REDUCING UNCERTAINTY OF SIMULATED INTERNAL VARIABILITY OF ARCTIC SEA ICE

by

CHRISTOPHER R. WYBURN-POWELL

B.Sc., University of East Anglia, 2019

M.S., University of Colorado, 2022

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Atmospheric and Oceanic Sciences

2024

Committee Members:

Prof. Alexandra Jahn

Prof. Jennifer Kay

Prof. Kristopher Karnauskas

Dr. Marika Holland

Dr. David Bailey

Reducing Uncertainty of Simulated Internal Variability of Arctic Sea Ice

Wyburn-Powell, Christopher R. (Ph.D., Atmospheric and Oceanic Sciences)

Thesis directed by Prof. Alexandra Jahn

Abstract

Arctic sea ice is currently undergoing a period of rapid decline and concurrent high variability. Future projection uncertainty is dominated by internal variability in the short- and medium-term. In this thesis we use a large number of climate model simulations, which together can explain drivers and characteristics of variability, whilst diagnosing model limitations when compared with observations. We find that large ensembles from the Coupled Model Intercomparison Project - Phase 5 (CMIP5) generally simulate interannual variability within observational uncertainty. However, model realism is only achieved due to large observational uncertainties and spatially heterogeneous model biases. To reduce uncertainty of low-frequency variability, we find that climate modes of variability can be used skillfully to explain regional Arctic sea ice variability. The primary drivers of a decadal decline in Arctic sea ice are above trend global temperatures, a negative Interdecadal Pacific Oscillation (IPO) and a positive Atlantic Meridional Oscillation (AMO), which currently are aligned as such. The IPO across CMIP6 climate models is highly important for up to 15 years, but with large variations in magnitude and sign between models and realizations. With further investigation, we find the decadal predictive skill of the IPO is often dependent on the modulation by the slowly-changing Pacific-Arctic (PARC) atmospheric teleconnection. Considering the current negative phases of the IPO and PARC, an increased likelihood of an accelerated transition to an ice-free Arctic is expected by most climate models.

DEDICATION

To my parents who always encouraged and enabled me to follow my interests and passions, and to my husband Sean whose love and support has guided me through my studies.

ACKNOWLEDGEMENTS

Firstly, I would like to acknowledge the science reporters whose writing on the record minimum Arctic sea ice extent of 2012 first sparked my interest in the research being conducted in Boulder. I would also like to thank teachers throughout my academic career who guided me toward the pursuit of science.

I would like to thank my committee members who have asked interesting and thought-provoking questions, which have immeasurably improved my work over the years. I would also like to acknowledge my former committee member Mark England who was instrumental in the research project which forms Chapter 2 of this dissertation.

I would like to especially thank my advisor Alexandra Jahn, whose research ideas led me to join her research group and focus on important questions in the field. I thank her for proposing this interesting scope of work, and trusting me to carry out this research. Professor Jahn's mentorship and dedication to her students is exemplary. I am grateful to have worked with her and to have been a member of The Polar and Paleoclimate Modeling Group. The research group has granted me mentorship from the senior scientists in the group: Hannah Zanowski, Abby Smith and Patricia DeRepentigny. I am also grateful for my friend and colleague Jed Lenetsky and our many collaborations over the years. More broadly, I would like to thank the ATOC and INSTAAR community, who have made working here a rewarding experience.

This work was supported by the National Science Foundation (grant number 1847398). We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. We also acknowledge the US CLIVAR Working Group on Large Ensembles and the modeling centers that contributed to the CLIVAR Large Ensemble project. We would also like to acknowledge high-performance computing on Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation.

