. Almost identical results are obtained if using the CSF20C ensemble. This clearly suggests that the r2 correlation of -0.34 is perfectly consistent with the SPNA-JetSpeed link diagnosed in ASF20C/CSF20C. It also suggests that the r1 correlation of -0.63 is larger than expected, though not impossible, with a handful of concatenated ASF20C/CSF20C members producing correlations of around -0.6. Since the AMOC oscillations are particularly strong in the piControl runs, it is possible that this translates into SPNA-JetSpeed correlations which may be unrealistically large. The r1 member may also just be a random outlier.

Based on the above analysis we conclude that AMOC oscillations alone can easily produce SPNA variations associated with an SPNA-JetSpeed link of the same strength as in the forecast models, even in the absence of any aerosol forcing. It is therefore natural to speculate that the AMOC is responsible for the coherent SPNA and jet speed variability seen in the early 20th century (Figure 15), and likely contributes to the variability later on as well.

6.3 The role of stochastic atmospheric forcing

It has been suggested that the variability of the SPNA may be largely driven by stochastic atmospheric forcing from the midlatitude circulation (Clement et al., 2015). We have already discussed at length the question of causality in the SPNA-JetSpeed link, providing several arguments for the assertion that the SPNA variability in turn affects the midlatitudes (see Sections 4.3 and 55.2). Perhaps most fundamentally, a rejection of this assertion begs the question of where the forecast skill is coming from, since our analysis shows that the SPNA is the only viable source of skill in SSTs, and the 2-week decorrelation timescale of the tropospheric circulation precludes it from generating decadal timescale predictability. Section 6.1 shows that sulphate aerosol emissions also cannot explain the entirety of the skill.

The relatively weak link between the stochastic component of the atmosphere with the SPNA discussed in Section 4.3 (Figure 6 and B3) seems consistent with an interpretation that on decadal timescales there is no overall coherence in how the stochastic atmospheric variability forces the ocean, at least in the forecast model. However, Ma et al. (2020) argued that the response of the ocean to the NAO is lagged, or in other words that the NAO forces the SST tendencies rather than the SSTs themselves. Our instantaneous correlations may therefore underestimate the forcing. For this reason, our results are probably still consistent with the hypothesis of O'Reilly et al. (2019) that the AMV is effectively the integral of the NAO-induced forcing

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

over time. Viewing this hypothesis in the context of the present paper, the accumulated jet speed forcing across a season would essentially 'store memory' in the SPNA, which would in turn force the jet speed both this season and, potentially, the next.

While our results do not seem able to rule out the SPNA variability being purely jet-driven, there are several lines of evidence from earlier work pointing to a role for the ocean circulation. Simpson et al. (2018) showed that the DPLE forecasts can skillfully predict decadal SPNA SST variability, but fails to predict shorter timescale North Atlantic winds of an appreciable magnitude, which would seem to rule these winds out as the source of the SST skill. This was further corroborated by Yeager (2020), which found evidence that the SST skill has an abyssal origin. Finally, Delworth et al. (2017) highlighted that the North Atlantic SST processes taking place on multi-decadal timescales are more consistent with forcing by the oceanic circulation than by the atmosphere.

It seems highly likely that the SPNA-JetSpeed system is a genuinely coupled one, and it would be of clear interest to understand the relative importance of the two directions of forcing at setting the total observed variability in both the real world and forecast models. Given the subtleties introduced by the multiple timescales involved, even when restricting to just the SPNA (Khatri et al., 2022), this is not pursued at present.

7 Conclusions

We briefly summarise the main results and arguments.

1. Decadal forecast skill of the winter NAO appears to be entirely due to the decadal predictability of the speed of the North Atlantic jet. There is no apparent predictability of decadal variations in the latitude of the jet, even in seasonal hindcasts with prescribed SSTs (Figures 1 and 2).

2. Initialised seasonal hindcasts can skillfully reproduce decadal variations in jet speed all the way back to 1900, and match the behaviour of a genuine decadal forecast in the period 1954-2010. We argued that this justifies using such seasonal hindcasts to diagnose decadal forecast signals using the full period 1900-2010, effectively doubling the available years compared to existing decadal forecasts. In addition, this implies that the decadal signal in the atmosphere must be skilfully captured already on seasonal timescales.

3. The only potential source of interannual-to-decadal jet speed skill coming from SSTs is the SPNA region (Figures 3 and 4). SSTs in this region enjoy large and statistically significant correlations with the jet speed, which we argue are due at least in part to a causal link from the ocean to the atmosphere (Figure 5). We also show that the majority of the decadal SPNA-JetSpeed link in the forecast models, along with the associated forecast skill, can be explained by a small but consistent seasonal timescale forcing from the SSTs (Figure 7).

4. The pathway from SPNA SSTs to the jet speed is argued to be tropospheric in nature: the surface heating anomaly extends relatively deeply into the troposphere (Figure 9) and is optimally situated to perturb the jet speed (Figures 8 and 11) both by direct adjustments consistent with thermal wind balance and subsequent reinforcements by the eddies.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

5. The forecast models are found to significantly underestimate the link between SSTs and heatfluxes, and hence surface temperatures, in the SPNA (Figures 12, 13 and 14). This may lead to the forecast models underestimating the atmospheric response to SPNA anomalies, thereby contributing to the 'signal-to-noise paradox'. We speculate that this underestimation may be related to deficiencies in the parameterizations of surface roughness lengths.

6. Sulphate aerosol emissions in North America may explain part of the 20th century variability in the SPNA, but do not explain the early 20th century (Figure 15). AMOC oscillations alone are shown to be capable of inducing SST shifts of strength comparable to those connected with jet speed variability (Figure 16).

Note that the importance of the SPNA in driving predictable decadal jet variability was emphasised in Simpson et al. (2018) using a different metric of North Atlantic jet variability in March. Our work here is clearly closely related and was directly inspired by theirs.

