Eurasian snow cover variability and links to winter climate in the CMIP5 models

Jason C. Furtado · Judah L. Cohen · Amy H. Butler · Emily E. Riddle · Arun Kumar

Abstract Observational studies and modeling experiments illustrate that variability in October Eurasian snow cover extent impacts boreal wintertime conditions over the Northern Hemisphere (NH) through a dynamical pathway involving the stratosphere and changes in the surface-based Arctic Oscillation (AO). In this paper, we conduct a comprehensive study of the Eurasian snow–AO relationship in twenty coupled climate models run under pre-industrial conditions from the Coupled Model Intercomparison Project Phase 5 (CMIP5). Our analyses indicate that the coupled climate models, individually and collectively, do not capture well the observed snow–AO relationship. The models lack a robust lagged response between October Eurasian snow cover and several NH wintertime variables (e.g., vertically propagating waves and geopotential heights). Additionally, the CMIP5 models do not simulate the observed spatial distribution and statistics of boreal fall snow cover across the NH including Eurasia. However, when analyzing individual 40-year time slices of the models, there are periods of time in select models when the observed snow–AO relationship emerges. This finding suggests that internal variability may play a significant role in the observed relationship. Further analysis demonstrates that the models poorly capture the downward propagation of stratospheric anomalies into the troposphere, a key facet of NH wintertime climate variability irrespective of the influence of Eurasian snow cover. A weak downward propagation signal may be related to several factors including too few stratospheric vortex disruptions and weaker-than-observed tropospheric wave driving. The analyses presented can be used as a roadmap for model evaluations in future studies involving NH wintertime climate variability, including those considering future climate change.

Keywords Arctic Oscillation · Large-scale extratropical climate variability · Stratosphere–troposphere coupling · Eurasian snow cover

J. C. Furtado (✉) · J. L. Cohen
Atmospheric and Environmental Research, Inc.,
131 Hartwell Ave., Lexington, MA 02421, USA
e-mail: jfurtado@aer.com

A. H. Butler
Earth Systems Research Laboratory, Chemical Sciences Division,
Cooperative Institute for Research in Environmental Sciences,
University of Colorado, Boulder, CO, USA

E. E. Riddle · A. Kumar
Climate Prediction Center, NOAA, College Park, MD, USA

E. E. Riddle
INNOVIM, LLC, Greenbelt, MD, USA

1 Introduction

The Arctic Oscillation (AO) (also referred to as the Northern Annular Mode) is the leading mode of Northern Hemisphere (NH) wintertime climate variability and describes meridional shifts in the polar jet stream and changes in storm tracks and temperatures across the NH middle and high latitudes (Thompson and Wallace 1998, 2000). Because the AO typifies wintertime weather regimes, predicting its phase is an active goal for seasonal forecasts (e.g., Baldwin et al. 2003; Cohen 2003; Orsolini and Kindem 2011a; Maidens et al. 2013). Moreover, detection and prediction of decadal- and longer-scale trends in the AO resulting from, for example, radiative warming is important for future climate projections (e.g., Gillett and Fyfe 2013).

The existence of the AO is primarily the result of internal tropospheric dynamics, arising from large-scale tropospheric eddy-zonal mean feedbacks (e.g. Lorenz and Hartmann 2003; Gerber and Vallis 2007). But, variability in the AO can be modulated from forcings and interactions with other large-scale climate modes. Several hypotheses on teleconnections and dynamical mechanisms influencing the AO exist, including: (1) Indo-Pacific SST variability (e.g., Hoerling et al. 2001; Schneider et al. 2003; Hurrell et al. 2004), particularly the El Niño-Southern Oscillation (ENSO; e.g., Garfinkel and Hartmann 2007; Ineson and Scaife 2009); (2) the quasi-biennial oscillation (e.g., Baldwin and Dunkerton 1999); (3) solar variability (e.g., Ineson et al. 2011); and (4) influences of autumn Eurasian snow cover variability on the extratropical atmospheric circulation in the fall and following winter (e.g., Foster et al. 1983; Cohen and Entekabi 1999). The focus of this study is on the last hypothesis.

The dynamical mechanism linking October Eurasian snow cover to the following wintertime climate is schematically presented in Fig. 1 (Cohen et al. 2007). Step 1 involves
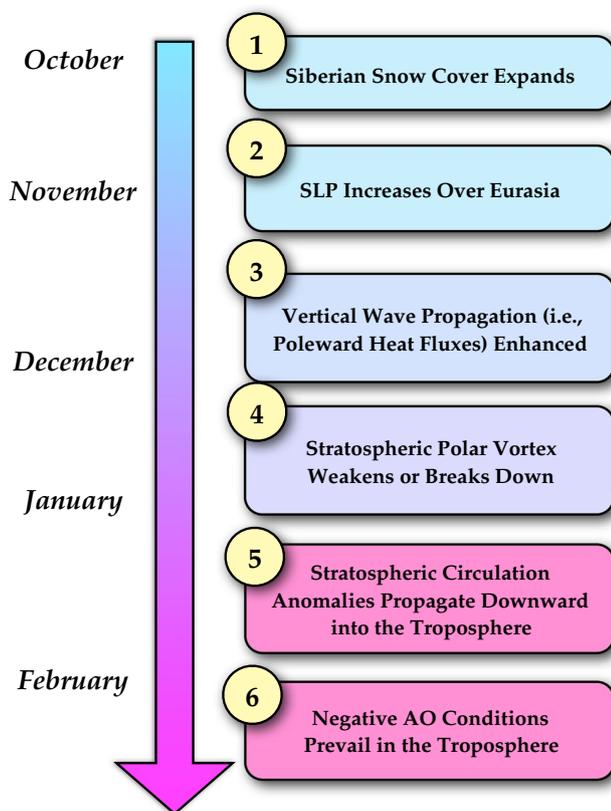
the expansion of the Eurasian snow cover (primarily in Siberia during most of October and expanding westward and southward from late October into November). Anomalously high Eurasian snow cover cools the surface and increases surface pressure (Step 2). The lower tropospheric anomalies amplify the downstream standing wave pattern and enhance vertical wave propagation into the polar stratosphere (Step 3). Anomalous momentum and heat fluxes from these vertically propagating waves are deposited in the stratosphere once the waves "break", disturbing the mean stratospheric circulation (Step 4). The imposed stratospheric anomalies eventually descend into the troposphere from continued wave-breaking at consecutively lower levels (e.g., Haynes et al. 1991), introducing same-signed anomalies there (Step 5). The tropospheric manifestation of these downward propagating anomalies is the negative phase of the AO; i.e., an equatorward shift of the polar jet and colder weather in the middle latitudes (Step 6).

Though the Eurasian snow–AO relationship was initially shown statistically, modeling experiments forced with Eurasian snow cover anomalies also reproduce this connection (e.g., Gong et al. 2003; Fletcher et al. 2007, 2009; Allen and Zender 2010, 2011; Peings et al. 2012). This relationship can therefore offer improvement in winter seasonal forecasts (e.g., Fletcher et al. 2007; Cohen and Fletcher 2007). Modeling experiment success, however, arises only when prescribing observed snow cover variability; i.e., models cannot recover the Eurasian snow–AO relationship with internally-generated snow cover (e.g., Hardiman et al. 2008; Allen and Zender 2011; Riddle et al. 2013).

While the dynamical mechanism depicted in Fig. 1 evolves on seasonal timescales, long-term trends in autumn temperature and high-latitude snowfall could amplify or reduce the impacts of autumn climate variability on wintertime climate. Coupled-climate models have long projected a positive trend in the AO (i.e., strong warming in the NH mid-latitudes and a poleward contraction of the polar jet stream) due to increases in greenhouse gases and warming of the tropical oceans (e.g., Shindell et al. 1999; Hurrell et al. 2004; Scaife et al. 2012), though that consensus has diminished in the latest set of coupled climate models (e.g., Gillett and Fyfe 2013; Cattiaux and Cassou 2013). Yet, several studies suggest that less Arctic sea ice in the fall may increase latent heat fluxes at high latitudes and thus generate higher snowfall across portions of Eurasia (e.g., Deser et al. 2010; Ghatak et al. 2010; Orsolini et al. 2011; Cohen et al. 2012). With potentially higher autumn snowfall in Eurasia, the mechanism outlined in Fig. 1 could mitigate a projected positive trend in the AO and associated NH climate changes, at least on shorter time scales.

This paper diagnoses the presence of the Eurasian snow–AO mechanism in the latest set of coupled climate models in the Coupled Model Intercomparison Project (CMIP)

Fig. 1 The six-step process describing how Eurasian snow cover in the fall (October) can impact the NH tropospheric circulation in the winter (December–February) via changes in the stratospheric circulation and changes in the surface AO pattern. Diagram adapted from Fig. 6 of Cohen et al. (2007)

Phase 5 (CMIP5). We will demonstrate that these models are unable to reproduce the snow–AO dynamical links, echoing similar results from Hardiman et al. (2008) for the CMIP3 models. Like the Hardiman et al. (2008) study, we will illustrate that the CMIP5 models underestimate the variability of October Eurasian snow cover and do not simulate lagged atmospheric responses during the winter to October Eurasian snow cover variability. However, we will also illustrate that the relationship is present in select models during select periods of time, suggesting that internal variability may play a dominant role in the observed snow–AO relationship. The paper is organized as follows. Section 2 describes the observational data and model output analyzed and compared in this study. Relationships derived from the six-step framework in Fig. 1 are then explored in Sect. 3. Next we offer some hypotheses on why the models remain unable to reproduce the Eurasian snow–AO relationship through examination of stratosphere–troposphere coupling dynamics in Sect. 4. A discussion of the findings and conclusions follow.

