ADVANCES IN ATMOSPHERIC SCIENCES, 2024

• Original Paper •

Projecting Wintertime Newly Formed Arctic Sea Ice through Weighting CMIP6 Model Performance and Independence※

Jiazhen ZHAO[1], Shengping HE[*2,4,5], Ke FAN[3], Huijun WANG[1,5,6], and Fei LI[2]

[1]*Key Laboratory of Meteorological Disaster*, *Ministry of Education/Joint International Research Laboratory of Climate and Environment Change (ILCEC)/Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD)*, *Nanjing University of Information Science & Technology*, *Nanjing* 210044, *China*

[2]*Geophysical Institute*, *University of Bergen and Bjerknes Centre for Climate Research*, *Bergen* 5007, *Norway*

[3]*School of Atmospheric Science*, *Sun Yat-sen University*, *and Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai)*, *Zhuhai* 519082, *China*

[4]*Nansen Environmental and Remote Sensing Center and Bjerknes Centre for Climate Research*, *Bergen* 5007, *Norway*

[5]*Nansen-Zhu International Research Center*, *Institute of Atmospheric Physics*, *Chinese Academy of Sciences*, *Beijing* 100029, *China*

[6]*Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai)*, *Zhuhai* 519082, *China*

(Received 29 December 2022; revised 19 April 2023; accepted 6 May 2023)

ABSTRACT

Precipitous Arctic sea-ice decline and the corresponding increase in Arctic open-water areas in summer months give more space for sea-ice growth in the subsequent cold seasons. Compared to the decline of the entire Arctic multiyear sea ice, changes in newly formed sea ice indicate more thermodynamic and dynamic information on Arctic atmosphere–ocean–ice interaction and northern mid–high latitude atmospheric teleconnections. Here, we use a large multimodel ensemble from phase 6 of the Coupled Model Intercomparison Project (CMIP6) to investigate future changes in wintertime newly formed Arctic sea ice. The commonly used model-democracy approach that gives equal weight to each model essentially assumes that all models are independent and equally plausible, which contradicts with the fact that there are large interdependencies in the ensemble and discrepancies in models' performances in reproducing observations. Therefore, instead of using the arithmetic mean of well-performing models or all available models for projections like in previous studies, we employ a newly developed model weighting scheme that weights all models in the ensemble with consideration of their performance and independence to provide more reliable projections. Model democracy leads to evident bias and large intermodel spread in CMIP6 projections of newly formed Arctic sea ice. However, we show that both the bias and the intermodel spread can be effectively reduced by the weighting scheme. Projections from the weighted models indicate that wintertime newly formed Arctic sea ice is likely to increase dramatically until the middle of this century regardless of the emissions scenario. Thereafter, it may decrease (or remain stable) if the Arctic warming crosses a threshold (or is extensively constrained).

Key words: wintertime newly formed Arctic sea ice, model democracy, model weighting scheme, model performance, model independence

Citation: Zhao, J. Z., S. P. He, K. Fan, H. J. Wang, and F. Li, 2024: Projecting wintertime newly formed Arctic sea ice through weighting CMIP6 model performance and independence. *Adv. Atmos. Sci.*, https://doi.org/10.1007/s00376-023-2393-2.

Article Highlights:

- CMIP6 projections of wintertime newly formed sea ice are subject to large bias and uncertainty.
- Both the bias and uncertainty can be effectively constrained by weighting models by their performance and independence.
- Weighted projections indicate that newly formed sea ice will likely increase continuously from the mid-2000s to the mid-21st century.
- Thereafter, newly formed sea ice may decrease (or stabilize) if the Arctic warming crosses a threshold (or is constrained).

※ This paper is a contribution to the special issue on the Ocean, Sea Ice and Northern Hemisphere Climate: In Remembrance of Professor Yongqi GAO's key contributions.

* Corresponding author: Shengping HE
Email: shengping.he@uib.no

1. Introduction

Arctic sea ice has declined precipitously in recent decades, which has been one of the most visible signals of climate change (Stroeve and Notz, 2018). This has manifested as a decrease in ice cover, extent, thickness, volume, age, and length of the ice-covered season (Laxon et al., 2013; Renner et al., 2014; Stroeve et al., 2014; Comiso et al., 2017; Kwok, 2018). The pronounced decline in sea-ice cover and extent since the beginning of the modern satellite record has been a prominent topic in climate research (Ding et al., 2019; Liu et al., 2021). While the decline in ice extent occurs in all months, the rates of decline maximize at the end of the melting season in September. In the past few decades, scientists have made great efforts to examine both the causes and climatic effects of the decline in September Arctic sea-ice extent (Kim et al., 2014; Ding et al., 2017; Liu et al., 2022).

One of the most direct consequences of the decline in September sea-ice extent is an increase in the open-water areas in the Arctic, which gives more space for sea-ice growth during the subsequent freeze-up season. One recent study noted that there was an increasing trend in the total extent of newly formed Arctic sea ice in the winters of 1979/80–2016/17 (Hegyi and Taylor, 2018). In addition, the substantial thinning of Arctic sea ice has led to the total thickness and total volume of wintertime newly formed sea ice having also experienced an increasing trend in recent years (Petty et al., 2018; Ricker et al., 2021). The underlying physical mechanism is associated with a negative feedback process during the growth of ice thickness, in which thinner ice grows faster than thicker ice owing to its decreased insulation (Bitz and Roe, 2004). With the dramatic decline in summer Arctic sea ice, we can imagine that wintertime newly formed Arctic sea ice will continue to increase in the future, which would suggest that newly formed sea ice may play an increasingly important role in the Arctic climate system.

The continuous Arctic sea-ice decline could lead to an ice-free Arctic (i.e., sea-ice extent or area below 1×10^6 km^2) in summer over the coming decades (Jahn, 2018; Niederdrenk and Notz, 2018; Zhao et al., 2022b). If the future Arctic reaches an ice-free condition in September, the sea ice in a year will be completely composed of wintertime newly formed sea ice. Compared with ice that lasts multiple years (i.e., multiyear sea ice), newly formed sea ice is thinner, more breakable, easier to move, and more sensitive to climate variability. The dramatic decline in multiyear sea-ice coverage and the increase in newly formed sea ice demonstrates that the Arctic cryosphere is transitioning from a multiyear ice–dominated state to a newly formed ice–dominated state. Such a transition may lead to essential changes in atmosphere–ocean–ice interaction in the Arctic, which would affect the northern mid–high latitude climate (Deng and Dai, 2022; Zhao et al., 2023). This implies an urgency and necessity to improve assessments of future changes in wintertime newly formed sea ice. While previous studies have explored the changes and variability of total Arctic sea ice (including both newly formed sea ice and multiyear sea ice) (Wernli and Papritz, 2018; Olonscheck et al., 2019; Petty et al., 2020), changes in newly formed sea ice alone have received relatively less attention.

However, projections of wintertime newly formed Arctic sea ice are subject to large uncertainty (Petty et al., 2018), and this uncertainty needs to be effectively constrained to better understand and predict future climate changes over the northern mid–high latitudes. Climate projections are associated with three main sources of uncertainty: natural internal variability, emissions scenario uncertainty, and structural model uncertainty (Tebaldi and Knutti, 2007; Hawkins and Sutton, 2009; Maher et al., 2020). Internal variability is highly dependent on the variables, time period, and regions of interest, and arises from the chaotic behavior of the climate system so as to be generally regarded as irreducible on time scales of more than a few years (Hawkins and Sutton, 2009; Deser et al., 2012; Fatichi et al., 2016). Emissions scenario uncertainty is also hard to reduce since it is mainly determined by political decisions and technological development, which of course do not follow the laws of physics and are largely choices driven by society (van Vuuren et al., 2011; Lorenz et al., 2018). Therefore, reducing the model uncertainty that arises from differences in the response of models to specified forcing agents is the best choice if we want to reduce uncertainty in climate projections.

A common approach to quantify model uncertainty is to use multimodel ensembles that contain multiple models running common experiments but using different ways to characterize the physical processes, as in the Coupled Model Intercomparison Projects (CMIPs) (Eyring et al., 2016). When working with a multimodel ensemble, a prevailing approach is to give each model the same weight and use their arithmetic mean to estimate possible future changes of climate (Collins et al., 2013; IPCC, 2021). This approach is known as model democracy, and is based on the assumption that all models are independent of each other and have identical ability in providing reliable projections (Sanderson et al., 2015a, b; Knutti et al., 2017). However, such an assumption may not be tenable, especially in cases when large multimodel ensembles, such as those used in the CMIPs, are used to project future climate. On the one hand, models in a large ensemble normally differ greatly in their performances in reproducing the observed climate. It is hard to expect that a model having bad performance can provide projections as reliable as a model having good performance. On the other hand, there is large interdependence in large multimodel ensembles (Knutti et al., 2010, 2013; Masson and Knutti, 2013). Taking the CMIPs, for example, there are many versions of the same model that differ only in some minor aspects such as their horizontal resolution (e.g., MPI-ESM-HR and MPI-ESM-LR), and there are also numerous models from different institutes that share ideas and codes (e.g., CESM1.2.2 and TaiESM1) (Lee et al., 2020). These models have many similarities. They may have common structural

limitations, miss the same processes, or make similar approximations (Knutti et al., 2017). Therefore, many recent studies have proposed to consider both the performance and the independence of each model when using large multimodel ensembles for climate projections, arguing that the prevailing model democracy in many cases may not necessarily be the best choice (Annan and Hargreaves, 2011; Knutti et al., 2013; Herger et al., 2018).

In the past several decades, scientists have tried using sub-selection of well-performing models to quantify the actual uncertainty in climate projections (Herger et al., 2019; Notz and SIMIP Community, 2020; Docquier and Koenigk, 2021). This approach considers model performance in the model sub-selection process. However, it employs model democracy when dealing with results derived from the selected well-performing models (i.e., each model is equally weighted). This means that this approach ignores the interdependence of the selected models. Well-performing models produce outputs that are close to observations, and some of them are extremely close to each other, which usually results from the fact that they are from the same institute or share large amounts of code. These models have large interdependence, and ignoring their interdependence may bias the projections substantially. Taking an extreme case as an example, repeatedly using the same model would not provide any new information, but would instead amplify (double) the bias of that model from the observations. Ignoring the interdependence of these models may also lead to poor estimation of the actual uncertainty in climate projections (Herger et al., 2018; Brunner et al., 2020; Zhao et al., 2022b). In addition, using only a few models for projections may lead to loss of some new and useful information and thus produce overconfident results, which motivates us to use all available models for projections. The aforementioned limitations of the model sub-selection method imply an urgency and necessity to use advanced approaches that can consider both the model performance and independence when dealing with multimodel projections.

In the present study, we use a multimodel weighting scheme developed by Knutti et al. (2017) based on the work of Sanderson et al. (2015a, b), which weights models with consideration of both model performance and model independence, to improve the skill and quantify the actual uncertainty in projections from phase 6 of CMIP (CMIP6) of wintertime newly formed Arctic sea ice. It has been proven that such a weighting scheme is effective at reducing the uncertainty in projections of many variables, such as September Arctic sea ice (Zhao et al., 2022b), summer European temperature and precipitation (Brunner et al., 2019), and summer runoff in the Yangtze River Basin (Zhao et al., 2022a). The specific aim of the study is to use such a weighting scheme to provide reliable projections of wintertime newly formed Arctic sea ice for the scientific community.