CONTENTS

CHAPTER

TABLES

Table

FIGURES

Figure

CHAPTER 1

INTRODUCTION

Due to anthropogenic emissions, the Arctic is warming more rapidly than any other region of the planet, with much of this Arctic Amplification caused by sea ice loss (Dai and Bloecker, 2019; England et al., 2021). Sea ice responds quickly to a changing climate, with the ice-albedo feedback driving much of this sensitivity to global temperatures (Kashiwase et al., 2017). The Arctic will likely be seasonally ice-free for the first time in the next two decades (Jahn et al., 2024). This will dramatically alter the way humans interact with the Arctic Ocean, from resource exploitation and shipping (e.g. Petrick et al., 2017; Chen et al., 2020), to uses by indigenous communities (Christensen and Nilsson, 2017), and with large biosphere effects (Kovacs et al., 2011; Lim et al., 2022). With reduced sea ice cover, there are also associated local temperature, moisture and circulation effects (e.g. Ogi and Wallace, 2007; Boeke and Taylor, 2018; Smith et al., 2022), with subsequent remote impacts for the global climate (Li et al., 2013; Sévellec et al., 2017). The decline of sea ice cover is occurring gradually over decadal timescales as a response to increasing CO_2 forcing (Notz and Stroeve, 2016), with dramatic changes in sea ice cover due to internal variability also occurring on decadal and interannual timescales (Kay et al., 2011; Swart et al., 2015; Mioduszewski et al., 2019). Thus, understanding the future evolution of Arctic sea ice and reducing uncertainty in its prediction is the research focus of this dissertation, which is both timely and has a wide range of applications.

The most dramatic declines and variations in Arctic sea ice occur at the seasonal minimum in September (Årthun et al., 2021), which is the primary focus of the research in this thesis. At a mean state similar, or slightly lower than current conditions, peak interannual variability is expected (Goosse et al., 2009; Jahn et al., 2016; Olonscheck and Notz, 2017; Massonnet et al., 2018; Landrum and Holland, 2020). This is due to the average sea ice edge in September at current warming levels being away from the coasts, facilitating large variations in meridional ice edge without constraint (Eisenman, 2010). This manifests as peak variability in the shelf sea regions along the Siberian and Alaskan coasts; which are also coincident with the regions of fastest ice retreat and current locations of shipping interests (England et al., 2019; Smith and Stephenson, 2013).

Projection uncertainty of climate phenomena can be separated into three distinct drivers: scenario, model, and internal variability, which act at different timescales (Hawkins and Sutton, 2009). In order to separate these distinct forms of projection uncertainty we must rely on global climate model (GCM) simulations. Projected sea ice has a high degree of uncertainty in both the mean state and in the response to the effect of anthropogenic emissions of greenhouse gases and aerosols (Notz and SIMIP-Community, 2020). Sea ice components of GCMs can be sensitive to model physics and parameterizations (Kay et al., 2022), but other components of the coupled climate system and forcing also vary (Derepentigny et al., 2022; Kuma et al., 2023), further contributing to the range of model performance for Arctic sea ice. Thus, model uncertainty contributes substantially to projection uncertainty at all time periods (Keen et al., 2021). However, despite these model uncertainties, internal variability is the dominant source of uncertainty in Arctic sea ice extent in the medium-term ($\sim$5-20 years) for the shelf seas in September (Bonan et al., 2021). Scenario uncertainty does not alter the trajectories for September SIE until after the expected first ice-free years Notz and SIMIP-Community (2020). This leaves the period of up to several decades in which reducing internal variability prediction uncertainty can be highly impactful.

High quality Arctic sea ice observations started with continuous satellite observations in 1978 (Meier and Stewart, 2019), which has also been a period of rapid climate change (Stroeve et al., 2012). Such single realizations, from observations or one climate model simulation, contain both the forced response from climate change and internal variability. Single model initial condition large ensembles (SMILEs) (e.g. Kay et al., 2015; Maher et al., 2019; Lehner et al., 2020), are ideally suited to isolate the internal variability by providing many realizations from a single model. This is achieved by each of the realizations having the same forcing, differing only in small perturbations at initialisation. Thus, internal variability can be explicitly studied, as the difference between realizations within a SMILE is due only to internal variability, with the ensemble mean representing the forced response (Thompson et al., 2015; Deser et al., 2020; Suarez-Gutierrez et al., 2021). It is only in the last several years with the two latest coupled model intercomparison projects (CMIP5 and CMIP6) that SMILEs have become ubiquitous, with sufficiently large numbers of ensemble members to characterize and compare different internal variability simulations (Milinski et al., 2020). As a result, it has been shown that the observed sea ice loss has been enhanced by internal variability (Kay et al., 2011). Approximately 43-53% the rapid decline in Arctic sea ice over over this time period has been attributed to internal variability enhancing the response to warming (Stroeve et al., 2007; Kay et al., 2011; Ding et al., 2019), with the period 2000-2012 particularly dominating these calculations (Dörr et al., 2023). However, it is difficult to assess how 'typical' our observations have been based on models, while accounting for model biases, due to our single realization of reality (Notz, 2015; Frankcombe et al., 2018).

