

# Geophysical Research Letters<sup>®</sup>



## RESEARCH LETTER

10.1029/2022GL098212

### Key Points:

- Climate models exhibit systematically weak winter multidecadal North Atlantic (NA) atmospheric variability relative to reanalysis output
- The subseasonal timing of 20th century winter variability is sensitive to the choice of atmospheric circulation index
- Onset of strong variability occurs earliest for NA jet speed (December), followed by NAO (January) and eastern-NA jet speed (February)

### Supporting Information:

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

### Correspondence to:

T. J. Bracegirdle,  
[tjbra@bas.ac.uk](mailto:tjbra@bas.ac.uk)

### Citation:

Bracegirdle, T. J. (2022). Early-to-late winter 20th century North Atlantic multidecadal atmospheric variability in observations, CMIP5 and CMIP6. *Geophysical Research Letters*, 49, e2022GL098212. <https://doi.org/10.1029/2022GL098212>

Received 4 FEB 2022  
Accepted 17 MAY 2022

## Early-to-Late Winter 20th Century North Atlantic Multidecadal Atmospheric Variability in Observations, CMIP5 and CMIP6

Thomas J. Bracegirdle<sup>1</sup> 

<sup>1</sup>British Antarctic Survey, Cambridge, UK

**Abstract** The strong multidecadal variability in North Atlantic (NA) winter atmospheric circulation is poorly understood and appears too weak in climate models. Recent research has shown peak atmospheric multidecadal variability over the NA in late winter, particularly March, linked to Atlantic multidecadal variability (AMV) of the ocean. Here a range of NA atmospheric circulation indices are assessed to provide a comprehensive picture of early-to-late winter low-frequency variability and its representation in the latest generation of climate models (Coupled Model Intercomparison Project Phase 6 [CMIP6]). As found for CMIP5, CMIP6 models exhibit too-weak multidecadal NA atmospheric variability compared to reanalysis data over the period 1862–2005. Consistent with previous research, the eastern part of the NA westerly jet (U700NA) exhibits peak low-frequency variability in March. However, for NA-wide jet speed and the NAO, low-frequency variability and model-reanalysis discrepancies are strongest in January and February, associated with too-weak NA ocean-atmosphere linkages.

**Plain Language Summary** The strong low-frequency multidecadal variability of large-scale atmospheric circulation in the North Atlantic (NA) drives variability in climate conditions over western Europe. Reliably reproducing this variability in climate models is of high importance for estimating future climatic conditions. However, a current issue in climate modeling is that models generally exhibit weak winter multidecadal variability compared to estimates based on observationally constrained reconstructions (reanalyses). In particular recent research has shown that strong low-frequency atmospheric variability and associated discrepancies between models and reanalyses are most pronounced in late winter, particularly March. The first step in this study is to establish that the current generation of climate models still exhibit the long-standing issue that simulated multidecadal variability of NA atmospheric circulation is weaker than in reanalyses. A subseasonal analysis shows that, consistent with previous research, the eastern part of the NA westerly jet exhibits peak low-frequency variability in March. However, for NA-wide indices, tropospheric westerly jet speed and the NAO, low-frequency variability is strongest in January and February. Evidence is shown suggesting that these inter-index differences in subseasonal variability are associated with differences in monthly correlations with NA subpolar gyre SSTs.

## 1. Introduction

In the North Atlantic (NA) region, significant low-frequency (multidecadal) atmospheric variability is evident in reanalysis data sets over the historical period from the mid-19th century to present day (e.g., Woollings et al., 2014). A comprehensive understanding of this variability is of fundamental importance to reliable estimates of how climate may change over the coming decades and beyond (Deser et al., 2017; Simpson et al., 2019). However, a key issue in current research is that climate models exhibit weaker-than-observed low-frequency winter atmospheric variability over the NA (Bracegirdle et al., 2018; Cheung et al., 2017; Kravtsov, 2017; Simpson et al., 2018; Wang et al., 2017; Zhang & Wang, 2013).

Although many factors are known to contribute to NA atmospheric multidecadal variability, including internal climate variability and a range of external forcings (e.g., Sutton et al., 2018), evidence currently suggests that a key factor in the missing low-frequency variability in climate models is too-weak atmospheric linkages to NA sea-surface temperature (SST) anomalies (Kim et al., 2018; Peings et al., 2016; Wang et al., 2017). Consistent with this, Simpson et al. (2018) showed that the World Climate Research Program's (WCRP) Coupled Model Intercomparison Phase 5 (CMIP5) climate models (Taylor et al., 2012) do not broadly reproduce a reanalysis-derived

© 2022. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

strong late-winter (March) peak in multidecadal variability at the eastern part of the NA jet and its connection to indices of Atlantic Multidecadal Variability (AMV).

The analysis here is broadened to include basin-wide lower-tropospheric westerly jet speed and an assessment of multidecadal timescales characteristic of AMV. The main motivation for this is that jet speed explains a large part of the observed low-frequency component of winter NAO variability on multidecadal timescales (Woollings et al., 2012) and that Bracegirdle et al. (2018) found a clear model-reanalysis discrepancy in winter-mean low-frequency power of jet speed on AMV timescales. Therefore, three indices of atmospheric variability are included: speed of the Atlantic lower-tropospheric westerly jet, speed of the eastern part of this jet (the “U700NA” diagnostic of Simpson et al. (2018)) and the NAO. Jet latitude is not included since it exhibits weak power on multidecadal timescales (Bracegirdle et al., 2018). With respect to assessing links to ocean variability, SST variability across the NA basin is considered along with two subregions. The NA subpolar gyre (SPG), which emerges as a dominant region on multidecadal timescales that has been identified as important in correctly capturing NAO-SST coupling (Peings et al., 2016; Woollings et al., 2015) and the subtropical NA, which has been found to influence decadal variability of midlatitude circulation via wave propagation (Gastineau & Frankignoul, 2015; Peings et al., 2016).