2 Data and methods

2.1 Observations

The observational dataset used for comparison with the model output analyses is the European Centre for Medium-Range Weather Forecasts Interim Reanalysis (ERA-Interim; Dee et al. 2011). ERA-Interim data are spaced on a 1.5° by 1.5° longitude/latitude grid globally, with 23 vertical pressure levels ranging non-uniformly from 1 to 1,000 hPa. However, for comparison with the coupled climate models, we restrict analyses to below 10 hPa. Monthly-mean values from 1979 to 2013 of several atmospheric variables [geopotential height, zonal and meridional winds, air temperature, and sea level pressure (SLP)] are studied with focus on the extended NH cold season (October through March). Additionally, to test the robustness of our conclusions, we also performed comparisons with the National Centers of Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) Reanalysis project (Kistler et al. 2001) and the National Aeronautics and Space Administration Modern Era Retrospective Analysis for Research and Applications (NASA MERRA; Rienecker et al. 2011) and received qualitatively similar results (not shown).

For snow cover, we use observed October snow cover extent (SCE) from the Rutgers Global Snow Lab (Robinson et al. 1993) from 1979 to 2012 (available for download at http://climate.rutgers.edu/snowcover). The SCE data are a merged product from reanalysis of observational snow cover maps and remote sensing measurements. Since 1999, the Rutgers snow cover data rely on the daily Interactive Multisensor Snow and Ice Mapping System (IMS) high-resolution (24 km) product. All SCE data are on a 89×89 grid. The data source also provides a Eurasian-sector snow cover index, in km^2, which is used as the basis for several regression and correlation analyses presented in this study. Brown and Robinson (2011) suggest uncertainties in the measurements for the Rutgers SCE are up to 5 % over the Eurasian sector, but including these uncertainties do not impact our results or conclusions.

2.2 Coupled climate model output

Models examined in this study originate from the CMIP5 multi-model archive, which are available for download from the Program for Climate Model Diagnosis and Intercomparison (PCMDI) at the Lawrence Livermore National Laboratory (more information on the program is available at http://cmip-pcmdi.llnl.gov/index.html). We select the pre-industrial control (piControl) scenario (i.e., prescribed, non-evolving greenhouse gas concentrations and aerosols mimicking conditions prior to 1850 are the primary forcings) for analysis in this work. The advantages of choosing this scenario over others available in CMIP5 are its long integrations (hundreds of years of model output) and its exclusion of anthropogenic effects which could influence the studied relationship. That said, we also repeated analyses for the historical scenario of the models (i.e., a continuation of the piControl run from 1850 onwards with anthropogenic and aerosol forcing representative of the late 19th/20th century included). Analyses with this scenario test whether the observed relationship relies on modern-day forcings. Results with the historical scenario of the model runs were quite similar to those from the piControl runs, indicating that the effects of anthropogenic greenhouse gas or aerosol forcing in the 20th century do not impact the simulated Eurasian snow–AO relationship versus observations. Particular models for the study are chosen based on the availability of monthly-mean output for the piControl scenario of all required variables, including SCE. Twenty models meet this criterion and are listed in Table 1. All model output are re-gridded to a common 2.5° by 2.5° longitude/latitude grid to facilitate inter-model comparison and comparison with the observations.

Model atmospheric variables analyzed are identical to those from observations. Snow cover from the model output is provided as the monthly-mean fractional area (0–1) of a grid cell covered in snow. To get an areal extent of snow cover, the fractional area is multiplied by the size of the grid box and then summed over the area from 0° to 170°W, 20°N to 75°N, which coincides roughly with the same area used for the Rutgers Global Snow Lab Eurasian SCE index (T. Estilow, personal communication).

Table 1 List of coupled climate models from the CMIP5 model archive analyzed in this study, along with total length of the piControl run for each model (years)

| Institution, Country	Model name	Length of piControl run (years)
Beijing Climate Center (BCC), China	BCC-CSM1.1	500
Beijing Normal University (BNU), China	BNU-ESM	559
Canadian Centre for Climate Modeling and Analysis (CCCma), Canada	CanESM2	996
National Center for Atmospheric Research (NCAR), United States	CCSM4	501
National Center for Atmospheric Research (NCAR), United States	CESM1-WACCM	200
Centre National de Recherches Météorologiques/Centre Européen de Recherche et Formation Avancées en Calcul Scientifique (CNRM-CERFACS), France	CNRM-CM5	850
Commonwealth Scientific and Industrial Research Organisation (CSIRO) in collaboration with the Queensland Climate Change Centre of Excellence (QCCCE), Australia	CSIRO-Mk3-6-0	500
Institute of Atmospheric Physics (IAP) and Tsinghua University, China	FGOALS-g2	700
National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies (GISS), United States of America	GISS-E2-H	240
National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies (GISS), United States of America	GISS-E2-R	401
Institute of Numerical Mathematics (INM), Russia	INMCM4	500
Atmosphere and Ocean Research Institute (The University of Tokyo)/National Institute for Environmental Studies/Japan Agency for Marine-Earth Science and Technology, Japan	MIROC4h	100
Atmosphere and Ocean Research Institute (The University of Tokyo)/National Institute for Environmental Studies/Japan Agency for Marine-Earth Science and Technology, Japan	MIROC5	670
Atmosphere and Ocean Research Institute (The University of Tokyo)/National Institute for Environmental Studies/Japan Agency for Marine-Earth Science and Technology, Japan	MIROC-ESM	531
Atmosphere and Ocean Research Institute (The University of Tokyo)/National Institute for Environmental Studies/Japan Agency for Marine-Earth Science and Technology, Japan	MIROC-ESM-CHEM	255
Max Planck Institute for Meteorology (MPI-M), Germany	MPI-ESM-LR	1000
Max Planck Institute for Meteorology (MPI-M), Germany	MPI-ESM-MR	1000
Max Planck Institute for Meteorology (MPI-M), Germany	MPI-ESM-P	1156
Meteorological Research Institute (MRI), Japan	MRI-CGCM3	500
Norwegian Climate Centre (NCC), Norway	NorESM1-M	501

2.3 Statistical methods

Lag regression and correlation are the primary statistical methods used for studying relationships between SCE and atmospheric variables. Significance testing for temporal correlations and regressions is done through a two-sided Student t test with one degree of freedom per year. We also calculated effective degrees of freedom using the formulation from Bretherton et al. (1999), but the differences were small (~1–2). To complement the Student t test method, we also conducted significance testing using a Monte Carlo approach. Under this method, we generate 1,500 pairs of random time series representing the indices of interest. The random time series were formed in two ways: (1) using a simple autoregressive order-1 model with random numbers; and (2) taking the index and randomly shuffling the values to form a new time series. The determined significance level was not sensitive to the method chosen. Similar methodology was used to assess the significance of similarities

between the observed and modeled correlation/regression maps. In this case, we calculate a null distribution of pattern correlation values by (1) forming regression/correlation maps between each synthetic time series described above and the variable of interest and (2) calculating pattern correlations between the resulting synthetic maps and the observed map.

Monthly-mean anomalies are used throughout the study. For the observations, anomalies are calculated by removing the long-term climatological mean for each month in the record. For the models, however, hundreds of years of output exist for computing anomalies and relevant statistics. With 10–25 times more samples in the models than observations for statistical analyses, we choose an alternative approach to test the robustness of the results. The output from the long model integrations are subdivided into 40-year long sub-intervals, and monthly-mean anomalies are computed based on each 40-year climatology, along with the relevant statistics

Fig. 2 a Regression of observed November–March (NDJFM) SLPa (hPa) onto the principal component time series of NDJFM SLPa poleward of 20°N (i.e., the surface AO index). **b** As in **a** but for the CMIP5 ensemble-mean regression pattern. Sign convention of the anomalies chosen to represent the negative phase of the AO

Regression of NDJFM SLPa onto PC-1 of SLPa

(a) *Observations* **(b)** *CMIP5 Ensemble-Mean*

[hPa]

-7   0   7
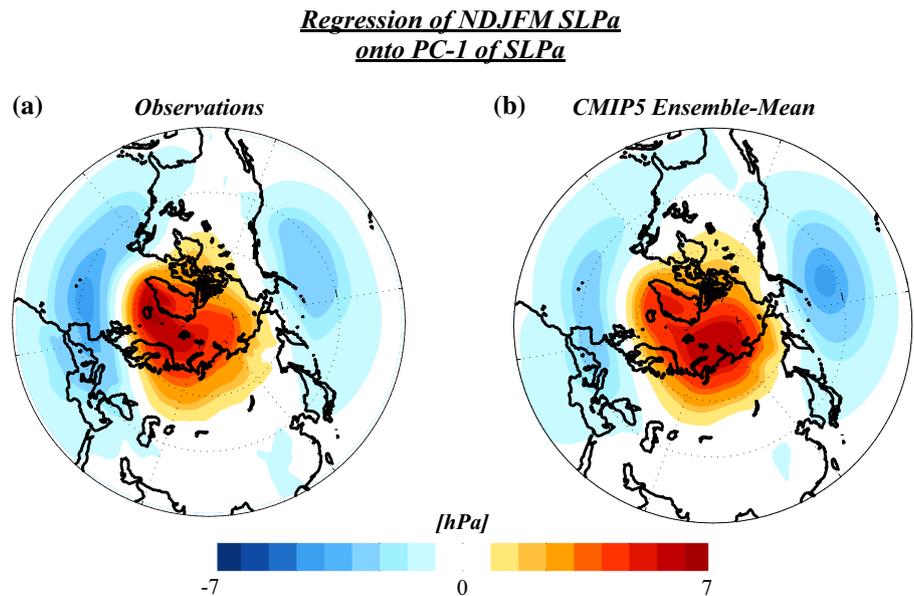
on that sub-interval. This parsing method was also done with 30-year and 50-year segments of the model output, along with performing no parsing of the record (i.e., using the full record). The aggregated results, including the multi-model ensemble-mean statistics, are highly similar, increasing confidence in our methodology. Before statistics are calculated in the observations and models, all anomalies are linearly detrended.

When comparing model performance to observations, we rely primarily on ensemble-mean statistics for the comparison, though we highlight individual model performance as well. For the individual models, we aggregate statistics among the several sub-intervals of a model to get representative statistics for that model. Multi-model ensemble-mean statistics are then calculated by averaging a particular statistic over all models.