Large ensemble simulations provide many equally likely realizations of the climate system, resulting in a climatology of interannual variability, as assessed in Chapter 2. In order to understand the drivers of interannual variability, various modeling techniques have be deployed where aspects of the short-term chaotic nature of the climate system are controlled. Through a 'decoupling' analysis, temperature fluctuations were found to be highly influential, explaining nearly three quarters of variability (Olonscheck et al., 2019). Through 'nudging' experiments, wind patterns have also

been found to explain around approximately a third of internal variability (Roach and Blanchard-Wrigglesworth, 2022; Ding et al., 2022), which can be partly encapsulated by high-frequency modes of variability in sea level pressure (Ukita et al., 2007; Serreze et al., 2007; L'Heureux et al., 2008; Zhang et al., 2019; Liu et al., 2021). These high-frequency and coupled drivers of variability manifest themselves in large heterogeneity of summer sea ice, with variable thermodynamic and dynamic properties due to sea ice with varying snow cover and melt ponds adjacent to open water (Nicolaus et al., 2022). Much of this highly regional and high-frequency variability is dampened for large-scale metrics such as pan-Arctic monthly sea ice extent. However, for applications such as the navigability of specific shipping routes, the specific manifestation of local variability has great importance Winton et al. (2022).

For timescales longer than 2 years there is little predictive skill that can be gained from initial state prediction (Blanchard-Wrigglesworth et al., 2011; Yeager et al., 2015; Bushuk and Giannakis, 2017; Bonan et al., 2019; Gregory et al., 2021). At multi-decadal timescales climate forcing dominates projections, but at approximately 2-20 year timescales some predictability is retained due to slowly changing modes of variability within the climate system (Guemas et al., 2016), which is the focus of Chapter 3. Dominant modes of decadal variability with the potential to affect Arctic sea ice are primarily associated with Pacific sea surface temperatures (SSTs) (Hu et al., 2016; Li et al., 2018a; Screen and Deser, 2019; Ding et al., 2019; Kim et al., 2020; Clancy et al., 2021; Jeong et al., 2022b; Simon et al., 2022), but also with Atlantic Ocean variability (Day et al., 2012; Miles et al., 2014; Meehl et al., 2018; Li et al., 2018b). These patterns of SST anomalies are thought to affect the circulation, heat and moisture fluxes between the tropics and Arctic, affecting sea ice (Yuan et al., 2018; Baxter et al., 2019). Further, these pathways of tropical-Arctic connection are thought to have variability themselves (e.g. Bonan and Blanchard-Wrigglesworth, 2020). For example, the Pacific-Arctic (PARC) teleconnection is thought to greatly modulate the transmission of Rossby waves between the tropical Pacific and affect the summer melting of Arctic sea ice by favouring certain upper tropospheric pressure patterns (Ding et al., 2019).

The drivers of decadal variability of Arctic sea ice have not been coherently understood, especially at decadal to multi-decadal timescales. This is due to a number of reasons. Firstly, it is uncertain whether climate models are able to accurately simulate large-scale decadal variability in the climate system (Fasullo et al., 2020; Dalelane et al., 2023). This is partly due to the fact that our observations cover a short period of time compared to decadal to multi-decadal modes of variability, leading to uncertainties in processes (e.g. Capotondi et al., 2023), and uncertainty in how observations can be compared with the full distribution of what could have occurred (Notz, 2015). Secondly, large-scale variability is expected to be sensitive to changes in forcing, such as the spatial pattern of the Pacific Decadal Oscillation (Werb and Rudnick, 2023), the Atlantic Multidacadal Oscillation (Klavans et al., 2022), or Interdecadal Pacific Oscllation (IPO) (e.g. Meehl et al., 2013). Thirdly, the teleconnections between different climate modes of variability and between the tropics and the Arctic are still poorly understood at decadal timescales (Liu et al., 2004; Blanchard-Wrigglesworth and Ding, 2019). Given these limitations it is therefore difficult to evaluate model performance and isolate the causes of model differences. Nonetheless, prior work has shown that there is promise in using these decadal and multidecadal modes of variability to reduce the projection uncertainty of an ice-free Arctic Screen and Deser (2019). Thus, in Chapter 4, we assess the impact of tropical Pacific climate modes of variability on Arctic sea ice predictability in a suite of large ensembles.