The main aims of this study are to (a) assess whether previous climate model-reanalysis discrepancies in NA low-frequency variability are reduced in the latest generation of climate models (CMIP6, Eyring et al., 2016) and the recently released version of the 20th Century Reanalysis (20CRv3, Slivinski et al., 2019) and (b) broaden previous studies to assess low-frequency variability across individual winter months and a range of indices of NA atmospheric circulation.

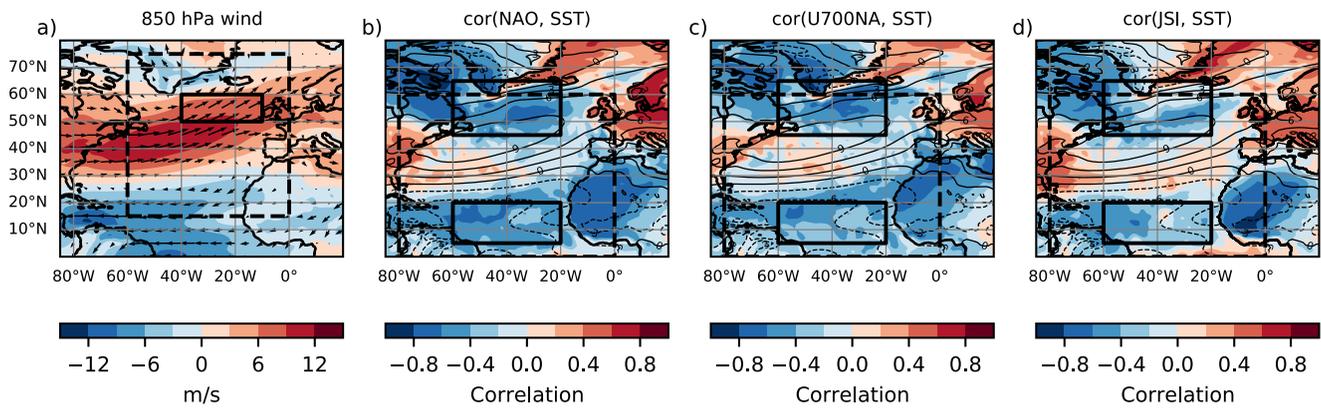
## 2. Data and Methods

### 2.1. Reanalysis Data

The recently released 20CRv3 was used here. The only input data used in 20CRv3 are surface pressure observations in combination with a state-of-the-art data assimilation and modeling system with the aim of reconstructing atmospheric conditions in a consistent manner over the period of available observations (Slivinski et al., 2019). Compared to the previous version, 20CRv2, version 3 extends further back in time to the early 19th century and has an increased ensemble size (80 members compared to 56). Due to uncertainty relating to bias corrections of mid-19th century ship observations (Slivinski et al., 2019), and to aid direct comparisons to results presented in Bracegirdle et al. (2018), analysis was restricted to 1861 onwards. The variables used were (all monthly mean): zonal wind on 850 and 700 hPa, atmospheric pressure at mean sea level and skin temperature (which over open ocean is SST). A land mask was used to exclude land areas in calculating Atlantic SST indices.

### 2.2. Climate Model Data

Equivalent variables to those extracted from the 20CRv3 data set were assessed in output from *historical* simulations of CMIP5 and CMIP6 models. The historical simulations include known major anthropogenic and natural climate forcings (such as greenhouse gas concentration increases and major volcanic eruptions) over the period since the mid nineteenth century (Eyring et al., 2016). They are free-running and therefore do not in general follow the observed phases of natural internal variability. For most CMIP models, the historical simulations were repeated a number of times with slightly different initial conditions. These “realizations” provide a measure of the role of internal climate variability in simulated and observed trends over the historical period. The CMIP variables used here are “ua” (zonal wind on pressure levels), “psl” (atmospheric pressure at mean sea level), “ts” (skin temperature of the lower boundary of the atmosphere), and “sftlf” (land area fraction, used to mask out land areas). All available realizations for which output was found with the necessary variables required for this study are listed in Tables S3 and S4 in Supporting Information S1. This comprises 312 realizations from 81 models across both data sets. Four fewer CMIP5 models were available with required data for the SST diagnostics, therefore results using these diagnostics are based on a subset of 77 CMIP models (see Tables S1 and S2 in Supporting Information S1).



**Figure 1.** (a) January 850 hPa vectors of climatological time-mean wind (arrows) and the zonal component (shading) from 20CRv3 over the period 1862–2005. The region used to define the jet speed index (JSI) is shown by the dashed-border box and the region used to define U700NA by the solid-border box. (b–d) Spatial correlation maps between low-pass (>10 years) filtered time series of atmospheric indices and grid-point surface temperature derived from January monthly mean data from the 20CRv3 reanalysis for (b) NAO, (c) U700NA, and (d) JSI. All time series are linearly detrended before calculating correlations. The dashed-border boxes in (b–d) show the region used to define SST<sub>AMV</sub>, the northern solid-border boxes define the SST<sub>SPG</sub> region and the southern solid-border boxes SST<sub>ST</sub>.

### 2.3. Atmospheric Circulation Indices

An EOF-based definition of the NAO is used in this study. It is calculated from mean sea level pressure (MSLP) over the NA, with latitude and longitude ranges following Hurrell (1995) (20–80°N, 90°W–40°E). The principal component time series of the leading EOF of year-to-year anomalies in winter-mean (DJF) or monthly mean MSLP is used to define the time series of NAO variability.

