Climate Dynamics (2021) 56:1467–1485
https://doi.org/10.1007/s00382-020-05543-5

Summertime atmosphere–sea ice coupling in the Arctic simulated by CMIP5/6 models: Importance of large-scale circulation

Rui Luo[1,2] · Qinghua Ding[2] · Zhiwei Wu[1] · Ian Baxter[2] · Mitchell Bushuk[3] · Yiyi Huang[4] · Xiquan Dong[4]

Received: 6 July 2020 / Accepted: 16 November 2020 / Published online: 2 January 2021

Abstract

Summertime barotropic high pressure in the Arctic and its induced warmer and wetter atmosphere over sea ice are suggested to be important contributors to September sea ice loss on interannual and interdecadal time scales in the past decades. Using ERA5 and other reanalysis data, we find that atmospheric warming and moistening in the Arctic, synchronized by high latitude atmospheric circulation variability, work in tandem to melt sea ice through increasing downwelling longwave radiation at the surface. To what extent this atmosphere-longwave radiation-sea ice relationship can be captured in CMIP5 and 6 remains unknown and thus addressing this question is the objective of this study. To achieve this goal, we construct a process-oriented metric emphasizing the statistical relationship between atmospheric temperature and humidity with sea ice, which can effectively rank and differentiate the performance of 30 CMIP5 climate models in reproducing the observed connection. Based on our evaluation, we suggest that most available models in CMIP5 and 6 have a limitation in reproducing the full strength of the observed atmosphere–sea ice connection. This limitation likely stems from a weak impact of downwelling longwave radiation in linking sea ice with circulation associated with the weak sensitivity of the temperature and humidity fields to circulation variability in the Arctic. Thus, further efforts should be devoted to understanding the sources of these models' limitations and improve skill in simulating the effects of atmospheric circulation in coupling temperature, humidity, surface radiation and sea ice together during Arctic summer.

1 Introduction

In the past four decades, the Arctic has experienced substantial warming at a rate of twice the global average (Cohen et al. 2020), characterized by dramatic sea ice retreat and land ice melting in summer (Serreze et al. 2007; Comiso et al. 2008; Stroeve and Notz 2018; Mouginot et al. 2019). This warming has commonly been attributed to anthropogenic forcing and its influences on heat transport toward the Arctic through oceanic and atmospheric circulations and a number of local positive feedbacks collectively known as "Arctic Amplification" (Screen and Simmonds 2010; Notz and Marotzke 2012; Taylor et al. 2013; Pithan and Mauritsen 2014; Notz and Stroeve 2016; Stroeve and Notz 2018). Recent studies further suggest that atmospheric internal variability is an essential contributor to Arctic warming in the past decades. Atmospheric temperature fluctuations associated with moist-static energy (MSE) transport is a key driver for sea ice variability (Olonscheck et al. 2019). Low frequency summertime Arctic high pressure is demonstrated to explain 30–50% of the sea ice decline since the 1980s (Kay et al. 2011; Swart et al. 2015; Zhang 2015; Ding et al. 2017, 2019).

A key mechanism of this low-frequency high pressure forcing on sea ice variability is manifested as an atmosphere–sea ice coupling process: barotropic high pressure over Greenland and the Arctic Ocean in summer can

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00382-020-05543-5.

✉ Qinghua Ding
qinghua@uscb.edu

✉ Zhiwei Wu
zhiweiwu@fudan.edu.cn

[1] Department of Atmospheric and Oceanic Sciences, Institute of Atmospheric Sciences, Fudan University, Shanghai 200438, China

[2] Department of Geography, and Earth Research Institute, University of California, Santa Barbara, CA 93106, USA

[3] Geophysical Fluid Dynamics Laboratory, Princeton, NJ 08540, USA

[4] Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ 85721, USA

significantly cause sea ice decline from June to September through a circulation-driven adiabatic warming and moistening of the lower troposphere. By emitting more downward longwave radiation over sea ice, the warmer and wetter atmosphere is able to thermally melt sea ice (Ding et al. 2017, 2019). This mechanism is suggested to exist and operate over broad temporal scales ranging from interannual to interdecadal time scales. In addition, this anticyclonic circulation may also reduce sea ice cover through a dynamical ice drifting effect (Rampal et al. 2009, 2011; Ogi et al. 2010; Kwok 2011) and a thermodynamic impact from cloud changes in the Arctic (Sedlar and Tjernström 2017; Kay et al. 2016).

Given the importance of this atmosphere–sea ice linkage in contributing significant sea ice melt on year-to-year and low–frequency time scales, it is important to assess model fidelity in simulating the observed sea ice-atmospheric linkage. Previous studies have examined this linkage in climate models (Swart et al. 2015; Ding et al. 2017, 2019; Topál et al. 2020), however, owing to the lack of an efficient metric to evaluate models, model performance in reproducing this connection remains unclear. Thus, the main goal of this study is to perform a statistical-based evaluation of Phase 5/6 of the Coupled Model Inter-comparison Project (CMIP5/6) models to examine how well these state-of-the-art global climate models can replicate the observed impacts of summertime atmospheric forcing on September sea ice and whether this atmospheric forcing can be measured by changes in MSE of the atmosphere. We aim to design an effective metric to assess each individual model's skill in simulating the connection process and rank the metric scores of all CMIP5/6 models. Models in the highest and lowest ranking groups will allow us to better quantify model successes and insufficiencies in reproducing the observed atmosphere–sea ice connection and better characterize model biases and uncertainties in simulating observed physical processes determining recent sea ice melt.

2 Data

2.1 Atmospheric reanalysis and observational sea ice data

To fully consider the uncertainty associated with reanalysis data in the Arctic, we use five available satellite-assimilated monthly reanalysis datasets, including the fifth generation European Centre for Medium-Range Weather Forecasts Re-Analysis (ERA5) (Hersbach et al. 2020), ERA-Interim (Dee et al. 2011), MERRA-2 (Gelaro et al. 2017), the National Center for Environmental Prediction Re-analysis 2 (NCEP2) (Kanamitsu et al. 2002), and Japanese 55-yr Re-Analysis (JRA55) (Kobayashi et al. 2015), to understand circulation, temperature, humidity and radiation over the period of 1979–2018.

Monthly sea ice concentration observations are derived from the National Snow and Ice Data Center (NSIDC, Cavalieri et al. 1996). We use a total sea ice area (SIA) index to characterize sea ice variability in September. To construct this index, in each ocean grid cell, SIA is defined as the product of grid element area and sea ice concentration that is equal to or larger than 15%. The total SIA index is then calculated as the SIA sum of all grid cells in the Arctic. We also use this index to measure September sea ice viability in climate models.

2.2 CMIP5 and CMIP6 simulations

Our model evaluation will focus on all CMIP5 models (Taylor et al. 2012). To obtain a robust and reliable evaluation of model results, we calculate our metric primarily based on monthly long preindustrial control runs (PI), which contain long integrations spanning more than 200 years for all models (Table 1). Since reanalysis data in the satellite era only spans over a 40-year period, to make an appropriate and comprehensive comparison between models and reanalyses, a so-called pseudo-ensemble method is applied to pre-process model simulations before the evaluation. We generate many short 40-year periods (same size as the reanalysis record) by extracting any consecutive 40-year period from a long PI integration. In this way, we can form a large ensemble of 40-year short periods out of any sufficiently long integration although each period is not completely independent from other periods in the ensemble. By examining the model evaluation across all pseudo-ensemble members, we can understand how large the spreads of the evaluation scores are and how sensitive the model performance is to the length of simulations. Since variability simulated by these PI experiments is only due to internal variability, no temporal filtering is used to process PI model results. We also examine models' performance using the whole integration period of each PI run and find that they are very similar to the average scores of all pseudo-ensemble members derived from the corresponding run. Thus, we primarily use the pseudo-ensemble method to evaluate models in the following parts of the study.

To further understand whether these models behave differently in capturing the observed relationship when anthropogenic forcing is present, we also examine the monthly forced runs (Forced) over a 40-year period from 1979 to 2018 by combining the historical experiment (1979–2005) with the RCP8.5 experiment (2006–2018; Table 2). All datasets over 1979–2018 are linearly detrended before the metrics are calculated to focus on a direct connection between the atmosphere and sea ice. In addition, fifteen long PI runs (Table 3) from the

Table 1 List of 30 climate models of preindustrial control runs (PI runs) in CMIP5 and their temporal coverage

	CMIP5 model designation preindustrial (PI) long control run	Years	Pseudo-ensemble members
1	ACCESS1.0	500	461
2	ACCESS1.3	500	461
3	CanESM2	996	957
4	CMCC-CESM	277	238
5	CMCC-CM	330	291
6	CMCC-CMS	500	461
7	CNRM-CM5-2	490	451
8	CNRM-CM5	850	811
9	CSIRO-Mk3.6.0	500	461
10	FGOALS-g2	700	661
11	GFDL-CM3	500	461
12	GFDL-ESM2G	500	461
13	GFDL-ESM2M	500	461
14	GISS-E2-H-CC	251	212
15	GISS-E2-H	780	741
16	GISS-E2-R-CC	251	212
17	GISS-E2-R	850	811
18	HadGEM2-CC	240	201
19	HadGEM2-ES	314	275
20	INM-CM4	500	461
21	IPSL-CM5B-LR	300	261
22	MIROC5	870	831
23	MIROC-ESM-CHEM	255	216
24	MIROC-ESM	630	591
25	MPI-ESM-LR	1000	961
26	MPI-ESM-MR	1000	961
27	MPI-ESM-P	1156	1117
28	MRI-CGCM3	500	461
29	NorESM1-M	501	462
30	NorESM1-ME	252	213

The size of pseudo-ensemble members is calculated based on a sample size of 40-year

Table 2 List of 30 climate models of forced runs from CMIP5 (40-year, historical of 1979–2005 and RCP8.5 of 2006–2018)

	CMIP5 model designation Historical + RCP8.5 simulation of 1979–2018
1	ACCESS1.0
2	ACCESS1.3
3	CanESM2
4	CMCC-CESM
5	CMCC-CM
6	CMCC-CMS
7	CNRM-CM5
8	CSIRO-Mk3.6.0
9	* FGOALS-s2
10	GFDL-CM3
11	GFDL-ESM2G
12	GFDL-ESM2M
13	GISS-E2-H-CC
14	GISS-E2-H
15	GISS-E2-R-CC
16	GISS-E2-R
17	* HadGEM2-AO
18	HadGEM2-CC
19	HadGEM2-ES
20	INM-CM4
21	IPSL-CM5B-LR
22	MIROC5
23	MIROC-ESM-CHEM
24	MIROC-ESM
25	MPI-ESM-LR
26	MPI-ESM-MR
27	MRI-CGCM3
28	* MRI-ESM1
29	NorESM1-M
30	NorESM1-ME

Atmospheric variables and sea ice data of both Forced and PI runs are available in 27 models. Asterisks indicate models in which Forced runs are available while PI runs are not

latest CMIP6 models are also used to investigate the performance of these new generation models (Eyring et al. 2016). Due to limited diversity of the latest CMIP6 models that provide all necessary atmospheric and sea ice variables available at the time of this study, we only discuss the mean performance of these models. All model output in CMIP5 and 6 data and reanalysis data are regridded to the same horizontal (1.5° x 1.5°) and vertical resolutions (19 levels from 1000 to 100 hPa with an interval of 50 hPa) to facilitate a comparison between model results and reanalysis through an evaluation using a well-designed metric, which will be calculated and discussed in detail in Sect. 4.

2.3 Significance of correlation

The statistical significance of the correlation coefficient between two variables is examined by considering the effective sample size (Bretherton et al. 1999), which is calculated in the following manner:

$$N^* = N\frac{1 - r_1 r_2}{1 + r_1 r_2} \quad (1)$$

Table 3 List of 15 climate models of preindustrial long control runs (PI runs) in CMIP6 and their temporal coverage

	CMIP6 model designation	Years	Pseudo-ensemble members
1	CAMS-CSM1-0	500	461
2	CanESM5	1000	961
3	CESM2	1200	1161
4	CESM2-WACCM	499	460
5	E3SM-1-0	500	461
6	EC-Earth3	501	462
7	EC-Earth3-Veg	500	461
8	GISS-E2-1-H	801	762
9	HadGEM3-GC31-LL	500	461
10	IPSL-CM6A-LR	1200	1161
11	MIROC6	800	761
12	MRI-ESM2-0	701	662
13	NorCPM1	500	461
14	NorESM2-LM	301	262
15	SAM0-UNICON	699	660

The size of pseudo-ensemble members is calculated based on a sample size of 40-year

where N* is the effective sample size, N is the number of available time steps, r_1 and r_2 are lag-one autocorrelation coefficients of each variable. For the PI runs, the pseudo-ensemble method will yield an ensemble of multiple correlations and corresponding effective sample sizes from each 40-year long member. These correlations are averaged to represent an overall coupling strength for the model and their significance is examined using the two-tailed Student's *t*-test and the averaged effective sample size (Eq. 1) from all pseudo-ensemble members belonging to an individual model.

