ABSTRACT: To examine the atmospheric responses to Arctic sea ice variability in the Northern Hemisphere cold season (from October to the following March), this study uses a coordinated set of large-ensemble experiments of nine atmospheric general circulation models (AGCMs) forced with observed daily varying sea ice, sea surface temperature, and radiative forcings prescribed during the 1979–2014 period, together with a parallel set of experiments where Arctic sea ice is substituted by its climatology. The simulations of the former set reproduce the near-surface temperature trends in reanalysis data, with similar amplitude, and their multimodel ensemble mean (MMEM) shows decreasing sea level pressure over much of the polar cap and Eurasia in boreal autumn. The MMEM difference between the two experiments allows isolating the effects of Arctic sea ice loss, which explain a large portion of the Arctic warming trends in the lower troposphere and drive a small but statistically significant weakening of the wintertime Arctic Oscillation. The observed interannual covariability between sea ice extent in the Barents–Kara Seas and lagged atmospheric circulation is distinguished from the effects of confounding factors based on multiple regression, and quantitatively compared to the covariability in MMEMs. The interannual sea ice decline followed by a negative North Atlantic Oscillation–like anomaly found in observations is also seen in the MMEM differences, with consistent spatial structure but much smaller amplitude. This result suggests that the sea ice impacts on trends and interannual atmospheric variability simulated by AGCMs could be underestimated, but caution is needed because internal atmospheric variability may have affected the observed relationship.

KEYWORDS: Arctic; Sea ice; Atmospheric circulation; Climate models

1. Introduction

The Arctic climate has experienced profound changes over the past decades, as its surface air temperature has risen 2–3 times faster than the global averaged temperature (Serreze et al. 2009; Screen and Simmonds 2010; Cowtan and Way 2013) and sea ice extent and thickness have significantly decreased (Meehl et al. 2007; Serreze et al. 2007; Stroeve et al. 2012). The Arctic sea ice decline has been shown to exert strong impacts on local weather, ecosystem, human communities, industrial activities, and polar commercial shipping routes (e.g., Jung et al. 2016). The Arctic warming and sea ice decline are expected to continue in response to increasing greenhouse gas (GHG) concentration and multiple reinforcing feedbacks (Pithan and Mauritsen 2014; Döscher et al. 2014; Wendisch et al. 2017; Stuecker et al. 2018; Dai et al. 2019). A summer ice-free Arctic could be reached around the middle of the twenty-first century based on the consensus of climate model projections (e.g., Overland and Wang 2013; Stroeve et al. 2012; Screen and Deser 2019).

In contrast, the remote influence of Arctic sea ice changes on the midlatitudes and the tropics remains controversial.
cases using the same time-varying evolution of sea surface temperature (SST), greenhouse gases (GHGs), aerosol, and other forcings [often called the Atmospheric Model Intercomparison Project (AMIP)-style simulations] (e.g., Screen et al. 2014; Perlwitz et al. 2015; Sun et al. 2016). If models behave sufficiently linearly and enough ensemble members are available, the differences of atmospheric trends between the former and the latter should reflect the effects of the sea ice decline. These studies suggested a large local near-surface influence of the Arctic sea ice loss, but little impact on the large-scale atmospheric circulation. However, they only used two different AGCMs, so that the number of ensemble members was limited. Note that a larger multimodel ensemble was used by Ogawa et al. (2018), but as they compared the response to the time-varying SST evolution with that to the SST climatological cycle while the SIC evolution and transient forcings were prescribed, the impact of sea ice loss could not be directly estimated. Hence, further assessment of the performance of the state-of-the-art AGCMs in simulating the effects of Arctic sea ice decline on the atmospheric circulation is needed.

Because it is difficult to single out the Arctic sea ice effects on the long-term trends in observations, one can attempt to establish model fidelity by comparing their response to interannual SIC fluctuations to that estimated from observations, using statistical methods. Indeed, because the long-term trends can be largely removed before analysis, the influence of GHGs, aerosols, and other slowly varying external forcings is likely greatly reduced, and attribution of atmospheric circulation fluctuations to interannual sea ice changes should be easier, provided cause and effect can be distinguished despite large internal atmospheric variability, and possible confounding factors are taken into account (e.g., Simon et al. 2020).

Many previous studies have suggested an influence of prior Arctic SIC anomalies, in particular over the Barents–Kara (BK), Greenland, and Labrador Seas, on the observed atmospheric circulation in the following winter. For example, using lagged regression analysis to distinguish cause and effect, Honda et al. (2009) found that SIC anomalies averaged along the Siberian coast were followed by significant near-surface cooling in Eurasia. This circulation feature was associated with an intensified Siberian high and a negative NAO. Using lagged maximum covariance analysis (MCA) to separate the atmosphere driving sea ice (the stronger signal) from sea ice driving the atmosphere (a weaker signal), Frankignoul et al. (2014) found that an NAO-driven sea ice seesaw between the Greenland–BK Seas and the Labrador Sea was influencing the NAO later in the season, in addition to an Aleutian–Icelandic low seesaw-like response to SIC changes in the Bering and Okhotsk Seas. Both signals seemed to be primarily driven by SIC anomalies rather than by concomitant SST anomalies. Lagged MCA was also used in García-Serrano et al. (2015) to show that a BK SIC reduction in November was followed by a significant negative NAO signal in winter via a stratospheric pathway, broadly similar to other studies using lagged regressions (King et al. 2016; Nakamura et al. 2016; Yang et al. 2016; Koenigk et al. 2016). On the other hand, Blackport et al. (2019), using the turbulent heat flux to infer causality, found only minimal influence of sea ice fluctuations on the midlatitude wintertime circulation at a one-month lag.
Interannual SIC fluctuations are often synchronous with significant anomalies associated with other forcing agents. The snow cover, for example, may enhance (e.g., Furtado et al. 2016) or even dominate the SIC link with lagged atmospheric circulation (e.g., Gastineau et al. 2017). Ural blocking during boreal autumn can increase upward planetary wave propagation into the stratosphere, which weakens the stratospheric polar vortex and affects the atmospheric circulation in the following winter (Garfinkle et al. 2010). Peings (2019) argued that, as Ural blocking also affects the BK SIC (e.g., Chen et al. 2018; Gong and Luo 2017; Luo et al. 2017), Ural blocking–induced variability in autumn BK SIC and later atmospheric circulation could be incorrectly interpreted as a BK SIC impact on the atmospheric circulation. Other studies argued that, as the atmospheric variability can drive SIC variations at multiple time scales, causality is questionable (e.g., Sorokina et al. 2016; Blackport et al. 2019; Peings 2019; Screen and Blackport 2019; Zappa et al. 2021). In addition, the sea ice–atmospheric circulation linkage presents nonstationarity and intermittency (Kolstad and Screen 2019; Siew et al. 2020). This suggests that the influence of SIC variability on the atmospheric circulation needs to be further clarified.