Speed of the NA tropospheric westerly jet follows the definition of Bracegirdle et al. (2018), referred to hereinafter as the jet speed index (JSI). Key aspects of this definition are as follows: (a) In order to maximize the number of CMIP models assessed, it is based on monthly means rather than the more standard daily mean fields. (b) It draws from the definition of Woollings et al. (2015) with zonal averaging between 60°W and 0°W at the 850 hPa level (region shown in Figure 1a). (c) For seasonal winter means, jet diagnostics are first calculated for each month before averaging over the standard winter months December–February.

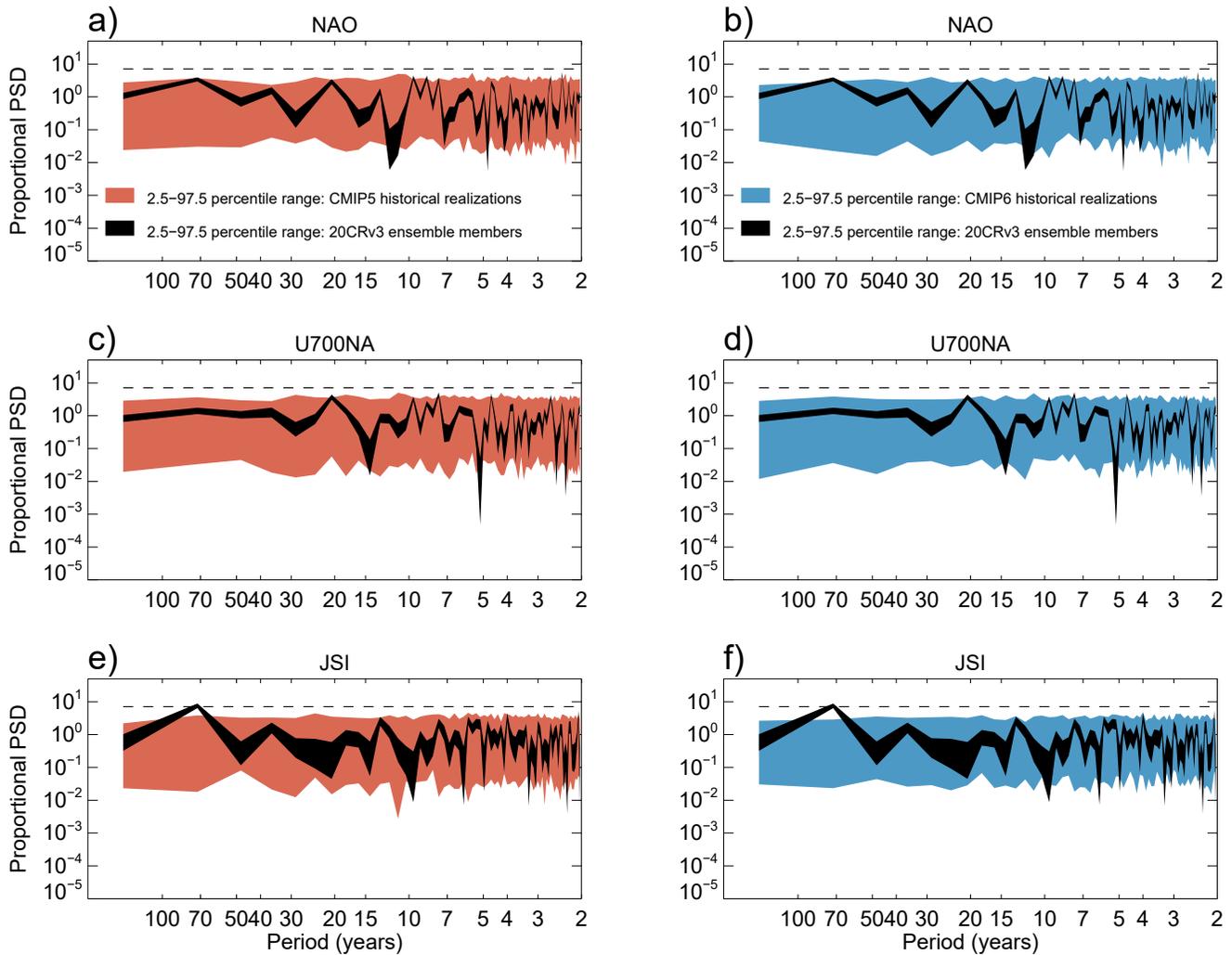
The westerly wind index used by Simpson et al. (2018), U700NA, is also included here. It is defined as the area-weighted spatial mean westerly wind component at 700 hPa in box 50°–60°N and 40°–10°W (region shown in Figure 1a).

### 2.4. Sea-Surface Temperature Diagnostics

Three indices of SSTs in the NA are used, the locations of which are indicated in Figures 1b–1d. For linking SST and atmospheric variability, Peings et al. (2016) suggest two regions of the NA that are important for climate models to reproduce the observed strength of NAO–SST coupling, the SPG and subtropical NA. They are:

- (1) A basin-wide index of NA SSTs (SST<sub>AMV</sub>), which is an area-weighted spatial mean (with land areas masked out) over the region traditionally used to define AMV indices (0°–60°N and 80°W–0°W) (Qasmi et al., 2017).
- (2) An index of NA subpolar gyre SSTs (SST<sub>SPG</sub>), which is an area-weighted spatial mean over the region 45°–65°N, 60°–20°W. This definition is chosen based on the region of the gyre as identified by Biri and Klein (2019).
- (3) An index of NA subtropical SSTs (SST<sub>ST</sub>). This is a spatial mean over the region 5°–20°N, 60°–20°W and broadly captures the southern section of the characteristic AMV horseshoe pattern (Gastineau & Frankignoul, 2015), which is evident in Figure 1b.

To generate time series of these indices, each calendar month is treated separately (e.g., time series of all the Januaries over the range of years) after which a low-pass filter is applied to retain multidecadal timescales (>10 years). Each time series is linearly detrended before application of the filter.

