2021-09

EVALUATION OF THE COUPLED MODEL INTERCOMPARISON PROJECT PHASE 6 (CMIP6) HISTORICAL SIMULATIONS OF THE ARCTIC SEA ICE: PROGRESS, LIMITATIONS AND THEIR CAUSES

Watts, Matthew N.

Monterey, CA; Naval Postgraduate School

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

DISSERTATION

EVALUATION OF THE COUPLED MODEL INTERCOMPARISON PROJECT PHASE 6 (CMIP6) HISTORICAL SIMULATIONS OF THE ARCTIC SEA ICE: PROGRESS, LIMITATIONS AND THEIR CAUSES

by

Matthew N. Watts

September 2021

Dissertation Supervisor: Wieslaw Maslowski

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2021	**3. REPORT TYPE AND DATES COVERED** Dissertation

4. TITLE AND SUBTITLE
EVALUATION OF THE COUPLED MODEL INTERCOMPARISON PROJECT PHASE 6 (CMIP6) HISTORICAL SIMULATIONS OF THE ARCTIC SEA ICE: PROGRESS, LIMITATIONS AND THEIR CAUSES

5. FUNDING NUMBERS

6. AUTHOR(S) Matthew N. Watts

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release. Distribution is unlimited.

12b. DISTRIBUTION CODE
A

13. ABSTRACT (maximum 200 words)

A decline of the Arctic sea ice in response to a warming climate is assessed in the historical sea ice simulations from state-of-the-art global climate models participating in Phase 6 of the Coupled Model Intercomparison Project (CMIP6). Accurate simulations of sea ice are important for projections of its regional and global effects on the air-sea exchanges, weather, and climate. The timing and rate of simulated sea ice decline is compared with available observations for sea ice area and volume. Analysis indicates multi-model means and long-term trends for these common sea ice metrics are well represented, but the individual CMIP6 model ability to represent the observed accelerated rate in sea ice decline remains a challenge. Local and regional sea ice biases are identified through spatial analysis metrics, like sea ice thickness distribution pattern and sea ice edge analysis. Large model spatial errors imply limitations in or lack of representation of some key physical processes. The oceanic heat transport (OHT) and its forcing of the pan-Arctic sea ice decline are examined as possible model limitations. CMIP6 models show a strong correlation between increasing OHT and decreasing sea ice trends but likely underestimate the northward OHT over the polar cap (70°–90°N). Isolating specific model limitations and identifying possible processes affecting them will guide future model improvements critical to our understanding and projection of Arctic climate change.

14. SUBJECT TERMS
Arctic Ocean, Arctic sea ice, Arctic amplification, Earth system models, global climate models, oceanic forcing, ocean heat convergence, Coupled Model Intercomparison Project, CMIP6, model biases, sea ice thickness, SIT, oceanic heat transport, OHT, sea ice area, SIA, sea ice extent, SIE, sea ice volume, SIV

15. NUMBER OF PAGES
171

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

EVALUATION OF THE COUPLED MODEL INTERCOMPARISON PROJECT PHASE 6 (CMIP6) HISTORICAL SIMULATIONS OF THE ARCTIC SEA ICE: PROGRESS, LIMITATIONS AND THEIR CAUSES

Matthew N. Watts
Commander, United States Navy
BS, U.S. Naval Academy, 2004
MS, Massachusetts Institute of Technology, 2006

Submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY IN PHYSICAL OCEANOGRAPHY

from the

NAVAL POSTGRADUATE SCHOOL
September 2021

Approved by:

Wieslaw Maslowski
Department of Oceanography
Dissertation Supervisor

Younjoo Lee
Department of Oceanography

Tom Murphree
Department of Meteorology

Mara S. Orescanin
Department of Oceanography

Timothy P. Stanton
Department of Oceanography

Wieslaw Maslowski
Department of Oceanography
Dissertation Chair

Approved by:

Peter C. Chu
Chair, Department of Oceanography

Michael E. Freeman
Vice Provost of Academic Affairs

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

A decline of the Arctic sea ice in response to a warming climate is assessed in the historical sea ice simulations from state-of-the-art global climate models participating in Phase 6 of the Coupled Model Intercomparison Project (CMIP6). Accurate simulations of sea ice are important for projections of its regional and global effects on the air-sea exchanges, weather, and climate. The timing and rate of simulated sea ice decline is compared with available observations for sea ice area and volume. Analysis indicates multi-model means and long-term trends for these common sea ice metrics are well represented, but the individual CMIP6 model ability to represent the observed accelerated rate in sea ice decline remains a challenge. Local and regional sea ice biases are identified through spatial analysis metrics, like sea ice thickness distribution pattern and sea ice edge analysis. Large model spatial errors imply limitations in or lack of representation of some key physical processes. The oceanic heat transport (OHT) and its forcing of the pan-Arctic sea ice decline are examined as possible model limitations. CMIP6 models show a strong correlation between increasing OHT and decreasing sea ice trends but likely underestimate the northward OHT over the polar cap (70°–90°N). Isolating specific model limitations and identifying possible processes affecting them will guide future model improvements critical to our understanding and projection of Arctic climate change.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AGP	Arctic and Global Prediction
AO	Arctic Oscillation
AR1	lag-1 temporal autocorrelation
AR6	Sixth Assessment Report
ARCSS	Arctic System Science
BS	Bootstrap
CAA	Canadian Arctic Archipelago
CDR	climate data record
CGLORS	Centro Euro-Mediterraneo sui Cambiamenti Climatici coupled ocean-sea ice reanalysis
CMIP	Coupled Model Intercomparison Project
DOD	Department of Defense
DOE	Department of Energy
ESM	Earth System Model
ESGF	Earth System Grid Federation
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
GHG	greenhouse gas
HPCMP	High Performance Computer Modernization Program
IIEE	integrated ice-edge error
IPCC	Intergovernmental Panel on Climate Change
ITP	Ice-Tethered Profiler
JRA-55	Japanese 55-year atmospheric reanalysis
LIDAR	Light Detection and Ranging
MIZ	marginal ice zone
MM	multimodel mean
MYI	multiyear ice
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
NSF	National Science Foundation

NSIDC	National Snow and Ice Data Center
NT	National Aeronautics and Space Administration Team
ONR	Office of Naval Research
ORAS5	Ocean Reanalysis System 5
OSI-SAF	Ocean and Sea Ice Satellite Application Facility
PIOMAS	Pan-Arctic Ice Ocean Modeling and Assimilation System
RASM	Regional Arctic System Model
RASM-G	Regional Arctic System Model sea ice-ocean simulation
RGMA	Regional and Global Model Analysis
RMSE	root-mean-square error
SAT	surface air temperature
SIC	sea ice concentration
SIE	sea ice extent
SIMIP	Sea Ice Model Intercomparison Project
SIPS	Sea Ice Polar Stereographic
SIT	sea ice thickness
SIV	sea ice volume
SPS	Spatial Probability Score
SPS-O	Spatial Probability Score - overestimation
SPS-U	Spatial Probability Score – underestimation
TOA	top of the atmosphere
WCRP	World Climate Research Programme
uRMSD	unbiased root-mean-square difference

ACKNOWLEDGMENTS

First and foremost, I would like to thank my wonderful wife and children, Dawn, Jacob, and Lila, for their encouragement, sacrifice, and love. The life of a military family is filled with fantastic experiences during the best of times, and immense challenges during the worst. The last few years have seen both, and this personal accomplishment would not be possible without them. Thank you.

I am grateful to the U.S. Navy for supporting my education and academic growth. I understand that time is a precious resource for any naval officers' career and granting five years of postgraduate education out of the last 17 is no small act. Thank you to the faculty and staff at Naval Postgraduate School (NPS), particularly every member in the Department of Oceanography. They are committed to the best graduate student experience for all. A special thanks to my PhD committee members for taking on the task. My advisor, Dr. Wieslaw Maslowski, provided great mentorship and steady encouragement throughout the program. Dr. Younjoo Lee was a constant sounding board for new ideas and fleshing out scientific approaches to my research questions. I am grateful to Dr. Tim Stanton for the opportunity to complete Arctic field research onboard USCGC Healy with fellow NPS scientists Chris Miller and Keith Wyckoff. Thank you to Dr. Mara Orescanin and Dr. Tom Murphree for the thoughtful questions and commitment to my development.

The following acknowledgements are but the top of a long list of professional and academic leaders who have left a lasting impact and offered me encouragement to complete this ultimate level of education. Thank you to Vice Admiral Jon Alexander (Ret.), Rear Admiral Jon White (Ret.), Captain Nick Vincent (Ret.), Captain Rip Coke, Captain Dwight Smith, and Dr. Bruce Morris. Thank you to my master's degree thesis advisors Dr. Michael Triantafyllou and Dr. Franz Hover and to my undergraduate advisors Dr. Kevin Smith and Dr. Ted Maksym.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

The Arctic climate is changing in response to greenhouse gas (GHG) forcing and the sea ice cover is declining at a rate faster than has been predicted by most state-of-the-art Earth System and global climate models (hereafter called climate models) (IPCC 2014, 2019). With the potential for increased economic and military activity as a result of extended periods of ice-free Arctic and subarctic waterways in the future, it is of strategic importance to the United States to obtain credible information from climate predictions and projections (Department of Defense 2019; Department of the Navy 2021). To accomplish this, it is first necessary to understand and represent the underlying key physical processes and resulting feedbacks in climate model simulations of the historical climate record. Improved climate model simulations of past states are a first step toward improving model projections of future states. Many studies have documented the representation of Arctic sea ice in different phases of the Coupled Model Intercomparison Project (CMIP) simulations, but few have focused on diagnosing the causality of climate model local and regional biases (Stroeve et al. 2012; SIMIP Community 2020; Shu et al. 2020; Shen et al. 2021).

A. MOTIVATION

1. Earth's Climate Is Changing

Earth's climate is changing at an unprecedented rate, driven largely by increased levels of carbon dioxide and other GHGs (IPCC 2014). According to the Intergovernmental Panel on Climate Change (IPCC), anthropogenic radiative forcing is manifested in a number of observed physical changes in the environment (IPCC 2014). Since around 1900, nearly our whole planet has experienced surface warming, and the combined land and ocean surface air temperature (SAT) has warmed by 0.85 °C between 1880-2012 (IPCC 2014). Over roughly the same period, global mean sea level has risen by about 20 cm, resulting mostly from glacier melt and ocean thermal expansion. Additionally, the global near surface ocean temperature (i.e., upper 75 m) has warmed by 0.11 °C during the period 1971-2010 (IPCC 2014). It is the ocean which accounts for more than 90% of the additional

energy stored within Earth's climate system, with 60% of the net energy increase stored within the upper 700m during the period 1971-2010 (IPCC 2014).

Even more striking than the general global warming trends is the much larger surface warming experienced in the Arctic. Here the warming rate exceeds the rest of the planet by a factor of 2 to 3 (Serreze and Barry 2011). This so-called Arctic amplification (Serreze and Francis 2006) of SATs is prominently linked to the rapidly diminishing Arctic sea ice cover by way of positive feedback processes. One such process is the ice-albedo feedback, whereby the loss of highly reflective sea ice reduces the surface albedo, thus increasing absorption of shortwave radiation and available heat in the upper ocean commanding further sea ice melt (Perovich et al. 2008; Serreze and Barry 2011). The net result is a larger ocean surface area containing ice-free conditions over a greater number of days. Longer periods of seasonally ice-free conditions in the Arctic Ocean and its approaches allow for increased access to the Arctic region (Melia et al. 2016). Global competition for control and influence across the region is of strategic importance to the United States.

2. Persistent *in situ* Arctic Observing Systems Are Scarce

Observational studies describing the rapid changes of the Arctic climate are critical to advance knowledge of causality and predict the future state. However, a number of challenges impede *in situ* collection. Access to much of the Arctic Ocean and Northern hemisphere cryosphere is limited by way of geopolitical obstacles (e.g., Russian policy), as well as the inhospitable climate itself (Uotila et al. 2019). For instance, ubiquitous winter sea ice cover over the Arctic Ocean limits shipboard access and inhibits widespread oceanographic data collection (e.g., Argo-buoys) during a large portion of the year (Uotila et al. 2019). Extreme temperatures, weather, and high risk to personnel make manned ice camps largely untenable, so the limited observation stations for ocean and atmosphere data must operate autonomously for extensive periods.

Most Arctic *in situ* observations are limited and discrete in both time and space. But they are the geophysical ground truth for the physical processes occurring within the environment. As well they are essential in constraining historical climate model simulations and building confidence in the complex simulated processes therein. Satellite-

derived observations may serve as a better reference for intercomparisons among sea ice model simulations (Markus et al. 2009), but some of the products are only available after 2010 during the freeze-up season normally from October to April.

3. Climate Models Fill Data Gaps

Global climate models are uniquely capable to complement many of the observational systems listed above. They provide "access" to all regions of the Arctic on dependable time and spatial scales (NOAA 2021). Climate models inform the projections of future climate states, weather, and environmental hazards (NOAA 2021). Their role is also central in shaping policy and infrastructure preparedness (e.g., IPCC 2014). In the United States alone, no less than four government organizations operate global climate models (i.e., NOAA, NASA, DOE, NSF), commanding a great number of resources (Golaz et al. 2019; Held et al. 2019; Danabasoglu et al. 2020; Kelley et al. 2020). Climate models have long been used as a tool to better understand and predict the inner workings of the Earth system (e.g., Mahlstein and Knutti 2011; DeRepentigny et al. 2020; Im et al. 2021). But the physics of the Earth system are complicated and complex (e.g., Kirtman et al. 2012; Maslowski et al. 2012). Accordingly, simplifications must be made, and parameters manipulated to best represent the physical reality (e.g., Roach et al. 2018). Some of these choices and limitations result in systematic model biases, e.g., cold SAT bias in high latitudes (e.g., Davy and Esau 2014). Furthermore, many climate model simulations are not necessarily tuned for the Arctic climate, and therefore may not represent observed feedback processes (e.g., Urrego-Blanco et al. 2016).

B. NAVAL RELEVANCE

1. United States Arctic Strategy

The Arctic is defined as the region north of the Arctic Circle and is depicted along with the pan-Arctic regional geography in Figure 1. The United States is an Arctic nation by way of Alaska and by treaty with Canada (Department of the Navy 2021). Strategic documentation calling for preparations for climate change and to operate United States forces in a more accessible Arctic region has increased over the last decade. The Arctic region was specifically included in the National Security Strategy 2010 and 2015, as well as

making new entries as standalone guidance in the National Strategy for the Arctic Region 2013, and Department of Defense (DOD) Arctic Strategy 2016 and 2019. When questioned by United States Congress as to the need for such rapid updates to Arctic strategies, the Secretary of the Navy Richard Spencer replied, "The damn thing melted" (Werrell and Femia 2018). The urgency reflected by Secretary Spencer's comment has echoed across all service components, with each branch having published a recent Arctic strategy to date.

Figure 1. Map of pan-Arctic regional geography. Source: Department of the Navy (2021).

The DOD Arctic Strategy 2019 outlined three ways to support the end-state for the Arctic as "a secure and stable region in which United States national security interests are safeguarded, the United States homeland is defended, and nations work cooperatively to address shared challenges" (Department of Defense 2019). They include (Department of Defense 2019)

- Building Arctic awareness;
- Enhancing Arctic operations; and,
- Strengthening the rules-based order in the Arctic.

2. United States Navy Arctic Strategy

According to O'Rourke et al. (2019), "in support of National and DOD objectives for the Arctic, the Navy will… defend United States sovereignty" and preserve freedom of the seas. The United States Navy strategy, *A Blue Arctic*, nested the following three naval objectives in 2021 to meet those listed above (Department of the Navy 2021)

- Maintain enhanced presence;
- Strengthen cooperative partnerships; and,
- Build a more capable Arctic naval force.

The Navy must operate further north and with increasing regularity in the coming decades. The lengthening of summer melt seasons (Markus et al. 2009) increases the navigability of the Arctic sea routes (i.e., Northern Sea Route, Northwest Passage, and Transpolar Route) and results in more human activity in and around the Arctic regions (Melia et al. 2016). Figure 2 illustrates the Arctic transit sea routes and the anticipated sea ice extent minima over the next decade. These Arctic shipping routes are most commercially enticing between the European and Asian markets, whereby transit time and distance may be reduced by as much as 40% (e.g., from 30 days to 18 days) compared to the primary alternative Suez canal route (Melia et al. 2016). Furthermore, natural resource exploration and commercial fishing over the vast Arctic continental shelf will likely increase over the near future (Department of the Navy 2021; Task Force Climate Change 2014). Thus, knowledge of the timing for access to these regions holds strategic importance.

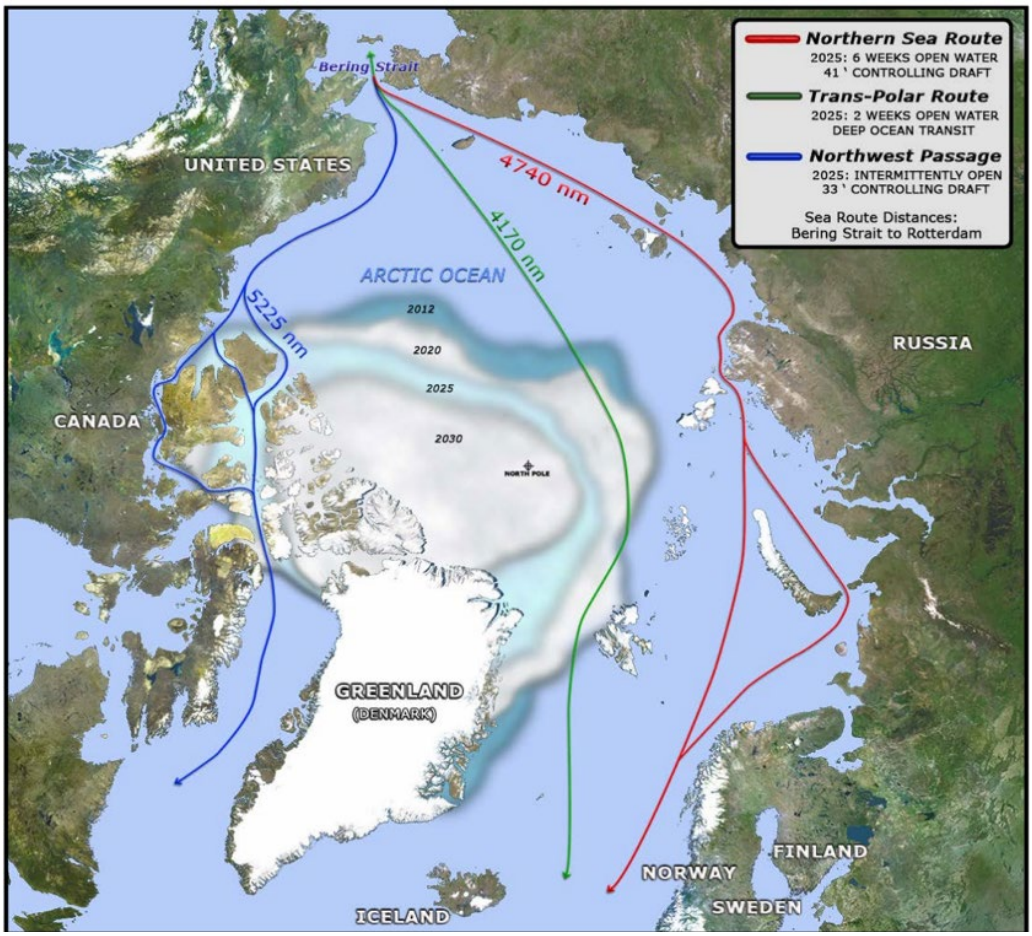


Figure 2. Arctic sea routes and the anticipated sea ice minima through 2030. Source: Task Force Climate Change (2014).

The Navy strategy clearly articulates the expected speed at which great power competition will jockey for advantage in the region. For example, large investments have already been made in Russia's Arctic defense, as well as posturing for unlawful regulation of the Northern Sea Route (Department of the Navy 2021). Additionally, Chinese investments in polar-capable cargo ships, icebreakers, and accompanying port infrastructure foreshadow their own Arctic ambition (Department of the Navy 2021). Because of the potential for future competition and conflict in the more accessible Arctic, the United States must work closely with allies and partners, both military and civilian, to improve regional understanding of the environment and the operating capabilities of the force (Department of the Navy 2021).

According to the Department of the Navy (2021), "understanding and predicting the physical environment from sea floor to space, today and for decades, is critical for mission advantage." To build a more Arctic capable naval force, the Navy emphasizes professional military education to "deepen our knowledge of current and future challenges" conducting Arctic operations (Department of the Navy 2021). Additionally, research is needed to understand the underlying key physical processes which will reduce model uncertainty in the seasonal and sub-seasonal sea ice predictions, and future projections of climate change impacts on sea ice.

C. RESEARCH GOALS

Our research was conducted to support United States Navy objectives and address science questions from CMIP6 Sea Ice Model Intercomparison Project (SIMIP) (Notz et al. 2016). To do so, we assessed CMIP6 simulations of sea ice with a focus on the role of oceanic forcing in the generation of sea ice biases. Understanding the causality of sea ice biases may inform modeling groups of focused improvement areas and caution against the use of particular models for certain process studies. Questions that were posed by SIMIP (Notz et al. 2016) and foundational for this research include the following: (a) Why do CMIP6 sea ice simulations differ from each other and/or the observational records? (b) Do CMIP6 sea ice simulations have systematic regional biases in the distribution of the sea ice pack as a result of oceanographic influence? This dissertation has two primary research objectives and several underlying science questions:

1. Assess progress and outstanding limitations in the historical sea ice simulations from the state-of-the-art global climate models participating in the CMIP6.
    - How well do CMIP6 models simulate the 1979-2014 observed mean sea ice state and trends?
    - Do CMIP6 models simulate the apparent accelerated rate of sea ice decline shown in observations?

- Can new spatial analysis metrics of sea ice provide model bias information beyond that from standard time series metrics?

2. Examine the role of oceanic forcing (e.g., oceanic heat transport (OHT) and convergence) in causing local and regional biases in Arctic sea ice in individual CMIP6 models.

    - How well do CMIP6 models simulate the OHT shown in the limited *in situ* observations and in ocean reanalyses?

    - Are biases in CMIP6 sea ice simulations related to the simulated OHT?

D. STRUCTURE OF THIS DISSERTATION

Following this introduction is a background chapter covering the Arctic region, the observed decline in sea ice, and a review of recent CMIP6 sea ice studies. Next, three "results" chapters are presented in journal article format, each with its own distinct sections including an introduction, data and methods, results, discussion and conclusion. The first paper, "An Assessment of the Pan-Arctic Accelerated Rate of Sea Ice Decline in CMIP6 Historical Simulations" (in review), examines the full suite of CMIP6 models and addresses temporal variability in sea ice area and volume simulations. The second paper, "A Spatial Evaluation of Arctic Sea Ice and Regional Limitations in CMIP6 Historical Simulations" (Watts et al. 2021), evaluates spatial variability in sea ice thickness and extent in order to identify persistent regional biases in CMIP6 models. Chapter V is my contribution to a coauthored paper in preparation, "Assessment of Oceanic Heat Transport and Its Effects on Pan-Arctic Sea Ice Decline in CMIP6 Historical Simulations", which examines the role of OHT on the sea ice cover in CMIP6 models. A summary of conclusions and recommendations for future research is presented in Chapter VI.

II. BACKGROUND

A. ARCTIC SEA ICE

The region of the Earth that is characterized by water in the frozen form is called the cryosphere, which includes glaciers, permafrost, snow, and sea ice. Sea ice is the frozen water that is created by cooling of the ocean sea surface to the freezing temperature. The presence of sea ice floating on the ocean surface substantially modifies the surface albedo (i.e., the ratio of the solar energy reflected by a surface to the total solar energy received by the surface) (NSIDC 2020) and alters the momentum, heat, and mass exchanges between the atmosphere and ocean (Parkinson et al. 1987; Rampal et al. 2011). The level of impact that sea ice has on these exchanges is not static, rather it evolves based on sea ice physical characteristics such as age, sea ice thickness (SIT), sea ice concentration (i.e., percentage of ice-covered area; SIC), and presence of snow cover (Tschudi et al. 2011; Stroeve et al. 2012; Bokhorst et al. 2016). For instance, as young sea ice forms, it traps brine pockets within which makes the ice darker in appearance than pure ice. Therefore, it reflects a smaller amount of shortwave radiation. As the sea ice thickens and ages, brine pockets drain and the sea ice becomes lighter in color, thus more reflective. Additionally, the presence of clean snow cover (i.e., with no contaminants) on sea ice results in the most reflective surfaces on Earth (i.e., albedo ~0.9) (NSIDC 2020).

The Arctic sea ice cover experiences a strong mean annual cycle in both sea ice extent (i.e., area covered by at least 15% or more SIC; SIE) and sea ice volume (SIV). Over the period 1981-2010, the monthly mean SIE maximum in March and minimum in September averaged 15.4×10^6 km^2 and 6.4×10^6 km^2, respectively (Stroeve and Notz 2018). The seasonal evolution of sea ice is controlled by both thermodynamic growth or melt and by dynamic redistribution. Growth of sea ice occurs if the surface mixed layer is cooled down to a freezing temperature (which is a function of salinity) and SATs are below the freezing temperature of sea water. If SATs remain below the sea water freezing temperature sea ice will continue to form. While Arctic sea ice grows almost exclusively as basal growth, it melts as the result of different basal, lateral, and surface processes. One such melting process driven by the ocean is when sea water temperature increases above

the freezing point and basal and/or lateral melt occurs. On the other hand, one of several atmospheric driven melting processes is surface melt of sea ice by SAT increasing above the freezing point.

The sea ice that forms following the September minimum is first year ice by definition (NSIDC 2021). First year ice that forms and melts out every year is called seasonal sea ice (thinner and weaker), and is most commonly found in Arctic marginal seas (e.g., Barents and Chukchi seas) (Serreze and Stroeve 2015). In contrast to seasonal sea ice, the perennial sea ice, or multiyear ice (MYI), survives at least one melt season. MYI is usually the thickest and oldest sea ice. Additional components of sea ice are polar cap ice (which forms a large portion of the MYI), pack ice (i.e., seasonal), and land fast ice (i.e., connected with the shore ocean bottom).

Arctic sea ice is in constant motion (Serreze and Meier 2019). The general circulation of the sea ice is driven mainly by the anticyclonic Beaufort gyre and the transpolar drift (Figure 3). High atmospheric pressure over the Beaufort Sea sets up anticyclonic surface wind stress, which leads to Ekman transport convergence and sea ice convergence within the Beaufort gyre and the buildup of sea ice along the north Greenland and Canadian Arctic Archipelago coasts (Kwok 2018). On the other hand, the transpolar drift results in sea ice export from the Arctic, primarily through the Fram Strait. On shorter time scales, variability in sea ice cover is dominated by transient atmospheric forcing (e.g., low pressure weather systems). These transient systems tend to occur over the Arctic marginal seas during the colder winter months and disrupt the general circulation and introduce warmer subarctic air which can inhibit sea ice growth (Woods and Caballero 2016).

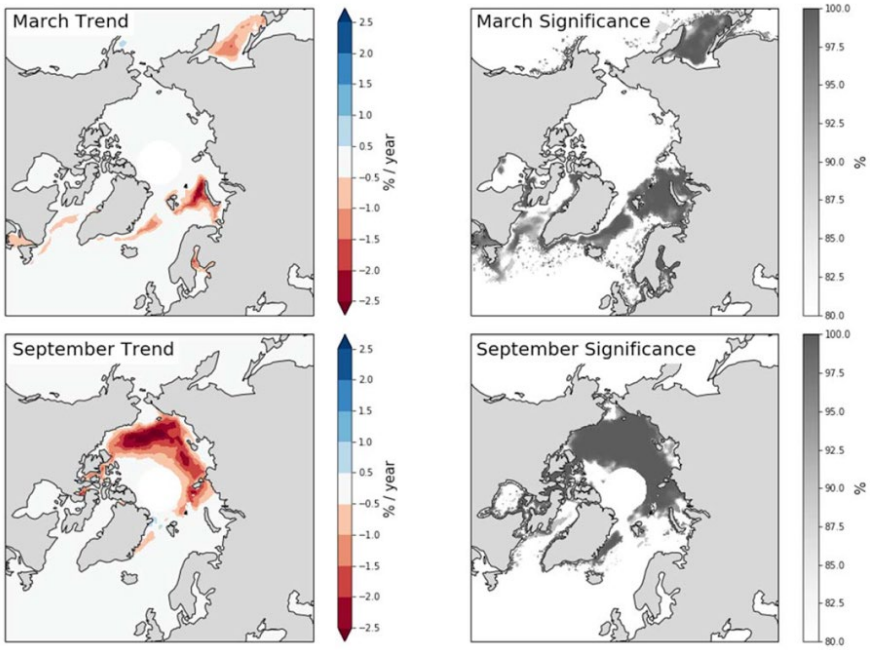
Figure 3. Schematic of Arctic surface currents. Adapted from AMAP (2007).

B. OBSERVED SEA ICE CHANGES

Remote sensing of the Arctic sea ice began late in 1978 using passive microwave measurements of SIC. This observational data, extending through today, provides the most established and continuous measure of Arctic sea ice, albeit only a two-dimensional data set with grid size of about 25 km x 25 km. Interpreting the observational passive microwave data does come with its own algorithm and internal uncertainties (Screen 2011; Eisenman et al. 2014; Meier et al. 2014). According to Ivanova et al. (2014), these uncertainties peak during the summer season, largely due to extensive melt pond coverage and regions of low SIC, and especially in the marginal ice zone (MIZ; SIC between 20-80%). Using long-term passive microwave SIE timeseries reduces some of the observational uncertainty triggered by relatively low-resolution observations, difficulty distinguishing thin ice from open ocean, and changing pole hole size. The pole hole is an area not observed by satellite due to orbit paths, whose size has changed three times over the satellite record. We note that measurements of SIT over the period of satellite observations are discontinuous, and

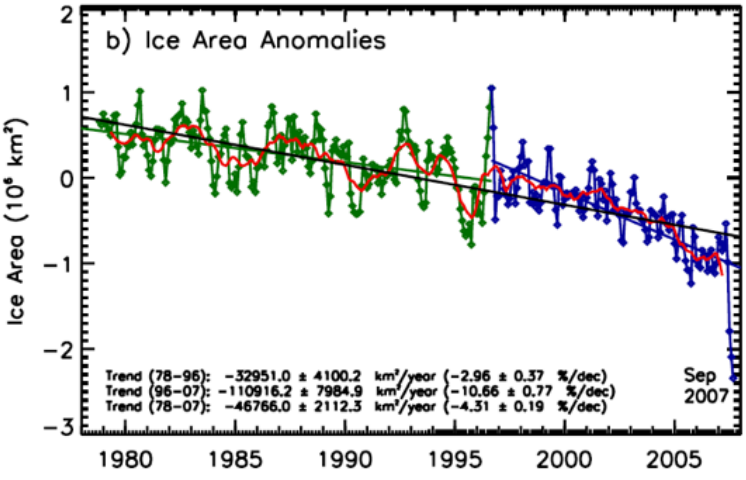
include collection by submarine, aircraft, ships, and more recently Light Detection and Ranging (LIDAR) satellite observations (i.e., ICESat from 2003 to 2009, CryoSat-2 since 2010, and ICESat-2 launched in 2018). Therefore, uncertainty in modeled SIV and SIT is not as well constrained as SIE (Zygmuntowska et al. 2013; Kwok 2018).

SIE has decreased across all months between 1979-2018, with September sea ice showing the largest decline, reducing by 12.8 ±2.3% per decade (IPCC 2019). This decreasing trend is about 83,000 km^2 per year, an area equivalent to about the size of Maine. Figure 4 shows the Arctic regions most impacted by the negative SIC trends for March and September (Stroeve and Notz 2018). Additionally, an accelerated rate in SIA and SIE decline was reported by Comiso et al. (2008) subsequent to the record setting September SIE minimum in 2007 (which has since been eclipsed by September 2012; Figure 5), and has continued through present day (Stroeve and Notz 2018). The cause of this accelerated rate of sea ice decline is subject of active research, but can at least partly be attributed to lengthening of the melt season, increasing inflow of warm air during winter months, and increased net heat transport into the Arctic (e.g., Burgard and Notz 2017; Stroeve and Notz 2018).



SIC trends (% per year) and statistical significance (%) for March (1979-2018) and September (1979-2017).

Figure 4. Arctic sea ice concentration trends. Source: Stroeve and Notz (2018).



Monthly SIA anomaly for November 1978 to September 2007 (green and blue), 12-month running mean (red) and linear trend lines for the full record (black) and 1978-1996 (green) and 1996-2007 (blue). An anomaly is the difference between the reference and observed values. Here the reference is the 1979-2007 SIA monthly mean annual cycle.

Figure 5. Observed accelerated decline in Arctic sea ice area. Source: Comiso et al. (2008).

However, this two-dimensional view of sea ice loss does not report the full story. SIT, and consequently SIV, are also well reported to be in rapid decline (Kwok and Rothrock 2009; Schweiger et al. 2011; Maslowski et al. 2012; Stroeve et al. 2014). Most notably, the Arctic sea ice is thinning and becoming much younger, having lost more than 50% of its MYI during the period 1999–2017, according to Kwok (2018). This is illustrated in Figure 6 showing the sea ice age in late April for 1984 and 2018, and by time series of the percentage among first year and MYI (Stroeve and Notz 2018). Not only is the sea ice thinning due to a warming climate, but as older ice is exported from the Arctic, it is not replenished (Serreze and Meier 2019). Thus, the general SIT state of the Arctic is dramatically changing.

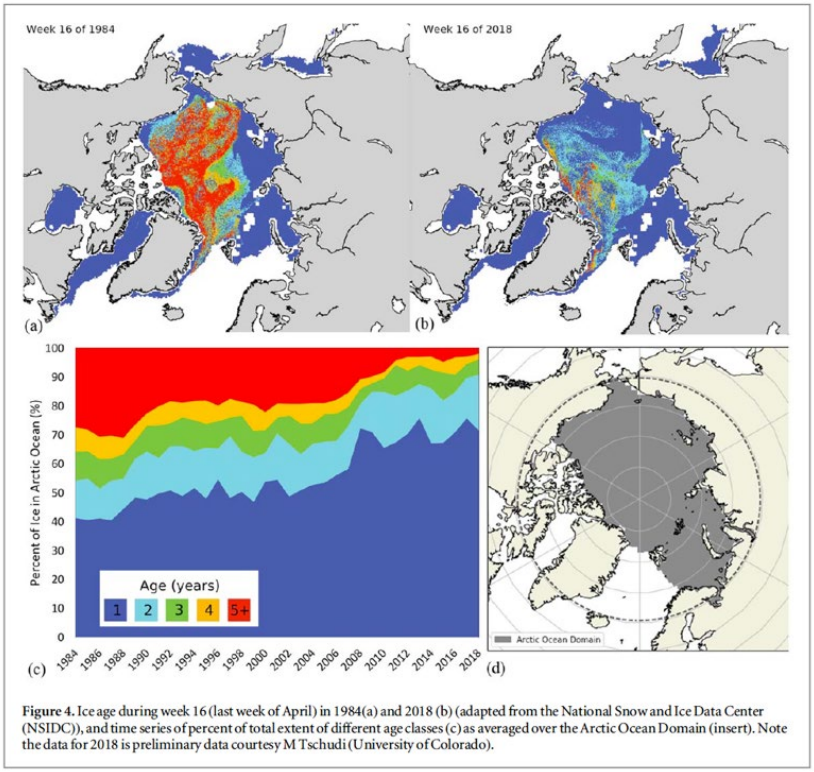


Figure 4. Ice age during week 16 (last week of April) in 1984(a) and 2018 (b) (adapted from the National Snow and Ice Data Center (NSIDC)), and time series of percent of total extent of different age classes (c) as averaged over the Arctic Ocean Domain (insert). Note the data for 2018 is preliminary data courtesy M Tschudi (University of Colorado).

Ice age during the last week of April in 1984 and 2018 (top left and right, respectively). Time series of different age classes given as the percent of total extent (bottom left) averaged over the Arctic Ocean Domain (bottom right).

Figure 6. Arctic sea ice age from 1984 to 2018. Source: Stroeve and Notz (2018).

C. PHYSICAL PROCESSES IMPACTING SEA ICE

Sea ice advection and deformation are driven by winds and ocean currents (Kwok et al. 2013; Serreze and Meier 2019). For any defined region or grid cell, the sea ice mass can be changed when ice is transported into or out of the region, or undergoes convergence (i.e., ridging or rafting) which results in thickening. The ice thickness is important as it alters the response of the sea ice to atmospheric and oceanic forcing (Kwok et al. 2013). Thick ice has deep keels that expose large surface area to surface ocean currents. On the other hand, thin sea ice is more susceptible to atmospheric forcing. In a study by Kwok et al. (2013), thinner sea ice had faster drift speeds than thicker sea ice in response to similar changes in the Arctic Oscillation (AO) during the period 1982-2009. Additionally, faster sea ice drift speeds have consequence on the efficiency of sea ice export through Fram Strait, particularly during Arctic dipole anomaly during 2001-2009 (Kwok et al. 2013).

The Arctic is changing in response to a range of forcing induced by increased levels of GHG in the atmosphere. One example is the forcing associated with increases in Arctic SATs, which have increased at a rate greater than those at lower latitudes (Ballinger et al. 2020). Warmer Arctic SATs reduce sea ice mass and initiate the annual summer melt earlier, which results in a higher percentage of heat-absorbing open water area during the summer (Stroeve and Notz 2018). A warmer upper ocean, in turn, melts adjacent sea ice as well as requiring more time to cool before freeze-up. This positive ice-albedo feedback is a key process in what is called the Arctic amplification (Serreze et al. 2009).

Variability in the sea ice cover is largely driven by climate scale variations in atmospheric and oceanic forcing (Garuba et al. 2020). Diagnosing the relative roles of the atmospheric and oceanic processes in the sea ice melting is challenging as they are not independent processes (Serreze and Meier 2019). Several studies have focused on the atmospheric processes driving interannual to decadal variability (Rigor et al. 2002; Watanabe and Hasumi 2005; Deser and Teng 2008). Key among the atmospheric variations are those involving large-scale changes in Arctic winds, SATs, humidities, and other atmospheric variables that occur on scales of months to years. Examples include the AO, North Atlantic Oscillation, and the Arctic dipole anomaly (Serreze and Meier 2019). For instance, during the positive AO, the prevailing sea level pressure anomaly over the central

Arctic is negative and results in anomalously cyclonic surface wind stresses (Rigor et al. 2002; Watanabe and Hasumi 2005). Thus, sea ice: a) advection from the western Arctic to eastern Arctic decreases, b) export through Fram Strait increases, and c) advection away from the Siberian coast results in more thin ice production in coastal leads (Rigor et al. 2002). Anomalous wind patterns also impact the surface energy budget. Winds from the south usually bring warmer temperatures and more humid air, thereby inhibiting sea ice growth or hastening its melt depending on the season (Serreze and Meier 2019).

The role of oceanic forcing on the declining trends of Arctic sea ice has been the subject of a number of observational and modeling studies (Schlichtholz 2011; Årthun et al. 2012; Woodgate et al. 2012; Schlichtholz 2019). The consensus from these studies is that ocean forcing is involved in the recent sea ice melt in the regions exposed to the northward flow of Pacific and Atlantic water. There, sea ice is strongly impacted by the advection of warm water from the lower latitudes, including variations in ocean heat and volume fluxes (Maslowski et al. 2012). These findings are consistent with an observed increasing trend in OHT (i.e., advection of ocean heat content) from the Pacific and Atlantic ocean basins during approximately 2000-2018 (Woodgate et al. 2012; Tsubouchi et al. 2021) when sea ice rate of decline is accelerated (Comiso et al. 2008).

D. COUPLED MODEL INTERCOMPARISON PROJECTS

1. Over 20 Years of Experience

The CMIPs are a foundational element of today's climate research (Eyring et al. 2016). Tightly associated with the World Climate Research Programme (WCRP), the different phases of CMIP have provided invaluable insight into the past, current, and future climate for over 20 years (Eyring et al. 2016). All climate models participating in CMIP have a global domain. Results and analyses from previous CMIP phases have contributed extensively to international assessments of climate change (e.g., IPCC 2014) and policy recommendations. In the sixth phase of CMIP (CMIP6), the focus is on answering the WCRP's Grand Science Questions, including: a) "How does the Earth system respond to forcing?; [and b)] What are the origins and consequences of systematic model biases?" (Eyring et al. 2016). The CMIP6 simulations of Earth's climate include simulations of the

responses to historical forcing and simulations of the future responses, including a) an abrupt quadrupling of CO_2 concentration (relative to the 1850 pre-industrial control value; a high emissions scenario) and b) a 1% per year CO_2 concentration increase. The latter is the standard CO_2 rate increase used since CMIP2, and the former is continued from CMIP5. CMIP6 model simulations and their studies will be used to inform the IPCC's Sixth Assessment Report (AR6), which is currently in preparation.

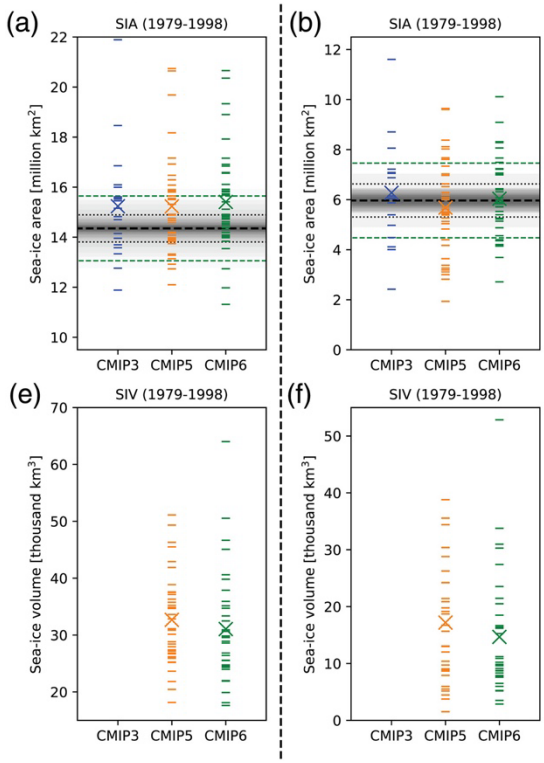
Climate models have long been used as a tool to better understand, predict, and project the complex innerworkings of the Earth system. They also offer an opportunity to improve our process level understanding of climate and to complement *in situ* historical observation gaps from data sparse environments. For example, in the case of reanalysis data, model simulations are constrained by observations or assimilated to provide best estimates. Climate models are often the best tool to quantify the internal variability of observed trends and whether a simulation falls within the observed uncertainty level (Kay et al. 2011). Fully coupled models commonly include individual climate system components for at least the atmosphere, ocean, land, and sea ice (Kauffman et al. 2004). The components are joined by a coupler that exchanges heat, momentum, and mass fluxes across the interfaces (e.g., ocean-atmosphere and ocean-sea ice) at set time intervals. The coupler interface allows for the boundary interface computation to be completed once and ensures conservation of fluxed quantities (Kauffman et al. 2004).

2. CMIP Simulations of Sea Ice

CMIP phases 3, 5, and 6 provide several model variables used to describe the simulated sea ice. The number and variety of variables, as well as the number of participating modeling centers and individual ensemble members has increased through the phases. There have been many studies documenting the representation of Arctic sea ice in these CMIP simulations, but few have focused on diagnosing the causality of local and regional model biases. Provided here is a short review of the most common sea ice variable metrics: SIE or SIA, SIV, and SIT.

The spread in CMIP simulations of SIE and SIA has remained relatively constant throughout all CMIP phases (Stroeve et al. 2012; Shu et al. 2020; SIMIP Community

2020). Figure 7 illustrates the point, showing large model spread in simulating the 1979-1998 mean SIA and SIV among individual models for CMIP phases 3, 5, and 6 (SIMIP Community 2020). Also evident is the relatively consistent skill, relative to observations or SIV reference, of the CMIP multimodel mean (MM). For example, the SIA MM remains largely unchanged and around the observed value for each of the CMIP phases. On the other hand, the SIV MM noticeably decreased from CMIP5 to CMIP6 indicating a generally thinner sea ice cover, and closer to the commonly used Pan-Arctic Ice Ocean Modeling and Assimilation System (PIOMAS) SIV reanalysis (Davy and Outten 2020; SIMIP Community 2020).



The 1979-1998 mean SIA and SIV for March (on left) and September (on right) for CMIP3 (blue), CMIP5 (orange), and CMIP6 (green). Horizontal dashes represent model first ensemble simulation and crosses represent multimodel ensemble mean. The thick dashed lines denote SIA observations.

Figure 7. Comparison of simulated mean sea ice area and volume in different phases of CMIP. Source: SIMIP Community (2020).

Other improvements in the sea ice simulations relative to observations include a progressively more realistically negative SIE decline in the CMIP MM and an increase in the number of individual models exceeding the observed rate (Stroeve et al. 2012; Shu et al. 2020; Davy and Outten 2020; Shen et al. 2021). Davy and Outten (2020) reported that CMIP6 simulations improved on the CMIP5 simulations in better representing the observed ice edge retreat in the Barents Sea, for which the CMIP5 simulations showed too slow a decline in SIE.

Stroeve et al. (2012) suggests internal variability has a strong impact on the September SIE trends. In a study of the CMIP5 models, they found that 16% of the ensemble members have trends that are indistinguishable from zero. However, in a comparison study between phases 5 and 6, Shen et al. (2021) indicates a reduction in the contribution of internal variability to the 1979-2014 September SIE trend, from 33% to 22%, respectively.

The first CMIP phase to enable a comprehensive intercomparison of SIT and SIV was CMIP5. While Stroeve et al. (2014) found good agreement in Arctic mean SIT in CMIP5 compared to available observations and reanalysis, but substantial differences in the thickness patterns were revealed (e.g., location and extent of thickest ice). For example, a large number of sea ice simulations failed to locate the thickest sea ice along the Canadian Arctic Archipelago and northern Greenland coasts (Stroeve et al. 2014). These failures are largely attributed to biases in the simulated general atmospheric circulation. Correctly simulating the SIT pattern is important in determining local surface heat fluxes, which have impacts on both sea ice mass and the rate of sea ice loss (Stroeve et al. 2014).

The SIMIP community (2020) reported sea ice sensitivity improvements in CMIP6 models over previous CMIPs. For example a larger number of models display "the observed sensitivity of Arctic sea ice to anthropogenic CO_2 emissions (i.e., 2.73×10^6 km^2 of sea ice loss per ton of CO_2)… [or] to global warming" (i.e., 4.01×10^6 km^2 of sea ice loss per degree of warming) over the period 1979-2014 (SIMIP Community 2020). However, only a small number of models simulate plausible values (i.e., within two standard deviations) for both (SIMIP Community 2020).

Advances in sea ice models, like improved ice rheology, parametrizations, or higher resolution, could be responsible for some of the improvements discussed here. However, factors other than model structure have also been reported in a number of studies. For example, the inclusion of volcanic emissions in the historical forcing for all models from CMIP3 to CMIP5 strongly influenced the sea ice trends (Rosenblum and Eisenman 2016). In another study, Rosenblum and Eisenman (2017) discovered that CMIP5 simulations with rates of sea ice decline comparable to observations (1979-2013) also have positive global warming biases, which when bias corrected resulted in the models underestimating the observed negative sea ice trends. Rosenblum and Eisenman (2017) "leverage[d] the approximately linear relationship between SIE and global-mean surface temperature" for bias correction of CMIP5 models. According to Rosenblum and Eisenman (2017) "this suggests that the models may be getting the right sea ice trends for the wrong reasons." General underestimation of the simulated historical declining trends is concerning and raises concerns about the accuracy and reliability of model projections of the future Arctic sea ice state.

The consensus among climate scientists is that a seasonally ice-free Arctic will occur when the SIE falls below 1.0×10^6 km^2. The Arctic SIE minimum occurs in September, so projections of September SIE are used to estimate when an ice-free Arctic is likely to occur in the future. While there is a large spread among models, the majority of CMIP6 sea ice projections indicate this will likely occur before the year 2050 in the high emissions scenario (Eyring et al. 2016; Davy and Outten 2020; SIMIP Community 2020). However, in the low anthropogenic CO_2 emissions scenario the Arctic is predicted to stabilize around 2040 at about 2.5×10^6 km^2, and there is a better than 50% chance the Arctic will not become seasonally ice free (Eyring et al. 2016; Davy and Outten 2020). The uncertainty in the timing of the first ice-free September has not appreciably improved over the CMIP phases (Stroeve et al. 2012; SIMIP Community 2020).

III. ASSESSMENT OF THE PAN-ARCTIC ACCELERATED RATE OF SEA ICE DECLINE IN CMIP6 HISTORICAL SIMULATIONS

This chapter was submitted to *Journal of Climate* for publication (14 July 2021). Except for figure numbering, the formatting has been retained as submitted. As the main author of the work, I made the major contributions to the research and writing. Co-authors include W. Maslowski[1],Y.J. Lee[1], J. Clement Kinney[1], and R. Osinski[2].

A. INTRODUCTION

Relative to the global average, the Arctic surface climate represents an amplified response to forced greenhouse warming, largely due to several positive feedbacks within the atmosphere-ocean-sea ice system. One of the most exemplary reflections of Arctic amplification (Serreze and Francis 2006) is the negative trend in sea ice cover observed over all months of the satellite record since 1978 (Serreze and Barry 2011; Notz and Stroeve 2018). Under a reduced sea ice cover regime, the Arctic region not only absorbs more shortwave radiation, but also exchanges additional momentum and heat across the ocean-atmosphere interface (Parkinson et al. 1987; Rampal et al. 2011; Proshutinsky et al. 2019). Thus, the surface energy budget is significantly altered (Jackson et al. 2011; Timmermans et al. 2018). Climate change has been experienced most dramatically in the high latitudes (e.g., sea ice loss and warmer SATs) (Serreze et al. 2009; Taylor et al. 2013; IPCC 2019), potential linkages with mid-latitude weather and climate are subjects of active observations and model research (e.g., Cohen et al. 2020). It is therefore imperative that global climate model simulations reflect the changing Arctic sea ice state.