The overarching research question that this dissertation addresses is: **How can we best understand internal variability of Arctic sea ice both in models and observations in order to reduce our projection uncertainty of future sea ice conditions?** Three specific research questions are explored in the subsequent chapters of this thesis which stem from, and seeks to expand upon, the current literature. Firstly, **are large ensembles in the CMIP5 archive consistent with observations in their representation of historical Arctic sea ice interannual variability?** Secondly, **how do dominant modes of climate variability have remote impacts on regional Arctic sea ice anomalies at yearly to decadal lead**

times, and can we explain future sea ice trajectories based solely on these modes of variability? Thirdly, **can we reduce uncertainty in decadal Arctic sea ice trajectories on the way to an ice-free Arctic based on the initial state of Pacific decadal variability, namely the IPO and PARC teleconnections?**

CHAPTER 2

MODELED INTERANNUAL VARIABILITY OF ARCTIC SEA ICE COVER IS WITHIN OBSERVATIONAL UNCERTAINTY

Preface

This chapter is published as cited below, and is reproduced here based on the last author version. It is identical in content to the published version, except for formatting and copy editing:

Wyburn-Powell, C., Jahn, A., England, M. (2022). Modeled Interannual Variability of Arctic Sea Ice Cover is Within Observational Uncertainty. Journal of Climate. 35 (20) 3227–3242. doi:10.1175/JCLI-D-21-0958.1.

Abstract

Internal variability is the dominant cause of projection uncertainty of Arctic sea ice in the short and medium term. However, it is difficult to determine the realism of simulated internal variability in climate models, as observations only provide one possible realization while climate models can provide numerous different realizations. To enable a robust assessment of simulated internal variability of Arctic sea ice, we use a resampling technique to build synthetic ensembles for both observations and climate models, focusing on interannual variability which is the dominant timescale of Arctic sea ice internal variability. We assess the realism of the interannual variability of Arctic sea ice cover as simulated by six models from the Coupled Model Intercomparison Project 5 (CMIP5) that provide large ensembles compared to four observational datasets. We augment the standard definition of model and observational consistency by representing the full distribution of resamplings, analogous to the distribution of variability which could have randomly occurred. We

find that modeled interannual variability typically lies within observational uncertainty. The three models with the smallest mean state biases are the only ones consistent in the pan-Arctic for all months, but no model is consistent for all regions and seasons. Hence, choosing the right model for a given task as well as using internal variability as additional metric to assess sea ice simulations is important. The fact that CMIP5 large ensembles broadly simulate interannual variability consistent within observational uncertainty gives confidence in the internal projection uncertainty for Arctic sea ice based on these models.

Significance Statement

The purpose of this study is to evaluate the historical simulated internal variability of Arctic sea ice in climate models. Determining model realism is important to have confidence in the projected sea ice evolution from these models, but so far only mean state and trends are commonly assessed metrics. Here we assess internal variability with a focus on the interannual variability, which is the dominant timescale for internal variability. We find that, in general, models agree well with observations, but as no model is within observational uncertainty for all months and locations, choosing the right model for a given task is crucial. Further refinement of internal variability realism assessments will require reduced observational uncertainty.

2.1 Introduction

Arctic sea ice has declined precipitously since 1979, at a faster rate than at any time over the last millennium (Brennan and Hakim, 2022), with less than half the summer area and one quarter the summer volume remaining (Schweiger et al., 2011; Notz and Stroeve, 2018). This observed decline is due to both anthropogenic climate change and internal variability, which can act to amplify or dampen the trend from external forcing alone (Kay et al., 2011; Notz and Marotzke, 2012). The relative contribution of internal variability to the observed September sea ice area decline remains uncertain but has been estimated at 43-53% (Stroeve et al., 2007; Kay et al., 2011; Ding et al., 2019).

Internal variability also influences future sea ice projections, leading to large internal variability uncertainty, especially for the next few decades (Kay et al., 2011; Jahn et al., 2016; Bonan et al., 2021). As internal variability is such a large contributor to the observed and projected changes in Arctic sea ice cover, but Global Climate Models (GCMs) differ in the magnitude of their simulated sea ice internal variability (Olonscheck and Notz, 2017), it is imperative that we understand how realistically models simulate internal variability.