3 Observed and simulated atmosphere–sea ice connections

3.1 Role of circulations in the connections

To obtain a benchmark for the comparison with model simulations and evaluation, five reanalysis data are used to illustrate the strong observed connections of the detrended September SIA index with detrended summertime (June–July–August, JJA) geopotential height (Z), air temperature (T) and specific humidity (Q) for the period of 1979 to 2018 (Fig. 1 of ERA5 and Supplementary Figs. 1–4 of ERA-Interim, NCEP2, MERRA-2 and JRA55, respectively). The correlations using raw data exhibit more significant values, indicating that all these connections are evident on both year-to-year and low-frequency timescales. Since these five reanalysis datasets show very similar patterns and magnitudes (Supplementary Figs. 1–4), we focus our following evaluation on the patterns derived from ERA5.

Negative correlations in Fig. 1 suggest that decreased total September SIA is significantly related to a high pressure dominated Arctic with a warmer and moister atmosphere in JJA. The coupling process underlying these statistical relationships can be understood as the impacts of a barotropic anticyclonic circulation located over Greenland and the Arctic Ocean that warms and moistens the lower atmosphere adiabatically, hence melting sea ice through regulating atmospheric downward radiation over sea ice. This mechanism was deduced from a series of experiments in Ding et al. (2017), which show that the JJA atmospheric response to simultaneous sea ice melting is very weak in the Arctic. However, when the model is nudged to observed tropospheric winds, observed changes in JJA temperature, humidity, radiation and sea ice fields can partially be induced, indicating a circulation-driven process. Thus, the causality of the observed atmosphere–sea ice connections in Fig. 1 should be interpreted as a forcing of the JJA atmosphere on September sea ice melting rather than a surface-driven response.

Correlations of the September SIA index with JJA geopotential height, temperature and specific humidity have nearly identical spatial distributions over the Arctic Ocean and Greenland (Fig. 1a–c), indicating that large scale circulation plays a key role in linking the temperature and humidity fields together to influence sea ice as suggested by Ding et al. (2017, 2019). To illustrate the role of the large-scale circulation in coupling all these processes, atmospheric variables over regions with significant correlation in Fig. 1d–f are averaged to better reflect the air–sea ice connection. Thus we construct atmospheric indices as averaged zonal mean upper level (400–200 hPa) geopotential height (Fig. 1d red box), tropospheric air temperature (1000–300 hPa, Fig. 1e red box), specific humidity (1000–200 hPa, Fig. 1f red box), and averaged surface DLR north of 70° N. We examine relationships of detrended Arctic averaged indices of JJA temperature, JJA specific humidity and September SIA with JJA upper level (200–400 hPa) winds and geopotential height (Fig. 2). The results suggest that increased temperature, specific humidity and sea ice reduction are all associated with a similar circulation pattern characterized by a geopotential height increase and anticyclonic winds over the Arctic Ocean and Greenland (Fig. 2a, c). However, specific humidity related winds show a slightly weaker pattern. This is likely because moisture variability is determined by a large number of processes, such as evaporation, precipitation, moisture transport and water storage in the air and clouds, some of which are not intimately controlled by large scale circulation in summer.

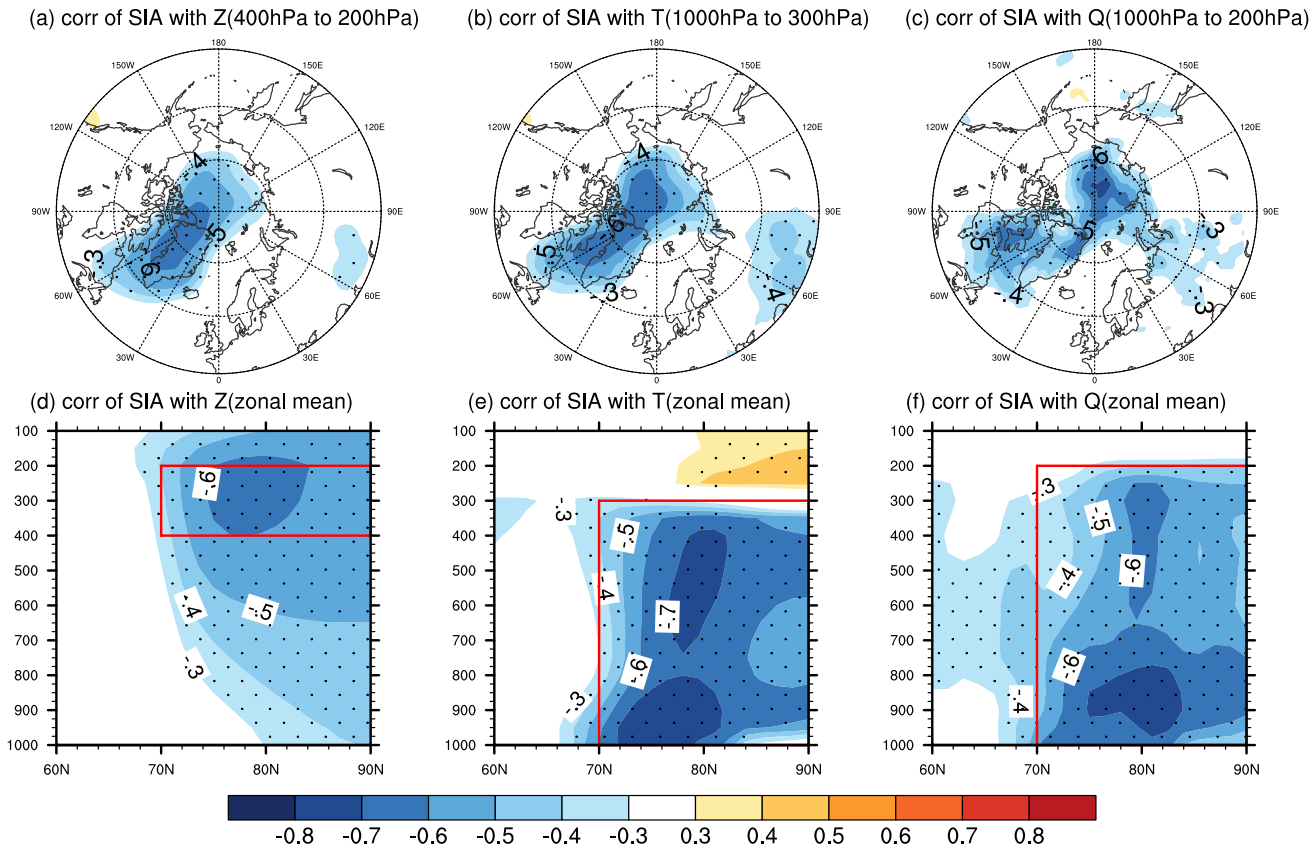

Fig. 1 Correlations of detrended NSIDC September sea ice area (SIA) index with ERA5 JJA detrended **a** geopotential height (400–200 hPa), **b** air temperature (1000–300 hPa), **c** specific humidity (1000–200 hPa) and zonal mean detrended **d** geopotential height, **e** air temperature, **f** specific humidity over the Arctic region during the period of 1979–2018. Stippling indicates statistical significance at the 95% confidence level based on a two-tailed Student's *t*-test considering the effective sample size. Red boxes in **d**–**f** indicate key atmosphere–sea ice coupling regions

3.2 The importance of downwelling longwave radiation (DLR) in determining the net heat flux at the surface

It is known that sea ice is the most sensitive to the net heat flux at the surface. Therefore, we decompose surface net heat flux into its primary components to examine the processes by which circulation forcing is translated into sea ice changes. A high correlation (− 0.75) is observed between the September SIA index and JJA Q_{net} index (averaged over north of 70° N with land masked in ERA5) on interannual time scales over the past four decades. Q_{net} is calculated as:

$$Q_{net} = DLR - ULR + DSR - USR + SH(\downarrow) + LH(\downarrow) \quad (2)$$

where DLR is downwelling longwave radiation, ULR is upwelling longwave radiation, DSR is downwelling shortwave radiation, USR is upwelling shortwave radiation, SH(↓) is sensible heat flux (downward) and LH(↓) is latent heat flux (downward). Q_{net} is positive downward and represents heat transferred from the atmosphere to the surface. To examine which components in Eq. (2) dominate the net heat flux variability, correlations are calculated between indices of JJA Q_{net} and the individual components over the past 40 years with linear trends removed from all indices averaged over the Arctic (north of 70° N) with land masked. These correlations represent the degree to which changes in Q_{net} can be explained by each component. The same calculation is also performed on CMIP5 simulations to examine the extent to which each radiative flux contributes to Q_{net} in the models.

As shown in Fig. 3, DLR and USR are two major contributors to Q_{net} in ERA5 on interannual time scales given their high correlations with Q_{net} and larger magnitude of variances (standard deviations of each flux: DLR:2.1 W/m**2, ULR:0.8 W/m**2, DSR:3.2 W/m**2, USR:3.3 W/m**2, SH:0.5 W/m**2, LH:1.0 W/m**2). This suggests that high JJA Q_{net} in the Arctic, that favors more sea ice melting, primarily results from more DLR discharged by the overlying atmosphere and less USR reflected by the darker surface. In most of the CMIP5 models, the relative importance of each component in explaining Q_{net} can be reasonably captured,

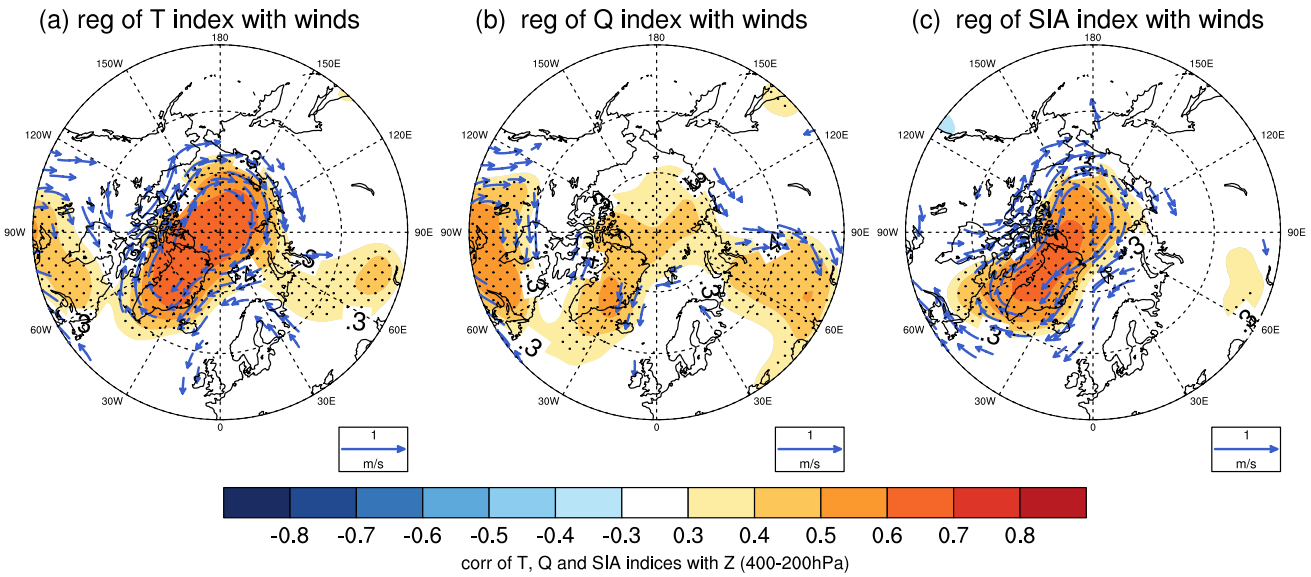

Fig. 2 Correlations of JJA detrended **a** air temperature index (averaged over 70° N–90° N, 0-360o, 1000–300 hPa, denoted as red box in Fig. 1e); **b** specific humidity index (averaged over 70° N–90° N, 0°–360°, 1000–200 hPa, denoted as red box in Fig. 1f); **c** September SIA index (averaged over 70° N–90° N, 0–360°, sign reversed) with JJA detrended high-level (400–200 hPa) geopotential height (shading). Regressions of JJA detrended **a** air temperature index; **b** specific humidity index and **c** September SIA index (sign reversed) with JJA detrended high-level (400–200 hPa) winds are superposed (vectors). All indices and winds fields are standardized before calculating regressions to show wind changes associated with one standard deviation change of each index. Vectors are displayed if either component is significant above the 95% confidence level. Stippling indicates statistical significance at the 95% confidence level