In this contribution to Phase 6 of the CMIP, we assess historical model simulations of the observed accelerated sea ice trend reflected by a gradual loss of pan-Arctic SIA and SIE prior to the late 1990s, followed by an enhanced rate of decline through present day (Comiso et al. 2008; Serreze and Stroeve 2015). All months of SIA and SIV in the available CMIP6 historical simulations are examined. According to observational and model

[1] Department of Oceanography, Naval Postgraduate School

[2] Institute of Oceanology of Polish Academy of Sciences, Sopot, Poland

reconstructed estimates, the negative trend in SIT and SIV has been even stronger than that in SIA and SIE (Kwok and Rothrock 2009; Schweiger et al. 2011; Maslowski et al. 2012; Stroeve et al. 2014).

Simulations of the historical Arctic sea ice over the past several phases of CMIP have shown similar performance (e.g., spread in mean values, seasonal cycles, and long-term trends) with respect to the large-scale integrated measures of SIA, SIE, and SIV (SIMIP Community 2020; Shu et al. 2020; Davy and Outten 2020; Shen et al. 2021). For example, large model spread in these integrated measures endures whereas the CMIP multi-model ensemble mean continues to outperform all individual models. While some improvements have been reported in CMIP6, including a larger percentage of models simulating the observed sensitivity of Arctic sea ice to anthropogenic CO_2 emissions and global warming (SIMIP Community 2020), September SIE trends closer to observations (Shu et al. 2020; Davy and Outten 2020; Shen et al. 2021), and a modest reduction in September SIE internal variability (Shen et al. 2021), much uncertainty remains as to whether the improvements are attributed to upgraded model physics, forcing, or 'by chance' variability. Still CMIP6 models lack the skill to simulate the faster decline in perennial (i.e., September) SIE observed for 1979-2014 (Davy and Outten 2020; Shu et al. 2020; Shen et al. 2021), and especially after 2000 (Shu et al. 2020).

In this study, we examine both the full Arctic SIA and SIV time series (i.e., seasonal and perennial) instead of the single months of SIA maximum and minimum (March and September, respectively) and focus on the acceleration of the negative trend in recent decades. This paper is organized as follows, data and methods are presented in Section B. Results in Section C emphasize a) the pan-Arctic sea ice state for 1979-2014, b) an acceleration of trends in SIA and SIV decline, and c) the utility of CMIP6 SIV MM as SIV reference. These are followed with the discussion in Section D and conclusions in Section E.

B. DATA AND METHODS

1. Model Output and Observation Data

Our study used the CMIP6 historical experiment data to evaluate Arctic sea ice during the recent past in climate models (Notz et al. 2016; Eyring et al. 2016). Sea ice

outputs for 42 CMIP6 models, including all 268 of their ensemble members, were retrieved from the Earth System Grid Federation (ESGF) repository (https://esgf-node.llnl.gov/search/cmip6/). Additionally, depth-integrated northward net OHT outputs (*hfbasin*) available from 17 out of 42 CMIP6 models were retrieved from the ESGF (Griffies et al. 2016). Table 1 summarizes the model metadata and the variables used for this analysis.

To compare model simulations of sea ice against observations, we first retrieved monthly mean SIC data for the period 1979-2014 from the National Oceanic and Atmospheric Administration (NOAA)/National Snow and Ice Data Center (NSIDC) and the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice Satellite Application Facility (OSI-SAF). Following SIMIP Community (2020; Olason and Notz 2014), a mean observational reference SIC was determined by combining the National Aeronautics and Space Administration (NASA) Team (NT) and Bootstrap (BS) SIC algorithms (Cavalieri et al. (1984) and Comiso (1986), respectively) from NOAA/NSIDC climate data record (CDR) of Passive Microwave SIC, Version 3 (Meier et al. 2017; ftp://sidads.colorado.edu/pub/DATASETS/NOAA/G02202_V3) and the OSI-SAF team SIC algorithm (OSI-450; Lavergne et al. 2019; https://doi.org/10.15770/EUM_SAF_OSI_0008). The OSI-SAF data was first linearly interpolated onto the NSIDC grid, and then averaged with the NT and BT data to create our primary observational reference. For the NT and BT data, the pole hole was filled by the average SIC around the pole hole edges (Olason and Notz 2014). Monthly mean SIA was calculated by multiplying the grid cell area by SIC. The spread in observational estimates (i.e., standard deviation) as the result of algorithm differences can be interpreted as the observational uncertainty (Meier and Stewart 2019; SIMIP Community 2020).

Table 1. Coupled Model Intercomparison Project Phase 6 (CMIP6) model simulations used in this analysis

Model	Ens No.	DOI	siconc	sivol	simass	siarean	sivoln	hfbasin
ACCESS-CM2	3	10.22033/ESGF/CMIP6.4271; v20200817	x	x				
ACCESS-ESM1.5	20	10.22033/ESGF/CMIP6.4272; v20200817	x	x				
AWI-CM 1.1 MR	5	10.22033/ESGF/CMIP6.2686; v20181218				x	x	
AWI-ESM 1.1 LR	1	10.22033/ESGF/CMIP6.9328; v20200212				x	x	
BCC-CSM 2 MR	3	10.22033/ESGF/CMIP6.2948; v20200218	x	x				
BCC-ESM 1	3	10.22033/ESGF/CMIP6.2949; v20200218	x	x				
CAMS-CSM 1.0	2	10.22033/ESGF/CMIP6.9754; v20190708	x	x				
CanESM5	25	10.22033/ESGF/CMIP6.3610; v20190429	x		x			x
CanESM5-CanOE	3	10.22033/ESGF/CMIP6.10260; v20190429	x		x			x
CAS-ESM2-0	4	10.22033/ESGF/CMIP6.3353; v20201225	x	x				
CESM2	11	10.22033/ESGF/CMIP6.7627; v20190308	x	x				
CESM2-FV2	3	10.22033/ESGF/CMIP6.11297; v20191120	x	x				
CESM2-WACCM	3	10.22033/ESGF/CMIP6.10071; v20190227	x	x				
CESM2-WACCM-FV2	3	10.22033/ESGF/CMIP6.11298; v20191120	x	x				
CIESM	3	10.22033/ESGF/CMIP6.8843; v20200420	x	x				
CNRM-CM6-1	20	10.22033/ESGF/CMIP6.4066; v20181126	x	x				
CNRM-CM6-1-HR	1	10.22033/ESGF/CMIP6.4067; v20191021				x	x	
CNRM-ESM2-1	5	10.22033/ESGF/CMIP6.4068; v20181206	x	x				
E3SM 1.0	5	10.22033/ESGF/CMIP6.4497; v20190926	x		x			
E3SM 1.1	1	10.22033/ESGF/CMIP6.11485; v20191210	x		x			
E3SM 1.1 ECA	1	10.22033/ESGF/CMIP6.11486; v20200127	x		x			
EC-Earth3	9	10.22033/ESGF/CMIP6.4700; v20200403		x				x
EC-Earth3-Veg	7	10.22033/ESGF/CMIP6.4706; v20200312		x				x
FGOALS-f3-L	3	10.22033/ESGF/CMIP6.3355; v20191031	x	x				x
GFDL-CM4	1	10.22033/ESGF/CMIP6.8594; v20180701	x	x				
GFDL-ESM4	1	10.22033/ESGF/CMIP6.8597; v20180701	x	x				
GISS-E2.1H	10	10.22033/ESGF/CMIP6.7128; v20191003	x	x				
HadGEM3-GC31-LL	3	10.22033/ESGF/CMIP6.6109; v20200330	x	x				x
HadGEM3-GC31-MM	4	10.22033/ESGF/CMIP6.6112; v20191207	x	x				x
IPSL-CM6A-LR	32	10.22033/ESGF/CMIP6.5195; v20180803	x		x			x
KIOST-ESM	1	10.22033/ESGF/CMIP6.5296; v20201211	x		x			
MIROC6	10	10.22033/ESGF/CMIP6.5603; v20181212	x		x			x
MPI-ESM1.2-HAM	3	10.22033/ESGF/CMIP6.5016; v20190627	x	x				x
MPI-ESM1.2-HR	10	10.22033/ESGF/CMIP6.6595; v20190710	x	x				x
MPI-ESM1.2-LR	10	10.22033/ESGF/CMIP6.6595; v20190710	x	x				x
MRI-ESM2.0	5	10.22033/ESGF/CMIP6.6842; v20190904				x	x	x
NorCPM1	30	10.22033/ESGF/CMIP6.10894; v20190914	x	x				x
NorESM2-LM	3	10.22033/ESGF/CMIP6.8036; v20190920	x	x				x
NorESM2-MM	3	10.22033/ESGF/CMIP6.8040; v20191108	x	x				x
SAM0-UNICON	1	10.22033/ESGF/CMIP6.7789; v20190323	x	x				x
TaiESM1	1	10.22033/ESGF/CMIP6.9755; v20200630	x	x				
UKESM1.0-LL	14	10.22033/ESGF/CMIP6.6113; v20200310	x	x				x

The gridded variables used for this study are sea ice area fraction (siconc), sea ice volume per area (sivol), and sea ice mass per area (simass). Integrated variables used for this study are Northern Hemisphere sea ice area (siarean), extent (siextentn), and volume (sivoln), and northward oceanic heat transport (hfbasin).

A long-term observational time series of SIV does not exist due to the lack of persistent SIT observations over the Arctic. Instead, we use the CMIP6 SIV MM as the primary reference for model intercomparison. Additionally, two SIV reanalyses are included for further comparison against CMIP6 models. These include the PIOMAS (Zhang and Rothrock 2003) and an ice-ocean version of the Regional Arctic System Model (RASM-G; Maslowski et al. 2012; Roberts et al. 2015; Hamman et al. 2016; Cassano et al. 2017) forced with the Japanese 55-year atmospheric reanalysis (JRA-55) data. The monthly mean PIOMAS version 2.1 SIV reanalysis data for the period 1979-2014 was retrieved from the Polar Science Center at University of Washington (http://psc.apl.uw.edu/research/projects/arctic-sea-ice-volume-anomaly/data/). The RASM-G is used as an alternative SIV model reanalysis and is available from the Naval Postgraduate School.

2. Methods

Time series linear trend analysis was used to examine the monthly mean pan-Arctic SIA and SIV for the period of 1979-2014 and the two selected subperiods 1979-1996 and 1997-2014. The loss of sea ice is generally best described by linear regression models for at least the last 30 years (Peng et al. 2020). Here, the simulated SIA time series for CMIP6 models was preferably calculated by the product of SIC (siconc) and grid cell area (areacell) for all northern hemisphere ocean grid cells. If these variables were not provided, we used the variable called Northern Hemisphere SIA (siarean). Analogously, we computed simulated SIV time series by the product of SIV (sivol) or sea ice mass (simass) divided by density of sea ice (ρ_{si}) and areacell for all Northern Hemisphere ocean grid cells. If not provided, instead we used the Northern Hemisphere SIV (sivoln). Following SIMIP Community (2020), SIA was analyzed as our two dimensional sea ice evaluation metric over SIE (i.e., area of SIC greater than at least 15%) which is a strongly grid dependent metric (Notz 2014). Nevertheless, we also calculated SIE and determined that the average difference in sea ice trends between the two metrics to be less than 5% for the CMIP6 models (individual model differences ranges from 0.1% to 24%; Table 2).

Table 2. Difference in trends between the choice of CMIP6 sea ice extent and area

Model	Full			P2		
	SIE	SIA	Abs Error %	SIE	SIA	Abs Error %
ACCESS-CM2	-0.32	-0.33	5.91	-0.21	-0.23	9.69
ACCESS-ESM1.5	-0.39	-0.37	4.99	-0.33	-0.33	1.71
AWI-CM 1.1 MR	-0.39	-0.37	5.24	-0.50	-0.46	7.27
AWI-ESM 1.1 LR	-0.35	-0.32	7.76	-0.47	-0.38	18.78
BCC-CSM 2 MR	-0.53	-0.52	1.25	-0.59	-0.63	8.17
BCC-ESM 1	-0.42	-0.44	3.37	-0.44	-0.49	11.67
CAMS-CSM 1.0	-0.26	-0.24	7.06	-0.19	-0.18	5.01
CanESM5	-0.84	-0.82	2.99	-1.07	-1.04	2.80
CanESM5-CanOE	-0.79	-0.78	1.26	-1.25	-1.27	1.23
CAS-ESM2-0	-0.45	-0.41	8.53	-0.57	-0.51	11.33
CESM2	-0.52	-0.52	0.13	-0.48	-0.47	2.51
CESM2-FV2	-0.47	-0.49	2.62	-0.49	-0.52	7.29
CESM2-WACCM	-0.52	-0.56	6.72	-0.44	-0.49	11.00
CESM2-WACCM-FV2	-0.56	-0.59	4.95	-0.51	-0.57	12.52
CIESM	-0.25	-0.23	7.13	-0.16	-0.17	4.37
CNRM-CM6-1	-0.25	-0.24	2.61	-0.21	-0.21	3.38
CNRM-CM6-1-HR	-0.29	-0.29	0.20	-0.42	-0.41	2.32
CNRM-ESM2-1	-0.36	-0.35	2.99	-0.38	-0.38	0.46
E3SM 1.0	-1.06	-1.01	4.93	-1.46	-1.41	3.64
E3SM 1.1	-0.76	-0.69	10.02	-0.85	-0.84	1.05
E3SM 1.1 ECA	-0.69	-0.66	3.52	-0.78	-0.83	6.72
EC-Earth3	-0.68	-0.65	4.06	-0.87	-0.83	5.18
EC-Earth3-Veg	-0.61	-0.59	4.01	-0.94	-0.91	2.90
FGOALS-f3-L	-0.36	-0.36	0.61	-0.31	-0.33	5.08
GFDL-CM4	-0.50	-0.49	2.90	-0.76	-0.76	0.37
GFDL-ESM4	-0.36	-0.35	2.62	-0.23	-0.21	9.82
GISS-E2.1H	-0.70	-0.74	4.77	-0.80	-0.84	5.26
HadGEM3-GC31-LL	-0.53	-0.54	1.75	-0.92	-0.92	0.45
HadGEM3-GC31-MM	-0.45	-0.47	6.06	-0.51	-0.59	15.00
IPSL-CM6A-LR	-0.53	-0.54	1.44	-0.57	-0.57	0.35
KIOST-ESM	-0.25	-0.25	1.62	-0.53	-0.54	1.86
MIROC6	-0.31	-0.32	3.53	-0.45	-0.44	1.74
MPI-ESM1.2-HAM	-0.39	-0.38	1.09	-0.42	-0.31	25.66
MPI-ESM1.2-HR	-0.44	-0.41	5.95	-0.55	-0.51	6.07
MPI-ESM1.2-LR	-0.38	-0.36	6.75	-0.44	-0.40	8.59
MRI-ESM2.0	-0.42	-0.43	0.85	-0.36	-0.38	4.59
NorCPM1	-0.19	-0.20	4.27	-0.19	-0.21	8.47
NorESM2-LM	-0.22	-0.25	13.26	0.04*	-0.02	-
NorESM2-MM	-0.10	-0.13	24.03	0.01*	-0.04	-
SAM0-UNICON	-0.45	-0.42	7.81	-1.00	-0.90	10.63
TaiESM1	-0.64	-0.61	5.19	-1.00	-0.97	2.42
UKESM1.0-LL	-0.61	-0.59	3.68	-0.74	-0.72	2.70
CMIP6 MM	-0.47	-0.47	1.57	-0.56	-0.55	0.53
CMIP6 average error	-	-	4.77	-	-	6.25
Observations	-0.53	-0.51	3.77	-0.65	-0.67	3.08

Absolute (Abs) error resulting from the choice between CMIP6 SIE and SIA trends for the period 1979-2014 (Full) and 1997-2014 (P2) [(SIE-SIA)/SIE]. Note that * indicates positive P2 SIE trend which produces too large error % and is excluded (-) from CMIP6 average error.

Characteristic anomalies were all calculated relative to the 1979-2014 monthly mean for the individual models and the observational references by removing their own mean annual cycles. Assessing monthly anomalies allowed us to remove the seasonal signal that could otherwise distort the statistics and reduce the influence of individual model biases when examining trend behaviors of the simulated ice pack. Since the majority of CMIP6 models provided multiple ensemble members, up to 32 (Table 1), all results are model ensemble means unless otherwise noted. Also, a CMIP6 multi-model mean (MM) was calculated for SIA, SIV, and OHT; we averaged the individual model ensemble means instead of all ensemble members in order to avoid biasing the MM towards the large ensemble models.

Model and observed characteristic trends (β_m and β_o, respectively) were determined by least-squares linear regression of the monthly mean anomaly time series. Anomaly trends, hereafter referred to as trends, are shown as the 'physical amount' of change (e.g., SIA), or the percentage change relative to the 1979-2014 mean, per decade. Trend uncertainties were calculated following Santer et al. (2008) and Stroeve et al. (2012), whereby we adjusted the modeled and observed standard errors ($s(\beta_m)$ and $s(\beta_o)$, respectively) using an effective sample size (n_{eff}) to account for large lag-1 temporal autocorrelation (AR1) of the trend residuals:

$$n_{eff} = n_{tot}(1 - AR1)/(1 + AR1)$$

where n_{tot} is the number of total months over which the trend is calculated. This is necessary because many geophysical data show pronounced month-to-month persistence, and as such are not statistically independent. Essentially, larger AR1 reduces the number of statistically independent samples, therefore decreasing the statistical degrees of freedom and increasing trend uncertainty.

Furthermore, assuming a normal distribution (Santer et al. 2008; Stroeve et al. 2012), we applied a paired two-tailed *t*-test (*d*) against p=0.10, in which the observed trend is tested against each model realization trend to reject the null hypothesis, which states that the model trend is no different than the observed trend.

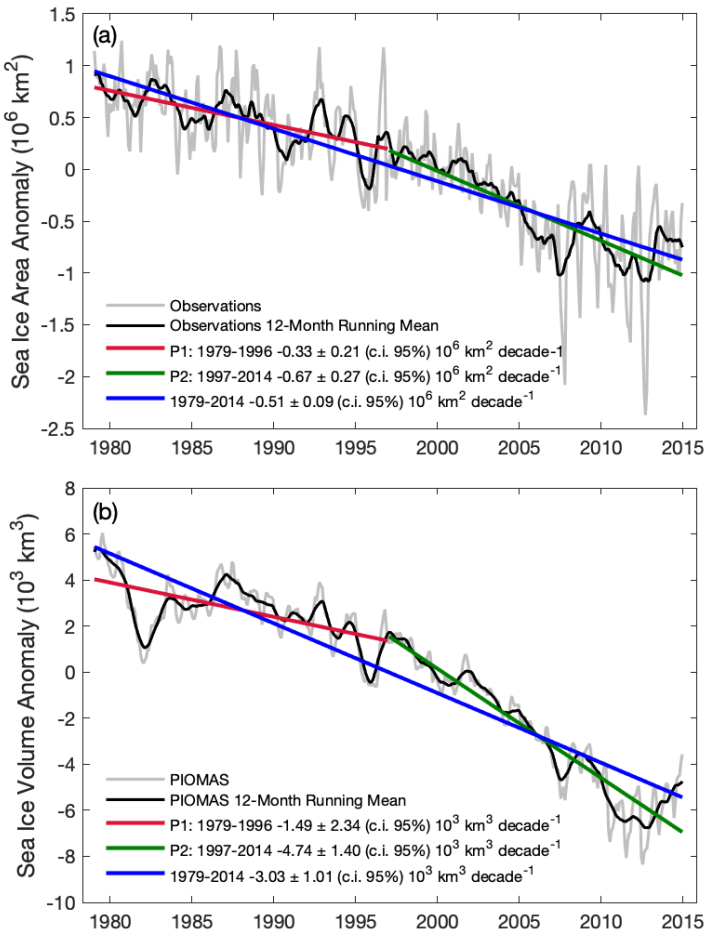
$$d = (\beta_m - \beta_o)/\sqrt{s(\beta_m)^2 + s(\beta_o)^2}$$

C. RESULTS

Figure 8a shows the apparent accelerated decline of SIA from the combined passive microwave observations that are similarly reported in previous studies (Comiso et al. 2008; Stroeve et al. 2012b; Serreze and Stroeve 2015). In this contribution, we examine the linear trends of SIA anomalies for the mutual period of satellite observations and the historical CMIP6 simulations (1979-2014) and two sub-periods: 1979-1996 (P1) and 1997-2014 (P2). The P2 rate of decline in observations, -0.67×10^6 km^2 (–6.61%) decade^{-1}, represents a 100% increase of the P1 rate, -0.33×10^6 km^2 (–3.25%) decade^{-1} (Figure 8a). Our analysis also examines an accelerated rate in SIV decline shown in the PIOMAS reanalysis (Figure 8b). Here the P2 rate, -4.74×10^3 km^3 (–23.48%) decade^{-1}, represents around 200% increase of the P1 rate, -1.49×10^3 km^3 (–7.38%) decade^{-1}.

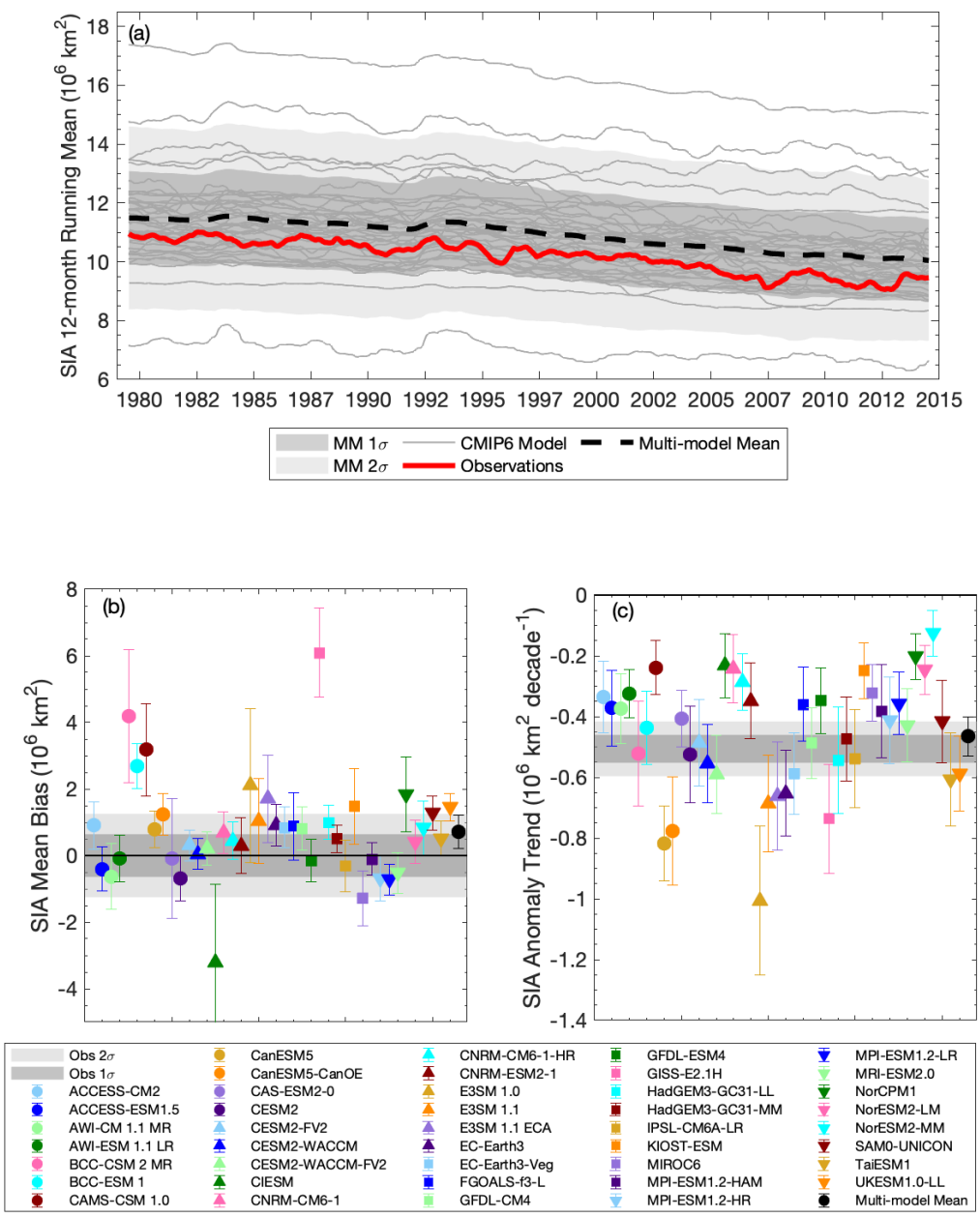
1. 1979-2014 Mean State and Trends of Simulated Arctic Sea Ice

First, we examine the CMIP6 long-term Arctic sea ice statistics and trends. Figure 9a shows a time series of the 12-month running mean SIA for the CMIP6 models and the combined passive microwave observations. Collectively, the MM SIA time series shows approximately 0.5×10^6 km^2 positive bias relative to the observed SIA time series (Figure 9a,b; see also Figure 10a for individual model SIA time series). For the majority of CMIP6 models, mean SIA ranges between 9 and 14×10^6 km^2 in P1 and 8 and 13×10^6 km^2 (Figure 9a) in P2, and their time series remain within one standard deviation (σ) from the CMIP6 MM. However, six CMIP6 models are notably biased, showing time series outside of the $\pm1\sigma$ range from the CMIP6 MM for the total period (Figure 9a). Furthermore, GISS-E2.1H and BCC-CSM 2 MR are largest positively biased (greater than 2σ over the MM) and CIESM is largest negatively biased (2σ under the MM).



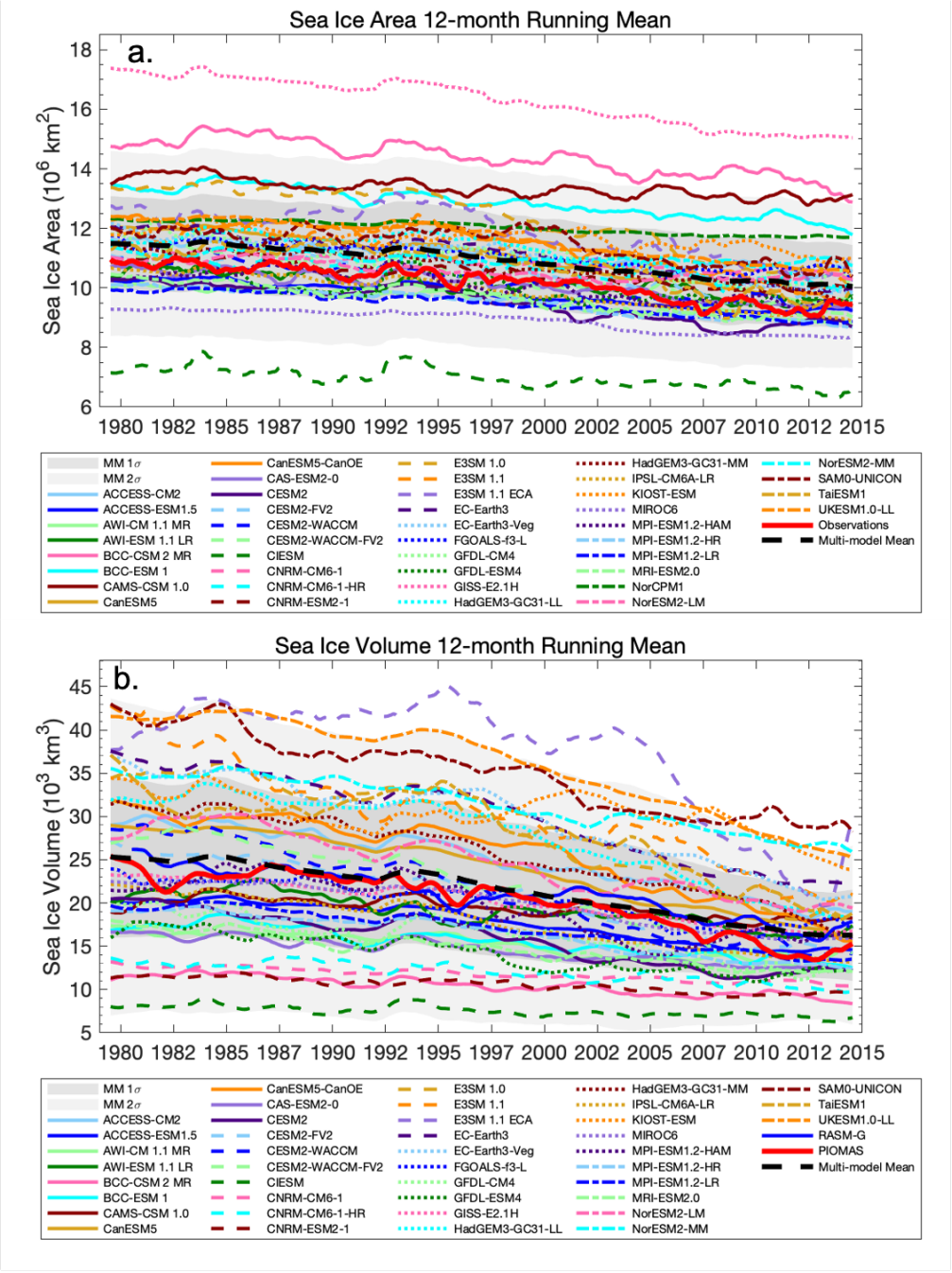
(a) Monthly combined passive microwave observations SIA anomaly (gray), and (b) PIOMAS SIV anomaly (gray) referenced to the period 1979 to 2014. The anomalies based on 12-month running mean (black) are used to determine a linear fit with 95% confidence intervals (c.i.) for the periods 1979 to 2014 (blue), 1979 to 1996 (P1, red), and 1997-2014 (P2, green).

Figure 8. Anomaly linear trends of observed sea ice area and PIOMAS sea ice volume



CMIP6 model simulations and the combined passive microwave observations (obs) (a) 12-month running means of SIA. Gray shading indicates one and two standard deviation (s.d.; σ) from the multi-model mean (MM), respectively. (b) SIA mean bias (i.e., model SIA minus observed SIA) and one s.d. (error bar) for models in reference to the observed SIA. Gray shading indicates one and two s.d. for observed SIA anomaly. (c) SIA anomaly trend and 95% confidence interval (c.i.; 2σ error bar) for models. Two gray shading indicates 68% (darker) and 95% (lighter) c.i. for observed SIA anomaly trend.

Figure 9. CMIP6 sea ice area statistics and trends for 1979-2014



(a) CMIP6 models and the combined passive microwave observations: 12-month running means of SIA ($\times 10^6$ km^2). (b) CMIP6 models and the PIOMAS and RASM-G reanalyses: 12-month running means of SIV ($\times 10^3$ km^3). Gray shading indicates one standard deviation (σ) from the multi-model mean (MM). Note that NorCPM1 is not shown because it is out of the axes range in (b).

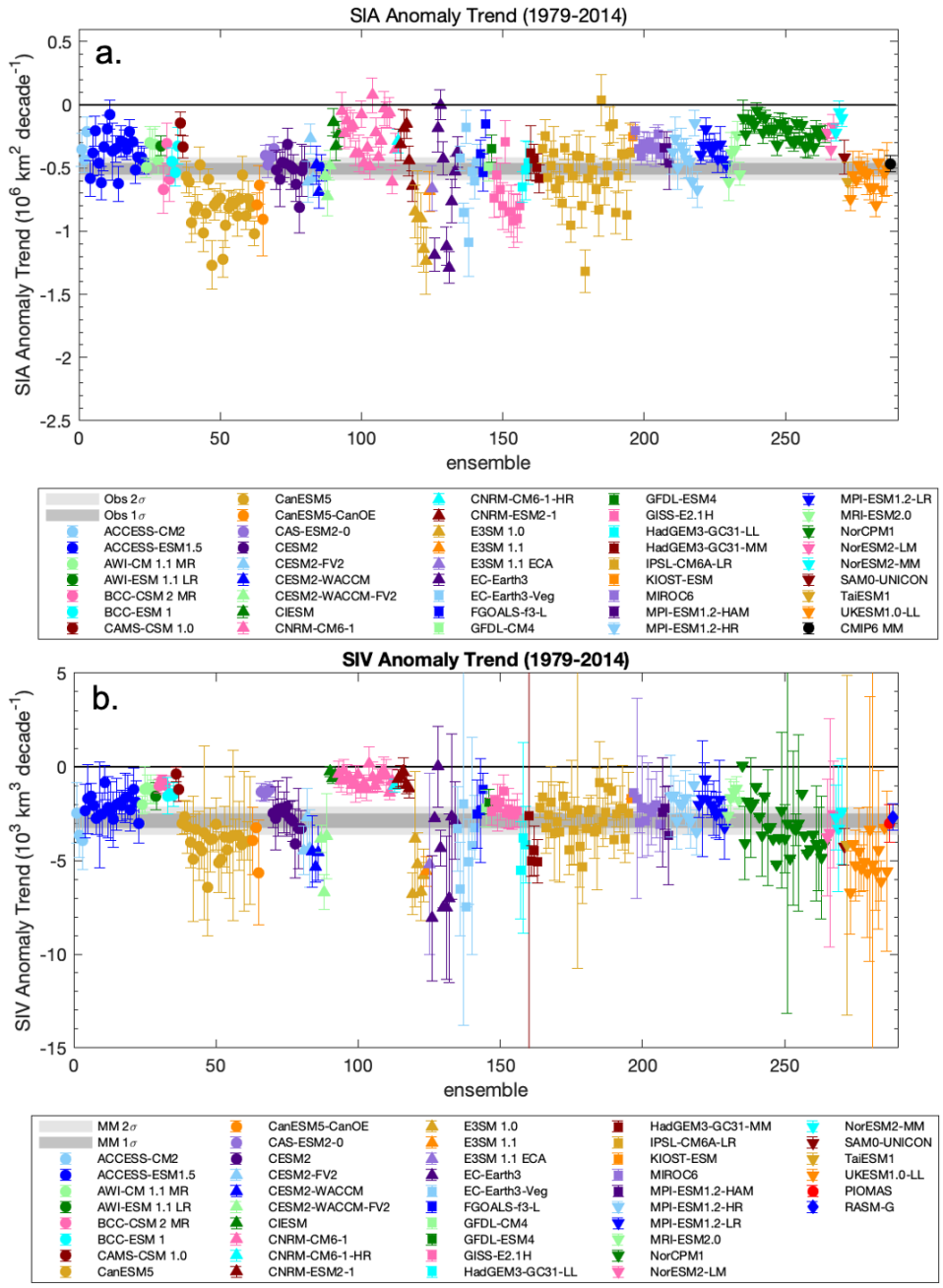
Figure 10. Individual CMIP6 ensembles 12-month running mean for sea ice area and volume

All models simulate a decline in SIA with varying intensity, in general agreement with the observed historical decline for the 36-year period (Figure 9c and Table 3). However, the majority of CMIP6 models underestimate the SIA trend, and those which exceed it usually have positive SIA bias (Figure 9b,c). But slower rates of SIA decline are not unique to the models with low SIA bias (e.g., CAMS-CSM 1.0). The individual model trends are spread relatively wide, ranging from less than one quarter (e.g., NorESM2-MM at -0.13×10^6 km^2 decade^{-1}) to twice (e.g., E3SM 1.0 at -1.01×10^6 km^2 decade^{-1}) the trend observed. About 25% of CMIP6 models show trends within $\pm2\sigma$ of the observed trend, and about 60% have overlapping error bars (Figure 9c), which indicates that the simulated and observed values are from the same population distribution. Despite large intermodel variability, the CMIP6 SIA MM rate of decline, -0.47×10^6 km^2 decade^{-1}, underestimates the observed trend (-0.53×10^6 km^2 decade^{-1}) only by about 10%. Individual ensemble member trends are shown in Figure 11a (Shen et al. 2021).

Table 3. CMIP6 sea ice area anomaly trends

Model	Full		P1		P2		Accelerated
	Mean	c.i.	Mean	c.i.	Mean	c.i.	Trend (P2/P1)
ACCESS-CM2	-0.33	0.12	-0.22	0.29	-0.23	0.33	1.04
ACCESS-ESM1.5	-0.37	0.13	-0.20	0.31	-0.33	0.31	1.67
AWI-CM 1.1 MR	-0.37	0.12	-0.18	0.29	-0.46	0.26	2.62
AWI-ESM 1.1 LR	-0.32	0.08	-0.46	0.22	-0.38	0.23	0.83
BCC-CSM 2 MR	-0.52	0.18	-0.29	0.49	-0.63	0.45	2.18
BCC-ESM 1	-0.44	0.12	-0.31	0.37	-0.49	0.28	1.56
CAMS-CSM 1.0	-0.24	0.09	-0.34	0.24	-0.18	0.25	0.52
CanESM5	-0.82	0.13	-0.50	0.28	-1.04	0.26	2.09
CanESM5-CanOE	-0.78	0.18	-0.16	0.29	-1.27	0.24	7.82
CAS-ESM2-0	-0.41	0.10	-0.19	0.28	-0.51	0.22	2.65
CESM2	-0.52	0.16	-0.20	0.38	-0.47	0.48	2.38
CESM2-FV2	-0.49	0.15	-0.08	0.38	-0.52	0.33	6.84
CESM2-WACCM	-0.56	0.13	-0.34	0.33	-0.49	0.38	1.44
CESM2-WACCM-FV2	-0.59	0.13	-0.38	0.33	-0.57	0.41	1.50
CIESM	-0.23	0.11	-0.06	0.33	-0.17	0.25	2.80
CNRM-CM6-1	-0.24	0.11	-0.22	0.29	-0.21	0.30	0.98
CNRM-CM6-1-HR	-0.29	0.09	-0.01	0.23	-0.41	0.24	46.80
CNRM-ESM2-1	-0.35	0.13	-0.25	0.36	-0.38	0.35	1.51
E3SM 1.0	-1.01	0.25	-0.25	0.42	-1.41	0.51	5.56
E3SM 1.1	-0.69	0.16	0.04	0.34	-0.84	0.25	-23.00**
E3SM 1.1 ECA	-0.66	0.18	0.19	0.26	-0.83	0.36	-4.34**
EC-Earth3	-0.65	0.14	-0.40	0.35	-0.83	0.34	2.06
EC-Earth3-Veg	-0.59	0.14	-0.39	0.30	-0.91	0.26	2.36
FGOALS-f3-L	-0.36	0.12	-0.25	0.27	-0.33	0.32	1.32
GFDL-CM4	-0.49	0.12	-0.52	0.34	-0.76	0.28	1.46
GFDL-ESM4	-0.35	0.11	-0.30	0.46	-0.21	0.18	0.69
GISS-E2.1H	-0.74	0.18	-0.36	0.44	-0.84	0.45	2.34
HadGEM3-GC31-LL	-0.54	0.18	0.06	0.32	-0.92	0.24	-15.64**
HadGEM3-GC31-MM	-0.47	0.14	-0.24	0.28	-0.59	0.30	2.39
IPSL-CM6A-LR	-0.54	0.16	-0.31	0.40	-0.57	0.41	1.81
KIOST-ESM	-0.25	0.09	-0.24	0.22	-0.54	0.22	2.28
MIROC6	-0.32	0.09	-0.10	0.20	-0.44	0.24	4.59
MPI-ESM1.2-HAM	-0.38	0.16	0.00	0.40	-0.31	0.35	-
MPI-ESM1.2-HR	-0.41	0.15	-0.26	0.35	-0.51	0.34	1.99
MPI-ESM1.2-LR	-0.36	0.10	-0.24	0.27	-0.40	0.25	1.66
MRI-ESM2.0	-0.43	0.12	-0.29	0.29	-0.38	0.36	1.29
NorCPM1	-0.20	0.08	-0.08	0.18	-0.21	0.19	2.67
NorESM2-LM	-0.25	0.08	-0.27	0.21	-0.02	0.20	0.06
NorESM2-MM	-0.13	0.08	-0.16	0.19	-0.04	0.23	0.23
SAM0-UNICON	-0.42	0.14	-0.02	0.23	-0.90	0.32	49.09
TaiESM1	-0.61	0.16	-0.08	0.24	-0.97	0.42	11.91
UKESM1.0-LL	-0.59	0.13	-0.25	0.26	-0.72	0.24	2.83
CMIP6 SIA MM	-0.47	0.07	-0.22	0.15	-0.56	0.10	2.55
Observations	-0.51	0.09	-0.33	0.21	-0.67	0.27	2.03

Linear trend ($\times 10^6$ km^2 decade^{-1}) and 95% confidence interval ($\times 10^6$ km^2 decade^{-1}; c.i.) of monthly mean sea ice area anomalies (in reference to individual model or observation 1979-2014 climatology) for the periods 1979-2014 (Full), 1979 to 1996 (P1), and 1997 to 2014 (P2). Accelerated trend is determined by the ratio of P2 trend to P1 trend. Note that **negative ratio is the result of a positive P1 trend, the magnitude should not be compared.



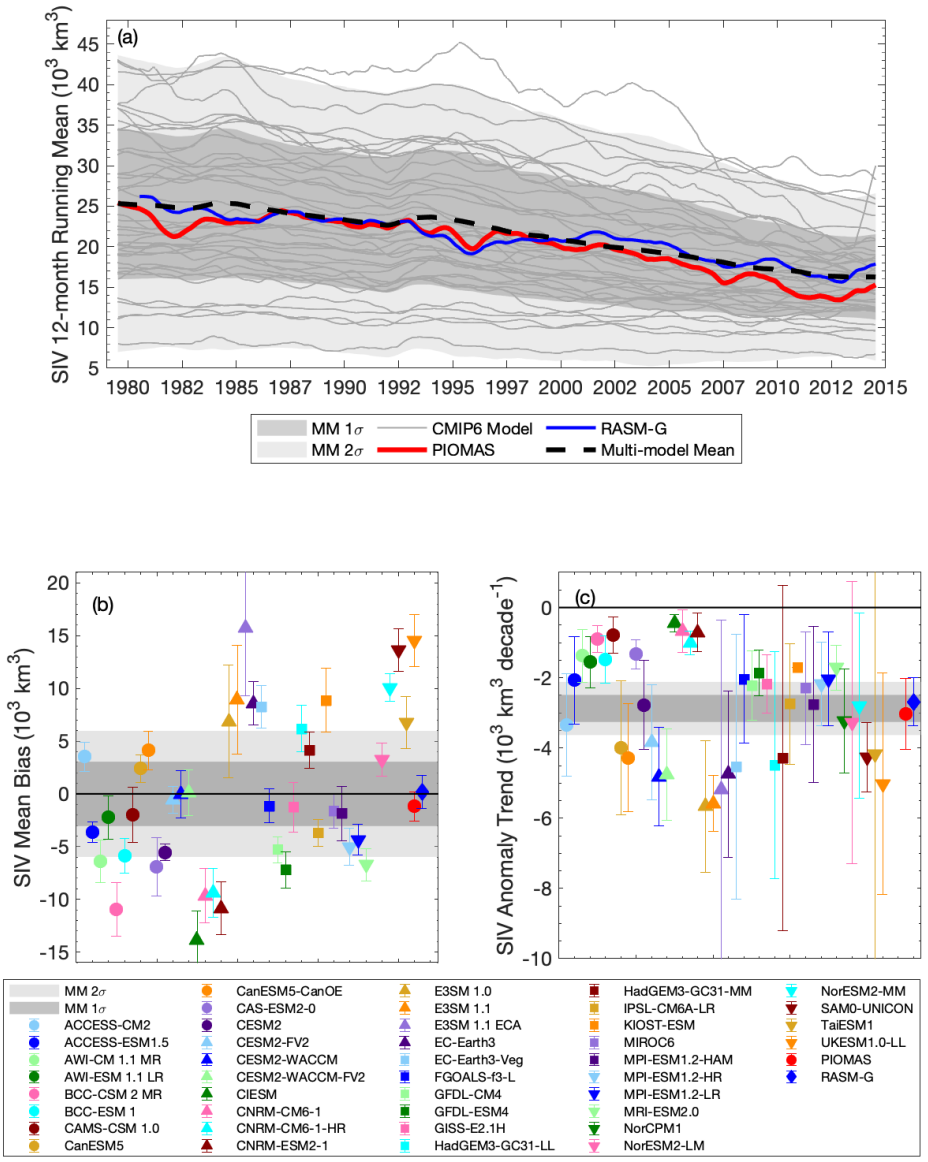
(a) SIA and (b) SIV trends from 1979-2014 for all individual model ensemble members and the CMIP6 multi-model mean. Error bars indicate the 95% confidence intervals (c.i.). The dark and light gray horizontal shading indicates the adjusted c.i.: ± one standard deviation (σ) (68%) and ±2σ (95%) from the observed SIA trend in (a) and CMIP6 SIV MM trend in (b), respectively (following Santer et al. (2008) and Stroeve et al. (2012)).

Figure 11. CMIP6 model ensemble member sea ice area and sea ice volume trends for 1979-2014

Next, we follow with a similar analysis for SIV. Figure 12a shows a time series of 12-month running mean SIV for the CMIP6 models and PIOMAS and RASM-G SIV reanalyses. One model, NorCPM1, has such a high bias (40×10^3 km^3; more than twice as large as any other model; Figure 12b) that we excluded it from the CMIP6 SIV MM. For the majority of CMIP6 models, SIV ranges within ±1σ from the MM (i.e. 11-35×10^3 km^3; Figure 12), but models are more evenly spread out across the broader range of ±2σ of the SIV MM compared to SIA shown in Figure 9a. Beyond around the year 2000, the model spread in simulated SIV is reduced 45% to a minimum range of 11-22×10^3 km^3 (Figure 12a). This indicates that large simulated ice thickness uncertainty exists during the early portion of our analysis period. Models with largest bias are E3SM 1.1 ECA, largest positively biased, and CIESM, largest negatively biased (Figure 12b; see Figure 10b for individual model SIV time series).

Collectively, the CMIP6 SIV MM mean of 21.4×10^3 km^3 and trend of -2.88×10^3 km^3 decade^{-1} for 1979-2014 pairs well with the two SIV reanalyses (Figure 12b,c and Table 4). Relative to the CMIP6 SIV MM mean, PIOMAS has a slight negative difference in mean (20.2×10^3 km^3) and trend (-3.03×10^3 km^3 decade^{-1}), while RASM-G has no difference in mean and a slower trend (-2.52×10^3 km^3 decade^{-1}). All models simulate a decline in SIV with varying intensity, which qualitatively matches the reanalyses and satellite estimated (Kwok 2018) SIV trends. Individual CMIP6 SIV trend ranges from -0.68×10^3 km^3 decade^{-1} (CNRM-CM6-1) to -5.66×10^3 km^3 decade^{-1} (E3SM 1.0). About 30% of the CMIP6 models show SIV trends within ±2σ of the CMIP6 MM trend (about 40% for PIOMAS trend), and about 80% have overlapping error bars with the MM (Figure 12c). Individual ensemble member trends are shown in Figure 11b.

Only four models (CESM2, GFDL-CM4, IPSL-CM6A-LR, MPI-ESM1.2-HR) are within two confidence intervals for both SIA and SIV respective reference means and trends for 1979-2014.



CMIP6 model simulations and the PIOMAS and RASM-G reanalyses (a) 12-month running means of SIV (×10^3 km3). Gray shading indicates one and two standard deviation (s.d.; σ) from the CMIP6 SIV multi-model mean (MM), respectively. (b) SIV mean bias (i.e., model SIV minus MM SIV) and one s.d. (error bars) for models in reference to the MM SIV. Two gray shading indicates one and two s.d. from the CMIP6 MM mean. (c) SIV mean trend and 95% confidence interval (c.i.; error bar) for models. Two gray shading indicates 68% and 95% c.i. for the MM SIV anomaly trend, respectively. Note that NorCPM1 is not shown because it is out of the axes' range except in (c).