Internal variability of Arctic sea ice has been shown to be spatially heterogeneous (England et al., 2019), and act on multiple time scales from annual to multi-decadal (Zhang and Wallace, 2015; Ding et al., 2017, 2019; Brennan and Hakim, 2022). Over the historical period, internal variability has been the dominant cause of sea ice decline in many regions, most notably parts of the Kara Sea in summer and the Barents Sea in winter (Li et al., 2017; England et al., 2019; Dörr et al., 2021). Sea ice loss in recent decades has been most rapid and expansive in the summer, particularly in the shelf seas which have transitioned from mainly ice-covered to ice-free for more of the summer, facilitating high internal variability (Onarheim et al., 2018; Mioduszewski et al., 2019). These areas of rapid and unpredictable change coincide with the most impactful areas for a range of stakeholders from shipping and oil interests to indigenous peoples and biodiversity (Kovacs et al., 2011; Petrick et al., 2017; Christensen and Nilsson, 2017; Chen et al., 2020).

The established way to estimate internal variability in GCMs is to use multiple realizations of single model initial-condition large ensembles (SMILEs) or long constant forcing model runs to assess the ensemble spread or standard deviation (Olonscheck and Notz, 2017; Lehner et al., 2020; Maher et al., 2020). SMILEs have successfully been used to study internal variability in the context of polar temperatures (England, 2021), precipitation trends (Dai and Bloecker, 2019), and regional trends (McKinnon and Deser, 2018; Hu et al., 2019). However, such analysis can not be done on observations, due to only one realization of reality and a limited length of the observational record. It is this single realization of reality over a relatively short period of time that has

previously prevented direct assessment of internal variability of Arctic sea ice in models compared to observations. Hence, previous sea ice model assessments have been focused on the trends (e.g. Swart et al., 2015; Rosenblum and Eisenman, 2017), sensitivity to warming (e.g. Winton, 2011; Niederdrenk and Notz, 2018), and mean state (e.g. Davy and Outten, 2020). Furthermore, even if we were able to precisely disentangle internal variability from the forced response in observations, comparisons with GCMs are still challenging because we do not know where the one realization seen in the observations falls within the probability distribution obtained from a model ensemble (Notz, 2015).

Here, we provide the first direct comparison of internal variability of Arctic sea ice from a suite of SMILEs from the Coupled Model Intercomparison Project Phase 5 (CMIP5) with observations, by using a statistical technique to construct a 'synthetic ensemble' of Arctic sea ice variability, following McKinnon et al. (2017). Synthetic ensembles have been used for several climate variability questions such as for sea surface temperature (Chan et al., 2020), climate extremes (Deser et al., 2020), precipitation (McKinnon and Deser, 2021), ocean chlorophyll concentration (Elsworth et al., 2021), and Antarctic sea ice trends (Chemke and Polvani, 2020). Here we present the first use of a synthetic ensemble for studying Arctic sea ice, specifically to assess the realism of internal variability on interannual timescales. Using the synthetic ensemble method, we are able to show that generally the simulated interannual variability fits within the observational uncertainty derived from different datasets, but that there are considerable seasonal and spatial differences, and that some models perform better than others for a given task. We also show that internannual variability makes up approximately three quarters of the total Arctic sea ice internal variability, and hence the majority of the sea ice internal variability over the past 42 years.

2.2 Data Sources

2.2.1 Observational data

We primarily use two observational datasets for sea ice concentrations (SIC), the National Snow and Ice Data Center (NSIDC) Climate Data Record (CDR) version 4 (Meier et al., 2021), and the Hadley Centre Sea Ice and Sea Surface Temperature dataset (HadISST1) (Rayner et al., 2003). In order to further test the sensitivity of our results to the observational dataset used, we also utilize datasets derived from the satellite algorithums NASA Team (NT) (Cavalieri et al., 1984) and NASA Bootstrap (BT) (Comiso, 1986). Together, these datasets are a representative sample of interpretations of past sea ice conditions, with both the mean state and variability differing between the datasets due to observational uncertainties (Comiso et al., 2017; Kern et al., 2019). Sea ice area (SIA) was chosen over sea ice extent (SIE) as the variability of SIA is more independent of satellite algorithms and is intrinsically more precise and thus better for comparing internal variability between models (Notz, 2014). All analysis is performed using monthly data for 1979 to 2020. Missing data for NSIDC datasets, and discontinuities in the HadISST1 dataset, were filled using the same month's data in a different year, instead of interpolating to avoid unrealistic SIC values (see Table S2.1) for the specific replacements used).