Following this introduction, section 2 details the datasets used, the definitions of wintertime newly formed Arctic sea ice, and the weighting scheme employed. Section 3 evaluates the weighting scheme and presents future projections of wintertime newly formed Arctic sea ice. Section 4 provides some further discussion and our conclusions based on the results of this study.

2. Data and methods

2.1. *Observations and model datasets*

Monthly sea-ice concentration data on a 1° latitude × 1° longitude grid were obtained from the Met Office Hadley Centre Sea Ice and Sea Surface Temperature dataset, version 1.1 (HadISST 1.1; Rayner et al., 2003). Note that the main results of this study do not change appreciably if other observational sea-ice concentration datasets are used. The monthly sea-ice extent was obtained from the National Snow and Ice Data Center (NSIDC) Sea Ice Index, version 3 (Fetterer et al., 2017). To obtain the time series of monthly sea-ice area, we multiplied every grid-cell area by its sea-ice concentration and summed up all the grid values over the Northern Hemisphere.

The multimodel ensemble used in this study comprises all currently available CMIP6 models that provide monthly sea-ice concentration and surface air temperature in both historical simulations during 1850–2014 and future projections during 2015–2099 under four different Shared Socioeconomic Pathways (SSPs): SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 (Eyring et al., 2016; O'Neill et al., 2016). These four pathways have radiative forcings in 2100 of 2.6, 4.5, 7.0 and 8.5 W m^{-2}, respectively. In total, 29 CMIP6 models are used in this study, and for each model only the first available ensemble member is used (see Table 1 for a full list). This is because (1) the number of total ensemble members of most CMIP6 models is small, which makes it hard to analyze the impacts of internal variability, and (2) it has been proven that internal variability has a negligible impact on sea-ice projections obtained by the weighting scheme (Zhao et al., 2022b). To facilitate comparison, all model outputs are regridded to a regular 1° × 1° grid using bilinear remapping. For each model, the sea-ice extent time series was estimated as the sum of all areas in the Northern Hemisphere covered by sea ice with a concentration of at least 0.15. The time series of sea-ice area was estimated using the same approach as for the observations based on sea-ice concentration data.

2.2. *Definitions of wintertime newly formed Arctic sea ice*

In this study, wintertime newly formed Arctic sea ice is quantified as the difference between the sea-ice state at the beginning and the end of the freeze-up season (latter minus former). The freeze-up season in year i can be defined in two distinct ways. One is a dynamic way, by which the freeze-up season is defined as the period from the time in year i when the sea-ice state reaches the minimum to the time in year $i + 1$ when the sea-ice state reaches the maximum. In this case, the length of the freeze-up season is dependent on the variables used to characterize the sea-ice state,

Table 1. Details of the 29 CMIP6 models used in this study, including model ID, institution ID, country, ocean component model, and ocean model grid numbers (number of *x*-direction grids × number of *y*-direction grid × number of vertical layers).

Model no.	Model ID	Institution ID/country	Ocean model/grid numbers (lon.×lat.×lev.)
1	ACCESS-CM2	CSIRO-ARCCSS/Australia	ACCESS-OM2/360×300×50
2	ACCESS-ESM1-5	CSIRO/Australia	ACCESS-OM2/360×300×50
3	BCC-CSM2-MR	BCC-CMA/China	MOM4/360×232×40
4	CAMS-CSM1-0	CAMS-CMA/China	MOM4/360×200×50
5	CanESM5	CCCMA/Canada	NEMO3.4.1/361×290×45
6	CanESM5-CanOE	CCCMA/Canada	NEMO3.4.1/361×290×45
7	CAS-ESM2-0	CAS/China	LICOM2.0/362×196×30
8	CESM2	NCAR/USA	POP2/320×384×60
9	CESM2-WACCM	NCAR/USA	POP2/320×384×60
10	CNRM-CM6-1	CNRM-CERFACS/France	NEMO3.6/362×294×75
11	CNRM-CM6-1-HR	CNRM-CERFACS/France	NEMO3.6/1442×1050×75
12	CNRM-ESM2-1	CNRM-CERFACS/France	NEMO3.6/362×294×75
13	EC-Earth3	EC-Earth-Consortium/EU	NEMO3.6/362×292×75
14	EC-Earth3-Veg	EC-Earth-Consortium/EU	NEMO3.6/362×292×75
15	EC-Earth3-Veg-LR	EC-Earth-Consortium/EU	NEMO3.6/362×292×75
16	FGOALS-f3-L	LASG-IAP/China	LICOM3.0/360×218×30
17	FGOALS-g3	LASG-IAP/China	LICOM3.0/360×218×30
18	INM-CM4-8	INM/Russia	INM-OM5/360×318×40
19	INM-CM5-0	INM/Russia	INM-OM5/720×720×40
20	IPSL-CM6A-LR	IPSL/France	NEMO-OPA/362×332×75
21	MIROC6	MIROC/Japan	COCO4.9/360×256×63
22	MIROC-ES2L	MIROC/Japan	COCO4.9/360×256×63
23	MPI-ESM1-2-HR	MPI-M/Germany	MPIOM1.63/802×404×40
24	MPI-ESM1-2-LR	MPI-M/Germany	MPIOM1.63/256×220×40
25	MRI-ESM2-0	MRI/Japan	MRI.COM4.4/360×364×61
26	NorESM2-LM	NCC/Norway	MICOM/360×384×70
27	NorESM2-MM	NCC/Norway	MICOM/360×384×70
28	TaiESM1	AS-RCEC/Taiwan	POP2/320×384×60
29	UKESM1-0-LL	MOHC/UK	NEMO-HadGEM3-GO6.0/360×330×75

and may not always be the same in all years. The other is a fixed way, by which the freeze-up season is defined as the period that conventionally spans from September of year i to March of year $i + 1$. In this case, the length of the freeze-up season does not depend on the variables used to characterize the sea-ice state, and is identical in all years. Note that, in this paper, years are labeled as the calendar year of the freeze-up season. The focus of this study was on the sea-ice growth in the horizontal direction (i.e., cover, extent, area) rather than the vertical direction (i.e., thickness). When analyzing the temporal variability of newly formed sea ice, the monthly sea-ice extent and monthly sea-ice area are used to characterize the sea-ice state, enabling us to obtain a total of four different time series of newly formed sea ice based on the two definitions of the freeze-up season. When analyzing the spatial patterns of the newly formed sea ice, the monthly sea-ice concentration is used to characterize the sea-ice state, and this is applied only to the second definition of the freeze-up season (the fixed way). The newly formed sea ice quantified in this way is termed as "the newly formed sea-ice concentration" hereafter.

2.3. *Multimodel weighting scheme*

In this study, a weighting scheme is applied to the CMIP6 large multimodel ensemble (containing 29 models in total; see Table 1 for details) to improve the skill and quantify the real uncertainty in the projections of wintertime newly formed Arctic sea ice. The weighting scheme was developed by Knutti et al. (2017), following on from the work of Sanderson et al. (2015a, b). One of the most attractive advantages of this scheme is that it considers the model interdependence in multimodel ensembles, which has been commonly overlooked in previous studies, including the IPCC reports (Collins et al., 2013; IPCC, 2021). The weighting scheme weights each CMIP6 model according to its performance and its independence, following the principle that models agreeing well with observations (i.e., having good performance) will be up-weighted, and models having rare duplicates within the multimodel ensemble (i.e., having large independence) will also be up-weighted. The performance of each model is evaluated by the distance of that model from the observation, with a smaller distance indicative of better performance. The independence of each model is evaluated

by the distances of that model from other models, with larger distances indicative of larger independence. The distance metric is represented by the root-mean-square-error (RMSE). Such a metric is not the only option, and some other skill scores may also be applicable (Perkins et al., 2007; Chen et al., 2011). Distances are calculated using observations and model simulations of the historical "climate" related to the projected target. The "climate" is characterized by basic climate features (e.g., climatology, trend, interannual variability) of a set of variables that are relevant to the projected target (see section 2.4). Thus, it can be seen that the performance and independence of each model are comprehensively evaluated using multiple diagnostics instead of solely the projected target. This can reduce the risk that unreliable models are unreasonably up-weighted just because they accidently agree well with the projected target observations. After all models are weighted, the mean of the multimodel ensemble is utilized to estimate a relatively accurate and representative future change of wintertime newly formed Arctic sea ice. A weighted standard deviation or percentiles across the weighted models are calculated to quantify the actual uncertainty in future projections of the newly formed sea ice.

The main equation calculates a single weight w_i for model i as a product of its performance weight (numerator) and its independence weight (1/denominator):

$$w_i = \frac{e^{-\frac{D_i^2}{\sigma_D^2}}}{1 + \sum_{j \neq i}^{M} e^{-\frac{S_{ij}^2}{\sigma_s^2}}}, \quad (1)$$

with D_i being the distance of model i to the observations, σ_D a constant parameter determining the strength of the performance weighting, M the total number of models, $S_{ij} (j \neq i)$ the distance of model i to another model j, and σ_s a constant parameter determining the strength of the independence weighting. To facilitate comparisons between the projections derived from the weighted and unweighted (i.e., each model can be seen to have an equal weight of $1/M$) multimodel ensemble, which can help evaluate the effect of the weighting scheme, all models' weights are multiplied by a constant to ensure that their sum is equal to one. Diagnostics play equal roles in the weight estimation. D_i and S_{ij} are calculated as the mean distance across all diagnostics. Note that the magnitude of the RMSE raw values in terms of different diagnostics may have large discrepancies, which can lead to unequal roles of the diagnostics in the weight estimation. To address this issue, for each diagnostic, both the RMSE distances between models and observations and between model pairs are put together and normalized by their median before the diagnostic distance is averaged.

The two shape parameters [σ_D and σ_s in Eq. (1)] determine the typical distance by which a model would be considered to be "close" to observations (i.e., setting the strength of the performance weighting) and to another model (i.e., setting the strength of the independence weighting), respectively. It has been suggested that results derived by the weighting scheme show relatively low sensitivity to the choice of the σ_s value, but has high sensitivity to the choice of the σ_D value (Knutti et al., 2017). A small σ_D value results in aggressive weighting with few models receiving most of the weights and therefore would produce narrow projections, while a large σ_D value leads to an approximation of equal weighting and therefore would produce wide projections (see Knutti et al., 2017 for a discussion). It is inappropriate to determine σ_D as a too large or a too small value. To choose a proper pair of σ_D and σ_s values, we follow the same principle as Knutti et al. (2017) by using a "perfect model setup", which is explained in detail in section 2.5.