Figure 12. Sea ice volume statistics and trends for 1979-2014

Table 4. CMIP6 sea ice volume anomaly trends

Model	Full		P1		P2		Accelerated
	Mean	c.i.	Mean	c.i.	Mean	c.i.	Trend (P2/P1)
ACCESS-CM2	-3.34	0.74	-2.01	1.78	-3.76	3.80	1.87
ACCESS-ESM1.5	-2.07	0.64	-0.91	1.75	-2.21	2.16	2.42
AWI-CM 1.1 MR	-1.37	0.38	-0.54	1.18	-1.48	0.72	2.72
AWI-ESM 1.1 LR	-1.55	0.37	-1.99	1.92	-1.82	0.53	0.92
BCC-CSM 2 MR	-0.90	0.20	-0.69	0.50	-0.86	0.46	1.24
BCC-ESM 1	-1.48	0.34	-1.17	1.79	-1.54	0.56	1.31
CAMS-CSM 1.0	-0.79	0.26	0.03	0.87	-0.86	0.59	-32.39**
CanESM5	-3.99	0.98	-2.16	2.69	-4.95	1.98	2.29
CanESM5-CanOE	-4.28	0.78	-2.05	2.19	-6.70	1.36	3.27
CAS-ESM2-0	-1.33	0.21	-0.87	0.57	-1.21	0.44	1.40
CESM2	-2.78	0.65	-1.90	2.33	-2.52	1.99	1.33
CESM2-FV2	-3.84	0.84	-2.64	2.54	-4.37	1.71	1.66
CESM2-WACCM	-4.82	0.72	-3.99	6.14	-4.01	1.11	1.00
CESM2-WACCM-FV2	-4.75	0.67	-3.16	2.31	-5.27	0.95	1.67
CIESM	-0.44	0.13	-0.16	0.37	-0.30	0.31	1.83
CNRM-CM6-1	-0.68	0.31	-0.63	1.00	-0.61	0.86	0.96
CNRM-CM6-1-HR	-1.00	0.17	-0.31	0.48	-1.14	0.34	3.68
CNRM-ESM2-1	-0.71	0.28	-0.52	0.91	-0.80	0.87	1.55
E3SM 1.0	-5.66	0.96	-2.25	1.86	-6.64	1.07	2.94
E3SM 1.1	-5.58	0.41	-7.30	1.26	-6.22	0.74	0.85
E3SM 1.1 ECA	-5.19	2.46	2.16	0.84	-12.22	5.70	-5.66**
EC-Earth3	-4.75	1.20	-3.03	3.23	-5.65	4.91	1.87
EC-Earth3-Veg	-4.54	1.92	-2.67	1.95	-6.56	24.54	2.46
FGOALS-f3-L	-2.03	0.93	-1.42	10.64	-2.41	1.81	1.69
GFDL-CM4	-2.22	0.50	-1.67	5.41	-3.33	0.52	1.99
GFDL-ESM4	-1.86	0.34	-1.49	0.51	-0.72	0.65	0.48
GISS-E2.1H	-2.17	0.42	-1.44	1.34	-2.22	0.80	1.54
HadGEM3-GC31-LL	-4.49	1.65	-1.01	1.47	-6.84	1.22	6.75
HadGEM3-GC31-MM	-4.29	2.51	-2.55	1.35	-5.01	2.07	1.97
IPSL-CM6A-LR	-2.75	0.88	-1.85	2.50	-2.46	1.22	1.33
KIOST-ESM	-1.70	0.00	-3.19	0.00	-4.27	3.55	1.34
MIROC6	-2.30	0.82	-0.69	0.48	-3.11	2.85	4.49
MPI-ESM1.2-HAM	-2.76	1.14	-0.07	1.83	-1.52	3.29	22.12
MPI-ESM1.2-HR	-2.17	0.61	-1.14	1.18	-2.64	0.98	2.32
MPI-ESM1.2-LR	-2.03	0.68	-1.23	1.17	-2.28	0.94	1.86
MRI-ESM2.0	-1.71	0.33	-1.65	0.93	-1.68	0.91	1.02
NorCPM1	-3.23	0.75	-1.78	1.43	-4.74	1.55	2.67
NorESM2-LM	-3.27	2.05	-2.21	3.66	-2.91	0.64	1.32
NorESM2-MM	-2.80	1.35	-1.59	0.49	-2.83	0.81	1.78
SAM0-UNICON	-4.27	0.51	-4.07	1.50	-4.04	2.14	0.99
TaiESM1	-4.17	4.62	-0.18	2.43	-7.04	1.01	39.48
UKESM1.0-LL	-5.02	1.61	-1.88	1.73	-7.19	1.83	3.83
CMIP6 SIV MM	-2.88	0.89	-1.66	1.92	-3.55	2.06	2.13
PIOMAS	-3.03	1.01	-1.49	2.34	-4.74	1.40	3.18
RASM-G*	-2.52	0.36	-3.07	0.86	-3.11	1.12	1.01

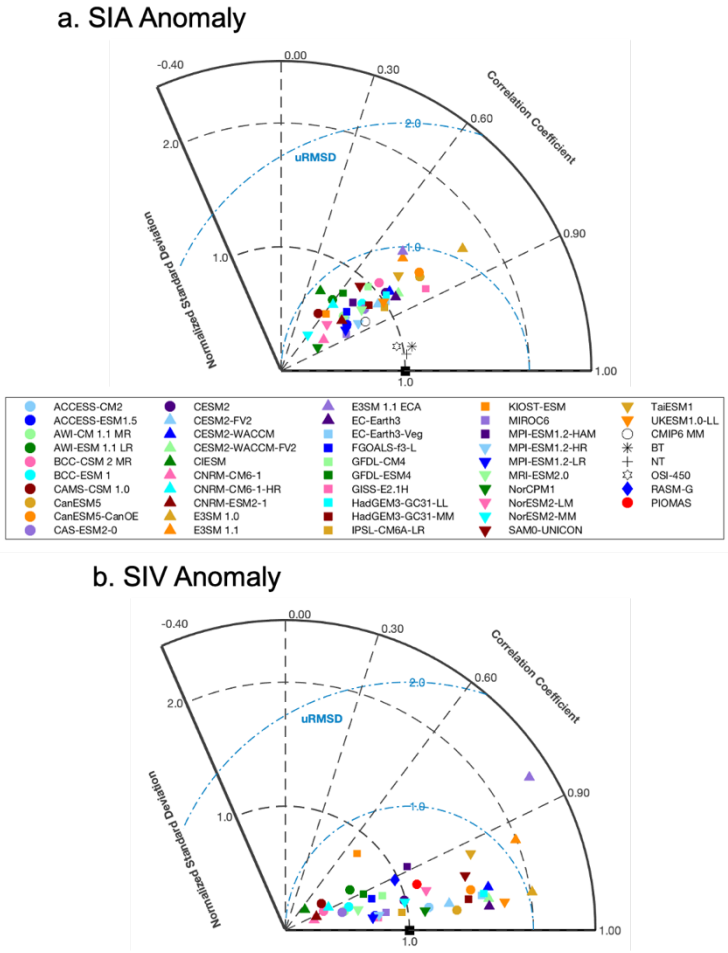
Linear decadal trends ($\times 10^3$ km^3 decade^{-1}) and 95% confidence interval ($\times 10^3$ km^3 decade^{-1}; c.i.) of monthly mean SIV anomalies (in reference to individual model 1979-2014 climatology) for the periods 1979-2014 (Full), 1979 to 1996 (P1; *except RASM-G for 1980 to 1997), and 1997 to 2014 (P2). Accelerated trend is determined by the ratio of P2 trend to P1 trend. Note that **negative ratio is the result of a positive P1 trend, the magnitude should not be compared.

2. Sea Ice Variability

In order to evaluate CMIP6 model skill in simulating sea ice variability, Taylor diagrams (Taylor 2001) are used to quantify the statistics of CMIP6 sea ice anomalies against observations and the CMIP6 SIV MM reference (Figure 13). This analysis essentially removes the individual model bias and presents the following statistics: the correlation coefficient (r) measures the strength of linear relationship, the unbiased root-mean-square difference (uRMSD) describes the difference between the observed and simulated values, and the normalized standard deviation (σ) indicate the ratio of model variability against the reference. Figure 13a shows the model SIA anomaly against the observed SIA anomaly for the period 1979-2014. Additionally, individual SIA observational estimates (i.e., BT, NT, and OSI-450) are included to quantify uncertainties resulting from choice of algorithm. Relative to the combined SIA observations, all individual observational estimates have strong correlation coefficients of 0.98 (p<0.01), and BT / NT has a slight positive / negative bias with larger standard deviation while OSI-450 has a slight positive bias and smaller standard deviation (Figure 13a).

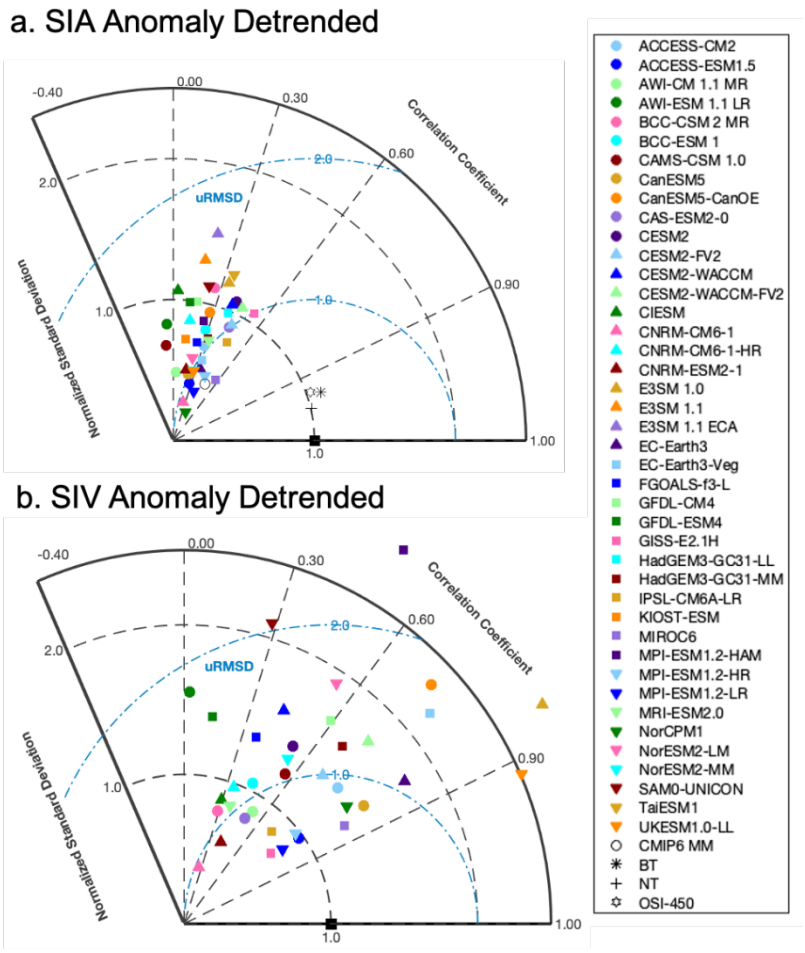
About 80% of the CMIP6 SIA simulations have anomaly correlation coefficients greater than 0.7 (the lower end of strong correlation, p<0.05), and about 43% also have normalized standard deviations between 0.75 and 1.25 (Figure 13). The majority of the CMIP6 models (64%) show lower SIA variability (i.e., normalized standard deviation is less than one) than in observations, which is manifested in the CMIP6 SIA MM (Figure 13a). We also examined the detrended SIA anomalies (Figure 14) to evaluate model interannual variability without the influence of the negative long-term SIA trend. When detrended, all CMIP6 models shift left in the Taylor diagram (indicating lower correlation) and only 12 models (about 30%; CAS-ESM2-0, CESM2, CESM2-FV2, CESM2-WACCM, CESM2-WACCM-FV2, E3SM 1.0, HadGEM3-GC31-LL, HadGEM3-GC31-MM, GISS-E2.1H, IPSL-CM6A-LR, MRI-ESM2.0, TaiESM1) exceed weak correlation (r>0.3, p<0.05) against the observed SIA anomaly and have normalized standard deviations between 0.75 and 1.25. These models (except GISS-E2.1H, E3SM 1.0) also show low bias and have realistic trends in SIA (Figure 9b,c). This suggests that the moderate model 'skill' identified in Figure 13a for nearly half the CMIP6 models is mostly

controlled by larger correlation and standard deviation values resulting from long-term trends in the SIA decline.



Model skill in representing the monthly anomalies of (a) SIA and (b) SIV reference over the period 1979 to 2014. Taylor diagram displays the unbiased root-mean-square difference (uRMSD), correlation coefficient (*r*), and normalized standard deviation (s.d.) divided by the observed s.d. (0.64×10^6 km^2) for SIA and the CMIP6 multi-model mean (MM) s.d. (3.05×10^3 k m^3) for SIV. The square marker indicates the perfect model.

Figure 13. Taylor diagrams illustrating CMIP6 skill in simulating sea ice variability



Model skill in representing (a) the monthly anomalies of (a) SIA and (b) SIV reference over the period 1979 to 2014. Taylor diagram displays the unbiased root-mean-square difference (uRMSD), correlation coefficient (*r*), and normalized standard deviation (s.d.) divided by the observed s.d. (0.37×10^6 km^2) for SIA and the CMIP6 multi-model mean (MM) s.d. (0.61×10^3 km^3) for SIV. The square marker indicates the perfect model.

Figure 14. Taylor diagrams illustrating CMIP6 skill in simulating detrended sea ice variability

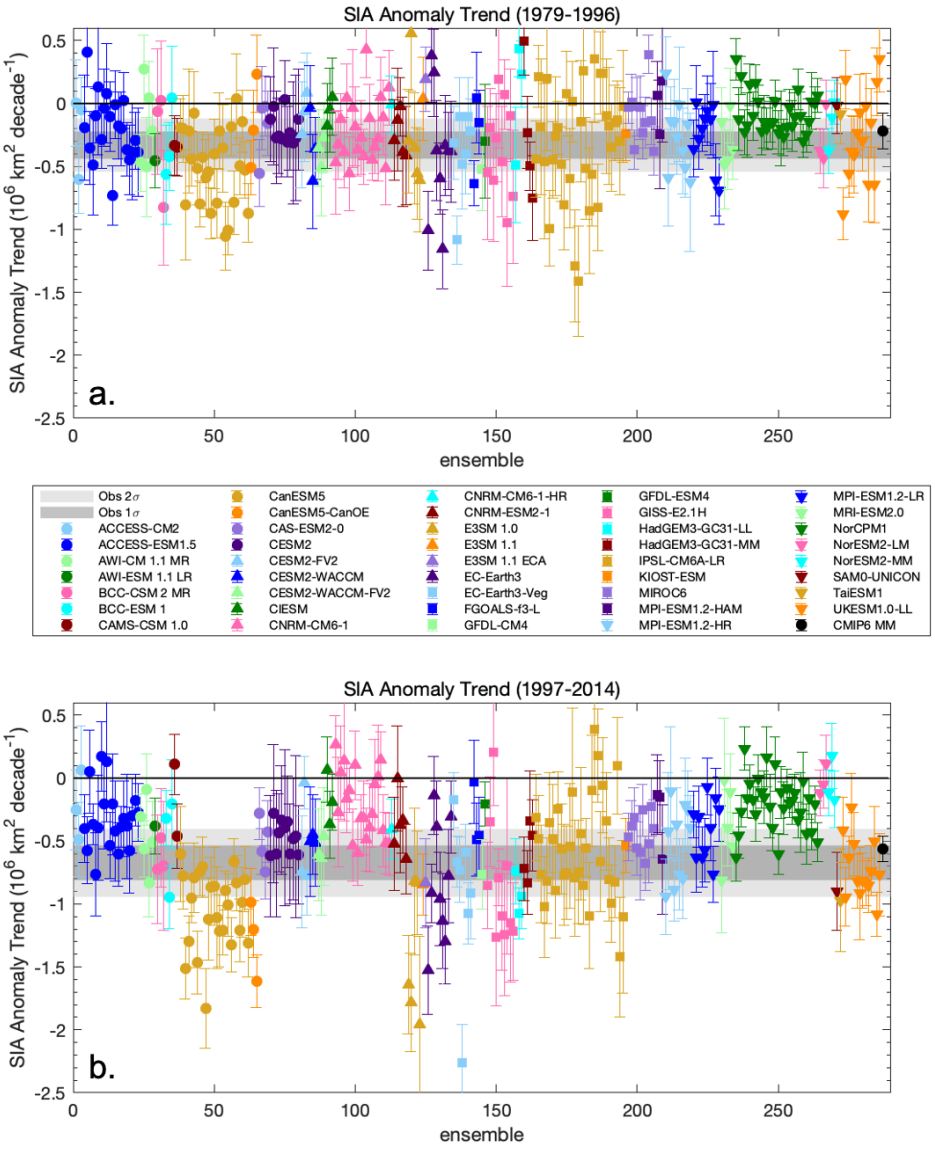
For SIV anomaly (Figure 13b), we evaluated CMIP6 models against the CMIP6 SIV MM anomaly for the period 1979-2014. Also shown are the two SIV reanalyses, PIOMAS and RASM-G, which both have strong correlation and comparable standard deviation against the CMIP6 MM. All models, except CIESM and KIOST-ESM, are strongly correlated ($r>0.7$, $p<0.05$) against the CMIP6 MM, but also show larger spread in SIV variability compared to the SIA spread (Figure 13b). When the same normalized standard deviation criteria used for SIA is applied (between 0.75 and 1.25), we find 11 CMIP6 models (about

26%; ACCESS-CM2, CESM2, GFDL-CM4, GISS-E2.1H, IPSL-CM6A-LR, MIROC6, MPI-ESM1.2-HAM, MPI-ESM1.2-HR, NorCPM1, NorESM2-LM, NorESM2-MM) reasonably simulate the interannual variability statistics of the CMIP6 SIV MM. When detrended, there is large model spread against the CMIP6 SIV MM (Figure 14b).

3. Accelerated Rates of Sea Ice Decline

Next, we gauged the ability of CMIP6 models to simulate multidecadal trends on a shorter scale than shown in section 3a. Our reference is the apparent accelerated decline of SIA from observations (Figure 8a) based on linear trends for two sub-periods: P1 (1979-1996) and P2 (1997-2014). However, for shorter periods (i.e., 18 versus 36 year periods) we must consider that the contribution of internal climate variability increases and makes the forced signal more difficult to detect (Kay et al. 2011). Nevertheless, both P1 and P2 observational trends are statistically different from zero trend at $\pm 2\sigma$ confidence interval.

We evaluated both the CMIP6 individual ensemble member simulations and the model ensemble mean of SIA trends, and whether they are consistent with observations at a level of statistical significance. Following Santer et al. (2008) and Stroeve et al. (2012), Figure 15 shows the simulated SIA trend for all individual CMIP6 ensemble members for periods P1 and P2. There is a large range in SIA trends among models as well as within individual model's ensemble members, the latter indicating sizeable internal variability. The model spread is much larger when analyzing the shorter climate periods (i.e., P1 and P2; Figure 15a,b) rather than the full historical simulation period (Figure 11a). Another consequence of examining the shorter time periods P1 and P2 is the sizable trend uncertainty (i.e., large error bars) that results from substantial lag autocorrelation in the SIA time series (i.e., AR1 ranges from 0.52 to 0.89) and the subsequent adjustment of independent samples size.



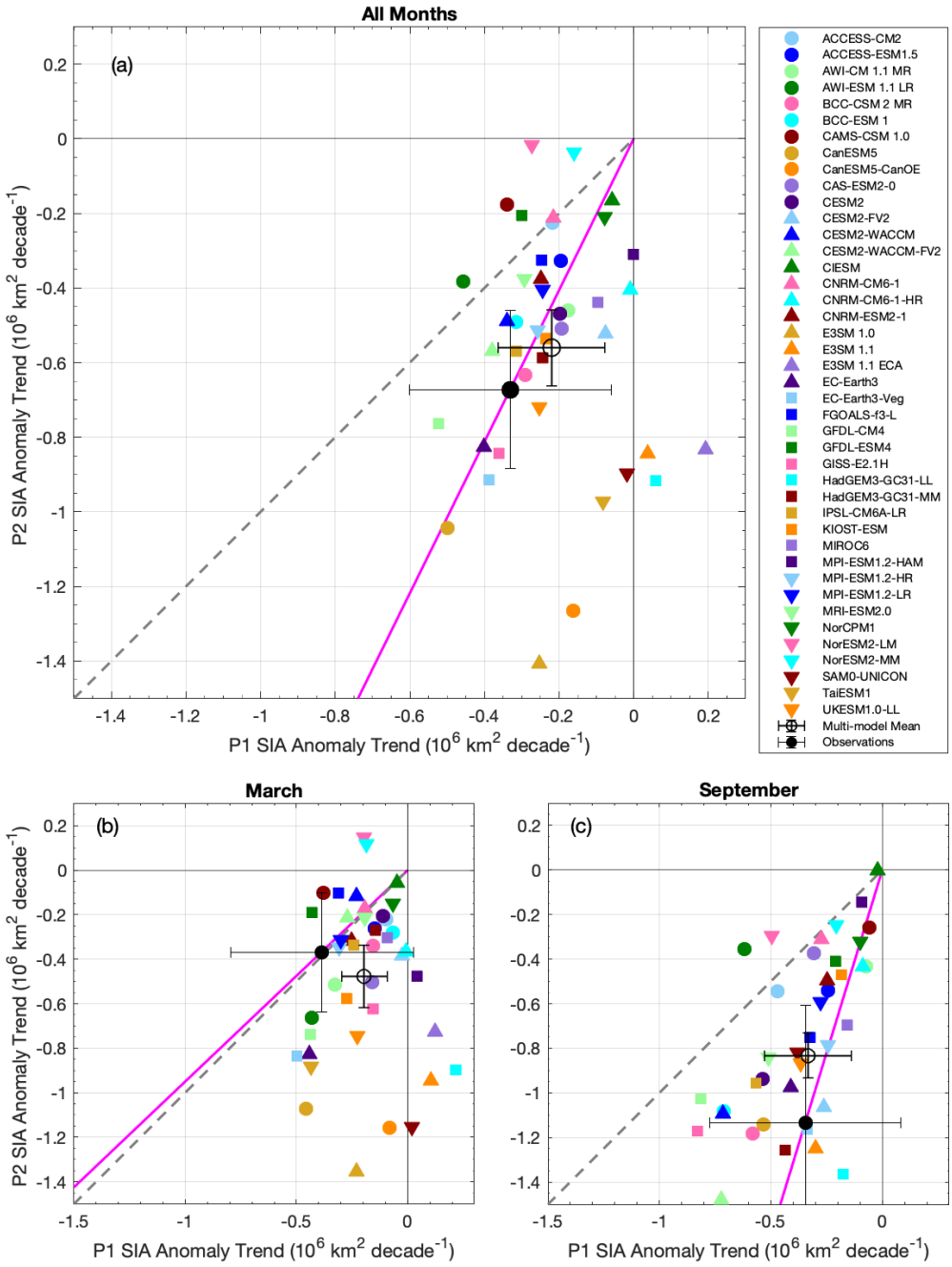
SIA trends from (a) 1979-1996 and (b) 1997-2014 for all individual model ensemble members and the CMIP6 multi-model mean (MM). Error bars indicate the model ±2σ (95%) confidence intervals. The dark and light gray horizontal shading indicates the adjusted ±1σ (68%) and ±2σ observed SIA trend, respectively (following Santer et al. [2008] and Stroeve et al. [2012]).

Figure 15. CMIP6 ensemble member sea ice area trends for periods P1 and P2

Of the total of 268 CMIP6 ensemble members examined here, trends in 145 (54%) for P1 and 68 (25%) for P2 are positive or not statistically different from zero trend, while trends in 128 (48%) and 149 (56%) members fall outside ±2σ of the observed P1 and P2 trend, respectively (Figure 15a,b, respectively). 48 (18%) members from 24 CMIP6 models

have at least one realization with a positive P1 trend, and despite the larger observed negative trend signal of P2, 20 (7%) members from 10 CMIP6 models show a positive P2 SIA trend (Figure 15a,b, respectively).

The ensemble mean SIA anomaly trends for P1 and P2 are shown alongside observations in Figure 16a and Table 3. We consider a model which replicates the observed SIA accelerated trend (i.e., ratio of P2/P1 equals ~2.0, to possess some skill in representing the multi-decadal variability). In view of SIA observational algorithm uncertainty, we note that the individual accelerated trends are all about 2.0 (not shown). The P1 CMIP6 SIA MM trend is -0.22×10^6 km^2 decade^{-1}, with individual models ranging from 0.19×10^6 km^2 decade^{-1} (E3SM 1.1 ECA) to -0.52×10^6 km^2 decade^{-1} (GFDL-CM4). The P2 MM trend is -0.56×10^6 km^2 decade^{-1}, with individual models ranging from -0.02×10^6 km^2 decade^{-1} (NorESM2-LM) to more than twice the observed rate, -1.41×10^6 km^2 decade^{-1} (E3SM 1.0). Collectively, the MM underestimates the observed trend in SIA by 33% during P1 and 16% during P2, yet its accelerated trend of 2.6 is about 25% larger (Table 3). All but six CMIP6 models (AWI-ESM 1.1 LR, CAMS-CSM 1.0, CNRM-CM6-1, GFDL-ESM4, NorESM2-LM, NorESM2-MM) show an accelerated rate of sea ice decline between periods P1 and P2, and therefore qualitatively match the observations (Figure 16). However, the rate of decline during P2 in 29 (69%) CMIP6 models is slower than that of observations; albeit notably faster (i.e., greater than -1.0×10^6 km^2 decade^{-1}) in CanESM5, CanESM5-CanOE, and E3SM 1.0. Still, the accelerated trend in 45% of CMIP6 models is greater than 2.0 (Table 3).



Decadal SIA anomaly trends for the periods 1979 to 1996 (P1 in x-axis) and 1997-2014 (P2 in y-axis) for (a) all months, (b) March, and (c) September for CMIP6 models and passive microwave combined observations. Error bars indicate two standard deviation (95% confidence interval) for each period. The solid magenta line illustrates the observed SIA acceleration ratio (slope of P2/P1= (a) 2.03, (b) 0.95, and (c) 3.28) and the gray dashed line illustrates an acceleration ratio of 1.

Figure 16. CMIP6 ensembles accelerated trends in sea ice area decline

The SIA anomaly linear trends are sensitive to the choice of break point (i.e., year) used to define the period P1 and P2 (Table 5). The absolute error from using an alternate later subperiod (e.g., break point between 1999 and 2000, cf. Shu et al. 2020) for the observed SIA accelerated trend is about 6%, while models show greater sensitivity with an average error of about 37% (ranges from 5% to 90%). Of the 42 CMIP6 models examined here, only six (CAMS-CSM 1.0, EC-Earth3-Veg, GFDL-CM4, KIOST-ESM, MPI-ESM1.2-HAM, NorESM2-MM) show a stronger accelerated trend when we split the time series at year 2000, which indicates these models simulate a stronger decline of sea ice later in the analysis period (Table 5). Thirty-six of 42 models have an identifiable break point that separates a relatively weaker SIA trend followed by a stronger SIA trend at some point during the 1990's and the year 2014 (not shown). And for 30 CMIP6 models, this occurs before our defined break point between P1 and P2. As such, models with break points in the early 1990's will include more negative trend during the defined P1 period than those with break points later in the 1990's. The result is that the accelerated trend in these models tends to be lower (i.e., left of the solid magenta line; Figure 16a) than those with later break points during P1. Only six models (AWI-ESM 1.1 LR, BCC-CSM 2 MR, EC-Earth3-Veg, KIOST-ESM, SAM0-UNICON, TaiESM1) have break points occurring between 1996 to 2000, which best coincides with observations (Comiso et al. 2008; Stroeve et al. 2012b; Serreze and Stroeve 2015).

Table 5. Sensitivity of sea ice area accelerated trend ratio to choice of break point

Model	SP1	SP2	Absolute Error %
ACCESS-CM2	1.04	0.53	49.68
ACCESS-ESM1.5	1.67	1.01	39.77
AWI-CM 1.1 MR	2.62	1.69	35.24
AWI-ESM 1.1 LR	0.83	1.19	42.82
BCC-CSM 2 MR	2.18	1.90	13.11
BCC-ESM 1	1.56	1.38	11.24
CAMS-CSM 1.0	0.52	0.65	23.88
CanESM5	2.09	1.80	13.87
CanESM5-CanOE	7.82	4.55	41.79
CAS-ESM2-0	2.65	2.03	23.42
CESM2	2.38	0.98	58.79
CESM2-FV2	6.84	1.96	71.41

Model	SP1	SP2	Absolute Error %
CESM2-WACCM	1.44	0.94	34.74
CESM2-WACCM-FV2	1.50	1.03	31.42
CIESM	2.80	1.25	55.28
CNRM-CM6-1	0.98	0.88	11.07
CNRM-CM6-1-HR	46.80	4.21	91.01
CNRM-ESM2-1	1.51	1.14	24.93
E3SM 1.0	5.56	2.99	46.17
E3SM 1.1	-23.00**	3.29	-
E3SM 1.1 ECA	-4.34**	6.36	-
EC-Earth3	2.06	1.72	16.40
EC-Earth3-Veg	2.36	2.48	5.11
FGOALS-f3-L	1.32	0.87	34.31
GFDL-CM4	1.46	1.88	29.16
GFDL-ESM4	0.69	0.11	84.36
GISS-E2.1H	2.34	1.66	29.13
HadGEM3-GC31-LL	-15.64**	15.95	-
HadGEM3-GC31-MM	2.39	2.01	16.04
IPSL-CM6A-LR	1.81	1.37	24.25
KIOST-ESM	2.28	3.98	74.80
MIROC6	4.59	2.73	40.45
MPI-ESM1.2-HAM	-	1.51	-
MPI-ESM1.2-HR	1.99	1.67	16.18
MPI-ESM1.2-LR	1.66	1.44	12.94
MRI-ESM2.0	1.29	0.80	37.69
NorCPM1	2.67	1.40	47.64
NorESM2-LM	0.06	-0.29**	-
NorESM2-MM	0.23	0.34	49.15
SAM0-UNICON	49.09	-30.28**	-
TaiESM1	11.91	4.90	58.84
UKESM1.0-LL	2.83	1.84	34.92
CMIP6 MM	2.55	1.80	29.43
CMIP6 average error	-	-	36.97
Observations	2.04	1.91	6.31

SIA anomaly absolute errors between acceleration ratios (P2/P1) resulting from the choice of subperiod years 1979-1996 and 1997-2014 (SP1) and subperiod years 1979-1999 and 2000-2014 (SP2) [(SP1-SP2)/SP1]. SP1 and SP2 acceleration ratios are determined from model ensemble means. The alternate choice of SP2 follows Shu et al. (2020). Note that ** indicates a negative ratio that is the result of a positive P1 trend, the magnitude should not be compared and it is not used to determine % error (-).

Figure 16b,c show SIA accelerated trend analysis for the months of March and September, respectively. March SIA observations indicate a near steady rate of sea ice decline exists for both P1 and P2, -0.38×10^6 km^2 decade^{-1} (Figure 16b). For September, observations show a substantial accelerated rate of SIA decline (P2/P1=3.3) from -0.33×10^6 km^2 decade^{-1} during P1 to -1.12×10^6 km^2 decade^{-1} during P2 (Figure 16c). However, such a strong seasonal distinction is not shown in the CMIP6 SIA MM. Instead,

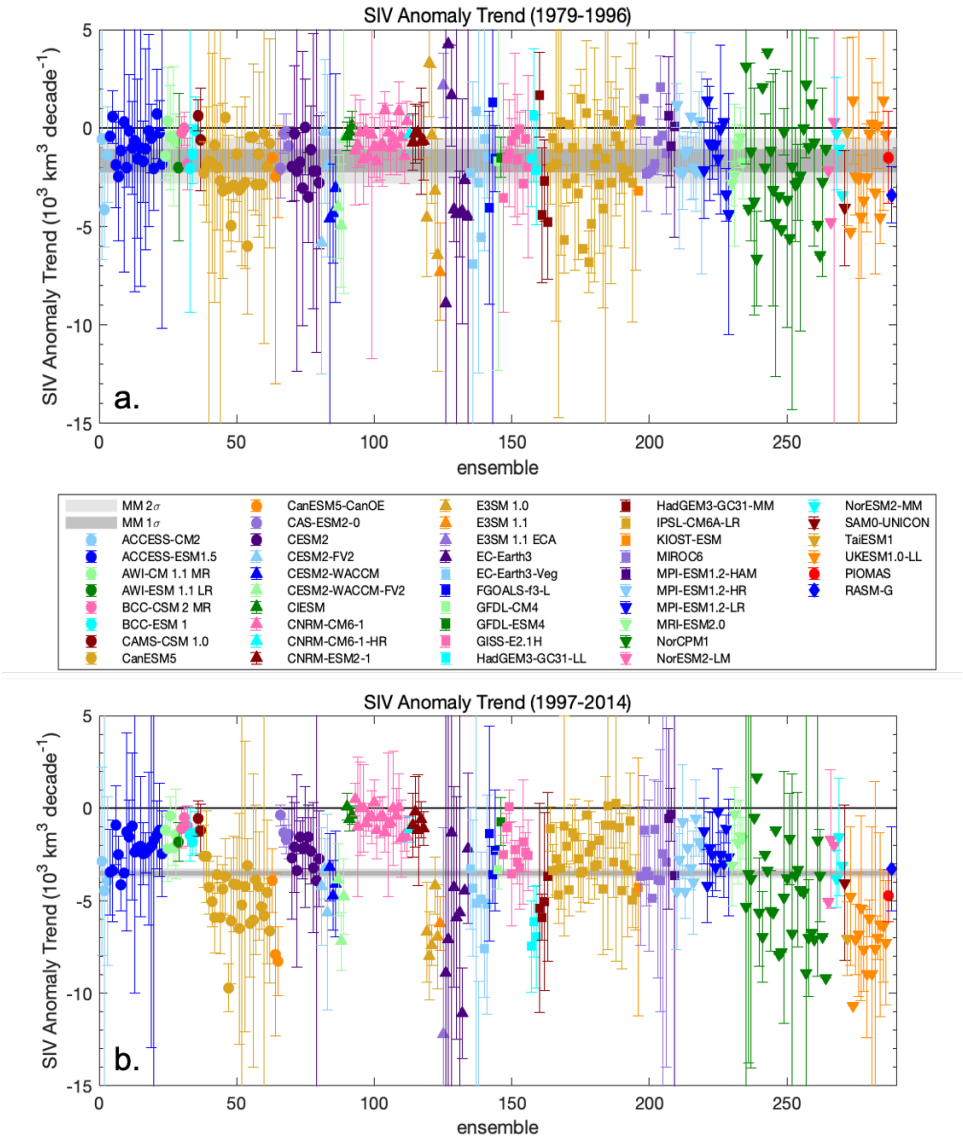
the relative magnitude of accelerated trend is basically the same (P2/P1=2.4) for March and September, though the latter does have larger magnitude negative trends. It is clear the September SIA decline drives the accelerated rate observed for the full ice cover (Stroeve and Notz 2018). However, this is not ubiquitous in CMIP6 models, of which many show a more dampened response meaning smaller September and larger March rates of SIA decline.

Next, we apply similar analysis to evaluate CMIP6 SIV trends and examine whether an accelerated rate of SIV decline is shown (as in the case for PIOMAS, Figure 8b). The CMIP6 SIV MM trend for P1 is -1.66×10^3 km^3 decade^{-1}, about 10% stronger than PIOMAS (-1.49×10^3 km^3), and for P2 the MM trend is -3.55×10^3 km^3 decade^{-1}, which is about 25% weaker than PIOMAS (-4.74×10^3 km^3, Table 2). The result is the CMIP6 SIV MM shows an accelerated trend (P2/P1=2.1) that is two thirds the rate of PIOMAS (P2/P1=3.2). On the other hand, the CMIP6 MM accelerated trend is about twice that of RASM-G, which has no appreciable change in SIV trend between P1 and P2, and can be interpreted as a lower bound (Table 4).

In Figure 17, SIV trends for all CMIP6 individual ensemble members and both SIV reanalyses are shown for periods P1 and P2. As was the case for SIA, large internal variability of simulated SIV trends is exhibited between models and within a model's ensemble members (Figure 17). Also, the model spread is larger when analyzing the shorter climate periods (i.e., P1 and P2) rather than the full SIV historical simulation period (Figure 11b). Trend uncertainty (i.e., error bars) for SIV is even larger than for SIA due to large autocorrelation; SIV AR1 is greater than 0.85 for P1 and greater than 0.80 for P2 in all models (not shown). Both P1 and P2 CMIP6 SIV MM and RASM-G trends are statistically different than zero, while the PIOMAS SIV P1 trend is not (Figure 17). For PIOMAS, this is due in part to large trend uncertainty as result of P1 AR1=0.95.

Of the 268 CMIP6 ensemble members examined here, 247 (92%) members for P1 and 202 (75%) members for P2 have 2σ trend uncertainties within their respective CMIP6 SIV MM 2σ trend uncertainties (Figure 17a,b, respectively). Forty-two (16%) members from 19 CMIP6 models have at least one realization with a positive P1 trend, and despite

the larger negative trend signal of P2, seven members from five CMIP6 models show a positive P2 trend.



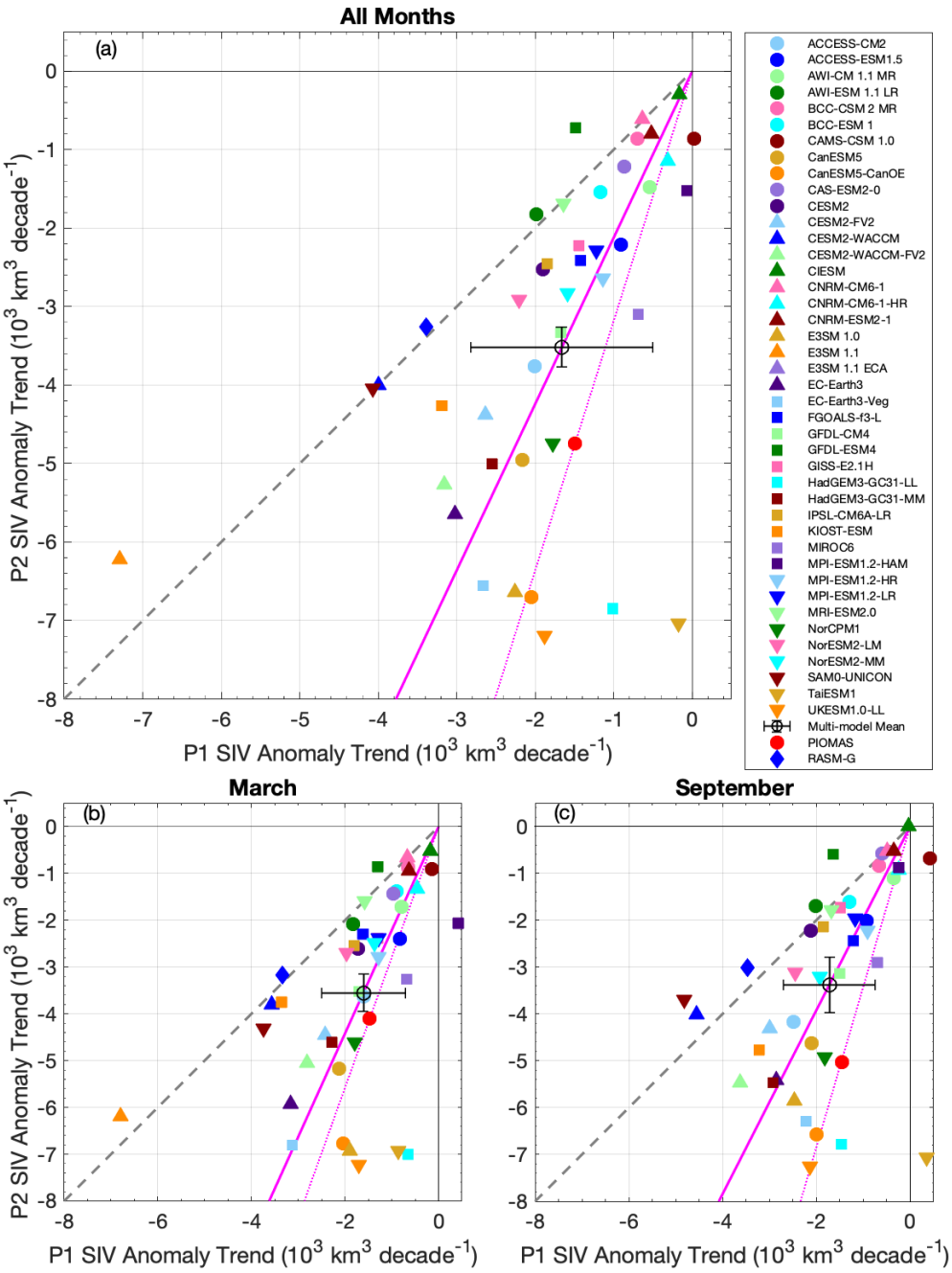
SIV trends from (a) 1979-1996 and (b) 1997-2014 for all individual model ensemble members and PIOMAS and RASM-G reanalyses. Error bars indicate the model ±2σ (95%) confidence intervals. The dark and light gray horizontal shading indicates the adjusted ±1σ (68%) and ±2σ CMIP6 SIV multi-model mean (MM), respectively (following Santer et al. (2008) and Stroeve et al. (2012)).

Figure 17. CMIP6 ensemble member sea ice volume trends for periods P1 and P2

In Figure 18, SIV anomaly trends for CMIP6 ensemble means and the SIV reanalyses are shown for P1 and P2 and listed in Table 4. Individual CMIP6 model trends for P1 range between 2.16×10^3 km^3 decade^{-1} (E3SM 1.1 ECA) and -7.30×10^3 km^3 decade^{-1} (E3SM 1.1). For P2, CMIP6 model ranges between -0.30×10^3 km^{3}decade^{-1} (CIESM) and -12.22×10^3 km^3 decade^{-1} (E3SM 1.1 ECA). All but seven CMIP6 models (AWI-ESM 1.1 LR, CESM-WACCM, CNRM-CM6-1, E3SM 1.0, GFDL-ESM4, MRI-ESM2, SAM0-UNICON) show an accelerated rate of sea ice decline between the periods P1 and P2, and therefore qualitatively match the PIOMAS SIV tendency (Table 4). The rate of decline during P2 in 30 (71%) CMIP6 models is slower than that of PIOMAS; albeit notably faster in nine models (i.e., greater than -6.0×10^3 km^3 decade^{-1}): CanESM5-CanOE, E3SM 1.0, E3SM 1.1, E3SM 1.1 ECA, EC-Earth3-Veg, HadGEM3-GC31-LL, SAMO-UNICON, TaiESM1, and UKESM1.0-LL. The SIV accelerated trend is larger than PIOMAS for about 20% of CMIP6 models, while 29 (69%) CMIP6 models range between the accelerated trend values given by PIOMAS and RASM-G, 3.2 and 1.0, respectively.

Of the 42 CMIP6 models examined here, 37 have an identifiable break point that separates a relatively weaker SIV trend followed by a stronger SIV trend at some point during the 1990's and the year 2014 (not shown). For 33 CMIP6 models, this occurs before our defined break point between P1 and P2 (i.e., 1997). Only four models (AWI-ESM 1.1 LR, BCC-CSM 2 MR, CanESM5-CanOE, NorCPM1) have a break point occurring between 1996 and 2000 and corresponding with the period of SIA accelerated decline from observations.

Figure 18b and 18c show SIV accelerated rate analysis for the months of March and September, respectively. The CMIP6 MM SIV trends for both March and September are similar, showing P1 rates of sea ice decline around -1.6×10^3 km^2 decade^{-1} and P2 rates about -3.5×10^3 km^2 decade^{-1}. The same is true for the PIOMAS P1 trends for both March and September (about -1.6×10^3 km^2 decade^{-1}), but the PIOMAS September P2 trend is 22% stronger than the March P2 trend (-5.0×10^3 km^2 decade^{-1} and -4.1×10^3 km^2 decade^{-1}, respectively). Thus, SIV accelerated trends for the CMIP6 SIV MM are about P2/P1=2.1 and show little seasonality, while PIOMAS accelerated trends range between P2/P1=2.8 and 3.5 (March and September, respectively). So PIOMAS suggests some seasonal enhancement of the SIV decline during P2 while the CMIP6 MM does not.



Decadal SIV anomaly trends for the periods 1979 to 1996 (P1 in x-axis) and 1997-2014 (P2 in y-axis) for (a) all months, (b) March, and (c) September for CMIP6 models and PIOMAS and RASM-G reanalyses. Error bars indicate 2σ (95%) confidence interval for each period. The solid magenta line illustrates the CMIP6 multi-model mean (MM) SIV acceleration ratio (slope of P2/P1= (a) 2.13, (b) 2.21, and (c) 1.96), the dotted magenta line illustrates the PIOMAS SIV acceleration ratio (slope of P2/P1=3.18 (b) 2.71, and (c) 3.42), and the gray dashed line illustrates an acceleration ratio of 1 (e.g., RASM-G). Note that E3SM 1.1 ECA is not shown because it is out of the axes range.

Figure 18. CMIP6 ensembles accelerated trends in sea ice volume decline

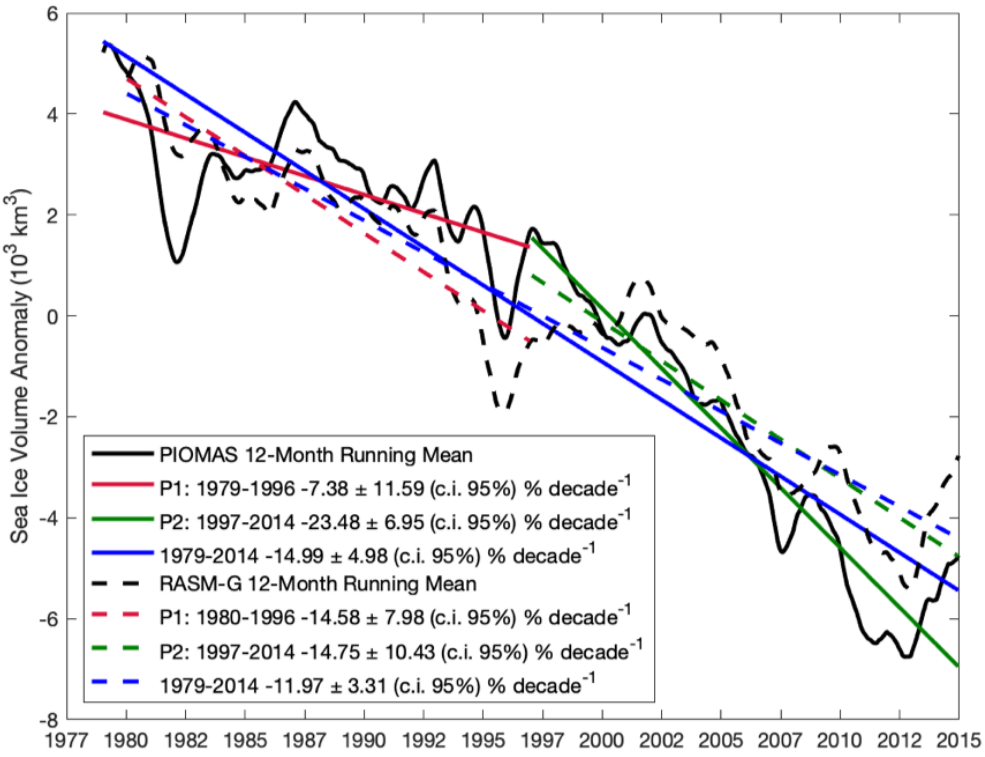
D. DISCUSSION

Incorporating relevant observationally constrained metrics to qualify sea ice simulations is important to help identify models suited for further process-level analysis. Here we used an analysis of the accelerated rate of sea ice decline between the periods P1 and P2 (Figure 16 and Figure 18) to help qualify the CMIP6 models' capability to represent the complex climate interactions that have contributed to the Arctic amplified response to global climate warming. Additionally, fully coupled simulations from the better performing models may provide further insights into understanding potential linkages between the Arctic and the mid-latitude weather and climate. While SIA simulations are reasonably well constrained against passive microwave observations, simulations of the three dimensional sea ice state (i.e., SIV) are not (Zygmuntowska et al. 2013). As such, our criteria for identifying quality simulations of SIV is less certain and requires some discussion.

Internal variability of the climate, as well as of individual climate model simulations, must always be considered when comparing against observations. Specifically, it is known that internal climate variability permits a range of possible outcomes of Arctic sea ice states, of which the observed state is but one realization (England et al. 2019). Such variability can account for as much as 50% of the September SIE trend in the pan-Arctic sea ice loss since 1979 (Stroeve et al. 2007; Kay et al. 2011; Stroeve et al. 2012b), and from 10% to 75% of regional SIC trends (England et al. 2019). For CMIP6 models, Shen et al. (2021) inferred about 22% of the 1979-2014 September SIE trend can be attributed to model internal variability, assuming no bias in model response to external forcing. While not quantified here, we demonstrated that analysis of shorter time series (e.g., less than 20 years in the analysis of accelerated rate of SIA/SIV decline) contributed to large trend uncertainties (Kay et al. 2011) in CMIP6. However, the next iteration of CMIP should have a longer overlapping period of historical experiment and SIA observations, and therefore reduced trend uncertainty. Additionally, we intentionally analyzed CMIP6 ensembles and the CMIP6 MM in order to reduce the impact from internal model variability.