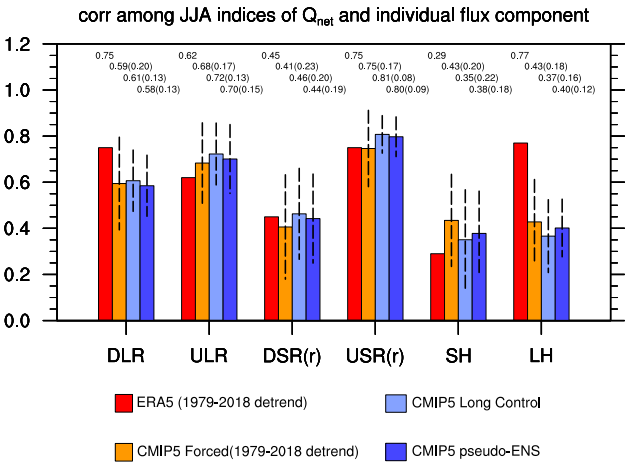

Fig. 3 Correlation coefficients among indices (all averaged over 70° N–90° N, 0–360° with land masked) of JJA Qnet index (Qnet = DLR-ULR + DSR-USR + RH + LH) and JJA DLR, ULR, DSR, USL, RH and LH in ERA5(1979–2018, detrended), 30-CMIP5-model mean of detrended forced runs of 1979–2018 (historical of 1979–2005 and RCP8.5 of 2006–2018, see Table 2), 30-CMIP5-model mean of the PI control runs using the whole period of the long control runs and the pseudo-ensemble method based on an effective sample size of 40-year (Table 1). The coefficient signs of corr(Qnet,DSR) and corr(Qnet,USR) are reversed for easier comparison. One standard deviation of all model correlations is calculated to denote CMIP5 models' spread and are displayed as dashed lines with their corresponding values given in parentheses

except that ULR's contribution is overestimated in the simulations, due to warmer ice-free ocean across models. Meanwhile, the role of DLR and LH in determining Q_{net} has been underestimated relative to ERA5. Given that DLR has a stronger connection with Q_{net} and larger variability, we thus focus on the bias of DLR in the following analysis to understand why CMIP5 models have a limitation to capture the significance of DLR in contributing Q_{net}. Although models also generally underestimate the correlation of LH with Q_{net}, we don't focus on the cause of this bias here since LH is very small compared with other fluxes and determined by complex interactions at the interface between the atmosphere and the ocean-ice mixed surface that current climate models may not be able to fully capture (Persson et al. 2002; Westermann et al. 2009).

3.3 Role of DLR in linking the atmosphere and sea ice

The importance of DLR in linking large scale atmospheric variables with sea ice has also been raised in previous studies (Kapsch et al. 2016; Ding et al. 2017; Lee et al. 2017; Wernli and Papritz 2018). In Fig. 4, we correlate JJA Arctic averaged temperature, specific humidity and September SIA indices with JJA DLR at the surface. Given that DLR is highly coupled with temperature and humidity, the strong correlation magnitude in Fig. 4a, b is expected. While the

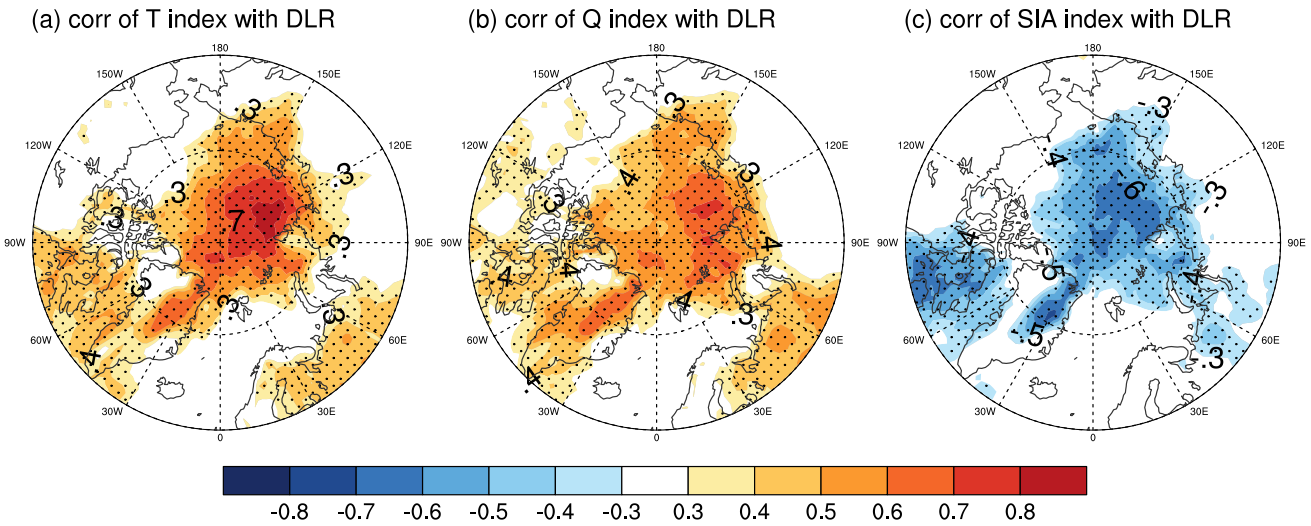

Fig. 4 Correlations of JJA detrended **a** air temperature index; **b** specific humidity index; **c** September SIA index with JJA detrended downward longwave radiation (DLR) of 1979–2018. Stippling indicates statistical significance at the 95% confidence level

spatial distributions all show a similar longwave radiation pattern throughout the Arctic, with a maximum mostly in the Laptev Sea and central Arctic Basin (Fig. 4a–c), indicating DLR's critical role in linking temperature and humidity with sea ice variability. Statistically, correlations of the September SIA index with the JJA temperature and humidity indices are − 0.73 (r(SIA-T)) and − 0.72 (r(SIA-Q), Fig. 5), respectively. The corresponding partial correlations with the effects of longwave radiation removed are − 0.37 ($r_{partial}$(SIA-T, $DLR_{removed}$) and − 0.31 ($r_{partial}$(SIA-Q, $DLR_{removed}$), which are only marginally significant at the 95% confidence level. Namely, DLR, appearing as the essential factor linking the atmosphere and sea ice, is jointly driven by circulation, temperature and specific humidity above the sea ice in the reanalysis.

3.4 Simulated atmosphere–sea ice connections in CMIP5

To quantify the strength of the observed relationship, we use simple correlations between the September SIA index and JJA atmospheric indices on interannual time scales to represent coupling strengths in ERA5 and across different models. Considering that these correlations have been found

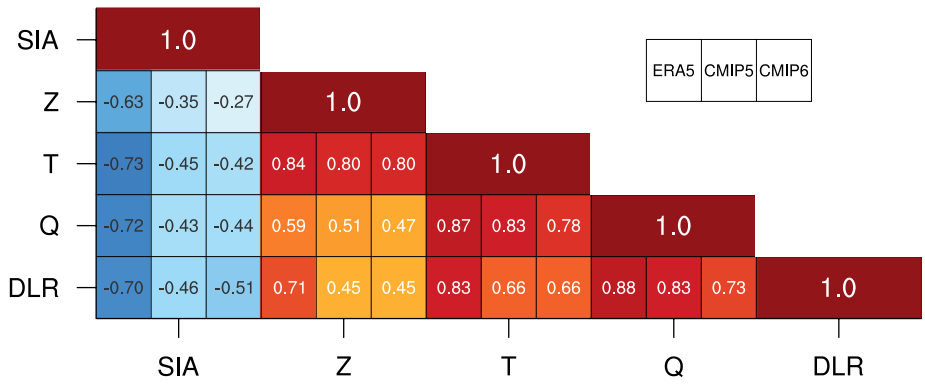

Fig. 5 Correlation coefficients among indices of September sea ice area (SIA, averaged over 0° N–90° N, 0–360°), JJA geopotential height (Z, averaged over 70° N–90° N, 0–360°, 400–200 hPa, denoted as red box in Fig. 1d), air temperature (T, averaged over 70° N–90° N, 0–360°, 1000–300 hPa, denoted as red box in Fig. 1e), specific humidity (Q, averaged over 70° N–90° N, 0–360°, 1000–200 hPa, denoted as red box in Fig. 1f) and surface downward longwave radiation (DLR, averaged over 70oN-90oN, 0-360o) from ERA5 (1979–2018, detrended), 30-CMIP5-model mean and 15-CMIP6-model mean of the PI control runs. Correlations are calculated for each model using the pseudo-ensemble method based on running 40-year periods, representative of each model's ability to capture the observed air–sea ice connection. Shown here are the model-ensemble averages of the 30 CMIP5 and 15 CMIP6 models calculated as the general coupling strengths in CMIP5 and CMIP6, respectively

to indicate strong physical interactions, models with higher credibility in simulating sea ice variability should reasonably capture these observed connection features.

Figure 5 displays air–sea ice relationships in ERA5, as well as the 30-CMIP5 and 15-CMIP6 model means based on their PI runs using a pseudo-ensemble approach. Observed detrended JJA T, Q and DLR indices are correlated with each other at $r \approx -0.8$, reflecting strong coupling processes in the atmosphere. Meanwhile, the correlation between detrended JJA DLR and September SIA is − 0.70, which suggests longwave variability can explain as much as ~ 50% of observed sea ice variability over the past 40 years. However, models are weaker in their ability to replicate these processes with respect to the coupling strength among T, Q and DLR. Compared with their observed counterpart (~ 50%,), about 20% (r(SIA-DLR) = − 0.46) of September sea ice variability is associated with JJA DLR in CMIP5 simulations.

In terms of spatial distribution (Fig. 6), the simulated connection between total September SIA and JJA geopotential height, temperature and specific humidity primarily occurs in the interior of the polar region (Fig. 6a–c), reflecting a locally confined coupling mechanism in CMIP5 models. In contrast, observed air–sea ice coupling processes occupy a broad region extending from the Arctic Ocean to Greenland (Fig. 1a–c) that is possibly due to a Pacific–Arctic teleconnection propagating from the east-central tropical Pacific and tropical Atlantic toward Greenland (Meehl et al. 2018; Baxter et al. 2019; Luo et al. 2019). Despite differences in the horizontal patterns, models do a reasonable job in replicating some basic features of the zonal mean pattern that are identified in the zonal-vertical transects (Fig. 1d–f), however the correlation values are notably lower. This indicates that CMIP5 models are able to capture the adiabatic warming process associated with the local barotropic high pressure in the interior of the Arctic. As such, we will mainly use the correlation of the zonal mean component of atmospheric variables with SIA to evaluate each CMIP5 model's performance in the following parts of this study.