3 Model evaluation of October Eurasian SCE and wintertime surface climate relationship

3.1 The AO signature in the models

Before exploring particular statistical relationships represented in Fig. 1, we examine how the CMIP5 models reproduce fundamental characteristics of the AO. Figure 2 presents the spatial signature of the AO (shown in the negative phase) in observations (Fig. 2a) and for the CMIP5 ensemble-mean (Fig. 2b). Here, the AO is defined as the leading empirical orthogonal function (EOF) of detrended November–March (NDJFM) monthly SLP anomalies (SLPa) poleward of 20 N. The observed AO pattern explains ~23 % of the total variance of the NDJFM NH SLPa field and

resembles that shown in prior studies, including the maximum of the positive SLPa loading center situated in the North Atlantic (Fig. 2a). The CMIP5 ensemble-mean AO pattern explains nearly the same fraction of variance as the observed (~25 %) and resembles quite closely the observed pattern. The pattern correlation between Fig. 2a, b is $r = 0.88$ ($p < 0.01$). This close correspondence in the AO spatial patterns is notable considering the background climate and boundary conditions are different in the piControl runs from the observations. Inter-model spread in the AO signature is quite small as well, with pattern correlations between the observed and model AO patterns ranging from to $r = 0.88$ to $r = 0.97$. Temporally, the standard deviation of the December–February (DJF) surface AO index (σ_{AO}) for the observations is $\sigma_{AO} = 1.70$, which is slightly less than that of the CMIP5 ensemble-mean ($\sigma_{AO} = 2.02$). For individual models, the standard deviation ranges from $\sigma_{AO} = 1.16$ in the FGOALS-g2 model to a notably high $\sigma_{AO} = 3.23$ in the BNU-ESM model.

3.2 Assessing the snow-AO dynamical framework in the models

Seeing that the CMIP5 models represent the surface-based AO sufficiently well, we now explore pieces of the six-step Eurasian snow–AO framework (Fig. 1). Figure 3 presents the mean October NH fractional snow cover for observations (Fig. 3a) and in the multi-model ensemble-mean (Fig. 3b). The spatial extent of snow cover has notable differences between the observations and the models. In particular, the observations show higher fractional snow cover (nearly 30–40 % higher) over northeastern Siberia and Canada. The ensemble-mean field exhibits slightly higher

(a) *Mean October Snow Cover (Observations)*

(b) *Mean October Snow Cover (CMIP5 Ensemble-Mean)*

[Percent Snow Cover]

1   50   100

(c) *Mean October Eurasian Areal Snow Cover*

(d) *Standard Deviation of October Eurasian Snow Cover Index*
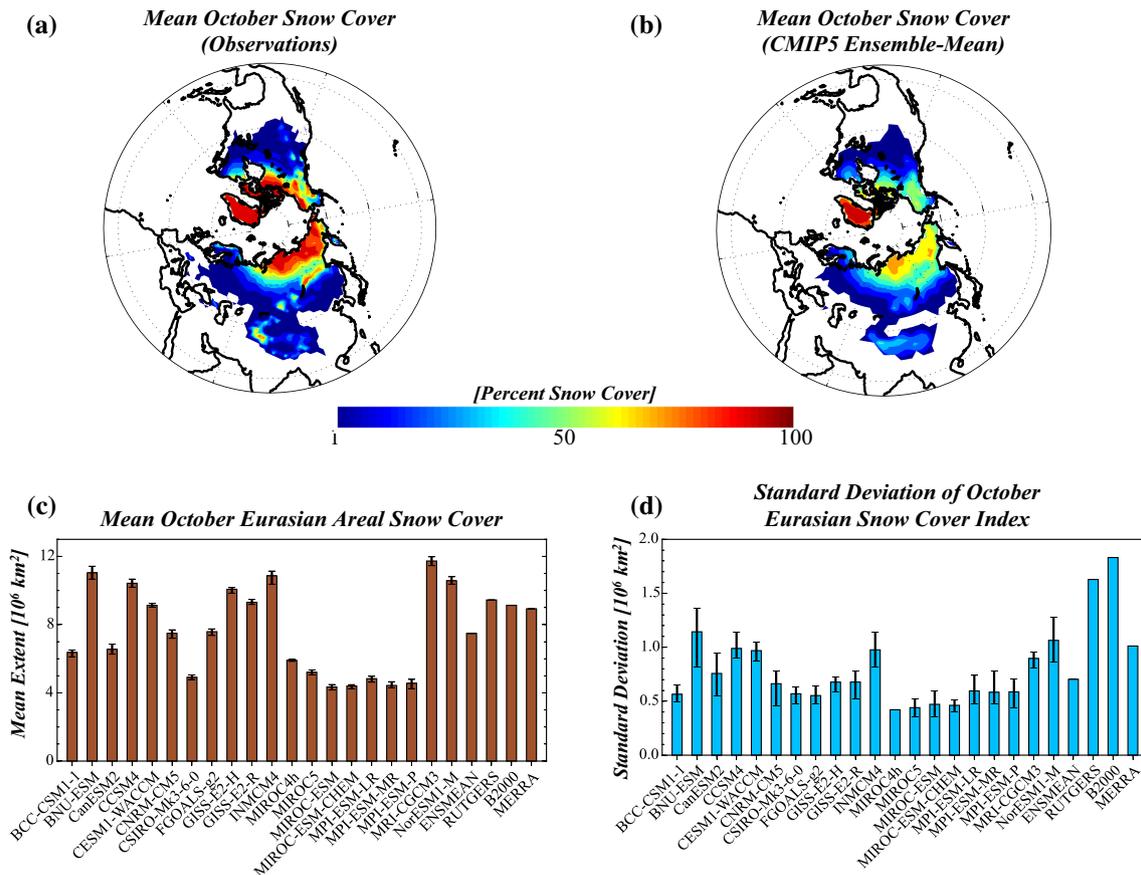
Fig. 3 a The mean October fractional snow cover (in percent) from 1979 to 2012 from the Rutgers snow cover dataset. Only coverage exceeding 1 % is plotted. **b** As in **a** but from the multi-model ensemble-mean. **c** The October monthly-mean Eurasian snow cover extent (10^6 km^2) for each model, the multi-model ensemble-mean, and three versions of observations: the Rutgers Eurasian SCE index, the long-term Eurasian SCE index from Brown (2000), and the MERRA Eurasian SCE index. *Black vertical bars* indicate the total spread in mean October SCE among the 40-year sub-intervals from the models analyzed. **d** The standard deviation of the October Eurasian SCE index (10^6 km^2) for each model, the multi-model ensemble-mean, and the three observational datasets. *Black vertical bars as in* **c**

October snow cover over eastern Europe and Scandinavia than the observations, but these larger snow cover amounts do not offset the deficits over Siberia when aggregated for the Eurasian SCE index.

Figure 3c illustrates the long-term mean October Eurasian SCE in observations and the models. The observed value from the Rutgers SCE dataset (9.45×10^6 km^2) is nearly 20 % larger than the ensemble-mean (7.64×10^6 km^2), but the inter-model spread is fairly large. Some models (i.e., the MIROC and MPI models) have nearly half the mean October Eurasian SCE, but others have a significantly higher SCE. However, spread within 40-year sub-intervals within the same model (denoted by the vertical bars in Fig. 3c) is very small. The standard deviation of the Rutgers October Eurasian SCE index is $\sigma = 1.63 \times 10^6$ km^2, which is significantly larger than that seen from all of the CMIP5 models, including all sub-intervals (Fig. 3d). The ensemble-mean standard deviation of the October Eurasian SCE index is 0.72×10^6 km^2,

less than half of that from the Rutgers snow cover dataset (Fig. 3d). The models with the closest level of variability to observations are the BNU-ESM and NorESM1-M models. The reduced interannual variability in October Eurasian SCE in the models agrees with Derkson and Brown (2012), who also found an underestimate of interannual spring SCE in the CMIP5 models. Coefficients of variation are fairly consistent among all models, with an absolute range from ~7 to 13 % and a mean coefficient of variation among all models of ~9.7 %. By comparison, the Rutgers snow cover dataset has a coefficient of variation of about 17 %. Ultimately, the lack of high variance in the CMIP5 October Eurasian SCE indices compared to observations mirrors results from other studies (e.g., Hardiman et al. 2008; Allen and Zender 2011). Hence, the issues with diagnosed snow cover in the coupled climate models persist with the latest generation of models.

Recently, Brown and Derkson (2013) identified different trends in several observational snow cover datasets,

including the NOAA satellite record (used for the Rutgers SCE index calculation), hence questioning the accuracy of the Rutgers snow record. To check for potential differences in the mean and standard deviation of the October Eurasian SCE in multiple datasets, Fig. 3c, d also include statistics from two other observational datasets: (1) A reconstruction of Eurasian October SCE from station data from 1922 to 1997 (Brown 2000), and (2) a computed October Eurasian SCE index using MERRA snow cover data (we elected not to use the ERA-Interim snow depth data given known issues with its fidelity. See, for example, http://old.ecmwf.int/research/era/do/get/index/QualityIssues). The statistics from the Brown (2000) dataset was computed by randomly choosing 34 years from the entire period and computing the mean and standard deviation for those years. This procedure was repeated 30 times and then averaged. For the mean, note that the values across all three observational datasets are highly comparable (Fig. 3c). In terms of standard deviation, both the Rutgers and Brown (2000) October Eurasian SCE indices are comparably high, with the latter the largest of all the datasets and models examined. However, the standard deviation of the MERRA October Eurasian SCE index is 80–100 % less than the other two observational datasets (Fig. 3d). The coefficient of variation for the Brown (2000) October Eurasian SCE index is about 21 %, or slightly higher than that from the Rutgers Eurasian SCE index. The MERRA Eurasian SCE index, however, has a coefficient of variation of only 10 %, which is closer to that from the CMIP5 models than the two other observational datasets.