2.2.2 Model data

Six models from the Climate Variability and Predictability Program (CLIVAR) Multi-Model Large Ensemble Archive (Deser et al., 2020) are utilized in this analysis, as detailed in Table 2.1. EC-Earth was excluded from this analysis due to no available SIC output. All models are CMIP5-class and use historical and representative concentration pathway (RCP) 8.5 forcing, the high emissions CMIP5 scenario. The models from the Multi-Model Large Ensemble Archive are diverse in their mean state and trends, spanning nearly the full range of CMIP5 sea ice projections (see Figure 1 in Bonan et al., 2021). In winter, model mean-state biases are typically smaller in absolute and relative terms than summer (see Table S2.2). The notable outliers in summer are

Table 2.1: Models used in this analysis from the CLIVAR Multi-Model Large Ensemble Archive (Deser et al., 2020)

Modeling Center	Model	Members	Years	Reference
CCCma	CanESM2	50	1950-2100	Kirchmeier-Young et al. (2017)
NCAR	CESM1	40	1920-2100	Kay et al. (2015)
CSIRO	MK3.6	30	1850-2100	Jeffrey et al. (2013)
GFDL	CM3	20	1920-2100	Sun et al. (2018)
GFDL	ESM2M	30	1950-2100	Rodgers et al. (2015)
MPI	ESM1	100	1850-2100	Maher et al. (2019)

CanESM2 with the largest negative mean-state bias of -54% in September and CSIRO MK3.6 being an extreme positive outlier for all seasons and $+83\%$ in September. Although GFDL CM3 is not as large an outlier in mean state in September, its SIA loss over the period 1979-2020 is by far the most rapid. The six models range in ensemble size between 20 and 100 (see Table 2.1). We present results for all members of the SMILEs to assess each GCM's ability to realistically simulate the observed interannual variability. We also provide subsampled results, scaled to 20 members, the size of the smallest large ensemble, for model inter-comparison with our consistency metric. Subsampling is discussed in more detail in section 2.3.3.

2.3 Methods

2.3.1 Resampling technique

We estimate interannual variability in a single model member or observational time series by assuming the forced response is represented by an ordinary least squares regression linear trend. This assumption is deemed appropriate for 1979-2020, but may not be applicable for time periods extending further back (England et al., 2019; England, 2021), and allows us to follow the methodology from McKinnon et al. (2017). Anomalies from this linear trend are therefore considered largely due to interannual variability alone. Typically the ensemble mean is a more accurate measure of the forced response, but as discussed in section 2.3.4 and 2.4.3, detrending using the individual member produces similar results and is chosen here for reasons detailed in 2.3.4.

By using this technique we can calculate a consistent metric of interannual variability to directly compare model members and observations. We consider all months in the pan-Arctic and present spatial results for the minimum and maximum SIA months September and March respectively. We resample the anomalies from the linear trend 10,000 times for SIA and 1000 times for each SIC grid box, with replacement, and use a 2 year bootstrap block size. This can be considered analogous to shuffling independent anomalies to produce a range of alternative scenarios which would have been equally likely to occur, allowing us to calculate metrics of interannual variability for a representative sample of all possible scenarios (see Figure 2.1). As suggested by McKinnon et al. (2017) and McKinnon and Deser (2018), we retain spatial coherence by resampling in the time dimension for all grid boxes at once. 10,000 resamplings in the pan-Arctic were chosen for increased reliability of consistency classifications, whereas spatially 1000 was determined sufficient as each grid box has a lower impact on results if a classification were to change from rerunning the experiment. A 2 year block size is chosen because normalized autocorrelation frequently exceeds 0.4 for a lag of 1 year, and a marked drop-off in autocorrelation between a lag of 1 and 2 years occurs in comparison with years 2 and 3 (not shown), occurring both spatially and in the pan-Arctic time series. Resampling with a 1 year or 2 year block size leads to almost identical results (not shown).

We focus our analysis on the standard deviation of sea ice state over the 42-year period 1979-2020, not the trends, as we want to assess the realism of the models' simulated interannual variability, rather than the realism of the simulated trends (see Swart et al. (2015) for a discussion of simulated trends compared to observations). The standard deviation with respect to time is computed either for the 10,000 pan-Arctic SIA resampling or the 1000 SIC resamplings in each grid cell. To represent the distribution of these resamplings or ensemble members we use the standard deviation (σ) and mean (μ). Here, σ can be considered analogous to the range of interannual variability which could have occurred, given the underlying data; μ is analogous to the typical interannual variability represented in the resamplings.