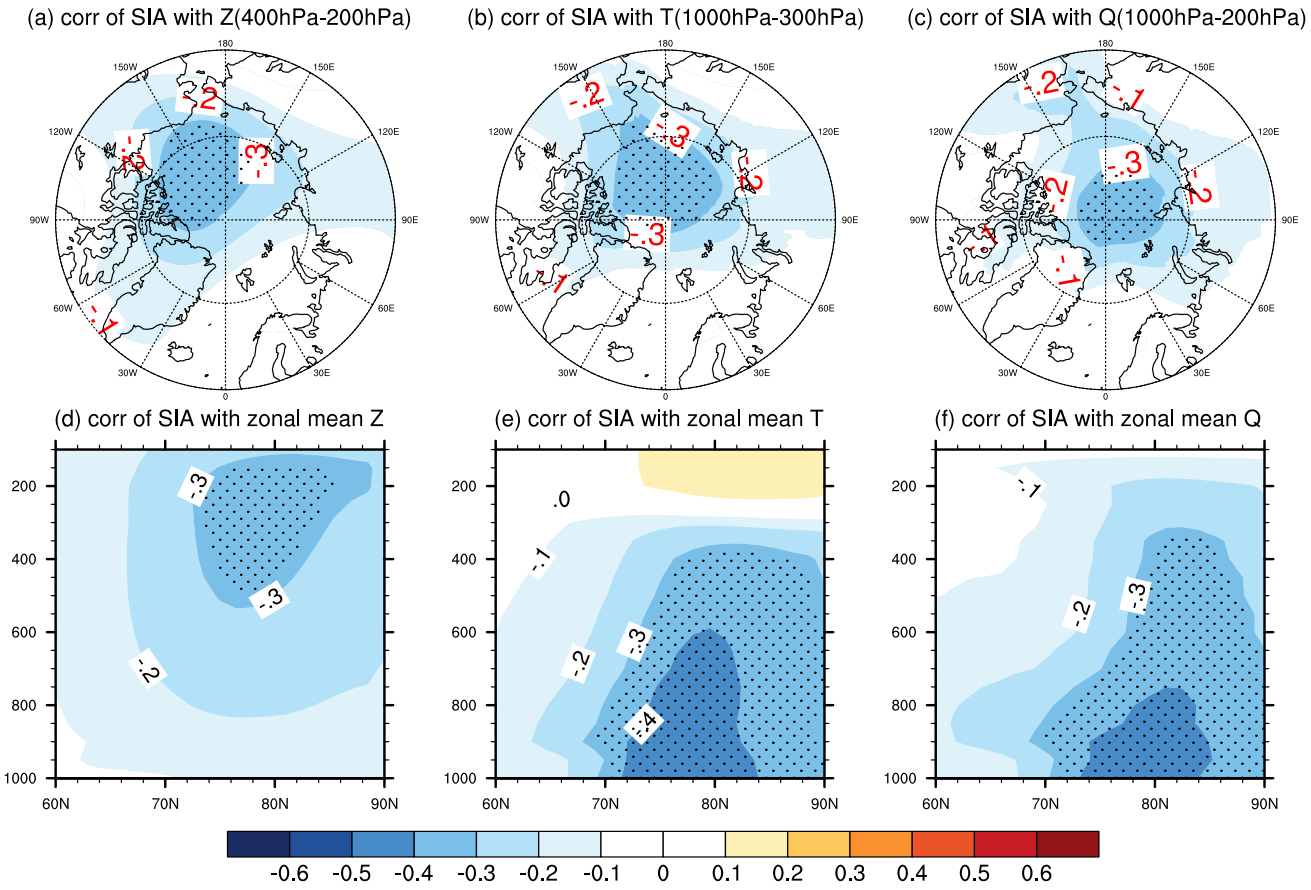

Fig. 6 Simulated correlations of total September SIA and JJA **a** geopotential height (400–200 hPa), **b** air temperature (1000–300 hPa), **c** specific humidity (1000–200 hPa) and zonal mean **d** geopotential height, **e** air temperature, **f** specific humidity of 30-model mean of the PI control runs using the pseudo-ensemble method based on a sample size of 40-year. Stippling indicates statistical significance at the 95% confidence level considering the effective sample size. Note the different color bar limits compared with Fig. 1 due to simulated weak connections

4 Process-oriented metric

4.1 Definition of the metric

In the large-scale atmosphere–sea ice connection, observed temperature and specific humidity play an important role in melting sea ice by modulating DLR. Therefore, we define the metric as:

$$Metric = [cor(SIA, Q) + cor(SIA, T)]/2 \quad (3)$$

where JJA temperature (T), specific humidity (Q) indices and the September SIA index have been previously defined in Sects. 2.1 and 3.1, respectively. Figure 7 shows the ranking of the 30 CMIP5 models based on the pseudo-ensemble method and all reanalyses based on the metric. According to this metric, a model with higher (lower) ranking will be considered as a strong (weak) air–sea ice coupling model.

4.2 CMIP5 model ranking based on the metric

Figure 7 displays metric rankings based on 30 CMIP5 PI control runs (bars) and their externally Forced counterparts (only 27 models are available, red asterisks). The magnitudes of the metric calculated using the CMIP5 Forced runs (1979–2018) are generally within one standard deviation range of those scores derived from their respective PI control runs based on the pseudo-ensemble method. However, there are models such as GISS-E2-H-CC and inm-cm4 that exhibit large disparities in the values between the Forced- and PI-derived metric scores. In these two cases, when using the short integration length of the 40-year Forced period, the metric scores seem to be more variable.

Furthermore, it is important to note that the average of the 27-model metric scores from the PI control runs based on the pseudo-ensemble method (r = − 0.44, long control run r = − 0.46; note that the average of the 30-model metric scores is also − 0.44) is indistinguishable from that of the Forced runs (r = − 0.43, detrended), indicating that the evaluation of air–sea ice relationship in CMIP5 is generally unaffected by external forcing from the perspective of the ensemble mean of

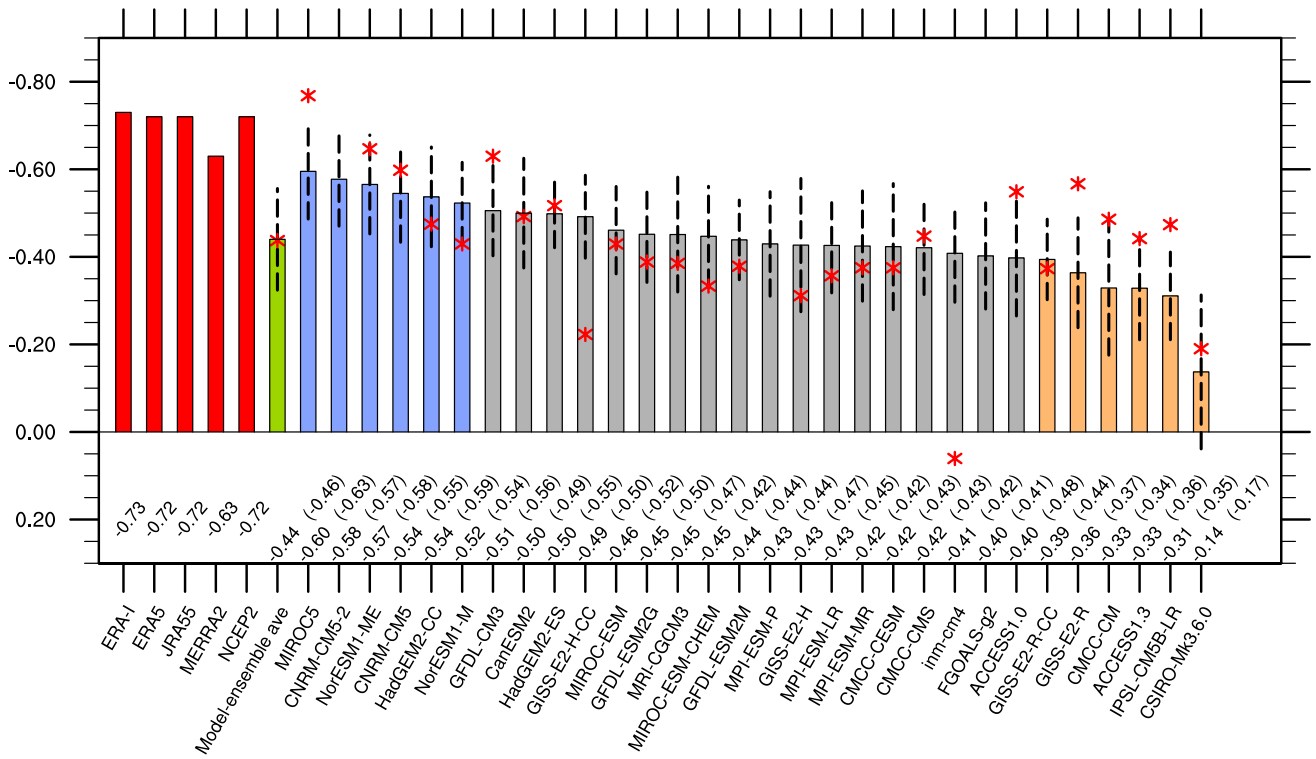

Fig. 7 Metric ranking of reanalysis, pseudo-ensemble mean of 30 CMIP5 models' PI runs based on a sample size of 40-year (bars) and 27 detrended Forced runs of 1979–2018 (red asterisks; historical of 1979–2005 and RCP8.5 of 2006–2018, only 27 models have both PI and Forced runs, see Table 2). Bars in red, green, blue and orange represent the 5 detrended reanalyses, CMIP5-ensemble average, and strong and weak coupling models, respectively. The metric is calculated using the pseudo-ensemble method based on a sample size of 40-year, generating a series of metric scores corresponding to consecutive 40-year periods in each model. One standard deviation is calculated to denote each model's metric spread and is displayed as dashed lines. Metric scores of the pseudo-ensemble means are given above the X-axis. In parentheses are metric scores calculated using the whole period of each long control run for reference

multiple models. Thus, to calculate each individual model's metric and ranking, the long PI control run is recommended to ensure a more stable and reliable evaluation.

4.3 Physical implications of the metric

Evidence has suggested that MSE, defined as the sum of internal energy, latent heat and potential energy (Eq. 4), is an important cause of atmospheric temperature fluctuation that directly control sea ice variability (Olonscheck et al. 2019).

$$MSE = Cp * T + g * z + Lv * q \quad (4)$$

where the product of specific heat capacity at constant pressure (Cp) and absolute air temperature (T) represents the heat content of moist air, the product of the gravitational constant (g) and the height above the surface (z) is gravitational potential energy, and the product of the latent heat of vaporization (Lv) and water vapor specific humidity (q) indicates latent energy associated with water vapor.

Each term in the MSE equation reflects one component of the atmospheric coupling process displayed in Fig. 1. To understand the relationship between MSE with sea ice, a JJA Arctic mean MSE index is constructed using detrended Z, T and Q from 1000 to 200 hPa within the Arctic (north of 70° N). The correlation of JJA MSE with the September SIA index is − 0.73, which is very similar to the ERA5 metric score (− 0.72) based on Eq. (3). Thus, our metric can be considered as a quantitative approach to examining how a model replicates the observed relationship between JJA MSE and September sea ice. To further support this, we use corr(September SIA,JJA MSE) as a metric to rank the CMIP5 PI control runs, using the same methodology as the evaluation carried out in Fig. 7 and obtain an identical ranking as those using Eq. (3). The consistency indicates that the thermodynamic state of the JJA atmosphere represented by the magnitude of the MSE serves as one important source to influence summertime sea ice as suggested by Olonscheck et al. (2019). Our metric, though defined as simple correlations, not only represents the large-scale coupling process between temperature, humidity and sea ice, but also reflects the state of the total MSE in the atmosphere that could be partially released to melt sea ice in summer. To maintain a reasonable degree of simplicity in our evaluation, we use the metric in Eq. (3) to evaluate the air–sea ice connection across CMIP5/6 models.

5 Effectiveness of the metric in distinguishing models' skill

5.1 Strong and weak coupling groups in CMIP5

The metrics derived from 5 reanalysis datasets show very similar strengths of around -0.72, with MERRA2 yielding the lowest score (− 0.63). In contrast, CMIP5 models show a very diversified performance in replicating this coupling strength ranging from − 0.6 to − 0.14. To better characterize this difference, first and last 20th percentiles of the 30 models are selected as strong (blue bars in Fig. 7) and weak (orange bars in Fig. 7) coupling groups, respectively, based on the metric ranking. Compared with the 30-model mean of − 0.44 (green bar in Fig. 7), the metric score of the top 6 ranked models is relatively stronger, but in the 6 weakest coupled models, the summertime atmospheric linkage with September sea ice is considerably underestimated and is almost absent in the weakest coupled model of CSIRO-Mk3.6.0 (− 0.14).

The barotropic structure of high pressure in the strong coupling group (Fig. 8a) shows similar features to what is observed in ERA5 (Fig. 1d). The greatest connection occurs in the upper troposphere between 400 and 200 hPa but is shifted slightly southward relative to the reanalysis. In the lower troposphere, a warm and humid JJA atmosphere is well correlated with total September SIA at ~ − 0.5 (Fig. 8b, c), suggesting strong air–sea ice interactions in these models. However, it is difficult to detect comparable coupling strength in the 6 lower ranking models (Fig. 8d–f). Anomalous high pressure in the deep Arctic is accompanied by weak increases of temperature and moisture above the sea ice at $r \approx -0.3$.

Analyses using historical and RCP8.5 simulations from 1979 to 2018 are also conducted to assess CMIP5 model performance under anthropogenic forcing. We find the contrast between strong and weak coupling groups (Fig. 9) is similar with that in the PI runs (Fig. 8), suggesting that CMIP5 models can generally capture the vertical connection patterns found in ERA5, but there are still deficiencies in realistically simulating the JJA atmospheric connection with September sea ice change with a wide spread amongst these models, regardless of anthropogenic forcing.