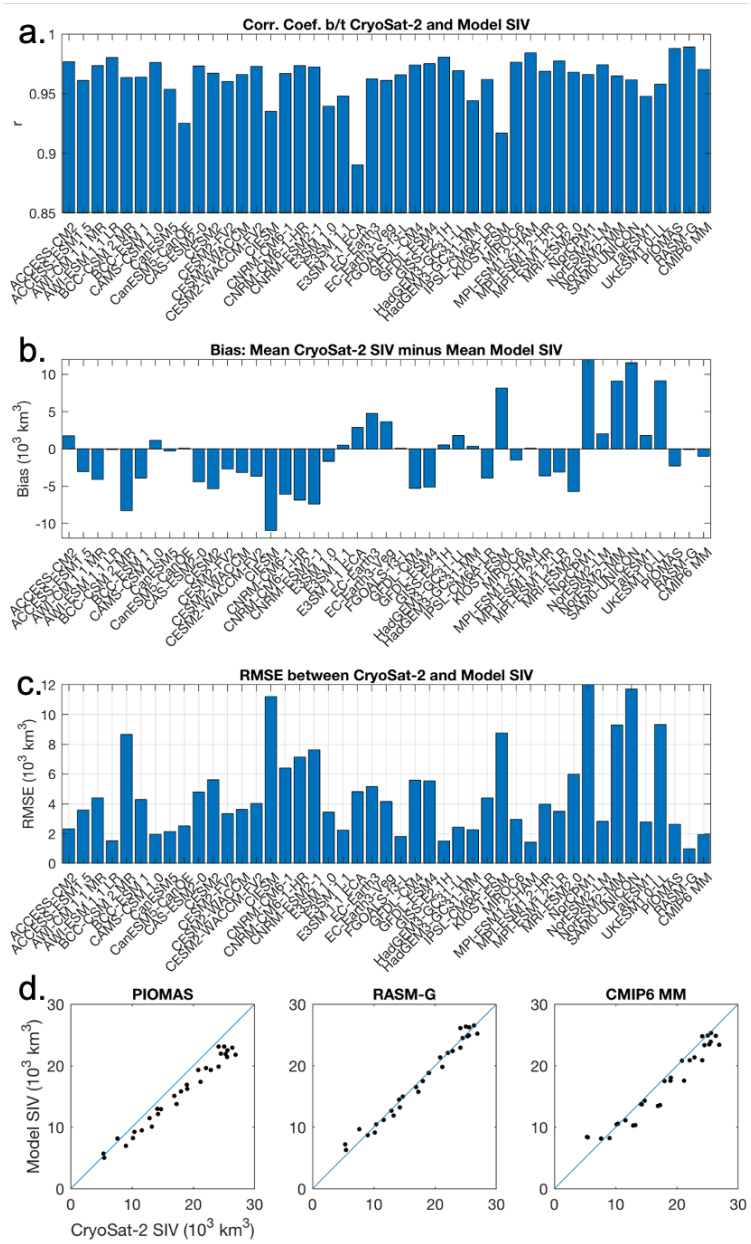
Due to large uncertainties in estimated SIT observations and the corresponding SIV time series, we examined two SIV reanalyses, RASM-G and PIOMAS, alongside the CMIP6 models. We note that the spatial and temporal SIT uncertainties vary widely within PIOMAS, with Schweiger et al. (2011) reporting for 1979-2010 a conservative volume trend uncertainty estimate of 1.0×10^3 km^3 decade^{-1} based on three PIOMAS integration runs. The PIOMAS domain-wide SIV bias relative to the SIT CDR (Lindsay 2010) has been estimated at -2.8×10^3 km^3 for March and -1.5×10^3 km^3 for October, or about 10% of the total SIV over the same period. In Figure 19, we show the 12-month running means of PIOMAS and RASM-G SIV anomaly with linear trends. The RASM-G SIV mean is 21.0×10^3 km^3 (Figure 12b) and linear trend is -2.52×10^3 km^3 decade^{-1} for 1979-2014 (Table 4), which is 4% above and 17% slower compared to the PIOMAS mean and trend, respectively. The RASM-G simulated range of 15-27$\times10^3$ km^3 (Figure 12a) is another expression of slightly thicker ice in P2, yet well correlated SIV evolution compared to the PIOMAS range, with a 20% smaller standard deviation (Figure 13b).

A short period of CryoSat-2 SIV observations (October to April during 2010-2014) overlap the CMIP6 historical period and offer an observational constraint for a portion of period P2, albeit too short for more than a qualitative interpretation here. Against the CryoSat-2 SIV time series, RASM-G shows little bias and root-mean-square error (RMSE), PIOMAS shows notable bias of -2.3×10^3 km^3 and RMSE 2.6×10^3 km^3, and the CMIP6 SIV MM splits the difference (i.e., -1.0×10^3 km^3 bias and RMSE 1.9×10^3 km^3; Figure 20b,c). The PIOMAS negative SIV bias increases against higher observed SIV values (Figure 20d), thus the bias is most pronounced during the months of largest SIT (e.g., March-May). This likely contributes to a P2 trend that is too strong resulting from overly deep troughs in the SIV anomaly during 2010-2014 (Figure 8b and Figure 19).



The 12-month running mean of SIV anomalies for PIOMAS (solid black) and RASM-G (dashed black) referenced to the period 1979-2014 and 1980-2014, respectively. The SIV anomaly is used to determine linear fit with 95% confidence intervals (c.i.) for the periods 1979 to 2014 (blue), 1979 to 1996 (P1, red), and 1997-2014 (P2, green).

Figure 19. 12-month running means of PIOMAS and RASM-G sea ice volume anomaly



Comparison of SIV (a) correlation coefficient (*r*), (b) mean bias, and (c) root-mean-square error (RMSE) for CMIP6 models and PIOMAS and RASM-G reanalyses against available months of CryoSat-2 satellite observations in October-April during the period of 2000-2014. (d) Scatter plots of PIOMAS, RASM-G and CMIP6 multi-model mean (MM) against CryoSat-2. The blue diagonal line indicates a 1 to 1 ratio between model and CryoSat-2.

Figure 20. Comparison of models against CryoSat-2 sea ice volume observations (October to April during 2010-2014)

The P1 SIV trend for RASM-G is twice as strong as that for PIOMAS, and it is larger than the CMIP6 MM trend, while the P2 trend for RASM-G is 12% smaller than the

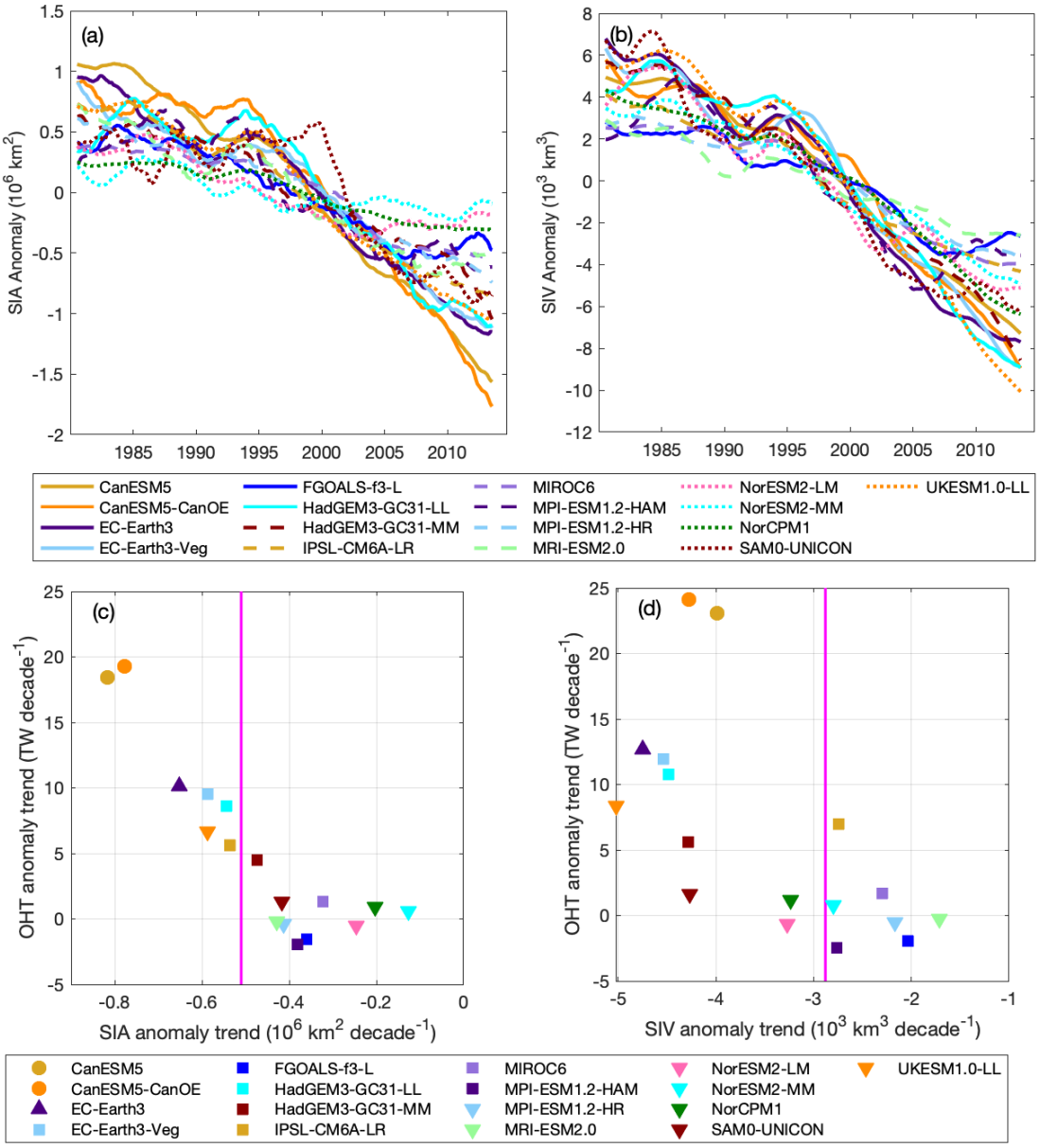
MM trend but 34% less than PIOMAS trend for P2 (Table 4). The overall result is only a slight increase in strength between P1 and P2 in RASM-G, which is half of the CMIP6 MM acceleration ratio but close to 3 times smaller than the one for PIOMAS. Given the reasonable RASM-G skill in simulating both SIV time series and SIT spatial patterns (Watts et al. 2021) and the reported bias in PIOMAS SIV combined with uncertainty in its trend estimates, RASM-G offers an alternative to the PIOMAS perspective on the mean state and evolution of SIV in the Arctic. Further constraints on the historical evolution of SIT and SIV (i.e., longer period of CryoSat-2 combined with ICASat-2 observations) will assist future model intercomparison projects.

The Arctic sea ice decreases near linearly to global mean temperature rise (Gregory et al. 2002; Rosenblum and Eisenman 2016) and cumulative CO_2 emissions (Zickfeld et al. 2012; Notz and Stroeve 2016), revealing temperature as the primary driver for sea ice decline. For CMIP6, 70% of SIA simulations have CO_2 sensitivity within plausible ranges, while only 28% have plausible sea ice loss per degree of warming (SIMIP Community 2020) for the period 1979-2014. With respect to CO_2 sensitivity, the percent of plausible models is similar to the number of plausible SIA and SIV trends in our analysis (60% and 80%, respectively; Figure 9c and Figure 12c). It might be informative to examine in a future separate study whether the plausible models from SIMIP Community (2020) respond with the same magnitudes over the shorter periods P1 and P2 used in our accelerated trend analysis. Of the 13 models identified by SIMIP Community (2020) with plausible sea ice sensitivity to warming and a plausible amount of decline, only three have accelerated trend rates matching observations (our Table 1; BCC-CSM 2 MR, CNRM-CM6-1-HR, GISS-E2.1H).

The regions of largest observed SIC decline before about 2000 was over the East Siberian, Chukchi, and Beaufort seas, but has since shifted to the Laptev, Kara, and Barents seas region. This spatial pattern explains the accelerated decline in sea ice cover, but is not well represented by CMIP6 models (Shu et al. 2020), thus indicating insufficient representation of physical processes there. For example, a number of studies describe an increase in Atlantic water OHT into the Arctic Ocean beginning around 2000 (Tsubouchi et al. 2021), so called Atlantification (Årthun et al. 2012). Prior to this, Arctic energy

budget studies show near equilibrium heat exchanges (i.e., meridional OHT and vertical heat exchange between ocean-atmosphere) over the period 1979-2001 (Serreze et al. 2007). In contrast, Mayer et al. (2016) showed that increased OHT preconditioned the Arctic sea ice minimum of 2007, after which positive radiative flux anomalies drove the large energy signal through 2015. While detailed analysis of possible causality of the sea ice decline is beyond the scope of this paper, preliminary evidence from our analysis suggests increased northward OHT in CMIP6 models could be closely linked to their simulations sharing the characteristic accelerated sea ice decline.

About half of the CMIP6 models analyzed here provided the latitude- and depth-integrated variable northward OHT (i.e., *hfbasin*) from all ocean processes (e.g., resolved advective transport, diffusion, etc.; Griffies et al. 2016) which we used for a first order examination of oceanic forcing on sea ice (Figure 21). A reasonably clear direct relationship is shown between models with positive OHT anomaly trends at 80° N (CanESM5, CanESM5-CanOE, EC-Earth3, EC-Earth3-Veg, HadGEM3-GC31-LL, HadGEM3-GC31-MM, IPSL-CM6A-LR, and UKESM1.0-LL) and the magnitude of negative SIA anomaly trends (Figs. 9a and 9c, respectively; $r \leq -0.95$, $p < 0.05$). For these models, more Atlantic Ocean heat is carried poleward across 80 ºN from the shallow shelf regions of the Barents and Kara seas as well as through Fram Strait and into the Arctic Eurasian basin. This suggests that the magnitude of Atlantic OHT may be a key process underlying CMIP6 models' ability to simulate the observed rates of SIA decline. Additionally, these models have SIA anomaly trends at or exceeding observations (Figure 9c). The relationship between OHT anomaly and SIV anomaly is not as clear as for SIA (Figure 21b,d, respectively). We speculate that accelerated loss of SIV in CMIP6 may result from other dominant processes, such as the forced greenhouse warming. We propose further process-level examination of CMIP6 model's simulated OHT may advance understanding of its impact on sea ice decline.



36-month running means of (a) SIA anomaly and (b) SIV anomaly for a subset of CMIP6 model simulations. (c) Scatter plot of northward global OHT anomaly linear trends at 80° N (y-axis) and pan-Arctic SIA anomaly linear trends (x-axis) for the period 1979-2014. The vertical magenta line illustrates the combined passive microwave observations trend (-0.51×10^6 km^2 decade^{-1}). (d) Scatter plot of northward global OHT anomaly linear trends at 80° N (y-axis) and pan-Arctic SIV anomaly linear trends (x-axis) for the period 1979-2014. The vertical magenta line illustrates the CMIP6 multi-model mean trend (-2.8×10^3 km^3 decade^{-1}).

Figure 21. Oceanic heat forcing on pan-Arctic sea ice

E. CONCLUSIONS

Most CMIP6 models simulate an accelerated rate of decline in SIA (86%) and SIV (88%) starting in the mid-1990s, but it starts earlier than SIA observations in a majority of models. The majority of CMIP6 models underestimate the SIA trend, and those that do exceed the rate of decline mostly have positive SIA bias. The majority of CMIP6 models do not simulate the SIV accelerated rate of decline shown in PIOMAS reanalysis. Given that SIV simulations are not well constrained by observations, we offer that CMIP6 SIV MM trends may be a suitable reference estimate. Internal variability is large among individual CMIP6 ensemble member simulations, particularly when examining shorter time periods (i.e., 18 years versus 36 years).

F. DATA AVAILABILITY

CMIP6, EUMETSAT, NSIDC, and PIOMAS data used for this study can be acquired from the links provided in Section III.B.1. The RASM-G data can be acquired from Naval Postgraduate School (https://nps.app.box.com/folder/139647168752?s=xyp563ee40w6lffn4718zr52a6m7rtcc).

G. ACKNOWLEDGMENTS

We acknowledge the WCRP, which, through its Working Group on Coupled Modeling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the ESGF for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. We acknowledge partial support from the following programs: US Navy (to MW), the Department of Energy (DOE) Regional and Global Model Analysis (RGMA), the Office of Naval Research (ONR) Arctic and Global Prediction (AGP) and National Science Foundation (NSF) Arctic System Science (ARCSS) (to WM, YL, JCK), and Ministry of Science and Higher Education in Poland (to RO).

IV. A SPATIAL EVALUATION OF ARCTIC SEA ICE AND REGIONAL LIMITATIONS IN CMIP6 HISTORICAL SIMULATIONS

This chapter was published in *Journal of Climate* (1 August 2021), DOI: 10.1175/JCLI-D-20-0491.1. Except for figure numbering, the formatting has been retained as submitted. As the main author of the work, I made the major contributions to the research and writing. Co-authors include W. Maslowski[3],Y.J. Lee[3], J. Clement Kinney[3], and R. Osinski[4].

A. INTRODUCTION

The Arctic is warming at twice the rate of the rest of the planet, evidenced by rising SATs in response to GHGs (Serreze et al. 2009; Serreze and Barry 2011; Taylor et al. 2013; IPCC 2019). One of the most striking reflections of this Arctic amplification (Serreze and Francis 2006) is the accelerated decrease in SIE (Meier et al. 2017) observed for each month of the year over the satellite record since 1978 (Serreze and Barry 2011; Stroeve and Notz 2018). Changes in the sea ice cover alter the surface albedo, the upper ocean heat content, and thus the surface energy budget of the Arctic Ocean (Jackson et al. 2011; Timmermans et al. 2018). In addition, a diminishing Arctic sea ice cover increases the air-sea exchange of momentum, surface buoyancy flux, and freshwater content (Parkinson et al. 1987; Rampal et al. 2011; Proshutinsky et al. 2019). Hence, a better understanding of the sea ice reduction is needed to improve climate predictions and projections.

The primary objective of this study is to assess and guide improvements of outstanding pan-Arctic as well as regional limitations in historical simulations of sea ice, by employing a combination of common and new metrics on a subset of state-of-the-art Earth System Models (ESMs) participating in CMIP6 (Eyring et al. 2016). This is motivated in part by the need to better understand the complex operation of the Earth system under climate forcing and in part to convey confidence in model skill to project the future. The latter is based on the argument that a model's ability to simulate 'known' mean climate state, trends, variability and extremes raises confidence in its projections (Randall

[3] Department of Oceanography, Naval Postgraduate School

[4] Institute of Oceanology of Polish Academy of Sciences, Sopot, Poland

et al. 2007; Massonnet et al. 2012). However, past model performance of particular observables (e.g. pan-Arctic SIE) alone is not sufficient to describe the quality of a model's future projection, due in large part to internal variability, observational uncertainty, and model tuning (Notz 2015). Internal climate variability itself allows for a range of possible outcomes of Arctic sea ice states, of which the observed state is but one realization (Notz 2015; England et al. 2019).

While modern ESMs generally capture much of the physics and the downward trends of the observed Arctic SIE, they have so far underestimated its acceleration in response to GHG forcing and increasing global SAT (e.g. Winton 2011; Massonnet et al. 2012; Stroeve and Notz 2015; Rosenblum and Eisenman 2017; SIMIP Community 2020). In addition, according to observational and model reconstructed estimates, the negative trend in SIT and SIV has been even stronger than that in SIE (Kwok and Rothrock 2009; Schweiger et al. 2011; Maslowski et al. 2012; Stroeve et al. 2014). This aspect alone corroborates the need for observationally-constrained metrics of SIT and SIV for model evaluation as they allow for additional insights into regional and seasonal biases and overall quality of sea ice simulations. At the same time, observations of sea ice from passive microwave satellites contain internal and algorithm uncertainties (Screen 2011; Eisenman et al. 2014; Ivanova et al. 2014; Meier et al. 2014), while pan-Arctic and long-term estimates of SIT and SIV are not readily available from satellites as they cover relatively short time period. In particular, uncertainty in satellite-derived estimates of pan-Arctic SIT distribution and summer SIC pose considerable challenges (Zygmuntowska et al. 2013; Kwok 2018).

In this study, we expand on the published CMIP6 Arctic sea ice analyses (Shu et al. 2020; SIMIP Community 2020; Shen et al. 2021) to isolate specific spatial model limitations. In particular, we examine SIE, SIT and SIV from a subset of 12 CMIP6 models, that provides a good representation of the whole set (as discussed in Section B), for the period 1979-2014. The integrated ice-edge error (IIEE; Goessling et al. 2016) and Spatial Probability Score (SPS; Goessling and Jung 2018) analyses, referred to collectively as ice edge analysis, are introduced to identify regions commonly challenging for the majority or individual CMIP6 models to accurately replicate sea ice conditions. The rest of this paper

is organized as follows: in Section B we describe data and methods; in Section C results are presented, emphasizing a) the mean state and decline in pan-Arctic sea ice, b) the simulated SIT spatial distribution, and c) a regional ice edge analysis; and in Section D are the discussion and conclusions.

B. DATA AND METHODS

1. Model Sea Ice Output

Sea ice outputs for a subset of CMIP6 models (Table 6) were retrieved from the ESGF repository (https://esgf-node.llnl.gov/search/cmip6/). Our study used the historical experiment data to evaluate the capability of the participating CMIP6 models to represent Arctic sea ice during the recent past. These experiments were initialized and forced with common time-dependent observations, including anthropogenic short-lived climate forcing, carbon emissions, land use, and GHG historical concentrations, for the time period of 1850 to 2014 (Notz et al. 2016; Eyring et al. 2016).

The 12 models selected for our study range from ± two standard deviations of the CMIP6 multi-model mean SIA and SIV, as shown in Table S3 of the online supplementary material of SIMIP Community (2020). We chose a limited number of models, representative of the whole with respect to multi-model mean and spread (Table 7), to highlight the range of simulated sea ice biases and limitations and a workable subset to present the utility of ice edge analysis. Additional criteria for model selection included the availability of sea ice variables for spatial analysis (e.g., SIT and SIC) and a representation of different sea ice model components used in CMIP6 simulations (e.g., NEMO-LIM, CICE, MPAS-Sea ice, GELATO).

Table 6. CMIP6 model and RASM metadata used in this study

Model Name	Label (No. of ensemble members)	Organization	Components	Resolution (km) A/O/I	Model Variable
			Atmosphere (A) / Ocean (O) / Sea Ice (I)		
Canadian Earth System Model (CanESM) 5	CanESM5 (25)	Canadian Centre for Climate Modelling and Analysis	Canadian Atmosphere Model (CanAM5)/ Nucleus for European Modelling of the Ocean (NEMO) v3.4.1/NEMO-Louvain-la-Neuve Sea Ice Model (LIM2) DOI: 10.22033/ESGF/CMIP6.3610; v20190429	500/100/100	siconc simass
Community Earth System Model (CESM) 2	CESM2 (11)	National Center for Atmospheric Research (NCAR)	Community Atmosphere Model (CAM6)/ Parallel Ocean Program version 2 (POP2)/Los Alamos National Laboratory Sea Ice Model version 5.1 (CICE5.1) DOI: 10.22033/ESGF/CMIP6.7627; v20190308	100/100/100	siconc sivol
CESM Whole Atmosphere Community Climate Model (WACCM)	CESM2-WACCM (3)	NCAR	Whole Atmosphere Community Climate Model (WACCM6)/POP2/CICE5.1 DOI: 10.22033/ESGF/CMIP6.10071; v20190227	100/100/100	siconc sivol
National Center Meteorological Research (CNRM) Earth System Model 2	CNRM-ESM2-1 (5)	CNRM-CERFACS	Action de Recherche Petite Echelle Grande Echelle (ARPEGE 6.3)/NEMO3.6/Global Experimental Leads and Ice for Atmosphere and Ocean (GELATO 6.1) DOI: 10.22033/ESGF/CMIP6.4068; v20181206	250/100/100	siconc sivol
Energy Exascale ESM (E3SM)	E3SM 1.0 (5)	Department of Energy E3SM Project	E3SM Atmosphere Model (EAM)/Model for Prediction Across Scales (MPAS)-Ocean/ MPAS-Sea Ice DOI: 10.22033/ESGF/CMIP6.4497; v20190926	100/50/50	siconc* simass*
Flexible Global Ocean-Atmosphere-Land System Model (FGOALS)	FGOALS-f3-L (3)	Institute of Atmospheric Physics (IAP)	FAMIL2.2/State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG)/IAP Climate Ocean Model, (LICOM3.0)/CICE4.0 DOI: 10.22033/ESGF/CMIP6.3355; v20191031	100/100/100	siconc sivol

Model Name	Label (No. of ensemble members)	Organization	Components	Resolution (km) A/O/I	Model Variable
			Atmosphere (A) / Ocean (O) / Sea Ice (I)		
Geophysical Fluid Dynamics Laboratory (GFDL) ESM 4	GFDL-ESM4 (1)	National Oceanic and Atmospheric Administration	GFDL-AM4.1/GFDL-OM4p5 Modular Ocean Model (MOM6)/GFDL-SIM4p5 Sea Ice Simulator (SIS2.0) DOI: 10.22033/ESGF/CMIP6.8597; v20180701	100/50/50	siconc sivol
Goddard Institute of Space Studies (GISS)	GISS-E2.1H (10)	National Aeronautics and Space Administration	GISS-E2.1/Hybrid Coordinate Ocean Model (HYCOM)/GISS Sea Ice DOI: 10.22033/ESGF/CMIP6.7128; v20191003	250/100/100	siconc* areacello sivol
Institut Pierre Simon Laplace (IPSL) Climate Model 6	IPSL-CM6A-LR (32)	IPSL Climate Modelling Centre	LMDz/NEMO-OPA/NEMO-Louvain-la-Neuve Sea Ice Model (LIM3) DOI: 10.22033/ESGF/CMIP6.5195; v20180803	250/100/100	siconc simass
Model for Interdisciplinary Research on Climate (MIROC) 6	MIROC6 (10)	Japan Agency for Marine-Earth Science and Technology	Center for Climate System Research (CCSR) Atmospheric General Circulation model (AGCM)/CCSR Ocean Component model (COCO4.9)/COCO4.9 DOI: 10.22033/ESGF/CMIP6.5603; v20181212	250/100/100	siconc simass
Max Planck Institute (MPI)-ESM	MPI-ESM1.2-HR (10)	Max Planck Institute for Meteorology	ECHAM6.3/MPIOM1.63/unnamed DOI: 10.22033/ESGF/CMIP6.6595; v20190710	100/50/100	siconc sivol
United Kingdom ESM (UKESM) 1	UKESM1.0-LL (14)	Met Office Hadley Centre	MetUM Hadley Centre Global Environment Model (HadGEM3-GA7.1)/NEMO-HadGEM3-GO6.0/CICE-HadGEM3-GSI8 DOI: 10.22033/ESGF/CMIP6.6113; v20200310	250/100/100	siconc sivol
Regional Arctic System Model (RASM) Sea Ice-Ocean	RASM-G (1)	Naval Postgraduate School	Japanese 55-year Reanalysis (JRA55-do)/ POP2/CICE6	50/9/9	siconc sivol

CMIP6 model and Regional Arctic System Model (RASM-G) simulations used in this analysis. The number of ensemble members is indicated in parentheses under the column Label. Native model grids are used unless otherwise annotated with * indicating re-grid. The gridded variables used for this study are sea ice area fraction (siconc), sea ice volume per area (sivol), and sea ice mass per area (simass).

Table 6 summarizes the model metadata, climate system components with nominal resolutions, and the model variables used for analysis. The horizontal resolution of the CMIP6 sea ice model component varies between 50 km to 500 km, with the majority still using a relatively coarse resolution of one-degree (~100 km). All but one CMIP6 model in our study, GFDL-ESM4, produced output from multiple ensemble members (up to 32). Unless otherwise indicated, all values presented in figures and tables show ensemble means of individual models.

In addition to our CMIP6 analysis, we analyze sea ice model output from the Regional Arctic System Model (RASM; Maslowski et al. 2012; Roberts et al. 2015; Hamman et al. 2016; Cassano et al. 2017). The forced sea ice – ocean model configuration, which we term RASM-G (Figure 22 shows RASM-G domain), was used for a high spatial resolution (~9 km) hindcast simulation, results of which are presented here. It was initialized after a 57-year spin up and forced with the Japanese 55-year atmospheric surface reanalysis data for driving ocean-sea ice models (JRA55-do; Tsujino et al. 2018). This RASM-G hindcast simulation provides a complementary reconstruction of multidecadal sea ice conditions for the period of 1980 to 2014, in addition to the remotely-sensed observations and PIOMAS (Zhang and Rothrock 2003) reanalysis data.

Due to the lack of persistent SIT observations over the Arctic, we use the PIOMAS sea ice reanalysis as a SIT 'observational' proxy reference, following a number of previous studies favorably comparing PIOMAS results against thickness observations from submarines, satellites and airborne (Zhang and Rothrock 2003; Schweiger et al. 2011; Stroeve et al. 2014). The PIOMAS version 2.1 SIV and effective SIT were retrieved from the Polar Science Center at University of Washington (http://psc.apl.uw.edu/research/projects/arctic-sea-ice-volume-anomaly/data/) in order to evaluate SIT simulations of CMIP6 models. Here we use SIV time series and monthly mean gridded effective SIT (Zhang and Rothrock 2003; Schweiger et al. 2011).

Table 7. Sea ice area and volume mean and standard deviation for the period 1979-1998 for models and observational references

Model	March SIA Mean	s.d.	September SIA Mean	s.d.	March SIV Mean	s.d.	September SIV Mean	s.d.
CanESM5 (25)	15.97	0.29	6.81	0.36	34.46	1.44	18.51	1.47
CESM2 (11)	13.96	0.11	4.30	0.50	24.90	1.24	9.69	1.50
CESM2-WACCM (3)	14.31	0.20	5.81	0.58	31.56	2.42	17.34	3.01
CNRM-ESM2-1 (5)	15.41	0.18	4.82	0.30	18.65	0.43	3.25	0.38
E3SM 1.0 (5)	20.57	0.23	5.44	0.38	41.71	1.92	32.74	2.54
FGOALS-f3-L (3)	16.48	0.22	4.90	0.34	30.08	1.00	11.75	0.95
GFDL-ESM4 (1)	14.10	0.42	5.88	0.48	21.90	1.09	8.34	1.44
GISS-E2.1H (10)	21.03	0.14	10.58	0.64	31.46	0.93	10.24	1.16
IPSL-CM6A-LR (32)	15.14	0.19	4.96	0.48	28.61	1.31	10.04	1.44
MIROC6 (10)	12.07	0.07	5.37	0.17	27.03	0.56	14.79	0.68
MPI-ESM1.2HR (10)	14.20	0.18	4.36	0.22	26.95	0.86	7.88	0.75
UKESM1.0-LL (14)	16.26	0.17	7.54	0.32	47.41	1.32	31.03	1.73
CMIP6 subset mean	15.79	2.63	5.90	1.75	30.42	7.51	14.47	8.56
CMIP6 multi-model mean*	15.46	2.01	6.07	1.55	30.99	9.5	14.55	10.47
RASM-G (1)	14.58	0.24	5.16	0.59	30.73	1.71	12.46	1.84
Observations* (3)	14.35	0.54	5.97	0.66	-	-	-	-
PIOMAS	-	-	-	-	29.28	1.26	14.20	1.49

Monthly mean ($\times 10^6$ km^2) and standard deviation (s.d.) for SIA ($\times 10^6$ km^2) and SIV ($\times 10^3$ km^3) for the period 1979-1998 for March and September of all ensemble members (# in parentheses) per model of our subset of CMIP6 models. The CMIP6 multi-model mean for all CMIP6 models and SIA passive microwave observations from SIMIP community (2020) for the same period are denoted by *.

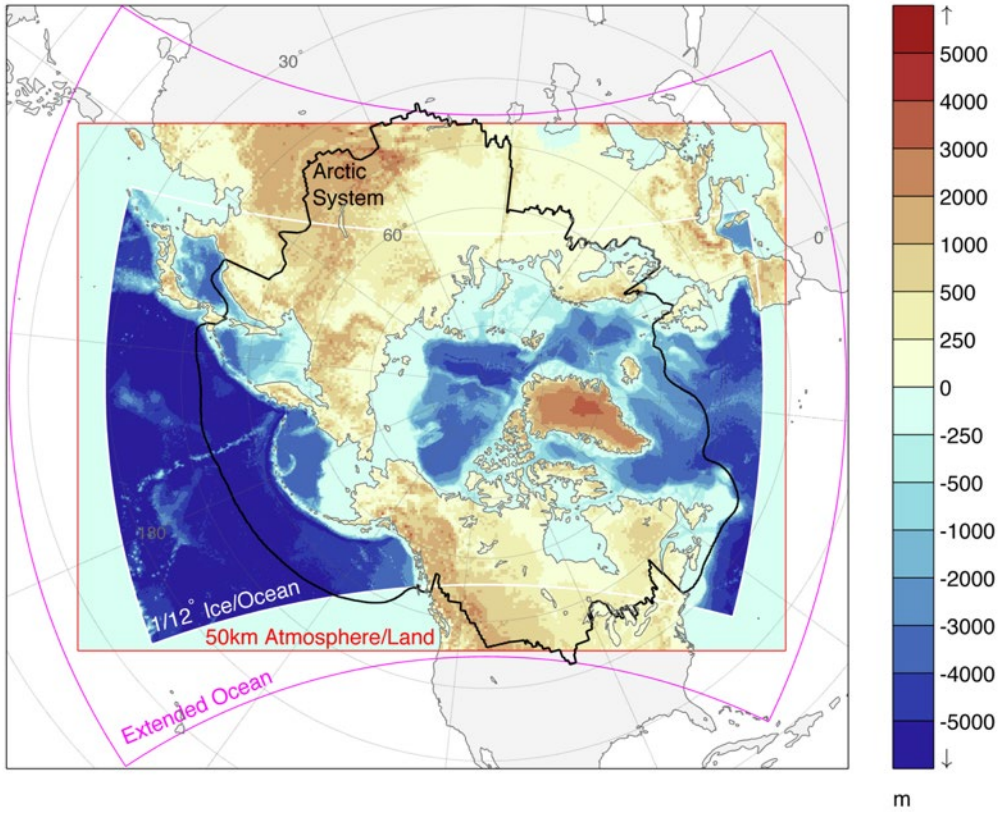


Figure 22. Domain map of Regional Arctic System Model. Adapted from Maslowski et al. (2012).

2. Observational Sea Ice Data

Monthly mean SIC data for the period of 1979-2014 were retrieved from the NOAA)/NSIDC and the EUMETSAT OSI-SAF. A mean observational SIC was determined by combining the NT and BS SIC from CDR of Passive Microwave SIC, Version 3 (Meier et al. 2017; ftp://sidads.colorado.edu/pub/DATASETS/NOAA/G02202_V3) and the OSI-SAF team SIC estimate (OSI-450; Lavergne et al. 2019; https://doi.org/10.15770/EUM_SAF_OSI_0008). We use this combined SIC estimate as the primary observational reference for SIE time series analysis. For the ice edge analysis, we use the merged NT/BS SIC as observational reference data. The spread in observational estimates as a result of algorithm differences can be interpreted as the observational uncertainty, or absolute uncertainty (Meier and Stewart 2019; SIMIP Community 2020).

Monthly mean SIT estimates from ICESat were retrieved from NSIDC (Yi and Zwally 2009; https://nsidc.org/data/NSIDC-0393/versions/1) for the period 2003-2008, and CryoSat-2 from the Alfred Wegener Institute (Hendricks and Ricker 2019; https://spaces.awi.de/display/SIRAL) for the period 2010-2014. Satellite SIT observations are available only during the colder months from October to April, which is why we limit our SIT analysis to the month of March. For ICESat, SIT data are available for shorter campaign periods than CryoSat-2 and does not align seamlessly with our selected month (but CyroSat-2 does), so some temporal sampling bias is expected in the model comparison.

3. Sea Ice Extent, Volume, and Thickness

We computed simulated SIE time series using the CMIP6 variables of SIC (siconc) and grid cell area (areacell) by calculating the total area of all grid cells with SIC ≥ 15%. SIE is a prevalent metric used for model comparisons and benefits from the availability of a long-term passive microwave satellite record and the reduction of uncertainties in SIC associated with the pole hole, melt ponds, thin ice, and MIZs. We chose the SIE metric, as opposed to SIA, to show results consistent with the ice edge spatial analyses methods described below (section B.4). The primary shortfall of the SIE metric is that it is strongly grid-dependent (Notz 2014) as compared to SIA, and that both SIA and SIE afford only a limited two-dimensional sea ice evaluation. In determining an ensemble model mean SIE, SIE time series for each ensemble member is first calculated before averaging.

The SIV metric incorporates the vertical dimension, i.e. thickness, and it provides a more complete measure of the state and rapid change of the Arctic sea ice (Kwok and Rothrock 2009; Stroeve et al. 2012). For each model, we computed simulated ensemble mean SIV time series using the CMIP6 variables (e.g., SIV (sivol) or sea ice mass (simass), and areacell provided by the modeling groups) (see Table 6). A CMIP6 MM SIE and SIV for our study subset is determined by averaging the 12 individual model ensemble means. Sea ice anomalies are calculated relative to the 1979-2014 monthly mean for the individual models and observational references.

We also examined model simulated SIT in order to assess its spatial pattern distributions. Monthly ensemble mean SIT was calculated on the model's native grid (where available) using the CMIP6 variables: e.g., sivol or simass multiplied by density of

sea ice (ρsi), per areacell (see Table 6). A SIT cutoff of 6 m, informed by Melling (2002), was applied to all models when determining correlation and RMSE to correct for erroneous values of too thick sea ice simulated in portions of the Canadian Arctic Archipelago (CAA) by some models.

4. Ice Edge Analysis

Two ice edge spatial analysis techniques were evaluated to compare the satellite observed SIE against model simulations of SIE. Firstly, the Spatial Probability Score (SPS) is a probabilistic verification score for contours (Goessling and Jung 2018). For the sea ice edge contour, SPS is defined as

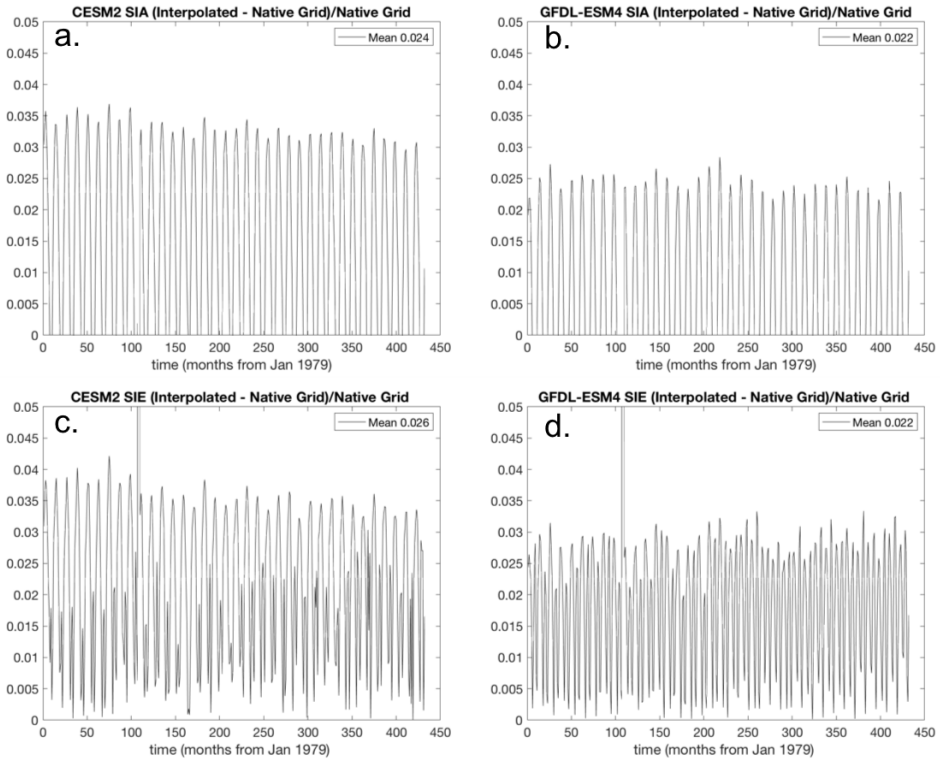
$$SPS = \int_x \int_y \left\{ P[sic > 0.15]_f(x,y) - P[sic > 0.15]_o(x,y) \right\}^2 dx\, dy$$

where $P[sic > 0.15]_f$ is the ensemble probability of having $sic > 0.15$ and $P[sic > 0.15]_o$ is the binary field (1 or 0) representing 'perfect' SIE observations. Simply put, the SPS is a summation of the areal differences between the 'true' ice edge and the probabilistic modeled ice edge. The total SPS is therefore spatial integration of local areas that are both overestimated and underestimated by model SIE. Additionally, the average of the squared difference term in the SPS equation is the local Brier score (Brier 1950), and is used here in spatial maps to evaluate each grid cell skill between zero and one in representing the ice edge (cf. Wayand et al. 2019). A Brier score of zero represents a perfect prediction of SIE, and the score of one represents the alternative extreme.

The other spatial analysis technique we apply is the integrated ice-edge error (IIEE) (Goessling et al. 2016). The IIEE is a special case of SPS whereby $P[sic > 0.15]_f$ is replaced by a binary deterministic value, as is the case for models with single member simulations. While the majority of modeling centers participating in CMIP6 do provide at least a limited number of ensemble members, there is a large range (e.g., 1-32). It is through this lens we evaluate the mean IIEE against the SPS. For IIEE, each individual ensemble member was first treated as a single deterministic simulation then averaged across all

individual model ensembles to determine the individual model mean. Here we calculated CMIP6 and RASM-G monthly mean SPS and IIEE over several Arctic regions.

Sea ice observations and model data used for spatial analyses were linearly regridded, as needed, onto the NSIDC Sea Ice Polar Stereographic (SIPS) North 25 km × 25 km grid to allow for cell by cell comparison between the simulated and the observed values in both time and space between models. The absolute mean errors resulting from the grid interpolation are conservatively estimated at less than 4% for SIE or SIA (up to 3% for 50 km resolution GFDL-ESM4 and up to 3.6% for 100 km resolution CESM2; Figure 23).



Absolute mean error as a result of interpolating (a,b) SIA and (c,d) SIE from nominal native grid (a,c) 100 km (using CESM2) and (b,d) 50 km (using GFDL-ESM4) to the NSIDC Sea Ice Polar Stereographic North 25 km × 25 km grid.

Figure 23. Mean error as the result of interpolating sea ice area and extent from native model grid to NSIDC 25 km × 25 km grid

C. RESULTS

Prior to presenting spatial evaluation results, we first introduce time series analysis of the commonly used SIE and SIV metrics. These results reveal the simulated sea ice spread of our selected CMIP6 models and set the stage for spatial analyses. For both SIE and SIV, we find that CMIP6 MM outperforms any single model in representing the mean state and trend of the historical sea ice cover.

1. Sea Ice Extent and Volume

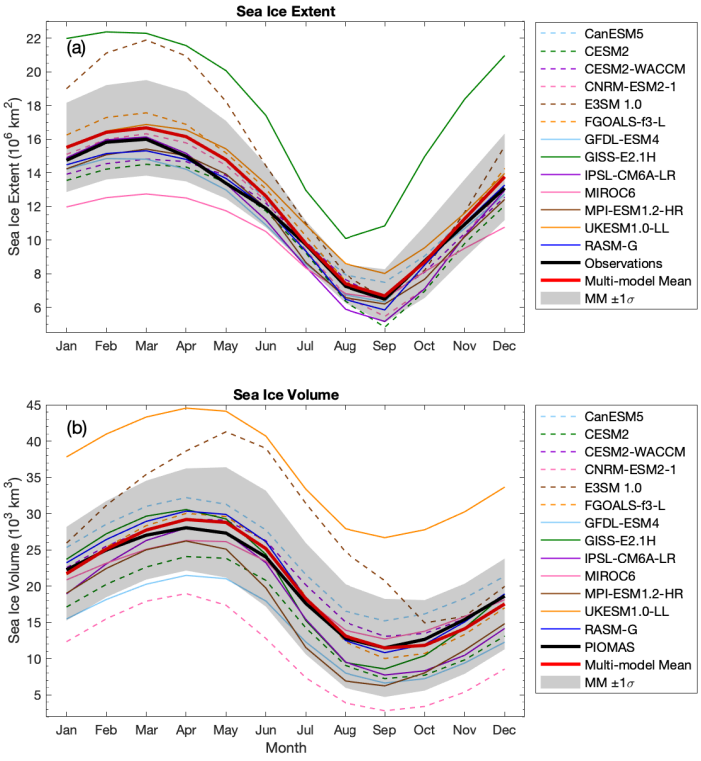
Figure 24 shows the SIE mean annual cycle for CMIP6 and RASM-G models and the combined passive microwave observations, and Table 8 summarizes the SIE 12-month running mean, standard deviation, and linear trends from 1979 to 2014. We can clearly see that GISS-E2.1H is highest biased across all months, while E3SM 1.0 and MIROC6 both exhibit biases in winter months outside one standard deviation of the MM spread (Figure 24a) relative to the observed mean SIE annual cycle. Additionally, all models, except GISS-E2.1H, exhibit a realistic seasonal cycle consisting of SIE maximum in March and minimum in September. Consistent with Shu et al. (2020), the month of March has a slightly larger spread among the CMIP6 models than September (Figure 24a), suggesting larger winter sea ice edge variability across models.

All models simulate SIE negative trends with varying intensity (Table 8), in general agreement with the observed historical SIE decline for the 36-year period (1979-2014). The CMIP6 MM rate of decline, –0.55×106 km2 decade-1, matches closely the observed trend (–0.53×106 km2 decade-1). However, individual CMIP6 model trends are spread relatively wide, with only four of 12 CMIP6 models falling within two standard deviations of the observed trend. The discrepancy here can be mostly explained by model bias, i.e., positive (negative) bias models tend to have stronger (weaker) declining trends (Table 8).

Table 8. Mean and standard deviations (s.d.) of 12-month running mean sea ice extent and anomaly linear trends

Model	Sea Ice Extent			
	Mean ($\times 10^6$ km^2)	s.d. ($\times 10^6$ km^2)	Linear Trend ($\times 10^6$ km^2 decade^{-1})	s.d. ($\times 10^6$ km^2 decade^{-1})
CanESM5	12.62	0.87	-0.84	0.06
CESM2	10.93	0.58	-0.52	0.07
CESM2-WACCM	11.60	0.55	-0.52	0.05
CNRM-ESM2-1	11.85	0.39	-0.36	0.05
E3SM 1.0	14.73	1.14	-1.06	0.11
FGOALS-f3-L	12.88	0.38	-0.36	0.05
GFDL-ESM4	11.32	0.46	-0.36	0.05
GISS-E2.1H	17.83	0.74	-0.70	0.08
IPSL-CM6A-LR	11.38	0.55	-0.53	0.07
MIROC6	10.18	0.34	-0.31	0.04
MPI-ESM1.2HR	11.45	0.46	-0.44	0.06
UKESM1.0-LL	13.07	0.65	-0.61	0.06
CMIP6 MM	12.49	2.06	-0.55	0.06
RASM-G*	11.65	0.47	-0.44	0.04
Observations	11.96	0.58	-0.53	0.04
PIOMAS	-	-	-	-

Mean and standard deviations (s.d.) of 12-month running mean SIE and their anomaly linear trend and s.d. for CMIP6 models, RASM-G, and the combined passive microwave observations for the period 1979 to 2014 (*except RASM-G for 1980 to 2014).



Seasonal cycle of (a) SIE and (b) SIV for CMIP6 models and RASM-G with combined passive microwave observations and PIOMAS for the period 1979 to 2014. The CMIP6 multi-model mean and standard deviation are displayed by a black line and gray shading.

Figure 24. Mean annual cycles of sea ice extent and volume for the period 1979-2014

Figure 24b shows the mean SIV annual cycle for CMIP6 and RASM-G models and PIOMAS reanalysis, while Table 9 summarizes their respective SIV 12-month running means, standard deviations, and linear trends. Compared to SIE, the SIV shows a relatively larger model spread, with the largest bias in UKESM1.0-LL and CNRM-ESM2-1 (Figure 24b). All models, except E3SM 1.0, have a realistic seasonal cycle consisting of the SIV maximum in April and the SIV minimum in September.

As in the case for SIE, all models simulate a declining SIV trend with varying intensity, which qualitatively matches the PIOMAS and satellite estimated (Kwok 2018) SIV trends. The CMIP6 MM SIV mean of 20.36×10^3 km^3 and trend of -3.02×10^3 km^3 decade^{-1} for 1979-2014 pairs very well with the PIOMAS mean of 20.18×10^3 km^3 and trend

of -3.03×10^3 km^3 decade^{-1} (Table 9). Here seven of 12 CMIP6 models fall within two standard deviations of the PIOMAS SIV trend.