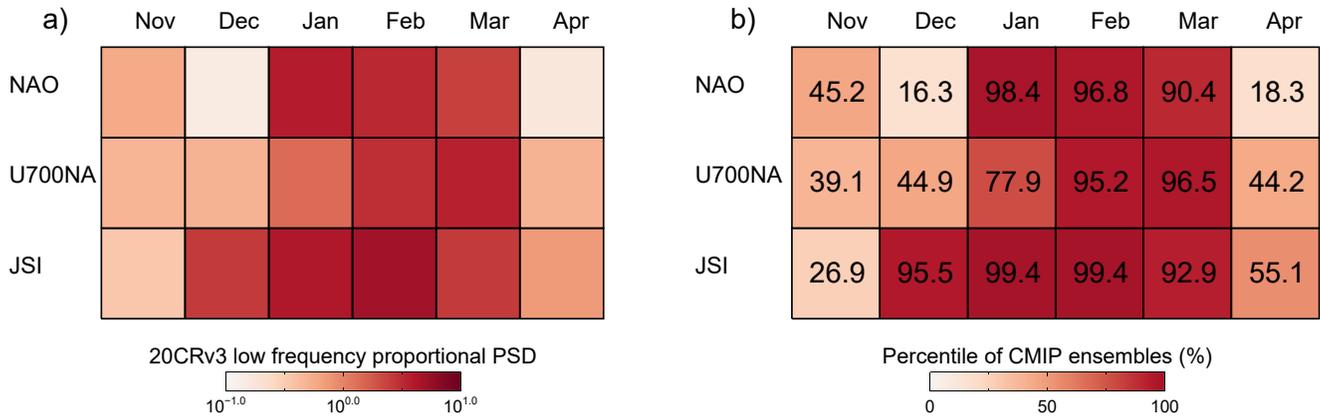


**Figure 2.** Proportional power spectral density from (a, b) winter (DJF) NAO index, (c, d) U700NA, and (e, f) jet speed index (JSI) from Coupled Model Intercomparison Project Phase 5 (CMIP5) (left column) and CMIP6 (right column) model realizations. The 2.5th–97.5th percentile ranges across each ensemble are shown by the colored shading (red for CMIP5 and blue for CMIP6). The black shading shows the 2.5th–97.5th percentile range across all 80 ensemble members of 20CRv3. Each time series was linearly detrended prior to the spectral analysis, and the time range was restricted to the common period of 1862–2005 for comparison of the same frequency bins (note that the year refers to January of the winter in question). The dashed line indicates the  $p = 0.05$  threshold for the significance of a peak in any particular frequency bin using Fisher’s  $g$ -statistic.

### 3. Results

In the first part of this section, relevant winter-mean diagnostics from Bracegirdle et al. (2018) are repeated for the updated climate models (CMIP6 in place of CMIP5) and reanalysis (20CRv3 in place of 20CRv2c). This shows whether model-reanalysis discrepancies in low-frequency NA variability have changed significantly in updated versions of climate models and reanalyses. This is followed by a month-by-month early-to-late winter assessment of reanalysis-derived multidecadal variability and its representation in CMIP climate models. To maximize the number of realizations available, this further analysis combines historical simulations from both CMIP5 and CMIP6 (referred to hereinafter as CMIP).

Figure 2 shows proportional power spectral density (PSD) periodograms from time series of winter-mean (DJF) atmospheric circulation indices (to aid comparison between periodograms of different diagnostics, the displayed “proportional PSD” values are proportional to the mean of the PSD values across the frequency bins in each periodogram). This shows that the new CMIP data set (CMIP6) produces very similar results to those based on the preceding generation (CMIP5). The main point to note is that, as found for CMIP5, the CMIP6 historical