5.2 Improvement in CMIP6?

The new generation of climate models available in CMIP6 offers an opportunity to explore whether updated versions of these models are more capable of replicating the observed atmosphere–sea ice connection considering that improvements in the representation of sea ice volume and

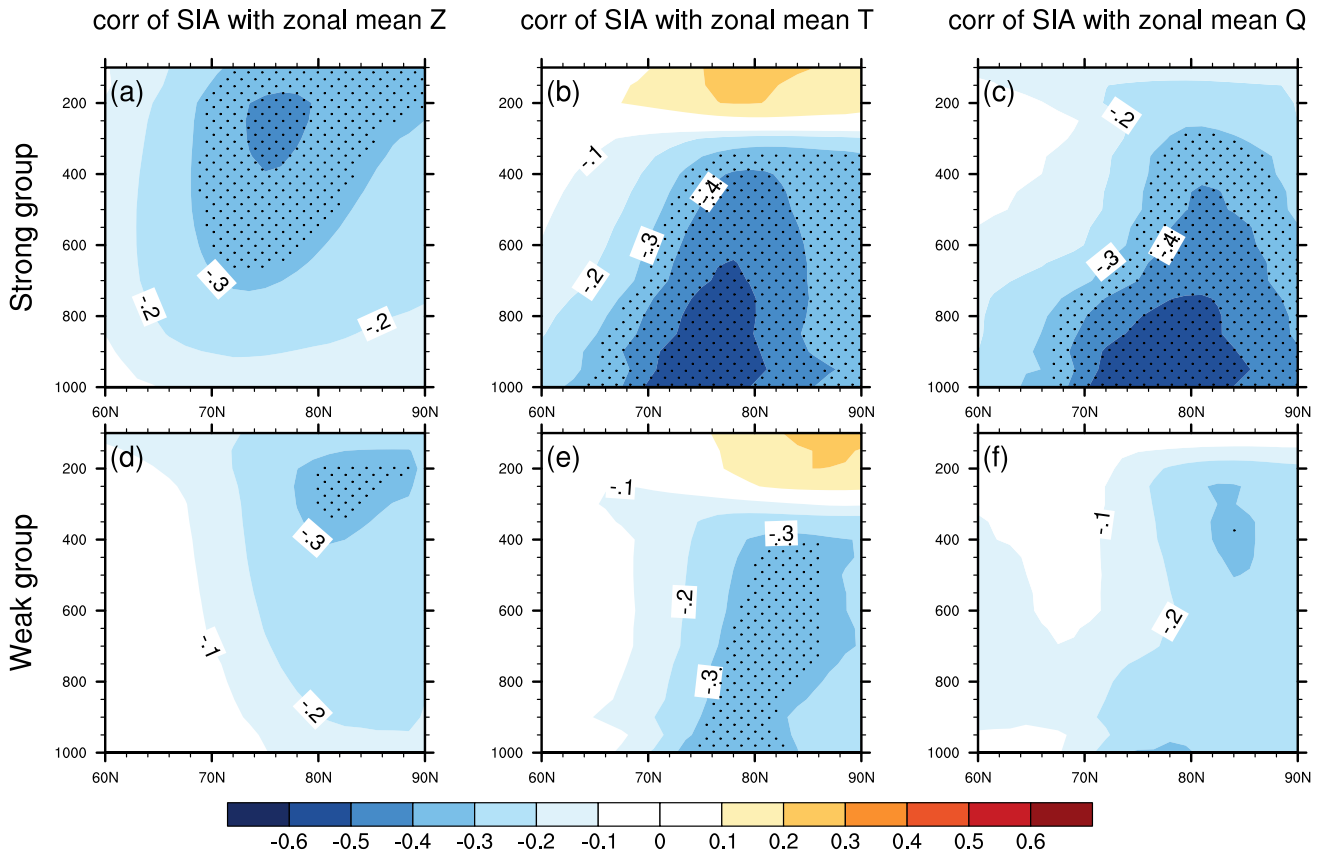

Fig. 8 Simulated correlations of total September SIA and JJA zonal mean geopotential height (left column), air temperature (middle column), specific humidity (right column) of 6 strong (blue bars in Fig. 7) and 6 weak (orange bars in Fig. 7) couplings models' mean from the PI control runs in CMIP5 using the pseudo-ensemble method based on a sample size of 40-year. Stippling indicates statistical significance at the 95% confidence level

extent have been applied on their physical representation of coupling processes in the Arctic (Davy and Outten 2019; SIMIP Community 2020). To do this, PI control runs from 15 CMIP6 models are analyzed in this study. Due to the limited number of CMIP6 models that have all atmospheric and sea ice data available, we only analyze the average of these models' skill. Air–sea ice connections in CMIP6 (Fig. 10a–c) bear a close resemblance to that in CMIP5 (Fig. 6a–c) both in spatial distribution and coupling strength. In particular, metric scores of MIROC6 (− 0.45) and CanESM5 (− 0.48, Fig. 11) in CMIP6 reflect an even weaker air–sea ice connection than their CMIP5 counterparts of MIROC5 (− 0.6) and CanESM2 (− 0.5, Fig. 7). The small size of currently available models in CMIP6 hinders us from making a more thorough comparison between CMIP5 and CMIP6. Based on the metric ranking of the 15 CMIP6 models, it seems that the models as a group have no significant improvement in their ability to simulate the observed atmosphere–sea ice relationship. Continuous model output releases will help us more thoroughly evaluate the performance of CMIP6 in the future.

6 Possible sources of model biases in simulating the atmosphere–sea ice coupling

6.1 A weak simulated forcing of circulation in the Arctic

The evaluation of the air–sea ice connection in CMIP5 has demonstrated model limitations (Fig. 6), but sources of these limitations remain unknown. Given that the observed relationship is primarily driven by large-scale circulation variability, we first examine whether CMIP5 models can exhibit a similar feature in the Arctic.

Similar to the calculation performed in Fig. 2, simulated JJA Arctic-averaged temperature and humidity and September SIA in PI control runs are correlated with JJA winds and geopotential height, respectively. The related winds and geopotential heights exhibit a weak signal in the high latitudes (Fig. 12a–c). These patterns differ from the observed counterparts in both the magnitude and spatial distribution (Fig. 2), suggesting that simulated atmospheric temperature

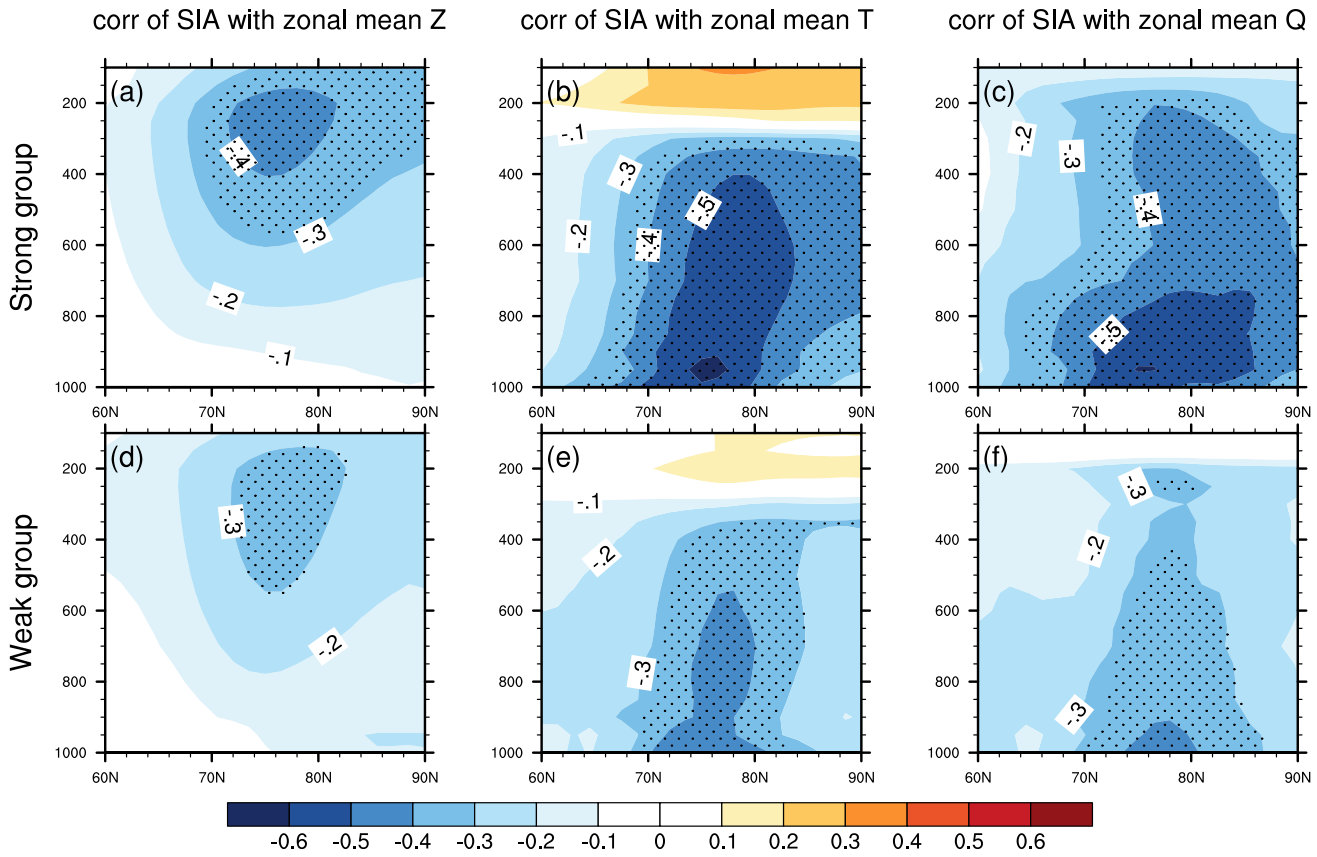

Fig. 9 Simulated correlations of detrended September SIA index and JJA detrended zonal mean geopotential height (left column), air temperature (middle column), specific humidity (right column) of 5 strong (blue bars in Fig. 7) and 6 weak (orange bars in Fig. 7) coupling models' means from the CMIP5 historical and RCP8.5 simulations of 1979–2018. Note that one of the 6 strong coupling models (CNRM-CM5) does not have an available Forced run, therefore the composites of the strong group in **a**–**c** are 5-model means. Stippling indicates statistical significance at the 95% confidence level

and humidity fields are less responsive to advection and adiabatic heating or cooling processes associated with circulation variability in the Arctic. Thus, it is not surprising to observe a relatively weak coupling between sea ice and large-scale circulation variability in CMIP5 models. The weak sensitivity of temperature and humidity to large-scale wind forcing in the Arctic over a year-to-year time scale could partially be responsible for model bias in replicating the sea ice-atmosphere coupling from reanalysis.

6.2 Important role of DLR in linking atmosphere and sea ice in models

To further understand how other linkages in the air–sea ice process differentiate the performance of CMIP5 models in reproducing the observed coupling, we examine relationships between DLR with temperature, specific humidity and sea ice in the ensemble mean, the best and worst groups on the basis of the metric. We speculate that the important linkages, which really determine model skill, should show a strong difference in their performance between the best and worst groups. However, we do not observe such contrast in the T-DLR relationships between the two groups (Fig. 13c, e). The Q-DLR connections exhibit a weak contrast in their magnitudes but share a very similar spatial pattern between the two groups (Fig. 13d, f). This suggests that no matter how a model replicates the observed sea ice-atmosphere connection, most models have a similar sensitivity of DLR to the temperature and humidity of the overlying atmosphere, which is likely because there is little diversity amongst radiative transfer models employed in these models (Mlawer et al. 1997).