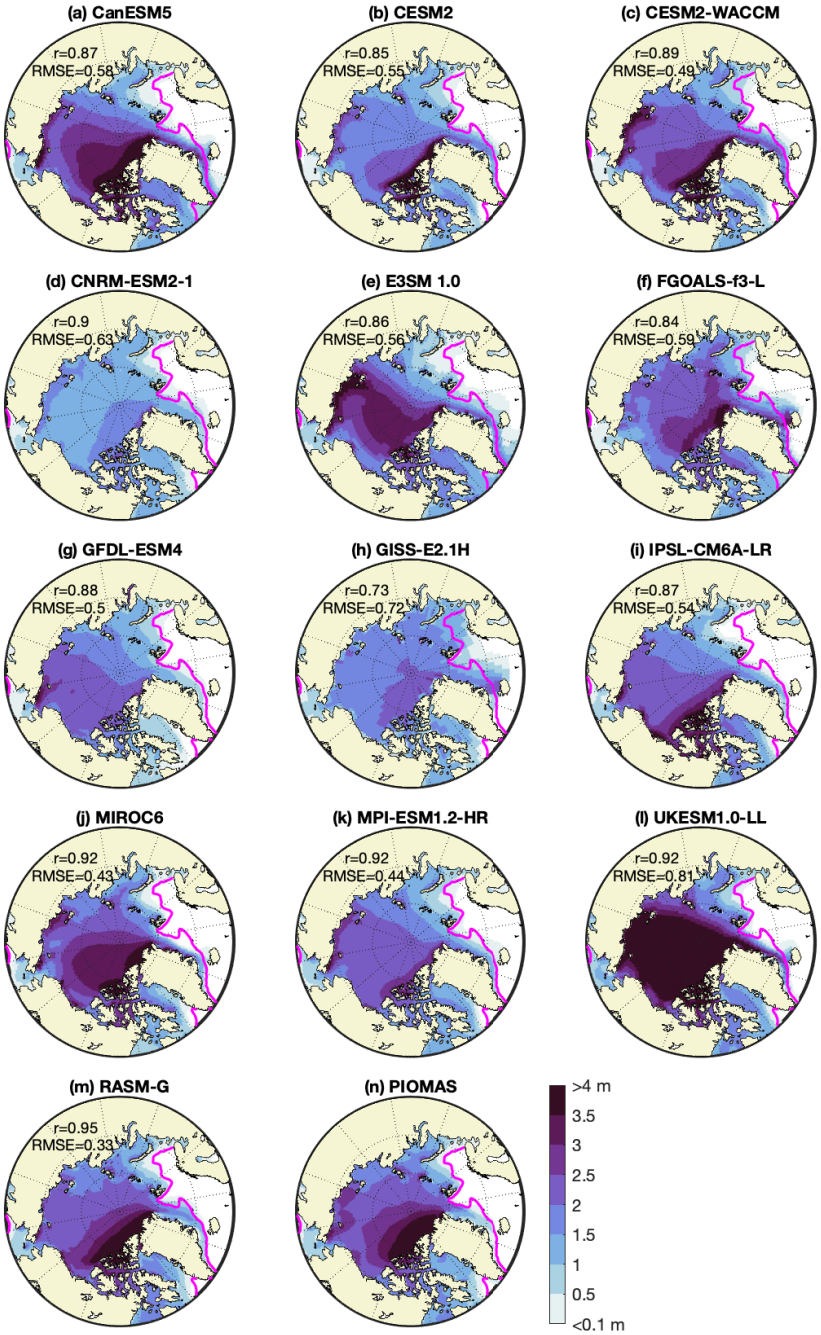
Table 9. Mean and standard deviations (s.d.) of 12-month running mean sea ice volume and anomaly linear trends

Model	Sea Ice Volume			
	Mean ($\times10^3$ km^3)	s.d. ($\times10^3$ km^3)	Linear Trend ($\times10^3$ km^3 decade^{-1})	s.d. ($\times10^3$ km^3 decade^{-1})
CanESM5	23.82	4.15	-3.99	0.62
CESM2	15.83	2.96	-2.78	0.48
CESM2-WACCM	21.37	5.01	-4.82	0.37
CNRM-ESM2-1	10.54	0.79	-0.71	0.10
E3SM 1.0	28.30	6.04	-5.66	0.55
FGOALS-f3-L	20.25	2.20	-2.03	0.37
GFDL-ESM4	14.19	2.09	-1.86	0.34
GISS-E2.1H	20.14	2.26	-2.17	0.12
IPSL-CM6A-LR	17.72	2.84	-2.75	0.21
MIROC6	19.80	2.48	-2.30	1.68
MPI-ESM1.2HR	16.38	2.29	-2.17	0.20
UKESM1.0-LL	35.99	5.32	-5.02	2.19
CMIP6 MM	20.36	3.16	-3.02	0.60
RASM-G*	21.03	2.64	-2.52	0.36
Observations	-	-	-	-
PIOMAS	20.18	3.31	-3.03	0.51

Mean and standard deviations (s.d.) of 12-month running mean SIV and their anomaly linear trend and s.d. for CMIP6 models, RASM-G, and PIOMAS for the period 1979 to 2014 (*except RASM-G for 1980 to 2014).

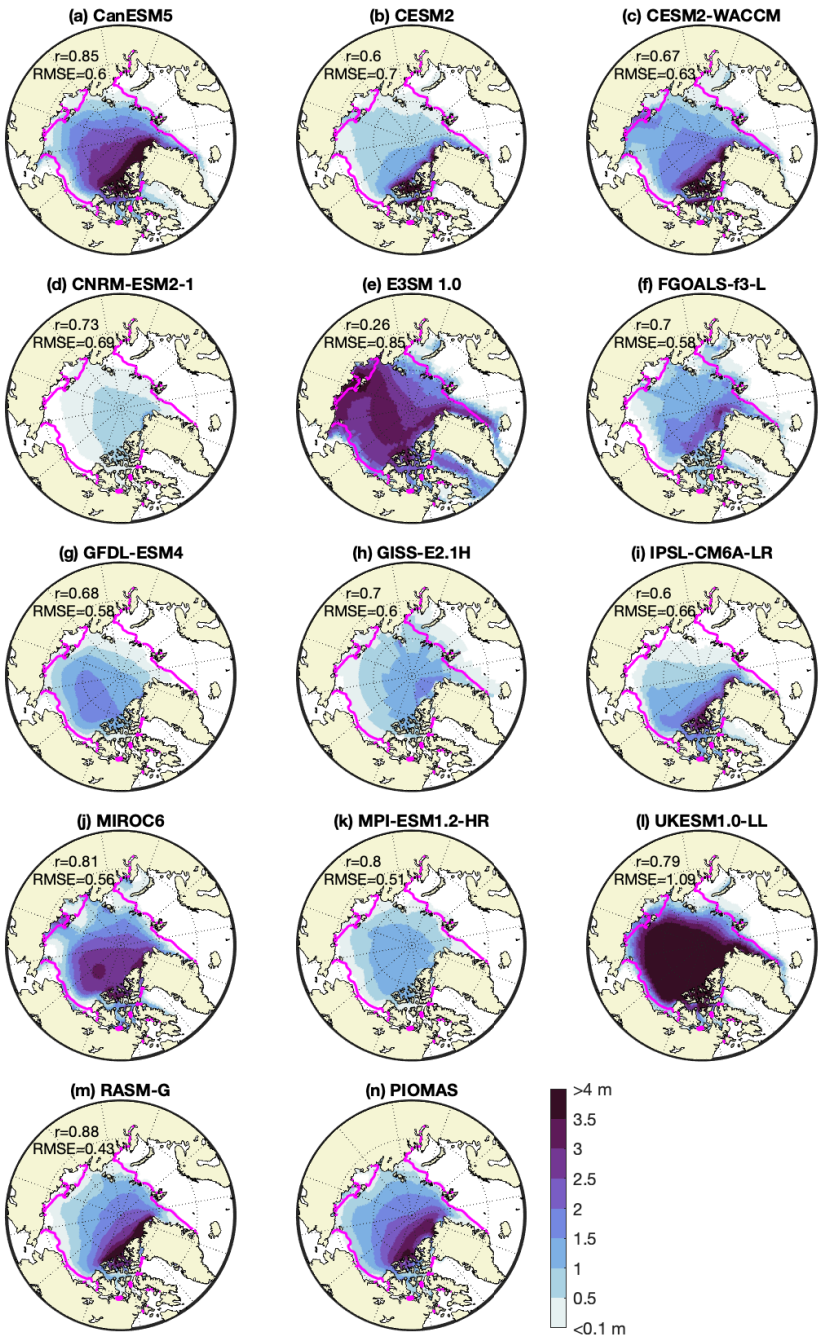
2. Sea Ice Thickness

An accurate spatial distribution of SIT is key to estimates of SIV and it reflects the skill in simulation of local processes, coupled interactions and energy transfer between the ocean below, the sea ice and the atmosphere above (Stroeve et al. 2014). We first assess whether the CMIP6 models accurately simulate the spatial distribution of SIT by focusing on the months of mean SIE maximum (March, Figure 25) and minimum (September, Figure 26) for the period 1979-2014. The mean satellite observed ice edge, determined from the gridded NSIDC monthly SIE, is included on each SIT image. However, because of the satellite limitations in differentiating thin ice (at least up to 0.2 m) from open water (personal communication, W. Meier, NSIDC), we impose the limit SIT > 0.1 m in order to provide a conservative and comparable estimate of the simulated ice edge.



March mean SIT distribution for (a-l) CMIP6 models (1979-2014), (m) the RASM-G simulation (1980-2014), and (n) PIOMAS (1979-2014). Magenta contours indicate the averaged March NSIDC sea ice edge for the same period. Spatial pattern correlation coefficients (r) and root-mean-square error (RMSE) for individual models against PIOMAS reanalysis are included in the upper-left corner of each panel.

Figure 25. March mean sea ice thickness distribution for the period 1979-2014



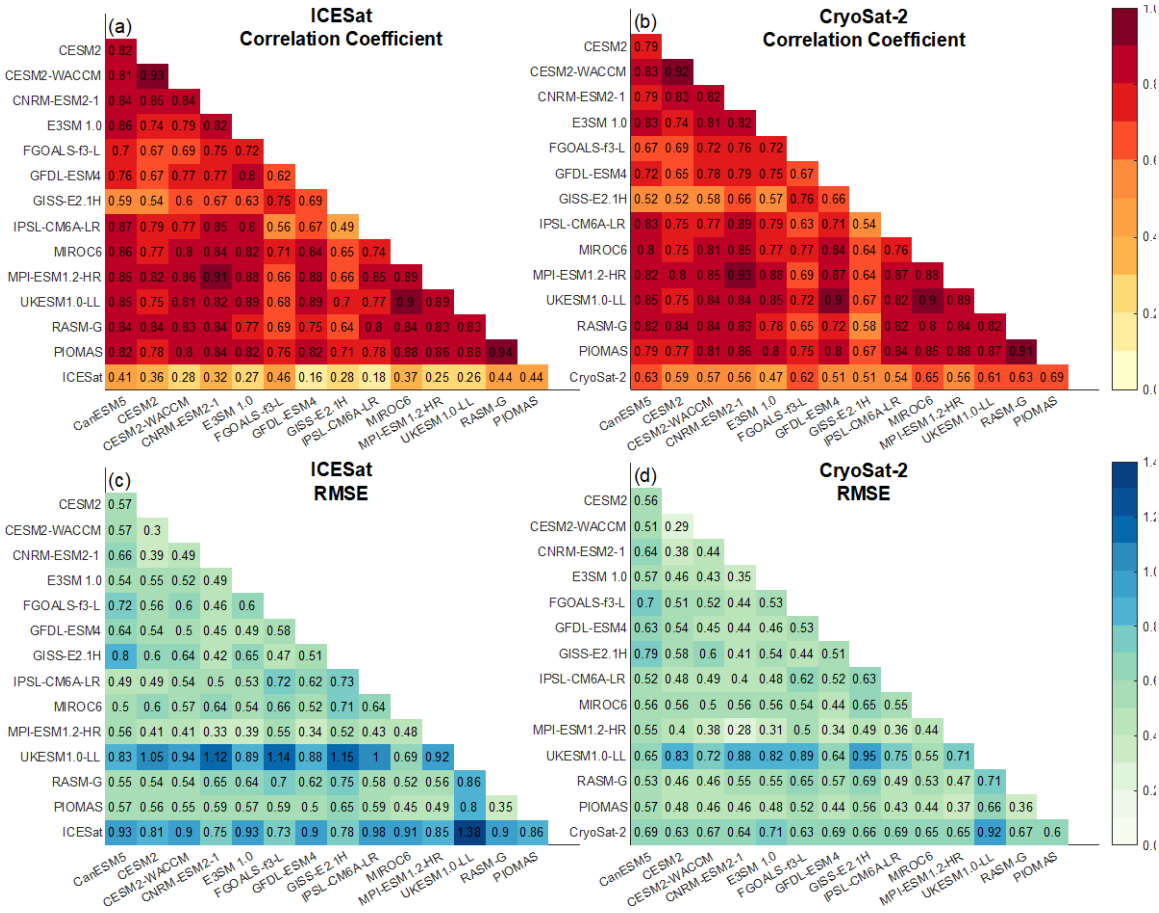
Same as Figure 6 but for September mean SIT and magenta contours indicate averaged September NSIDC sea ice edge.

Figure 26. September mean sea ice thickness distribution for the period 1979-2014

The PIOMAS simulated monthly mean ice thickness (Figure 25n and Figure 26n) is often used as an historical observational 'proxy' (e.g. Labe et al. 2018) with the following caveats: PIOMAS has a general tendency to underestimate SIT in regions of thick (> 3.5 m) ice (e.g., near the CAA and north of Greenland), and overestimate ice thickness in regions of thin ice. Additionally, PIOMAS has a tongue of ~2.5 m ice that extends across the Arctic to the Chukchi and East Siberian seas in March that is not depicted in *in situ* observations (Schweiger et al. 2011a; Stroeve et al. 2014). Here we include a spatial pattern correlation coefficient (r, at 99% confidence interval) and RMSE against PIOMAS on each figure. However, as was the case for CMIP5 (Stroeve et al. 2014), nearly all CMIP6 models show high correlation (r>0.84) with PIOMAS for March SIT (except GISS-E2.1H, r=0.73; Figure 25) which makes identification of poor performers less effective, despite some models clearly underrepresenting general SIT patterns described below.

Six of 12 CMIP6 models (CanESM5, CESM2, CESM2-WACCM, FGOALS-f3-L, IPSL-CM6A-LR, and MIROC6) demonstrate a reasonable pattern of SIT relative to the PIOMAS SIT distribution reference. This means that generally, they correctly locate the thickest (at least 3.5 m) Arctic sea ice along the CAA and north of Greenland, as well as thinner March sea ice located along the Eurasian shelf (Figure 25). For September, the same six models also maintain an appreciable amount of thick ice (at least 3.0 m) along the CAA and northern Greenland. In contrast, four of the 12 CMIP6 models (CNRM-ESM2-1, GFDL-ESM4, GISS-E2.1H, and MPI-ESM1.2-HR) fail to maintain an appreciable amount of sea ice which is greater than 2.0 m at the September sea ice minimum (Figure 26). On the other hand, UKESM1.0-LL (Figure 25l and Figure 26l) is laden with sea ice greater than 3.5 m throughout the analysis period and it covers a much larger area.

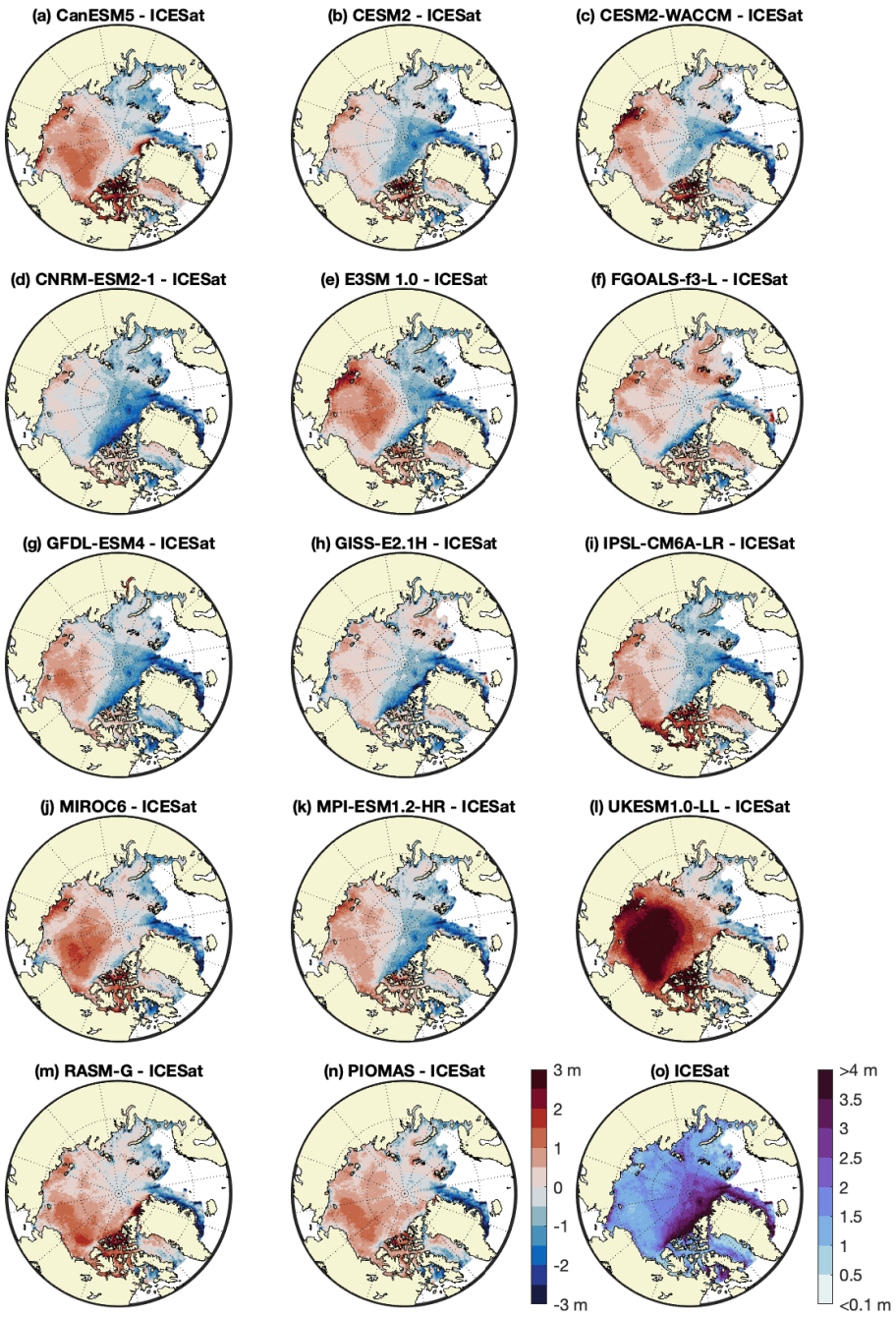
Turning towards the period of satellite-derived SIT observations and following Stroeve et al. (2014), we show spatial pattern correlations and RMSE between models, ICESat, and CyroSat-2 in Figure 27, and difference plots in Figure 28 and Figure 29. For the period 2003-2008, models show low correlations against ICESat (all of which are significant at the 99% confidence interval) ranging from r=0.18 to 0.46 and RMSE from 0.73 m to 1.38 m (Figure 27a and Figure 27c).



Spatial pattern correlations (a and c) and RMSE (b and d) of March SIT between CMIP6 models, PIOMAS and RASM-G with ICESat for the period of 2003-2008 and CryoSat-2 for the period of 2011-2014, respectively (following Stroeve et al. [2014]).

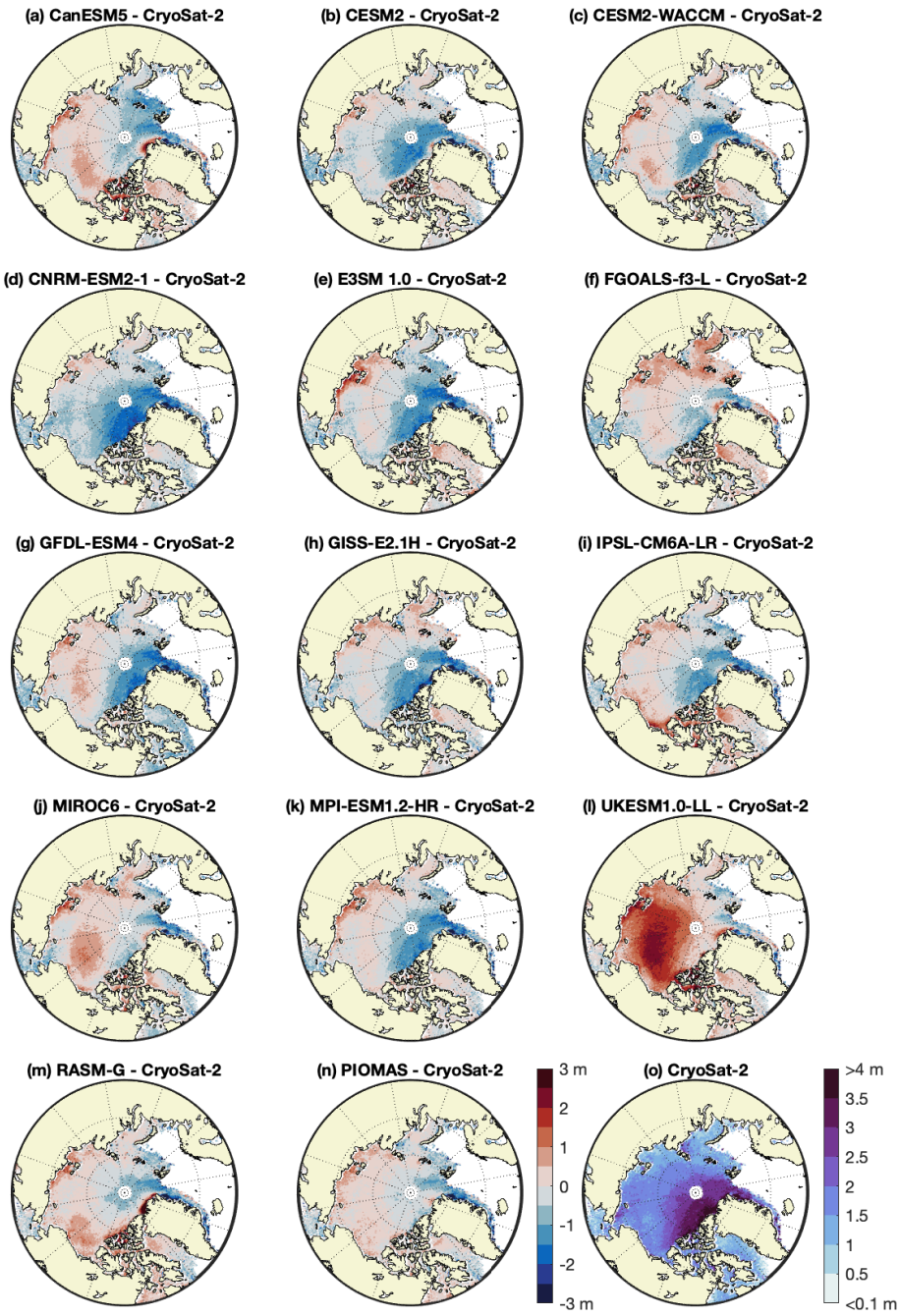
Figure 27. Spatial pattern correlations and root-mean-square error of March sea ice thickness between models and ICEsat and CryoSat-2

Overall, models tend to underestimate SIT along the CAA and north and east of Greenland, and overestimate SIT over the Beaufort and Chukchi seas and along the Siberian coast (Figure 28). Nine of 12 CMIP6 models exhibit overestimation of SIT throughout the western Arctic (Figure 28), which may be in part the result of the timing of ICESat collected observations (i.e., late February through early March) and the historically low sea ice cover during the collection years (Stroeve et al. 2014). In comparison against CryoSat-2 for the later period 2011-2014, all models show larger correlation ranging from r=0.47 to 0.65 and smaller RMSE from 0.63 m to 0.92 m than against ICESat (Figure 27b and Figure 27d). Regional SIT differences between models and CryoSat-2 are consistent with those against ICESat but are less pronounced (Figure 29).



March SIT differences for the period 2003-2008 between (a-l) CMIP6 models, (m) RASM-G, (n) PIOMAS and ICESat. ICESat data are shown in (o).

Figure 28. March sea ice thickness differences for the period 2003-2008 between models and ICESsat



March SIT differences for the period 2011-2014 between (a-l) CMIP6 models, (m) RASM-G, (n) PIOMAS and CryoSat-2. CryoSat-2 data are shown in (o).

Figure 29. March sea ice thickness differences for the period 2011-2014 between models and CryoSat-2.

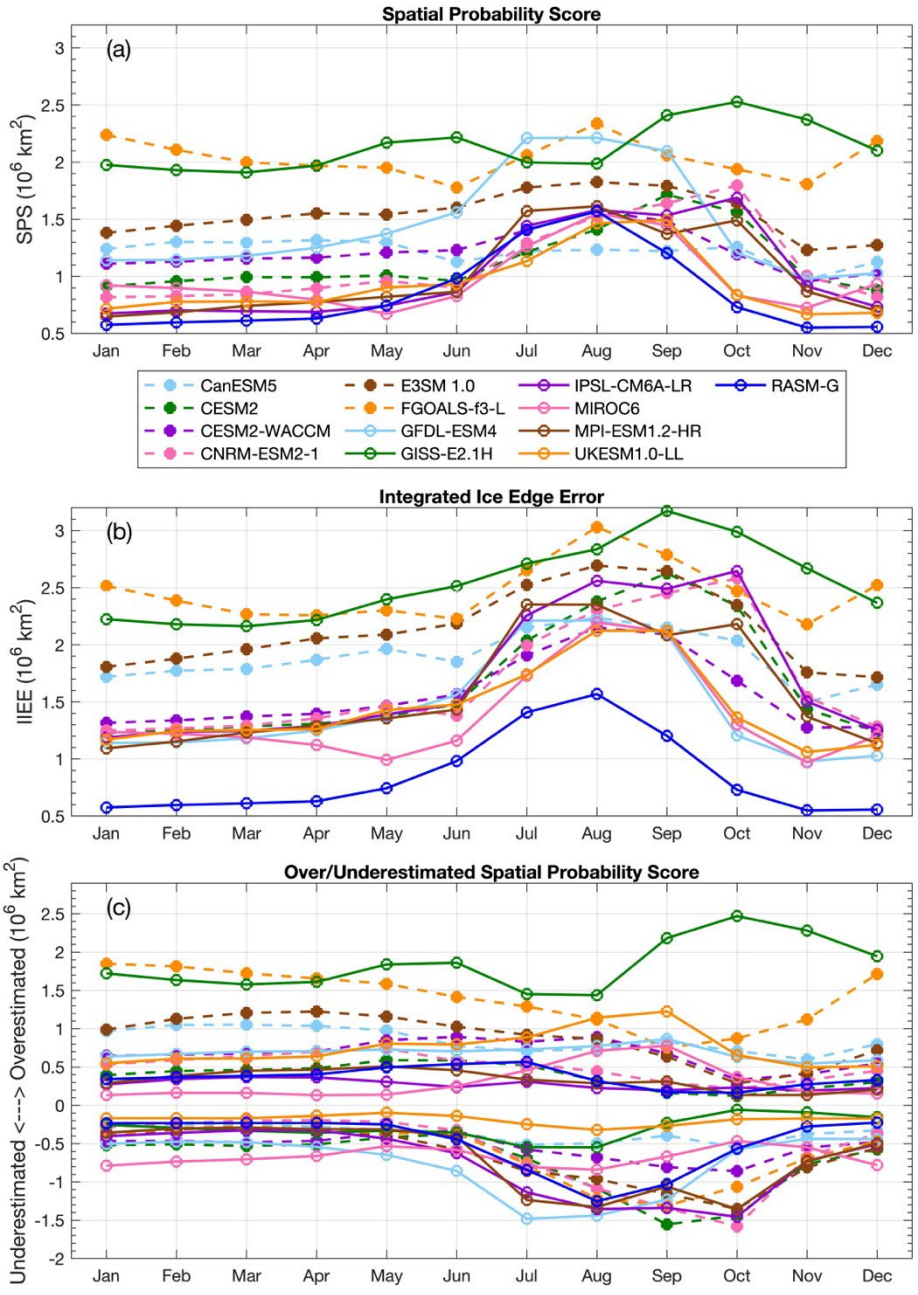
The RASM-G simulated SIT agrees well with PIOMAS (Figure 25m and Figure 26m), and ICESat and CryoSat-2 (Figure 28 and Figure 29, respectively). In comparison

with PIOMAS, RASM-G mean SIT along the Eurasian shelf is thinner for both March and September. Additionally, RASM-G maintains a larger area of thick (>3.5 m) September sea ice along and within the CAA and north of Greenland. Evaluated against ICESat and CryoSat-2 (Figure 27), RASM-G spatial pattern (correlation and RMSE) is highly correlated with PIOMAS reanalysis and is similar to the better performing CMIP6 models. Overall, we find that the RASM-G simulated SIT distribution represents a comparable skill to PIOMAS.

3. Sea Ice Edge

The final metrics we present for CMIP6 sea ice analysis are the SPS and the IIEE. For our subset of CMIP6 models, all but GFDL-ESM4 includes multiple ensemble members, and therefore we present the majority of our main text findings and figures from the SPS analysis. However, we show that IIEE is also an appropriate technique to compare the simulated ice edge across a spectrum of ensemble model classes (i.e., single through small to large ensemble models). Additional IIEE results are shown in the supporting material.

First, we analyze the SPS for the whole pan-Arctic (Figure 30), defined here as all areas within the NSIDC SIPS North 25 km × 25 km grid with the exception of Hudson Bay, and for the 11 sub-regions (Figure 31 and Figure 32). Second, we apply the SPS analysis and examine the local Brier score in order to identify regions where CMIP6 models commonly have difficulty in reproducing sea ice coverage skillfully (Figure 33 and Figure 34). We restrict our SPS analysis to the long-term mean errors, avoiding interannual comparisons, of which no CMIP6 models from our subset show predictive SIE skill (not shown).



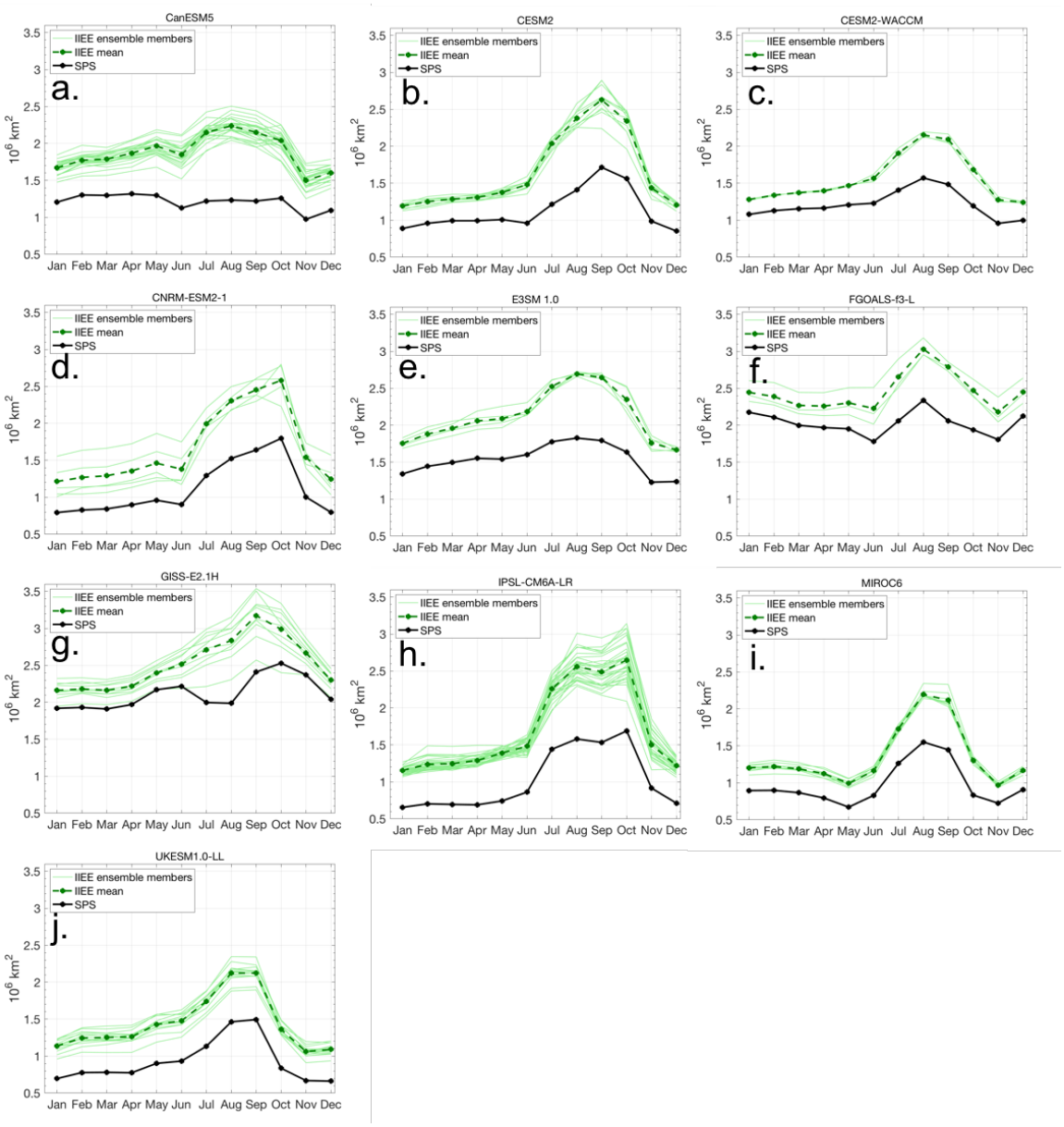
Pan-Arctic monthly mean (a) SPS and (b) IIEE for the period 1980-2014 for CMIP6 models and RASM-G. (c) Pan-Arctic monthly mean SPS subdivided into two components: overestimation (positive) and underestimation (negative).

Figure 30. Pan-Arctic monthly mean sea ice edge analysis metrics for the period 1980-2014

The resulting pan-Arctic SPS exhibits a strong mean annual cycle (Figure 30a). Over the 35-year record (1980-2014), most CMIP6 models have a relatively steady SPS

during the colder months of December through April, with relatively better performing models showing SPS ranges between 0.7×10^6 km^2 to 1.0×10^6 km^2 (CESM2, CNRM-ESM2-1, IPSL-CM6A-LR, MIROC6, MPI-ESM1.2-HR, and UKESM1.0-LL). All models, except CanESM5, experience increased SPS during the warmer summer/autumn months (July through October) ranging from 0.8×10^6 km^2 to 2.5×10^6 km^2. FGOALS-f3-L and GISS-E2.1H have less dramatic seasonal cycles due to their colder season SPS remaining high, when their minimum ranges from 1.9×10^6 km^2 to 2.3×10^6 km^2. The timing of individual model SPS peak ranges from July through October, with the majority of CMIP6 models peaking in August (five models) or October (three models). GISS-E2.1H has the largest single month SPS, near 2.5×10^6 km^2 in October. The lowest single month SPS values, near 0.7×10^6 km^2, occur in May (MIROC6) and November (UKESM1.0-LL).

Compared against the SPS, the pan-Arctic mean IIEE magnitude is always larger, by annual mean factor ranging between 1.2 (GISS-E2.1H) to 1.7 (IPSL-CM6A-LR), for all models with multiple ensemble members (Figure 30b). Additionally, the shape of a model's mean annual cycle and relative performance against other models remains largely consistent between the chosen metrics (except the shape of GISS-E2-1H). Note that the single ensemble member GFDL-ESM4 SPS and IIEE are the same.



Pan-Arctic monthly mean IIEE for individual ensemble members, IIEE ensemble mean, and SPS for the period 1980-2014. Note that GFDL-ESM4 and RASM-G have single ensemble members are therefore not included.

Figure 31. Ensemble member spread in integrated ice-edge error and comparison against Spatial Probability Score for individual CMIP6 models

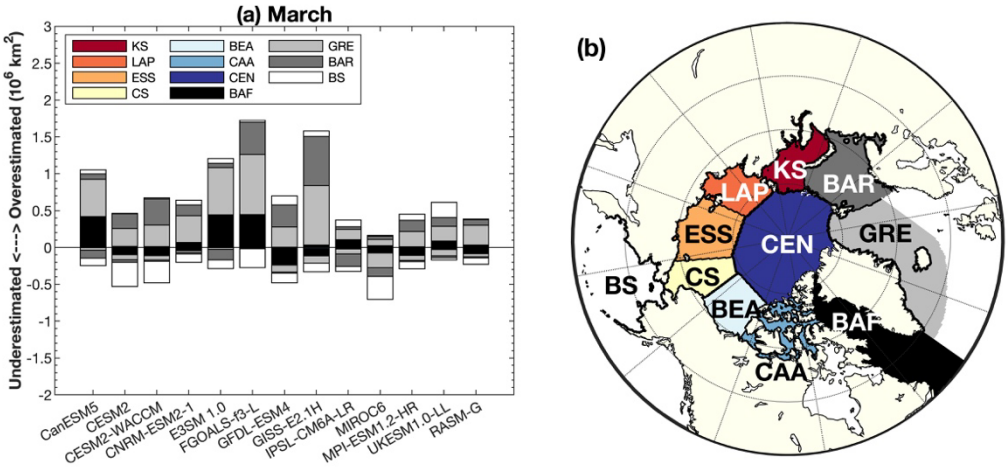
In Figure 31, we show the individual model ensemble member's IIEE alongside their mean IIEE and SPS. This detailed look illustrates the IIEE spread about the model's mean across the whole seasonal cycle, which is not evident from the SPS alone. The

individual model spread between ensemble realizations can be interpreted as the model internal variability in simulating the sea ice edge.

In Figure 30c, the two separate components of pan-Arctic SPS, overestimation (SPS-O) and underestimation (SPS-U), are further examined. For December through April, SPS-O has a greater inter-model spread than SPS-U. Neither component of SPS has much change from month to month during this period. However, during the spring melt through autumn freeze up, individual models do show considerable temporal variability in SPS (i.e., variability in the position of the sea ice edge). Magnitudes of SPS-O in five of 12 CMIP6 models (CanESM5, CNRM-ESM2-1, GFDL-ESM4, IPSL-CM6A-LR, and MPI-ESM1.2-HR) exhibit relatively small ($<0.4\times10^6$ km^2) changes from June through November (Figure 30c). However, MIROC6, UKESM1.0-LL, and GISS-E2.1H have increased values of SPS-O during the same period, suggesting that the simulated sea ice is not melting in the right regions and/or as fast as observed. On the other hand, E3SM 1.0, FGOALS-f3-L, and CESM2-WACCM models have decreased values of SPS-O. In contrast, most CMIP6 models (except UKESM1.0-LL and CanESM5) have distinct seasonal surges of SPS-U for June through November. Four of 12 CMIP6 models have peak SPS-U in July/August and six models peak in September/October. Models with larger SPS-U early in the melt season (July/August) suggests that sea ice is removed by melting or advection faster than observed. Models with peak underestimation occurring later into the freeze up season (September/October) suggests that sea ice is not growing quick enough in their simulations.

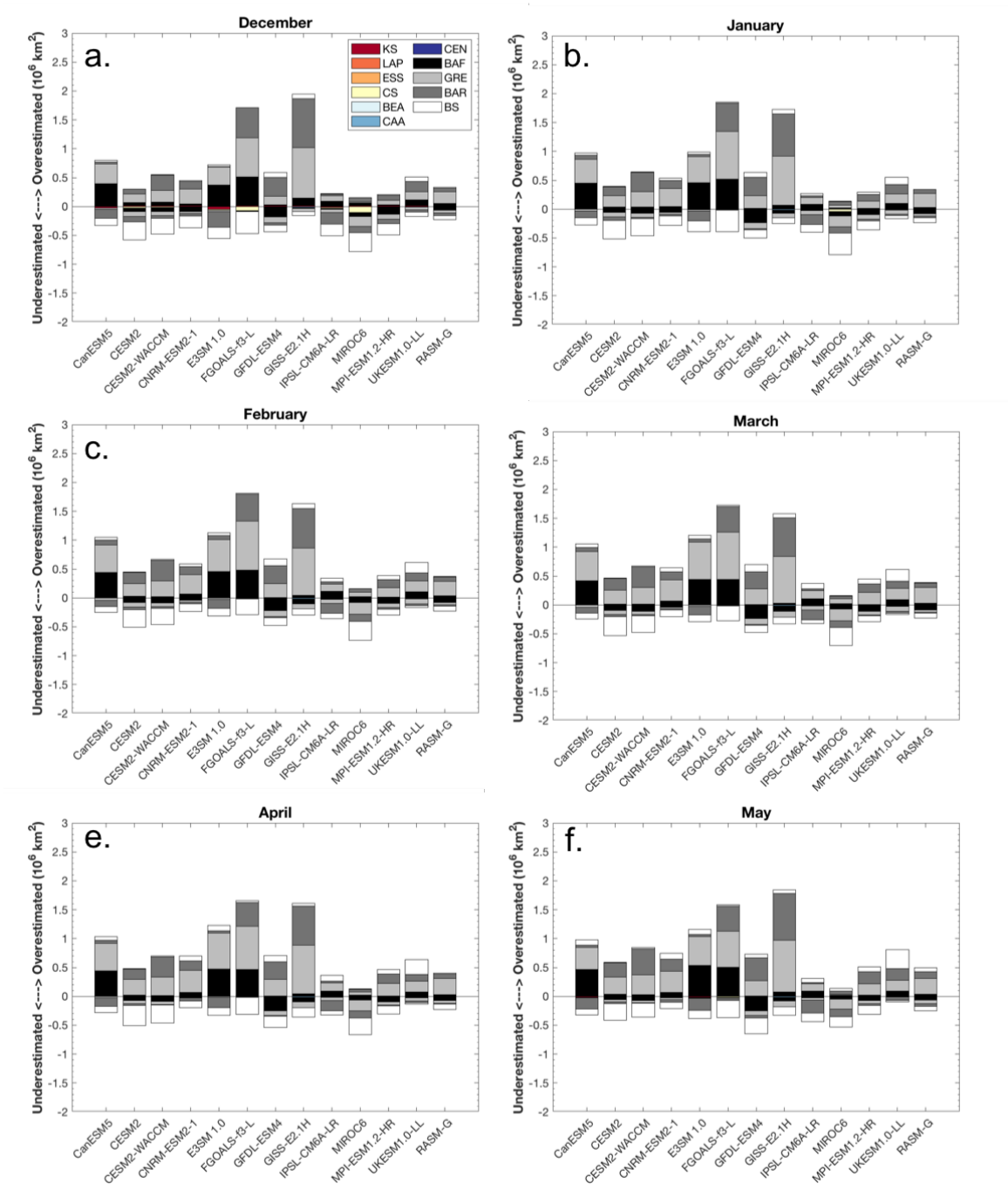
Figure 32a shows the 35-year mean monthly SPS-O/U for individual models and the relative contribution per Arctic sub-region (defined in Figure 32b) for March. During both winter and spring (December through May; Figure 33), the areas which contribute to SPS are predominantly limited to the sub-Arctic seas (i.e., Bering Sea, Nordic Seas, and Baffin Bay), because the ice-covered Arctic interior is all frozen and well represented by all CMIP6 models during this period. The majority of the SPS is composed of overestimations across the Greenland and Barents seas and underestimations for the Bering Sea. We use stacked SPS bar charts in Figures 13-15 to identify the common regions of large SPS-O/U, as well as diagnose differences between the seasonal evolution of individual model SPS. For example, CanESM5 and GISS-E2.1H have the same region of

primary overestimation over the Greenland Sea, but have different secondary error regions; for CanESM5 it is the Baffin Bay, whereas for GISS-E2.1H it is the Barents Sea (Figure 32a).



(a) March monthly mean SPS overestimation and underestimation for CMIP6 models and RASM-G and the relative contribution of the Arctic sub-regions. (b) Arctic regional mask defined in this study: KS = Kara Sea; LAP = Laptev Sea; ESS = East Siberian Sea; CS = Chukchi Sea; BEA = Beaufort Sea; CAA = Canadian Arctic Archipelago; CEN = Central Arctic; BAF = Baffin Bay; GRE = Greenland Sea; BAR = Barents Sea; BS = Bering Sea). (adapted from Meier et al. (2007))

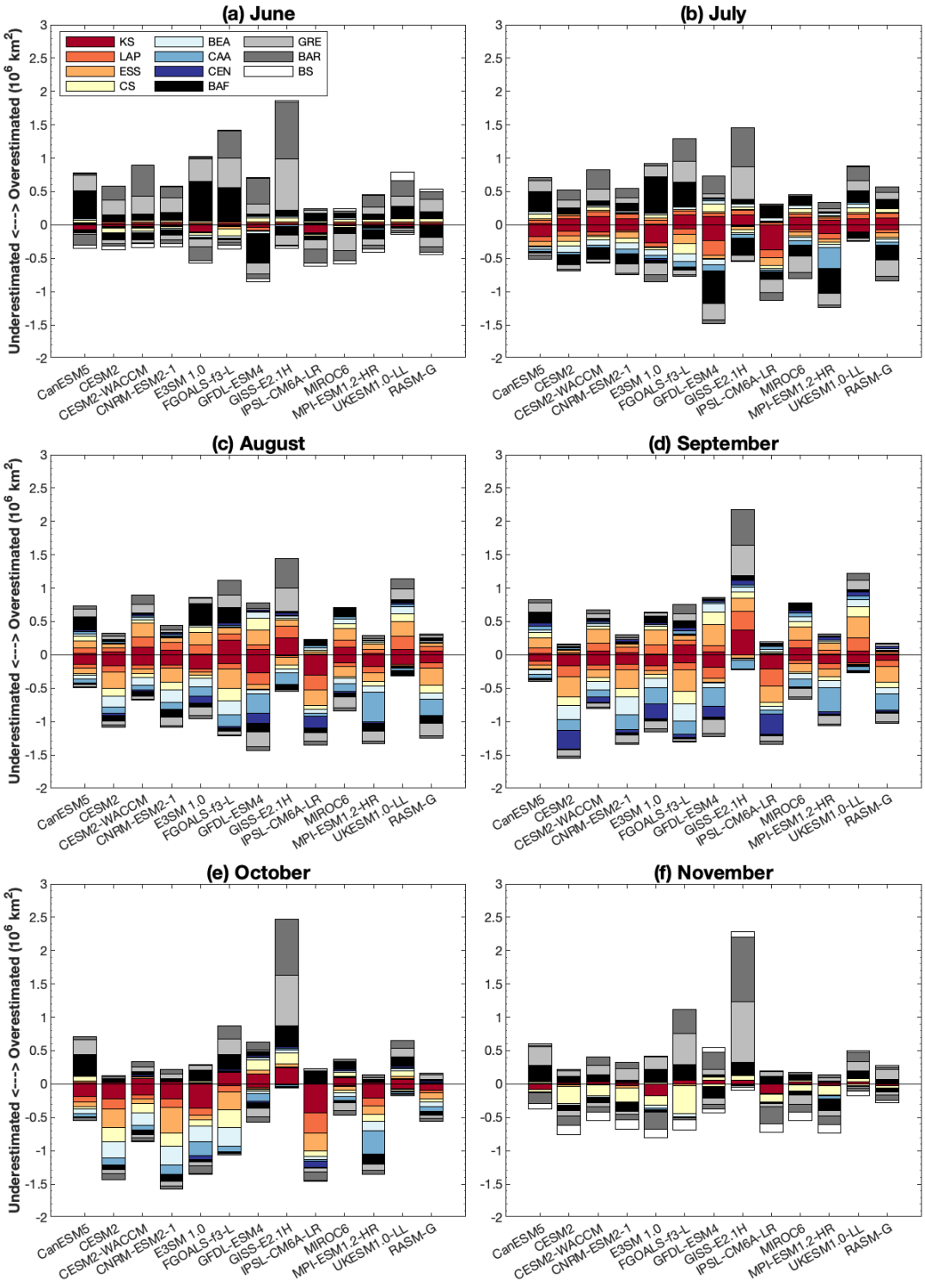
Figure 32. Regional monthly mean Spatial Probability Score overestimation and underestimation for the period 1980-2014



Regional monthly mean SPS for the period 1980-2014 from CMIP6 models and RASM-G subdivided between Arctic regions from December to May (a-f).

Figure 33. Regional monthly mean Spatial Probability Score for the period 1980-2014 from December to May

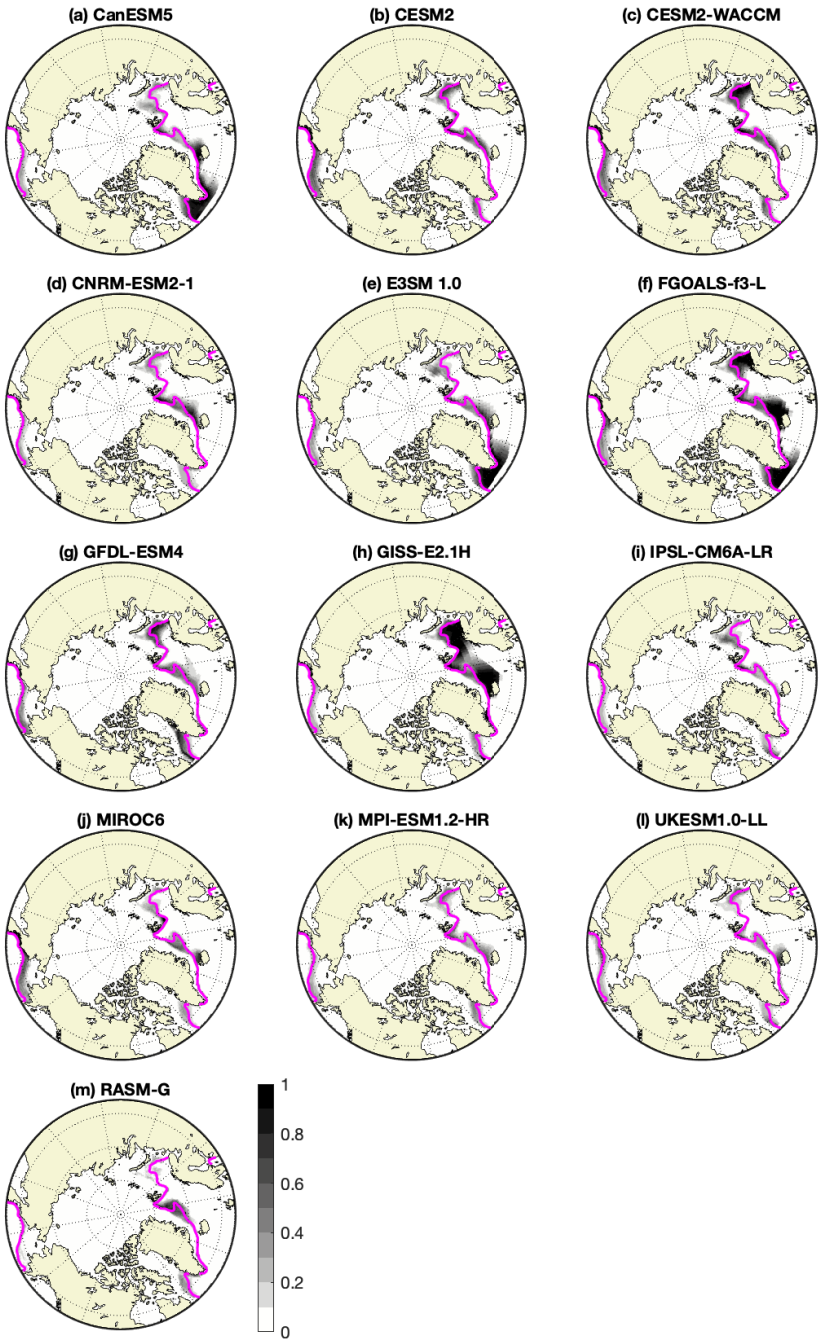
Figure 35 displays the mean local Brier score for March and clearly depicts individual model regions contributing to the total SPS. However, because the Brier score is the squared difference, it does not distinguish between overestimation or underestimation of the ice edge. GISS-E2.1H, E3SM 1.0, FGOALS-f3-L, and CanESM5 have particularly high winter SPS values, which can be attributed to a substantial overestimation of the ice edge position in the Greenland Sea (Figure 32a and Figure 35). Additionally, GISS-E2.1H and FGOALS-f3-L notably overestimate the Barents Sea ice edge, and E3SM 1.0, FGOALS-f3-L, and CanESM5 overestimate the position of the Baffin Bay ice edge. On the other hand, MIROC6 is identified as having the largest SPS-U during this period, which can be attributed to a large underestimation of ice edge in both the Greenland and Bering seas (Figure 32a and Figure 35).