There are several shortcomings to our analysis, chief of which is the fact that we have only studied two independent forecast models. Our argument that the stratosphere is unlikely to be relevant due to the initialisation biases of the seasonal hindcasts may also be flawed, and it would be of clear interest to examine the role of the stratosphere more closely. We also interpreted the similar jet speed correlations in ASF20C/CSF20C and DPLE as evidence that the mechanisms involved are the same for these data sets. However, it may be that ASF20C/CSF20C obtain most of their 'skill' from the initialisation (lag 0), while DPLE obtains most of its skill from representing slower frequency processes better (lags > 0). If so, conclusions drawn using ASF20C/CSF20C may not carry over to a genuine forecast context. Nevertheless, our findings have some important implications.

To begin with, we have reinforced the conclusion of earlier studies such as Baker et al. (2017) that the speed and latitude of the jet should be considered separately, and that the use of indices like the NAO (which amalgamate the two) may be misleading. Our finding that the decadally averaged jet speed is predictable while the latitude is not stands in amusing contrast to work on *seasonal* forecasts, which suggest the opposite picture of a predictable latitude and unpredictable speed (Parker et al., 2019). This reinforces the analysis of Woollings et al. (2015), who showed that the nature of forced jet variability differs depending on the timescale. On interannual timescales, the variability is dominated by the jet latitude, in turn associated with meridional shifts in the transient eddies and location of blocking. On decadal timescales, the variability is dominated by the jet speed, in turn associated with changes both to the *strength* of the eddy forcing and to the occurrence of transient Rossby wave breaking on both sides of the jet. Our analysis corroborates the speculation in Woollings et al. (2015) that this may result in the nature and sources of predictability being different on the two timescales. In particular, our analysis is consistent with the following hypothesis:

(a) On seasonal timescales the forcing from the SPNA is too small to be visible (resulting in the appearance of no jet speed skill), with jet latitude skill arising from accurately predicting meridional shifts in the eddies;

(b) On decadal timescales the eddies are varying chaotically around their climatological position (resulting in no jet latitude predictability), but their intensity can vary predictably depending on the underlying SPNA SSTs, which strengthen or weaken the meridional temperature gradient.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

It is also possible that biases in the climatological jet latitude (i.e. in the model mean state) of the forecast models are contributing to the apparent lack of jet latitude predictability. The sensitivity analysis of Baker et al. (2017) showed that the response of the jet latitude to thermal forcing changes sign abruptly around the jet core, while the response of the jet speed is more uniform across the jet. The accurate response of a model jet to a given SST anomaly may therefore differ dramatically based on its climatological position. This may go some way to explaining the inconsistent jet latitude response in multimodel studies such as Ruggieri et al. (2021).

A novelty of our work is the use of the initialised 20th century seasonal hindcast products ASF20C and CSF20C. While Parker et al. (2019) considered the fast and slow variability of the jet speed and latitude in ASF20C, the possibility of utilising these hindcasts to study the mechanisms underpinning decadal predictability seems to have been overlooked. The fact that ASF20C covers the entire 20th century, is reinitialised every year, and is uncoupled, are all highly attractive properties for simplifying analysis which we made extensive use of. We believe that ASF20C may be similarly beneficial for the study of decadal predictability in many other contexts.

We add to a growing body of literature suggesting that the sub-polar gyre is the main source of North Atlantic jet forcing on decadal timescales, and that the use of a larger AMV or AMO pattern might be counterproductive for questions of predictability by mixing together different mechanisms and timescales. There are however some hints that the tropical Atlantic may contribute to a coupled ocean-atmosphere feedback. The coupled CSF20C hindcast exhibits higher skill than ASF20C at reproducing the 20th century jet speed variability. Comparing Figures 3(e) and 4(e) shows that the tropical Atlantic is a region of jet speed correlations common to both CSF20C and ERA20C which is not present in ASF20C. In observational data SPNA SST anomalies are known to precede tropical Atlantic anomalies (Zhang et al., 2019), and this may be an essentially coupled ocean-atmosphere phenomenon which reinforces the forcing on the jet exerted by the SPNA. This could explain both the higher skill of CSF20C and the fact that ERA20C appears to show a stronger SPNA-JetSpeed link than what is expected from the seasonal timescale link alone. It also provides a possible reconciliation of our results with those of Davini et al. (2015) who argued for the importance of the tropical Atlantic.

Finally, our analysis has important implications for the robustness of the decadal forecast skill. We have shown that the SST variability in the SPNA driving predictable jet speed shifts can in principle be due to both the AMOC and aerosol emissions, and it seems likely that the observed 20th century variability is a result of both of these. Knowledge of current aerosol emission trends and skillful forecasts of ocean dynamics are therefore both likely sources of decadal predictability in the future as well. We have also highlighted a model bias of an overly weak heatflux response to a fixed SST anomaly, which results in the atmospheric temperature response being far too weak in the forecast models already at 2 metres above sea, let alone at depth. The heatflux response depends sensitively on the roughness lengths of the SPNA (Kirk-Davidoff and Keith, 2008), the magnitude of which in ASF20C depends on a globally fixed 'Charnock constant'. A natural avenue of future work is therefore to assess whether the 'signal-to-noise paradox' could be alleviated with an improved parameterization of roughness lengths, for example by utilising machine-learning tools to learn an improved and locally varying scheme.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

Data availability. ASF20C and CSF20C data is freely available on CEDA (Weisheimer, 2020). ERA20C data is freely available via ECMWF at https://apps.ecmwf.int/datasets/data/era20c-daily/levtype=sfc/type=an/. DPLE data is available via the Earth System Grid Federation (Yeager, 2018). CMIP6 data is available for download via the Earth System Grid Federation.

Appendix A

Appendix B

Author contributions. KS led the writing of the manuscript and carried out the majority of the data analysis. PR and PD contributed data analysis related to idealised model results and the role of the AMOC in the pre-industrial CMIP6 simulations; aided with literature review; and helped interpret results. TW aided interpretation of results and provided expert guidance on physical mechanisms. IRS helped in the procurement and analysis of DPLE model data and interpretation of the results.

Competing interests. We declare that there are no competing interests.