A scatter plot of SIA-T correlations (X-axis) and SIA-Q correlations (Y-axis) shows a linear distribution across 30 CMIP5 models (gray diamonds in Fig. 13g), reflecting a close tie between effects of temperature and moisture fields of a model in regulating sea ice. The strength of SIA-T connection in a model is generally proportional to its strength in SIA-Q connection. By applying a partial correlation analysis to remove the effect of DLR from SIA-T and SIA-Q relationships (gray hollow circles in Fig. 13g), we find all these strong relationships drop to insignificant levels. Thus,

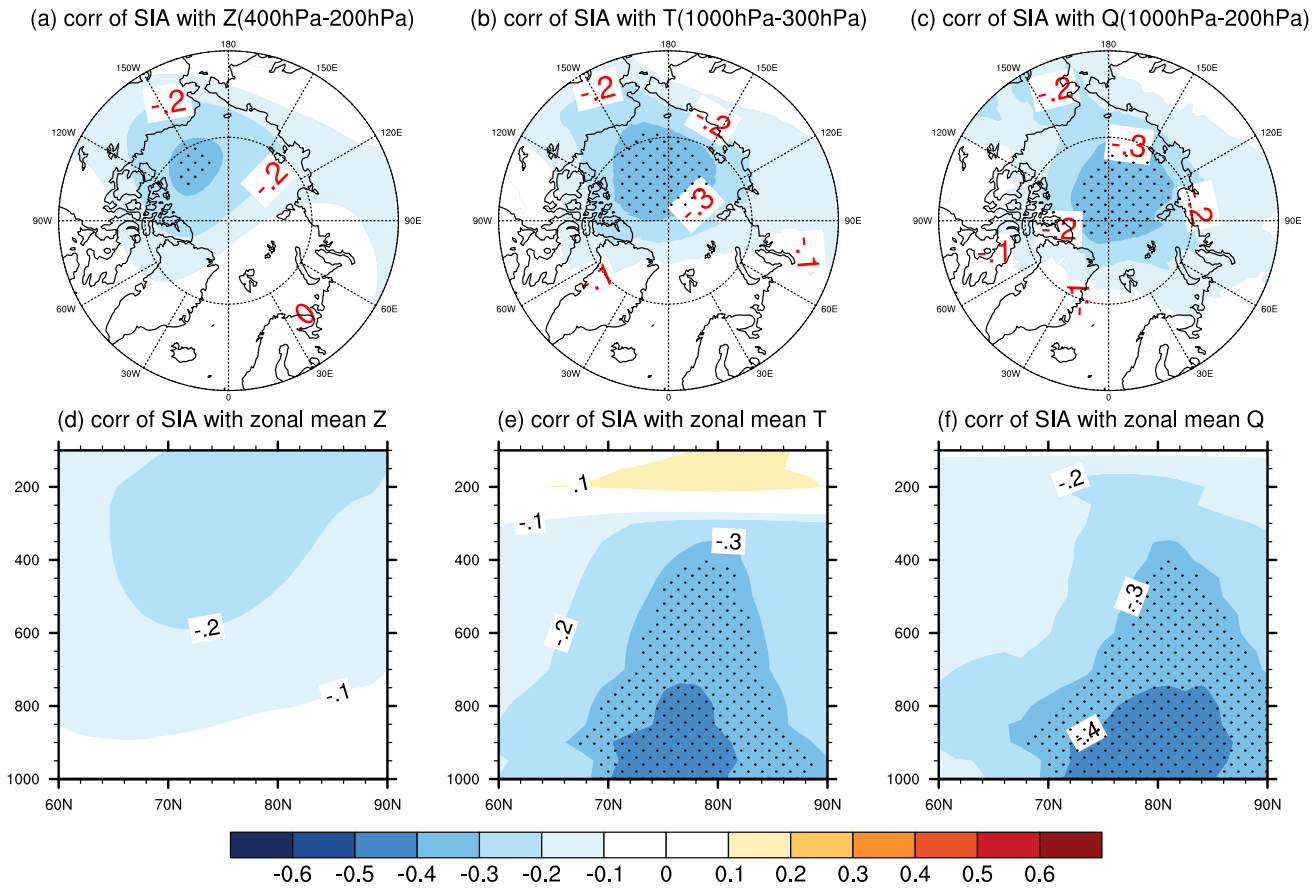

Fig. 10 Same as Fig. 6 but for 15-model mean of the CMIP6 PI control runs using the pseudo-ensemble method based on a sample size of 40-year. Stippling indicates statistical significance at the 95% confidence level

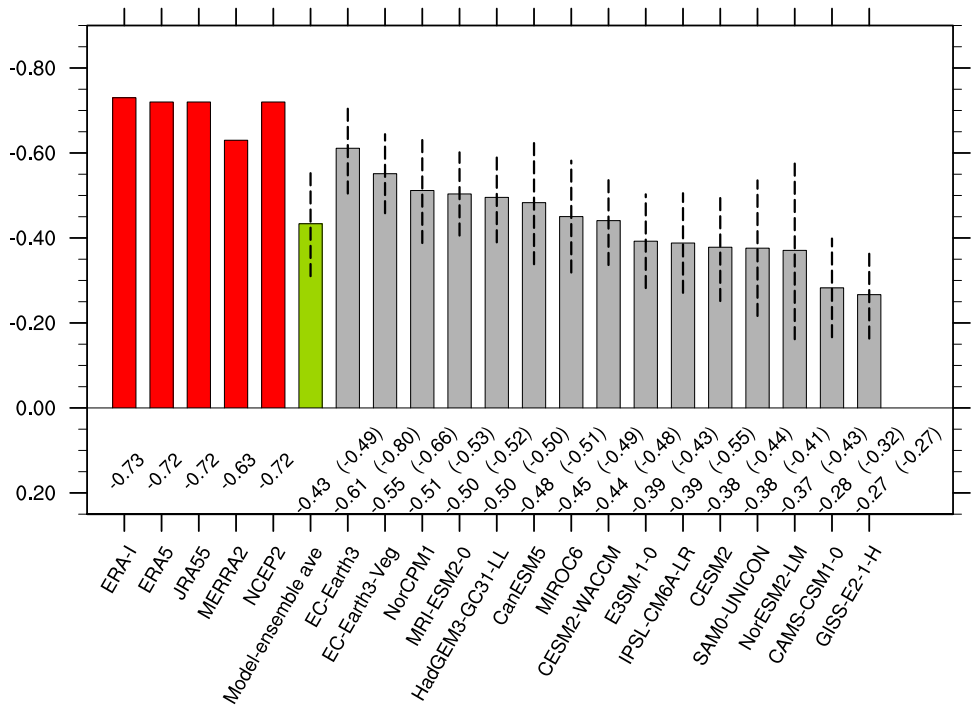

Fig. 11 Metric ranking of 5 detrended reanalyses (red bars), PI control runs of 15 CMIP6 models (gray bars) and their pseudo-ensemble mean (green bar). The metric is calculated using a pseudo-ensemble method based on a sample size of 40-year, generating a series of metric scores corresponding to consecutive 40-year periods in each model. One standard deviation is calculated to denote each model's metric spread and displayed as dashed lines. Metric scores of the pseudo-ensemble mean are given above the X-axis. In parentheses are metric scores calculated using the whole period of the long control runs for reference

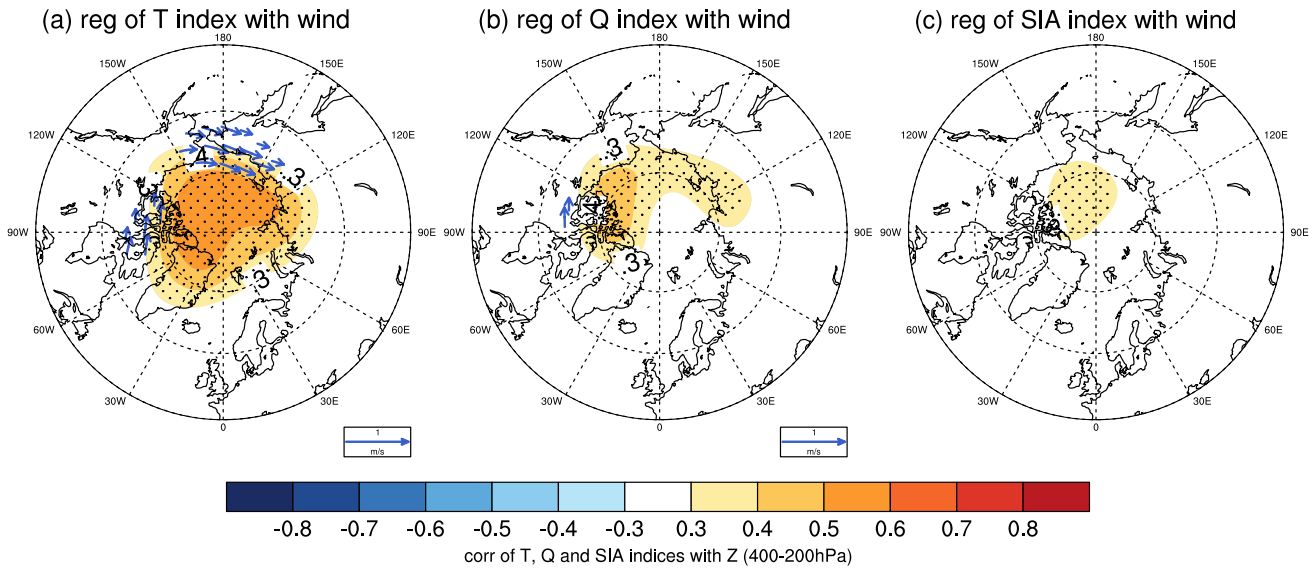

Fig. 12 Simulated correlations of JJA **a** air temperature index; **b** specific humidity index; **c** September SIA index (sign reversed) with JJA high-level (400–200 hPa) geopotential height (shading) of 30-model mean of the PI control run using a pseudo-ensemble method based on a sample size of 40-year. Regressions of JJA **a** air temperature index; **b** specific humidity index and **c** September SIA index with JJA high-level (400–200 hPa) winds are superposed (vectors). These indices and winds fields are standardized before calculating regression to show wind changes associated with one standard deviation change of each index. Vectors are displayed if either component is significant above the 95% confidence level. No vectors pass the significance test in (**c**). Stippling indicates statistical significance at the 95% confidence level

in CMIP5 models, the fact still holds that DLR, which is jointly modulated by temperature and humidity, also plays an essential role to link the atmosphere and sea ice as it does in reanalyses. Since DLR in CMIP5 models have a similar sensitivity to temperature and humidity variability in both the strong and weak coupling groups, an additional factor, rather than temperature and humidity, must be playing a key role to determine how much DLR could be efficiently emitted by the atmosphere to drive sea ice change below, and thus determine the performance of each model.

Unlike the weak contrast in T-DLR and Q-DLR linkages between the two coupling groups, the best group appears to capture a stronger sea ice-DLR connection than that of the worst group (Fig. 14c, d). The sea ice-DLR coupling in the strong coupling group occurs throughout the interior of the Arctic (Fig. 14c) as it does in ERA5 (Fig. 14a), while the coupling in the worst group is limited to a small region around the Kara Sea (Fig. 14d). Given that temperature and humidity are not the key factors differentiating DLR variability in the two coupling groups, we speculate that the contrast between the sea ice-DLR connections may be attributed to two possible reasons: I) clouds could act as an efficient pathway to translate large scale DLR forcing to sea ice; II) sea ice is more sensitive to atmosphere related DLR in the strong coupling group.

Previous studies have suggested that cloud variability that is strongly subject to large-scale dynamics and cloud microphysics is a crucial factor to modulate the surface energy balance and sea ice changes (Shupe and Intrieri 2004; Kay and Gettelman 2009; Sedlar and Tjernström 2009; Cesana et al. 2012; Bennartz et al. 2013; Hofer et al. 2019). Thus, cloud changes serve as a key factor determining how large scale circulation applies its related radiative forcing on sea ice (Kay et al. 2016). However, biases of Arctic cloud properties in reanalysis data have also been identified based on comparisons with satellite observations of NASA CERES-MODIS, CloudSat-CALIPSO (Huang et al. 2017). There are also complex relationships between clouds and relative humidity, surface heat flux and atmospheric circulations (Kay et al. 2008; Kay and L'Ecuyer 2013). Given these biases, uncertainties and complexities, further comparisons and analyses regarding the interannual linkage among the atmosphere, clouds and sea ice in observations, reanalysis and CMIP5/6 models are necessary and essential.