The relationship between Eurasian SCE, the lower tropospheric circulation pattern, and vertical wave propagation is summarized in Fig. 4. In observations, the November SLPa pattern linearly related to October Eurasian SCE anomalies resembles a wave-2-like pattern. The direct response over northwestern Eurasia is anomalously high pressure at the surface with anomalously low pressure downstream across the North Pacific (Fig. 4a, line contours). Upstream across the North Atlantic, there is negative SLPa across the North Atlantic and slightly positive SLPa near the Azores. The multi-model mean October SCE/November SLPa regression pattern bears some resemblance to that from the observations, especially across Eurasia and the northwest Pacific, but with much weaker coefficients (Fig. 4b). However, the ensemble-mean pattern also shows positive SLPa over the North Atlantic, which is opposite of what is observed.

Also shown in Fig. 4a, b is the lag regression of November SLPa onto December 100 hPa vertical component of the wave activity flux (WAFz; proportional to the meridional heat flux; Plumb 1985) area-averaged from 40 to 80°N (shaded contours). As anomalously positive WAFz represents wave forcing that weakens the stratospheric polar vortex, this latter lag regression analysis depicts favorable
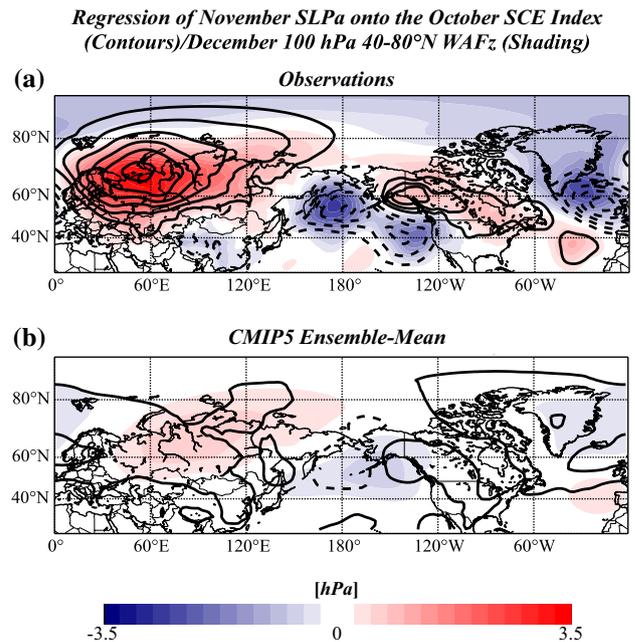

Fig. 4 **a** (*Contours*) Regression of observed November SLPa (hPa) onto the Rutgers October Eurasian SCE index. (*Shading*) Regression of November SLPa (hPa) onto December 100 hPa WAFz anomalies, area-averaged from 40 to 80°N. The pattern correlation between the two regression patterns is $r = 0.83$. Results similar to that in Cohen et al. (2014) but for ERA-Interim. **b** As in **a** but for the CMIP5 ensemble-mean. The pattern correlation between the two regression patterns is $r = 0.09$

near-surface circulation patterns that precede anomalously strong vertical wave propagation into the stratosphere in December. Note that the November SLPa/December 100 hPa WAFz regression analysis is done irrespective of the October Eurasian SCE. In observations, anomalously positive December 100 hPa WAFz is generally preceded by a hemispheric SLPa pattern that resembles the SLPa pattern linearly related to October SCE—i.e., positive SLPa throughout Eurasia and negative SLPa in the North Pacific and North Atlantic basins (Fig. 4a). Indeed, the spatial correlation between the two regression patterns in observations is $r = 0.83$ (see also Fig. 5a). Overall, this finding suggests that the November SLP pattern related to anomalously high October Eurasian SCE is spatially similar to a favorable precursor pattern to strong vertical wave propagation (e.g., Garfinkel and Hartmann 2010; Kolstad and Charlton-Perez 2011; Cohen and Jones 2011). For the CMIP5 multi-model ensemble-mean, the November SLPa pattern preceding anomalously large poleward heat flux in December matches well with that from the observations (the pattern correlation between shaded contours in Fig. 4a, b is $r = 0.65$). However, the spatial correlation between the line contours (i.e., the October Eurasian SCE-November SLPa relationship) and the shaded contours (i.e., the tropospheric precursor

(a)

(b)
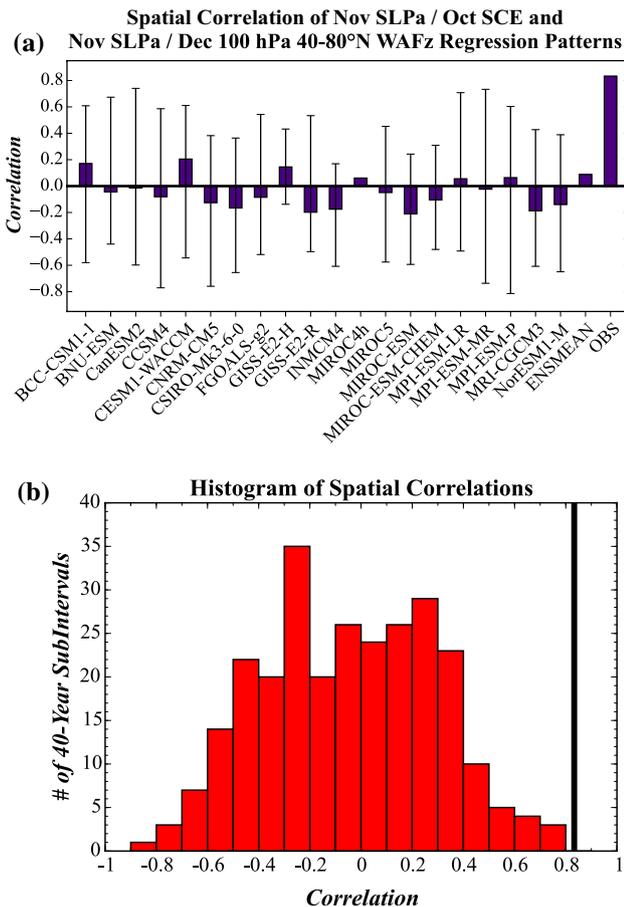
Fig. 5 a The spatial correlation between the November SLPa/October Eurasian SCE and November SLPa/December 100 hPa 40–80°N WAFz regression patterns for each model, the multi-model ensemble-mean, and the observations. *Vertical bars* indicate the range in spatial correlation values among the individual 40-year sub-intervals in the models. **b** Histogram of the spatial correlations between the November SLPa/October Eurasian SCE and November SLPa/December 100 hPa 40–80°N WAFz from the individual 40-year sub-intervals in the models. *Thick black line* denotes the correlation from observations (i.e., ERA-Interim)

pattern to strong vertical wave propagation into the stratosphere) in the ensemble-mean is only $r = 0.09$ (Fig. 4b). Most of the individual models also do not simulate the correspondence between the October Eurasian SCE/November SLPa association and the tropospheric precursor to strong vertical wave propagation (Fig. 5a). However, the spread among 40-year sub-intervals is large and spans both positive and negative values in almost all of the models (vertical lines in Fig. 5a). The spread is especially seen in the histogram of the spatial correlations from the 279 40-year sub-intervals of the models (Fig. 5b). Note that the observed correlation lies just outside the entire distribution of simulated spatial correlations. Indeed, more than half of the 40-year sub-intervals have a negative spatial correlation between the November SLPa pattern and the December

100 hPa WAFz pattern. By contrast, 45 of the 40-year sub-intervals (~17 %) have spatial correlations greater than 0.3 (Fig. 5b), indicating that one out of six 40-year sub-intervals from the CMIP5 multi-model suite reproduces the observed relationship decently well.

We now investigate more directly the observed relationship between October Eurasian SCE and impacts on the stratospheric circulation. Figure 6 shows the lag correlation between the October Eurasian SCE index and two quantities: WAFz anomalies, area-averaged from 40 to 80°N (Fig. 6a, b), and polar cap geopotential height anomalies (i.e., geopotential height anomalies area-averaged from 60 to 90°N; Fig. 6c, d). For WAFz, the observations show two significant regions of positive correlation (i.e., anomalous upward WAF/poleward heat flux associated with anomalously high October Eurasian SCE)—one in the lower and middle troposphere during November and another in the lower and middle stratosphere during December (Fig. 6a), representing Steps 3 and 4 in Fig. 1. The multi-model ensemble-mean WAFz correlation plot, however, exhibits weak contemporary correlations with negligible signals in November and December (Fig. 6b). This weak signal in the ensemble-mean exists likely because of the large spread in the relationship between individual models (and even 40-year sub-intervals within the model), both in timing (i.e., appropriate lags) and the sign of the relationship. Of all the 40-year sub-intervals from the multiple models, only 17 segments (i.e., ~6 % of all segments) show lag correlation plots with a pattern correlation with the observed plot exceeding $r = 0.35$, a threshold approximately equal to the 90 % significance level. Analysis of other statistical relationships (i.e., analyses in Fig. 4) for those 17 segments reveal no systematic pattern or other agreement with the observations or with each other, however. The weak October SCE/late fall WAFz relationship may also reflect some issues with stratosphere–troposphere coupling in the models, upon which we elaborate further in Sect. 4.