Acknowledgements. KS was funded by a Thomas Philips and Jocelyn Keene Junior Research Fellowship at Jesus College, Oxford. PR acknowledges the use of computational resources from the parallel computing cluster of the Open Physics Hub (https://site.unibo.it/openphysicshub/en) at the Department of Physics and Astronomy of the University of Bologna. IRS was supported by the National Center for Atmospheric Research which is a major facility sponsored by the National Science Foundation under the Cooperative Agreement 1852977.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

Model name	Ensemble member	Historical dates
ACCESS-CM2	r1i1p1f1	1950-2010
AWI-ESM-1-1-LR	r1i1p1f1	1900-2010
BCC-CSM2	r1i1p1f1	1950-2010
BCC-ESM1	r1i1p1f1	1950-2010
CanESM5	r1i1p1f1	1900-2010
CESM2-FV2	r1i1p1f1	1900-2010
CESM2	r1i1p1f1	1900-2010
CESM2-WACCM-FV2	r1i1p1f1	1900-2010
CESM2-WACCM	r1i1p1f1	1900-2010
CNRM-CM6-1	r1i1p1f2	1950-2010
CNRM-CM6-1-HR	r1i1p1f2	1900-2010
CNRM-ESM2	r1i1p1f2	1950-2010
EC-Earth3	r1i1p1f1	1900-2010
FGOALS-f3	r1i1p1f1	1950-2010
FGOALS-g3	r1i1p1f1	1900-2010
GFDL-CM4	r1i1p1f1	1910-2010
GISS-E2-1-G	r1i1p1f1	1910-2010
HadGEM3-GC31-LL	r1i1p1f3	1900-2010
HadGEM3-GC31-MM	r1i1p1f3	1900-2010
INM-CM4-8	r1i1p1f1	1900-2010
INM-CM5-0	r1i1p1f1	1900-2010
IPSL-CM6A-LR	r1i1p1f1	1950-2010
MIROC6	r1i1p1f1	1900-2010
MPI-ESM1-2-HAM	r1i1p1f1	1900-2010
MPI-ESM1-2-HR	r1i1p1f1	1900-2010
MPI-ESM1-2-LR	r1i1p1f1	1910-2010
MRI-ESM2-0	r1i1p1f1	1950-2010
NorESM2-LM	r1i1p1f1	1950-2010
NorESM2-MM	r1i1p1f1	1950-2010
TaiESM1	r1i1p1f1	1900-2010
UKESM1-0-LL	r1i1p1f2	1900-2010

Table A1. CMIP6 models used in this paper, with the range of historical data available.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

Model name	Ensemble members	Historical dates
ACCESS1-0	r1i1p1	1950-2005
ACCESS1-3	r1i1p1	1950-2005
BCC-CSM1-1	r1i1p1	1950-1998
BCC-CSM1-1-m	r1i1p1	1950-2005
BNU-ESM	r1i1p1	1950-2005
CanESM2	r[1,2,3,4,5]i1p1	1950-2005
CCSM4	r6i1p1	1950-2005
CMCC-CESM	r1i1p1	1950-2005
CMCC-CM	r1i1p1	1950-2005
CMCC-CMS	r1i1p1	1950-2005
CMCC-CM5	r1i1p1	1950-2005
EC-EARTH	r[1,2,7,9,12]i1p1	1950-2005
FGOALS-g2	r[1,3]i1p1	1950-2005
GFDL-CM3	r[1,2,3]i1p1	1950-2005
HadCM3	r[1,2,3,4,5,6,7,8,9,10]i1p1	1960-2005
HadGEM2	r[1,2,3]i1p1	1960-2005
IPSL-CM5A-LR	r[1,2,3,4,5,6]i1p1	1950-2005
IPSL-CM5A-MR	r[1,2,3]i1p1	1950-2005
IPSL-CM5B-LR	r1i1p1	1950-2005
MIROC5	r[1,2,3,4,5]i1p1	1950-2005
MIROC-ESM-CHEM	r1i1p1	1950-2005
MIROC-ESM	r[1,2,3]i1p1	1950-2005
MPI-ESM-LR	r[1,2,3]i1p1	1950-2005
MPI-ESM-MR	r[1,2,3]i1p1	1950-2005
MPI-ESM-P	r[1,2]i1p1	1950-2005
MRI-CGCM3	r1i1p1	1950-2005
MRI-ESM1	r1i1p1	1950-2005
NorESM1-M	r[1,2,3]i1p1	1950-2005

Table A2. CMIP5 models used in this paper, and the corresponding ensemble members and historical date range available.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

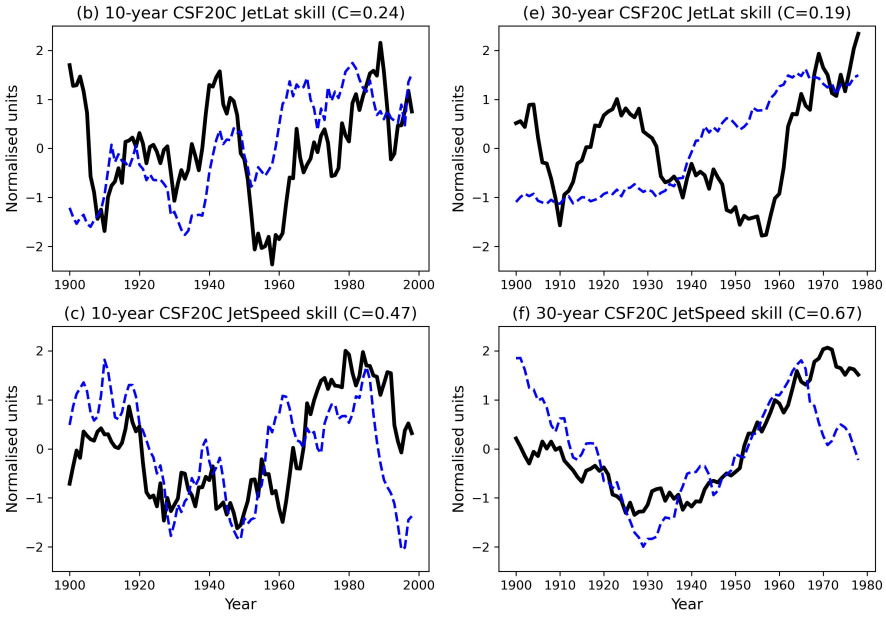

Figure B1. Timeseries of 10-year running DJF means of (a) the jet latitude, and (b) the jet speed. The same but with 30-year running means in (c) and (d). The thick black curves are always ERA20C and the dashed blue curves are always the CSF20C ensemble mean. The timeseries have been normalised to have mean 0 and standard deviation 1, for visualisation purposes. The value C in each subplot is the correlation between the two timeseries.