7 Conclusion and discussion

In this study, we develop a statistical-based metric to assess the skill of 30 CMIP5 and 15 CMIP6 models in simulating the observed atmosphere–sea ice linkage under different external radiative forcing scenarios, including the pre-industrial (PI), historical and RCP8.5 (Forced) simulations. This evaluation suggests that all CMIP5 and recently available

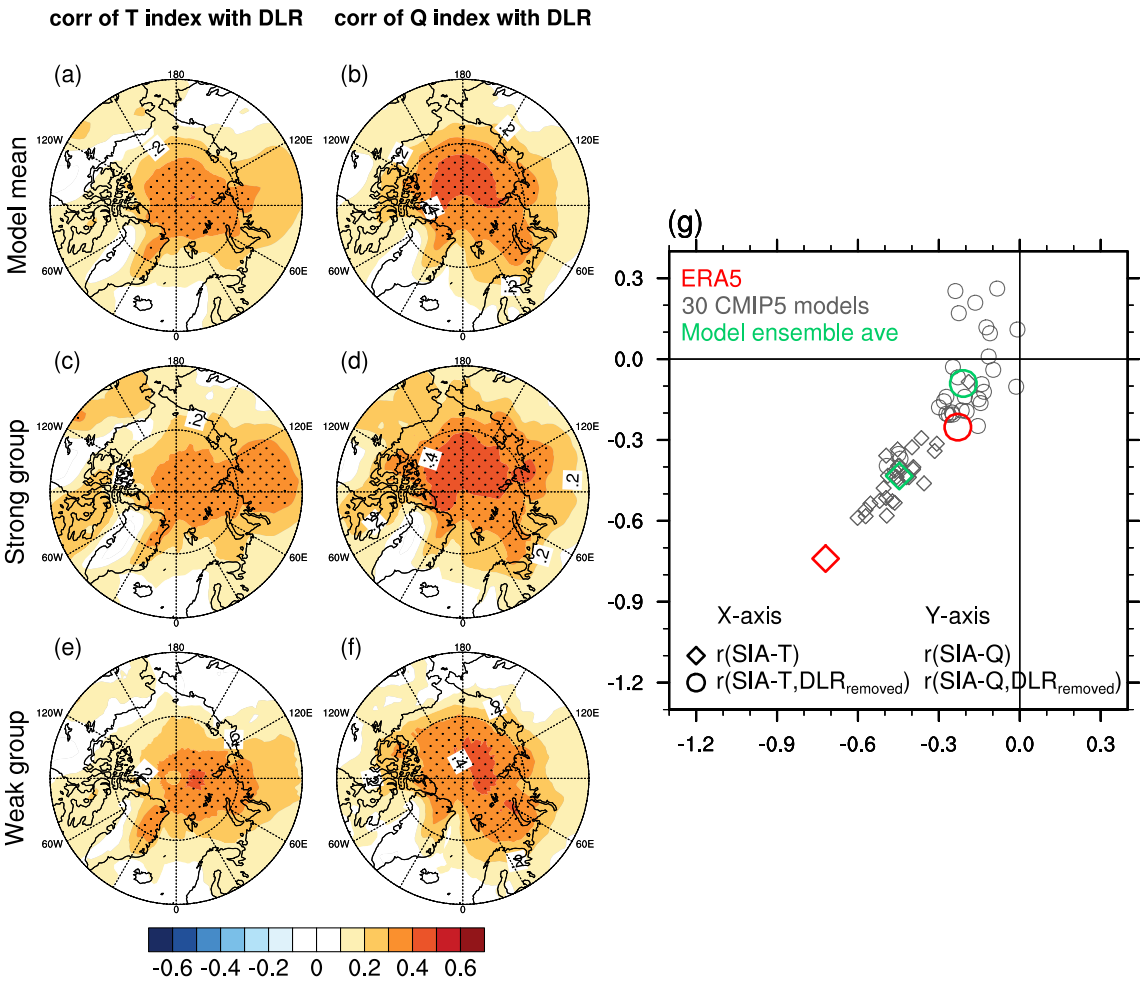

Fig. 13 Correlations of JJA air temperature index (**a**, **c**, **e**); specific humidity index (**b**, **d**, **f**) with JJA downward longwave radiation of CMIP5 30-model mean (top row), 6 strong (middle row) and 6 weak (bottom row) coupling models' mean of the PI control runs using the pseudo-ensemble method based on a sample size of 40-year. Stippling indicates statistical significance at the 95% confidence level considering the effective sample size. In **g**, correlations of the September SIA index and JJA air temperature index (X-axis, r(SIA-T)) versus correlations of the September SIA index and JJA specific humidity index (Y-axis, r(SIA-Q)). Diamonds represent values of ERA5 (red diamonds, detrended), 30 CMIP5 models (gray diamonds) and the 30-model average (green diamonds) of the PI control runs in CMIP5. Hollow circles are counterparts of correlations in diamonds when a partial correlation method is applied to remove the radiative effects of DLR (X-axis, r(SIA-T,$DLR_{removed}$); Y-axis, r(SIA-Q,$DLR_{removed}$))

CMIP6 models have a common limitation to reproduce the observed connection between JJA atmospheric circulation and September sea ice variability on interannual time scales. Models with the highest skill in replicating this relationship still have some limitations to capture the observed coupling strength revealed in reanalysis, indicating an apparent underestimation of sea ice sensitivity to atmospheric forcing across CMIP5 and CMIP6 models.

The metric we use to evaluate CMIP5 and 6 models in this study emphasizes the atmosphere–sea ice connection strength, which is simply defined as the average of two correlations (temperature with sea ice and specific humidity with sea ice) derived from detrended ERA5 reanalysis and NSIDC sea ice data. This metric can effectively rank CMIP5/6 model performance with respect to their skill in replicating the observed atmosphere–sea ice connection. In addition, a new ranking oriented by the moist static energy (MSE) perspective is found to be highly consistent with our metric. This indicates that although our metric is simply designed as an average of the two correlations, it physically represents how efficient sea ice responds to changes in MSE contained in the overlying atmosphere. Hence, model evaluation based on our simple metric is physically meaningful and supportive.

We also note that the metrics based on the PI control runs and Forced runs are similar in most models. Metric scores derived from detrended forced runs (1979–2018) generally stay within one standard deviation of the corresponding

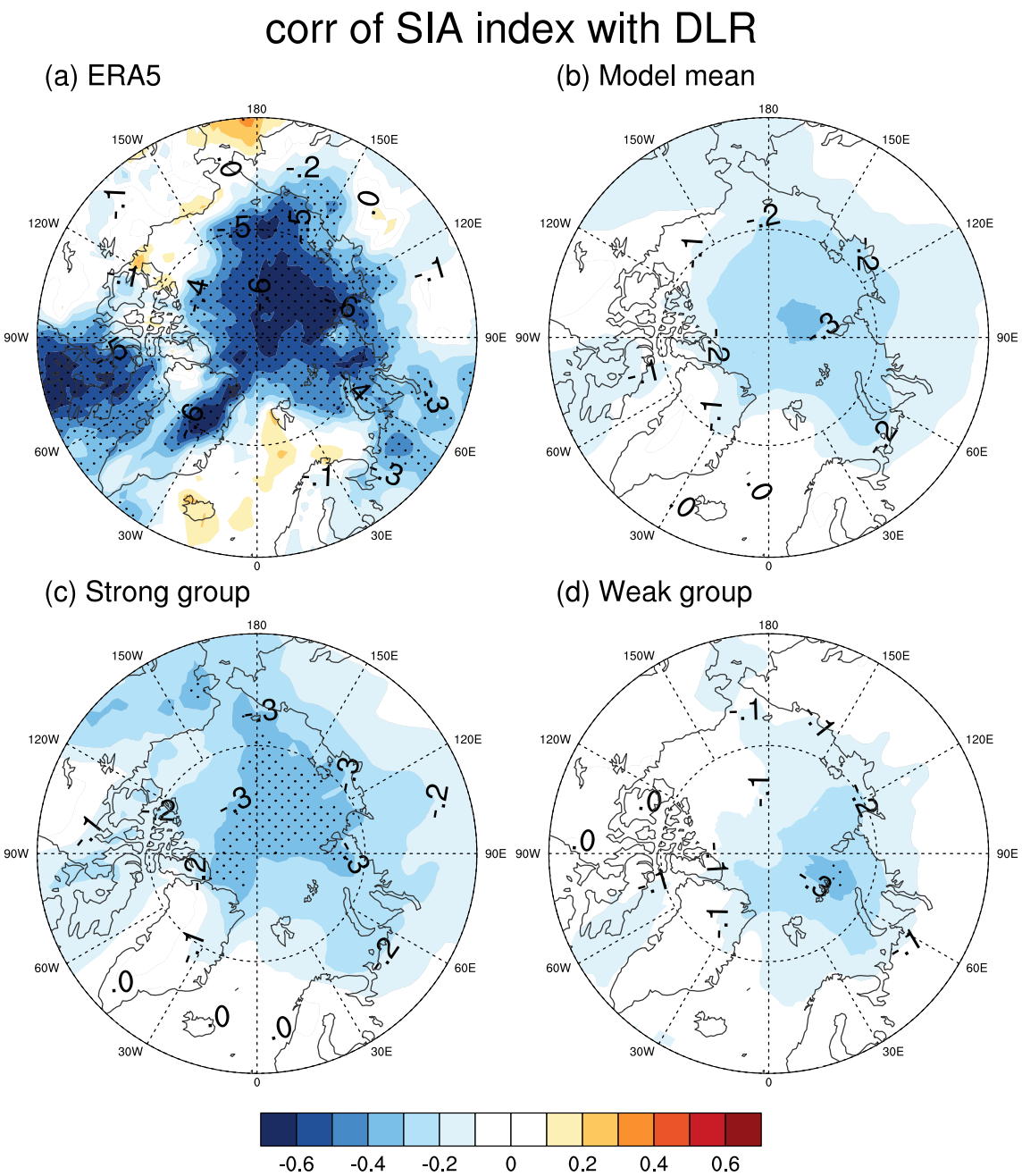

Fig. 14 Correlations of total September SIA with JJA downward longwave radiation of **a** detrended ERA5 data during the period of 1979–2018, **b** CMIP5 models' average, **c** 6 strong and **d** 6 weak coupling models' mean in the CMIP5 PI control runs using the pseudo-ensemble method based on a sample size of 40-year. Stippling indicates statistical significance at the 95% confidence level considering the effective sample size

scores derived from the pseudo-ensembles of PI control runs. Although our evaluation is not sensitive to the addition of external radiative forcing, using PI control runs for model evaluation is preferable considering that the evaluations using longer simulations are more reliable and robust.

Based on our metric, which only emphasizes the relationships between sea ice with temperature and humidity, we find that in general, the models' bias is partially attributable to the weak large-scale wind forcing in the Arctic across all models. This may indicate that simulations of the Arctic climate may not have sufficient sensitivity to internal dynamical variability, which is strongly associated with high latitude large-scale circulation changes and remote forcing originating outside of the Arctic (Baxter et al. 2019; Luo et al. 2019). Since most models show a good capability of simulating T-DLR and Q-DLR connections, the factor that

differentiates the skill of the high- and low-ranking groups is likely associated with a model's capability in reproducing the observed relationships between DLR and sea ice. In addition, we speculate that how each model simulates clouds in the Arctic may partially determine its performance in replicating the observed atmosphere–sea ice connection considering the importance of clouds in driving DLR at the surface. Large scale circulation forcing in the Arctic can not only regulate temperature and humidity fields through horizontal advection and adiabatic subsidence but also favors clouds variations in different levels, through a combination of processes, including strong subsidence (Young et al. 2018), increased moisture transport into the Arctic (Morrison et al. 2012), a strengthened near surface temperature inversion, enhanced turbulent mixing within the boundary layer and a strengthened longwave radiative cooling at the top of low-level clouds (Sedlar and Tjernström 2009; Morrison et al. 2012; Shupe 2011). How climate models capture the response of clouds to large scale circulation forcing remain an open question. In addition, many other processes, such as latent heat variability, sea ice dynamics, ocean heat content changes, synoptic scale atmospheric variability, are important to regulate sea ice conditions (Stroeve et al. 2014; Hori et al. 2015; Tokinaga et al. 2017; Lee et al. 2017; Labe et al. 2018). Thus, these processes and their interactions should be taken into account in the future to better understand models' biases.

Although further analysis is necessary to improve our understanding of the underlying mechanisms linking large-scale circulation with local temperature, humidity, clouds and sea ice, we believe our study forms a needed first step in advancing our understanding of model limitations in simulating key atmospheric processes that shape sea ice variability in summer. The evaluation of CMIP5 and 6 models performed in this study is also informative for the recent efforts and attempts of the community to improve seasonal prediction of Arctic sea ice because models used for seasonal sea ice prediction may also share the same limitations with those in CMIP5/6 in replicating the observed relationship. Therefore, models that can reasonably reproduce the observed atmosphere–sea ice connection may have higher credibility in sea ice prediction.