At least two empirical approaches have been used to address the effects of the confounding factors and the issue of causality. The first one is using causal effect networks to provide clear causal directions across each component of variability. For instance, Kretschmer et al. (2016) found that a low autumn BK SIC was an important driver of the winter negative AO response via tropospheric mechanisms, while Eurasian snow cover had no clear link with the AO, but Siew et al. (2020) found that the SIC influence is intermittent and often occurred through a stratospheric pathway. The second approach, adopted in this study, is using multiple lag regressions to statistically separate contributions from each driver (e.g., Gastineau et al. 2017; Simon et al. 2020). As some studies have investigated the possibility that GCMs could underestimate the SIC-driven variability in the atmosphere (e.g., Honda et al. 2009; Mori et al. 2014, 2019a,b), while other indicated that the simple statistical relationships in observations overestimate the connections (e.g., Blackport and Screen 2021), it is also of much interest to establish with carefully designed experimental designs whether or not the interannual relationship between SIC and lagged atmospheric circulation can be adequately simulated and singled out in AGCMs and CGCMs.

In this study, we use a coordinated set of multimodel historical AGCM simulations performed in the Blue-Action Project (http://blueaction.eu/) in order to enhance our understanding on the SIC impacts and address the causality issues. These experiments were designed to single out the influence of the Arctic sea ice variations on the atmospheric circulation during the 1979–2014 period, following a protocol developed by the Blue-Action Project, which is similar to that used by Perlwitz et al. (2015) and Sun et al. (2016), and recommended by Smith et al. (2019). Nine state-of-the-art AGCMs were used, each contributing 10–30 members for a total of 165 multimodel historical simulations (see Liang et al. 2020). This larger AGCM dataset than previously available is used in this study to investigate the impacts of observed Arctic sea ice variability, including long-term trends and interannual variability, with a focus on the cold season (from October to the following March) of the Northern Hemisphere. The manuscript is organized as follows. Datasets and analysis methods are described in section 2. A series of comparisons between observed and simulated trends and their link to Arctic sea ice loss are presented in section 3. Covariability between Arctic SIC and lagged atmospheric circulation with the effects of confounding factors addressed is discussed in section 4 to inform the capability of AGCMs in simulating observed lagged relationships at interannual time scales. We summarize and discuss the results in section 5.

2. Datasets and methods

a. Coordinated multimodel AGCM experiments

This study uses nine AGCMs, listed in Table 1, to conduct two sets of large-ensemble experiments during 1979–2014 following a protocol developed by the H2020 Blue-Action Project. The first set of experiments is forced with daily time-varying global SST, SIC, and CMIP6 radiative forcings (see the seventh column in Table 1; Eyring et al. 2016; Haarsma et al. 2016) to include the effects of all observed forcings on the atmospheric circulation. These experiments are AMIP-type simulations. We call this set of experiments ALL to reflect that it uses all observed forcings. The second set is identical to ALL except that the daily time-varying SIC field for the Northern Hemisphere is replaced by its daily climatological (1979–2014 average) values. We call this set of experiments SICclim to denote its Arctic sea ice forcing is replaced by climatological values. The atmospheric circulation in SICclim, therefore, is not directly affected by Arctic sea ice variability, and the atmospheric circulation response to Arctic sea ice variability can be estimated from the multimodel ensemble mean (MMEM) of ALL members minus the MMEM of SICclim members (S1 MMEM hereafter), assuming additivity between the SIC-driven changes and that driven by other forcings. All fields are regridded to 1.25° (longitude) × 0.94° (latitude) horizontal resolution before further analysis.

The SST and SIC boundary conditions used to force the AGCMs during the 1979–2014 period are obtained from the Met Office Hadley Centre Sea Ice and SST version 2.2.0 dataset (Kennedy et al. 2017), which was also used in the CMIP6 HighResMIP protocol (Haarsma et al. 2016). To make sure that the SST and SIC fields evolve consistently in SICclim and any unphysical SST or SIC values are removed, most modeling groups performed an adjustment following Hurrell et al. (2008), although its impacts on the atmospheric responses are quite small [see section 2.1 of Liang et al. (2020) for discussion]. Table 1 specifies whether this adjustment was implemented.

To minimize the effects of internal atmospheric variability, each modeling group has conducted ensembles of 10–30 members that differ only in their initial conditions for each model, resulting in a total of 165 members (Table 1). Because the EC-Earth3-NLeSC model only provides variables at the surface, 500-hPa, and 100-hPa levels due to their storage limitations, we do not include the EC-Earth3-NLeSC 10-member results in the
analyses at other levels or for vertical profiles. In other words, only 155 members are used in those analyses. How many members are used is specified in each figure caption. It is also noted that not every modeling group strictly followed all the protocol requirements (see Table 1). However, such departures are not expected to impact the results in any significant way [see comparison and discussion in Liang et al. (2020)]. We take MMEM over 165 members to reduce the effect of internal atmospheric variability (e.g., Deser et al. 2020) without applying any specific weight for each member of different models when taking MMEM.

b. Reanalysis datasets

In comparison to the simulated results, this study also uses reanalysis datasets. Sea level pressure, air temperature, geopotential height, and atmospheric zonal wind fields from the latest European Centre for Medium-Range Weather Forecasts reanalysis (ERA5) during 1979–2014 (Hersbach et al. 2020). To test robustness, we also analyzed four other reanalysis products: ERA-Interim (Dee et al. 2011), Japanese 55-Year Reanalysis (Kobayashi et al. 2015), Modern-Era Retrospective analysis for Research and Applications version 2 (Gelaro et al. 2017), and National Center for Atmospheric Research–Department of Energy Atmospheric Model Intercomparison Project (AMIP-II) reanalysis (Kanamitsu et al. 2002). As we find very similar results, we only present the ERA5 results, although a global mean cold bias is present in the lower stratosphere from 2000 to 2006 in the ERA5 temperature (Simmons et al. 2020). Weekly Northern Hemisphere continental snow cover extent (SCE) from 1979 to 2017 is obtained from the Rutgers University Global Snow Laboratory (Estilow et al. 2020). Weekly Northern Hemisphere continental snow cover extent (SCE) from 1979 to 2017 is obtained from the Rutgers University Global Snow Laboratory (Estilow et al. 2020) and aggregated into monthly data.

c. Trend analysis

The linear trends during 1979–2014 are calculated by a least squares fit of the linear regression $y_i = a + b t_i$, where $y_i$ is the variable of interest and $t_i$ is time. Their statistical significance is determined by a two-sided Student’s $t$ test with a null hypothesis that the slope $b$ is zero. When the corresponding $p$ value is less than 0.05, we consider that the trend is significant at the 5% significance level. The 95% confidence interval of the estimated trend $b$ in one single realization (e.g., ERA5) is given by

$$
\left( \hat{b} - \frac{t_{(1-0.05)/2}}{S} \hat{\sigma}_b, \hat{b} + \frac{t_{(1-0.05)/2}}{S} \hat{\sigma}_b \right),
$$

where $t_{(1-0.05)/2}$ is the $(1 + 0.95)/2$ quantile of the $t$ distribution with degree of freedom $n - 2$, $\hat{\sigma}_b$ is the residual standard deviation, and $S$ is the root-mean-square difference of the time $t_i$ from the time mean. A thorough derivation can be found in section 8.3.7 of von Storch and Zwiers (1999).