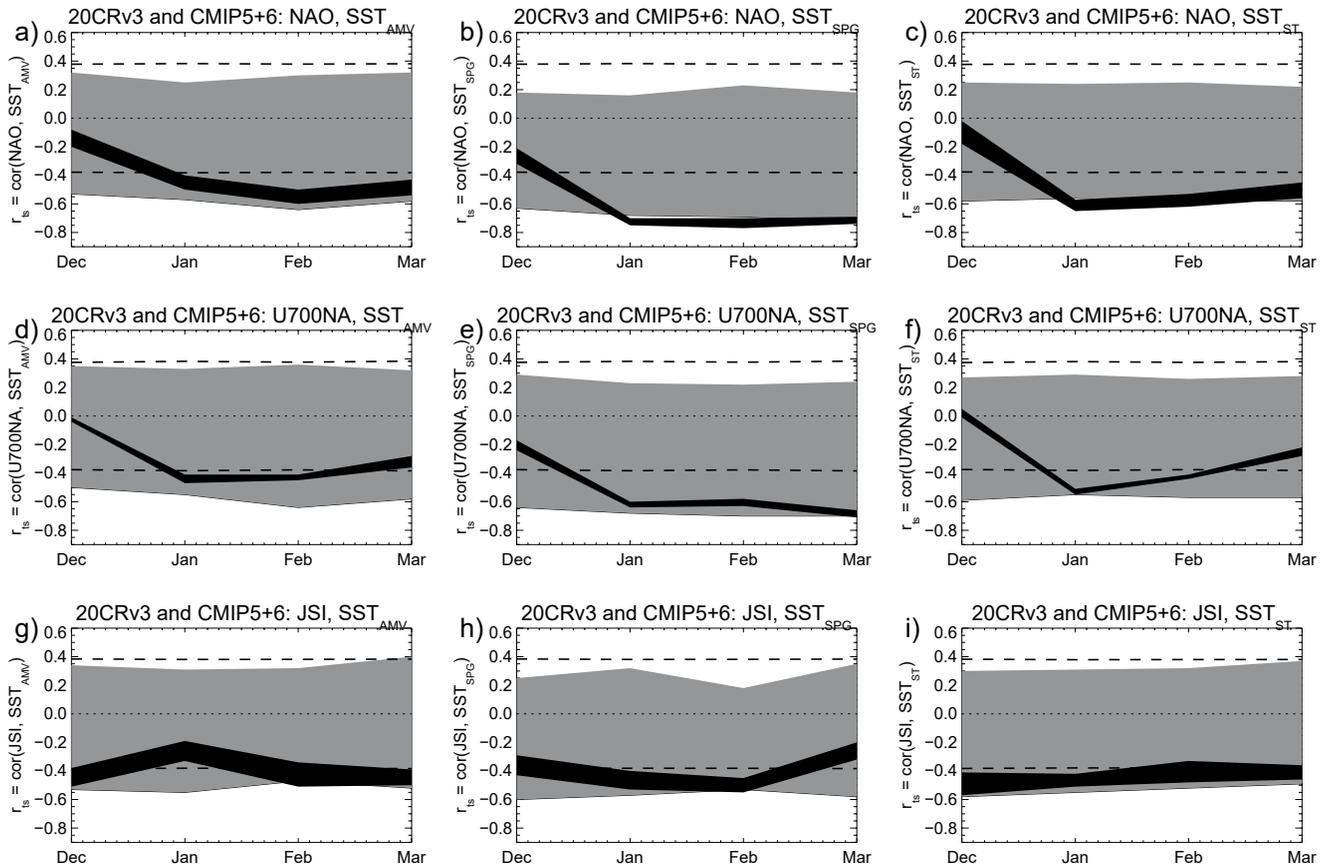
**Figure 3.** (a) Monthly 20CRv3 ensemble median low-frequency proportional power spectral density (PSD) for periods of  $\sim 70$  years (b) 20CRv3 low-frequency power relative to the combined Coupled Model Intercomparison Project Phase 5 (CMIP5) and CMIP6 ensembles, expressed in terms of percentiles of the distributions across the 312 realizations.

realizations do not reproduce the statistically significant low-frequency peak in JSI that is evident in current and previous versions of 20CR (Bracegirdle et al., 2018; Woollings et al., 2014). At this  $\sim 70$  year period there is a clear separation between the 2.5th and 97.5th percentile ranges of the CMIP climate model realizations and the 80 members of the 20CRv3 ensemble (Figures 2e and 2f). Indeed, it is notable that the reanalysis ensemble spread is particularly narrow for this timescale, demonstrating high confidence as was seen for the earlier 20CRv2c version used in Bracegirdle et al. (2018). For NAO and U700NA, this is less clear for these winter-mean diagnostics. A spectral analysis of the three AMV-related SST diagnostics shows that the observed significant multidecadal ocean variability is stronger than most CMIP ensemble realizations (consistent with power spectra assessments of CMIP5 shown in Zhang and Wang (2013) and Cheung et al. (2017)), but with no clear separation between the 20CRv3 and CMIP ensembles (Figure S1 in Supporting Information S1). This illustrates that although part of the 20CRv3-CMIP discrepancy in low-frequency JSI variability may be due to weak ocean variability, other factors are likely to play a role (discussed in Section 5).

Figure 3 reveals clear month-by-month contrasts in the strength of low-frequency variability exhibited by the three atmospheric circulation indices. Figure 3a shows that low-frequency power in 20CRv3 is strong in all winter months (DJFM) for JSI. For the NAO, low-frequency variability is also strongest in mid-to-late winter, but is weak in December. This largely explains the weaker low-frequency power in the NAO for the DJF means shown in Figure 2. For U700NA, the low-frequency power is strongest in late winter, peaking in March. In canonical winter months (December-February), the low-frequency power is stronger for JSI than for NAO or U700NA, but in March U700NA is the strongest.

The months and diagnostics with the strongest low-frequency power in 20CRv3 are also, perhaps unsurprisingly, the months for which CMIP models struggle most to reproduce the power seen in 20CRv3 (Figure 3b). In January, both the NAO and JSI reanalysis-estimated low-frequency power exceeds the 98th percentile of the distribution of the 312 CMIP realizations. For JSI it exceeds the 99th percentile. For both indices, the CMIP-relative power remains above the 90th percentile through March before dropping dramatically in April. For U700NA the 20CRv3 low-frequency power reaches the 96.5th percentile in March. The overall picture is then that the CMIP models do not in general reproduce the strength of reanalysis-derived low-frequency atmospheric variability in January, February and March. The onset of strong variability relative to the CMIP ensemble range occurs earlier for JSI (December) and later for NAO (January) and U700NA (February).

To assess potential links between climate model-reanalysis discrepancies in low-frequency variability and model representation of AMV-atmosphere linkages, Figure 4 shows reanalysis and CMIP-derived correlations between time series of NA SST and atmospheric circulation indices, where the time series have been low-pass filtered at decadal ( $>10$  years) timescales. They are shown for each month between December and March. Negative and statistically significant reanalysis-derived correlations between the three atmospheric indices and  $SST_{AMV}$  are evident in Figures 4a, 4d, and 4g, most consistently in February. In this month both NAO and JSI exhibit



**Figure 4.** Monthly correlations from low-pass filtered (>10 years) time series of atmospheric and sea-surface temperature (SST) indices. The left column shows correlations with  $SST_{AMV}$ , the middle column with  $SST_{SPG}$ , and the right column with  $SST_{ST}$ . Each time series was linearly detrended prior to filtering and the time range was restricted to the common period of 1862–2005. The black (gray) shading shows the 2.5th–97.5th percentile range of the 20CRv3 (Coupled Model Intercomparison Project Phase (CMIP) historical) ensemble members. The dashed lines show the threshold for statistically significant correlations at the  $p = 0.01$  level based on a Student's  $t$  test and taking into account lag 1 autocorrelation.

correlations that approach the lower part of the 2.5th–97.5th percentile range of the combined CMIP5 and CMIP6 data sets.