Like Figure 14 but for the months of June to November (a-f).

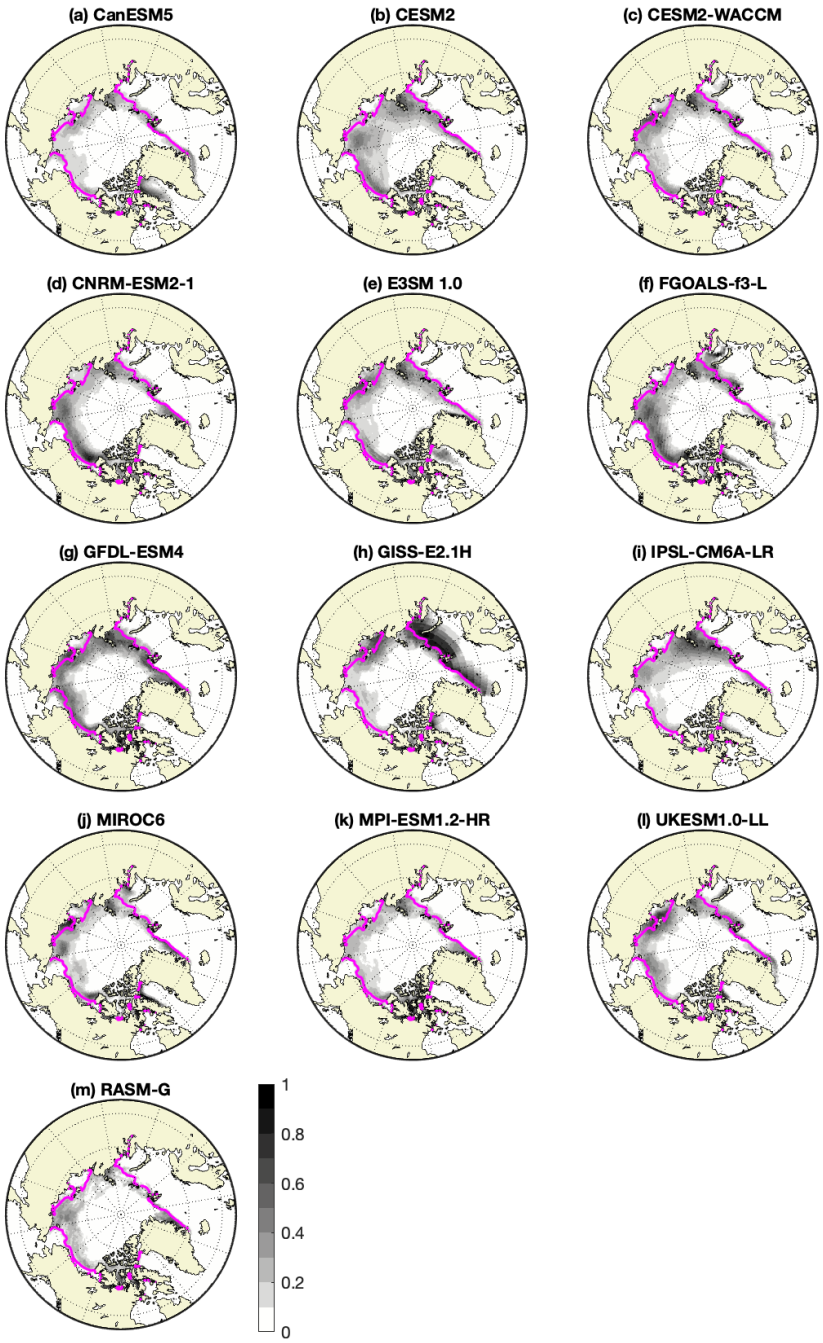
Figure 34. Regional monthly mean Spatial Probability Score for the period 1980-2014 from June to November

Next, we examine the remaining, mostly warmer Arctic months of June through November (Figure 34). Three principal regions can be identified as the largest contributors of SPS for each month. As sea ice retreats poleward, the number of regions which contribute to SPS increases and evolves in time. In June (Figure 34a), the Greenland and Barents seas and Baffin Bay are the regions of primary SPS (predominantly overestimation). This suggests that as the observed sea ice is retreating, the CMIP6 models tend to melt sea ice too slowly. In July (Figure 34b), Barents Sea overestimation decreases (but still remains a substantial error source) and the underestimation of Baffin Bay and the Kara Sea is increasing. This suggests that the simulated sea ice over the Barents Sea is catching up (i.e., accelerating modeled sea ice retreat, from a slow start of sea ice melt) and that sea ice over Baffin Bay and the Kara Sea reduces faster than the observed retreat. August is dominated by SPS-U across the East Siberian and Kara seas, and CAA (Figure 34c), implying that simulated sea ice along the Arctic periphery generally melts quicker than observed. In September (Figure 34d and Figure 36), the CAA and central Arctic Ocean ice edges are generally underestimated by CMIP6 models. On the other hand, the East Siberian Sea is nearly split among ice edge overestimation and underestimation. October ice edge errors are dominated by underestimation over the East Siberian, Kara, and Beaufort seas (Figure 34e), suggesting that sea ice growth is slower in a number of model simulations relative to observations. November SPS (Figure 34f) is dominated by overestimation of the Greenland and Barents seas, and underestimation of the Chukchi Sea. This suggests the CMIP6 models are over efficient in sea ice growth over the Greenland and Barents seas and have slow growth over the Chukchi Sea.



March mean local Brier score for (a-l) CMIP6 models and (m) RASM-G for the period 1980-2014. Magenta contours indicate the averaged March NSIDC sea ice edge for the same period.

Figure 35. March mean local Brier score for the period 1980-2014



Same as Figure 16 but for September and the magenta contours indicate the averaged September NSIDC sea ice edge.

Figure 36. September mean local Brier score for the period 1980-2014

Details of model specific regional SPS evaluation are summarized in Table 10, which shows the regional SPS-O/U values accumulated over the months of greatest SPS

spatial variability within the Arctic, June through November. The regions of the Barents Sea and Baffin Bay account for nine of 12 CMIP6 models largest SPS-O during this 6-month period, while the Greenland Sea accounts for the other three models (CESM2, CNRM-ESM2-1, and MIROC6). The regions with the largest SPS-U are the Kara Sea (CanESM5, CESM2-WACCM, E3SM 1.0, IPSL-CM6A-LR, and UKESM1.0-LL), the East Siberian Sea (CESM2 and CNRM-ESM2-1), and Baffin Bay (GFDL-ESM4 and GISS-E2.1H). Restricting analysis in Table 10 to the extended-interior Arctic (excluding the Barents and Greenland seas and Baffin Bay), we find that the Kara and East Siberian seas are the regions with the largest SPS, either for overestimation or underestimation, for most of the CMIP6 models. Of note, four of 12 CMIP6 models (CESM2, E3SM 1.0, MIROC6, and UKESM1.0-LL) have an inverse relationship between regions with largest SPS-O and largest SPS-U in the Kara and East Siberian seas (i.e., the Kara (East Siberian) Sea is largest overestimated and the East Siberian (Kara) Sea is largest underestimated). The other regions of peak SPS are the Chukchi Sea (CanESM5 and GFDL-ESM4) and the Beaufort Sea (IPSL-CM6A-LR) for SPS-O, and CAA (GFDL-ESM4 and MPI-ESM1.2-HR) for peak SPS-U.

The RASM-G sea ice simulation shares the strong SPS and IIEE seasonality observed in the majority of CMIP6 models and has a peak in August (Figure 30). However, compared to those models its IIEE is notably lower, by 0.5-1.0×10^6 km^2, throughout the year (Figure 30b). The RASM-G simulation has a relatively steady SPS-O of 0.3-0.4×10^6 km^2 throughout the year, and a slightly smaller baseline SPS-U from December through April (Figure 30c). The region which contributes the greatest ice edge error for the RASM-G simulation from June through November is the Greenland Sea (Table 10). Limiting our scope to the extended-interior Arctic (defined above), we find the regions with greatest errors in RASM-G are SPS-O of 0.22×10^6 km^2 across the Kara Sea, and SPS-U of 0.68×10^6 km^2 across the East Siberian Sea (Table 10).

Table 10. Accumulated 35-year (1980-2014) monthly mean SPS ($\times 10^6$ km^2, SPS) for the months of the greatest SPS spatial variability within the Arctic, June through November.

Region		CanESM5	CESM2	CESM2-WACCM	CNRM-ESM2-1	E3SM 1.0	FGOALS-f3-L	GFDL-ESM4	GISS-E2.1H	IPSL-CM6A-LR	MIROC6	MPI-ESM1.2HR	UKESM 1.0-LL	RASM-G
KS	O	0.11	0.22	0.45	0.28	0.05	0.78	0.47	*1.10	0.02	0.48	0.17	0.35	0.22
	U	***0.75***	0.69	***0.61***	0.62	***1.30***	0.34	0.78	0.03	****1.57***	0.32	0.67	***0.46***	0.40
LS	O	0.22	0.15	0.35	0.30	0.37	0.24	0.25	*0.55	0.07	0.31	0.17	0.50	0.20
	U	0.28	0.47	0.24	0.29	0.27	0.37	0.62	0.03	*0.90	0.15	0.41	0.07	0.25
ESS	O	0.31	0.12	0.48	0.15	0.46	0.19	*0.61	0.37	0.08	0.41	0.23	0.61	0.07
	U	0.27	***0.94***	0.30	***1.00***	0.20	*1.10	0.19	0.21	0.90	0.40	0.40	0.04	0.68
CS	O	0.36	0.04	0.18	0.06	0.26	0.03	*0.75	0.43	0.11	0.33	0.11	0.48	0.16
	U	0.15	0.87	0.50	0.76	0.39	****1.30***	0.12	0.14	0.38	0.22	0.42	0.09	0.31
BEA	O	0.14	0.05	0.09	0.05	0.08	0.08	0.32	0.21	0.14	0.21	0.10	*0.34	0.07
	U	0.24	0.76	0.56	0.86	0.66	*0.95	0.28	0.10	0.21	0.28	0.41	0.04	0.30
CAA	O	0.11	0.06	0.09	0.04	0.04	0.09	0.07	0.10	0.11	0.12	0.01	*0.13	0.03
	U	0.21	0.39	0.30	0.57	0.68	0.55	0.80	0.41	0.14	0.37	****1.55***	0.13	0.63
CEN	O	0.09	0.06	0.08	0.10	0.05	0.12	0.08	*0.14	0.02	0.08	0.09	0.12	0.08
	U	0.02	0.33	0.11	0.06	0.40	0.08	0.34	0.00	*0.60	0.05	0.06	0.01	0.04
BB	O	**1.60**	0.27	0.31	0.24	***1.85**	1.60	0.13	0.73	**0.60**	**0.56**	0.08	0.68	0.25
	U	0.05	0.47	0.50	0.58	0.23	0.15	****1.33***	***0.50***	0.26	0.40	0.98	0.16	0.58
GRE	O	1.14	**0.51**	0.86	**0.59**	0.89	**1.77**	0.37	*3.74	0.20	0.05	0.25	0.84	**0.59**
	U	0.11	0.42	0.29	0.56	0.62	0.21	*1.07	0.28	0.65	***1.04***	0.72	0.13	***0.80***
BAR	O	0.24	0.47	**1.15**	0.59	0.10	1.67	**1.12**	***4.20**	0.05	0.16	**0.44**	**0.98**	0.31
	U	0.46	0.37	0.21	0.25	0.79	0.16	0.39	0.03	*0.83	0.52	0.32	0.12	0.34
BS	O	0.03	0.00	0.01	0.01	0.02	0.00	0.08	0.11	0.00	0.04	0.02	*0.17	0.06
	U	0.13	0.20	0.18	0.18	0.17	*0.20	0.12	0.09	0.18	0.16	0.18	0.09	0.07
Pan-Arctic	O	4.36	1.94	4.03	2.41	4.16	6.57	4.25	11.69	1.40	2.74	1.67	5.20	2.04
	U	2.68	5.91	3.80	5.75	5.71	5.41	6.02	1.83	6.62	3.90	6.11	1.32	4.40

Bold values indicate individual model maximum SPS overestimation (O) and bold italics indicate the maximum SPS underestimation (U) for each region (see Figure 32b; KS = Kara Sea; LAP = Laptev Sea; ESS = East Siberian Sea; CS = Chukchi Sea; BEA = Beaufort Sea; CAA = Canadian Arctic Archipelago; CEN = Central Arctic; BAF = Baffin Bay; GRE = Greenland Sea; BAR = Barents Sea; BS = Bering Sea). The largest SPS-O and SPS-U in each region are indicated by *.

D. DISCUSSION AND CONCLUSIONS

In light of modest improvements of sea ice simulation in CMIP6 models over previous CMIP phases (Davy and Outten 2020; Shu et al. 2020; SIMIP Community 2020; Shen et al. 2021), we investigated how these models represent the spatial patterns of ice thickness and ice edge. While standard sea ice integrated analyses are good in identifying highly biased models, we caution against the potential for seemingly 'good' models arriving at the 'right' answer for the wrong reasons, especially without ensuring that the responsible physics to get the state correct is reasonably resolved and not by chance of internal variability (Stroeve and Notz 2015; Jahn et al. 2016). We argue that basic comparisons of time series of integrated SIE and SIV are not enough to qualify model performance and diagnose biases. Hence, we present spatial analysis techniques to assist.

Qualitative examination of the SIT spatial distributions showed that half of the analyzed CMIP6 models reasonably simulated March and September SIT (Figure 25 and Figure 26) for the period 1979-2014 against PIOMAS reanalysis. However, spatial pattern correlation and RMSE (Figure 27) does not clearly distinguish models with poor SIT patterns or lend itself to regional diagnostics. Because of the narrow range of values, e.g., PIOMAS ranges r=0.73 to 0.92 and RMSE=0.43 m to 0.81 m (Figure 25) and CryoSat-2 ranges r=0.47 to 0.65 and RMSE=0.63 m to 0.92 m (Figure 27b and Figure 27d), it is difficult to classify meaningful SIT differences between models.

We performed an ice edge analysis and identified several common 'problem' regions, which may benefit from further regional analyses at the individual model level. Note that the analyses with SIA carry similar, if not larger, uncertainties related to coarse resolution in both models and satellite measurements, especially due to melt pond coverage during the melt season, detection of thin ice (by passive microwave) and the MIZs (Ivanova et al. 2014; Yang et al. 2016; Comiso et al. 2017). To our understanding, the limitation with differentiating melt ponds and low SIC from open ocean surface by passive microwave satellites is the primary reason why the NSIDC and other centers choose SIE as their primary and less uncertain diagnostics of summer-time ice cover. Given the above concerns with SIA, we find SPS and IIEE as reasonable options to quantify regional errors

in the simulated sea ice cover and to narrow possible oceanic/atmospheric drivers of such errors. Hence, given the continued challenges with accurate SIA estimates the ice edge analysis appears to be a promising methodology for diagnosing model biases related to specific regional processes.

Our spatial analysis metrics identified Arctic regions with large ice edge and ice thickness errors, therefore pointing to limitations in or lack of representation of some physical processes within individual CMIP6 models. We postulate that many of them could be related to the oceanic forcing in the marginal and shelf seas. For example, during the warmer months, the SPS in the interior Arctic is generally largest over the East Siberian and Kara seas (Table 10), with many sea ice simulations tending to respond slowly to the seasonal changes (i.e., slow to melt or/and slow to grow). In case of errors of ice growth in these regions, we hypothesize this may be the result of models taking additional time to ventilate the excess heat accumulated in the upper ocean through the summer, to cool water down to freezing temperature in order to begin the ice growth. Another related cause of such errors might be an over estimation of the surface mixed layer depth, which would also take more time to cool down before freezing. Note that these two issues would affect a model representation of Arctic amplification hence more than just simulation of sea ice.

In the case of models that consistently underestimate the position of the sea ice edge for the Barents and Kara seas (e.g., GFDL-ESM4 and IPSL-CM6A-LR in Figure 26g and Figure 26i), possible causes may be an overestimated transport of warm Atlantic Water across or underestimated cooling over the Barents Sea (Maslowski et al. 2012a). The resulting excess of OHT could be accelerating ice melt and delaying freeze-up over the Barents Sea and further downstream, over the northern Kara Sea (and beyond).

During the colder months, the majority of SPS are overestimations over the Greenland and Barents seas and underestimations over the Bering Sea. On the European side, this could again be related to variability in volume and heat fluxes from the North Atlantic Ocean and their distribution between the Labrador, Greenland, and Barents seas, which may lead to interrelated inaccuracies in simulated sea ice over those regions. For example, if too much warm Atlantic water is diverted from entering the Labrador or eastern Greenland seas, and instead enters the Barents Sea, we could expect an overestimation of

ice melt in the latter region with underestimation of melt in the former two regions (e.g., IPSL-CM6A-LR in Figure 25i). Onarheim et al. (2018) found that the Barents Sea, considered a hotspot for recent climate change, contributes the largest fraction of winter sea ice loss through 2016. Possible explanations of the significant loss in the Barents Sea include a decreased sea ice import and the impact on upper ocean stratification (Lind et al. 2018) and greater control of the sea ice cover after around 2003 by Atlantic water ocean heat anomalies across the Nordic and Barents seas (Schlichtholz 2019). We found that the majority of CMIP6 models, consistent with Shu et al. (2020), underrepresent the strength of March SIC decline over the Barents Sea compared to NSIDC (not shown), which is indicative of potential model limitations in representing these dynamics.

Sea ice edge and thickness distribution is also determined by atmospheric forcing. Previous CMIP studies have shown that inaccurate SIT patterns are often associated with model deficiencies in simulating the observed Arctic sea level pressure patterns and associated geostrophic winds (Schweiger et al. 2011; Stroeve et al. 2014). Inaccurate simulations of the position and/or strength of large-scale circulation patterns, like the Beaufort High, can allow for sea ice convergence in regions not generally observed (DeRepentigny et al. 2016). Additionally, strong cyclones and anomalous circulation patterns can also contribute to anomalous SIT and SIE patterns in model simulations. For example, atmospheric heat and moisture drawn into the high-latitudes can increase downward longwave radiation which inhibits sea ice growth (Cai et al. 2020). Also, synoptic storms can excite inertial oscillations in sea ice (Roberts et al. 2015), which can lead to increased deformation rates and hence change in SIT.

Modeled SIT in the Bering Sea generally replicates observed ice thickness patterns well, but models tend to underestimate the SIE there. Given that sea ice in the Bering Sea melts completely every summer, most of the CMIP6 models get the SIT correct as long as they can replicate that seasonal retreat. However, first-year sea ice is more susceptible to atmospheric forcing and its variability (Rampal et al. 2011), hence the problems with simulation of the ice edge there. Model limitations in representing the Bering Sea dynamics also affect the Chukchi and East Siberian seas downstream, as those marginal seas are tightly connected via the advection of Pacific water (Maslowski et al. 2014). The

northeastward transport of warm Pacific summer water across Bering Strait into the East Siberian Sea is another process likely to be difficult to represent in the CMIP6 models with coarse horizontal resolution (e.g., the width of the Bering Strait is less than the single grid cell size for all but three models examined here). Yet the oceanic forcing of sea ice in the East Siberian Sea might play a role in its negative concentration trend, which Onarheim et al. (2018) identified as the largest in observed summer ice loss in the Arctic through 2016.

The choice of CMIP6 ice edge analysis technique as a model bias diagnostic was shown to be mostly arbitrary between the SPS and the mean IIEE. The exception to this is the case of single member simulations. Because not all CMIP6 models provide multiple ensemble realizations, the deterministic models may appear less skillful by the SPS in representing the ice edge than their ensemble model counterparts (see GFDL-ESM4, Figure 30a and Figure 30b). While the benefits of using ensemble models are beyond argument, our application of the mean IIEE identified the same relative model biases as the SPS, albeit with larger magnitudes. As such, the IIEE approach may provide a more equitable comparison of an individual model's skill in randomly representing the sea ice edge in any single realization (Figure 31).

RASM-G is examined alongside the CMIP6 models and achieves favorable results in all categories, so it is reasonable to consider it as a realistic alternative sea ice reference simulation of the multi-decadal evolution of the Arctic sea ice, especially with regard to SIT and SIV. The RASM-G configuration is similar to PIOMAS, in which both models use atmospheric reanalysis to force the ocean and sea ice model components. However, PIOMAS assimilates SIC and sea surface temperature data in order to improve model performance of sea ice characteristics. Yet, assimilated SIC fields come with passive microwave uncertainties (Yang et al. 2016), described earlier, which might contribute to the reported PIOMAS SIT and SIV biases and uncertainty in its trend estimates (Lindsay 2010; Schweiger et al. 2011a). RASM-G does not use data assimilation for its components and still performs remarkably well in replicating observed SIE, as well as comparable to PIOMAS SIT distribution and SIV time series over multiple decades. We attribute this in part to the realistic atmospheric forcing, as well as to a more realistic representation of sea ice relevant processes (e.g., oceanic forcing along MIZs) and high spatial and temporal

resolution. This lends confidence in the RASM-G model physics yielding correct depiction of the mean state and evolution of the Arctic ice pack.

E. ACKNOWLEDGEMENTS

We acknowledge the WVRP, which, through its Working Group on Coupled Modeling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the ESGF for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. We acknowledge partial support from the following programs: US Navy (to MW), the DOE RGMA, the ONR AGP and NSF ARCSS (to WM, YL, JCK), and Ministry of Science and Higher Education in Poland (to RO). The DOD HPCMP provided computer resources for RASM-G simulations analyzed here.

F. DATA AVAILABILITY STATEMENT

CMIP6, CryoSat-2, EUMETSAT, ICESat, NSIDC, and PIOMAS data used for this study can be acquired from the links provided in Sections IV.B.1 and IV.B.2. The RASM-G data can be acquired from the Naval Postgraduate School (https://nps.app.box.com/folder/139647168752?s=xyp563ee40w6lffn4718zr52a6m7rtcc).

V. ASSESSMENT OF OCEANIC HEAT TRANSPORT AND ITS EFFECTS ON PAN-ARCTIC SEA ICE DECLINE IN CMIP6 HISTORICAL SIMULATIONS

A. INTRODUCTION

The Earth's energy balance is characterized by a surplus of incoming solar energy over the tropics and midlatitudes and a solar energy deficit at and near the poles (von Schuckmann et al. 2016). Thus, poleward atmospheric and oceanic heat transport acts to distribute the radiative imbalance (Czaja and Marshall 2006). Averaged over the Arctic polar cap (i.e., 70°–90°N) since 2000, this amounts to annual mean atmospheric and oceanic net meridional transports of energy into the Arctic of about 1400 TW and 190 TW, respectively (Mayer et al. 2019). For long-term mean energy conservation, the net radiation at the top of the atmosphere (TOA) must balance this regional influx of heat (Mayer et al. 2019). However, less energy is leaving via the TOA than is being transported in, and this discrepancy is explained by energy storage associated with ocean warming and ice melting. The Arctic energy imbalance is on the order of about 12 TW (~1 Wm^{-2}) reported by Mayer et al (2019), with 2/3 warming the ocean and 1/3 melting the sea ice. Variability in the amount of OHT has been shown to impact the Arctic sea ice cover, particularly in winter over the Barents and Greenland seas during the periods of maximum SIE (Francis and Hunter 2007; Polyakov et al. 2010; Årthun et al. 2012; Smedsrud et al. 2013). However, the magnitude of oceanic heat convergence in such regions as the Barents and Kara seas might be significantly underestimated in many coarse resolution global climate models (Smedsrud et al. 2013).

In this study, OHT and its impact on the historical simulations of pan-Arctic sea ice cover is examined in a subset of models participating in CMIP6. Earlier in Chapter III, we showed that CMIP6 model simulations do not generally replicate the strength of the accelerated rate in SIA decline beginning around the year 2000. In Chapter IV, we identified Arctic subregions with large ice edge errors, suggesting these may be the result of model limitations in representing some physical processes. Collectively, the Arctic regions that are responsible for the accelerated rate in the observed SIC decline (Shu et al.

2020) are also the same regions for which many CMIP6 models showed large ice edge errors. These regions are the Kara, Laptev, and Barents seas (Figure 32). They are also closely connected to the Atlantic Water inflow and are most impacted by recent Atlantification of the Arctic (Årthun et al. 2012).

Studies of *in situ* oceanic volume and heat transports have mostly been limited to individual Arctic gateways and for relatively short time periods. Table 11 shows the observational estimates for the four primary Arctic gateways from a number of observation studies (Schauer et al. 2008; Smedsrud et al. 2013; Curry et al. 2014; Roach et al. 1995; Woodgate et al. 2012). The results from individual gates have large variability and uncertainty and collectively fail to 'close' the Arctic Ocean volume (or mass) budget. Additionally, a closed volume is prerequisite to determine the amount of ocean heat gained or lost within (Schauer et al. 2008), defined here as OHT convergence (OHTC). Thus, confidence in the OHT estimates for individual gateways is questionable (Schauer et al. 2008). However, recent studies by Tsubouchi et al. (2018, 2021) overcome the mass budget discrepancy by using the box inverse model along with *in situ* data from moored and shipboard instruments. This mass constrained reconstruction offers a holistic approach to the Arctic mass and energy budget, and to improve OHT and OHTC estimates (Table 11). Also, a significant OHT increase of 21 TW into Arctic Mediterranean is reported to have occurred since 2001 (Tsubouchi et al. 2021).

In this study, we analyze OHT analysis in CMIP6 models using the model variable northward OHT, *hfbasin* (Griffies et al. 2016). To our knowledge, this is the first CMIP6 study to present such analyses, and thus may be a useful baseline for the CMIP6 models, and for future discovery and comparison. We first examine the global OHT mean annual cycle against OHT from reanalysis data sets and estimates for the Arctic OHTC. Next, we evaluate the 1979-2014 time series, focusing on OHT mean and trends, and on the role of Atlantic OHT forcing on the pan-Arctic sea ice cover.

Table 11. Ocean volume and heat transport for the four primary pan-Arctic gateways

	Volume Transport (Sv)								Oceanic Heat Transport (TW)							
	FS		BSO		DS		BE		FS		BSO		DS		BE	
	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
Observations	−2.0[1]	2.7[1]	2.3[2]	1.0[2]	−1.6[3]	0.7[3]	0.8[4]	0.7[4]	36[5]	6[5]	70[2]	21[2]	20[3]	9[3]	10-20[6]	-
Tsubouchi et al. 2018	−1.1	1.2	2.3	1.2	−2.1	0.7	0.7	0.7	^58	^10	*64	*33	^29	^10	**8	**13
ORA MM[7]	−1.9	-	3.0	-	−2.3	-	1.1	-	21	-	72	-	5	-	5	-

Ocean volume and heat transport (OHT) annual mean and standard deviation (s.d.) through the Fram Strait (FS), the Barents Sea Opening (BSO), the Davis Strait (DS), and the Bering Strait (BE). OHT for Tsubouchi et al. (2018) are shown as referenced to the temperature used in the observational study annotated (0.0C*,−0.1C^,−1.9C**) . Observations cover a range of time periods (see reference texts for details), while the ocean and sea ice reanalysis multi-model mean (ORA MM) is for 1993–2010 (Uotila et al. 2019). [1]Schauer et al. (2008), [2]Smedsrud et al. (2013), [3]Curry et al. (2014), [4]Roach et al. (1995), [5]Schauer and Beszczynska-Möller (1995), [6]Woodgate et al. (2012), [7]Uotila et al. (2019)

B. DATA AND METHODS

1. Model Output and Observational Data

Our study used the CMIP6 historical experiment data to evaluate the Arctic sea ice response to OHT for the period 1979-2014 (Notz et al. 2016; Eyring et al. 2016; Griffies et al. 2016). Models were selected based on the availability of required variables at the time of manuscript preparation. Model outputs for 17 CMIP6 models, including 165 individual ensemble members, were retrieved from the ESGF repository (https://esgf-node.llnl.gov/search/cmip6/). Table 1 summarizes the metadata and the variables used for this model evaluation. We only analyzed outputs from models that provide data for the variable *hfbasin*. This variable is the "depth-integrated northward net OHT integrated within [the Atlantic, Indo-Pacific, and global] basins" (Griffies et al. 2016).

We chose the same SIA and SIV observational reference data sets for 1979-2014 described in Chapters III and IV. The SIA observational reference was calculated after averaging data from three SIC algorithms (i.e., NSIDC NT and BT, and OSI-SAF). For SIV, we used the CMIP6 SIV multi-model mean (MM) as our primary temporal analysis reference, and include the PIOMAS (Zhang and Rothrock 2003) and the RASM-G for comparison. The CMIP6 SIV MM time series was calculated by averaging the 42 CMIP6 model ensemble means.

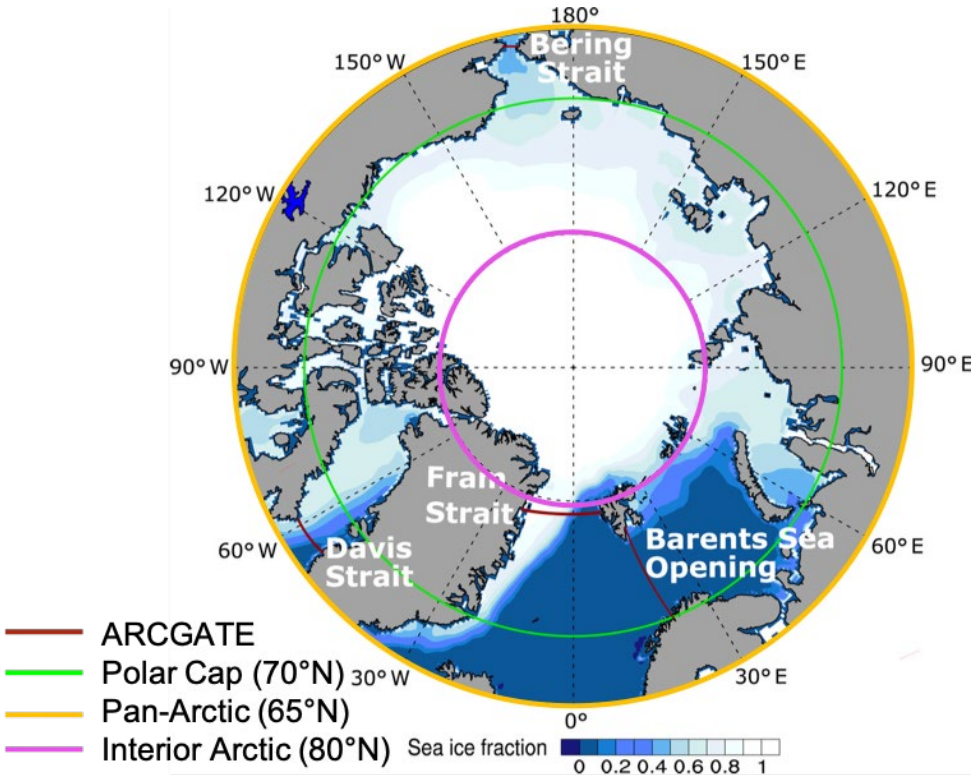
To compare model simulations of OHT against recent reanalyses, we retrieved the mass-balanced monthly mean OHT from Norwegian Marine Data Centre (Tsubouchi et al. 2021; http://metadata.nmdc.no/metadata-api/landingpage/0a2ae0e42ef7af767a920811e83784b1). Additionally, mean monthly OHT (Mayer et al. 2019) was retrieved from ECMWF Ocean Reanalysis System 5 (ORAS5; Zuo et al. 2019; https://icdc.cen.uni-hamburg.de/daten/reanalysis-ocean/easy-init-ocean/ecmwf-oras5.html) and Centro Euro-Mediterraneo sui Cambiamenti Climatici coupled ocean-sea ice reanalysis (C-GLORS v7; Storto and Masina 2016; http://c-glors.cmcc.it/index/index-7.html?sec=7).

2. Methods

We examined CMIP6 ensemble means for all models and all variables, except for the single realization SAMO-UNICON. Use of the ensemble mean instead of a single realization emphasizes our interest in the mean forced signal and reduces the influence of model internal variability. For temporal analysis, individual model ensemble mean time series are calculated by averaging their ensemble members. The pan-Arctic SIA and SIV time series is calculated as described in the previous chapter. The CMIP6 OHT time series is determined for specific lines of latitude directly from the CMIP6 variable northward OHT (*hfbasin*).

Sea ice and OHT anomalies were determined by removing the respective mean annual cycle for individual models, observations, and reanalyses. Unless otherwise indicated, the anomaly period is 1979-2014. Characteristic trends and their uncertainties were calculated by simple linear regression and following Santer et al. (2008) to adjust for the effects of autocorrelation in the number of independent time samples (see Chapter III for our calculation of confidence interval).

Figure 37 shows the approximate locations of the Arctic gateways and the southernmost lines of latitude used to determine OHT and OHTC in observations and reanalyses (Tsubouchi et al. 2018; Mayer et al. 2019) and in the CMIP6 models for this study. We chose 65°N to examine the global pan-Arctic OHT and its component pathways, which include the Bering Strait (i.e., Pacific Ocean pathway) and the Davis Strait and Nordic Seas (Atlantic Ocean pathways). The line of latitude 80°N encompasses the predominant Atlantic-Arctic Ocean communication pathways, namely the Fram Strait between Greenland and Svalbard, and the land arc including Svalbard, Franz Josef Land, and Severnaya Zemlya. Additionally, 80°N encompasses the Nares Strait, located between the Canadian Arctic Archipelago and Greenland, an important pathway for Arctic water export.



Approximate locations of the gateways and lines of latitude where OHT was determined by observations or calculated for CMIP6 simulations and ocean reanalyses. The ARCGATE region (Tsubouchi et al. 2018) is defined by Bering Strait, Barents Sea Opening, Fram Strait, and Davis Strait gateways. The polar cap (70°N), pan-Arctic (65°N), and interior Arctic (80°N) regions indicate the southernmost lines of latitude used for OHT convergence with 90°N.

Figure 37. Map of pan-Arctic gateways and lines of latitude used in this study. Adapted from Mayer et al. (2019).

The northward OHT (in TW) across a line of latitude is the total amount of poleward OHT supplied from the lower latitudes to the region at and poleward of the latitude. For the global basin, this is the OHTC in TW. However, this definition of the OHTC does not account for other processes that determine the total energy content of the ocean in the defined region, e.g., surface absorption of shortwave radiation or surface heat exchanges with the atmosphere. Still, this definition is helpful as we examine the magnitude of ocean heat supplied to the Arctic, particularly during the periods of total darkness and very cold SATs. We also assess the ocean basin specific northward OHT (i.e., Atlantic and Pacific basins). Unlike the global basin, the Atlantic and Pacific basins do not explicitly close the mass and heat budget required to determine basin specific OHTC. In

this case, convergence must be determined for a specified region within two lines of latitude, e.g., between 65°N and 80°N in this study. Thus, we differentiate between the flux variable OHT and OHTC.

For spatial analysis of SIA and SIT, we first linearly regridded CMIP6 model SIC and SIV onto the NSIDC SIPS North 25 km x 25 km grid (following Watts et al. 2021) before averaging. When plotting SIC or SIT mean and linear trend figures, a mask was applied to pixels with SIC less than 10% or SIT less than 10, respectively.

C. RESULTS

1. Oceanic Heat Transport Mean and Seasonal Cycle

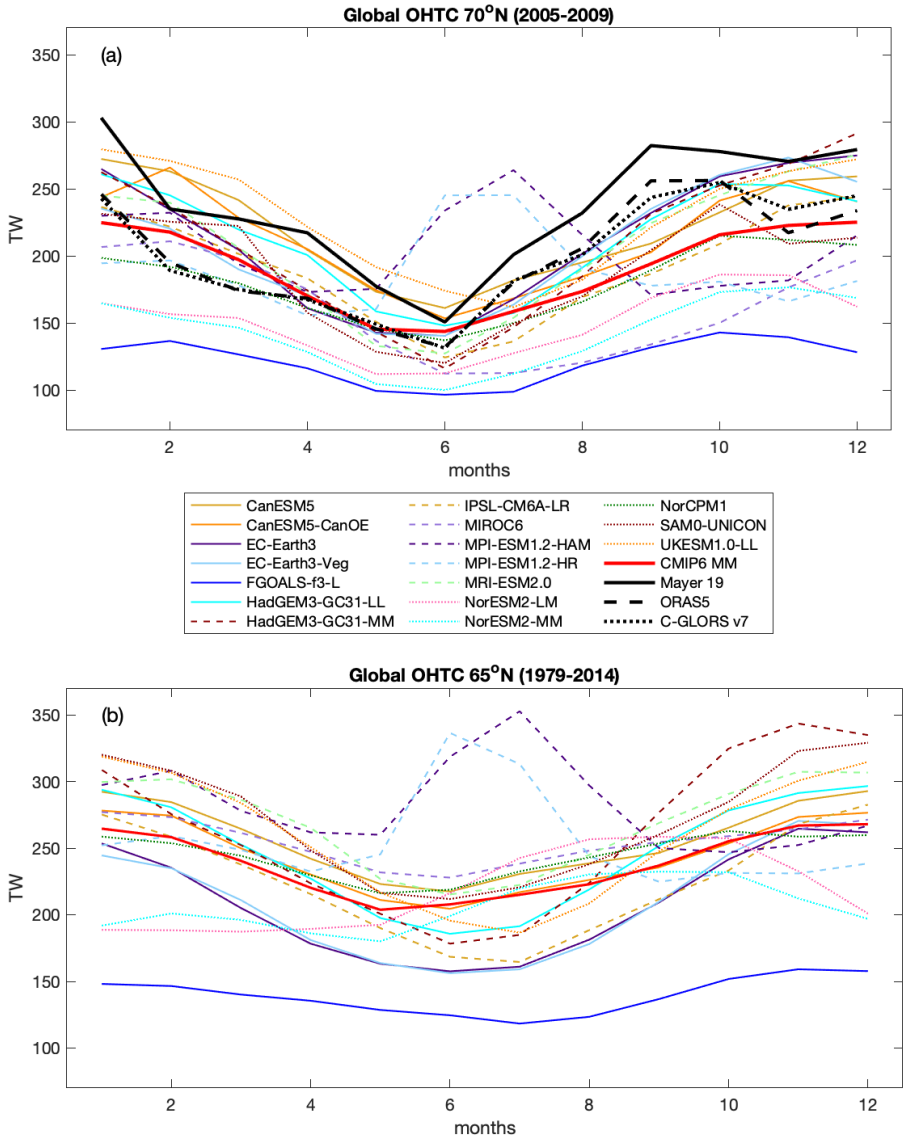
First, we compare the CMIP6 model simulations of northward OHT against recently published reanalyses. Table 12 compares ARCGATE observations (Tsubouchi et al. 2018) of the 2005 to 2009 annual mean OHTC to estimates from ocean reanalyses (e.g., ORAS5 and C-GLOS v7) and an estimate from Mayer et al. (2019) using a variational approach to enforce budget closure. The ORAS5 and C-GLORS v7 both underestimate ARCGATE observations by 9% and 15%, respectively, while Mayer et al. (2019) overestimates the same by about 5%. Note that the Barents Sea Opening (included in ARCGATE observations) is a meridional gateway and does not lend itself to CMIP6 northward OHT analysis, thus no CMIP6 MM estimate is shown (Table 12). However, the polar cap (70-90°N) results from Mayer et al. (2019) provide good comparison for CMIP6 models. For 2005–09, the mean polar cap OHTC for both ORAS5 and C-GLORS v7 underestimate the Mayer et al. (2019) by about 15%, and CMIP6 OHTC MM underestimates it by about 20%. For the longer period 2001–14, the CMIP6 MM shows similar mean OHTC as reanalyses (Table 12).

Table 12. Estimates of oceanic heat transport. Adapted from Mayer et al. (2019).

	Mean (ARCGATE)		Mean (Polar Cap)		Trend
	2005-2009	2001-2017	2005-2009	2001-2017	2001-2017
Tsubouchi et al. 2018	156	-	-	-	-
Mayer et al. 2019	163	-	238	-	-
ORAS5	141	131	201	194	1.11
C-GLORS v7	132	128	204	192	1.41
CMIP6 MM	-	-	191	193	*0.72

OHT convergence annual mean (TW) and monthly mean anomaly linear trend (TW per year) for Arctic gates (ARCGATE) observations, ocean reanalyses, and CMIP6 multi-model mean (MM) for the periods indicated. ARCGATE are defined in Tsubouchi et al. (2018). * indicates 2001-2014 for CMIP6 MM historical simulation.

Figure 38a shows the 2005 to 2009 CMIP6, Mayer et al. (2019), and ORAS5 polar cap mean annual cycle for Global OHTC. The CMIP6 OHTC MM shares a similar seasonal cycle with Mayer et al. (2019), ORAS5, and C-GLORS v7 reanalyses consisting of January maximum and June minimum. However, the magnitude of the CMIP6 OHTC MM mean is 20% less (Table 12) and each month is smaller than Mayer et al. (2019), particularly for July–January (Figure 38a). Individual CMIP6 simulations of OHTC have large model spread, and all but the two MPI-ESM1.2 models generally simulate the estimated mean annual cycle. Both MPI-ESM1.2 models show peak OHTC in summer (Figure 38a).



Seasonal cycles of global OHT convergence (OHTC) across (a) 70°N for CMIP6 models and ocean reanalyses and estimate from Mayer et al. (2019) for the period 2005 to 2009, and (b) 65°N for CMIP6 models for the period 1979-2014. The CMIP6 multi-model mean is displayed by a red line.

Figure 38. Global oceanic heat transport convergence mean annual cycle

For our study, we expand the area of interest from 70°N to 65°N (i.e., pan-Arctic 65°–90°N) to examine a larger portion of the Nordic Seas and incorporate the Pacific inflow through the Bering Strait. Figure 38b shows the global OHTC pan-Arctic mean annual cycle for CMIP6 models for 1979 to 2014. The OHTC calculated at this more southerly latitude results in a larger magnitude than in Figure 38a, but the mean annual

cycles are similar. Here, the CMIP6 MM is maximum in December and minimum in May, and simulates OHTC of about 270 TW and 200 TW, respectively. With the exception of the two MPI-ESM1.2 models and FGOALS-f3-L, the simulated peak OHTC ranges from 230 TW (NorESM2-MM) to 350 TW (HadGEM3-GC31-MM) and minimum OHTC ranges from 160 TW (both EC-Earth3 models) to 240 TW (MIROC6). FGOALS-f3-L simulates a relatively weak OHTC seasonal cycle and has the largest negative bias (about –100 TW) against the CMIP6 MM (Table 13). Both MPI-ESM1.2 models show a unique OHTC peak around 350 TW during June–July.

Table 13. Oceanic heat transport and convergence across lines of latitude 65°N and 80°N

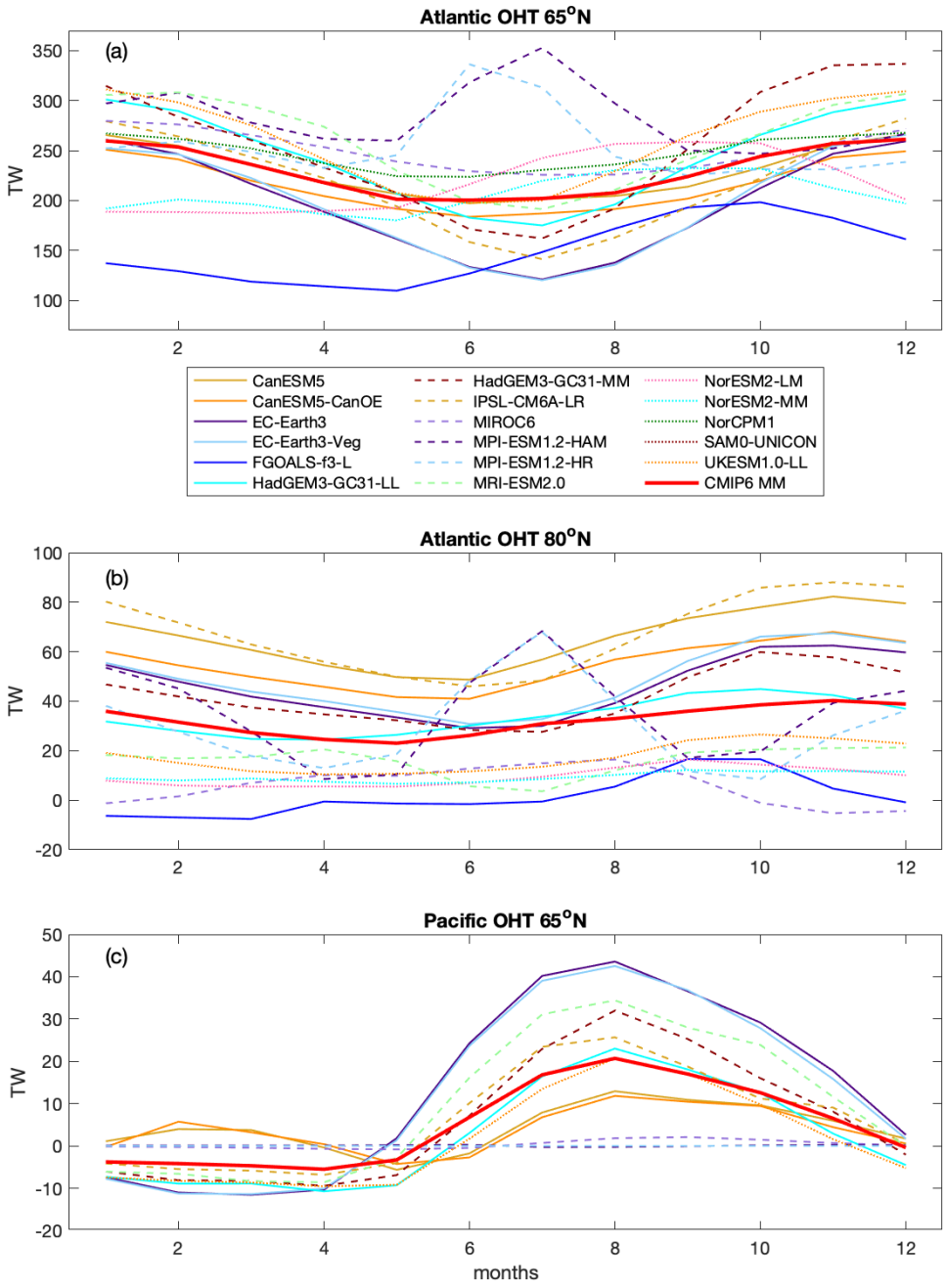
	Oceanic Heat Transport										Oceanic Heat Convergence			
	65°N						80°N				65°N minus 80°N			
	Atlantic		Pacific		Global		Atlantic		Global		Atlantic		Global	
	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
CanESM5	229	38	4	6	257	37	66	29	67	28	164	21	190	21
CanESM5-CanOE	215	43	4	7	244	42	55	32	56	31	161	26	188	26
EC-Earth3	196	55	13	21	209	47	46	19	37	20	151	43	172	30
EC-Earth3-Veg	198	55	12	21	210	47	49	19	39	20	149	43	171	30
FGOALS-f3-L	149	34	-	-	139	19	1	11	-9	23	148	28	148	28
HadGEM3-GC31-LL	245	52	2	12	247	49	34	15	24	15	211	48	223	41
HadGEM3-GC31-MM	255	63	6	15	261	59	42	15	31	17	213	53	230	46
IPSL-CM6A-LR	218	47	6	12	224	40	68	17	63	18	151	39	161	29
MIROC6	250	21	0	1	254	18	6	9	6	9	244	27	249	24
MPI-ESM1.2-HAM	282	39	0	0	282	39	35	28	35	28	247	36	247	35
MPI-ESM1.2-HR	255	36	0	0	255	36	30	19	30	19	225	28	225	28
MRI-ESM2.0	260	46	9	17	270	37	16	7	14	8	244	41	256	34
NorESM2-LM	218	35	-	-	218	35	9	6	9	6	208	32	208	32
NorESM2-MM	206	25	-	-	206	25	9	4	9	4	197	25	197	25
NorCPM1	248	17	-	-	244	17	-	-	11	5	-	-	233	16
SAM0-UNICON	-	-	-	-	271	49	-	-	9	6	-	-	261	51
UKESM1.0-LL	260	46	1	11	259	49	17	11	27	11	243	40	233	48
CMIP6 MM	230	25	5	9	238	25	32	10	27	8	197	21	211	19

Mean and standard deviation (s.d.) of oceanic heat transport (TW) across lines of Latitude 65°N and 80°N, and difference between them, oceanic heat convergence (TW), for CMIP6 models for the period 1979-2014.