For simulated fields, we present the trends of MMEM, except for the histograms of the individual member trends (as explicitly noted in the figure captions). Note that the MMEM trends are likely more significant than their observational counterparts, because the influence of internal atmospheric variability is minimized in the former, while often dominating...
in the latter. Note that we consider linear trends because they are more easily comparable between datasets than quadratic trends, even though the latter may slightly better represent Arctic sea ice loss (e.g., Dirksen et al. 2017). In the spatial maps of trends, statistical significance may be overestimated if simply based on statistical tests at each grid point (hereafter “local” significance; e.g., Livezy and Chen 1983; Wilks 2016).

The “field significance” approach (Wilks 2016), which constrains the false discovery rate (FDR) of the regression estimates, is used to address this issue. We choose $\alpha_{\text{FDR}} = 0.1$ to achieve a global test level $\alpha_{\text{Global}} = 0.05$, assuming a spatial decorrelation scale of $1.54 \times 10^7 \text{km}$, as estimated by the spatial autocorrelation of the Northern Hemisphere geopotential height field at 500 hPa in Polyak (1996) [see Fig. 4 in Wilks (2016)]. In the figures presented in this study, we use black stippling to denote that the trend has field significance, and use cyan stippling (or contour line) to indicate that the trend is 5% significant locally.

d. Maximum covariance analysis

We use maximum covariance analysis (MCA) to examine the lagged atmospheric circulation response to Arctic SIC variability at interannual time scale. MCA applies singular value decomposition (Bretherton et al. 1992) on the covariance matrix of any two fields and retrieves modes of covariability characterized by corresponding spatial patterns and associated time series. Here the covariance matrix is calculated using area-weighting for both SIC and SLP fields. We apply the MCA approach of García-Serrano et al. (2017) to SIC and sea level pressure (SLP) anomaly fields and investigate the leading modes of covariability when SIC leads by 1–6 months. The spatial SIC patterns are obtained by regressions on the SIC time series normalized by its standard deviation (i.e., homogeneous covariance map), and the SLP patterns by regression on the same SIC time series (i.e., heterogeneous covariance map). Using the heterogeneous maps for SLP avoids the limitations pointed out by Zappa et al. (2021), while preserving orthogonality (Czaja and Frankignoul 2002).

For ERA5 or individual AGCM members, we obtain the monthly SIC and SLP anomalies by subtracting their climatological monthly mean for the 1979–2014 period and then remove quadratic trends in order to reduce long-term variations. Here we remove the quadratic trend because the Arctic sea ice loss in Northern Hemisphere autumn and winter in the past decades shows an accelerating rate, which is closer to a quadratic structure [see Dirkson et al. (2017) and Liang et al. (2020) for discussions], although some MCA analyses have shown that results are similar when removing instead a linear or cubic trend (García-Serrano et al. 2017). For MCA using AGCM MMEM results, the same steps are applied to the MMEM SLP to obtain MMEM SLP anomalies.

Each MCA mode is characterized by its squared covariance (SC; the corresponding squared singular value of the covariance matrix), a correlation coefficient ($R$) between the two MCA time series, and a squared covariance fraction (SCF; the percentage of explained squared covariance). SCF is intrinsically the same measure as SC, so we only discuss SC. To assess the significance of SC and $R$, we repeat MCAs with 100 random permutations in time dimension for SLP to obtain 100 realizations of SC and $R$ reflecting the effect of internal variability as in previous studies (e.g., Gastineau et al. 2017). The number of SC and $R$ values that are larger than target SC and $R$ values provides an estimate of their level of significance.

e. Indices

In section 4, the Siberian snow index used is defined as the SCE time series of the leading MCA mode between November (December) SCE anomalies over northern Eurasia ($0^\circ–180^\circ$ and $40^\circ–65^\circ N$) and December (February) SLP anomalies over the North Atlantic–Eurasian domain ($90^\circ W–40^\circ E$, $20^\circ–90^\circ N$), following Gastineau et al. (2017). The North Atlantic SST index is similarly defined as the SST time series of the leading MCA mode between SST anomalies over the North Atlantic ($90^\circ W–40^\circ E$, $10^\circ S–60^\circ N$) and subsequent SLP anomalies over the North Atlantic–Eurasian domain ($90^\circ W–40^\circ E$, $20^\circ–90^\circ N$), which characterize the North Atlantic horseshoe (tripolar) SST pattern (Czaja and Frankignoul 1999, 2002). We define the Ural blocking index with two metrics: 1) the number of days with blocking occurring in the Ural sector ($0^\circ–80^\circ E$ and $40^\circ–75^\circ N$) during the 1979–2014 period based on the blocking definition by Scherrer et al. (2006), and 2) the area-averaged Z500 anomalies over the eastern Europe and Ural sectors ($10^\circ W–80^\circ E$, $45^\circ–80^\circ N$) following Peings (2019). It is noted that the thresholds in determining Ural blocking events for metric 1 and differences in geographic domain between the two metrics may result in a discrepancy in the blocking time series based on the two metrics. In section 4, we mainly show results using metric 1, but also used metric 2 for comparison (shown later in Fig. 13). The NAO index used in section 4 is obtained from NOAA Climate Prediction Center based on the rotated EOF method (Barnston and Livezy 1987).

f. Multiple regression analysis

To quantify relative importance of potential confounding factors on wintertime atmospheric circulation using the indices above mentioned, we use multivariate least squares regressions in section 4 as in previous studies (e.g., Gastineau et al. 2017; Simon et al. 2020). We address the multicollinearity with the variance inflation factor (VIF; Kendall 1946). The VIFs in our cases are at most 1.7, much smaller than 5, which is generally considered a sign of severe multicollinearity (Judge et al. 1988). The level of statistical significance is tested with 100 random permutations of the atmospheric fields in time. The number of permutations that exceeds the observed value in the permuted samples provides the $p$ value. We do not test all lead–lag relationships across all variables but only focus on specific timing and variables with possible physical linkages, including Eurasian SCE (Gastineau et al. 2017), North Atlantic SST (Czaja and Frankignoul 2002), and Ural blocking (Peings 2019).

3. Long-term linear trends during the 1979–2014 period

a. Observed and simulated trends

As the accelerated Arctic warming is the dominant feature of the Arctic sea ice decline in the past few decades, we first
analyze the linear trend of air temperature $T$ averaged over the polar cap domain ($65^\circ-90^\circ$N) during 1979–2014. The monthly evolution of the polar-cap $T$ trends reveals warming trends in the lower troposphere in every month with the strongest trends from autumn to early spring in both ERA5 and ALL MMEM (Figs. 1a,b). However, there are some differences. In April, a strong warming trend in lower troposphere occurs in ERA5, which is not seen in ALL MMEM, and from January to April large differences occur in middle and higher troposphere and stratosphere. The $T$ trends associated with the Arctic sea ice loss are derived from SI MMEM $T$ and they show the largest warming patches in October–December (OND) and January–March (JFM) (Fig. 1c). The OND warming trends are mostly confined below the 850-hPa level, while the JFM warming trends reach the middle troposphere up to the 500-hPa level. There is also a signal in the stratosphere, albeit not field significant. These results suggest that the Arctic sea ice changes explain more than half of the simulated and observed near-surface temperature trends in the polar cap during boreal fall and winter. Thus, we focus on the OND and JFM means in the following trend analyses.