Consistent with this, broadly stronger correlations are seen for the NA SST subregions ( $SST_{SPG}$  and  $SST_{ST}$ ) (Figures 4b, 4c, 4e, 4f, 4h, and 4i). For the NAO, reanalysis-derived correlations with  $SST_{SPG}$  and  $SST_{ST}$  are strongest in January and February, although only marginally compared to March (Figures 4b and 4c). Notably, in January the reanalysis ensemble lies outside the 2.5th–97.5th percentile range of the CMIP model realizations, particularly for  $SST_{SPG}$ . For JSI, the reanalysis correlations are strongest in January and February for  $SST_{SPG}$ , but do not clearly exceed the CMIP range (Figure 4h). For  $SST_{ST}$  correlations with JSI are strongest and most significant in December and January (Figure 4f), suggesting more of a subtropical linkage for the strong early season JSI variability seen in Figure 3. For U700NA, the strongest links to SST variability are for the SPG region in March, when reanalysis-derived correlations also lie closest to the edge of the CMIP ensemble range (Figure 4e). The strong correlations between U700NA and  $SST_{SPG}$  are not unexpected since there is considerable overlap between the two regions (Figure 1).

Overall Figures 3 and 4 present a picture of strong low-frequency atmospheric variability in mid-to-late winter that is associated with strong atmosphere-SST correlations in the same months, particularly for the SPG region, but with significant links to the subtropical NA.

#### 4. Conclusions and Discussion

In this paper, an assessment of winter low-frequency atmospheric variability over the NA is presented comparing historical simulations from CMIP5 and CMIP6 with reanalysis data from 20CRv3 over the period 1862–2005. The aims are (a) to assess whether previous climate model–reanalysis discrepancies in NA low-frequency variability are reduced in the current generation of climate models and (b) broaden previous studies to assess low-frequency variability across individual winter months and a range of indices of NA atmospheric circulation.

With regard to the first aim, it is shown that the previously documented climate model–reanalysis (CMIP5–20CRv2c) discrepancy in low-frequency variability in winter-mean JSI (Bracegirdle et al., 2018) is still apparent for the current generation data sets (CMIP6 and 20CRv3). It therefore remains important to develop a clearer characterization of this discrepancy, which is a key motivation for the second aim of conducting a broader month-by-month analysis.

Considering individual months from December to March, it was found that reanalysis-derived jet speed (JSI), NAO, and U700NA all exhibit strong multidecadal variability at different points during winter; clearest in both January and February for JSI and NAO and in March for U700NA. Overall JSI variability is strong and constant through the winter from December to March, which helps to explain its dominance in winter-mean diagnostics such as the DJF means used by Bracegirdle et al. (2018). For the NAO, reanalysis-derived multidecadal low-frequency power and associated CMIP model discrepancies are similar in magnitude to that seen for JSI. For the CMIP ensemble mean there is no significant month-by-month variability in low-frequency power and large interrealization spreads for specific models (not shown). This provides evidence for a major role for internal variability in the specific timing of low-frequency power in the historical record. However, the caveat being that the relatively weak low-frequency power in climate models could be a consequence of underrepresentation of forced variability over the historical period.

One puzzle of the late-winter peak in eastern-NA (U700NA) atmospheric variability identified by Simpson et al. (2018) is difficulty in explaining why it should occur preferentially in late winter rather than mid-winter. As they suggest, one possibility is that anomalies grow during the season in response to anomalous SSTs (e.g., Woollings et al., 2012). Taking a basin-wide perspective, the strong mid-to-late winter JSI variability found here is more concurrent with months of strong atmosphere–ocean interactions, with climatologically high ocean–atmosphere heat fluxes and large mixed layer depths widespread across the northern NA in December–March (e.g., de Boyer Montegut et al., 2004). Consistent with this, Gastineau and Frankignoul (2015) showed evidence that on decadal timescales NA atmosphere–SST links are strongest in mid-to-late winter (JFM) associated with a clearer annual reemergence of longer-term (decadal) ocean temperature anomalies below the seasonal thermocline as the mixed layer deepens in winter. This may help to explain the marked increase in strength in SST–atmosphere correlations for NAO and U700NA between December and January in Figure 4.

In mid-to-late winter months historical realizations of the CMIP5 and CMIP6 models generally fail to reproduce the reanalysis-derived low-frequency variability (Figure 3b) and associated links to SSTs (Figure 4). Since CMIP6 models still exhibit weak low-frequency variability compared to reanalysis data, the priority of explaining the general lack of NA multidecadal variability in climate models remains. The results of the spectral analysis and atmosphere–ocean time series correlations presented here are consistent with previous evidence for explanations involving both relatively weak CMIP-simulated NA SST variability (Cheung et al., 2017; Zhang & Wang, 2013) and too-weak model-simulated atmospheric responses to anomalous SSTs (Kim et al., 2018; Peings et al., 2016; Simpson et al., 2018; Wang et al., 2017). On the latter point a key piece of evidence is that even with the historical evolution of SSTs prescribed, models still do not reproduce the magnitude of observed NA atmospheric variability (Kim et al., 2018; Simpson et al., 2018). However, a clear causal link to specific model deficiencies has not yet been made. Equatorward jet latitude biases may play a role (Ruggieri et al., 2021), whereby models with more equatorward jets exhibit weaker atmospheric responses to AMV forcing. Equatorward NA storm track and jet latitude biases that were identified in CMIP5 do indeed persist in CMIP6, although reduced slightly (Bracegirdle et al., 2022; Priestley et al., 2020; Simpson et al., 2020), suggesting a potential role for this effect in explaining the continued weak low-frequency atmospheric variability in CMIP. However, no clear link was found here between the low-frequency jet speed variability shown in Figure 3 and jet latitude across the CMIP models (not shown), although the large internal variability in models with small ensemble sizes introduces significant uncertainty that could obscure any effect. Other suggested, and potentially related, contributory factors include model

resolution (Czaja et al., 2019; Scaife et al., 2019), the representation of stratosphere-troposphere interactions (Omrani et al., 2014), positive coupled ocean-atmosphere feedbacks (e.g., Osso et al., 2020) and teleconnections to the rest of the planet (Hardiman et al., 2019; Hurrell et al., 2004).