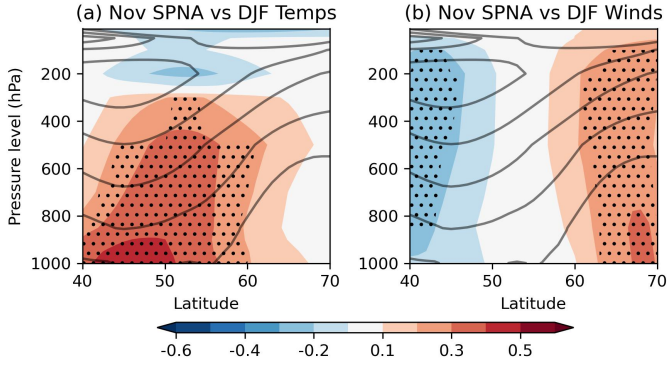

Figure B2. Correlations in ERA20C between November SPNA SSTs and (a) DJF zonally averaged air temperatures at different pressure levels; (b) DJF zonally averaged zonal winds at different pressure levels. The period 1900-2010 is used. The climatological zonal winds are shown in grey contours.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

Model name	Ensemble member	Historical dates
AWI-CM-1-1-LR	r1i1p1f002	1950-2014
AWI-CM-1-1-HR	r1i1p1f002	1950-2014
CMCC-CM2-HR4	r1i1p1f1	1950-2014
CMCC-CM2-VHR4	r1i1p1f2	1950-2014
CNRM-CM6-1	r[1,2]i1p1f2	1950-2014
CNRM-CM6-1-HR	r1i1p1f1	1950-2014
EC-Earth3	r[1,2,3]i1p2f1	1950-2014
EC-Earth3-HR	r[1,2,3]i1p2f1	1950-2014
ECMWF-IFS-LR	r[1,2,3,4,5,6,7,8]i1p1f1	1950-2014
ECMWF-IFS-MR	r[1,2,3]i1p1f1	1950-2014
ECMWF-IFS-HR	r[1,2,3,4,5,6]i1p1f1	1950-2014
HadGEM-GC31-LL	r1i[1,2,3,4,5,6,7]p1f1	1950-2014
HadGEM-GC31-MM	r1i1p1f1	1950-2014
HadGEM-GC31-HM	r1i[1,2,3]p1f1	1950-2014
HadGEM-GC31-HH	r1i1p1f1	1950-2014
MPI-ESM1-2-HR	r1i1p1f1	1950-2014
MPI-ESM1-2-XR	r1i1p1f1	1950-2014

Table A3. PRIMAVERA models used in this paper, with the range of historical data available.

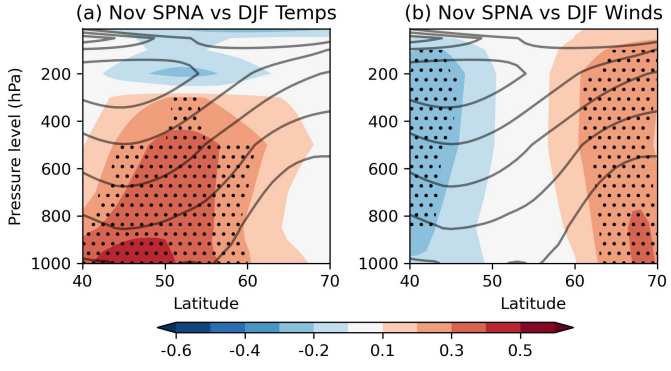

Figure B3. Correlations between the 'stochastic noise' component of the 10-year averaged DJF CSF20C jet speed and 10-year averaged DJF SSTs at each gridpoint, using the method of Ma et al. (2020). The period 1900-2010 is used.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

References

Athanasiadis, P. J., Yeager, S., Kwon, Y.-O., Bellucci, A., Smith, D. W., and Tibaldi, S.: Decadal predictability of North Atlantic blocking and the NAO, NPJ Climate and Atmospheric Science, 3, 1–10, 2020.

Baker, H. S., Woollings, T., and Mbengue, C.: Eddy-driven jet sensitivity to diabatic heating in an idealized GCM, Journal of Climate, 30, 6413–6431, 2017.

Baker, H. S., Woollings, T., Forest, C. E., and Allen, M. R.: The linear sensitivity of the North Atlantic Oscillation and eddy-driven jet to SSTs, Journal of Climate, 32, 6491–6511, 2019.

Barrier, N., Cassou, C., Deshayes, J., and Treguier, A.-M.: Response of North Atlantic Ocean circulation to atmospheric weather regimes, Journal of Physical Oceanography, 44, 179–201, 2014.

Bellomo, K., Murphy, L. N., Cane, M. A., Clement, A. C., and Polvani, L. M.: Historical forcings as main drivers of the Atlantic multidecadal variability in the CESM large ensemble, Climate Dynamics, 50, 3687–3698, 2018.

Bjerknes, J.: Atlantic air-sea interaction, in: Advances in geophysics, vol. 10, pp. 1–82, Elsevier, 1964.

Booth, B. B., Dunstone, N. J., Halloran, P. R., Andrews, T., and Bellouin, N.: Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability, Nature, 484, 228–232, 2012.