Acknowledgements This study was supported by Modeling, Analysis, Predictions and Projections (NA19OAR4310281) and Climate Variability & Predictability (NA18OAR4310424) programs as part of NOAA's Climate Program Office, and NSF's Polar Programs (OPP-1744598). R. L. was jointly supported by the Second Tibetan Plateau Scientific Expedition and Research (STEP) program (Grant No. 2019QZKK0102), National Natural Science Foundation of China (Grant Nos. 41790475 and 91937302) and the National Key R&D Program of China (Grant No. 2019YFC1509100), China Scholarship Council (CSC) scholarship (No.201806100093) and NA19OAR4310281. M. B. was supported by NOAA's Science Collaboration Program with funding administered by UCAR's Cooperative Programs for the Advancement of Earth System Science (CPAESS) under awards NA16NWS4620043 and NA18NWS4620043B. Y. Huang was supported by the NASA Earth and Space Science Fellowship program at the University of Arizona (80NSSC18K1339). X. Dong were supported by NASA CERES project through grant 80NSSC19K0172 at the University of Arizona. We thank Prof. Ed Blanchard for helpful discussions and comments on the manuscript

References

Baxter I et al (2019) How tropical pacific surface cooling contributed to accelerated sea ice melt from 2007 to 2012 as ice is thinned by anthropogenic forcing. J Clim 32:8583–8602. https://doi.org/10.1175/JCLI-D-18-0783.1

Bennartz R, Shupe MD, Turner DD, Walden VP, Steffen K, Cox CJ, Kulie MS, Miller NB, Pettersen C (2013) July 2012 Greenland melt extent enhanced by low-level liquid clouds. Nature 496(7443):83–86. https://doi.org/10.1038/nature12002

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

Cesana G, Kay JE, Chepfer H, English JM, de Boer G (2012) Ubiquitous low-level liquid-containing Arctic clouds: new observations and climate model constraints from CALIPSO-GOCCP. Geophys Res Lett. https://doi.org/10.1029/2012GL053385

Cavalieri D, Parkinson C, Gloersen P, Zwally HJ (1996, updated yearly) Sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I passive microwave data, Version 1. [Indicate subset used]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/8GQ8LZQVL0VL

Cohen J et al (2020) Divergent consensuses on arctic amplification influence on midlatitude severe winter weather. Nat Clim Change 10(1):20–29

Comiso JC, Parkinson CL, Gersten R, Stock L (2008) Accelerated decline in the Arctic sea ice cover. Geophys Res Lett 35:L01703. https://doi.org/10.1029/2007GL031972

Davy R, Outten S (2019) The Arctic surface climate in CMIP6: status and developments since CMIP5. J Clim. https://doi.org/10.1175/JCLI-D-19-0990.1

Dee D et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137:535–597

Ding Q et al (2017) Influence of high-latitude atmospheric circulation changes on summertime Arctic sea ice. Nat Clim Change 7:289–295. https://doi.org/10.1038/nclimate3241

Ding Q et al (2019) Fingerprints of internal drivers of Arctic sea ice loss in observations and model simulations. Nat Geosci 12:28–33. https://doi.org/10.1038/S41561-018-0256-8

Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, Taylor KE (2016) Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. Geosci Model Dev 9:1937–1958

Gelaro R, McCarty W, Suárez MJ et al (2017) The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). J Clim 30(14):5419–5454

Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A (2020) The ERA5 global reanalysis. Q J R Meteorol Soc 146(730):1999–2049

Hofer S, Tedstone AJ, Fettweis X, Bamber JL (2019) Cloud microphysics and circulation anomalies control differences in future Greenland melt. Nat Clim Change 9:523–528. https://doi.org/10.1038/s41558-019-0507-8

Hori ME, Inoue J, Kikuchi T (2015) The role of cyclone activity in the interannual variability of the summertime beaufort high. SOLA 11:104–107

Huang Y, Dong X, Xi B, Dolinar EK, Stanfield RE, Qiu S (2017) Quantifying the uncertainties of reanalyzed arctic cloud and radiation properties using satellite surface observations. J Clim 30(19):8007–8029

Kanamitsu M, Ebisuzaki W, Woolen J, Yang S-K, Hnilo JJ, Fiorino M, Potter GL (2002) NCEP-DOE AMIP-II reanalysis (R-2). Bull Am Meteorol Soc 83:1631–1643

Kapsch M-L, Graversen RG, Tjernström M, Bintanja R (2016) The effect of downwelling longwave and shortwave radiation on Arctic summer sea ice. J Clim 29:1143–1159. https://doi.org/10.1175/JCLI-D-15-0238.1

Kay JE, Gettelman A (2009) Cloud influence on and response to seasonal Arctic sea ice loss. J Geophys Res 114:D18204. https://doi.org/10.1029/2009JD011773

Kay JE, L'Ecuyer T (2013) Observational constraints on Arctic Ocean clouds and radiative fluxes during the early 21st century. J Geophys Res 118:7219–7236

Kay J, L'Ecuyer ET, Gettelman A, Stephens G, O'Dell C (2008) The contribution of cloud and radiation anomalies to the 2007 Arctic sea ice extent minimum. Geophys Res Let 35:L08503. https://doi.org/10.1029/2008GL033451

Kay J, Holland M, Jahn A (2011) Inter-annual to multi-decadal Arctic sea ice extent trends in a warming world. Geophys Res Lett 38:L15708. https://doi.org/10.1029/2011GL048008

Kay J, L'Ecuyer T, Chepfer H, Loeb N, Morrison A, Cesana G (2016) Recent advances in Arctic cloud and climate research. Curr Clim Change Rep 2:159–169. https://doi.org/10.1007/s40641-016-0051-9

Kobayashi S et al (2015) The JRA-55 reanalysis: general specifications and basic characteristics. J Meteorol Soc Jpn 93(1):5–48. https://doi.org/10.2151/jmsj.2015-001

Kwok R (2011) Observational assessment of Arctic Ocean sea ice motion, export, and thickness in CMIP3 climate simulations. J Geophys Res Oceans (1978–2012) 116(C8):C00D05

Labe Z, Peings Y, Magnusdottir G (2018) Contributions of ice thickness to the atmospheric response from projected Arctic sea ice loss. Geophys Res Lett 45(11):5635–5642

Lee HJ, MvO K, Yeh S, Kwon Y, Park W, Park J, Kim YH, Alexander MA (2017) Impact of poleward moisture transport from the North Pacific on the acceleration of sea ice loss in the Arctic since 2002. J Clim 30(17):6757–6769. https://doi.org/10.1175/JCLI-D-16-0461.1

Luo R, Wu Z, Zhang P, Dou J (2019) Potential influence of the developing La Niña on the sea–ice reduction in the Barents-Kara Sea. Atmos Ocean 57:182–194

Meehl G, Chung C, Arblaster J, Holland M, Bitz C (2018) Tropical decadal variability and the rate of Arctic sea ice decrease. Geophys Res Lett 45:11326–11333. https://doi.org/10.1029/2018GL079989

Mlawer EJ, Taubman SJ, Brown PD, Iacono MJ, Clough SA (1997) Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. J Geophys Res 102:16663–16682

Morrison H, de Boer G, Feingold G, Harrington J, Shupe M, Sulia K (2012) Resilience of persistent Arctic mixed-phase clouds. Nat Geosci 4:11–17. https://doi.org/10.1038/ngeo1332

Mouginot J, Rignot E, Bjørk AA, Van Den Broeke M, Millan R, Morlighem M, Noël B, Scheuchl B, Wood M (2019) Forty-six years of Greenland Ice Sheet mass balance from 1972 to 2018. Proc Natl Acad Sci 116(19):9239–9244

Notz D, Marotzke J (2012) Observations reveal external driver for Arctic sea-ice retreat. Geophys Res Lett 39:L08502. https://doi.org/10.1029/2012GL051094

Notz D, Stroeve J (2016) Observed Arctic sea–ice loss directly follows anthropogenic CO_2 emission. Science 354(6313):747–750. https://doi.org/10.1126/science.aag2345

Ogi M, Yamazaki K, Wallace JM (2010) Influence of winter and summer surface wind anomalies on summer Arctic sea ice extent. Geophys Res Lett 37:L07701. https://doi.org/10.1029/2009GL042356

Olonscheck D, Mauritsen T, Notz D (2019) Arctic sea-ice variability is primarily driven by atmospheric temperature fuctuations. Nat Geosci 12:430–434. https://doi.org/10.1038/s41561-019-0363-1

Persson POG, Fairall CW, Andreas EL, Guest PS, Perovich DK (2002) Measurements near the Atmospheric Surface Flux Group tower at SHEBA: near-surface conditions and surface energy budget. J Geophys Res 107(8045):1–21. https://doi.org/10.1029/2000JC000705

Pithan F, Mauritsen T (2014) Arctic amplification dominated by temperature feedbacks in contemporary climate models. Nat Geosci 7:181–184. https://doi.org/10.1038/ngeo2071

Rampal P, Weiss J, Marsan D (2009) Positive trend in the mean speed and deformation rate of Arctic sea ice, 1979–2007. J Geophys Res Oceans (1978–2012) 114(C5):C05013

Rampal P, Weiss J, Dubois C, Campin J-M (2011) IPCC climate models do not capture Arctic sea ice drift acceleration: consequences in terms of projected sea ice thinning and decline. J Geophys Res 116:C00D07. https://doi.org/10.1029/2011JC007110

Screen JA, Simmonds I (2010) The central role of diminishing sea ice in recent Arctic temperature amplification. Nature 464:1334–1337

Sedlar J, Tjernström M (2009) Stratiform cloud—Inversion characterization during the Arctic melt season. Bound Layer Meteorol 132:455–474. https://doi.org/10.1007/s10546-009-9407-1

Sedlar J, Tjernström M (2017) Clouds, warm air and a climate cooling signal over the summer Arctic. Geophys Res Lett 44:1095–1103. https://doi.org/10.1002/2016GL071959

Serreze MC, Holland MM, Stroeve J (2007) Perspectives on the Arctic's shrinking sea-ice cover. Science 315:1533–1536. https://doi.org/10.1126/science.1139426

Shupe MD (2011) Clouds at Arctic atmospheric observatories, Part II: thermodynamic phase characteristics. J Appl Meteorol Climatol 50(3):645–661. https://doi.org/10.1175/2010JAMC2468.1

Shupe MD, Intrieri JM (2004) Cloud radiative forcing of the Arctic surface: the influence of cloud properties, surface albedo, and solar zenith angle. J Clim 17:616–628

SIMIP Community (2020G) Arctic sea ice in CMIP6. Geophys Res Lett 47:086749. https://doi.org/10.1029/2019GL086749

Stroeve J, Notz D (2018) Changing state of Arctic sea ice across all seasons. Environ Res Lett 13:20

Stroeve J, Barrett A, Serreze M, Schweiger A (2014) Using records from submarine, aircraft and satellites to evaluate climate model simulations of Arctic sea ice thickness. Cryosphere 8(5):1839–1854

Swart NC, Fyfe JC, Hawkins E, Kay JE, Jahn A (2015) Influence of internal variability on Arctic sea-ice trends. Nat Clim Change 5(2):86–89

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Am Meteorol Soc 93(4):485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Taylor PC, Cai M, Hu A, Meehl J, Washington W, Zhang GJ (2013) A decomposition of feedback contributions to polar warming amplification. J Clim. https://doi.org/10.1175/jcli-d-12-00696.1

Tokinaga H, Xie SP, Mukougawa H (2017) Early 20th-century Arctic warming intensified by Pacific and Atlantic multidecadal variability. Proc Natl Acad Sci 114(24):6227–6232

Topál D, Ding Q, Mitchell J, Baxter I, Herein M, Haszpra T, Luo R, Li Q (2020) An internal atmospheric process determining summertime Arctic sea ice melting in the next three decades: lessons

learned from five large ensembles and multiple CMIP5 climate simulations. J Clim 33(17):7431–7454
Wernli H, Papritz L (2018) Role of polar anticyclones and mid-latitude cyclones for Arctic summertime sea-ice melting. Nat Geosci 11:108–113. https://doi.org/10.1038/s41561-017-0041-0
Westermann S, Lüers J, Langer M, Piel K, Boike J (2009) The annual surface energy budget of a high-arctic permafrost site on Svalbard, Norway. Cryosphere 3(2):245
Young G, Connolly P, Dearden C, Choularton T (2018) Relating large-scale subsidence to convection development in Arctic mixed-phase marine stratocumulus. Atmos Chem Phys 18:1475–1494. https://doi.org/10.5194/acp-18-1475-2018
Zhang R (2015) Mechanisms for low-frequency variability of summer Arctic sea ice extent. Proc Natl Acad Sci 112(15):4570–5457