The observed relationship between wintertime polar cap height anomalies and the October Eurasian SCE is also missing in the models (Fig. 6c, d). In observations, significant correlations between October Eurasian SCE and the polar cap height exist first in the lower troposphere during October and November (due to higher than normal heights over the Arctic Ocean initially that then build onto the Eurasian continent) and then in the lower stratosphere and throughout the troposphere in January and February (Fig. 6c). The multi-model mean polar cap height lag correlation plot, however, shows only a contemporaneous geopotential height signal throughout the atmospheric column associated with the October Eurasian SCE index and no lagged relationship into the late fall and winter months (Fig. 6d). A similar dissection of the analysis on the individual 40-year time slices from the models shows that only eight 40-year

Fig. 6 a Lag correlation of observed WAFz anomalies area-averaged between 40 and 80°N and the Rutgers October Eurasian SCE index. *Red line* denotes correlation values exceeding the $p < 0.1$ level. **b** As in **a** but for the multi-model ensemble-mean. **c** Lag correlation of geopotential height anomalies area-averaged between 60 and 90°N and the October Eurasian SCE index. *Red line* as in **a**. **d** As in **c** but for the multi-model ensemble-mean. Contour interval 0.1 (...,−0.15, −0.05, 0.05, 0.15,...). Positive (negative) correlation values in solid (*dashed*) contours. All *shaded* coefficients in **b** and **d** are statistically significant at the $p < 0.1$ level
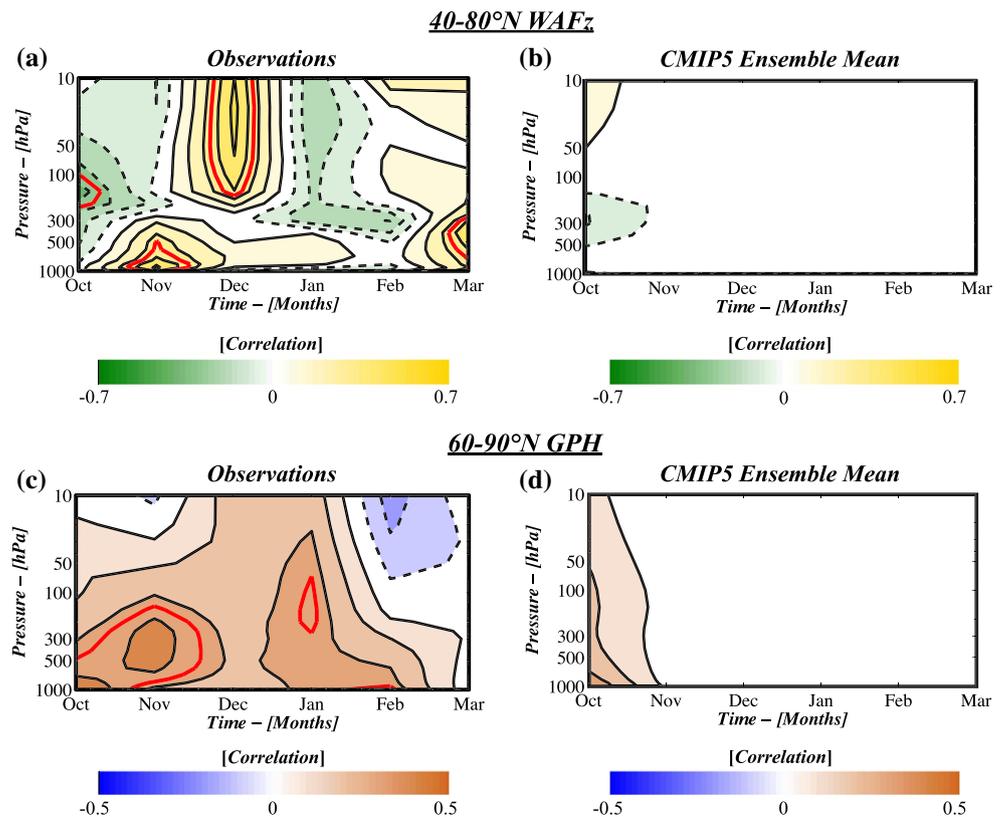
sub-intervals (~3 %) have pattern correlations with Fig. 6c exceeding the 90 % significance level. But, the significance arises mainly due to the October signal—i.e., those eight sub-intervals still lack a lagged relationship later in the fall and winter to October Eurasian SCE.

Based on the above results, we expect that the relationship between the October Eurasian SCE index and the DJF surface AO index in the models will also be weak; Fig. 7a supports this conclusion. The October Eurasian SCE–DJF AO index relationship in observations is significant and explains nearly 18 % of the variance in the DJF AO surface index over the last three decades ($r = -0.42$; $p < 0.05$). In considering the other two observational SCE indices included in Fig. 3c, d, for 1979–1997 (i.e., the common period for all three snow datasets), the correlations are $r = -0.60$ ($p < 0.05$) for the Rutgers SCE index, $r = -0.57$ ($p < 0.05$) for the Brown (2000) SCE index, and $r = -0.43$ ($p < 0.1$) for the MERRA SCE index. In considering the full period 1979–2012, the correlation between the DJF AO index and the MERRA SCE index falls to an insignificant $r = -0.13$. The relatively weak correlation between the DJF AO index and the MERRA October Eurasian SCE index for the full period is a function of the poor statistics of SCE noted before. When considering the CMIP5 models, mixed results arise—some models indicate a weakly negative correlation, while many have a *positive* correlation between October Eurasian SCE and

the DJF AO index. The vertical bars plotted in Fig. 7 represent the total range of correlation coefficients among the 40-year sub-intervals used for the analysis. Except for the MIROC4h (which has the shortest temporal record of the models; Table 1), all models show a very large range (spanning either side of zero) in correlation between the October Eurasian SCE index and the DJF surface AO index among the 40-year sub-intervals. The ensemble-mean correlation, as expected by this large spread, is close to 0 ($r = 0.01$). The distribution of correlations between all of the 40-year sub-intervals from the models appears nearly normal with notable outliers on both the negative and positive ends (Fig. 7b). Nearly 30 % (10 %) of the 40-year sub-intervals analyzed show an October Eurasian SCE–DJF AO index correlation less than −0.1 (−0.3), indicative that there is a subset of models and periods of time that capture the observed relationship. Conversely, 27 % of the 40-year sub-intervals show a positive correlation greater than 0.1.

4 Stratosphere–troposphere coupling dynamics in the models

Results from Sect. 3 overwhelmingly illustrate that the observed October Eurasian SCE–DJF AO connection (Fig. 1) is not present in most models but can be found in select 40-year sub-intervals of some of the models. The absence

(a)

(b)
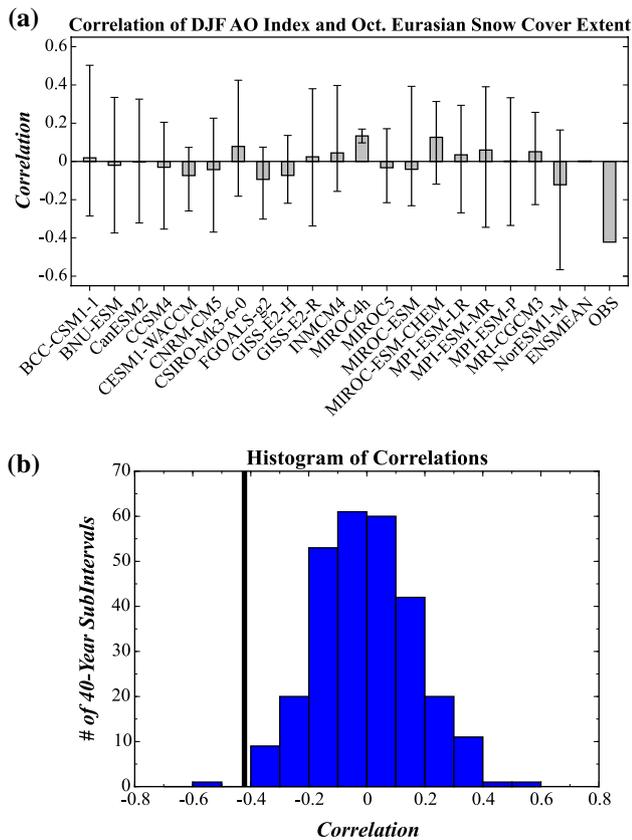
Fig. 7 a The correlation between the October Eurasian SCE index and the following DJF AO index in each model, the multi-model ensemble-mean, and the observations. *Vertical bars* indicate the range in correlation values among the individual 40-year sub-intervals in the models. **b** Histogram of the DJF AO—October Eurasian SCE correlations from the individual 40-year sub-intervals in the models. *Thick black line* denotes the correlation from observations

of this mechanism may be related to inherent internal variability associated with the observed relationship, issues with simulated snow cover, and/or issues with more fundamental stratosphere–troposphere coupled dynamics in the models. In this section, we select several diagnostics representing dynamical stratosphere–troposphere coupling to investigate whether the models can correctly capture the necessary dynamics needed for the troposphere-stratosphere–troposphere pathway associated with the October Eurasian SCE–AO connection.

The initiation and subsequent propagation of the waves that later impact the polar stratospheric circulation are related to background zonal-mean conditions in the troposphere and lower stratosphere (e.g., Charney and Drazin 1961). Figure 8 investigates the relationship between the zonal-mean circulation and October Eurasian SCE through regression of October zonal-mean temperatures, zonal-mean zonal winds, and Eliassen–Palm (EP) flux components (Eliassen and Palm 1961; Edmon et al. 1980) onto

the October Eurasian SCE index in both the observations (Fig. 8a, c) and the models (Fig. 8b, d). Because monthly-mean data and output are used, only the relationship with quasi-stationary eddies are investigated in this analysis. In observations, significant negative zonal-mean temperature anomalies (Ta) in the middle latitude troposphere are juxtaposed with significant positive tropospheric Ta in the high latitudes (Fig. 8a). This zonal-mean pattern is reminiscent of the "warm Arctic-cold continents" pattern (Overland et al. 2011; Cohen et al. 2013). The extratropical zonal-mean temperature pattern also reveals opposite-signed Ta in the troposphere and lower stratosphere. The resulting weaker meridional temperature gradient in the troposphere is associated with a weakening of the zonal-mean zonal jet at higher latitudes (Fig. 8c). Anomalously high Eurasian SCE is also associated with anomalously strong poleward heat fluxes (i.e., upward EP flux vectors) from about 25 to 40°N and again from 60 to 70°N and altogether anomalous poleward wave propagation in the middle latitudes in both the troposphere and stratosphere (Fig. 8c).

The ensemble-mean zonal-mean Ta regression pattern (Fig. 8b) captures the lower tropospheric cooling/warming dipole between middle and high latitudes, but its magnitude is weaker than in the observations, particularly in the Arctic. Consequently, the ensemble-mean zonal-mean zonal wind changes associated with the October Eurasian SCE is also weaker, especially in the high-latitude troposphere, and also exhibits a more equivalent barotropic structure than in the observations (Fig. 8d). Wave propagation as represented by the EP flux vectors is similar to observations in the troposphere but weaker. More mixed signals are observed in the polar stratosphere, with even a suggestion for equatorward heat flux anomalies (i.e., downward-pointing EP flux vectors) there.