Bracegirdle, T. J.: Early-to-Late Winter 20th Century North Atlantic Multidecadal Atmospheric Variability in Observations, CMIP5 and CMIP6, Geophysical Research Letters, 49, e2022GL098 212, 2022.

Ceppi, P. and Hartmann, D. L.: On the Speed of the Eddy-Driven Jet and the Width of the Hadley Cell in the Southern Hemisphere, Journal of Climate, 26, 3450 – 3465, https://doi.org/10.1175/JCLI-D-12-00414.1, 2013.

Cherchi, A., Fogli, P. G., Lovato, T., Peano, D., Iovino, D., Gualdi, S., Masina, S., Scoccimarro, E., Materia, S., Bellucci, A., and Navarra, A.: Global Mean Climate and Main Patterns of Variability in the CMCC-CM2 Coupled Model, Journal of Advances in Modeling Earth Systems, 11, 185–209, https://doi.org/10.1029/2018MS001369, 2019.

Clement, A., Bellomo, K., Murphy, L. N., Cane, M. A., Mauritsen, T., Rädel, G., and Stevens, B.: The Atlantic Multidecadal Oscillation without a role for ocean circulation, Science, 350, 320–324, 2015.

Danabasoglu, G., Bates, S. C., Briegleb, B. P., Jayne, S. R., Jochum, M., Large, W. G., Peacock, S., and Yeager, S. G.: The CCSM4 ocean component, Journal of Climate, 25, 1361–1389, 2012.

Davini, P., von Hardenberg, J., and Corti, S.: Tropical origin for the impacts of the Atlantic multidecadal variability on the Euro-Atlantic climate, Environmental Research Letters, 10, 094 010, 2015.

Dell'Aquila, A., Corti, S., Weisheimer, A., Hersbach, H., Peubey, C., Poli, P., Berrisford, P., Dee, D., and Simmons, A.: Benchmarking Northern Hemisphere midlatitude atmospheric synoptic variability in centennial reanalysis and numerical simulations, Geophysical Research Letters, 43, 5442–5449, 2016.

Delworth, T., Manabe, S., and Stouffer, R. J.: Interdecadal variations of the thermohaline circulation in a coupled ocean-atmosphere model, Journal of Climate, 6, 1993–2011, 1993.

Delworth, T. L., Zeng, F., Zhang, L., Zhang, R., Vecchi, G. A., and Yang, X.: The Central Role of Ocean Dynamics in Connecting the North Atlantic Oscillation to the Extratropical Component of the Atlantic Multidecadal Oscillation, Journal of Climate, 30, 3789 – 3805, https://doi.org/10.1175/JCLI-D-16-0358.1, 2017.

Deser, C., Tomas, R. A., and Peng, S.: The Transient Atmospheric Circulation Response to North Atlantic SST and Sea Ice Anomalies, Journal of Climate, 20, 4751 – 4767, https://doi.org/10.1175/JCLI4278.1, 2007.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

Deser, C., Alexander, M. A., Xie, S.-P., Phillips, A. S., et al.: Sea surface temperature variability: Patterns and mechanisms, Annu. Rev. Mar. Sci, 2, 115–143, 2010.

Dorrington, J., Strommen, K., and Fabiano, F.: Quantifying climate model representation of the wintertime Euro-Atlantic circulation using geopotential-jet regimes, Weather and Climate Dynamics, 3, 505–533, https://doi.org/10.5194/wcd-3-505-2022, 2022.

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arsouze, T., Bergman, T., Bernardello, R., Boussetta, S., Caron, L.-P., Carver, G., et al.: The EC-Earth3 earth system model for the coupled model intercomparison project 6, Geoscientific Model Development, 15, 2973–3020, 2022.

Ebisuzaki, W.: A method to estimate the statistical significance of a correlation when the data are serially correlated, Journal of climate, 10, 2147–2153, 1997.

ECMWF: IFS Documentation CY41R1 - Part IV: Physical Processes, no. 4 in IFS Documentation, ECMWF, https://doi.org/10.21957/p50qmwprw, 2015.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937–1958, https://doi.org/10.5194/GMD-9-1937-2016, 2016.

Fichefet, T. and Maqueda, M. M.: Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics, Journal of Geophysical Research: Oceans, 102, 12 609–12 646, 1997.

Frankignoul, C., Gastineau, G., and Kwon, Y.-O.: Wintertime Atmospheric Response to North Atlantic Ocean Circulation Variability in a Climate Model, Journal of Climate, 28, 7659 – 7677, https://doi.org/10.1175/JCLI-D-15-0007.1, 2015.

Gastineau, G. and Frankignoul, C.: Influence of the North Atlantic SST Variability on the Atmospheric Circulation during the Twentieth Century, Journal of Climate, 28, 1396 – 1416, https://doi.org/10.1175/JCLI-D-14-00424.1, 2015.

Gulev, S. K., Latif, M., Keenlyside, N., Park, W., and Koltermann, K. P.: North Atlantic Ocean control on surface heat flux on multidecadal timescales, Nature, 499, 464–467, 2013.

Gutjahr, O., Putrasahan, D., Lohmann, K., Jungclaus, J. H., von Storch, J.-S., Brüggemann, N., Haak, H., and Stössel, A.: Max Planck Institute Earth System Model (MPI-ESM1.2) for the High-Resolution Model Intercomparison Project (HighResMIP), Geoscientific Model Development, 12, 3241–3281, https://doi.org/10.5194/gmd-12-3241-2019, 2019.

Haarsma, R., Acosta, M., Bakhshi, R., Bretonnière, P.-A. B., Caron, L.-P., Castrillo, M., Corti, S., Davini, P., Exarchou, E., Fabiano, F., Fladrich, U., Fuentes Franco, R., García-Serrano, J., von Hardenberg, J., Koenigk, T., Levine, X., Meccia, V., van Noije, T., van den Oord, G., Palmeiro, F. M., Rodrigo, M., Ruprich-Robert, Y., Le Sager, P., Tourigny, E., Wang, S., van Weele, M., and Wyser, K.: HighResMIP versions of EC-Earth: EC-Earth3P and EC-Earth3P-HR. Description, model performance, data handling and validation, Geoscientific Model Development Discussions, 2020, 1–37, https://doi.org/10.5194/gmd-2019-350, 2020.