We next compare the ERA5 OND and JFM surface air temperature (SAT) trends in the Northern Hemisphere domain ($20^\circ-90^\circ$N) with the corresponding trends of ALL MMEM during 1979–2014. Consistent large warming trends in ERA5 and ALL MMEM SAT are found in the Arctic Siberian sector and BK Seas in OND (Figs. 2a,c) and in the BK, Labrador, and Okhotsk Seas in JFM (Figs. 2b,d), where large sea ice decline occurs (e.g., Onarheim et al. 2018). Furthermore, the trends of Arctic-averaged SAT from ERA5 and ALL MMEM (magenta and black vertical lines in Figs. 2e,f) exhibit similar amplitudes and are both significant at the 5% level, indicating a robust warming trend. We note that the width of 95% confidence intervals for the trends of ALL MMEM (0.035 K yr$^{-1}$ for OND and 0.034 K yr$^{-1}$ for JFM; magenta shadings in Figs. 2e,f) are smaller than those of ERA5 (0.058 K yr$^{-1}$ for OND and 0.065 K yr$^{-1}$ for JFM), but not as much as expected from the reduced contribution of internal atmospheric variability in MMEMs, showing that other uncertainty contributes to the spread, such as the model (and/or scenario) uncertainty (e.g., Lehner et al. 2020). We calculate the OND and JFM trends of Arctic-averaged SAT from each ALL member (bars in Figs. 2e,f) and obtain a Gaussian-like distribution of SAT trends, whose spread mostly overlaps with the 95% confidence intervals of the ERA5 SAT trends (gray shadings in Figs. 2e,f), except for the largest positive values. The contribution of model uncertainty can be estimated by comparing the full range of SAT trend among the 165 ALL

![FIG. 1. (a) Linear trends of ERA5 Arctic-averaged ($65^\circ-90^\circ$N) air temperature in each month. (b),(c) As in (a), but for ALL MMEM and SI MMEM, respectively. The black contour lines in (a) represent climatological (1979–2014) Arctic-averaged ERA5 air temperature values (K), and in (b) and (c) climatological (1979–2014) Arctic-averaged ALL air temperature values (K). The black dots indicate field significance, while the cyan dots local significance at 5% level. The ticks on the x axis indicate the middle of the month.](image-url)
members (0.068 K yr\(^{-1}\) for OND and 0.097 K yr\(^{-1}\) for JFM) with the range of estimated model internal variability (0.044 K yr\(^{-1}\) for OND and 0.074 K yr\(^{-1}\) for JFM), as obtained from 165-member trends after removing each corresponding model ensemble-mean trend. In the sense of explained portion of uncertainty [i.e., 1 – (internal variability/model uncertainty)\(^2\)] \times 100\%, we can estimate the fraction of model uncertainty for OND SAT to be about 58\%—that is, 1 – (0.044/0.068)\(^2\)] \times 100\%—and that for JFM SAT about 42\%, that is, 1 – (0.074/0.097)\(^2\)] \times 100\%.

Outside the Arctic, the ERA5 SAT trend patterns in the North Pacific and Atlantic sectors are mostly significant and realistically simulated by ALL MMEM, as expected from prescribed SSTs. Over the continents, ALL MMEM indicates a weak, but statistically significant warming trend, except in the northwestern United States and western Canada (Figs. 2c,d). In contrast, the trends over the continents in ERA5 are mostly not statistically significant (Figs. 2a,b), thus precluding a direct comparison between ALL MMEM and ERA5. Nonetheless, it should be noted that the warming trends in ERA5 are significant in a few continental regions, such as southern Europe, southeast Asia, and the southern United States and northern Mexico.

The same analysis is performed on the ERA5 and ALL MMEM SLP fields to examine the trends in near-surface atmospheric circulation (Figs. 3a–d). The comparison is difficult because the ERA5 SLP trends are mostly not significant, except locally for the positive trends over part of the North Pacific that are likely due to SST changes. The ALL MMEM SLP trends are not significant in JFM, but in OND SLP trends are significantly negative over much of Eurasia and the Arctic, and positive over the eastern North Pacific and northwestern North America. The OND and JFM trends of the Arctic-averaged SLP from each member of ALL present Gaussian-like distributions with the trends centering around zero, although the decreasing MMEM trend in OND is significant at the 5% level (Figs. 3e,f). Their spread (the range of vertical bars) is similar to the 95% confidence interval of the ERA5 SLP trends (the gray shading). The model uncertainties contribute little in OND and even less in JFM, as the ranges of estimated model internal variability (0.20 hPa yr\(^{-1}\) for OND...
and 0.28 hPa yr\(^{-1}\) for JFM) largely accounts for the full ranges of ALL members (0.23 hPa yr\(^{-1}\) for OND and 0.29 hPa yr\(^{-1}\) for JFM). We can estimate the fraction of model uncertainty for OND SLP to be about 24\%—that is, \([1 - (0.20/0.23)^2] \times 100\%\)—and that for JFM SLP about 7\%, that is, \([1 - (0.28/0.29)^2] \times 100\%\). The results indicate that, in single model simulations and in ERA5, the forced SLP trends within the Arctic are masked out by the internal atmospheric variability and too weak to be detected.

We also investigated the trends in the NAO, using for NAO index based on the difference between area-averaged SLP anomalies in the Azores (28°–20°W, 36°–40°N; southern red box in Figs. S1c,d in the online supplemental material) and Icelandic (25°–16°W, 63°–70°N; northern red box in Figs. S1c,d) regions after removing the time means (Smith et al. 2020). There is no significant NAO trend in ERA5 or in the MMEM, and the trend distribution shows a spread similar to that of Arctic-averaged SLP trends (Figs. S1e,f).

In the lower stratosphere, the trends of the ERA5 geopotential height at 50 hPa (Z50 hereafter) are not significant, except locally in OND above the British Isles and in JFM over central Eurasia (Figs. 4a,b). On the other hand, there is a positive trend in most regions of the Northern Hemisphere domain in the ALL MMEM in both OND and JFM, with large amplitudes over the northern North Atlantic (Figs. 4c,d).

The OND and JFM Arctic-averaged trends of each ALL member again show Gaussian-like distributions, and the spreads are largely comparable to the 95% confidence intervals of ERA5 (Figs. 4e,f). Again, the contribution of model uncertainties is small as the full ranges of ALL members (6.94 m yr\(^{-1}\) for OND and 12.97 m yr\(^{-1}\) for JFM) is only barely larger than that due to internal variability (6.14 m yr\(^{-1}\) for OND and 12.03 m yr\(^{-1}\) for JFM).

We finally examine the zonal-mean \(T\) trends considering height–latitude distributions (Fig. 5). Both ERA5 and ALL MMEM \(T\) have significant warming trends in the troposphere and cooling in the extrapolar stratosphere, which are expected features of global warming. In much of the troposphere between 20° and 60°N, the trends of ALL MMEM \(T\) are stronger than those of ERA5. The strongest lower-troposphere warming in the high latitudes is the signature of Arctic amplification (e.g., Lee et al. 2008; Alder et al. 2011; Deser et al. 2015, 2016b). It is noted that the Arctic warming trends in our simulations are closer to those of ERA5 than the warming trends presented in Cohen et al. (2020), in which the warming center is slightly displaced southward (see their Fig. 1c), perhaps because of significantly fewer ensemble members. However, the AGCM version and forcing applied also differ, which may also explain the discrepancies.