Overall, Fig. 8 shows that the CMIP5 models simulate well the observed relationships between October Eurasian SCE and the October tropospheric zonal-mean circulation pattern. But, later in the fall and early winter, what is the relationship between vertically propagating waves in the lower stratosphere and troposphere? Figure 9 presents the lag correlation between area-averaged (40–80°N) WAFz anomalies at 100 hPa during December (i.e., the same index used as the basis of regression for the shaded contours in Fig. 4a, b) and the inverted AO index at each pressure level. In this and subsequent analyses, the AO at a given pressure level is defined as the leading mode of detrended NDJFM geopotential height anomalies poleward of 20°N on that pressure level. Because a negative AO index is associated with anomalously high heights in the polar atmosphere, the inversion of the AO index allows us to relate these results to polar cap height anomalies (i.e., Fig. 6c, d). Hence, positive correlations in Fig. 9 indicate that anomalously positive WAFz anomalies (i.e., stronger vertical wave propagation)

(a) Oct Zonal-Mean Ta Regressed onto the Oct Eurasian SCE Index (Obs)

(b) Oct Zonal-Mean Ta Regressed onto the Oct Eurasian SCE Index (EnsMean)

(c) Oct Zonal-Mean Ua and EP Fluxes Regressed onto the Oct Eurasian SCE Index (Obs)

(d) Oct Zonal-Mean Ua and EP Fluxes Regressed onto the Oct Eurasian SCE Index (EnsMean)
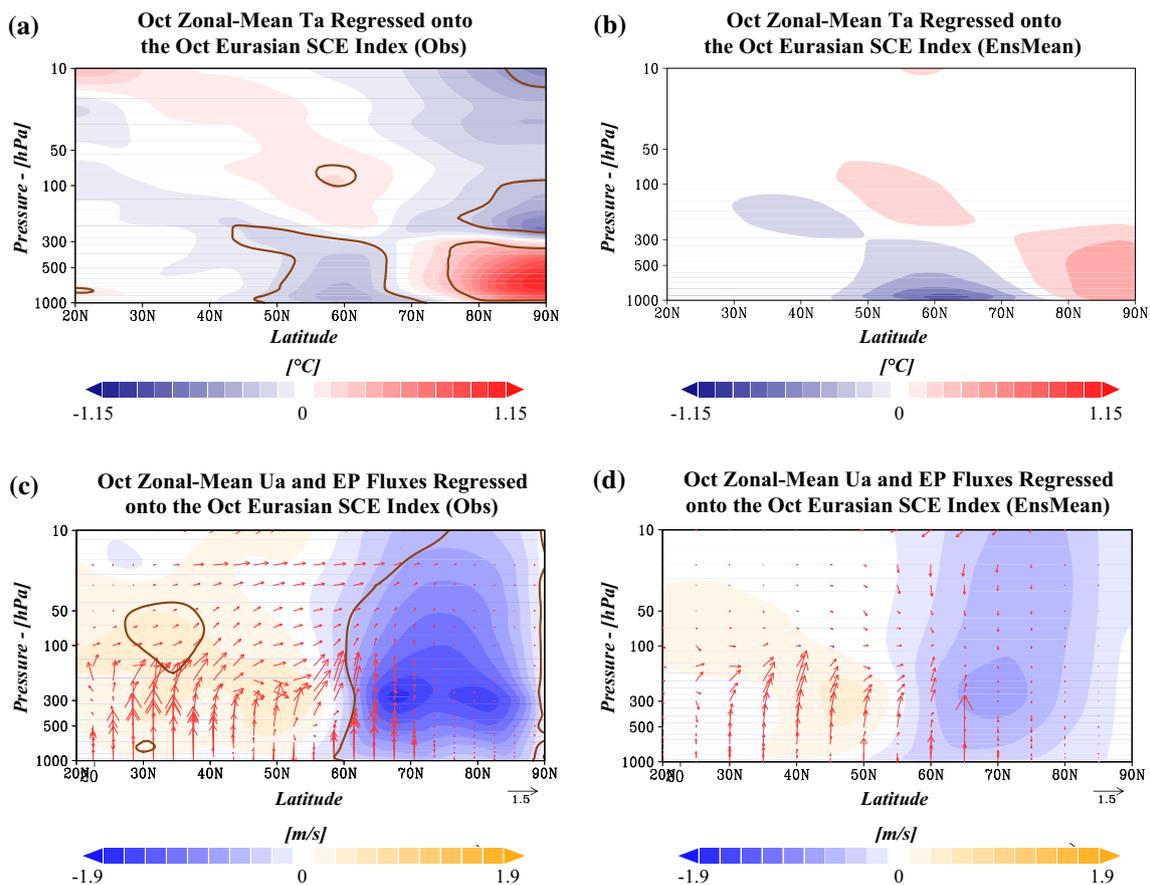
Fig. 8 a Regression of October zonal-mean Ta (°C) onto the Rutgers October Eurasian SCE index in the observations. Contour interval $0.1°C$ $(... , -0.15, -0.05, 0.05, 0.15, ...)$. Thick brown line denotes significant coefficients at the $p < 0.1$ level. **b** As in **a** but for the multi-model ensemble-mean regression pattern. **c** (*shading*) Regression of observed October zonal-mean zonal wind anomalies (Ua; m/s) onto the Rutgers October Eurasian SCE index. Contour interval 0.2 m/s $(..., -0.3, -0.1, 0.1, 0.3, ...)$. (*arrows*) Regression of the anomalous observed horizontal and vertical components of the EP flux $(F_\phi, F_p; J/kg)$ in October onto the Rutgers October Eurasian SCE index. *Arrows* scaled for plotting as: $\rho \cos\phi \left(\dfrac{F_\phi}{as_\phi}, \dfrac{F_p}{s_p} \right)$, where $\rho = \sqrt{\dfrac{1000}{pressure}}$, ϕ is latitude, a is the radius of Earth, $s_\phi = \pi$ radians, and $s_p = 1 \times 10^5$ Pa. Reference vector in scaled units included. *Thick brown line* as in **a**. **d** As in **c** but for the multi-model ensemble-mean regression pattern. In **b** and **d**, all *shaded* coefficients are significant at the $p < 0.1$ level

are related to more positive heights at the pole and hence a more *negative* AO index. In observations (Fig. 9a), two statistically significant ($p < 0.1$) features are seen: (a) positive correlations in the stratosphere during January and February, with some indication of downward propagation of the AO signal into the troposphere; and (b) a strong signal throughout the troposphere in October. For the multi-model ensemble-mean (Fig. 9b), the stratospheric signal in January and February is still apparent, although connections with the troposphere are much weaker than observed. The October signal is not replicated by the models, however. This feature may have ties to October Eurasian SCE variability (e.g., Cohen et al. 2014; also Fig. 6c), but how important its relationship is to late fall vertical wave propagation remains to be studied.

Finally, the models also do not simulate well the "downward propagation" of wave-induced stratospheric anomalies into the troposphere (e.g., Haynes et al. 1991; Baldwin and Dunkerton 1999). Figure 10 depicts this downward propagation signal via lag correlation of the January AO index at 10 hPa (AO_{10}) with the AO index at all pressure levels (e.g., a similar diagnostic done by Baldwin and Dunkerton (1999) except here we use monthly-mean values) for the observations (Fig. 10a) and the CMIP5 ensemble-mean (Fig. 10b). Observations illustrate strong downward propagation of changes in the January AO_{10}, as evidenced by positive, significant correlation coefficients extending throughout the atmospheric column below 10 hPa during January. The maximum correlation at the surface is $r \approx 0.4$ ($p < 0.05$) in January (Fig. 10a), while

**Lag Correlation of December 100 hPa WAFz (40-80°N)
with the Inverted AO Index at Each Level**
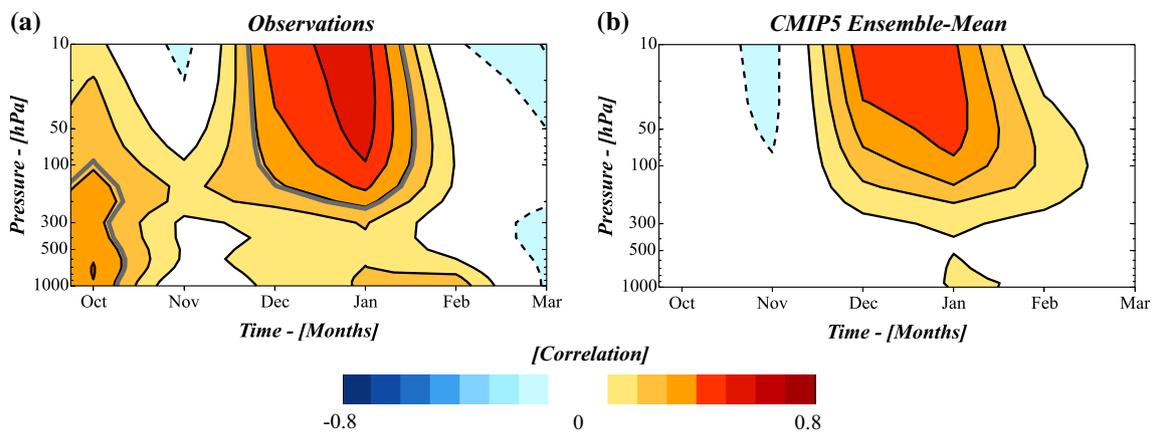
Fig. 9 **a** Lag correlation between the inverted AO index at each level and the December 100 hPa WAFz anomalies from 40 to 80°N. *Thick gray line* shows significance at the $p < 0.1$ level. **b** As in **a** but for the multi-model ensemble-mean