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N. S., Guemas, V., von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L. R., Lu, J., Luo, J.-J., Mao, J., Mizielinski, M. S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., and von Storch, J.-S.: High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6, Geoscientific Model Development, 9, 4185–4208, https://doi.org/10.5194/gmd-9-4185-2016, 2016.

Hardiman, S. C., Dunstone, N. J., Scaife, A. A., Smith, D. M., Comer, R., Nie, Y., and Ren, H.-L.: Missing eddy feedback may explain weak signal-to-noise ratios in climate predictions, npj Climate and Atmospheric Science, 5, 57, 2022.

Hassanzadeh, P. and Kuang, Z.: The linear response function of an idealized atmosphere. Part I: Construction using Green's functions and applications, Journal of the Atmospheric Sciences, 73, 3423–3439, 2016.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

Hasselmann, K.: Stochastic climate models part I. Theory, tellus, 28, 473–485, 1976.

Hunke, E. C., Lipscomb, W. H., Turner, A. K., Jeffery, N., and Elliott, S.: Cice: the los alamos sea ice model documentation and software user's manual version 4.1 la-cc-06-012, T-3 Fluid Dynamics Group, Los Alamos National Laboratory, 675, 500, 2010.

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., Lamarque, J.-F., Large, W. G., Lawrence, D., Lindsay, K., et al.: The community earth system model: a framework for collaborative research, Bulletin of the American Meteorological Society, 94, 1339–1360, 2013.

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., et al.: SEAS5: the new ECMWF seasonal forecast system, Geoscientific Model Development, 12, 1087–1117, 2019.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S., Danabasoglu, G., Edwards, J., et al.: The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability, Bulletin of the American Meteorological Society, 96, 1333–1349, 2015.

Khatri, H., Williams, R. G., Woollings, T., and Smith, D. M.: Fast and slow subpolar ocean responses to the North Atlantic Oscillation: Thermal and dynamical changes, Geophysical Research Letters, 49, e2022GL101 480, 2022.

Kim, W. M., Yeager, S., Chang, P., and Danabasoglu, G.: Low-frequency North Atlantic climate variability in the Community Earth System Model large ensemble, Journal of Climate, 31, 787–813, 2018.

Kirk-Davidoff, D. B. and Keith, D. W.: On the climate impact of surface roughness anomalies, Journal of the Atmospheric Sciences, 65, 2215–2234, 2008.

Kravtsov, S.: Pronounced differences between observed and CMIP5-simulated multidecadal climate variability in the twentieth century, Geophysical Research Letters, 44, 5749–5757, 2017.

Kushnir, Y.: Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions, Journal of Climate, 7, 141–157, 1994.

Kushnir, Y., Robinson, W., Bladé, I., Hall, N., Peng, S., and Sutton, R.: Atmospheric GCM response to extratropical SST anomalies: Synthesis and evaluation, Journal of Climate, 15, 2233–2256, 2002.

Kwon, Y.-O., Seo, H., Ummenhofer, C. C., and Joyce, T. M.: Impact of multidecadal variability in Atlantic SST on winter atmospheric blocking, Journal of Climate, 33, 867–892, 2020.

Laloyaux, P., de Boisseson, E., Balmaseda, M., Bidlot, J.-R., Broennimann, S., Buizza, R., Dalhgren, P., Dee, D., Haimberger, L., Hersbach, H., et al.: CERA-20C: A coupled reanalysis of the twentieth century, Journal of Advances in Modeling Earth Systems, 10, 1172–1195, 2018.

Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., et al.: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, Atmospheric Chemistry and Physics, 10, 7017–7039, 2010.

Ma, L., Woollings, T., Williams, R. G., Smith, D., and Dunstone, N.: How does the winter jet stream affect surface temperature, heat flux, and sea ice in the North Atlantic?, Journal of Climate, 33, 3711–3730, 2020.

Madec, G. and the NEMO team: NEMO ocean engine version 3.6 stable, Note du Pôle de modélisation de l'Institut Pierre-Simon Laplace, 27, 2016.

Meccia, V. L., Fuentes-Franco, R., Davini, P., Bellomo, K., Fabiano, F., Yang, S., and von Hardenberg, J.: Internal multi-centennial variability of the Atlantic Meridional Overturning Circulation simulated by EC-Earth3, Climate Dynamics, pp. 1–18, 2022.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

Msadek, R., Frankignoul, C., and Li, L. Z.: Mechanisms of the atmospheric response to North Atlantic multidecadal variability: A model study, Climate dynamics, 36, 1255–1276, 2011.

Omrani, N.-E., Keenlyside, N. S., Bader, J., and Manzini, E.: Stratosphere key for wintertime atmospheric response to warm Atlantic decadal conditions, Climate Dynamics, 42, 649–663, 2014.

O'Reilly, C. H., Weisheimer, A., Woollings, T., Gray, L. J., and MacLeod, D.: 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, 2019.

Ortega, P., Robson, J., Sutton, R. T., and Andrews, M. B.: Mechanisms of decadal variability in the Labrador Sea and the wider North Atlantic in a high-resolution climate model, Climate Dynamics, 49, 2625–2647, 2017.

O'Reilly, C. H., Zanna, L., and Woollings, T.: Assessing external and internal sources of Atlantic multidecadal variability using models, proxy data, and early instrumental indices, Journal of Climate, 32, 7727–7745, 2019.

Palmer, T. and Zhaobo, S.: A modelling and observational study of the relationship between sea surface temperature in the north-west Atlantic and the atmospheric general circulation, Quarterly Journal of the Royal Meteorological Society, 111, 947–975, 1985.

Parker, T., Woollings, T., Weisheimer, A., O'Reilly, C., Baker, L., and Shaffrey, L.: Seasonal Predictability of the Winter North Atlantic Oscillation From a Jet Stream Perspective, Geophysical Research Letters, 46, 10 159–10 167, https://doi.org/10.1029/2019GL084402, 2019.