In summary, the temperature trends of ERA5 and ALL MMEM, which result from the combined influence of SIC,
SST, and radiative forcings, show overall consistent features, in particular at high latitudes in the lower troposphere (Figs. 1, 2, and 5). In contrast, the consistency of the trends for the dynamical variables between ERA5 and ALL cannot be readily assessed because the effects of internal atmospheric variability dominate in ERA5.

b. Arctic sea ice influence on trends

Assuming that the atmospheric circulation response to Arctic sea ice variability is sufficiently additive, SI MMEM should best represent the response to Arctic sea ice changes, including trends and variability at interannual and longer time scales. In this section, we quantify the influence of Arctic sea ice loss on the long-term trends.

Figures 6 and 7 show the trends of SI MMEM for SAT, SLP, geopotential height at 500-hPa level (Z_500 hereafter), and zonal wind at 850-hPa level (U_850 hereafter) in OND and JFM, respectively. Strong SAT warming trends in OND and JFM appear in regions of large Arctic sea ice edge retreat and newly open waters, including the Labrador, Greenland, BK, East Siberian–Chukchi, and Okhotsk Seas (Figs. 6a and 7a). The similarity with the trends in Figs. 2a–d suggests that ERA5 and ALL MMEM warming trends in the Arctic are mostly driven by sea ice changes. In both seasons, the warming trends due to sea ice loss extend over northern Eurasia and North America, but their magnitude is small, thus contributing little to the continental warming trends in ALL MMEM. Consistent with the strong warming trends near the sea ice edge, there are strong negative SLP trends (Figs. 6b and 7b), likely as a result of the near-surface atmospheric circulation response to local heating that was also shown in previous modeling studies (e.g., Peings and Magnusdottir 2014; Deser et al. 2010).

Away from the local SLP response to the strong SAT warming near the sea ice edges, there is no field significant SLP trend in OND, except for a small decrease in northeastern North America and northwestern Asia, and only locally significant trends in Z_500 or U_850 (Figs. 6b–d). Conversely, there is a field significant dynamical large-scale response to Arctic sea ice loss in JFM, best seen in Z_500 (Fig. 7c). The trend has a negative AO-like pattern, particularly strong in the North Atlantic sector, with significant increasing trends over the Nordic seas, Greenland, northeast Asia, and northwestern North America, and a decreasing trend over southeastern America, western Europe, and the central Pacific. Correspondingly, the U_850 trends present a north–south dipolar structure (color shadings in Fig. 7d) that tends to shift the climatological Atlantic and Pacific jet (black contours in Fig. 7d) equatorward. The lower tropospheric zonal wind also decreases over northern Siberia. Note that the signature of the negative AO-like pattern also appears in the JFM SLP trends, but is masked by...
the local warming near the regions of sea ice retreat (Fig. 7b).
To quantify the trend uncertainties in SI MMEM, we further calculate the SLP, Z500, and U850 trends averaged over North Atlantic–northern Europe regions (red boxes in Figs. 7b–d) where the changes are statistically significant. The SLP and Z500 increase at rates of $0.021 \pm 0.018 \text{ hPa yr}^{-1}$ and $0.37 \pm 0.20 \text{ m yr}^{-1}$, respectively, whereas U850 decreases at a rate of $-0.015 \pm 0.011 \text{ ms}^{-1} \text{ yr}^{-1}$, where the 95% confidence intervals are indicated. These trends are small, however, implying a change of $0.76 \pm 0.65 \text{ hPa}$, $13.3 \pm 7.2 \text{ m}$, and $0.54 \pm 0.40 \text{ m}$, respectively, during the 36-yr period. Similarly, the JFM NAO index, as defined above, only decreases at a rate of $-0.014 \pm 0.031 \text{ hPa yr}^{-1}$ ($-0.44 \pm 0.34 \text{ m yr}^{-1}$ for the Z500 NAO index). In the stratosphere, there is no significant trend in OND (Figs. 8a,c), but sea ice loss induces a small weakening of the polar vortex in JFM, which is field significant over northeastern North America, as illustrated at 50hPa (Figs. 8b,d). Hence, these trends confirm that the Arctic sea ice decline affects, albeit weakly, the midlatitude atmospheric circulation in boreal winter.

The trends of zonally averaged $T$ and zonal velocity component $U$ in height–latitude space confirm that the JFM near-surface warming trends in high latitudes extend upward into the stratosphere (Fig. 9b) and are consistent with negative $U$ trends centered around 60°N (Fig. 8d). This is likely as a result

![Fig. 5. Zonally averaged air temperature trends from 1000 to 10 hPa. The black contour lines represent the corresponding climatological (1979–2014) zonally averaged air temperature (K) from (a),(b) ERA5 and (c),(d) ALL MMEM. The black dots indicate field significance, while cyan dots 5% local significance.](image-url)
of thermal wind adjustment and the associated eddy-driven response (e.g., Chemke et al. 2019; England et al. 2020; Hell et al. 2020). Altogether, these features indicate a negative tropospheric (Fig. 7c) and stratospheric (Fig. 8b) AO pattern. In lower latitudes, positive $U$ trends occur around 30°N in JFM (Fig. 9d), which corresponds to an intensification of the subtropical jet core. This may be caused by changes in the atmospheric meridional overturning circulation induced by the Arctic warming and subsequent modulations on $T$ trends in the lower latitudes (e.g., Deser et al. 2016b; Chemke et al. 2019; England et al. 2020; Hell et al. 2020; He et al. 2020). In contrast, in OND the warming $T$ trends are confined to the high latitudes below the 800-hPa level (Fig. 9a), while there is a small cooling trend in the lower latitudes, as well as in the high-latitude stratosphere. However, these trends are weak and lead to very little change in the zonal wind (Fig. 9c).

In summary, the SI MMEM results indicate that Arctic sea ice loss robustly contributed to the Arctic warming over the 1979–2014 period, extending to the 500-hPa level and the stratosphere during JFM. In this season, there is also a weak trend resembling a weak negative AO-like pattern, with a small southward shift of the midlatitude jet stream and a slowdown of the stratospheric circulation. In OND, however, there is no significant dynamical response to the sea ice loss.

![Fig. 6. OND trends from SI MMEM for (a) SAT, (b) SLP, (c) geopotential height at 500-hPa level (Z500), and (d) zonal wind at 850-hPa level (U850).](image-url)
These responses are not retrieved in ALL MMEM (Figs. 2 and 3), indicating that they are masked out by other forcing agents (e.g., SST).