2.4. *Projected target and weighting diagnostics*

Observations reveal that the time series of wintertime newly formed Arctic sea ice quantified by either monthly sea-ice extent (Fig. 1a) or monthly sea-ice area (Fig. 1b) based on the two different definitions of the freeze-up season (see section 2.2) have extremely high consistency in the present period. This indicates that the extent and area of Arctic sea ice in most years (except for a few years, such as 1986, 1997 and 2008) reach their minimum (maximum) value in September (March). In addition, high consistency can also be found between the time series of the newly formed sea ice quantified by the monthly sea-ice extent and the monthly sea-ice area (Fig. 1a vs Fig. 1b). Therefore, we use only one of these metrics to quantify the newly formed sea ice (i.e., the monthly sea-ice extent from the fixed way; termed as "newly formed sea-ice extent" hereafter) in the following analysis of its temporal variations. The wintertime newly formed sea-ice extent is chosen as the projected target in this study. Note that the results of this study do not change much when we use other metrics (e.g., monthly sea-ice area) to quantify the newly formed sea ice (not shown). Given that the sea-ice extent is directly tied to the sea-ice concentration (see section 2.1), as well as the fact that the climatology, trend, and interannual variability have been commonly regarded as basic features of climate relevant to the projected target in model assessments (Guo et al., 2021; Shiru and Chuang, 2021), we select six diagnostics to weight the CMIP6 models in this study, including the climatology, trend, and interannual variability (characterized by the standard deviation of the linearly detrended data) of (1) the newly formed sea-ice extent and (2) the newly formed sea-ice concentration (see section 2.2) in the Arctic region (60°–90°N). Such a number of diagnostics has been found to be a good number in the assessment of a model's ability to reproduce the observed climate (Lorenz et al., 2018). A larger number tends to converge the model performances, since models are unlikely to perform well across a large number of diagnostics, while a smaller number increases the risk of producing overconfident performance weighting. D_i and S_{ij} are calculated based on the six diagnostics according to models' simulations during the historical time period, which is determined as 1979–2013 with consideration of the avail-

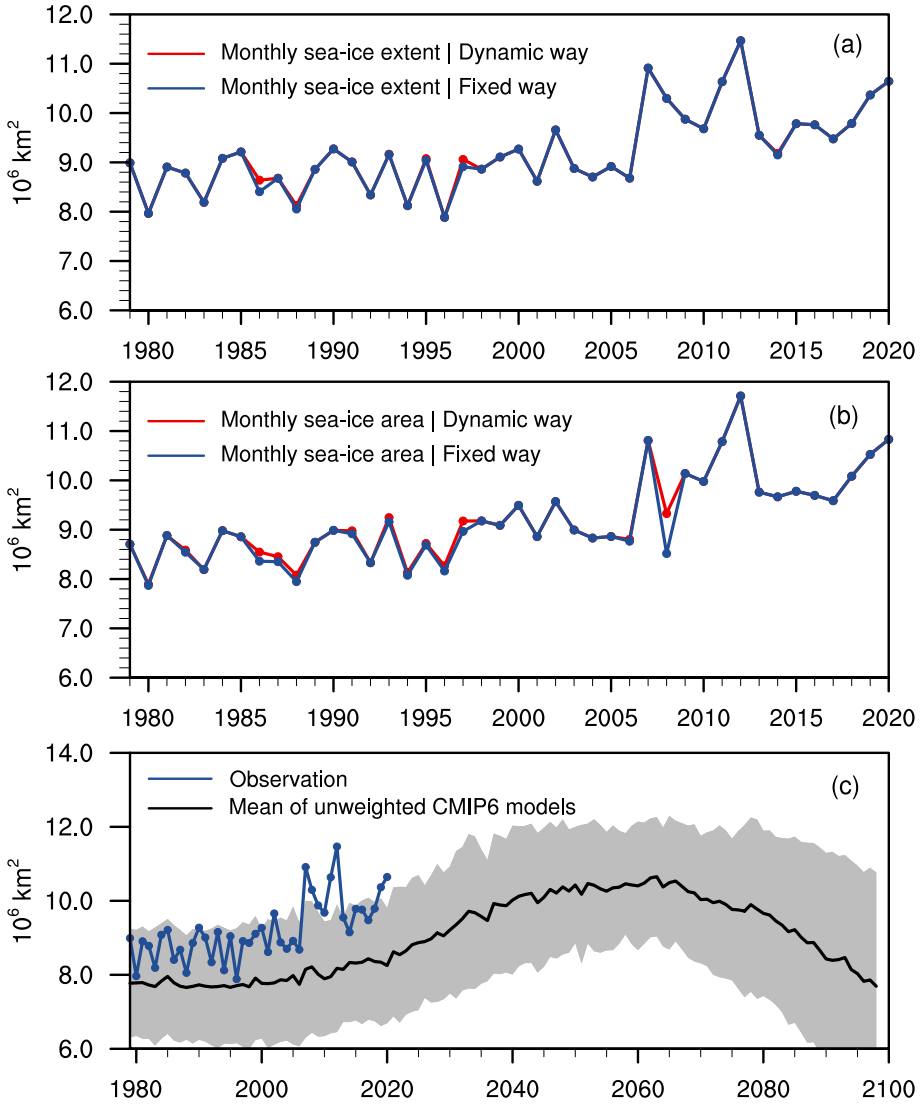

Fig. 1. (a) Observed time series of wintertime newly formed Arctic sea ice quantified by the monthly sea-ice extent from the dynamic way (red curve) and the fixed way (blue curve). (b) As in (a) except that the newly formed sea ice is quantified by the monthly sea-ice area. (c) Time series of wintertime newly formed Arctic sea ice in the unweighted CMIP6 multimodel projections under the SSP5-8.5 scenario. The newly formed sea ice is quantified by the monthly sea-ice extent from the fixed way. The black curve and gray band indicate the unweighted multimodel mean and intermodel spread, respectively. The blue curve represents the observation.

ability of observations and model outputs. A distance matrix, which contains both the D_i value (last row) and the S_{ij} value (first M rows, with M being the total number of models) of each model, is illustrated in Fig. 2. According to the last row of this matrix, MRI-ESM2-0 (FGOALS-g3) has the smallest (largest) distance from the observations, indicating that it has the best (worst) performance in simulating wintertime newly formed Arctic sea ice. The changes in two 30-year future periods, representing the mid-century (2039–2068) and end-of-century (2069–2098) conditions relative to the 1981–2010 baseline are analyzed to assess the effect of the weighting scheme on future projections.

2.5. *Estimation of shape parameters*

In this study, following Knutti et al. (2017), a perfect model setup is employed to estimate the values of the shape parameters (i.e., σ_D and σ_s) in Eq. (1). This setup solely uses model simulations without any observations. It first chooses one model in turn from the multimodel ensemble as a perfect model. Simulations of this perfect model during the historical time period (i.e., 1979–2013) are treated as "pseudo-observations ". The remaining models are then weighted based on Eq. (1) using models' distances from the "pseudo-observations" (performance weighting) and their distances from each other (independence weighting). After the

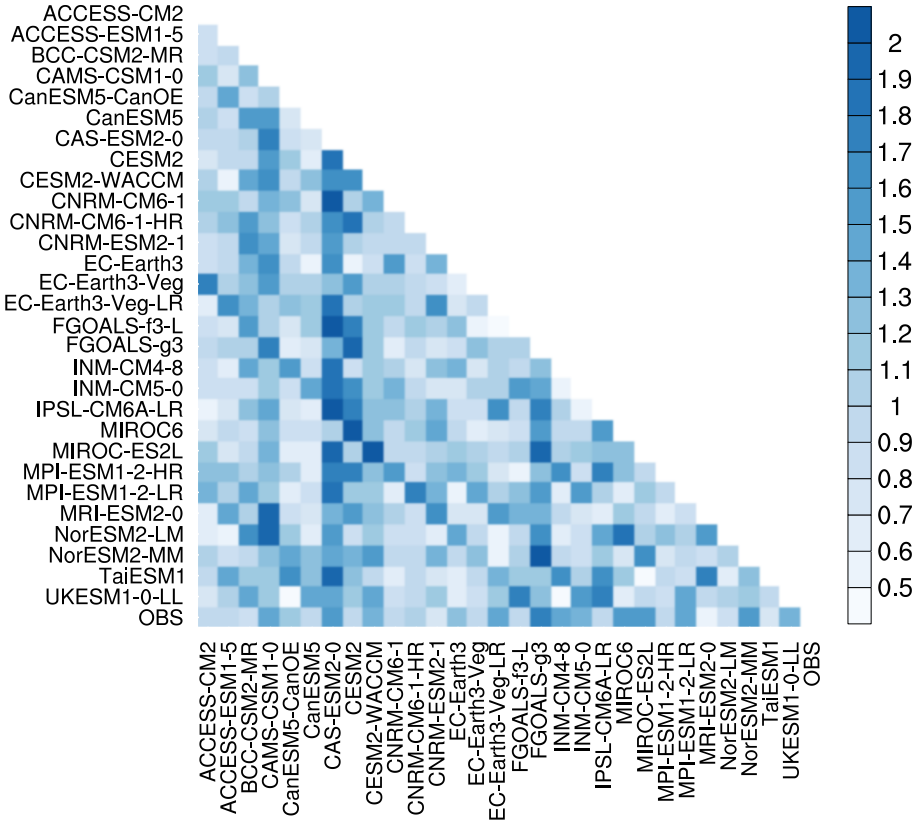

Fig. 2. Mean RMSE distance matrix for CMIP6 models averaged across the six weighting diagnostics (section 2.4). It contains both the model-to-model distances (S_{ij}; first M rows, with M being the total number of models) and the model-to-observation distances (D_i; the last row). Each row (column) represents a single climate model or observation. Each box represents a pairwise distance. Colors indicate the magnitude of the distance, with darker colors indicative of larger distance. Smaller distances mean the datasets are in better agreement than larger distances. Distances are calculated using observations and model simulations during 1979–2013. Each pairwise distance shown is measured as a fraction of the median of all pairwise distances.

remaining models are weighted, they are utilized to predict future projection results of that perfect model. For each perfect model, we can obtain one distribution across the individual weighted predictions (termed "weighted distribution" hereafter). The setup is conducted M times (M being the total number of models) to make sure that each model in the multimodel ensemble has been selected as the perfect model only once (so we can obtain M projection results and the corresponding M weighted distribution). After this process has been completed, we count the number of cases in which the future projection result of the perfect model lies within the 10th–90th percentile of the corresponding weighted distribution, and calculate a fraction case by dividing that number by M. Note that we use multiple combinations of σ_D and σ_s with each parameter ranging from 0.2 to 1.1 with an incremental bin of 0.025 (there are a total of 37 $\sigma_D \times 37$ σ_s combinations) for the model weighting, and as a result, we can obtain 37 × 37 different corresponding case fractions. This can help us analyze the dependence of the case fraction on σ_D and σ_s values, and the results are shown in Fig. 3. In this study, we focus on four future projection results of the perfect model, which are chosen as the mean newly formed sea-ice extent (i.e., the projected target; see section 2.4) averaged over the future period from 2015 to 2098 under four different emission scenarios (i.e., SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5 scenarios). The case fraction shown in Fig. 3 is the mean fraction averaged across the four scenarios. The best parameter combination should produce a weighted distribution that is neither too wide nor too narrow. To ensure the weighted distribution is not too narrow, which would lead to overconfident weighting, the case fraction needs to be high enough. Here, we followed the same principle as Lorenz et al. (2018) whereby the case fraction needs to reach at least 0.8 (indicated in Fig. 3 by the hatched areas). To avoid the weighted distribution being too wide, σ_D is determined here as the smallest value that can make the case fraction meet the requirement of 0.8. This is because a larger σ_D value tends to produce a wider weighted distribution (see section 2.3). Since the case fraction and the width of the weighted distribution is insensitive to the choice of the σ_s value (Fig. 3), the σ_s value is determined here as 0.6, and in this case the σ_D value is determined as 0.475 (see the blue dot in Fig. 3).