Peings, Y. and Magnusdottir, G.: Response of the Wintertime Northern Hemisphere Atmospheric Circulation to Current and Projected Arctic Sea Ice Decline: A Numerical Study with CAM5, Journal of Climate, 27, 244 – 264, https://doi.org/10.1175/JCLI-D-13-00272.1, 2014a.

Peings, Y. and Magnusdottir, G.: Forcing of the wintertime atmospheric circulation by the multidecadal fluctuations of the North Atlantic Ocean, Environmental Research Letters, 9, 034 018, 2014b.

Peings, Y. and Magnusdottir, G.: Wintertime atmospheric response to Atlantic multidecadal variability: Effect of stratospheric representation and ocean–atmosphere coupling, Climate dynamics, 47, 1029–1047, 2016.

Peings, Y., Simpkins, G., and Magnusdottir, G.: Multidecadal fluctuations of the North Atlantic Ocean and feedback on the winter climate in CMIP5 control simulations, Journal of Geophysical Research: Atmospheres, 121, 2571–2592, 2016.

Poli, P., Hersbach, H., Tan, D., Dee, D., Thépaut, J.-N., Simmons, A., Peubey, C., Laloyaux, P., Komori, T., Berrisford, P., and Dragani, R.: 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 series, 2013.

Portal, A., Pasquero, C., D'andrea, F., Davini, P., Hamouda, M. E., and Rivière, G.: Influence of Reduced Winter Land–Sea Contrast on the Midlatitude Atmospheric Circulation, Journal of Climate, 35, 2637–2651, 2022.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, Journal of Geophysical Research D: Atmospheres, 2003.

Roberts, C. D., Senan, R., Molteni, F., Boussetta, S., Mayer, M., and Keeley, S. P. E.: Climate model configurations of the ECMWF Integrated Forecasting System (ECMWF-IFS cycle 43r1) for HighResMIP, Geoscientific Model Development, 11, 3681–3712, https://doi.org/10.5194/gmd-11-3681-2018, 2018a.

Roberts, M. J., Vidale, P. L., Senior, C., Hewitt, H. T., Bates, C., Berthou, S., Chang, P., Christensen, H. M., Danilov, S., Demory, M.-E., Griffies, S. M., Haarsma, R., Jung, T., Martin, G., Minobe, S., Ringler, T., Satoh, M., Schiemann, R., Scoccimarro, E., Stephens, G., and Wehner, M. F.: The Benefits of Global High Resolution for Climate Simulation: Process Understanding and the Enabling of Stakeholder

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

Decisions at the Regional Scale, Bulletin of the American Meteorological Society, 99, 2341 – 2359, https://doi.org/10.1175/BAMS-D-15-00320.1, 2018b.

Robson, J., Menary, M. B., Sutton, R. T., Mecking, J., Gregory, J. M., Jones, C., Sinha, B., Stevens, D. P., and Wilcox, L. J.: The role of anthropogenic aerosol forcing in the 1850–1985 strengthening of the AMOC in CMIP6 historical simulations, Journal of Climate, 35, 3243–3263, 2022.

Ruggieri, P., Bellucci, A., Nicolí, D., Athanasiadis, P. J., Gualdi, S., Cassou, C., Castruccio, F., Danabasoglu, G., Davini, P., Dunstone, N., et al.: Atlantic multidecadal variability and North Atlantic jet: a multimodel view from the decadal climate prediction project, Journal of Climate, 34, 347–360, 2021.

Scaife, A. A. and Smith, D.: A signal-to-noise paradox in climate science, npj Climate and Atmospheric Science, 1, 28, 2018.

Seager, R., Kushnir, Y., Visbeck, M., Naik, N., Miller, J., Krahmann, G., and Cullen, H.: Causes of Atlantic Ocean climate variability between 1958 and 1998, Journal of Climate, 13, 2845–2862, 2000.

Sein, D. V., Koldunov, N. V., Danilov, S., Wang, Q., Sidorenko, D., Fast, I., Rackow, T., Cabos, W., and Jung, T.: Ocean Modeling on a Mesh With Resolution Following the Local Rossby Radius, Journal of Advances in Modeling Earth Systems, 9, 2601–2614, https://doi.org/10.1002/2017MS001099, 2017.

Simpson, I. R., Deser, C., McKinnon, K. A., and Barnes, E. A.: Modeled and Observed Multidecadal Variability in the North Atlantic Jet Stream and Its Connection to Sea Surface Temperatures, Journal of Climate, 31, 8313 – 8338, https://doi.org/10.1175/JCLI-D-18-0168.1, 2018.

Smith, D., Eade, R., Scaife, A. A., Caron, L.-P., Danabasoglu, G., DelSole, T., Delworth, T., Doblas-Reyes, F., Dunstone, N., Hermanson, L., et al.: Robust skill of decadal climate predictions, Npj Climate and Atmospheric Science, 2, 1–10, 2019.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, Bulletin of the American Meteorological Society, 93, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1, 2012.

Visbeck, M., Chassignet, E. P., Curry, R. G., Delworth, T. L., Dickson, R. R., and Krahmann, G.: The ocean's response to North Atlantic Oscillation variability, Geophysical Monograph-American Geophysical Union, 134, 113–146, 2003.

Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., Colin, J., Guérémy, J.-F., Michou, M., Moine, M.-P., Nabat, P., Roehrig, R., Salas y Mélia, D., Séférian, R., Valcke, S., Beau, I., Belamari, S., Berthet, S., Cassou, C., Cattiaux, J., Deshayes, J., Douville, H., Ethé, C., Franchistéguy, L., Geoffroy, O., Lévy, C., Madec, G., Meurdesoif, Y., Msadek, R., Ribes, A., Sanchez-Gomez, E., Terray, L., and Waldman, R.: Evaluation of CMIP6 DECK Experiments With CNRM-CM6-1, Journal of Advances in Modeling Earth Systems, 11, 2177–2213, https://doi.org/10.1029/2019MS001683, 2019.