Next, we show separately the CMIP6 mean annual cycle for the Atlantic and Pacific basins OHT. At 65°N, seasonal cycles of the Atlantic OHT (Figure 39a) are similar to Global OHTC (Figure 38b) for the CMIP6 MM and for individual models. Note that FGOALS-f3-L shows the largest difference between the Atlantic OHT and global OHTC mean annual cycles, with maximum Atlantic OHT about 50 TW larger than the Global OHTC (Figure 38b and Figure 39a). This suggests a substantial amount of ocean heat exits the Arctic by some alternate pathway not accounted for in the Atlantic basin OHT, e.g., possibly the Canadian Arctic Archipelago. For the Pacific OHT seasonal cycle at 65°N (Figure 39b), most models show maximum OHT in August (10–45 TW) and minimum in December–June (–5 TW), with the CMIP6 MM showing peak August OHT of about 20 TW. Note that only 12 of 17 CMIP6 models provide values for Pacific OHT at 65°N, and three more have zero annual mean (Table 13).

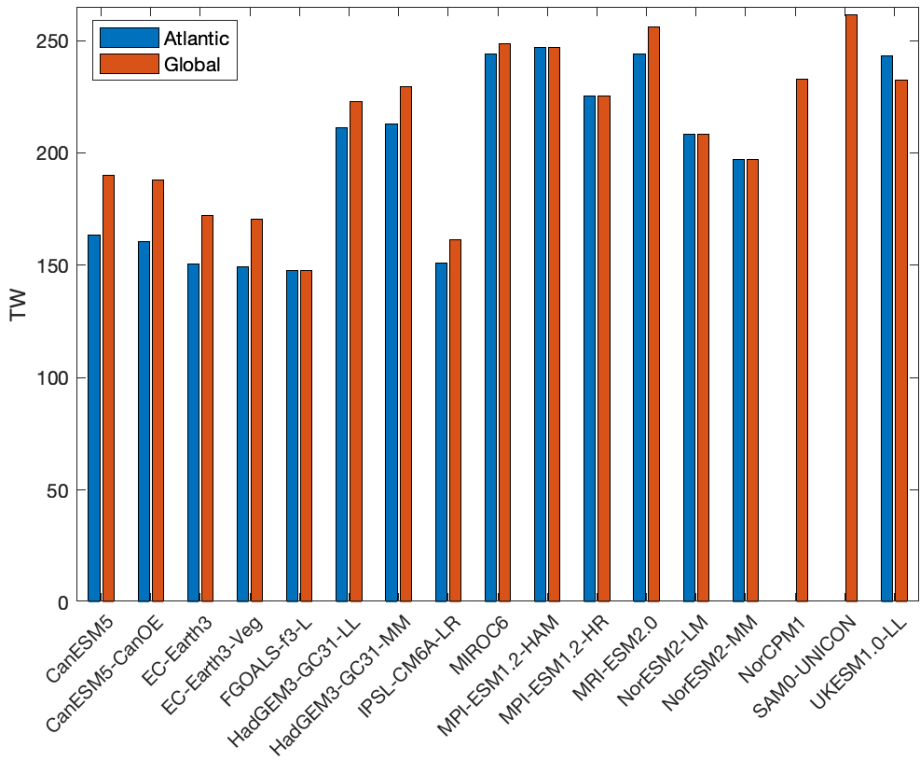
At 80°N, only the Atlantic basin OHT seasonal cycle is shown because the models do not differentiate Pacific basin transport north of the Bering Strait (Figure 39c). Here the CMIP6 Atlantic OHT MM seasonal cycle ranges between 20–40 TW with November maximum and June minimum with a mean value of 30 TW mean (Table 13). Individual CMIP6 models are evenly spread from about zero (FGOALS-f3-L) to 68 TW (IPSL-CM6A-LR), indicating a wide range in magnitude of Atlantic origin OHT entering the Arctic interior (i.e., poleward of 80°N).

Next, we investigate the OHTC between 65°–80°N in CMIP6 models (Figure 40 and Table 13). Most oceanic heat transported into the Arctic region at 65°N is of Atlantic Ocean origin and the CMIP6 OHT MM shows as much as 96% (Table 13). The relatively warm Atlantic water flows north of 65°N through the Nordic Seas and splits between the Fram Strait or the Barents Sea Opening gateways. As the Atlantic water flows poleward it exchanges energy with the atmosphere and sea ice, thus decreasing the magnitude of northward OHT (Skagseth et al. 2008). Alternatively, ocean heat enters the Arctic region from the Pacific Ocean by way of the Bering Strait gateway (Woodgate et al. 2010).



Mean seasonal cycle of Atlantic Water OHT across (a) 65°N and (b) 80°N, and Pacific Water OHT across (c) 65°N for CMIP6 models for the period 1979-2014. The CMIP6 multi-model mean are displayed by a red line.

Figure 39. Mean annual cycle of Atlantic and Pacific oceanic heat transport



Global and Atlantic mean OHT convergence (OHTC) between 65°N to 80°N for CMIP6 models for the period 1979-2014. No OHTC is calculated for NorCPM1 and SAM0-UNICON because no value for Atlantic OHT at 80°N is specified.

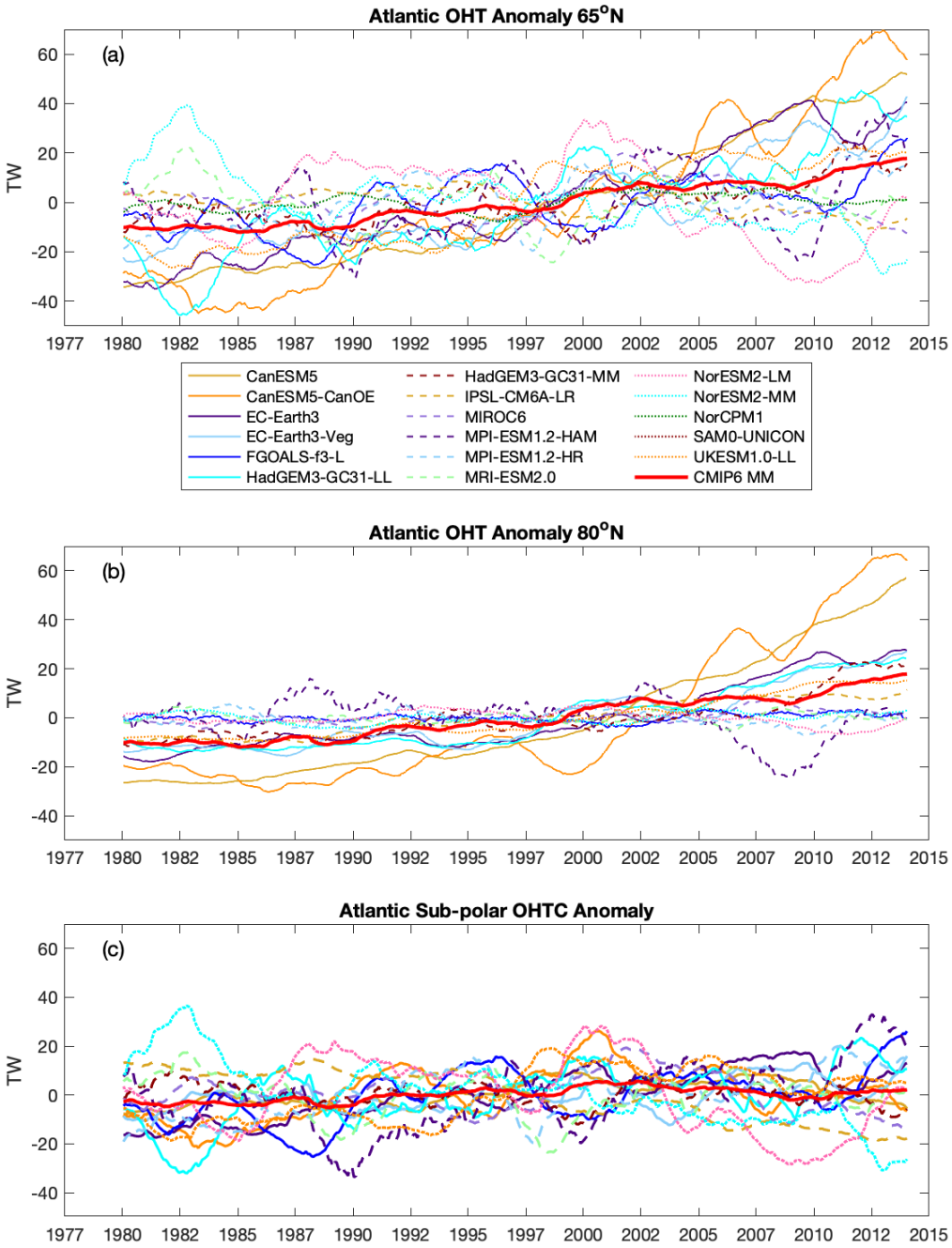
Figure 40. Global and Atlantic mean oceanic heat transport convergence for the period 1979-2014

The CMIP6 Atlantic sub-polar OHTC MM is 197 TW, while the CMIP6 global sub-polar OHTC MM is about 7% larger at 211 TW (Table 13). Individual CMIP6 models simulate mean Atlantic sub-polar OHTC ranges from 148 TW (FGOALS-f3-L) to 247 TW (MPI-ESM1.2-HAM). In terms of the percent of OHTC within the Atlantic sub-polar region, the CMIP6 models range from about 70% (CanESM5 and IPSL-CM6A-LR) to 99% (FGOALS-f3-L and MIROC), and the CMIP6 MM shows 86% convergence (Table 13). In this study, we are most interested in the impact of Atlantic OHT trend on the pan-Arctic sea ice cover, thus the remaining figures and results reflect this. Note that SAM0-UNICON data does not include separate Atlantic or Pacific OHT and is therefore excluded from basin-level analysis (Table 13).

In Figure 41 we examine the 24-month running mean Atlantic OHT and the Atlantic sub-polar OHTC anomalies for the CMIP6 models. The CMIP6 MM shows an increasing trend in OHT for both the pan-Arctic and Arctic interior (Figure 41a and Figure 41b, respectively) that noticeably increases after the year 2000. At 80°N, eight of 16 CMIP6 models (CanESM5, CanESM5-CanOE, EC-Earth3, EC-Earth3-Veg, HadGEM3-GC31-LL, HadGEM3-GC31-MM, IPSL-CM6A-LR, UKESM1.0-LL) show a change in OHT of about 20 TW or more over the study period. Similar OHT change is also evident at 65°N for these models (except for IPSL-CM6A-LR), indicating that the increased OHT signal seen at 80°N propagates through the sub-polar Atlantic region. However, this is not the case for IPSL-CM6A-LR, which shows a decrease in OHT at 65°N, thus less ocean heat converges within the sub-polar Atlantic over the analysis period. Note that MPI-ESM1.2-HAM also shows OHT trend inconsistency between their 65°N and 80°N OHT, indicating more ocean heat converges within the sub-polar Atlantic over time. Altogether, the CMIP6 MM Atlantic sub-polar OHTC anomaly shows little change over the analysis period (Figure 41c).

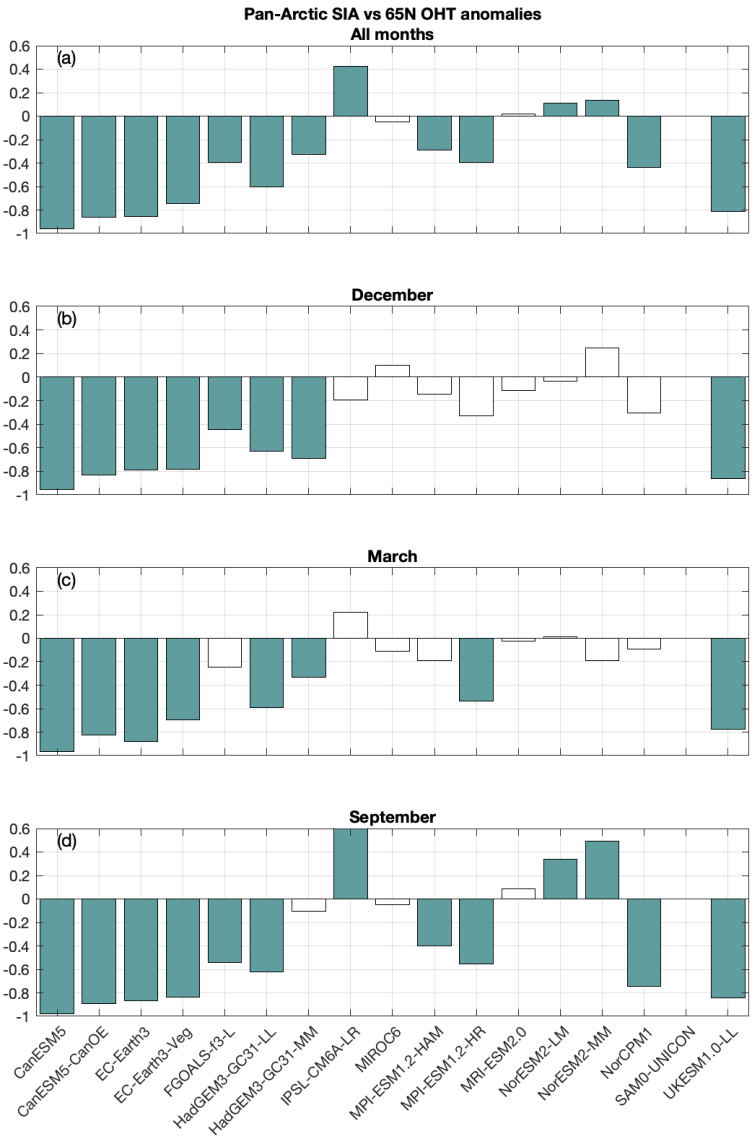
2. Temporal Analysis of OHT Forcing on Pan-Arctic Sea Ice

The expected response of sea ice to increasing OHT is for it to melt or decline. We examine this expected behavior by comparing CMIP6 simulations of monthly mean OHT anomalies against their monthly mean SIA (Figure 42 and Figure 43) and SIV (Figure 44 and Figure 45) anomalies. Here we show correlations for the full time series (i.e., month to month for January 1979 to December 2014) and the selected months of SIA maximum and minimum (March and September, respectively). December is also shown because it falls in the middle of the 80°N polar night, and therefore the atmosphere is generally below the freezing point of sea water and the region does not receive incoming shortwave radiation. As such, the atmospheric forcing on sea ice associated with shortwave radiation should be at or close to a minimum, and therefore somewhat isolate the oceanic forcing signal from the atmospheric.



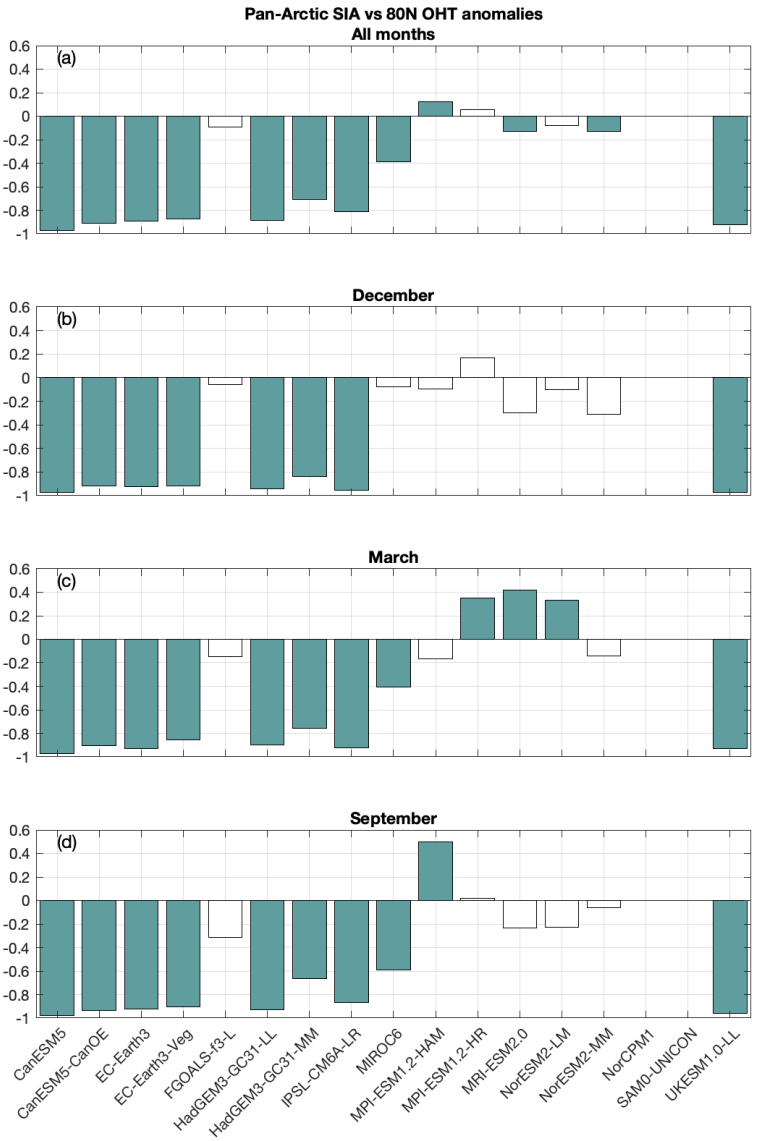
24-month running means of Atlantic OHT anomaly at (a) 65°N and (b) 80°N, and OHTC anomaly between 65°N to 80°N for CMIP6 models for the period 1979-2014. No OHTC anomaly is calculated for NorCPM1 and SAM0-UNICON because no value for Atlantic OHT at 80°N is specified. The CMIP6 multi-model mean are displayed by a red line.

Figure 41. 24-month running means of Atlantic oceanic heat transport anomaly



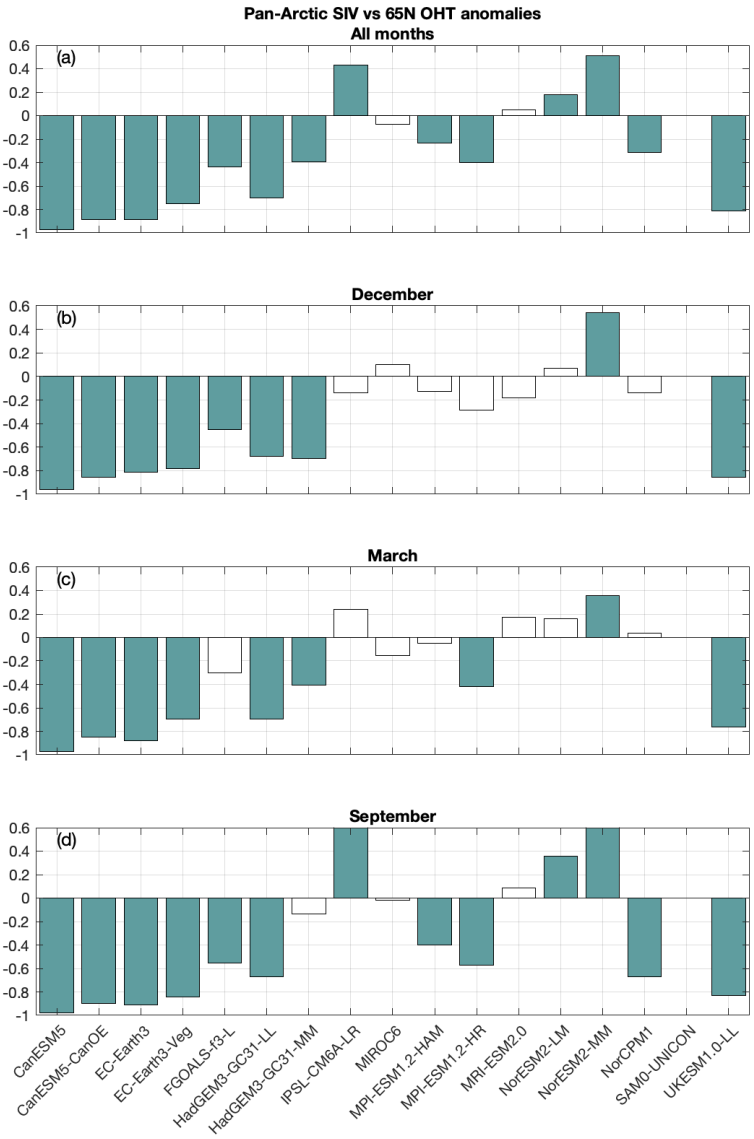
Correlation of OHT anomaly at 65°N against pan-Arctic sea ice area (SIA) anomaly for CMIP6 models for the period 1979-2014. Significant correlations ($p<0.01$) are indicated by filled bars. No OHT anomaly is calculated for SAM0-UNICON because no value for Atlantic OHT is specified.

Figure 42. Correlation of oceanic heat transport anomaly against pan-Arctic sea ice area anomaly at 65°N for the period 1979-2014



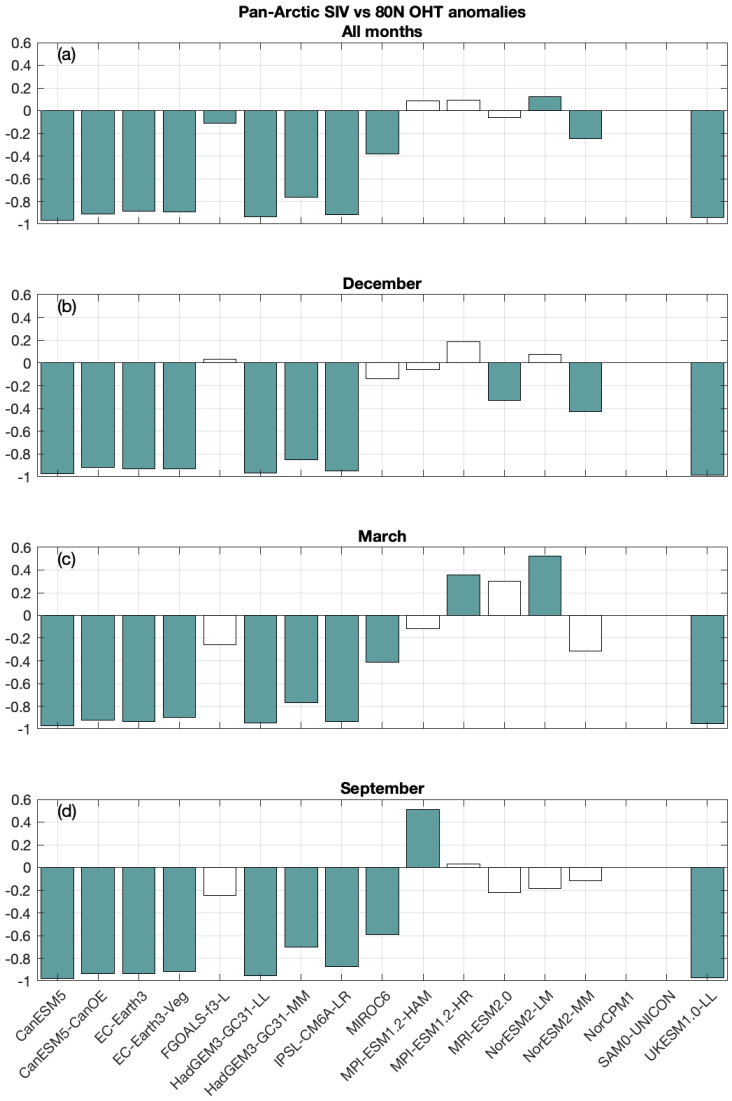
Correlation of OHT anomaly at 80°N against pan-Arctic sea ice area (SIA) anomaly for CMIP6 models for the period 1979-2014. Significant correlations ($p<0.01$) are indicated by filled bars. No OHT anomaly is calculated for NorCPM1 and SAM0-UNICON because no value for Atlantic OHT is specified.

Figure 43. Correlation of oceanic heat transport anomaly against pan-Arctic sea ice area anomaly at 80°N for the period 1979-2014



As in Figure 42 but for SIV anomaly.

Figure 44. Correlation of oceanic heat transport anomaly against pan-Arctic sea ice volume anomaly at 65°N for the period 1979-2014



As in Figure 43 but for SIV anomaly.

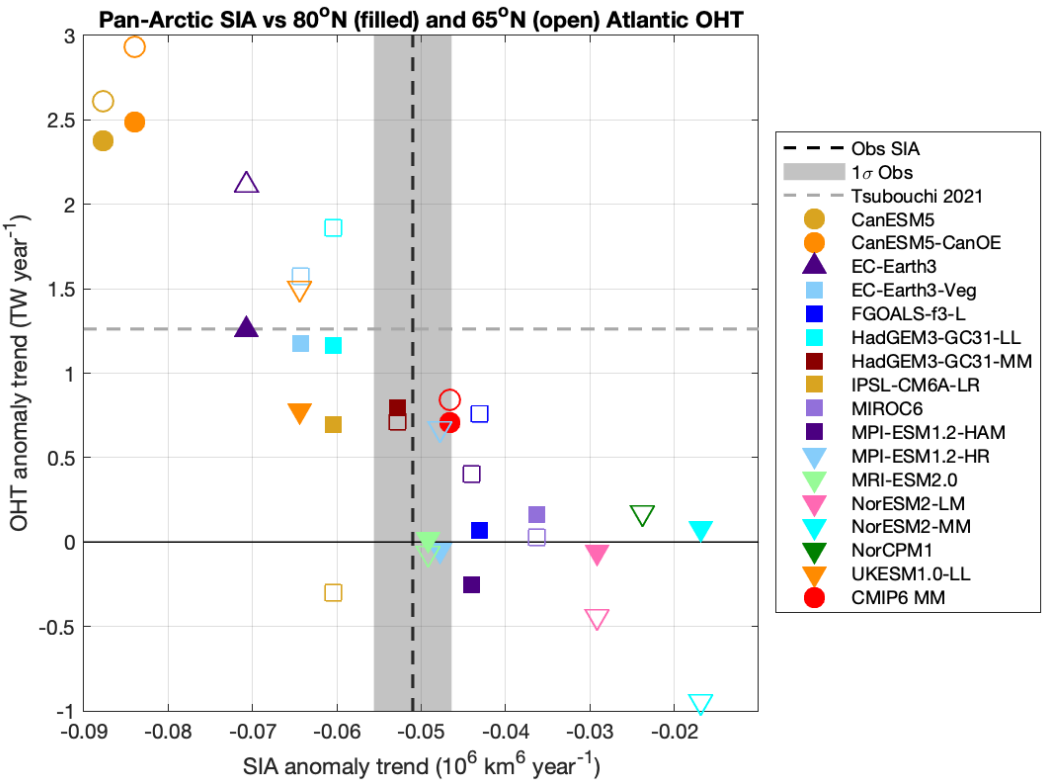
Figure 45. Correlation of oceanic heat transport anomaly against pan-Arctic sea ice volume anomaly at 80°N for the period 1979-2014

Correlation between pan-Arctic SIA and Atlantic OHT the CMIP6 simulations is generally stronger at 80°N (Figure 43) than for 65°N (Figure 42). For the full time series, 11 CMIP6 models show statistically significant anticorrelation between OHT and SIA anomalies at 80°N, meaning higher values of OHT have smaller values of SIA, and vice versa (Figure 43). The eight models identified earlier showing the greatest change in OHT anomaly (Figure 41b) are also identified here, with strongest anticorrelation across all

months examined (Figure 43). Differences between OHT correlation against pan-Arctic SIA at different latitudes is mostly in magnitude, but some model simulations do suggest opposing physical processes. For example, IPSL-CM6A-LR shows significant positive correlation for OHT at 65°N, i.e., decreasing heat transport and decreasing SIA anomaly, which is opposite to the OHT at 80°N results. In this case, the positive correlation is likely the result of the decreasing trends in Atlantic sub-polar OHTC (Figure 41c) and OHT at 65°N.

Correlation between CMIP6 model simulations of pan-Arctic SIV against OHT are very similar to the SIA results. Most models show stronger correlation against SIV for OHT at 80°N than for OHT at 65°N (Figure 45 and Figure 44, respectively). As was the case for SIA, IPSL-CM6A-LR shows some months of positive correlation between SIV and OHT at 65°N, as do the two NorESM2 models. Additionally, no meaningful lag correlation between monthly mean OHT and SIA or SIV is evident in the CMIP6 models (not shown).

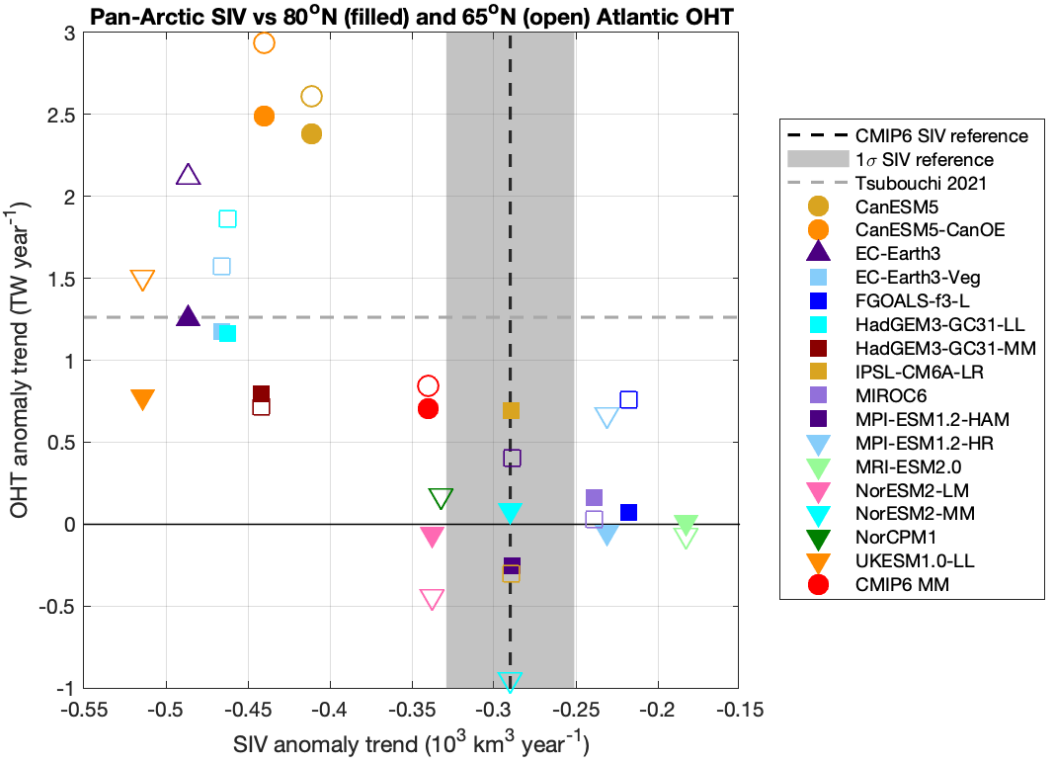
Next, we examine the relationship among CMIP6 simulations for OHT and pan-Arctic SIA anomaly trends for the period 1979-2014 in Figure 46. A strong linear negative relationship exists between model simulations with large positive OHT trends (e.g., both CanESM5 models) and their negative trends of SIA. This can be interpreted as showing that CMIP6 model simulations of sea ice are sensitive to the trend in OHT. The CMIP6 models are mostly divided between SIA simulations that show either stronger or weaker trends against the observed annual trend of -51×10^3 km^2 per year from the combined passive microwave observations (Figure 46). The CMIP6 SIA MM trend in this model subset is about 6% slower (-47×10^3 km^2 per year) than the observed trend, and its OHT MM trend ranges between 0.7 to 0.9 TW per year (at 80°N and 65°N, respectively). The CMIP6 models with the slowest SIA trends show little to no trend in OHT anomaly at 80°N (FGOALS-f3-L, MIROC6, and both NorESM2 models), and the OHT anomaly trend at 65°N for these models shows larger spread (Figure 46). A recent estimate of the OHT into the Arctic Mediterranean (i.e., similar to our OHT at 65°N) from Tsubouchi et al. (2021) for 1993-2016 reveals a trend near 1.3 TW per year (or about 45% larger than the CMIP6 OHT MM at 65°N).



Annual pan-Arctic SIA anomaly trends (in x-axis) and OHT anomaly trends (in y-axis) for the period 1979 to 2014 for CMIP6 models. OHT anomaly trends for 65°N and 80°N are indicated by open and filled symbols, respectively. The solid black line and gray shading indicates the passive microwave combined observations SIA decline (-51×10^3 km^2 year^{-1}) and standard deviation (66% uncertainty). The dashed gray line indicates the estimated OHT anomaly trend (1.3 TW year^{-1}) for the Arctic Mediterranean (Tsubouchi et al. 2021). No 80°N OHT anomaly trend for NorCPM1 is shown because no value for Atlantic OHT is specified.

Figure 46. Trends comparison between oceanic heat transport and sea ice area for the period 1979-2014

Similar analysis for OHT and pan-Arctic SIV is shown in Figure 47. CMIP6 models are again divided between a grouping with faster SIV decline than the CMIP6 SIV MM reference trend of –290 km^3 per year (from Chapter 3), and larger trends in OHT, and an opposite grouping. This 16-model CMIP6 subset has a SIV MM trend of –340 km^3 per year, which is about 17% faster than the reference from all CMIP6 models. Individual CMIP6 SIV trends generally share the same relationship between OHT as was shown for SIA (Figure 46). That is, models with larger trends in OHT tend to have larger negative trends in SIV.



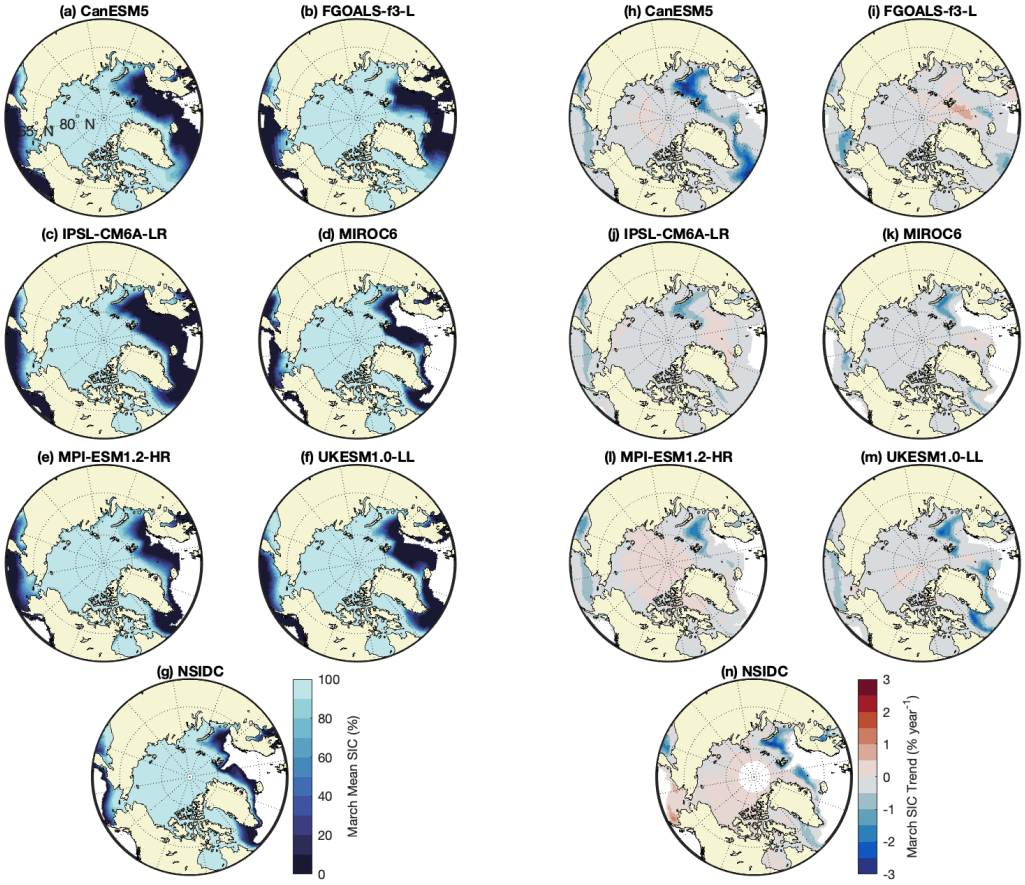
Annual pan-Arctic SIV anomaly trends (in x-axis) and OHT anomaly trends (in y-axis) for the period 1979 to 2014 for CMIP6 models. OHT anomaly trends for 65°N and 80°N are indicated by open and filled symbols, respectively. The solid black line and gray shading indicates the CMIP6 MM SIV decline (-290 km3 year-1, from previous chapter) and standard deviation (66% uncertainty). The dashed gray line indicates the estimated OHT anomaly trend (1.3 TW year-1) for the Arctic Mediterranean (Tsubouchi et al. 2021). No 80°N OHT anomaly trend for NorCPM1 is shown because no value for Atlantic OHT is specified.

Figure 47. Trends comparison between oceanic heat transport and sea ice volume for the period 1979-2014

3. Spatial Analysis of OHT Forcing on Pan-Arctic Sea Ice

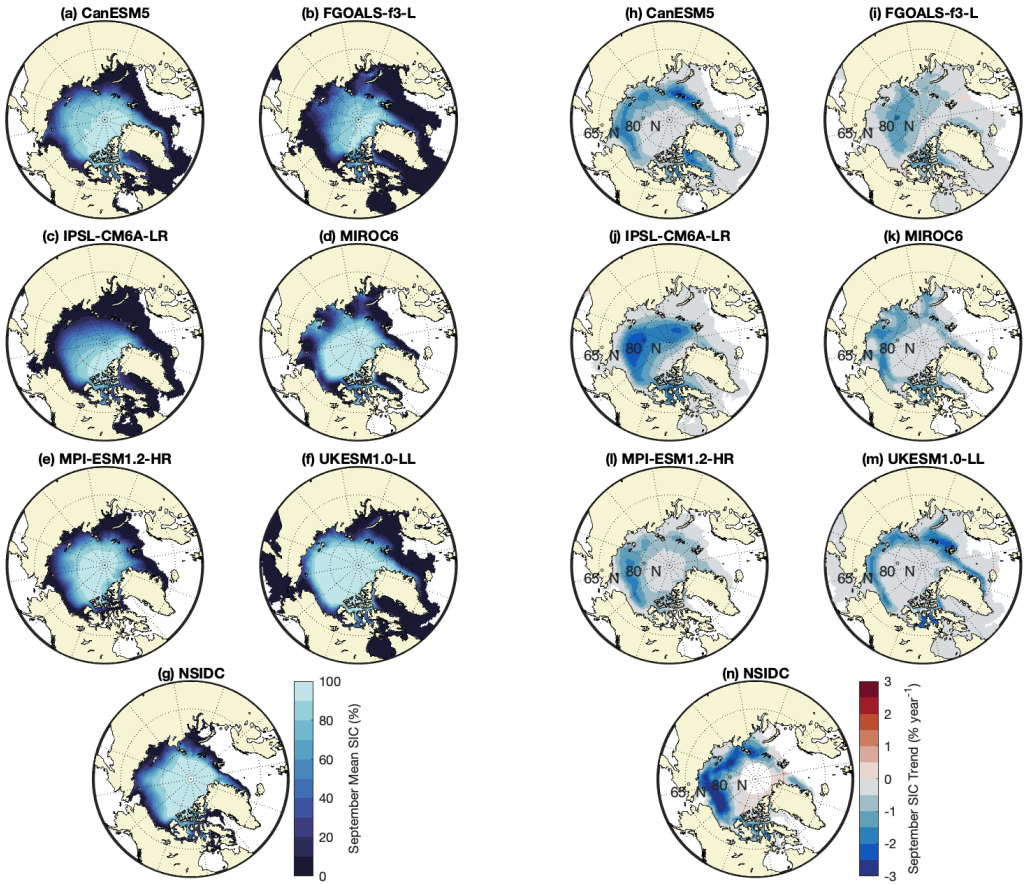
The 1979-2014 mean SIC and anomaly linear trend for selected CMIP6 models are shown for the months of March and September in Figure 48 and Figure 49, respectively. These models are evenly divided between larger and smaller trends than the observed pan-Arctic SIA anomaly trend of -510×10^3 km^2 per year range (Figure 46). They range from -320×10^3 km^2 per year (MIROC6) to -820×10^3 km^2 per year (CanESM5). In March, the Barents Sea region shows the largest negative SIC trends for all models except FGOALS-f3-L (Figure 48). CanESM5, MPI-ESM1.2-HR, and UKESM1.0-LL show the largest regional trends (Figure 48g,k,l) and also indicate large positive OHT trends at 65°N ranging

from 0.6 to 2.6 TW per year (Figure 46). In September, there is more variation among CMIP6 models as to the region with largest negative SIC trends. However, CanESM5 and UKESM1.0-LL clearly show strong SIC decline along the MIZ near the Barents and Nordic seas (Figure 49g,l), and both have substantial positive OHT trends at 80°N ranging from 0.7 to 2.4 TW per year (Figure 46). On the other hand, FGOALS-f3-L and IPSL-CM6A-LR show a more widely distributed SIC decline over much of the Arctic (Figure 49h,i), and show positive and negative OHT trends at 80°N, respectively.



March mean SIC for (a-f) CMIP6 models and (g) the NSIDC passive microwave observations for the period 1979-2014. March SIC annual trend for (h-m) CMIP6 models and (n) observations for the period 1979-2014.

Figure 48. March mean sea ice concentration and linear trends in CMIP6 models and observations for the period 1979-2014

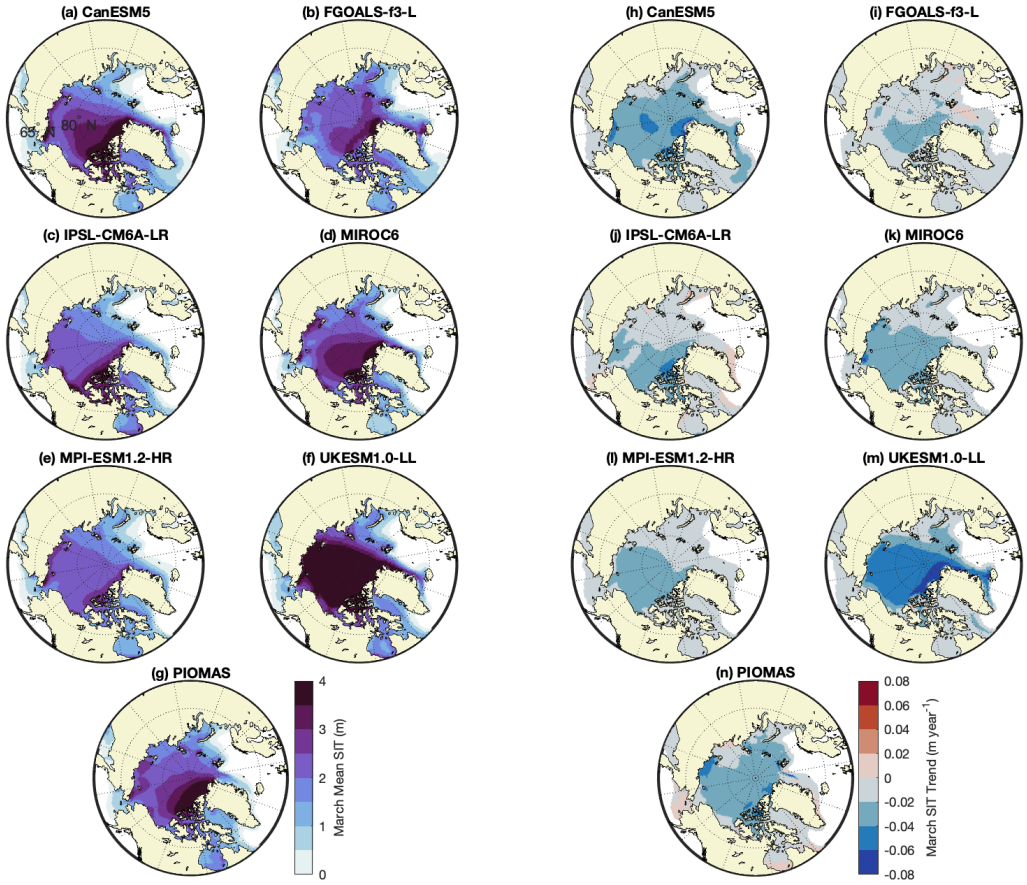


September mean SIC for (a-f) CMIP6 models and (g) the NSIDC passive microwave observations for the period 1979-2014. September SIC annual trend for (h-m) CMIP6 models and (n) observations for the period 1979-2014.

Figure 49. September mean sea ice concentration and linear trends in CMIP6 models and observations for the period 1979-2014

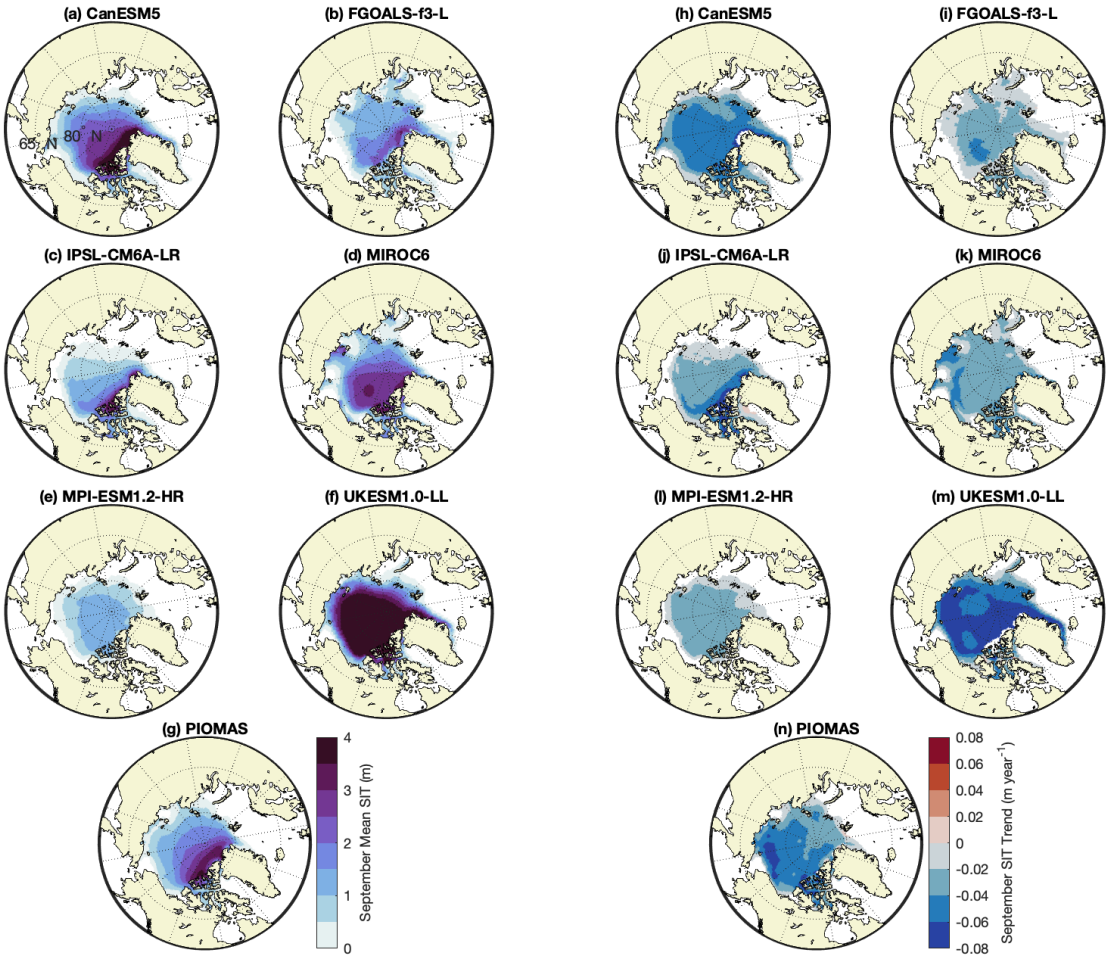
In Figure 50 and Figure 51, the same models selected for SIC spatial analysis are shown here for the 1979-2014 mean SIT and anomaly linear trend for March and September, respectively. As discussed previously in Chapter IV, CMIP6 models show a wide range in representing the observed and estimated spatial distribution of SIT. The CMIP6 models have SIV anomaly trend ranges from -0.52×10^3 km^3 per year (UKESM1.0-LL) to -0.20×10^3 km^3 per year (FGOALS-f3-L), compared against the 42-model CMIP6 SIV MM trend of -0.29×10^3 km^3 per year (Figure 47). In March, the regions of largest SIT decline in the CMIP6 models are largely coincident with their regions of thickest sea ice (Figure 50). This is most notable for CanESM5 and UKESM1.0-LL, which also have the

largest rates of SIV decline as well as strong positive Atlantic OHT trends at 65°N and 80°N (Figure 50h,m). In September, unlike the SIA decline along the MIZ, the SIT decline is almost ubiquitous for all regions and for all CMIP6 models (Figure 51).



March mean SIT for (a-f) CMIP6 models and (g) PIOMAS for the period 1979-2014. March SIT difference (2014 minus 1979) for (h-m) CMIP6 models and (n) PIOMAS.

Figure 50. March mean sea ice thickness and linear trends in CMIP6 models and PIOMAS for the period 1979-2014



September mean SIT for (a-f) CMIP6 models and (g) PIOMAS for the period 1979-2014. September SIT difference (2014 minus 1979) for (h-m) CMIP6 models and (n) PIOMAS.

Figure 51. September mean sea ice thickness and linear trends in CMIP6 models and PIOMAS for the period 1979-2014

D. DISCUSSION AND CONCLUSION

Although largely unaddressed in this study, atmospheric forcing is manifest in a number of processes driving the simulated Arctic sea ice decline (e.g., Liang et al. 2020). These likely include changes in large-scale atmospheric circulation (e.g., North Atlantic Oscillation, variations in the polar vortex), warming SATs caused by increased GHG emissions, and sea surface warming resulting from reduced sea ice cover (i.e., ice-albedo feedback). Because all CMIP6 models simulate negative SIA and SIV trends (Figure 46 and Figure 47) regardless of trends in the simulated OHT, the influence of atmospheric

forcing in CMIP6 models is clear. However, the majority of models with little to no Atlantic OHT trend at 80°N underestimate the rate of observed SIA decline (by up to 65% in NorESM2-MM). Thus, atmospheric forcing alone is likely insufficient to account for the observed sea ice decline (e.g., Deser and Teng 2008).