4. Interannual variability during 1979–2014

In this section, we use MCA to examine the relationships between Arctic sea ice and lagged atmospheric circulation fluctuations with a focus on interannual (year-to-year) variability. Following observational evidence (e.g., García-Serrano et al. 2015, 2017), we primarily consider the North Atlantic–Eurasian domain (30°W–120°E, 50°–90°N for SIC; 90°W–40°E, 20°–90°N for SLP). Note that using the Northern Hemisphere domain (20°–90°N) for SLP in MCA gives similar results with less statistical significance as expected from the inclusion of unrelated remote signals in the SLP fields (Figs. S2–S4). SIC fluctuations are largely driven by the atmosphere on the monthly time scale (e.g., Frankignoul et al. 2014; Blackport et al. 2019) and, as for the case of SST anomalies (e.g., Czaja and Frankignoul 2002; Wills and Thompson 2018), simultaneous correlations generally reflect the atmospheric forcing of the SIC anomalies. To exclude the atmospheric forcing, we only examine the MCA when the atmospheric anomalies lag the SIC anomalies. In view of the short time scale of the natural atmospheric variability at sea level and in the midtroposphere (negligible month-to-month persistence), lags of a month and longer could better separate a possible atmospheric response to changes in the surface conditions including SIC, SST, and SCE from the direct atmospheric forcing of these changes, as in numerous studies of SIC (or SST) influence on the atmospheric variability.
circulation (e.g., Frankignoul et al. 2014; King et al. 2016; Nakamura et al. 2016; Yang et al. 2016; Koenigk et al. 2016; García-Serrano et al. 2015, 2017). To further isolate the impact of SIC from that of other surface conditions, an apparent SIC impact, which might actually reflect the influence of concomitant SST, SCE or stratospheric fluctuations, should be separately quantified. Hence, possible confounding factors are also examined.

Figures 10a and 10b show SC and $R$ of the first MCA mode for ERA5 at varying lagged months from October to March. The most significant MCA mode is for November SIC and December SLP, with significant SC at 1% level and $R$ at 8% level. The November SIC spatial pattern presents a large SIC reduction in the BK Seas (Fig. 10c), and the SLP anomalies one month later have a negative NAO-like pattern in the North Atlantic sector (Fig. 10d). This result is consistent with the analyses by García-Serrano et al. (2017). SC is also 2% significant when the October SIC anomalies lead the January SLP anomalies, albeit with slightly less significant $R$ but similar patterns. Performing the MCA with Z500 anomalies instead of SLP anomalies gives very similar results and indicates that the December atmospheric circulation anomalies are largely barotropic in the North Atlantic sector. The signal extends to the stratosphere where it resembles a negative AO (highly significant at 50 hPa; not shown). However, Gastineau et al. (2017) found that a concomitant Siberian SCE increase in November (likely synchronously driven by the same atmospheric fluctuations that drove the SIC anomalies) also precedes a similar negative NAO-like signal. The link with SCE is twice stronger than with SIC, while November Ural blocking may have
negligible impact on these lag relations (their Fig. 13). Furthermore, Czaja and Frankignoul (2002) showed that a North Atlantic horseshoe pattern is also driving the NAO in early winter [November–January (NDJ)]. Hence, our MCA between November SIC and December SLP may be influenced by these confounding factors. Regressing the November global Z500, SST and SCE anomalies onto the November SIC MCA time series illustrate concomitant signals and shows that an eastern European Z500 increase, Siberian SCE increase, and a significant cooling in the subtropical North Atlantic may have indeed contributed to the relation between November SIC and December SLP (Figs. 11a,b). To estimate relative contributions, a multiple regression was performed using four regressors: the November SIC MCA time series, the November Siberian SCE index, the November North Atlantic SST index, and, for completeness, the November Ural blocking index. Collinearity is very limited, with a VIF of 1.3 at most. The results (not shown) indicate that although SIC still leads a significant negative NAO-like signal, its amplitude is about 3 times smaller than in Fig. 10d, which is dominated by the combined influence of Siberian SCE and North Atlantic SST.

FIG. 9. The trends of SI MMEM (155 members) zonally averaged air temperature in (a) OND and (b) JFM. (c),(d) As in (a) and (b), but for zonally averaged zonal winds. The black contour lines represent the corresponding climatological mean (1979–2014). The black dots indicate field significance, while the cyan dots 5% local significance.
Figure 10 also shows that there is a strong 2% significant SC when the December SIC anomalies lead the February SLP anomalies, although \( R \) is only 26% significant. This MCA mode shows that a SIC decrease in the Barents and Greenland Seas precedes the occurrence of a negative NAO-like pattern (Fig. 10e). The mode is also equivalent barotropic in the troposphere (but insignificant in the stratosphere), with similar statistical significance when based on Z500 anomalies. This mode was discussed by García-Serrano and Frankignoul (2016), who mentioned that the NAO-like response may also be affected by a concomitant North Atlantic SST tripolar anomaly, which is seen in the SST regression map (Fig. 11d). Figure 11d also shows that the December SIC decrease is accompanied by a SCE increase in central Asia and decrease in central Europe, which were found by Gastineau et al. (2017) to have a weakly significant link with the February atmospheric circulation. A concomitant statistically significant Z500 signal is also found in the Ural blocking region (Fig. 11c), suggesting that December
Ural blocking may be another confounding factor in analogy to the results that Peings (2019) suggested for November Ural blocking. To investigate the impacts of these possible confounding factors, we have performed a multiple regression with four predictors, the December SIC MCA PC1, the December SCE index, the North Atlantic SST index, and the December Ural blocking index. The collinearity is again limited (VIF of 1.7 at most). Although each regressor precedes a significant February SLP signal in the North Atlantic–Europe sector with substantial amplitude, the patterns differ (Fig. 12). Indeed, the SLP signal linked to December SIC most resembles a negative NAO, while that linked to the North Atlantic SST tripolar anomalies has similarity with a positive NAO with a southward shift of anomaly centers, and with associated anticyclonic anomalies over Scandinavia. Such a pattern is consistent with Han et al. (2016), who showed opposite SST and SIC impacts on the atmosphere in late winter. The SLP signal linked to December Ural blocking occurs west of the British Isles and does resemble the NAO. To quantify the influence of these factors on the NAO, we have performed the multivariate regression using the February NAO index as predictand. As shown by the coefficient of determinations ($R^2$) in Fig. 13a, December SIC, SCE, and SST contribute comparatively (e.g., $R^2 = 0.23$ for SIC and 0.15 for SCE) to the NAO variability, while December Ural blocking has a weaker influence. It is noted that the explained NAO variance is substantially increased when the four (or the first three) indices are used together as predictors. Similar results are found when using Peings’s (2019) area-averaged Z500 blocking index (Fig. 13b). Since Peings (2019) suggested that November Ural blocking influences the NAO, we also performed the multiple regression using the November Ural blocking indices instead of the December ones, but the influence on the NAO remained small (not shown).

In summary, our MCA results between the December SIC and February SLP is influenced by the concomitant SCE, SST, and, to a much lesser extent, Ural blocking. Quantitatively, the February signal that can be attributed to SIC change in our multiple regression model is about 2.5 hPa over the Icelandic region (area-averaged values over the green box in Fig. 12a. This box is chosen to cover the northern center of action for NAO), which accounts for 50% of that in the MCA results (about 5 hPa over the same green box).

We next perform the same MCA on ALL MMEM SLP fields to assess if the MCA modes found in ERA5 are reproduced in AGCM simulations with all forcings (Fig. 14). Even though MMEM should be less affected by internal variability than ERA5, there is no significant SC at 10% significance level. The highest statistical significance for the SC (20%) is reached in three cases for December SIC leading SLP, but in all the cases a SIC decline in the BK Seas leads a positive NAO-like SLP in March, opposite to the ERA5 results (perhaps primarily...
reflecting the influence of the concomitant SST tripolar anomalies; see Fig. S5).