**Lag Correlation of Jan AO₁₀ with the
AO Index at All Levels**

Lag Correlation of Jan AO_{10} with the AO Index at All Levels
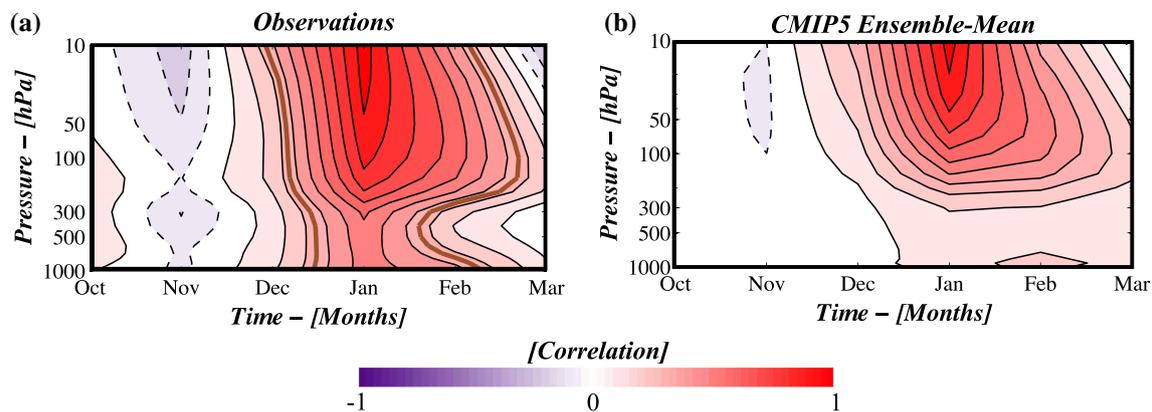
Fig. 10 **a** Lag correlation of the January AO index at 10 hPa (AO_{10}) with itself and the AO index at all other pressure levels. *Thick brown line* outlines correlation coefficients significant at the $p < 0.1$ level. **b** As in **a** but for the CMIP5 ensemble-mean correlation. All shaded coefficients in **b** are significant at the $p < 0.1$ level. Contour interval in both plots 0.1 (..., -0.15, -0.05, 0.05, 0.15, ...). Positive (negative) correlation values in solid (*dashed*) contours

positive, significant correlation values persist in the lower stratosphere into February. The downward propagation signature in the CMIP5 ensemble-mean differs from observations, especially in the troposphere (Fig. 10b), however. The most striking difference is the low correlation between the January AO_{10} and January AO_{1000} indices in the models. When examining individual models, the correlation coefficients between the January AO_{10} and January AO_{1000} indices vary from $r = -0.40$ in the INMCM4 model to $r = 0.31$ in the MPI-ESM-MR model. Instead of simultaneous correlation with the January AO_{10}, significant positive (albeit weak) correlation coefficients exist at 1,000 hPa

in February and March. Lower stratospheric memory is simulated well in the models, though, with positive correlation coefficients remaining from January into February. We also subdivided the models into "high–top" (i.e., the lid height of the model was at least 1 hPa) and "low–top" models (i.e., all other models) and re-constructed Fig. 10b for each of those subdivisions. The results were similar to the ensemble-mean with only slightly higher correlations at the surface for the high-top ensemble, but in February, not January (not shown). Figure 10 thus illustrates that downward propagating signals from the stratosphere into the troposphere are likely muted or not captured at all in the

models. Hence, more work needs to be done to investigate model deficiencies in capturing the physics and dynamics of stratosphere–troposphere dynamical coupling.

5 Discussion and conclusions

This study evaluated how the latest state-of-the-art coupled climate models simulate an observed relationship between October Eurasian SCE and the following wintertime atmospheric circulation regime. By systematically testing the components outlined in Fig. 1, we showed that the coupled climate models from CMIP5 do not consistently replicate the observed relationship. As in Hardiman et al. (2008), the CMIP5 models do not reproduce the lagged response in either late fall and early winter geopotential height or vertical wave propagation to variability in October Eurasian SCE (e.g., Fig. 6b, d). In examining the direct relationship between October Eurasian SCE and the DJF AO index, we find that the models exhibit a wide range of correlations both positive and negative (Fig. 7), indicating there are periods of time when most of the models will capture something close to the observed relationship. Therefore, internal variability of the climate system may play a role in setting the observed snow–AO relationship. However, the CMIP5 models capture the first-order response of the stratosphere to tropospheric wave driving (Fig. 9b), though the downward propagation of those anomalies is questionable in most models (Fig. 10). Implications of these latter dynamical analyses involving stratosphere–troposphere coupling are important for understanding how the models handle wintertime climate variability in both the stratosphere and the troposphere, irrespective of the existence of the snow–AO relationship.

A recurring issue seen with the CMIP5 models is that the model-generated snow cover over Eurasia (and the NH) overall is overall less and exhibits less interannual variability than the observations (Fig. 3; Derkson and Brown 2012). We examined the boreal fall mean and monthly climatological land surface temperatures in the models and found them to be sufficiently cold in Eurasia and North America for snowfall (not shown). Though unresolved, other potential problem areas associated with the snow deficiency may be in precipitation generation and/or land-surface parameterizations for accumulating snowfall. Irrespective of its implications for our dynamical framework, poor snowfall and snow cover representation in the coupled climate models results in weak or incorrect thermodynamical and dynamical responses in the atmosphere. Future modeling studies should focus on assessing this snow cover/snowfall issue, as its effects permeate to other important quantities not considered in this study (e.g., soil moisture, river runoff).

One important conclusion is that the CMIP5 models have weak lagged responses between fall anomalies and wintertime regimes. The simulated October Eurasian SCE is related to October zonal-mean temperatures and zonal winds in the models similar to the observations (Fig. 8). However, by November, the circulation response is already much weaker or non-existent in the models versus the observations (Fig. 6d). Moreover, the spatial pattern of the November surface circulation response to October SCE anomalies is not nearly as well-aligned with the optimal precursor pattern for vertical wave propagation (i.e., the main driver of extratropical stratospheric variability) as it is in the observations (Fig. 4). There is also little connection between late fall vertical wave propagation and October Eurasian SCE (Fig. 6b). Moreover, the subsequent impact of stratospheric circulation anomalies on the tropospheric circulation is much weaker in the models than in observations. The overall weaker wave propagation seen in the lower stratosphere in the models could be the result of poor eddy-mean flow interactions or weaker source regions, but this remains to be further explored.

This work builds upon a growing body of literature on the performance of CMIP5 models at simulating extratropical climate variability. Charlton-Perez et al. (2013) illustrate that the historical runs of the CMIP5 models simulate a lower frequency of sudden stratospheric warmings than observations and, relatedly, the variance of the polar vortex is also weaker. Similar findings are also seen in the piControl runs of those models in our study (not shown), suggesting that the low variance bias is systematic to the models themselves and not related to changes in external forcings. Weaker variability in the model stratospheric polar vortices versus observations may be tied to the strength of upward wave fluxes into the stratosphere and the tropospheric precursor patterns that initiate these waves (e.g., Fig. 4). There are also likely feedbacks between the weakened stratospheric polar vortex variability and poor downward propagation of stratospheric anomalies in the models (Fig. 10b). Given these results about stratosphere–troposphere coupling dynamics, projected NH winter climate from the latest models may be less certain than currently thought and should be evaluated more rigorously. Even operational forecasting models like the NOAA Coupled Forecast System version 2 (CFSv2) show weak lagged relations between fall boundary conditions and the circulation pattern the following winter (Riddle et al. 2013).

There are two main caveats to our study. First, testing any observed relationship in the CMIP5 model suite is problematic given the relatively short observational sample size (here, $N = 34$) versus the very large sample size from the models and their multiple ensemble members. We have attempted to account for this aspect by analyzing shorter subsets of the model runs and reporting on the variance in

the statistical relationships among individual 40-year sub-intervals within each model. The relatively large model spread in certain relationships (e.g., spatial correlations presented in Fig. 5 and the October Eurasian SCE index–DJF AO index correlations in Fig. 7) suggests that there may be periods of better agreement between the models and observations than other periods. Individual examination of these "good" periods, however, yield mixed signals in terms of the other elements of the snow–AO relationship (Fig. 1). Yet, the distributions of the spatial correlations of the November SLPa/December WAFz regression pattern (Fig. 5b) and the October Eurasian SCE–DJF AO index (Fig. 7b) illustrate some 40-year sub-intervals reproduce correlations near the observed value. This introduces the second caveat to this study—the role of internal variability in the observed snow–AO relationship. Indeed, Peings et al. (2013) find that the Siberian snow cover—AO connection is nonstationary in time and may have emerged in the latter part of the twentieth century only. Hence, internal variability may play a significant role in the observed relationship, and the century-long climate model runs may not be expected to faithfully replicate this relationship because of that. To better understand the contribution of the forced versus internal variability component of the Eurasian SCE–AO link, more studies in understanding the troposphere–stratosphere–troposphere pathway outlined in Fig. 1, particularly the early steps, are needed. Cohen et al. (2014) offers additional evidence of the dynamical forced pathway at work in observations, but more targeted modeling experiments are needed to confirm these findings.

Future model evaluation of wave generation and propagation related to October Eurasian SCE should rely on higher temporal resolution than presented here (i.e., daily-mean output). Observational analyses that track wave "pulses" or wave source regions associated with Eurasian snow cover would be particularly useful, accompanied with modeling experiment results. Such a study would further understanding on wave-mean flow interactions and feedbacks and may elucidate missing or incorrect model dynamics responsible for poorly recovering the snow–AO dynamical relationship.