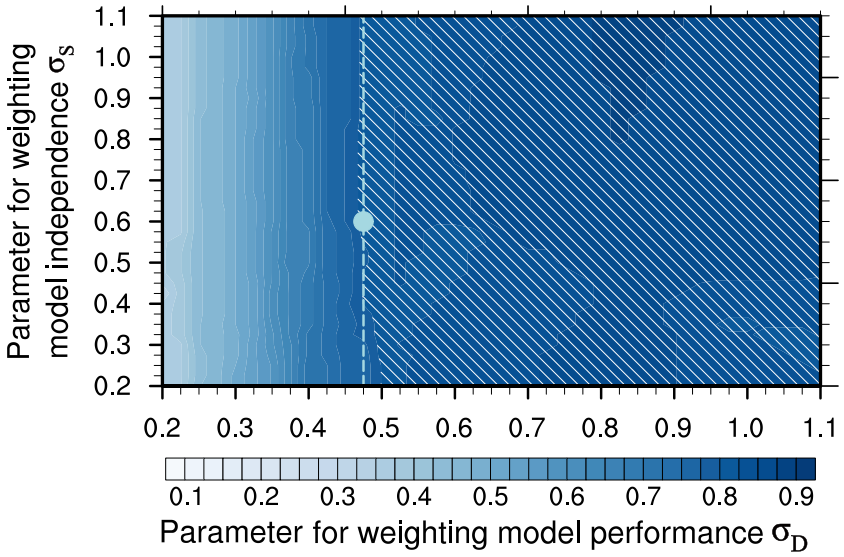

Fig. 3. Fraction of cases (shading) in which the actual projection result of the perfect model lies within the 10th–90th percentile of the weighted distribution in the perfect model setup, as a function of the σ_D value (x-axis) and the σ_s value (y-axis) in Eq. (1) (both range from 0.2 to 1.1 with an incremental bin of 0.025). Four projection results of the perfect model are considered, which are the mean newly formed sea-ice extent averaged over the future time period from 2015 to 2098 under the four different emission scenarios (i.e., SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5 scenarios). The case fraction shown here is the mean fraction averaged across the four scenarios. Hatched areas indicate that the case fraction exceeds the required value of 0.8. The vertical line indicates the minimum value of σ_D for the case faction to exceed the required value of 0.8, given a σ_s value of 0.6. The solid dot indicates the values of the two shape parameters in Eq. (1) (i.e., σ_D = 0.475, σ_s = 0.6) used in this study.

3. Results

3.1. *Evaluation of the weighting scheme*

Observations reveal that the wintertime newly formed sea-ice extent has experienced a precipitous increasing trend since the mid-2000s (Fig. 1c; blue curve). While the arithmetic mean of the unweighted CMIP6 multimodel ensemble indicates that such an increasing trend is likely to continue until the middle of this century under a high-emissions scenario (i.e., SSP5-8.5), it notably underestimates the observed climatology, trend, and interannual variability of the newly formed sea-ice extent during the present period (Fig. 1c; black curve). In addition, the spread across individual projections within the multimodel ensemble is also evident (Fig. 1c; grey band). The bad performance of the original outputs of CMIP6 models affects the reliability of their future projections, which motivates us to use the weighting scheme to improve the skill of the projections. To evaluate the effect of the weighting scheme, we present in Fig. 4 (top row) a comparison between the historical simulations of the newly formed sea-ice extent derived from the unweighted (left in each panel) and the weighted (right in each panel) CMIP6 multimodel ensemble. To help understand the simulation bias from the observations, we also compare their simulations of the monthly sea-ice extent at the beginning (i.e., September) and the end (i.e., March) of the freeze-up season, and the results are shown in the middle row and the last row in Fig. 4, respectively.

It is indicated that both the mean and median of the unweighted CMIP6 multimodel ensemble largely underestimate the observed climatology, increasing trend, and interannual variability of the wintertime newly formed sea-ice extent during the historical time period (Figs. 4a–c; grey bars and black dots). Bias in the climatology simulation is primarily attributed to an underestimation of the observed climatology of March sea-ice extent (Fig. 4g; grey bars and black dots). By comparison, the mean and median of the unweighted ensemble perform much better in simulating the climatology of September sea-ice extent (Fig. 4d; grey bars and black dots). In terms of the trend simulation, the situation becomes much different. The mean and median of the unweighted ensemble capture well the observed decreasing trend of March sea-ice extent (Fig. 4h; grey bars and black dots). In contrast, they underestimate the observed rates of decline of September sea-ice extent (Fig. 4e; grey bars and black dots), which leads to underestimation of the observed rates of increase in the newly formed sea-ice extent. In terms of bias in the interannual variability simulation, the bias in the simulations of September and March sea-ice extent may both play a role, so it is hard to attribute the bias of the newly formed sea ice simulation to one of them. The difficulty is that the interannual variability of newly formed sea-ice extent is not linearly correlated with the interannual

variability of either September or March sea-ice extent. We take two different extreme cases for examples. In each of the two cases, both the September and March sea-ice extent have large interannual variability. The difference between

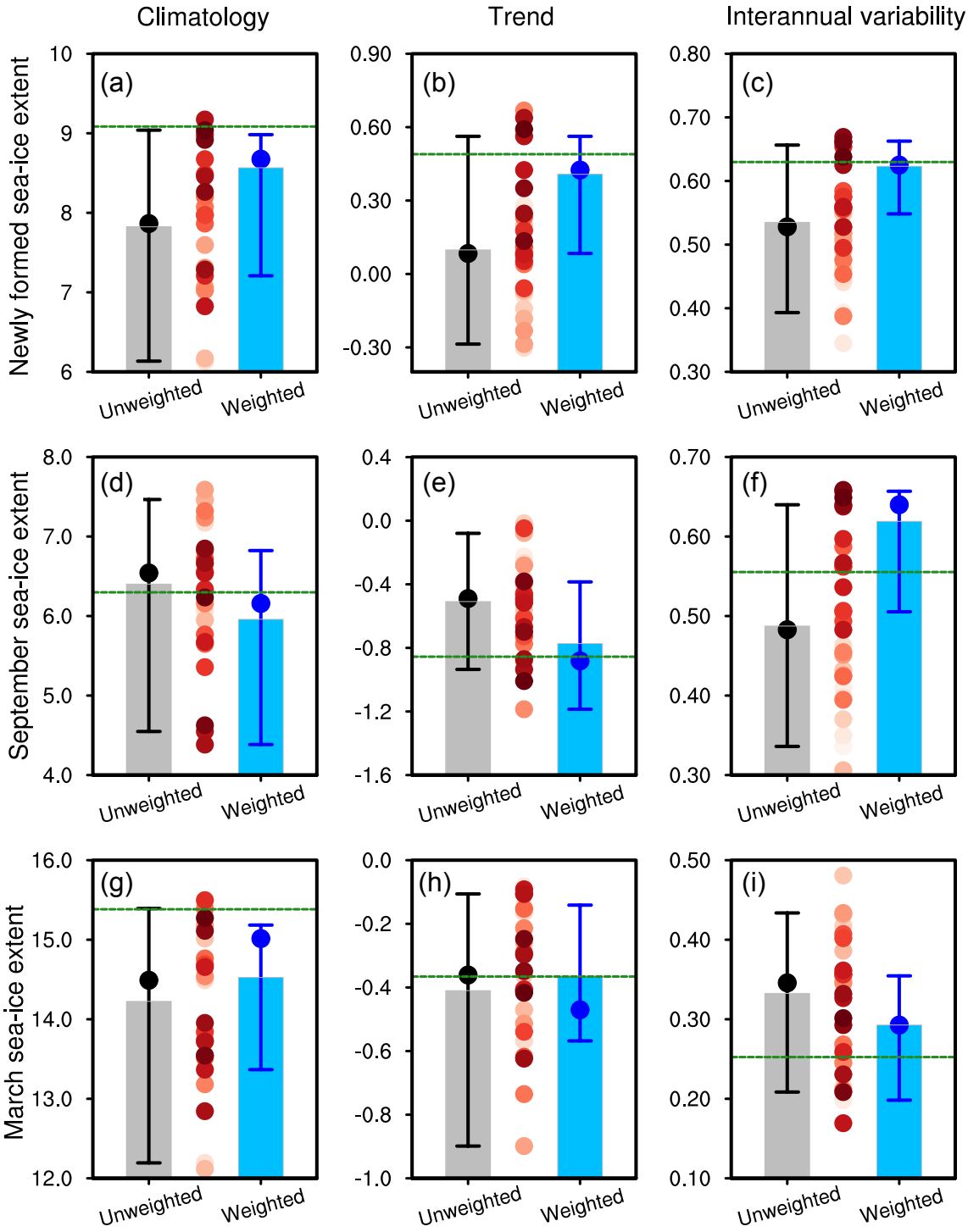

Fig. 4. Evaluation of the weighting scheme in terms of the (a) climatology (units: 10^6 km^2), (b) trend [units: 10^6 km^2 (10 yr)$^{-1}$], and (c) interannual variability (characterized by the standard deviation of the linearly detrended data; units: 10^6 km^2) of the wintertime newly formed sea-ice extent during the historical time period (i.e., 1979–2013) simulated by the unweighted (grey bars, black whiskers and dots) and weighted (blue bars, whiskers and dots) CMIP6 multimodel ensemble. (d–i) As in (a–c) but for simulations of (d–f) September sea-ice extent and (g–i) March sea-ice extent. In each panel, the bars (black or blue dot) represent the mean (median) of the CMIP6 multiple models, the whiskers denote the range between the 10th and 90th percentiles of the individual models, and the horizontal green line denotes the observation. The colored dots in the middle of each panel indicate the results derived from the individual models, with darker colors indicative of larger weights assigned by the weighting scheme.

the two cases is that the September and March sea-ice extent in the first (second) case are in-phase (out-of-phase) correlated, with a correlation coefficient of 1.0 (−1.0). Therefore, the newly formed sea-ice extent in the first (second) case would have small (large) interannual variability. This indicates that the interannual variability of the newly formed sea-ice extent is not only determined by the interannual variability of March and September sea-ice extent, but also affected by their relationship. We can see from Fig. 4f and Fig. 4i (grey bars and black dots) that the unweighted mean and unweighted median have poor skill in simulating the interannual variability of either September or March sea-ice extent. It is therefore unsurprising to see that they also have poor skill in simulating the interannual variability of the newly formed sea-ice extent.

Compared to the mean and median of the unweighted ensemble, the mean and median of the weighted ensemble have much less bias from the observations for most simulations shown in Fig. 4 (blue bars and dots). In addition, the ranges between the 10th and 90th percentiles of the individual models for most simulations become evidently smaller in the weighted case (blue whiskers in Fig. 4) than in the unweighted case (black whiskers in Fig. 4).This clearly indicates that the weighting scheme effectively reduces both the bias and the intermodel spread in CMIP6 historical simulations of the wintertime newly formed Arctic sea ice as well as the sea-ice state (i.e., extent) at the beginning and the end of the freeze-up season. This is because models having smaller distances from the observations are assigned larger weights by the weighting scheme, and vice versa (middle column in each panel of Fig. 4). The progress made by the weighting scheme increases our confidence that using the weighted ensemble instead of the unweighted ensemble can generate more reliable future projections of wintertime newly formed Arctic sea ice.

3.2. *Projections of wintertime newly formed Arctic sea ice*

Figure 5 illustrates projections of wintertime newly formed sea-ice extent derived from the unweighted (black curves and grey bands) and weighted (colored curves and bands) CMIP6 multimodel ensemble. The time series of the newly formed sea-ice extent simulated by the mean of the weighted ensemble (colored curves in Fig. 5) have much less bias from the observation relative to the series simulated by the mean of the unweighted ensemble (black curves in Fig. 5). This gives more credit to apply the weighting scheme to the CMIP6 multimodel ensemble to obtain more reliable future projections of newly formed sea-ice extent. Models that simulate less newly formed sea-ice extent get lower weights. As a consequence, projection results from the weighted mean shift upward in the whole future time period when compared with the results from the unweighted mean (black curves vs colored curves in Fig 5), indicating that there will be more newly formed sea-ice extent in the future. In addition, the spread of the projected newly formed sea-ice extent across the individual models is evidently smaller in the weighted case than in the unweighted case (grey bands vs colored bands in Fig. 5). This implies that the actual uncertainty in projections of the newly formed sea-ice extent may be overestimated by those traditional methods that employ model democracy without consideration of model performance and model independence.