More significant MCA results are found using SI MMEM (Fig. 15), very likely because the influence of interannual SST and radiative forcing has been sufficiently reduced in SI MMEM. The November SIC–December SLP is not statistically significant, perhaps because of the observed correlation between the BK SIC and Eurasian SCE is not reproduced in AMIP-type simulations. On the other hand, the December SIC–February SLP MCA mode has high statistical significance with 3% for SC and 0% for R (Figs. 15a,b). Its SIC and SLP spatial patterns (Figs. 15c,d) largely resemble those of ERA5 (Figs. 10e,f and 12) with a large sea ice decline in the BK Seas, albeit not in the Greenland Sea, and a negative NAO-like SLP pattern. Similar negative NAO-like dipolar patterns, albeit somewhat tilted, are also found in other MCA pairs, which include the Greenland Sea SIC decline [e.g., January SIC–February SLP (shown in Figs. 15e–f) and February SIC–March SLP (not shown), where either SC or R is 10% significant]. These SI MMEM MCA modes thus resemble the December–February mode in ERA5, but their SCs are one order of magnitude smaller. The amplitude of the regressed SLP patterns above the Icelandic region (e.g., area-averaged value over the green box in Fig. 15d) is about 0.46 hPa, less than 20% of the multivariate ERA5 estimate (2.5 hPa). To investigate if the smaller SLP values of AGCMs are sensitive to selected region, we also consider the values averaged over a larger spatial extent and obtain similar model underestimations: the SI MMEM SLP value averaged over the northern domain (60°–90°N, 90°W–0°) is about 0.47 hPa, smaller than the multivariate ERA5 estimate that is 1.37 hPa; the SI MMEM SLP value averaged over the southern domain (20°–50°N, 90°W–0°) is about −0.05 hPa, smaller than the multivariate ERA5 estimate that is −0.37 hPa. These results suggest that the AGCMs may underestimate the strength of SLP–SIC covariability.

5. Summary and discussion

This study uses two sets of coordinated AGCM experiments, one with observed Arctic SIC prescribed during 1979–2014 and the other with climatological Arctic SIC, to examine the impacts of Arctic sea ice on atmospheric circulation in the Northern Hemisphere cold seasons (i.e., October to March). We first compared the linear trends in the ALL MMEM (i.e., experiments with all the time-varying forcings, including SST, SIC, greenhouse gases, aerosols, and solar radiation) to those in ERA5. The warming SAT trends, which are large in regions of Arctic sea ice loss, are consistent within the Arctic Circle in terms of geographical pattern and magnitude between ALL MMEM and ERA5. The warming SAT trends, which are large in regions of Arctic sea ice loss, are consistent within the Arctic Circle in terms of geographical pattern and magnitude between ALL MMEM and ERA5. The Arctic warming trends in the lower troposphere are also similar, suggesting that the AGCM simulations are realistic in the Arctic and the observed trends are

![Figure 12. Multiple regression of the ERA5 SLP anomalies in February onto the normalized (a) December SIC MCA time series, (b) December Siberian snow cover index, (c) December North Atlantic SST index, and (d) December Ural blocking index. The black and cyan stippling indicates 5% local significance level and field significance, respectively. The green box in (a) delineates the region of 60°–70°N and 30°W–0°.](image-url)
(b) As in (a), but for December Ural blocking following the definition of Peings (2019).

The trends driven by the Arctic sea ice variability are singled out in SI MMEM, quantified by the difference between the MMEMs of ALL and SIC\textsubscript{clim} (corresponding experiments where the time-varying SIC is replaced by its mean seasonal cycle). The SI MMEM shows that in OND and JFM the SAT warming trends found in ALL MMEM and ERA5 in regions of sea ice decline are indeed mainly driven by sea ice loss, which account for more than 50% of the warming trends in ALL MMEM and ERA5. In both seasons, the warming trends due to sea ice loss extend over northern Eurasia and North America in SI MMEM, but their magnitude is small, even compared to ALL.

A strong local decrease dominates the SLP trends near the sea ice edge, as expected from a thermodynamical response to oceanic heat release, although it is somewhat masked in JFM by a significant dynamical barotropic response that resembles a negative AO-like circulation pattern and extends up to the stratosphere, weakening the circulation there. Correspondingly, sea ice loss induces a southward shift of the tropospheric jet stream. These JFM circulation changes may be related to the thermal wind adjustment to reduced meridional temperature gradient, the associated eddy feedback and/or troposphere–stratosphere coupling processes. These results support previous modeling studies that found a negative AO-like response in winter to an Arctic sea ice reduction (e.g., Seierstad and Bader 2009; Peings and Magnusdottir 2014). Our trend analysis is also qualitatively consistent with the observational study of Simon et al. (2020), which found that an Arctic sea ice loss should drive a negative NAO-like signal in winter, but has a negligible large-scale impact in autumn. However, in our analysis, the winter trends in SI MMEM are substantially smaller than those in Simon et al. (2020). For example, the winter SLP over a North Atlantic–northern Europe region increases at a rate of 0.21 ± 0.18 hPa decade\textsuperscript{−1} and Z500 increases at a rate of 3.7 ± 2.0 m decade\textsuperscript{−1}, while in the observational estimates the Icelandic high could increase by as much as 5 ± 4 hPa decade\textsuperscript{−1} for SLP and 60 ± 40 m decade\textsuperscript{−1} for Z500.
Although these observational estimates are upper bounds as they are based on perpetual winter conditions and assumption of linearity, they are much larger than in the MMEM. It is noted that these trend estimates are much larger than observed (Fig. 3; see also Blackport and Screen 2020b), but observed trends should also be influenced by SST changes and external forcings.

This discrepancy led us to investigate whether the AGCMs can represent the impact of Arctic sea ice variability on the cold season atmospheric circulation at interannual time scale. Such influence in observations has been extensively documented in the literature (e.g., Frankignoul et al. 2014; King et al. 2016; Nakamura et al. 2016; Yang et al. 2016; Koenigk et al. 2016; García-Serrano et al. 2015, 2017; and many others) and was more robustly characterized in Simon et al. (2020). For this purpose, MCA is used between detrended SIC anomalies and lagged SLP anomalies in the North Atlantic–Europe sector, and it is compared with ERA5 counterparts during the same 1979–2014 period. As in the studies above mentioned, we considered relation between SIC and the atmosphere lagging by at least one month, which should suffice to separate a possible atmospheric response to SIC changes from the direct atmospheric forcing of these changes (e.g., Frankignoul et al. 2014; Gastineau and Frankignoul 2015; Blackport et al. 2019; Screen and Blackport 2019; Blackport and Screen 2020b), in view of the short memory time scale of the internal atmospheric variability in the middle and lower troposphere. Possible confounding influence were also examined using multiple regression since, for example, Gastineau et al. (2017) showed that concomitant Siberian SCE anomalies had a similar but larger influence on the observed early winter atmospheric circulation compared to the response to the BK Seas SIC. Simon et al. (2020) found similar results for the atmospheric response in December to interannual SIC variability, but a SIC influence...
the NAO later in winter was not significantly affected by concomitant SST or SCE variability. Also, Peings (2019) suggested that an apparent SIC impact on the winter atmospheric circulation might arise from prior Ural blocking variability influencing both the BK SIC and the stratosphere, which in turn affected the NAO throughout the winter (see also Luo et al. 2016, 2018).