References

Allen RJ, Zender CS (2010) Effects of continental-scale snow albedo anomalies on the wintertime Arctic oscillation. J Geophys Res 115:D23105. doi:10.1029/2010JD014490

Allen RJ, Zender CS (2011) Forcing of the Arctic oscillation by Eurasian snow cover. J Clim 24:6528–6539

Baldwin MP, Dunkerton TJ (1999) Propagation of the Arctic oscillation from the stratosphere to the troposphere. J Geophys Res 104:30 937–30 946

Baldwin MP, Stephenson DB, Thompson DWJ, Dunkerton TJ, Charlton AJ, O'Neill A (2003) Stratospheric memory and extended-range weather forecasts. Science 301:636–640

Bretherton CS, Widmann M, Dymnikov VP, Wallace JM, Bladé I (1999) The effective number of spatial degrees of freedom of a time-varying field. J Clim 12:1990–2009

Brown RD (2000) Northern Hemisphere snow cover variability and change, 1915–1997. J Clim 13:2339–2355

Brown RD, Derkson C (2013) Is Eurasian October snow cover extent increasing? Env Res Lett 8:024006. doi:10.1088/1748-9326/8/2/024006

Brown RD, Robinson DA (2011) Northern Hemisphere spring snow cover variability and change over 1922–2010 including an assessment of uncertainty. Cryosphere 5:219–229

Cattiaux J, Cassou C (2013) Opposite CMIP3/CMIP5 trends in the wintertime Northern Annular Mode explained by combined local sea ice and remote tropical influences. Geophys Res Lett **40**, 3682–3687

Charlton-Perez AJ, Baldwin MP, Birner T, Black RX, Butler AH, Calvo N, Davis NA, Gerber EP, Gillett N, Hardiman S, Kim J, Krüger K, Lee YY, Manzini E, McDaniel BA, Polvani L, Reichler T, Shaw TA, Sigmond M, Son SW, Toohey M, Wilcox L, Yoden S, Christansen B, Lott F, Shindell D, Yukimoto S, Watanabe S (2013) On the lack of stratospheric dynamical variability in the low-top versions of the CMIP5 models. J Geophys Res 118:2494–2505

Charney JG, Drazin PG (1961) Propagation of planetary-scale disturbances from the lower into the upper atmosphere. J Geophys Res 66:83–109

Cohen J (2003) Introducing sub-seasonal and temporal resolution to winter climate prediction. Geophys Res Lett 30:1018. doi:10.1029/2002GL016066

Cohen J, Entekabi D (1999) Eurasian snow cover variability and Northern Hemisphere climate predictability. Geophys Res Lett 26:345–348

Cohen J, Fletcher CG (2007) Improved skill of Northern Hemisphere winter surface temperature predictions based on land-atmosphere fall anomalies. J Clim 20:4118–4132

Cohen J, Jones J (2011) Tropospheric precursors and stratospheric warmings. J Clim 24:6562–6572

Cohen J, Barlow M, Kushner PJ, Saito K (2007) Stratosphere–troposphere coupling and links with Eurasian land surface variability. J Clim 20:5335–5343

Cohen J, Furtado JC, Barlow MA, Alexeev VA, Cherry JE (2012) Arctic warming, increasing snow cover and widespread boreal winter cooling. Env Res Lett 7:014007. doi:10.1088/1748-9326/7/1/014007

Cohen J, Jones J, Furtado JC, Tziperman E (2013) Warm arctic, cold continents: a common pattern related to Arctic sea ice, snow advance, and extreme winter weather. Oceanography 26:150–160

Cohen J, Furtado JC, Jones J, Barlow M, Whittleston D, Entekhabi D (2014) Linking Siberian snow cover to precursors of stratospheric variability. J Clim 27:5422–5432

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C,

Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Holm EV, Isaksen L, Kallberg P, Kohler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thepaut JN, Vitart F (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137:553–597

Derkson C, Brown R (2012) Spring snow cover extent reductions in the 2008–2012 period exceeding climate model projections. Geophys Res Lett 39:L19504. doi:10.1029/2012GL053387

Deser C, Tomas R, Alexander M, Lawrence D (2010) The seasonal atmospheric response to projected Arctic sea ice loss in the late twenty-first century. J Clim 23:333–351. doi:10.1175/2009JCLI3053.1

Edmon HJ, Hoskins BJ, McIntyre ME (1980) Eliassen-Palm cross sections for the troposphere. J Atmos Sci 37:2600–2616

Eliassen A, Palm E (1961) On the transfer of energy in stationary mountain waves. Geofys Publ 22:1–23

Fletcher CG, Kushner PJ, Cohen J (2007) Stratospheric control of the extratropical circulation response to surface forcing. Geophys Res Lett 34:L21802. doi:10.1029/2007GL031626

Fletcher CG, Hardiman SC, Kushner PJ, Cohen J (2009) The dynamical response to snow cover perturbations in a large ensemble of atmospheric GCM integrations. J Clim 22:1208–1222. doi:10.1175/2008JCLI2505.1

Foster JL, Owe M, Rango A (1983) Snow cover and temperature relationships in North America and Eurasia. J Clim Appl Meteorol 22:460–469

Garfinkel CI, Hartmann DL (2007) Effects of the El Niño–Southern Oscillation and the quasi-biennial oscillation on polar temperatures. J Geophys Res 112:D19112. doi:10.1029/2007JD008481

Garfinkel CI, Hartmann DL (2010) Tropospheric precursors of anomalous Northern Hemisphere stratospheric polar vortices. J Clim 23:3282–3299

Gerber EP, Vallis GK (2007) Eddy-zonal flow interactions and the persistence of the zonal index. J Atmos Sci 64:3296–3311

Ghatak D, Frei A, Gong G, Stroeve J, Robinson D (2010) On the emergence of an Arctic amplification signal in terrestrial Arctic snow extent. J Geophys Res 115:D24105. doi:10.1029/2010JD014007

Gillett NP, Fyfe JC (2013) Annular mode changes in the CMIP5 simulations. Geophys Res Lett doi:10.1002/grl.50249

Gong G, Entekabi D, Cohen J (2003) Modeled Northern Hemisphere winter climate response to realistic Siberian snow anomalies. J Clim 16:3917–3931

Hardiman SC, Kushner PJ, Cohen J (2008) Investigating the ability of general circulation models to capture the effects of Eurasian snow cover on winter climate. J Geophys Res 113:D21123. doi:10.129/2008JD010623

Haynes PH, Marks CJ, McIntyre ME, Shepherd TG, Shines KP (1991) On the "downward control" of extratropical diabatic circulations by eddy-induced mean zonal forces. J Atmos Sci 48:651–678

Hoerling MP, Hurrell JW, Xu T (2001) Tropical origins for recent North Atlantic climate change. Science 292:90–92

Hurrell JW, Hoerling MP, Phillips AS, Xu T (2004) Twentieth century North Atlantic climate change. Part I: assessing determinism. Clim Dyn 23:371–389

Ineson S, Scaife AA (2009) The role of the stratosphere in the European climate response to El Niño. Nat Geosci 2:32–36

Ineson S, Scaife AA, Knight JR, Manners JC, Dunstone NJ, Gray LJ, Haigh JD (2011) Solar forcing of winter climate variability in the Northern Hemisphere. Nature Geosci 4:753–757

Kistler R, Collins W, Saha S, White G, Woollen J, Kalnay E, Chelliah M, Ebisuzaki W, Kanamitsu M, Kousky V, van den Dool H, Jenne R, Fiorino M (2001) The NCEP-NCAR 50-year reanalysis: monthly means CD-ROM and documentation. Bull Am Meteor Soc 82:247–267

Kolstad EW, Charlton-Perez AJ (2011) Observed and simulated precursors of stratospheric polar vortex anomalies in the Northern Hemisphere. Clim Dyn 37:1443–1456

Lorenz DJ, Hartmann DL (2003) Eddy-zonal flow feedback in the Northern Hemisphere winter. J Clim 16:1212–1227

Maidens A, Arribas A, Scaife AA, MacLachlan C, Peterson D, Knight J (2013) The influence of surface forcings on prediction of the North Atlantic Oscillation regime of Winter 2010/11. Mon Wea Rev 141:3801–3813

Orsolini YJ, Kindem IT, Kvamstø NG (2011a) On the potential impact of the stratosphere upon seasonal dynamical hindcasts of the North Atlantic Oscillation: a pilot study. Clim Dyn 36:579–588

Orsolini YJ, Senan R, Benestad RE, Melsom A (2011b) Autumn atmospheric response to the 2007 low Arctic sea ice extent in coupled ocean–atmosphere hindcasts. Clim Dyn 114:D19108. doi:10.1007/s00382-011-1169-z

Overland JE, Wood KR, Wang M (2011) Warm Arctic-cold continents: climate impacts of the newly open Arctic Sea. Polar Res 30:15787. doi:10.3402/polar.v30i0.15787

Peings Y, Saint-Martin D, Douville H (2012) A numerical sensitivity study of the influence of Siberian snow on the Northern Annular Mode. J Clim 25:592–607

Peings Y, Brun E, Mauvais V, Douville H (2013) How stationary is the relationship between Siberian snow and Arctic Oscillation over the 20th century? Geophys Res Lett 40:183–188

Plumb RA (1985) On the three-dimensional propagation of stationary waves. J Atmos Sci 42:217–229

Riddle EE, Butler AH, Furtado JC, Cohen JL, Kumar A (2013) CFSv2 ensemble prediction of the wintertime Arctic Oscillation. Clim Dyn 41:1099–1116

Rienecker MM, Suarez MJ, Gelaro R, Todling R, Bacmeister J, Liu E, Bosilovich MG, Schubert SD, Takacs L, Kim GK, Bloom S, Chen J, Collins D, Conaty A, da Silva A (2011) MERRA - NASA's modern-era retrospective analysis for research and applications. J Clim 24:3624–3648

Robinson DA, Dewey KF, Heim RR (1993) Global snow cover monitoring: an update. Bull Am Meteorol Soc 74:1689–1696

Scaife AA, Spangehl T, Fereday DR, Cubasch U, Langematz U, Akiyoshi H, Bekki S, Braesicke P, Butchart N, Chipperfield MP, Gettelman A, Hardiman SC, Michou M, Rozanov E, Shepherd TG (2012) Climate change projections and stratosphere–troposphere interaction. Clim Dyn 38:2089–2097

Schneider EK, Bengtsson L, Hu ZZ (2003) Forcing of Northern Hemisphere climate trends. J Atmos Sci 60:1504–1521

Shindell DT, Miller RL, Schmidt GA, Pandolfo L (1999) Simulation of recent northern winter climate trends by greenhouse-gas forcing. Nature 399:452–455

Thompson DWJ, Wallace JM (1998) The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. Geophys Res Lett 25:1297–1300

Thompson DWJ, Wallace JM (2000) Annular modes in the extratropical circulation. Part I: month-to-month variability. J Clim 13:1000–1016