The mean of the weighted ensemble compares well against the recent increasing trend of the newly formed sea-

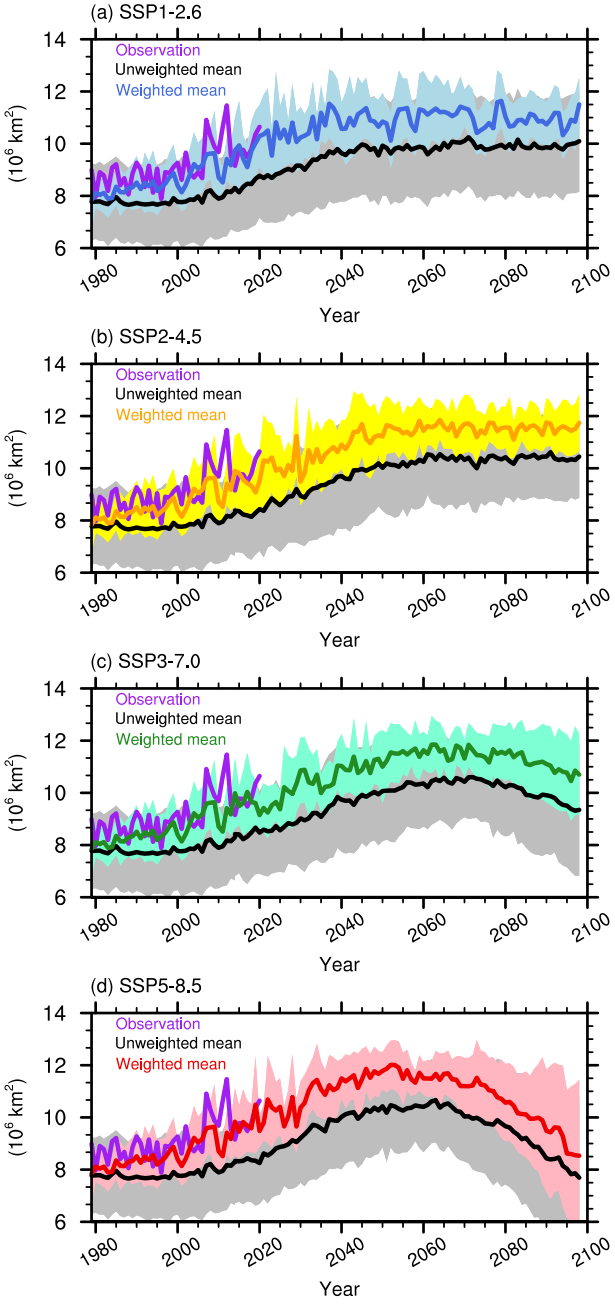

Fig. 5. Time series of wintertime newly formed sea-ice extent in the unweighted and weighted CMIP6 multimodel ensemble under the (a) SSP1-2.6, (b) SSP2-4.5, (c) SSP3-7.0, and (d) SSP5-8.5 scenarios. The colored curves and bands indicate the weighted multimodel mean and spread, respectively. The black curves and gray bands indicate the unweighted multimodel mean and spread, respectively. The purple solid curves represent observations.

ice extent shown in observations (purple curves in Fig. 5). Furthermore, projections derived from the mean of the weighted ensemble indicates that such an increasing trend is likely to continue until the middle of this century, regardless of the future emission scenario (colored curves in Fig. 5). However, the time when the newly formed sea-ice extent reaches a maximum in the 21st century is sensitive to the emission scenario, emerging around the year 2040, 2050, 2070 and 2050 under the SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5 scenario, respectively (colored curves in Fig. 5). Though the temperature in the Arctic region will rise dramatically in all months in the future, it will remain below the freezing point in the winter season over the coming decades, so the rising Arctic temperatures exert only weak suppression impacts on wintertime sea-ice growth before the middle of this century. By comparison, the precipitous decline in September Arctic sea ice and the corresponding increases in Arctic open-water areas exert more evident promotional impacts on wintertime sea-ice growth during the same period by providing more space for sea-ice growth in the horizontal direction. Therefore, the newly formed sea-ice extent will increase dramatically before the middle of this century. Afterwards, the rates of decline in September Arctic sea ice become smaller, and thus its promotional impacts on sea-ice growth would not change a lot over time. In contrast, the continuous increases in wintertime Arctic temperatures lead to their suppression impacts on sea-ice growth strengthening over time. Under low-emissions (high-emissions) scenarios, i.e., SSP1-2.6 and SSP2-4.5 (SSP3-7.0 and SSP5-8.5), the suppression impacts on sea-ice growth exerted by rising wintertime Arctic temperatures offset (overwhelm) the promotional impacts on the sea-ice growth exerted by the decreasing September Arctic sea ice, which leads to the newly formed sea-ice extent keeping stable (decreasing) at the end of this century (colored curves in Fig. 5).

Figure 6 illustrates the changes (relative to the climatology during 1981–2010) of the newly formed sea-ice extent during two future periods: the mid-century period (2039–2068; top row) and end-of-century period (2069–2098; bottom row) (section 2.4). The mean and median of both the unweighted and weighted ensemble projections indicate that there is an increase in the newly formed sea-ice extent during both the two periods under all scenarios (bars and dots in Fig. 6). In the weighted case (right in each panel in Fig. 6), the increase has larger magnitude (blue bars and dots), and the intermodel spread is smaller (blue whiskers). As outlined above, during the mid-century period, wintertime temperatures in the Arctic region are relatively low, so that most regions in the Arctic are still fully ice-covered at the end of the freeze-up season (i.e., March). Therefore, the September Arctic sea-ice extent plays a predominant role in the change of the newly formed sea-ice extent. Since the September Arctic sea-ice extent tends to be smaller under higher-emissions scenarios, thereby providing more space for sea-ice growth in winter, the increase in newly formed sea-ice extent projected by the weighted mean (median) has the largest magnitude under the highest-emissions scenario (i.e., the SSP5-8.5 scenario), with a value of 3.03×10^6 km^2 (3.10×10^6 km^2), while it has the smallest magnitude under the lowest-emissions scenario (i.e., the SSP1-2.6 scenario), with a value of 2.44×10^6 km^2 (2.46×10^6 km^2). During the end-of-century period, the situation becomes much different (bottom row in Fig. 6). During this period, the Arctic has already reached an ice-free condition (i.e., sea-ice extent below 1×10^6 km^2) in September under medium-to-high-emissions scenarios (i.e., SSP2-4.5, SSP3-7.0, and SSP5-8.5) (Zhao et al., 2022b). Therefore, the space for sea-ice growth in winter is nearly identical under these scenarios. Since wintertime temperatures in the Arctic tend to be higher under higher-emissions scenarios, which will more intensely suppress the sea-ice growth, the mean (median) of the weighted projections indicates that the newly formed sea-ice extent has the smallest increase of 1.64×10^6 km^2 (1.84×10^6 km^2) under the highest-emissions scenario (i.e., the SSP5-8.5 scenario), while, it has the largest increase of 2.97×10^6 km^2 (3.05×10^6 km^2) under a medium-emissions scenario (i.e., the SSP2-4.5 scenario). Note that although there is no solid evidence to support the choice of the σ_s value within the weighting scheme (see section 2.5), it is found that the future changes of newly formed sea-ice extent projected by the weighted ensemble are insensitive to the choice of the σ_s value (Fig. 7), thereby indicating the results of this study derived by the weighting scheme are robust and convincing.

The spatial distribution of wintertime newly formed Arctic sea ice in the future is also a concern of this study, which provides valuable reference for future planning of marine activities in the Arctic Ocean. Therefore, we use the mean of the weighted ensemble to investigate the spatial patterns of changes (relative to the climatology during 1981–2010) in the newly formed sea-ice concentration (see section 2.2 for more details) during the aforementioned two future periods (Fig. 8; shading). The results indicate that there will be a higher newly formed sea-ice concentration in the central Arctic Ocean region (north of 75°N) regardless of the future period or emissions scenario. In addition, the newly formed sea-ice concentration will decrease in the outermost areas in the Arctic Ocean, including the Barents Sea, Greenland Sea, and Bering Sea. The mean wintertime air temperature in the central Arctic region projected by the mean of the weighted ensemble is low and below the freezing point during both future periods under even the highest emissions scenario (i.e., SSP5-8.5) (Fig. 8; blue, black, and red contours). Therefore, the magnitude of the increase in the newly formed sea-ice concentration over this region during each future period is mainly determined by the amount of summertime open-water areas, which determine the space for sea-ice growth in the subsequent winter. During each future period, there will be more summertime open-water areas in the central Arctic region under higher-emissions scenarios (pink contours in Fig. 8). Therefore, the increase in newly formed sea-ice concentration relative to the 1981–2010 baseline over this region has larger magnitude during each future period under