Wagner, S. and Shishkina, O.: Heat flux enhancement by regular surface roughness in turbulent thermal convection, Journal of Fluid Mechanics, 763, 109–135, https://doi.org/10.1017/jfm.2014.665, 2015.

Wang, X., Li, J., Sun, C., and Liu, T.: NAO and its relationship with the Northern Hemisphere mean surface temperature in CMIP5 simulations, Journal of Geophysical Research: Atmospheres, 122, 4202–4227, 2017.

Weisheimer, A., Schaller, N., O'Reilly, C., MacLeod, D. A., and Palmer, T.: 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, https://doi.org/10.1002/qj.2976, 2017.

Weisheimer, A., Befort, D. J., MacLeod, D., Palmer, T., O'Reilly, C., and Strømmen, K.: Seasonal forecasts of the twentieth century, Bulletin of the American Meteorological Society, 101, E1413–E1426, 2020.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

Weisheimer, A.; O'Reilly, C.: Initialised seasonal forecast of the 20th Century., https://catalogue.ceda.ac.uk/uuid/6e1c3df49f644a0f812818080bed5e45, 2020.

Williams, K. D., Copsey, D., Blockley, E. W., Bodas-Salcedo, A., Calvert, D., Comer, R., Davis, P., Graham, T., Hewitt, H. T., Hill, R., Hyder, P., Ineson, S., Johns, T. C., Keen, A. B., Lee, R. W., Megann, A., Milton, S. F., Rae, J. G. L., Roberts, M. J., Scaife, A. A., Schiemann, R., Storkey, D., Thorpe, L., Watterson, I. G., Walters, D. N., West, A., Wood, R. A., Woollings, T., and Xavier, P. K.: The Met Office Global Coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) Configurations, Journal of Advances in Modeling Earth Systems, 10, 357–380, https://doi.org/10.1002/2017MS001115, 2018.

Wills, R. C., Armour, K. C., Battisti, D. S., and Hartmann, D. L.: Ocean–atmosphere dynamical coupling fundamental to the Atlantic multidecadal oscillation, Journal of Climate, 32, 251–272, 2019.

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

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

Woollings, T., Franzke, C., Hodson, D., Dong, B., Barnes, E. A., Raible, C., and Pinto, J.: Contrasting interannual and multidecadal NAO variability, Climate Dynamics, 45, 539–556, 2015.

Yeager, S.: Decadal Prediction Large Ensemble Project, https://doi.org/10.5065/D6DR2T8H, 2018.

Yeager, S.: The abyssal origins of North Atlantic decadal predictability, Climate Dynamics, 55, 2253–2271, 2020.

Yeager, S., Danabasoglu, G., Rosenbloom, N., Strand, W., Bates, S., Meehl, G., Karspeck, A., Lindsay, K., Long, M., Teng, H., et al.: Predicting near-term changes in the earth system: a large ensemble of initialized decadal prediction simulations using the community earth system model, Bulletin of the American Meteorological Society, 99, 1867–1886, 2018.

Zhang, R.: Anticorrelated multidecadal variations between surface and subsurface tropical North Atlantic, Geophysical Research Letters, 34, 2007.

Zhang, R., Delworth, T. L., Sutton, R., Hodson, D. L., Dixon, K. W., Held, I. M., Kushnir, Y., Marshall, J., Ming, Y., Msadek, R., et al.: Have aerosols caused the observed Atlantic multidecadal variability?, Journal of the Atmospheric Sciences, 70, 1135–1144, 2013.

Zhang, R., Sutton, R., Danabasoglu, G., Kwon, Y.-O., Marsh, R., Yeager, S. G., Amrhein, D. E., and Little, C. M.: A review of the role of the Atlantic meridional overturning circulation in Atlantic multidecadal variability and associated climate impacts, Reviews of Geophysics, 57, 316–375, 2019.

Zhang, W., Kirtman, B., Siqueira, L., Clement, A., and Xia, J.: Understanding the signal-to-noise paradox in decadal climate predictability from CMIP5 and an eddying global coupled model, Climate dynamics, 56, 2895–2913, 2021.

https://doi.org/10.5194/egusphere-2023-307
Preprint. Discussion started: 24 February 2023
© Author(s) 2023. CC BY 4.0 License.

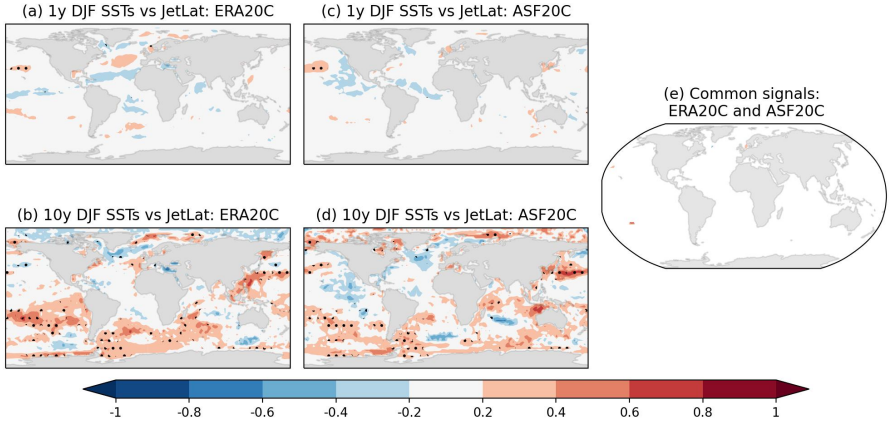

Figure B4. Correlations between the DJF jet latitude timeseries and DJF SSTs at each gridpoint. In (a) for ERA20C using seasonal data; (b) for ERA20C with a 10-year running mean applied; (c) for ASF20C using seasonal data; (d) for ASF20C with a 10-year running mean applied. In (e) are shown the correlations from (b) that are common across all subplots and significant in each subplot by itself. Stipling indicates significance using a two-tailed t-test. The period 1900-2010 is used.