Assuming the atmospheric forcing in the CMIP6 models similarly accounts only for a part of sea ice variability, then the oceanic forcing must account for the remaining external forcing of sea ice decline. A number of studies indicate increased OHT into the Arctic region, particularly AW through the Fram Strait and Barents Sea, which has impacts on the sea ice cover there (Polyakov et al. 2017; Tsubouchi et al. 2021). In contrast to the CMIP6 models which underestimate the SIA decline, those which exceed it exhibit strong positive OHT trends at 80°N (Figure 46). Thus, CMIP6 pan-Arctic sea ice simulations appear to be sensitive to OHT, but quantifying sensitivity is beyond the scope of this study. This also suggests that accurate delivery of ocean heat into the Arctic may be a key process for better representation of a) the 1979-2014 SIA trend, and b) the observed accelerated rate of sea ice decline (Comiso et al. 2008; Serreze and Stroeve 2015).

Determining a causal link between OHT and sea ice decline for the summer months (e.g., September) is complicated due to the atmospheric forcing discussed earlier. However, the winter sea ice decline is more straightforward. During the polar night, SATs are well below the sea water freezing temperature and there is no incoming shortwave radiation (e.g., December at 80°N), thus the OHT impacts on sea ice may be relatively pronounced in December.

Some limitations in our OHT analysis methods are worth discussing here. First, OHT is determined by both the volume transport and ocean temperature, but our methods did not permit analysis of these two variables separately. Such analysis could be a useful intercomparison for modeling centers to understand biases and disentangle systematic limitations. Second, while our analysis clearly illustrates the impact of increased OHT trends on pan-Arctic sea ice, particularly at 80°N, we were unable to differentiate the path of Atlantic OHT (i.e., West Spitsbergen Current or Barents Sea Opening). Regional OHT analysis better suited for differentiating these pathways is needed to diagnose model performance relative to observations.

A small number of CMIP6 models examined here in Chapter V are included in the Chapter IV spatial analysis and warrant qualitative discussion. Recall that Figure 35f,l show the March mean spatial Brier Scores for FGOALS-f3-L and UKESM1.0-LL. For FGOALS-f3-L, an increased trend in OHT at 80°N is not simulated well (Figure 41) and the position of the Bering and Greenland seas and Baffin Bay ice edge is poorly represented (Figure 35f). On the other hand, UKESM1.0-LL simulates an increased trend in OHT at 80°N and a much better overall representation of the sea ice edge (Figure 35l). Of the models analyzed here, the CanESM5 simulates the largest OHT trend at 80°N and well represents the March ice edge in Barents Sea (Figure 35l). However, the accompanying large Brier Score in the Greenland Sea and Baffin Bay may be an example of a poorly represented distribution of heat (e.g., bias towards BSO pathway). While not a conclusive finding, this does support our hypothesis that simulations of OHT may have a substantial role in representation of the sea ice.

In summary, compared to best estimates and recent reanalyses, the CMIP6 MM likely underestimates the northward OHT into the Arctic, particularly during the months July-November. About half of the CMIP6 models analyzed here have an increasing trend in Atlantic OHT into the pan-Arctic region after about 2000. This is consistent with recently published observational studies showing a similar increase (e.g., Tsubouchi et al. 2021). A strong linear negative relationship exists between those model simulations with large positive Atlantic OHT trends and their negative SIA trends. Additionally, those same models have strong correlation between Atlantic OHT anomalies and pan-Arctic SIA anomalies. While not a conclusive finding, this does suggest that model simulations of OHT are sensitive to the trend in OHT and may have a substantial role in representation of the sea ice.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. CONCLUSION

The research conducted as part of this dissertation focused on (a) improving the United States Navy's understanding of the changing Arctic environment under the effects of climate change, (b) identifying progress and limitations in global climate model simulations of Arctic sea ice, and (c) suggesting areas for future study and potential model improvements. At the time of research preparation, CMIP6 data were freshly filtering into the ESGF repository from an increasing number of modeling centers and no CMIP6 intercomparison studies had yet been published. As such, the primary direction for this research was clear, contribute to the base CMIP6 scientific knowledge through model intercomparison.

Since the observed climate is but one realization in nature, and given the countless physical interactions occurring in the environment, it is unreasonable to expect a 'perfect' representation of past observations (Notz 2015). In the case of sea ice simulations, Notz (2015) attributes the discrepancies against observations with model and climate internal variability, measured uncertainty in observations, and model tuning. While model skill in simulating the past may not translate directly into skillful projections, it is hard to argue that a model showing little skill against the past will perform considerably better for future projections (Randall et al. 2007; Massonnet et al. 2012). Thus, model intercomparison against available observations is a necessary first step to identify higher quality simulations and uncover biases that may compound future uncertainty in climate projections.

A. SUMMARY OF FINDINGS

Provided here is a summary of conclusions from the three results chapters. While sea ice simulations in CMIP6 show a wide range in representing the historical state of the Arctic sea ice, they do share a number of characteristics common to observations. Nearly all CMIP6 models have SIE and SIA mean annual cycles with September minimum and March maximum. They all have declining trends in SIE and SIA over the period of satellite observations, 1979-2014. The mean annual cycle for SIV is also well represented in CMIP6 with September minimum and April maximum. CMIP6 simulations of sea ice respond to

the increasing magnitude of cumulative CO_2 concentration in the atmosphere and to global warming (global mean SAT) by losing sea ice. Thus, models generally simulate the observed multidecadal decline of sea ice in a warming climate.

1. CMIP6 Multimodel Mean Time Series

Collectively, the CMIP6 MM tends to outperform individual CMIP6 ensembles and closely matches the existing sea ice observations. Examining the CMIP6 MM provides the best estimate of the average sea ice response to climate forcing for the period 1979-2014. Because individual models have structural biases and varied responses to different climate forcing, averaging therefore reduces signal variability and gives confidence that the forced signal is represented. Additionally, analysis of the MM spread (i.e., standard deviation) provides estimates for the internal climate variability in response to forcing. The CMIP6 SIA MM rate of decline, -0.47×10^6 km^2 decade^{-1}, underestimates the observed trend (-0.53×10^6 km^2 decade^{-1}) only by about 10%. And the 12-month running mean CMIP6 MM SIA has approximately 0.5×10^6 km^2 positive bias relative to the observed SIA, or about 5% relative to the total ice-covered area.

The CMIP6 SIA MM has an accelerated rate of decline between periods P1 and P2 that is about 25% larger than the observations. However, the magnitude of SIA MM decline in periods P1 and P2 are 33% and 16% slower than observations, respectively. The observed accelerated rate of SIA decline is predominantly associated with the strong negative trend in September SIA. During P1, the CMIP6 SIA MM simulates a similar September SIA rate of decline, but underestimates the P2 rate (by around 25%) and consequently the accelerated September SIA decline. On the other hand, the March CMIP6 SIA MM has an accelerated rate of decline that is not shown in SIA observations. Thus, the CMIP6 SIA MM dampens the observed seasonality in accelerated SIA decline, i.e., overestimates the March and underestimates the September P2 SIA trends.

The CMIP6 SIV MM mean of 21.4×10^3 km^3 and trend of -2.88×10^3 km^3 decade^{-1} for 1979-2014 pairs well with the SIV reanalyses. For example, relative to PIOMAS the SIV MM mean is about 6% overestimated and the trend is about 5% underestimated. The CMIP6 SIV MM has an accelerated rate of decline between periods P1 and P2. Compared

against PIOMAS SIV reanalysis, the SIV MM has a 10% stronger declining trend in P1 and a 25% weaker trend in P2. The CMIP6 SIV MM has a mostly constant decline in SIV across all months and both periods P1 and P2. In contrast, the PIOMAS reanalysis indicates a substantial seasonal signal in the September SIV decline during P2. This discrepancy may be caused by PIOMAS underestimation bias of thicker sea ice, particularly during the later P2 period, and by errors in data assimilation of SIC around the MIZ (Zygmuntowska et al. 2013).

2. CMIP6 Ensembles Time Series

Individual sea ice simulations for CMIP6 models have both positive and negative biases against observations, with several models exceeding ±2σ of the observed mean. These biases can be attributed to several causes, e.g., the model structure or parameter tuning, internal variability, sensitivity to external forcing, and lack of or limitation in the representation of some physical processes. For example, a bias in the surface energy budget may result in simulations of too thick or too thin sea ice. Such a model may nevertheless respond to external forcing consistent with observations and/or display comparable sea ice variability, and thus may still be a suitable representation for climate analysis. Also, large spread among CMIP6 models and across ensemble members within a model indicate large uncertainty due to internal variability.

The majority of CMIP6 models underestimate the 1979-2014 SIA trend, and those which overestimate it usually have positive SIA bias. About 25% of CMIP6 models show trends within ±2σ of the observed trend, and about 60% are within overlapping uncertainty estimates for the simulated and observed values, indicating they are from the same population distribution. About 70% of CMIP6 models have mean SIA within ±2σ of the observed SIA mean. SIA anomaly variability is lower than the observations in more than 60% of the CMIP6 models, which indicates less extreme SIA interannual variability. Most models simulate a somewhat faster rate of decline in SIA starting in the mid-1990s (86%), i.e., accelerating decline, with about half of the models at least doubling their period P2 SIA trends relative to those in P1.

For SIV, more models show trends that exceed the reference (i.e., CMIP6 MM for SIV) and PIOMAS than for SIA, albeit with larger model and observed trend uncertainty. About half of the CMIP6 models simulate mean SIV within ±2σ of the CMIP6 SIV MM mean, and about 30% are within ±2σ of the SIV MM trend. About 70% of CMIP6 models have a slower P2 rate of decline and slower accelerated trend than PIOMAS. Given that SIV is not well constrained in sea ice simulations and the large uncertainty in SIV reanalyses, we offer that CMIP6 SIV MM trends and acceleration rate for the time being may be a more suitable reference estimate.

3. Spatial Analysis of Sea Ice Edge and Thickness

The integrated measures of SIA and SIV time series provide relatively straightforward comparison among the CMIP6 sea ice simulations. However, they provide only a limited comparison and require a bit of caution. For example, SIA can mask underlying inaccuracies in the simulated sea ice due to cancellation errors in regions of positive and negative bias. Therefore, a simulation may compare favorably to observations, but not consistent with the physical processes under examination. In addition, knowledge of model bias, e.g., northern hemisphere SIA, but not of the region or process responsible may be of limited value to the individual modeling center and unactionable. Thus, comparison of simulated sea ice time series is not enough to diagnose regional and seasonal biases.

Examining the spatial distribution of SIT and the sea ice edge allows for a more thorough analysis of the sea ice state and change. This also lends itself to discerning potential causes of systematic sea ice biases in particular models. However, the tools and interpretation for such model comparison are not as well defined as for the time series. Narrowing the number of CMIP6 models analyzed to 12 was practically necessary to demonstrate spatial analysis with a workable subset. In particular, the ice edge analysis was proposed as a useful new metric to identify regional biases and guide further analyses, which should lead to improvements of common or individual model physics limitations in future CMIP simulations.

CMIP6 models have a wide range of SIT distribution patterns, as well as biases in ice edge representation. For SIT, the basic pattern of greatest SIT along the Canadian Archipelago and north of Greenland in March is mostly represented by the CMIP6 models. However, several models have significant pattern differences, likely associated with poorly represented general atmospheric circulation patterns, e.g., Beaufort High (Stroeve et al. 2014). Furthermore, the quantitative spatial pattern correlation for March does not clearly distinguish models with poor SIT patterns or lend itself to regional diagnostics. In contrast, the September SIT patterns may offer more meaningful differences between models, but have the issue of data sparse SIT observations in summer.

Because CMIP6 models correctly simulate the fully ice-covered Arctic interior in winter, the majority of ice edge errors during the colder months are in vicinity of the sub-Arctic marginal seas (e.g., Bering Sea, Nordic Seas, Barents Sea, Baffin Bay). Overestimations of the ice edge across the Greenland and Barents seas, and underestimations for the Bering Sea are dominant during this period. The months of greatest ice edge error variability among Arctic sub-regions occurs from June through November. Examination of the region and timing (i.e., months) of over- and underestimations of the ice edge points to potential limitations in or lack of representation of some physical processes within individual CMIP6 models. Collectively, the peak regional ice edge errors evaluated over these months are overestimation of the Barents Sea and Baffin Bay for nine of 12 CMIP6 models, and underestimation over the Kara Sea (5 models), and East Siberian Sea and Baffin Bay (each with 2 models).

4. Oceanic Heat Transport and Convergence

One such process that could contribute to discrepancies between the simulated and observed accelerated rate of decline in SIA, as well as the large ice edge errors in certain Arctic sub-regions, is the representation of warm ocean currents and their impact on the sea ice edge and thickness. In the majority of CMIP6 models, northward OHT in the polar cap region is likely underestimated when compared against reanalyses, but the seasonal cycles appear well represented. The CMIP6 OHT convergence (OHTC) MM shows about 96% of the total OHT into the pan-Arctic enters via the Atlantic pathway, compared against

the Pacific. There is strong correlation between increasing OHT and decreasing sea ice trends in CMIP6 models. And several models simulate an increased magnitude of Atlantic OHT into the pan-Arctic region after around 2000, which is consistent with recently published observations and reanalysis studies (Mayer et al. 2016; Tsubouchi et al. 2021).

B. RECOMMENDATIONS FOR FUTURE WORK

Continued research is needed on the heels of the accelerated rate analysis for SIA and SIV from this dissertation. A study investigating the future projections of SIA and SIV against more recent observations, i.e., 2015-21, would be interesting to examine whether past performance indeed reflects future model skill. Additionally, would sub-selecting models result in significant differences in future projections, e.g., year of seasonally ice-free Arctic? While the process and utility of sub selecting models is open for debate, the identification of plausible model simulations could improve projection uncertainties. Advocates of sub selection will point out grossly biased model simulations, e.g., NorCPM1 lacking 'skill' in SIV and providing questionable value to the MM mean, could be set aside. However, in the other extreme, too strict a criterion reduces the likelihood of the MMensemble spread representing internal climate variability.

Future studies are required to investigate processes responsible for the accelerated rate in sea ice decline in CMIP6 simulations. This dissertation identified the OHT as one potential driver, but a more thorough examination of the Arctic Ocean energy budget is needed. In particular, does the atmospheric surface flux or the meridional oceanic flux dominate in models which have accelerated rates of sea ice decline?

One effort to improve underlying physical processes of climate models is the application of increased temporal and spatial resolutions. As computational resources become ever more abundant, modeling centers and scientists turn towards higher resolution simulations to resolve and link connections between climate model components. Yet, higher resolution alone does not necessarily result in improved simulations, rather it provides a path forward for improvement, but still requiring evaluation of parameterizations and fine tuning. Nevertheless, given the example of RASM-G, higher model resolution is critical for improving not only the Arctic sea ice but also global climate

forecasts because of more realistic representation of the Arctic environments: e.g., ocean bathymetry, land topography, coastal geometry, coastal and shelf circulation, mesoscale eddies, sea ice deformations, polar lows and cyclones, clouds and atmospheric convection. The role of enhanced horizontal resolution in process level improvement for CMIP6 models is under active investigation in High Resolution Model Intercomparison Project HghResMIP v1.0 for CMIP6 (Haarsma et al. 2016).

Finally, the collection and availability of more observational data to evaluate and constrain model simulations of the environment is critical. For instance, a longer period of overlap among historical simulations and observations, e.g., SIT and SIA, will decrease uncertainty in the current sea ice state and trends. In future CMIP historical simulations, the observational overlap with passive microwave data will be extended, but also the scant Arctic SIT observations will also be increased following the launch of ICESat-2 in 2018 and joining CryoSat-2 (in orbit since 2010). Ice-Tethered Profilers (ITP) provide hydrographic observation data under the ice since 2004. Comparing CMIP models against ITPs located in purposefully sampled regions, e.g., near Pacific and Atlantic water inflows, may further illuminate model limitations and biases in the simulations of oceanic forcing on sea ice. Additionally, persistent hydrographic observations at the Arctic gateways are important to quantify changes in volume and heat transport and convergence.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

AMAP, 2007: Surface ocean currents in the Arctic. https://www.amap.no/documents/doc/surface-ocean-currents-in-the-arctic/566 (Accessed June 21, 2021).

Årthun, M., T. Eldevik, L. H. Smedsrud, Ø. Skagseth, and R. B. Ingvaldsen, 2012: Quantifying the influence of Atlantic heat on Barents Sea ice variability and retreat. *Journal of Climate*, **25**, 4736–4743, https://doi.org/10.1175/JCLI-D-11-00466.1.

Ballinger, T. J., and Coauthors, 2020: Arctic Report Card 2020: Surface Air Temperature. https://doi.org/10.25923/GCW8-2Z06.

Bokhorst, S., and Coauthors, 2016: Changing Arctic snow cover: A review of recent developments and assessment of future needs for observations, modelling, and impacts. *Ambio*, **45**, 516–537, https://doi.org/10.1007/s13280-016-0770-0.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, **78**, 1–3, https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2.

Burgard, C., and D. Notz, 2017: Drivers of Arctic Ocean warming in CMIP5 models. *Geophysical Research Letters*, **44**, 4263–4271, https://doi.org/10.1002/2016GL072342.

Cai, L., V. A. Alexeev, and J. E. Walsh, 2020: Arctic sea ice growth in response to synoptic- and large-scale atmospheric forcing from CMIP5 models. *J. Climate*, **33**, 6083–6099, https://doi.org/10.1175/JCLI-D-19-0326.1.

Cassano, J. J., and Coauthors, 2017: Development of the Regional Arctic System Model (RASM): Near-surface atmospheric climate sensitivity. *J. Climate*, **30**, 5729–5753, https://doi.org/10.1175/JCLI-D-15-0775.1.

Cavalieri, D. J., P. Gloersen, and W. J. Campbell, 1984: Determination of sea ice parameters with the NIMBUS 7 SMMR. *Journal of Geophysical Research: Atmospheres*, **89**, 5355–5369, https://doi.org/10.1029/JD089iD04p05355.

Cohen, J., and Coauthors, 2020: Divergent consensuses on Arctic amplification influence on midlatitude severe winter weather. *Nature Climate Change*, **10**, 20–29, https://doi.org/10.1038/s41558-019-0662-y.

Comiso, J. C., 1986: Characteristics of Arctic winter sea ice from satellite multispectral microwave observations. *Journal of Geophysical Research: Oceans*, **91**, 975–994, https://doi.org/10.1029/JC091iC01p00975.

Comiso, J. C., C. L. Parkinson, R. Gersten, and L. Stock, 2008: Accelerated decline in the Arctic sea ice cover. *Geophys. Res. Lett.*, **35**, L01703, https://doi.org/10.1029/2007GL031972.

——, W. N. Meier, and R. Gersten, 2017: Variability and trends in the Arctic sea ice cover: Results from different techniques. *Journal of Geophysical Research: Oceans*, **122**, 6883–6900, https://doi.org/10.1002/2017JC012768.

Curry, B., C. M. Lee, B. Petrie, R. E. Moritz, and R. Kwok, 2014: Multiyear volume, liquid freshwater, and sea ice transports through Davis Strait, 2004–10. *Journal of Physical Oceanography*, **44**, 1244–1266, https://doi.org/10.1175/JPO-D-13-0177.1.

Czaja, A., and J. Marshall, 2006: The Partitioning of Poleward Heat Transport between the Atmosphere and Ocean. *Journal of the Atmospheric Sciences*, **63**, 1498–1511, https://doi.org/10.1175/JAS3695.1.

Danabasoglu, G., and Coauthors, 2020: The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, **12**, e2019MS001916, https://doi.org/10.1029/2019MS001916.

Davy, R., and I. Esau, 2014: Global climate models' bias in surface temperature trends and variability. *Environ. Res. Lett.*, **9**, 114024, https://doi.org/10.1088/1748-9326/9/11/114024.

——, and S. Outten, 2020: The Arctic surface climate in CMIP6: status and developments since CMIP5. *J. Climate*, https://doi.org/10.1175/JCLI-D-19-0990.1.

Department of Defense, 2019: 2019 Department of Defense Arctic Strategy. https://media.defense.gov/2019/Jun/06/2002141657/-1/-1/1/2019-DOD-ARCTIC-STRATEGY.PDF

Department of the Navy, 2021: A Strategic Blueprint for the Arctic. https://www.globalsecurity.org/military/library/policy/navy/arctic-blueprint-2021-final.pdf

DeRepentigny, P., L. B. Tremblay, R. Newton, and S. Pfirman, 2016: Patterns of sea ice retreat in the transition to a seasonally ice-free Arctic. *Journal of Climate*, **29**, 6993–7008, https://doi.org/10.1175/JCLI-D-15-0733.1.

——, A. Jahn, M. M. Holland, and A. Smith, 2020: Arctic Sea Ice in Two Configurations of the CESM2 During the 20th and 21st Centuries. *Journal of Geophysical Research: Oceans*, **125**, e2020JC016133, https://doi.org/10.1029/2020JC016133.

Deser, C., and H. Teng, 2008: Evolution of Arctic sea ice concentration trends and the role of atmospheric circulation forcing, 1979–2007. *Geophysical Research Letters*, **35**, https://doi.org/10.1029/2007GL032023.

Eisenman, I., W. N. Meier, and J. R. Norris, 2014: A spurious jump in the satellite record: Has Antarctic sea ice expansion been overestimated? *Cryosphere*, **8**, 1289–1296, https://doi.org/10.5194/tc-8-1289-2014.

England, M., A. Jahn, and L. Polvani, 2019: Nonuniform contribution of internal variability to recent Arctic sea ice loss. *J. Climate*, **32**, 4039–4053, https://doi.org/10.1175/JCLI-D-18-0864.1.

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**, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016.

Francis, J. A., and E. Hunter, 2007: Drivers of declining sea ice in the Arctic winter: A tale of two seas. *Geophysical Research Letters*, **34**, https://doi.org/10.1029/2007GL030995.

Garuba, O. A., H. A. Singh, E. Hunke, and P. J. Rasch, 2020: Disentangling the Coupled Atmosphere-Ocean-Ice Interactions Driving Arctic Sea Ice Response to CO2 Increases. *Journal of Advances in Modeling Earth Systems*, **12**, e2019MS001902, https://doi.org/10.1029/2019MS001902.

Goessling, H. F., and T. Jung, 2018: A probabilistic verification score for contours: Methodology and application to Arctic ice-edge forecasts. *Quarterly Journal of the Royal Meteorological Society*, **144**, 735–743, https://doi.org/10.1002/qj.3242.

——, S. Tietsche, J. J. Day, E. Hawkins, and T. Jung, 2016: Predictability of the Arctic sea ice edge. *Geophys. Res. Lett.*, **43**, 1642–1650, https://doi.org/10.1002/2015GL067232.

Golaz, J.-C., and Coauthors, 2019: The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution. *Journal of Advances in Modeling Earth Systems*, **11**, 2089–2129, https://doi.org/10.1029/2018MS001603.

Gregory, J. M., P. A. Stott, D. J. Cresswell, N. A. Rayner, C. Gordon, and D. M. H. Sexton, 2002: Recent and future changes in Arctic sea ice simulated by the HadCM3 AOGCM. *Geophysical Research Letters*, **29**, 28-1-28–4, https://doi.org/10.1029/2001GL014575.

Griffies, S. M., and Coauthors, 2016: OMIP contribution to CMIP6: experimental and diagnostic protocol for the physical component of the Ocean Model Intercomparison Project. *Geoscientific Model Development*, **9**, 3231–3296, https://doi.org/10.5194/gmd-9-3231-2016.

Haarsma, R. J., and Coauthors, 2016: High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. *Geoscientific Model Development*, **9**, 4185–4208.

Hamman, J., and Coauthors, 2016: Land surface climate in the Regional Arctic System Model. *J. Climate*, **29**, 6543–6562, https://doi.org/10.1175/JCLI-D-15-0415.1.

Held, I. M., and Coauthors, 2019: Structure and Performance of GFDL's CM4.0 Climate Model. *Journal of Advances in Modeling Earth Systems*, **11**, 3691–3727, https://doi.org/10.1029/2019MS001829.

Hendricks, S., and R. Ricker, 2019: Product user guide & algorithm specification: AWI CryoSat-2 sea ice thickness (version 2.2). *EPIC3*,. https://epic.awi.de/id/eprint/50033/ (Accessed September 18, 2020).

Im, U., and Coauthors, 2021: Present and future aerosol impacts on Arctic climate change in the GISS-E2.1 Earth system model. *Atmospheric Chemistry and Physics*, **21**, 10413–10438, https://doi.org/10.5194/acp-21-10413-2021.

IPCC, 2014: *Climate Change 2013 – The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]*. Cambridge University Press, 1535 pp.

——, 2019: *Special report: The ocean and cryosphere in a changing climate [H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.)]*. https://www.ipcc.ch/srocc/cite-report/ In press. (Accessed June 15, 2020).

Ivanova, N., O. M. Johannessen, L. T. Pedersen, and R. T. Tonboe, 2014: Retrieval of Arctic sea ice parameters by satellite passive microwave sensors: A comparison of eleven sea ice concentration algorithms. *IEEE Trans. Geosci. Remote Sensing*, **52**, 7233–7246, https://doi.org/10.1109/TGRS.2014.2310136.

Jackson, J. M., S. E. Allen, F. A. McLaughlin, R. A. Woodgate, and E. C. Carmack, 2011: Changes to the near-surface waters in the Canada Basin, Arctic Ocean from 1993–2009: A basin in transition. *Journal of Geophysical Research: Oceans*, **116**, https://doi.org/10.1029/2011JC007069.

Jahn, A., J. E. Kay, M. M. Holland, and D. M. Hall, 2016: How predictable is the timing of a summer ice-free Arctic? *Geophysical Research Letters*, **43**, 9113–9120, https://doi.org/10.1002/2016GL070067.

Kauffman, B. G., R. Jacob, T. Craig, and W. G. Large, 2004: The CCSM Coupler Version 6.0. https://www.cesm.ucar.edu/models/ccsm3.0/cpl6/users_guide/users_guide.html (Accessed July 24, 2021).

Kay, J. E., M. M. Holland, and A. Jahn, 2011: Inter-annual to multi-decadal Arctic sea ice extent trends in a warming world. *Geophysical Research Letters*, **38**, https://doi.org/10.1029/2011GL048008.

Kelley, M., and Coauthors, 2020: GISS-E2.1: Configurations and Climatology. *Journal of Advances in Modeling Earth Systems*, **12**, e2019MS002025, https://doi.org/10.1029/2019MS002025.

Kirtman, B. P., and Coauthors, 2012: Impact of ocean model resolution on CCSM climate simulations. *Clim Dyn*, **39**, 1303–1328, https://doi.org/10.1007/s00382-012-1500-3.

Kwok, R., 2018: Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018). *Environ. Res. Lett.*, **13**, 105005, https://doi.org/10.1088/1748-9326/aae3ec.

——, and D. A. Rothrock, 2009: Decline in Arctic sea ice thickness from submarine and ICESat records: 1958–2008. *Geophysical Research Letters*, **36**, https://doi.org/10.1029/2009GL039035.

——, G. Spreen, and S. Pang, 2013: Arctic sea ice circulation and drift speed: Decadal trends and ocean currents. *Journal of Geophysical Research: Oceans*, **118**, 2408–2425, https://doi.org/10.1002/jgrc.20191.

Labe, Z., G. Magnusdottir, and H. Stern, 2018: Variability of Arctic sea ice thickness using PIOMAS and the CESM Large Ensemble. *J. Climate*, **31**, 3233–3247, https://doi.org/10.1175/JCLI-D-17-0436.1.

Lavergne, T., and Coauthors, 2019: Version 2 of the EUMETSAT OSI SAF and ESA CCI sea-ice concentration climate data records. *The Cryosphere*, **13**, 49–78, https://doi.org/10.5194/tc-13-49-2019.

Liang, Y., H. Bi, Y. Wang, Z. Zhang, and H. Huang, 2020: Role of atmospheric factors in forcing Arctic sea ice variability. *Acta Oceanol. Sin.*, **39**, 60–72, https://doi.org/10.1007/s13131-020-1629-6.

Lind, S., R. B. Ingvaldsen, and T. Furevik, 2018: Arctic warming hotspot in the northern Barents Sea linked to declining sea-ice import. *Nature Climate Change*, **8**, 634–639, https://doi.org/10.1038/s41558-018-0205-y.

Lindsay, R., 2010: New unified sea ice thickness climate data record. *Eos, Transactions American Geophysical Union*, **91**, 405–406, https://doi.org/10.1029/2010EO440001.

Mahlstein, I., and R. Knutti, 2011: Ocean Heat Transport as a Cause for Model Uncertainty in Projected Arctic Warming. *Journal of Climate*, **24**, 1451–1460, https://doi.org/10.1175/2010JCLI3713.1.

Markus, T., J. C. Stroeve, and J. Miller, 2009: Recent changes in Arctic sea ice melt onset, freezeup, and melt season length. *Journal of Geophysical Research: Oceans*, **114**, https://doi.org/10.1029/2009JC005436.

Maslowski, W., J. Clement Kinney, M. Higgins, and A. Roberts, 2012: The future of Arctic sea ice. *Annu. Rev. Earth Planet. Sci.*, **40**, 625–654, https://doi.org/10.1146/annurev-earth-042711-105345.

——, ——, S. R. Okkonen, R. Osinski, A. F. Roberts, and W. J. Williams, 2014: The large scale ocean circulation and physical processes controlling Pacific-Arctic interactions. *The Pacific Arctic Region: Ecosystem Status and Trends in a Rapidly Changing Environment*, J.M. Grebmeier and W. Maslowski, Eds., Springer Netherlands, 101–132.

Massonnet, F., T. Fichefet, H. Goosse, C. M. Bitz, G. Philippon-Berthier, M. M. Holland, and P.-Y. Barriat, 2012: Constraining projections of summer Arctic sea ice. *The Cryosphere*, **6**, 1383–1394, https://doi.org/10.5194/tc-6-1383-2012.

Mayer, M., L. Haimberger, M. Pietschnig, and A. Storto, 2016: Facets of Arctic energy accumulation based on observations and reanalyses 2000–2015. *Geophysical Research Letters*, **43**, 10,420-10,429, https://doi.org/10.1002/2016GL070557.

——, S. Tietsche, L. Haimberger, T. Tsubouchi, J. Mayer, and H. Zuo, 2019: An improved estimate of the coupled Arctic energy budget. *J. Climate*, **32**, 7915–7934, https://doi.org/10.1175/JCLI-D-19-0233.1.

Meier, W. N., and J. S. Stewart, 2019: Assessing uncertainties in sea ice extent climate indicators. *Environ. Res. Lett.*, **14**, 035005, https://doi.org/10.1088/1748-9326/aaf52c.

——, J. Stroeve, and F. Fetterer, 2007: Whither Arctic sea ice? A clear signal of decline regionally, seasonally and extending beyond the satellite record. *Annals of Glaciology*, **46**, 428–434, https://doi.org/10.3189/172756407782871170.

——, G. Peng, D. J. Scott, and M. H. Savoie, 2014: Verification of a new NOAA/NSIDC passive microwave sea-ice concentration climate record. *Polar Research*, **33**, 21004, https://doi.org/10.3402/polar.v33.21004.

——, F. Fetterer, M. Savoie, S. Mallory, R. Duerr, and J. Stroeve, 2017: NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 3. Accessed 24 June 2020. https://doi.org/10.7265/N59P2ZTG.

Melia, N., K. Haines, and E. Hawkins, 2016: Sea ice decline and 21st century trans-Arctic shipping routes. *Geophysical Research Letters*, **43**, 9720–9728, https://doi.org/10.1002/2016GL069315.

Melling, H., 2002: Sea ice of the northern Canadian Arctic Archipelago. *J. Geophys. Res.*, **107**, 3181, https://doi.org/10.1029/2001JC001102.

NOAA, 2021: Climate models. *NOAA Climate.gov*,. https://www.climate.gov/maps-data/primer/climate-models (Accessed July 25, 2021).

Notz, D., 2014: Sea-ice extent and its trend provide limited metrics of model performance. *The Cryosphere*, **8**, 229–243, https://doi.org/10.5194/tc-8-229-2014.

Notz, D., 2015: How well must climate models agree with observations? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **373**, 20140164, https://doi.org/10.1098/rsta.2014.0164.

——, and J. Stroeve, 2016: Observed Arctic sea-ice loss directly follows anthropogenic CO2 emission. *Science*, **354**, 747–750, https://doi.org/10.1126/science.aag2345.

——, and ——, 2018: The Trajectory Towards a Seasonally Ice-Free Arctic Ocean. *Current Climate Change Reports*, **4**, 407–416, https://doi.org/10.1007/s40641-018-0113-2.

——, A. Jahn, M. Holland, E. Hunke, F. Massonnet, J. Stroeve, B. Tremblay, and M. Vancoppenolle, 2016: The CMIP6 Sea-Ice Model Intercomparison Project (SIMIP): Understanding sea ice through climate-model simulations. *Geoscientific Model Development*, **9**, 3427–3446, https://doi.org/10.5194/gmd-9-3427-2016.

NSIDC, 2020: Thermodynamics: Albedo | National Snow and Ice Data Center. https://nsidc.org/cryosphere/seaice/processes/albedo.html (Accessed July 24, 2021).

——, 2021: First-year ice. https://nsidc.org/cryosphere/glossary/term/first-year-ice (Accessed July 25, 2021).

Olason, E., and D. Notz, 2014: Drivers of variability in Arctic sea-ice drift speed. *Journal of Geophysical Research: Oceans*, **119**, 5755–5775, https://doi.org/10.1002/2014JC009897.

Onarheim, I. H., T. Eldevik, L. H. Smedsrud, and J. C. Stroeve, 2018: Seasonal and regional manifestation of Arctic sea ice loss. *J. Climate*, **31**, 4917–4932, https://doi.org/10.1175/JCLI-D-17-0427.1.

O'Rourke, R., and Coauthors, 2019: Changes in the Arctic: Background and Issues for Congress.

Parkinson, C., J. Comiso, H. Zwally, D. Cavalieri, P. Gloersen, and W. Campbell, 1987: Arctic sea ice, 1973-1976: Satellite passive-microwave observations.

Peng, G., J. L. Matthews, M. Wang, R. Vose, and L. Sun, 2020: What Do Global Climate Models Tell Us about Future Arctic Sea Ice Coverage Changes? *Climate*, **8**, 15, https://doi.org/10.3390/cli8010015.

Perovich, D. K., J. A. Richter-Menge, K. F. Jones, and B. Light, 2008: Sunlight, water, and ice: Extreme Arctic sea ice melt during the summer of 2007. *Geophysical Research Letters*, **35**, https://doi.org/10.1029/2008GL034007.

Polyakov, I. V., and Coauthors, 2010: Arctic Ocean Warming Contributes to Reduced Polar Ice Cap. *Journal of Physical Oceanography*, **40**, 2743–2756, https://doi.org/10.1175/2010JPO4339.1.

——, and Coauthors, 2017: Greater role for Atlantic inflows on sea-ice loss in the Eurasian Basin of the Arctic Ocean. *Science*, **356**, 285–291, https://doi.org/10.1126/science.aai8204.

Proshutinsky, A., and Coauthors, 2019: Analysis of the Beaufort Gyre freshwater content in 2003–2018. *Journal of Geophysical Research: Oceans*, **124**, 9658–9689, https://doi.org/10.1029/2019JC015281.

Rampal, P., J. Weiss, C. Dubois, and J.-M. Campin, 2011: IPCC climate models do not capture Arctic sea ice drift acceleration: Consequences in terms of projected sea ice thinning and decline. *Journal of Geophysical Research: Oceans*, **116**, https://doi.org/10.1029/2011JC007110.

Randall, D. A., R. A. Wood, S. Bony, R. Colman, and T. Fichefet, 2007: Co-Authors. 2007. Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, and H.L. Miller, Eds., Cambridge University Press.

Rigor, I. G., J. M. Wallace, and R. L. Colony, 2002: Response of sea ice to the Arctic oscillation. *J. Climate*, **15**, 2648–2663, https://doi.org/10.1175/1520-0442(2002)015<2648:ROSITT>2.0.CO;2.

Roach, A. T., K. Aagaard, C. H. Pease, S. A. Salo, T. Weingartner, V. Pavlov, and M. Kulakov, 1995: Direct measurements of transport and water properties through the Bering Strait. *Journal of Geophysical Research: Oceans*, **100**, 18443–18457, https://doi.org/10.1029/95JC01673.

Roach, L. A., S. F. B. Tett, M. J. Mineter, K. Yamazaki, and C. D. Rae, 2018: Automated parameter tuning applied to sea ice in a global climate model. *Clim Dyn*, **50**, 51–65, https://doi.org/10.1007/s00382-017-3581-5.

Roberts, A., and Coauthors, 2015: Simulating transient ice-ocean Ekman transport in the Regional Arctic System Model and Community Earth System Model. *Annals of Glaciology*, **56**, 211–228, https://doi.org/10.3189/2015AoG69A760.

Rosenblum, E., and I. Eisenman, 2016: Faster Arctic sea ice retreat in CMIP5 than in CMIP3 due to volcanoes. *Journal of Climate*, **29**, 9179–9188, https://doi.org/10.1175/JCLI-D-16-0391.1.

——, and ——, 2017: Sea ice trends in climate models only accurate in runs with biased global warming. *J. Climate*, **30**, 6265–6278, https://doi.org/10.1175/JCLI-D-16-0455.1.

Santer, B. D., and Coauthors, 2008: Consistency of modelled and observed temperature trends in the tropical troposphere. *International Journal of Climatology*, **28**, 1703–1722, https://doi.org/10.1002/joc.1756.

Schauer, U., and A. Beszczynska-Möller, 2009: Problems with estimation and interpretation of oceanic heat transport – Conceptual remarks for the case of Fram Strait in the Arctic Ocean. *Ocean Science*, **5**, 487–494, https://doi.org/10.5194/os-5-487-2009.

Schauer, U., A. Beszczynska-Möller, W. Walczowski, E. Fahrbach, J. Piechura, and E. Hansen, 2008: Variation of measured heat flow through the Fram Strait between 1997 and 2006. *Arctic–Subarctic Ocean Fluxes: Defining the Role of the Northern Seas in Climate*, R.R. Dickson, J. Meincke, and P. Rhines, Eds., Springer Netherlands, 65–85.

Schlichtholz, P., 2011: Influence of oceanic heat variability on sea ice anomalies in the Nordic Seas. *Geophysical Research Letters*, **38**, https://doi.org/10.1029/2010GL045894.

Schlichtholz, P., 2019: Subsurface ocean flywheel of coupled climate variability in the Barents Sea hotspot of global warming. *Scientific Reports*, **9**, 13692, https://doi.org/10.1038/s41598-019-49965-6.

von Schuckmann, K., and Coauthors, 2016: An imperative to monitor Earth's energy imbalance. *Nature Climate Change*, **6**, 138–144, https://doi.org/10.1038/nclimate2876.

Schweiger, A., R. Lindsay, J. Zhang, M. Steele, H. Stern, and R. Kwok, 2011: Uncertainty in modeled Arctic sea ice volume. *Journal of Geophysical Research: Oceans*, **116**, https://doi.org/10.1029/2011JC007084.

Screen, J. A., 2011: Sudden increase in Antarctic sea ice: Fact or artifact? *Geophysical Research Letters*, **38**, https://doi.org/10.1029/2011GL047553.

Serreze, M. C., and J. A. Francis, 2006: The Arctic amplification debate. *Climatic Change*, **76**, 241–264, https://doi.org/10.1007/s10584-005-9017-y.

——, and R. G. Barry, 2011: Processes and impacts of Arctic amplification: A research synthesis. *Global and Planetary Change*, **77**, 85–96, https://doi.org/10.1016/j.gloplacha.2011.03.004.

——, and J. Stroeve, 2015: Arctic sea ice trends, variability and implications for seasonal ice forecasting. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **373**, 20140159, https://doi.org/10.1098/rsta.2014.0159.

——, and W. N. Meier, 2019: The Arctic's sea ice cover: trends, variability, predictability, and comparisons to the Antarctic. *Annals of the New York Academy of Sciences*, **1436**, 36–53, https://doi.org/10.1111/nyas.13856.

——, A. P. Barrett, A. G. Slater, M. Steele, J. Zhang, and K. E. Trenberth, 2007: The large-scale energy budget of the Arctic. *Journal of Geophysical Research: Atmospheres*, **112**, https://doi.org/10.1029/2006JD008230.

Serreze, M. C., A. P. Barrett, J. C. Stroeve, D. N. Kindig, and M. M. Holland, 2009: The emergence of surface-based Arctic amplification. *The Cryosphere*, 9.

Shen, Z., A. Duan, D. Li, and J. Li, 2021: Assessment and ranking of climate models in Arctic sea ice cover simulation: From CMIP5 to CMIP6. *J. Climate*, **34**, 3609–3627, https://doi.org/10.1175/JCLI-D-20-0294.1.

Shu, Q., Q. Wang, Z. Song, F. Qiao, J. Zhao, M. Chu, and X. Li, 2020: Assessment of sea ice extent in CMIP6 with comparison to observations and CMIP5. *Geophysical Research Letters*, **47**, e2020GL087965, https://doi.org/10.1029/2020GL087965.

SIMIP Community, 2020: Arctic Sea Ice in CMIP6. *Geophysical Research Letters*, **47**, e2019GL086749, https://doi.org/10.1029/2019GL086749.

Skagseth, Ø., T. Furevik, R. Ingvaldsen, H. Loeng, K. A. Mork, K. A. Orvik, and V. Ozhigin, 2008: Volume and Heat Transports to the Arctic Ocean Via the Norwegian and Barents Seas. *Arctic–Subarctic Ocean Fluxes: Defining the Role of the Northern Seas in Climate*, R.R. Dickson, J. Meincke, and P. Rhines, Eds., Springer Netherlands, 45–64.

Smedsrud, L. H., and Coauthors, 2013: The role of the Barents Sea in the Arctic climate system. *Reviews of Geophysics*, **51**, 415–449, https://doi.org/10.1002/rog.20017.

Storto, A., and S. Masina, 2016: C-GLORSv5: An improved multipurpose global ocean eddy-permitting physical reanalysis. *Earth System Science Data*, **8**, 679–696, https://doi.org/10.5194/essd-8-679-2016.

Stroeve, J., and D. Notz, 2015: Insights on past and future sea-ice evolution from combining observations and models. *Global and Planetary Change*, **135**, 119–132, https://doi.org/10.1016/j.gloplacha.2015.10.011.

——, and ——, 2018: Changing state of Arctic sea ice across all seasons. *Environ. Res. Lett.*, **13**, 103001, https://doi.org/10.1088/1748-9326/aade56.

——, M. M. Holland, W. Meier, T. Scambos, and M. Serreze, 2007: Arctic sea ice decline: Faster than forecast. *Geophysical Research Letters*, **34**, https://doi.org/10.1029/2007GL029703.

Stroeve, J., A. Barrett, M. Serreze, and A. Schweiger, 2014: Using records from submarine, aircraft and satellites to evaluate climate model simulations of Arctic sea ice thickness. *The Cryosphere*, **8**, 1839–1854, https://doi.org/10.5194/tc-8-1839-2014.

Stroeve, J. C., V. Kattsov, A. Barrett, M. Serreze, T. Pavlova, M. Holland, and W. N. Meier, 2012a: Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations. *Geophys. Res. Lett.*, **39**, n/a-n/a, https://doi.org/10.1029/2012GL052676.

——, M. C. Serreze, M. M. Holland, J. E. Kay, J. Malanik, and A. P. Barrett, 2012b: The Arctic's rapidly shrinking sea ice cover: A research synthesis. *Climatic Change*, **110**, 1005–1027, https://doi.org/10.1007/s10584-011-0101-1.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research Atmospheres*, https://doi.org/10.1029/2000JD900719.

Taylor, P. C., M. Cai, A. Hu, J. Meehl, W. Washington, and G. J. Zhang, 2013: A decomposition of feedback contributions to polar warming amplification. *Journal of Climate*, **26**, 7023–7043, https://doi.org/10.1175/JCLI-D-12-00696.1.

Timmermans, M.-L., J. Toole, and R. Krishfield, 2018: Warming of the interior Arctic Ocean linked to sea ice losses at the basin margins. *Science Advances*, **4**, eaat6773, https://doi.org/10.1126/sciadv.aat6773.

Tschudi, M., C. Fowler, J. Maslanik, and J. C. Stroeve, 2011: Tracking the Movement and Changing Surface Characteristics of Arctic Sea Ice. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, **3**, 536–540, https://doi.org/10.1109/JSTARS.2010.2048305.

Tsubouchi, T., and Coauthors, 2018: The Arctic Ocean Seasonal Cycles of Heat and Freshwater Fluxes: Observation-Based Inverse Estimates. *Journal of Physical Oceanography*, **48**, 2029–2055, https://doi.org/10.1175/JPO-D-17-0239.1.

——, K. Våge, B. Hansen, K. M. H. Larsen, S. Østerhus, C. Johnson, S. Jónsson, and H. Valdimarsson, 2021: Increased ocean heat transport into the Nordic Seas and Arctic Ocean over the period 1993–2016. *Nature Climate Change*, **11**, 21–26, https://doi.org/10.1038/s41558-020-00941-3.

Tsujino, H., and Coauthors, 2018: JRA-55 based surface dataset for driving ocean–sea-ice models (JRA55-do). *Ocean Modelling*, **130**, 79–139, https://doi.org/10.1016/j.ocemod.2018.07.002.

Uotila, P., and Coauthors, 2019: An assessment of ten ocean reanalyses in the polar regions. *Clim Dyn*, **52**, 1613–1650, https://doi.org/10.1007/s00382-018-4242-z.

Urrego-Blanco, J. R., N. M. Urban, E. C. Hunke, A. K. Turner, and N. Jeffery, 2016: Uncertainty quantification and global sensitivity analysis of the Los Alamos sea ice model. *Journal of Geophysical Research: Oceans*, **121**, 2709–2732, https://doi.org/10.1002/2015JC011558.

Watanabe, E., and H. Hasumi, 2005: Arctic sea ice response to wind stress variations. *Journal of Geophysical Research: Oceans*, **110**, https://doi.org/10.1029/2004JC002678.

Watts, M., W. Maslowski, Y. J. Lee, J. C. Kinney, and R. Osinski, 2021: A Spatial Evaluation of Arctic Sea Ice and Regional Limitations in CMIP6 Historical Simulations. *Journal of Climate*, **34**, 6399–6420, https://doi.org/10.1175/JCLI-D-20-0491.1.

Wayand, N. E., C. M. Bitz, and E. Blanchard-Wrigglesworth, 2019: A year-round subseasonal-to-seasonal sea ice prediction portal. *Geophysical Research Letters*, **46**, 3298–3307, https://doi.org/10.1029/2018GL081565.

Werrell, C., and F. Femia, 2018: Senior military leaders to Congress: Climate change presents strategic challenges in the Arctic. *The Center for Climate & Security*,. https://climateandsecurity.org/2018/04/senior-military-leaders-take-climate-change-in-the-arctic-seriously/ (Accessed April 26, 2021).

Winton, M., 2011: Do climate models underestimate the sensitivity of Northern Hemisphere sea ice cover? *Journal of Climate*, **24**, 3924–3934, https://doi.org/10.1175/2011JCLI4146.1.

Woodgate, R. A., T. Weingartner, and R. Lindsay, 2010: The 2007 Bering Strait oceanic heat flux and anomalous Arctic sea-ice retreat. *Geophysical Research Letters*, **37**, https://doi.org/10.1029/2009GL041621.

——, T. J. Weingartner, and R. Lindsay, 2012: Observed increases in Bering Strait oceanic fluxes from the Pacific to the Arctic from 2001 to 2011 and their impacts on the Arctic Ocean water column. *Geophysical Research Letters*, **39**, https://doi.org/10.1029/2012GL054092.

Woods, C., and R. Caballero, 2016: The Role of Moist Intrusions in Winter Arctic Warming and Sea Ice Decline. *Journal of Climate*, **29**, 4473–4485, https://doi.org/10.1175/JCLI-D-15-0773.1.

Yang, Q., M. Losch, S. N. Losa, T. Jung, L. Nerger, and T. Lavergne, 2016: Brief communication: The challenge and benefit of using sea ice concentration satellite data products with uncertainty estimates in summer sea ice data assimilation. *The Cryosphere*, **10**, 761–774, https://doi.org/10.5194/tc-10-761-2016.

Yi, D., and H. J. Zwally, 2009: Arctic sea ice freeboard and thickness, version 1. Accessed 29 October 2020. https://doi.org/10.5067/SXJVJ3A2XIZT.

Zhang, J., and D. A. Rothrock, 2003: Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates. *Monthly Weather Review*, https://doi.org/10.1175/1520-0493(2003)131<0845:MGSIWA>2.0.CO;2.

Zickfeld, K., V. K. Arora, and N. P. Gillett, 2012: Is the climate response to CO2 emissions path dependent? *Geophysical Research Letters*, **39**, https://doi.org/10.1029/2011GL050205.

Zuo, H., M. A. Balmaseda, S. Tietsche, K. Mogensen, and M. Mayer, 2019: The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: A description of the system and assessment. *Ocean Sci.*, **15**, 779–808, https://doi.org/10.5194/os-15-779-2019.

Zygmuntowska, M., P. Rampal, N. Ivanova, and L. H. Smedsrud, 2013: Uncertainties in Arctic sea ice thickness and volume: new estimates and implications for trends. *The Cryosphere Discuss.*, **7**, 5051–5095, https://doi.org/10.5194/tcd-7-5051-2013.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California