A major implication of the results presented in this study is that missing multidecadal variability in climate models potentially affects different regions at different points in the winter season depending on regional links to large-scale patterns of variability. It is shown that long-standing model-reanalysis discrepancies still persist in the latest generation of climate models (CMIP6). It is therefore imperative that research continues in this area to (a) better understand the natural and anthropogenic causes of NA multidecadal variability in the historical period, (b) evaluate model representation of these causes, (c) continue to refine long-term reanalyses and reconstructions, and (d) identify ways to improve models and/or adjust predictions to take account of their biases.

## Data Availability Statement

This study used existing openly available data sets. The data that support the findings of this study are openly available at the following URLs: CMIP5 from <https://esgf-node.llnl.gov/search/cmip5/>; CMIP6 from <https://esgf-index1.ceda.ac.uk/search/cmip6-ceda/>; and 20CRv3 from [https://portal.nerc.gov/project/20C\\_Reanalysis/](https://portal.nerc.gov/project/20C_Reanalysis/). The specific CMIP5 and CMIP6 run and version numbers of all 312 simulations that were analyzed are listed in Tables S3 and S4 in Supporting Information S1.

## Acknowledgments

T. J. B. was supported as part of the UK Natural Environment Research Council (NERC) ACSIS project (NE/N018028/1 and NE/N018001/1) and additionally through the NERC British Antarctic Survey research programme Polar Science for Planet Earth. The author acknowledges the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP5 and CMIP6. The climate modeling groups are also thanked for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP5, CMIP6, and ESGF. The Centre for Environmental Data Analysis (CEDA) and JASMIN provided the platform for much of the data analysis conducted. Support for the Twentieth Century Reanalysis Project version 3 data set is provided by the U.S. Department of Energy, Office of Science Biological and Environmental Research (BER), by the National Oceanic and Atmospheric Administration Climate Program Office, and by the NOAA Physical Sciences Laboratory.

## References

- Biri, S., & Klein, B. (2019). North Atlantic sub-polar Gyre climate index: A new approach. *Journal of Geophysical Research: Oceans*, *124*, 4222–4237. <https://doi.org/10.1029/2018JC014822>
- Bracegirdle, T. J., Lu, H., Eade, R., & Woollings, T. (2018). Do CMIP5 models reproduce observed low-frequency North Atlantic jet variability? *Geophysical Research Letters*, *45*, 7204–7212. <https://doi.org/10.1029/2018GL078965>
- Bracegirdle, T. J., Lu, H., & Robson, J. (2022). Early-winter North Atlantic low-level jet latitude biases in climate models: Implications for simulated regional atmosphere-ocean linkages. *Environmental Research Letters*, *17*, 014025. <https://doi.org/10.1088/1748-9326/ac417f>
- Cheung, A. H., Mann, M. E., Steinman, B. A., Frankcombe, L. M., England, M. H., & Miller, S. K. (2017). Comparison of low-frequency internal climate variability in CMIP5 models and observations. *Journal of Climate*, *30*(12), 4763–4776. <https://doi.org/10.1175/JCLI-D-16-0712.1>
- Czaja, A., Frankignoul, C., Minobe, S., & Vanniere, B. (2019). Simulating the midlatitude atmospheric circulation: What might we gain from high-resolution modeling of air-sea interactions? *Current Climate Change Reports*, *5*(4), 390–406. <https://doi.org/10.1007/s40641-019-00148-5>
- de Boyer Montegut, C., Madec, G., Fischer, A. S., Lazar, A., & Iudicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of Geophysical Research*, *109*, C12003. <https://doi.org/10.1029/2004JC002378>
- Deser, C., Hurrell, J. W., & Phillips, A. S. (2017). The role of the North Atlantic Oscillation in European climate projections. *Climate Dynamics*, *49*(9–10), 3141–3157. <https://doi.org/10.1007/s00382-016-3502-z>
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model Intercomparison project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. <https://doi.org/10.5194/gmd-9-1937-2016>
- Gastineau, G., & Frankignoul, C. (2015). Influence of the north Atlantic SST variability on the atmospheric circulation during the twentieth century. *Journal of Climate*, *28*(4), 1396–1416. <https://doi.org/10.1175/JCLI-D-14-00424.1>
- Hardiman, S. C., Dunstone, N. J., Scaife, A. A., Smith, D. M., Ineson, S., Lim, J., & Fereday, D. (2019). The impact of strong El Niño and La Niña events on the north Atlantic. *Geophysical Research Letters*, *46*, 2874–2883. <https://doi.org/10.1029/2018GL081776>
- Hurrell, J. W. (1995). Decadal trends in the North-Atlantic oscillation - Regional temperatures and precipitation. *Science*, *269*, 676–679. <https://doi.org/10.1126/science.269.5224.676>
- Hurrell, J. W., Hoerling, M. P., Phillips, A. S., & Xu, T. (2004). Twentieth century North Atlantic climate change. Part 1: Assessing determinism. *Climate Dynamics*, *23*(3–4), 371–389. <https://doi.org/10.1007/s00382-004-0432-y>
- Kim, W. M., Yeager, S., Chang, P., & Danabasoglu, G. (2018). Low-frequency North Atlantic climate variability in the Community Earth system model large ensemble. *Journal of Climate*, *31*, 787–813. <https://doi.org/10.1175/JCLI-D-17-0193.1>
- Kravtsov, S. (2017). Pronounced differences between observed and CMIP5-simulated multidecadal climate variability in the twentieth century. *Geophysical Research Letters*, *44*, 5749–5757. <https://doi.org/10.1002/2017GL074016>
- Omran, N. E., Keenlyside, N. S., Bader, J., & Manzini, E. (2014). Stratosphere key for wintertime atmospheric response to warm Atlantic decadal conditions. *Climate Dynamics*, *42*(3–4), 649–663. <https://doi.org/10.1007/s00382-013-1860-3>
- Osso, A., Sutton, R., Shaffrey, L., & Dong, B. W. (2020). Development, amplification, and decay of Atlantic/European summer weather patterns linked to spring North Atlantic sea surface temperatures. *Journal of Climate*, *33*(14), 5939–5951. <https://doi.org/10.1175/JCLI-D-19-0613.1>
- Peings, Y., Simpkins, G., & Magnusdottir, G. (2016). 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. <https://doi.org/10.1002/2015JD024107>
- Priestley, M. D. K., Ackerley, D., Catto, J. L., Hodges, K. I., McDonald, R. E., & Lee, R. W. (2020). An overview of the extratropical storm tracks in CMIP6 historical simulations. *Journal of Climate*, *33*(15), 6315–6343. <https://doi.org/10.1175/JCLI-D-19-0928.1>
- Qasmi, S., Cassou, C., & Boe, J. (2017). Teleconnection between Atlantic multidecadal variability and European temperature: Diversity and evaluation of the coupled model Intercomparison project phase 5 models. *Geophysical Research Letters*, *44*, 11140–11149. <https://doi.org/10.1002/2017GL074886>
- Ruggieri, P., Bellucci, A., Nicoli, D., Athanasiadis, P. J., Gualdi, S., Cassou, C., et al. (2021). Atlantic multidecadal variability and North Atlantic jet: A multimodel view from the decadal climate prediction project. *Journal of Climate*, *34*(1), 347–360. <https://doi.org/10.1175/JCLI-D-19-0981.1>