The most significant mode of lag covariability in ERA5 is between November SIC and the following December SLP, which is manifested as sea ice decline over the BK Seas preceding a negative NAO-like pattern. However, the negative NAO signal was found to be dominated by the influence of confounding factors, namely concomitant Siberian SCE anomalies, consistent with Gastineau et al. (2017), and North Atlantic SST anomalies, as in Czaja and Frankignoul (1999, 2002), which were presumably in part driven by the same atmospheric fluctuations as the BK SIC anomalies. On the other hand, November Ural blocking had negligible influence. In contrast to ERA5, no significant corresponding MCA mode was found between November SIC and the following December SLP in either ALL MMEM or SI MMEM, perhaps in part because the interannual snow cover changes in the AGCMs are essentially independent of the prescribed SIC variations.

Consistent with García-Serrano and Frankignoul (2016) and many others, another significant MCA mode was found in ERA5, showing that a SIC decrease in the Barents and Greenland Seas in December precedes the occurrence of a negative NAO-like pattern in February. The mode is equivalent barotropic but without significant signature in the stratosphere. The possible influence of concomitant SCE and SST anomalies was also investigated, along with that of December BK SIC, North Atlantic SST, and Siberian SCE anomalies significantly contributed to SLP fields in February, with the pattern...
linked to SIC most resembling the negative NAO, while that linked to SST had the opposite sign, consistent with Han et al. (2016). In contrast, the contribution from December (or November) Ural blocking to the NAO was smaller. Based on multiple regression, December SIC, Siberian snow cover and North Atlantic SST contributed substantially to the February NAO, and the signal that could be attributed to BK SIC in ERA5 reached an amplitude of about 2.5 hPa near the Icelandic region, about 50% of that suggested by the lagged MCA between SIC and SLP.

In ALL MMEM, there was no 10% significant mode between SIC and SLP in the same North Atlantic–Europe sector, and the only MCA mode that reaches the 20% significance level (December BK SIC leading March SLP) had an opposite polarity in SLP to the ERA5 one, possibly reflecting interannual North Atlantic SST tripolar forcing. On the other hand, a significant MCA mode emerges in late winter when using SI MMEM, with consistent SIC and SLP spatial patterns with ERA5. The most significant MCA mode is a SIC decrease in the BK Seas in December preceding a negative NAO-like signal in February, but there were also similar modes, albeit less significant, showing that a SIC decrease in the BK and Greenland Seas during January precedes a negative NAO-like SLP signal in February or March. However, the SLP amplitude in SI MMEM only reaches 0.5 hPa near the Icelandic region (or northern part of North Atlantic), which is only about 20% of our ERA5 estimate based on multiple regression model.

In summary, a negative NAO-like response to Arctic sea ice loss is found in SI MMEM during mid-to-late winter when considering both the trends and the interannual variability. This is broadly consistent with observation-based estimates, but the response to interannual SIC fluctuations is much weaker in AGCMs than in the reanalysis data, suggesting that the AGCMs might underestimate the response. Different AGCM background states, insufficient physics and parameterizations of turbulence, missing effects of leads on the ocean–atmosphere heat exchange (Marcq and Weiss 2012; Davy 2018), cloud microphysics (Screen et al. 2018), or sea ice configuration such as constant sea ice thickness in AGCMs (Lang et al. 2017), could be responsible for this underestimation. We do not find an indication that the spatial resolution could be a factor, since EC-Earth3-NLeSC and HadGEM3-GC3.1, with nominal 40- or 60-km horizontal resolutions that are higher than the other AGCMs, do not seem to behave differently from the other AGCMs, but other high-resolution models should be considered before ruling out the influence of spatial resolution.

The lack of active atmosphere–ocean coupling in AGCMs has been argued to play some roles in the weakness of the Arctic–midlatitude linkages at longer time scales (e.g., Deser et al. 2015, 2016a). Although García-Serrano et al. (2017) obtained lagged MCA results in coupled simulations rather comparable to observed ones, albeit with a smaller response amplitude, a growing body of studies with CGCMs show larger atmospheric circulation variation associated with Arctic sea ice changes (e.g., Blackport et al. 2019; Blackport and Screen 2020a,b, 2021; Cohen et al. 2020; Dai and Song 2020). More efforts are needed to investigate the role of atmosphere–ocean coupling, which could be done by comparing uncoupled and coupled simulations from the Polar Amplification Model Intercomparison Project (Smith et al. 2019) or multimodel large ensemble datasets (Deser et al. 2020). We also note that the weak response hinted at here could resonate with the so-called signal-to-noise paradox, in which the state-of-the-art climate models generally have much lower predicted variability (maybe relating to forced variability this study focuses on) than the observed one (Scaife and Smith 2018).

Finally, our suggestion that AGCMs underestimate the atmospheric response to Arctic sea ice changes must be viewed with caution since it relies on the analysis of 35 years of satellite-based SIC observations and their relation to the atmosphere, which is a relatively short period. Hence, it cannot be excluded that the results are intervened by the large internal atmospheric variability (e.g., Blackport and Screen 2021). Also, they may not be representative of earlier periods, since Kolstad and Screen (2019) suggested that the recent BK SIC–NAO relationship in observations is particularly high over the course of twentieth century. Conducting GCM experiments with SIC, albeit less accurate, and other forcings from earlier periods (thus with different background states and potential nonstationarity of the confounding factors) or applying causal effect network (e.g., Rehder et al. 2020) to observations with longer record period could help assess if AGCMs indeed underestimate the atmospheric response to Arctic SIC changes.

Acknowledgments. We acknowledge support by the Blue-Action Project (the European Union’s Horizon 2020 research and innovation programme, #727852, http://www.blue-action.eu/index.php?id=3498). The WHOI–NCAR group was supported by the U.S. National Science Foundation (NSF) Office of Polar Programs Grants 1736738 and 1737377. Their computing and data storage resources, including the Cheyenne supercomputer (doi:10.5065/D6RX99HX), were provided by the Computational and Information Systems Laboratory at NCAR. NCAR is a major facility sponsored by the U.S. NSF under Cooperative Agreement No. 1852977. Guillaume Gastineau was granted access to the HPC facility sponsored by the U.S. NSF under Cooperative Agreement No. 1852977. Guillaume Gastineau was granted access to the HPC facility sponsored by the U.S. NSF under Cooperative Agreement No. 1852977. Guillaume Gastineau was granted access to the HPC facility sponsored by the U.S. NSF under Cooperative Agreement No. 1852977.

Data availability statement. The monthly atmospheric fields were downloaded from the European Centre for Medium-Range Weather Forecasts website (https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5). AGCM simulation results will be available under Blue-Action Project data sharing platform and can be obtained upon request to the corresponding author. The SCE data are obtained from Rutgers University Global Snow Lab (https://climate.rutgers.edu/snowcover/docs.php?target=datareq). The NAO index is...

REFERENCES


Estilow, T. W., A. H. Young, and D. A. Robinson, 2015: A long-term Northern Hemisphere snow cover extent data record for


Seland, Ø., and Coauthors, 2020: The Norwegian Earth System Model, NorESM2—Evaluation of the CMIP6 DECK and Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and