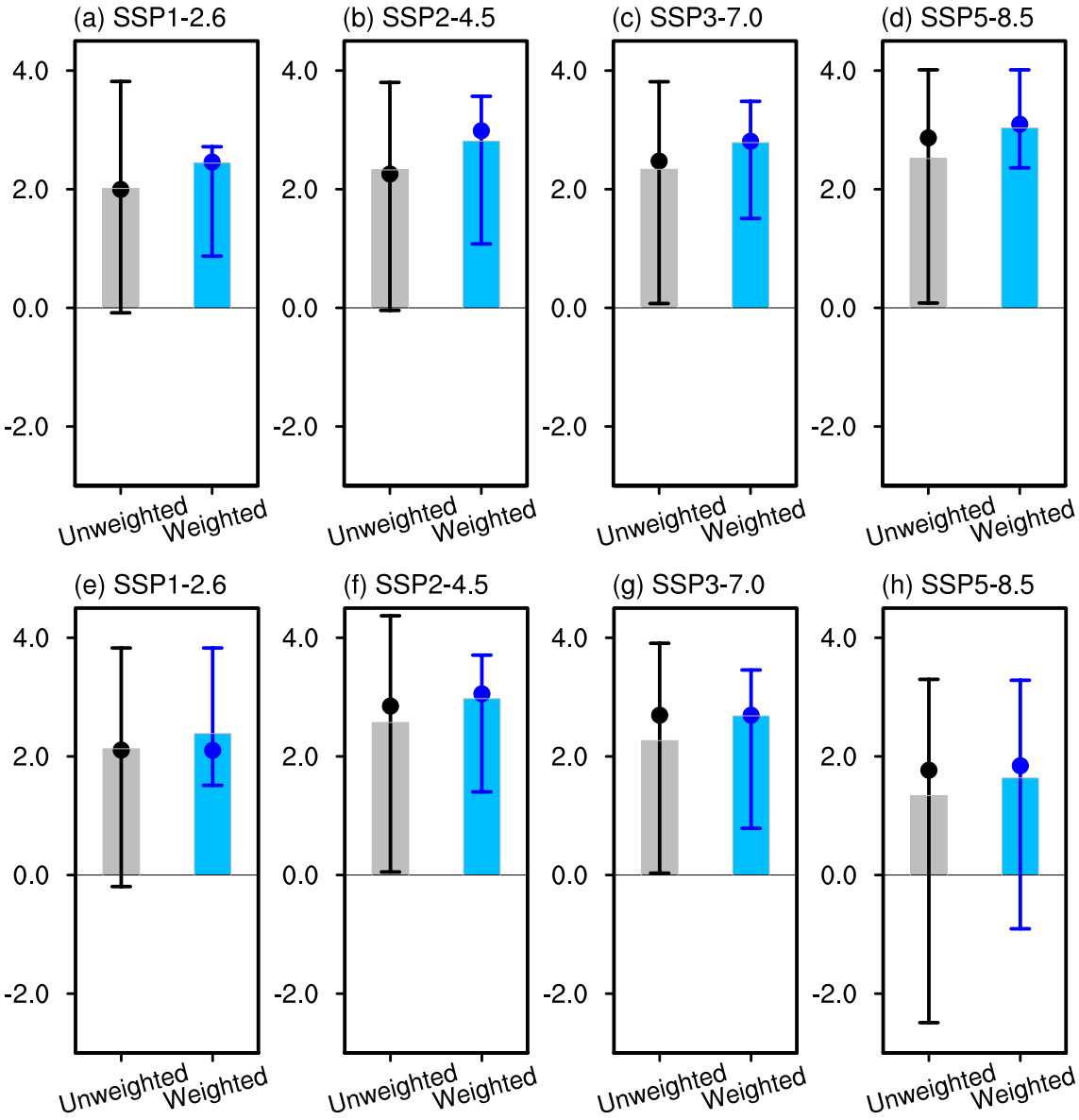

Fig. 6. Changes (relative to the climatology during 1981–2010) of wintertime newly formed sea-ice extent (units: 10^6 km^2) during (a–d) 2039–2068 and (e–h) 2069–2098 in the unweighted (grey bars, black whiskers and dots) and weighted (blue bars, whiskers and dots) CMIP6 multimodel projections under the SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 scenarios. In each panel, the bars (dots) represent the mean (median) of the multiple CMIP6 models, while the whiskers denote the range between the 10th and 90th percentiles of the individual models.

higher-emissions scenarios (Fig. 8; shading). One exception occurs during the end-of-century period under two high-emissions scenarios (i.e., SSP3-7.0 and SSP5-8.5), in which case the central Arctic Ocean is entirely free of sea ice in September (Figs. 8g and h). In this case, the space for sea-ice growth is identical under the two scenarios, and the difference in the magnitude of the increase in newly formed sea ice between the two scenarios comes mainly from their wintertime temperature differences. The higher winter temperatures under the SSP5-8.5 scenario lead to the increase in the newly formed sea-ice concentration having larger magnitude under the SSP3-7.0 scenario than under the SSP5-8.5 scenario (Figs. 8g and h; shading).

4. Discussion and conclusions

In this study, a weighting scheme, proposed to take into account both model performance and model independence, is applied to CMIP6 multimodel projections of wintertime newly formed Arctic sea ice. Results derived from the weighted multimodel ensemble are compared with those derived from the unweighted ensemble to evaluate the effect of the weighting scheme. Observations show that the wintertime newly formed Arctic sea-ice extent has experienced an interdecadal variation with a precipitous increasing trend since the mid-2000s. However, the climatology, the recent increasing trend, and the interannual variability of the newly

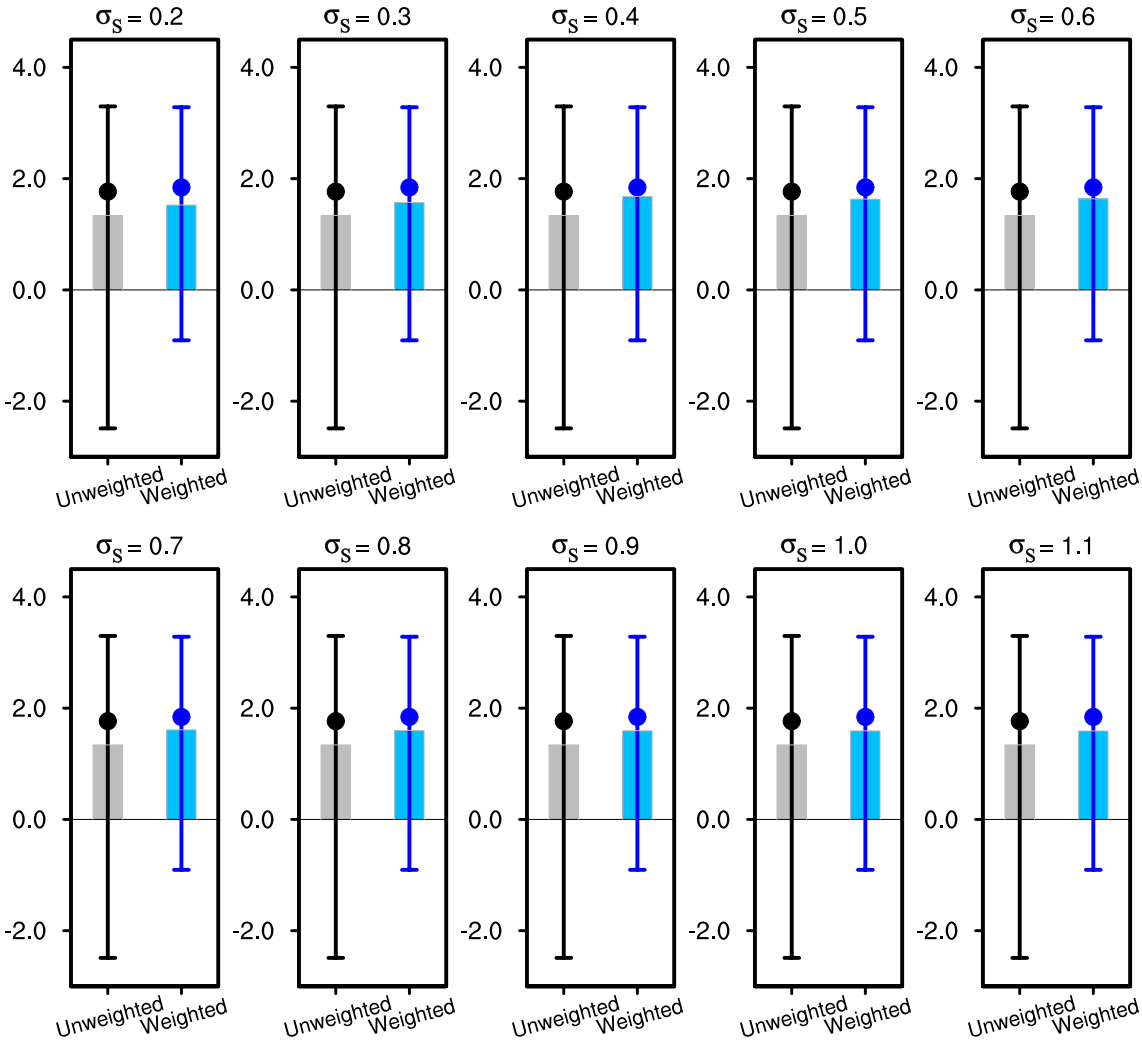

Fig. 7. Sensitivity of the results shown in Fig. 6h to the σ_s value. The σ_s used for each weighting is indicated in the title of each panel. For each σ_s, σ_D is determined as the smallest value that can make the case fraction outlined in section 2.5 meet the requirement of 0.8.

formed sea-ice extent shown in observations are all underestimated by the mean of the unweighted ensemble. This greatly affects the confidence level in future projections of the newly formed sea ice derived from the unweighted ensemble. It turns out that the bias in the unweighted simulations of the newly formed sea-ice extent is effectively reduced by the weighting scheme. The time series of the newly formed sea-ice extent simulated by the mean of the weighted ensemble agrees well with the observation. This is because models closer to the observation are reasonably assigned with larger weights by the scheme, and vice versa. The progress made by the weighting scheme gives more credit to using the weighted ensemble to project future changes of newly formed sea ice.

The intermodel spread is commonly used to quantify the uncertainty in climate projections. For the projections of newly formed sea-ice extent, the intermodel spread is evidently smaller in the weighted case than in the unweighted case, indicating that the uncertainty is efficiently constrained by the weighting scheme. This implies that the commonly used model democracy approach, i.e., simply giving equal weights to each model, may lead to overestimation of the actual uncertainty in projections of newly formed sea ice. The mean of the weighted ensemble projects that the recent increasing trend of the newly formed sea-ice extent is likely to continue until the middle of this century, regardless of the future emissions scenario. This is a result of increasing Arctic open-water areas in summer and autumn and the preserved freezing winter air temperatures in the Arctic. The magnitude of the increase in the newly formed sea-ice extent relative to the baseline period is sensitive to the emissions scenario. Before the middle of this century, the increase tends to be larger under higher-emissions scenarios, which is a result of a smaller September Arctic sea-ice extent. Afterwards, the aforementioned increasing trend of the newly formed sea-ice extent may remain stable if the Arctic warming is constrained significantly by the achievement of sustainable-development to middle-of-the-road scenarios (i.e., SSP1-2.6 and SSP2-4.5, respectively). Otherwise, the newly formed Arctic sea-ice extent may decrease if the Arctic warming crosses a