- Scaife, A. A., Camp, J., Comer, R., Davis, P., Dunstone, N., Gordon, M., et al. (2019). Does increased atmospheric resolution improve seasonal climate predictions? *Atmospheric Science Letters*, *20*(8), e922. <https://doi.org/10.1002/asl.922>
- Simpson, I. R., Bacmeister, J., Neale, R. B., Hannay, C., Gettelman, A., Garcia, R. R., et al. (2020). An evaluation of the large-scale Atmospheric circulation and its variability in CESM2 and other CMIP models. *Journal of Geophysical Research: Atmospheres*, *125*, e2020JD032835. <https://doi.org/10.1029/2020JD032835>
- Simpson, I. R., Deser, C., McKinnon, K. A., & Barnes, E. A. (2018). Modeled and observed multidecadal variability in the north Atlantic jet stream and its connection to sea surface temperatures. *Journal of Climate*, *31*(20), 8313–8338. <https://doi.org/10.1175/JCLI-D-18-0168.1>
- Simpson, I. R., Yeager, S. G., McKinnon, K. A., & Deser, C. (2019). Decadal predictability of late winter precipitation in Western Europe through an ocean-jet stream connection. *Nature Geoscience*, *12*(8), 613–619. <https://doi.org/10.1038/s41561-019-0391-x>
- Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., et al. (2019). Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, *145*(724), 2876–2908. <https://doi.org/10.1002/qj.3598>
- Sutton, R. T., McCarthy, G. D., Robson, J., Sinha, B., Archibald, A., & Gray, L. J. (2018). Atlantic multi-decadal variability and the UK ACSIS programme. *Bulletin of the American Meteorological Society*, *99*(2), 415–425. <https://doi.org/10.1175/bams-d-16-0266.1>
- Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). 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>
- Wang, X. F., Li, J. P., Sun, C., & Liu, T. (2017). NAO and its relationship with the Northern Hemisphere mean surface temperature in CMIP5 simulations. *Journal of Geophysical Research: Atmospheres*, *122*, 4202–4227. <https://doi.org/10.1002/2016JD025979>
- Woollings, T., Czuchnicki, C., & Franzke, C. (2014). Twentieth century North Atlantic jet variability. *Quarterly Journal of the Royal Meteorological Society*, *140*(680), 783–791. <https://doi.org/10.1002/qj.2197>
- Woollings, T., Franzke, C., Hodson, D. L. R., Dong, B., Barnes, E. A., Raible, C. C., & Pinto, J. G. (2015). Contrasting interannual and multidecadal NAO variability. *Climate Dynamics*, *45*(1–2), 539–556. <https://doi.org/10.1007/s00382-014-2237-y>
- Woollings, T., Gregory, J. M., Pinto, J. G., Reyers, M., & Brayshaw, D. J. (2012). Response of the North Atlantic storm track to climate change shaped by ocean-atmosphere coupling. *Nature Geoscience*, *5*(5), 313–317. <https://doi.org/10.1038/ngeo1438>
- Zhang, L. P., & Wang, C. Z. (2013). Multidecadal North Atlantic sea surface temperature and Atlantic meridional overturning circulation variability in CMIP5 historical simulations. *Journal of Geophysical Research: Oceans*, *118*, 5772–5791. <https://doi.org/10.1002/jgrc.20390>