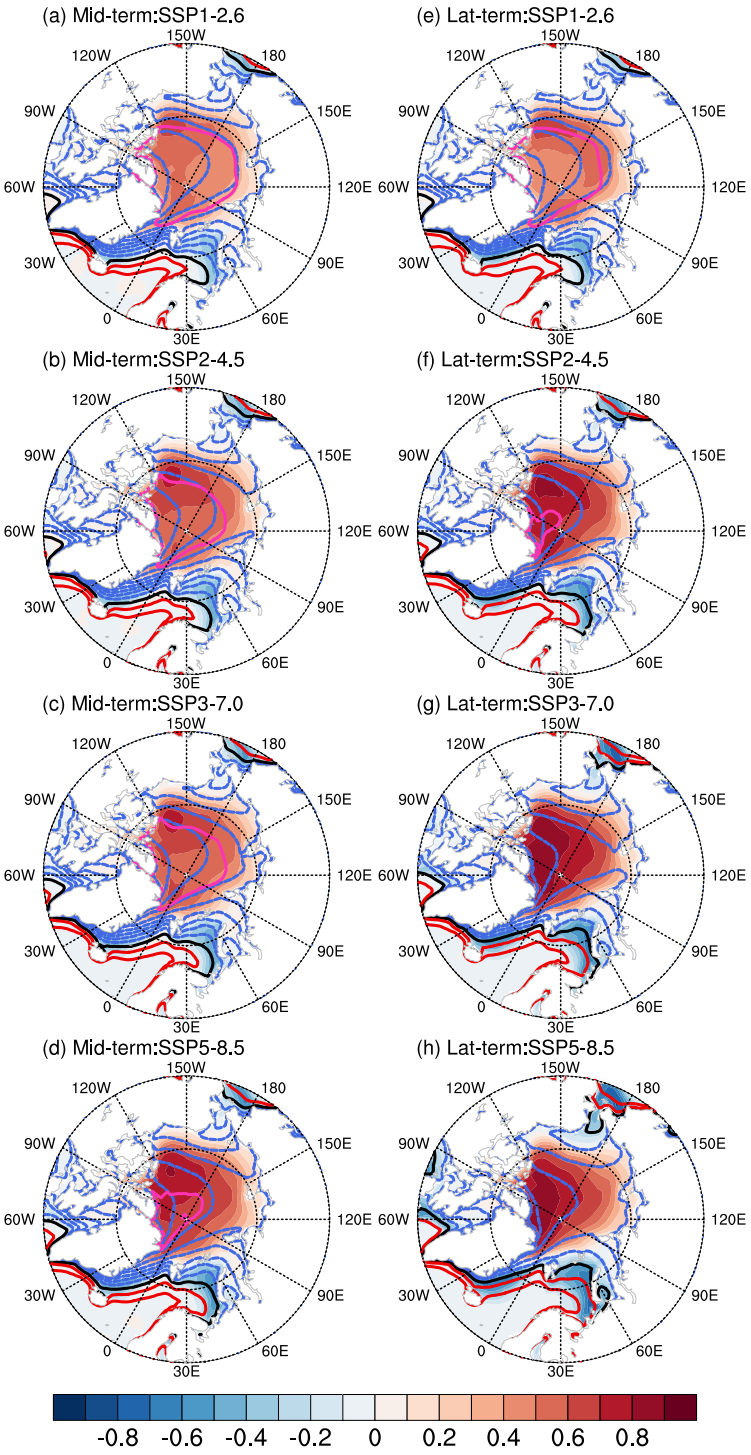

Fig. 8. (a–d) Spatial patterns of changes (relative to the climatology during 1981–2010) in the wintertime newly formed sea-ice concentration during 2039–2068 under the (a) SSP1-2.6, (b) SSP2-4.5, (c) SSP3-7.0, and (d) SSP5-8.5 scenarios (shading). (e–h) As in (a–d) but for changes during 2069–2098. Red, blue, and black contours in each panel denote the corresponding September–March mean surface air temperature averaged over the time period considered. Red (blue) contours indicate positive (negative) values. Black contours indicate the value of zero. The pink contour in each panel represents the ice edge that is represented by a sea-ice concentration value of 0.15. All results shown here (including both newly formed sea ice and air temperature) are derived from the mean of the weighted CMIP6 multimodel ensemble.

threshold. During the second half of this century, there is a smaller Arctic sea-ice extent in summer, and the rate of decline in summer Arctic sea-ice extent becomes much smaller than during the first half of this century. Therefore, the space for sea-ice growth provided by the summertime Arctic sea-ice decline remains stable over time. In contrast, the continued increases in wintertime temperatures in the Arctic lead to the suppression impacts exerted by the rising Arctic temperatures on sea-ice growth increasing over time. Under low-emissions (high-emissions) scenarios, the suppression impacts exerted by rising temperatures on sea-ice growth offset (overwhelm) the promotional impacts exerted by the summertime Arctic sea-ice decline on sea-ice growth, thereby leading to the newly formed sea-ice extent remaining stable (decreasing) at the end of this century. During this period, the increase in the newly formed sea-ice extent relative to the baseline period tends to be smaller under higher-emissions scenarios.

Future changes of the newly formed sea-ice concentration are further investigated using the mean of the weighted ensemble. Results show that there will be a higher (lower) newly formed sea-ice concentration over the central (outermost) areas in the Arctic Ocean in the future relative to the baseline period, regardless of the future period or emissions scenario. In the central Arctic region, the wintertime temperature remains low throughout the whole of the 21st century under even the highest emissions scenario, thereby exerting a weak suppression effect on sea-ice growth. Therefore, at the end of this century, unlike the newly formed sea-ice extent, which is partly affected by the rising wintertime temperature in the whole Arctic region, the newly formed sea-ice concentration in the central Arctic region is still mainly determined by the amount of open-water areas provided by the summertime sea-ice decline. Less Arctic sea ice in summer under higher-emissions scenarios leads to the future increase in the newly formed sea-ice concentration over this region tending to be larger under higher-emissions scenarios.

We clearly show in this work that the newly formed sea-ice projection skill can be evidently increased by weighting models by their related model components and their performance in reproducing observations. In our work, we have emphasized the importance of considering model independence when using large multimodel ensembles, such as those in the CMIPs, for climate projections. At this point, we use a recent study to help better understand the physical meaning of the model independence (Pan et al., 2023). This study pointed out that CMIP6 models that employ the same ocean model family (NEMO) produce similar future predictions of winter sea ice in the Arctic. Motivated by this study, we first selected all NEMO-family CMIP6 models used in our work (total of 10 models; first column in Table 2). Then, for each NEMO-family model, we identify a corresponding subset of this model's duplicates from the remaining models. The subset contains three models having the first, second and third smallest model-to-model distance S_{ij} value from this model (second, third and fourth column in Table 2, respectively). As shown in Table 2, all NEMO-family models except for IPSL-CM6A-LR have at least one NEMO-family model in their subset of duplicates (bold in Table 2). Note that the S_{ij} values are calculated based on models' historical simulations of wintertime newly formed sea ice (see section 2.4 for more details). Therefore, the results demonstrate that the NEMO-family models have large similarity in their simulations of wintertime newly formed sea ice, which supports the conclusions drawn by Pan et al. (2023). It also confirms the rationality of the weighting scheme used in our work.

Reliable projections of wintertime newly formed sea ice derived from the weighted ensemble indicate that there will be an increase in newly formed Arctic sea ice in the future. Increases in wintertime newly formed Arctic sea ice may have substantial impacts on northern mid–high latitude marine and climate systems. The enhanced brine rejection during the sea-ice formulation process may affect the mixed-layer properties of the Arctic Ocean and change the biogeochemical balance between the Arctic and the sub-Arctic marine system. Since newly formed sea ice has different physical properties from multiyear sea ice, an increase in the pro-

Table 2. Duplicates of models using the community ocean model "NEMO" as their ocean component (total of 10 models used in this study; first column). For each NEMO-family model, the ID of models having the first/second/third smallest model-to-model distance S_{ij} from this model are listed in the second/third/fourth column, respectively. The NEMO-family models are highlighted with bold font in their model ID.

Model name	Rank 1	Rank 2	Rank 3
CanESM5	**CanESM5-CanOE**	MPI-ESM1-2-HR	ACCESS-ESM1-5
CanESM5-CanOE	**CanESM5**	ACCESS-CM2	**CNRM-CM6-1-HR**
CNRM-CM6-1	MPI-ESM1-2-HR	FGOALS-f3-L	**CNRM-ESM2-1**
CNRM-CM6-1-HR	ACCESS-CM2	**CanESM5-CanOE**	**CanESM5**
CNRM-ESM2-1	FGOALS-f3-L	**CNRM-CM6-1**	**EC-Earth3-Veg**
EC-Earth3	INM-CM5-0	**CanESM5**	**UKESM1-0-LL**
EC-Earth3-Veg	TaiESM1	**CanESM5**	**EC-Earth3**
EC-Earth3-Veg-LR	**UKESM1-0-LL**	INM-CM4-8	INM-CM5-0
IPSL-CM6A-LR	MRI-ESM2-0	CESM2	FGOALS-f3-L
UKESM1-0-LL	**EC-Earth3-Veg-LR**	INM-CM5-0	INM-CM4-8

portion of newly formed sea ice in the total sea ice may make the Arctic sea ice in the future thinner, more breakable, and more sensitive to climate variability compared to the present-day conditions. This may lead to essential changes in Arctic atmosphere–ocean–ice interaction and large-scale atmospheric circulation, affecting mid–high latitude weather and climate. Furthermore, increases in newly formed sea ice can slow the long-term decline in summer total Arctic sea ice, thereby delaying the entrance of the Arctic into an ice-free state in summer. The results of our work (e.g., the weighted multimodel ensemble) lay a foundation for more follow-up research on the causes and consequences of future increases in wintertime newly formed Arctic sea ice.

Acknowledgements. This research was supported by the Chinese–Norwegian Collaboration Projects within Climate Systems jointly funded by the National Key Research and Development Program of China (Grant No. 2022YFE0106800) and the Research Council of Norway funded project, MAPARC (Grant No. 328943). We also acknowledge the support from the Research Council of Norway funded project, COMBINED (Grant No. 328935), the National Natural Science Foundation of China (Grant No. 42075030), and the Postgraduate Research and Practice Innovation Program of Jiangsu Province (KYCX23_1314).

Funding Note: Open Access funding provided by University of Bergen (incl Haukeland University Hospital) .

Author contributions S. P. HE and H. J. WANG designed the research. S. P. HE, J. Z. Zhao, K. FAN, and F. LI performed the research. J. Z. ZHAO prepared the manuscript with contributions from all co-authors.

Conflicts of interest The authors declare no conflicts of interest.

Data Availability Statement All the data analyzed in this study are openly available. The monthly mean sea-ice concentration is from the Met Office Hadley Centre Sea Ice and Sea Surface Temperature dataset at https://www.metoffice.gov.uk/hadobs/hadisst/data/download.html. Users should click on the link named "HadISST_ice.nc.gz" to download the compressed nc file. The monthly mean Arctic sea-ice extent is from the NSIDC Sea Ice Index, version 3, at https://nsidc.org/data/g02135/versions/3. CMIP6 simulations provided by ESGF can be found via the following open-source link: https://esgf-node.llnl.gov/search/cmip6/. Users should select the *Variable* as *siconc* and *tas*, which stand for sea-ice concentration and surface air temperature, respectively. Select the *Frequency* as *mon*; select the Table ID as *Simon* and *Amon*; select the *Experiment* ID as historical, ssp126, ssp245, ssp370 and ssp585; select the CMIP6 models employed in this study (see Table 1); and then download the nc files that appear as the search outputs.

REFERENCES

Annan, J. D., and J. C. Hargreaves, 2011: Understanding the CMIP3 multimodel ensemble. *J. Climate*, **24**(16), 4529–4538, https://doi.org/10.1175/2011JCLI3873.1.

Bitz, C. M., and G. H. Roe, 2004: A mechanism for the high rate of sea ice thinning in the Arctic Ocean. *J. Climate*, **17**(18), 3623–3632, https://doi.org/10.1175/1520-0442(2004)017<3623:AMFTHR>2.0.CO;2.

Brunner, L., R. Lorenz, M. Zumwald, and R. Knutti, 2019: Quantifying uncertainty in European climate projections using combined performance-independence weighting. *Environmental Research Letters*, **14**(12), 124010, https://doi.org/10.1088/1748-9326/ab492f.

Brunner, L., A. G. Pendergrass, F. Lehner, A. L. Merrifield, R. Lorenz, and R. Knutti, 2020: Reduced global warming from CMIP6 projections when weighting models by performance and independence. *Earth System Dynamics*, **11**(4), 995–1012, https://doi.org/10.5194/esd-11-995-2020.

Chen, W. L., Z. H. Jiang, and L. Li, 2011: Probabilistic projections of climate change over China under the SRES A1B scenario using 28 AOGCMs. *J. Climate*, **24**, 4741–4756, https://doi.org/10.1175/2011JCLI4102.1.

Collins, M., and Coauthors, 2013: Long-term climate change: Projections, commitments and irreversibility. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker et al., Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1029–1136.

Comiso, J. C., W. N. Meier, and R. Gersten, 2017: Variability and trends in the Arctic sea ice cover: Results from different techniques. *J. Geophys. Res.: Oceans*, **122**, 6883–6900, https://doi.org/10.1002/2017JC012768.

Deng, J. C., and A. G. Dai, 2022: Sea ice–air interactions amplify multidecadal variability in the North Atlantic and Arctic region. *Nature Communications*, **13**, 2100, https://doi.org/10.1038/s41467-022-29810-7.

Deser, C., A. Phillips, V. Bourdette, and H. Y. Teng, 2012: Uncertainty in climate change projections: The role of internal variability. *Climate Dyn.*, **38**, 527–546, https://doi.org/10.1007/s00382-010-0977-x.

Ding, Q. H., and Coauthors, 2017: Influence of high-latitude atmospheric circulation changes on summertime Arctic sea ice. *Nature Climate Change*, **7**(4), 289–295, https://doi.org/10.1038/nclimate3241.

Ding, Q. H., and Coauthors, 2019: Fingerprints of internal drivers of Arctic sea ice loss in observations and model simulations. *Nature Geoscience*, **12**(1), 28–33, https://doi.org/10.1038/s41561-018-0256-8.

Docquier, D., and T. Koenigk, 2021: Observation-based selection of climate models projects Arctic ice-free summers around 2035. *Communications Earth & Environment*, **2**(1), 144, https://doi.org/10.1038/s43247-021-00214-7.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, **9**(5), 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016.

Fatichi, S., and Coauthors, 2016: Uncertainty partition challenges the predictability of vital details of climate change. *Earth's Future*, **4**, 240–251, https://doi.org/10.1002/2015EF000336.

Fetterer, F., K. Knowles, W. N. Meier, M. Savoie, and A. K. Windnagel, 2017: Sea Ice Index, Version 3 [Data Set]. Boulder, Colorado USA. National Snow and Ice Data Center. https://doi.org/10.7265/N5K072F8.

Guo, H., A. M. Bao, T. Chen, G. X. Zheng, Y. Q. Wang, L. L. Jiang, and P. De Maeyer, 2021: Assessment of CMIP6 in simulating precipitation over arid Central Asia. *Atmospheric Research*, **252**, 105451, https://doi.org/10.1016/j.atmosres.2021.105451.

Hawkins, E., and R. Sutton, 2009: The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Meteor. Soc.*, **90**(8), 1095–1108, https://doi.org/10.1175/2009BAMS2607.1.

Hegyi, B. M., and P. C. Taylor, 2018: The unprecedented 2016-2017 Arctic sea ice growth season: The crucial role of atmospheric rivers and longwave fluxes. *Geophys. Res. Lett.*, **45**, 5204–5212, https://doi.org/10.1029/2017GL076717.

Herger, N., G. Abramowitz, R. Knutti, O. Angélil, K. Lehmann, and B. M. Sanderson, 2018: Selecting a climate model subset to optimise key ensemble properties. *Earth System Dynamics*, **9**(1), 135–151, https://doi.org/10.5194/esd-9-135-2018.

Herger, N., G. Abramowitz, S. Sherwood, R. Knutti, O. Angélil, and S. A. Sisson, 2019: Ensemble optimisation, multiple constraints and overconfidence: A case study with future Australian precipitation change. *Climate Dyn.*, **53**, 1581–1596, https://doi.org/10.1007/s00382-019-04690-8.

IPCC, 2021: Summary for policymakers. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, V. Masson-Delmotte et al., Eds., Cambridge University Press, 3–32.

Jahn, A., 2018: Reduced probability of ice-free summers for 1.5°C compared to 2°C warming. *Nature Climate Change*, **8**(5), 409–413, https://doi.org/10.1038/s41558-018-0127-8.

Kim, B. M., S. W. Son, S. K. Min, J. H. Jeong, S. J. Kim, X. D. Zhang, T. Shim, and J. H. Yoon, 2014: Weakening of the stratospheric polar vortex by Arctic sea-ice loss. *Nature Communications*, **5**, 4646, https://doi.org/10.1038/ncomms5646.

Knutti, R., D. Masson, and A. Gettelman, 2013: Climate model genealogy: Generation CMIP5 and how we got there. *Geophys. Res. Lett.*, **40**(6), 1194–1199, https://doi.org/10.1002/grl.50256.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl, 2010: Challenges in combining projections from multiple climate models. *J. Climate*, **23**(10), 2739–2758, https://doi.org/10.1175/2009JCLI3361.1.

Knutti, R., J. Sedláček, B. M. Sanderson, R. Lorenz, E. M. Fischer, and V. Eyring, 2017: A climate model projection weighting scheme accounting for performance and interdependence. *Geophys. Res. Lett.*, **44**(4), 1909–1918, https://doi.org/10.1002/2016GL072012.

Kwok, R., 2018: Arctic sea ice thickness, volume, and multiyear ice coverage: Losses and coupled variability (1958–2018). *Environmental Research Letters*, **13**(10), 105005, https://doi.org/10.1088/1748-9326/aae3ec.

Laxon, S. W., and Coauthors, 2013: CryoSat-2 estimates of Arctic sea ice thickness and volume. *Geophys. Res. Lett.*, **40**, 732–737, https://doi.org/10.1002/grl.50193.

Lee, W.-L., and Coauthors, 2020: Taiwan Earth System Model Version 1: Description and evaluation of mean state. *Geoscientific Model Development*, **13**(9), 3887–3904, https://doi.org/10.5194/gmd-13-3887-2020.

Liu, J. P., M. R. Song, Z. Zhu, R. M. Horton, Y. Y. Hu, and S. P. Xie, 2022: Arctic sea-ice loss is projected to lead to more frequent strong El Niño events. *Nature Communications*, **13**, 4952, https://doi.org/10.1038/s41467-022-32705-2.

Liu, Z. F., and Coauthors, 2021: Acceleration of western Arctic sea ice loss linked to the Pacific North American pattern. *Nature Communications*, **12**, 1519, https://doi.org/10.1038/s41467-021-21830-z.

Lorenz, R., N. Herger, J. Sedláček, V. Eyring, E. M. Fischer, and R. Knutti, 2018: Prospects and Caveats of Weighting Climate Models for Summer Maximum Temperature Projections Over North America. *J. Geophys. Res.: Atmos.*, **123**(9), 4509–4526, https://doi.org/10.1029/2017jd027992.

Maher, N., F. Lehner, and J. Marotzke, 2020: Quantifying the role of internal variability in the temperature we expect to observe in the coming decades. *Environmental Research Letters*, **15**, 054014, https://doi.org/10.1088/1748-9326/ab7d02.

Masson, D., and R. Knutti, 2013: Predictor screening, calibration, and observational constraints in climate model ensembles: An illustration using climate sensitivity. *J. Climate*, **26**(3), 887–898, https://doi.org/10.1175/JCLI-D-11-00540.1.

Niederdrenk, A. L., and D. Notz, 2018: Arctic sea ice in a 1.5°C warmer world. *Geophys. Res. Lett.*, **45**(4), 1963–1971, https://doi.org/10.1002/2017GL076159.

Notz, D., and SIMIP Community, 2020: Arctic sea ice in CMIP6. *Geophys. Res. Lett.*, **47**(10), e2019GL086749, https://doi.org/10.1029/2019gl086749.

Olonscheck, D., T. Mauritsen, and D. Notz, 2019: Arctic sea-ice variability is primarily driven by atmospheric temperature fluctuations. *Nature Geoscience*, **12**, 430–434, https://doi.org/10.1038/s41561-019-0363-1.

O'Neill, B. C., and Coauthors, 2016: The scenario model Intercomparison project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, **9**(9), 3461–3482, https://doi.org/10.5194/gmd-9-3461-2016.

Pan, R. R., Q. Shu, Q. Wang, S. Z. Wang, Z. Y. Song, Y. He, and F. L. Qiao, 2023: Future Arctic climate change in CMIP6 strikingly intensified by NEMO-family climate models. *Geophys. Res. Lett.*, **50**, e2022GL102077, https://doi.org/10.1029/2022GL102077.

Perkins, S. E., A. J. Pitman, N. J. Holbrook, and J. McAneney, 2007: Evaluation of the AR4 Climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *J. Climate*, **20**(17), 4356–4376, https://doi.org/10.1175/JCLI4253.1.

Petty, A. A., M. M. Holland, D. A. Bailey, and N. T. Kurtz, 2018: Warm Arctic, increased winter sea ice growth?. *Geophys. Res. Lett.*, **45**, 12 922–12 930, https://doi.org/10.1029/

2018GL079223.

Petty, A. A., N. T. Kurtz, R. Kwok, T. Markus, and T. A. Neumann, 2020: Winter Arctic sea ice thickness from ICESat-2 freeboards. *J. Geophys. Res.: Oceans*, **125**, e2019JC015764, https://doi.org/10.1029/2019JC015764.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.: Atmos.*, **108**, 4407, https://doi.org/10.1029/2002JD002670.

Renner, A. H. H., S. Gerland, C. Haas, G. Spreen, J. F. Beckers, E. Hansen, M. Nicolaus, and H. Goodwin, 2014: Evidence of Arctic sea ice thinning from direct observations. *Geophys. Res. Lett.*, **41**, 5029–5036, https://doi.org/10.1002/2014GL060369.

Ricker, R., F. Kauker, A. Schweiger, S. Hendricks, J. L. Zhang, and S. Paul, 2021: Evidence for an increasing role of ocean heat in Arctic winter sea ice growth. *J. Climate*, **34**(13), 5215–5227, https://doi.org/10.1175/JCLI-D-20-0848.1.

Sanderson, B. M., R. Knutti, and P. Caldwell, 2015a: Addressing interdependency in a multimodel ensemble by interpolation of model properties. *J. Climate*, **28**(13), 5150–5170, https://doi.org/10.1175/JCLI-D-14-00361.1.

Sanderson, B. M., R. Knutti, and P. Caldwell, 2015b: A representative democracy to reduce interdependency in a multimodel ensemble. *J. Climate*, **28**(13), 5171–5194, https://doi.org/10.1175/JCLI-D-14-00362.1.

Shiru, M. S., and E.-S. Chung, 2021: Performance evaluation of CMIP6 global climate models for selecting models for climate projection over Nigeria. *Theor. Appl. Climatol.*, **146**(1), 599–615, https://doi.org/10.1007/s00704-021-03746-2.

Stroeve, J., and D. Notz, 2018: Changing state of Arctic sea ice across all seasons. *Environmental Research Letters*, **13**(10), 103001, https://doi.org/10.1088/1748-9326/aade56.

Stroeve, J. C., T. Markus, L. Boisvert, J. Miller, and A. Barrett, 2014: Changes in Arctic melt season and implications for sea ice loss. *Geophys. Res. Lett.*, **41**, 1216–1225, https://doi.org/10.1002/2013GL058951.

Tebaldi, C., and R. Knutti, 2007: The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **365**(1857), 2053–2075, https://doi.org/10.1098/rsta.2007.2076.

van Vuuren, D. P., and Coauthors, 2011: The representative concentration pathways: An overview. *Climatic Change*, **109**, 5–31, https://doi.org/10.1007/s10584-011-0148-z.

Wernli, H., and L. Papritz, 2018: Role of polar anticyclones and mid-latitude cyclones for Arctic summertime sea-ice melting. *Nature Geoscience*, **11**, 108–113, https://doi.org/10.1038/s41561-017-0041-0.

Zhao, J. Z., S. P. He, and H. J. Wang, 2022a: Historical and future runoff changes in the Yangtze River Basin from CMIP6 models constrained by a weighting strategy. *Environmental Research Letters*, **17**(2), 024015, https://doi.org/10.1088/1748-9326/ac3f61.

Zhao, J. Z., S. P. He, and H. J. Wang, 2023: Role of atmosphere–ocean–ice interaction in the linkage between December Bering Sea ice and subsequent February surface air temperature over North America. *J. Climate*, **36**, 1679–1696, https://doi.org/10.1175/JCLI-D-22-0265.1.

Zhao, J. Z., S. P. He, H. J. Wang, and F. Li, 2022b: Constraining CMIP6 projections of an ice-free Arctic using a weighting scheme. *Earth's Future*, **10**, e2022EF002708, https://doi.org/10.1029/2022EF002708.