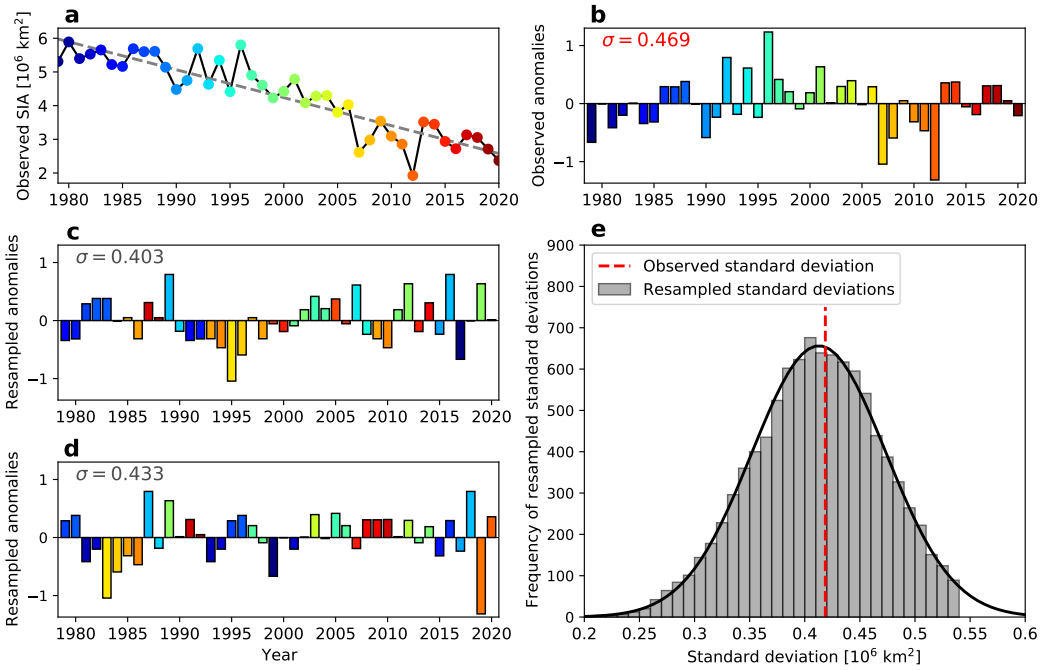

Figure 2.1: **Resampling methodology, applied to the observed September SIA.** (a): Observed sea ice area from CDR (dots) with linear trend (grey dashed line). (b): Anomalies from the linear trend. (c,d) Two randomly different resamplings of the anomalies in (b), color coded to match the year of anomaly. (e): Distribution of the standard deviation with respect to time for all 10,000 resamplings. The printed statistics represent σ for standard deviation. In (e) the red vertical line represents the standard deviation of the original data and gray refers to the distribution of standard deviations for the 10,000 resamplings.

In order to directly compare interannual variability between models and observations we define three measures of variability as follows, where σ_{LE} is internal variability in SMILEs and both σ_{mem} and σ_{obs} are the interannual variability within a synthetic ensemble:

- σ_{LE} and μ_{LE} - Standard deviation and mean of standard deviations within a single large ensemble, without resampling, an established measure of the full range of internal variability.

- σ_{mem} and μ_{mem} - Standard deviation and mean of the standard deviations of all resamplings of a single model member. The resampling process for a given ensemble member is equivalent to that of the observations in Figure 2.1. The median member's value across all members of the SMILE is denoted $\bar{\sigma}_{mem}$ and $\bar{\mu}_{mem}$.

- σ_{obs} and μ_{obs} - Standard deviations and mean of the standard deviation of all resamplings of the single realization of the observational dataset. These metrics relate to Figure 2.1 as the standard deviation and mean of the distribution in panel e.

2.3.2 Consistency

To assess the realism of simulated internal variability, we utilize a consistency metric to provide a binary classification as to whether the modeled variability is within or outside the range of observational uncertainty. For sea ice analysis in the past, consistency has typically been defined by at least one member of a large ensemble overlapping with observations (e.g. Notz, 2015; Swart et al., 2015; Jahn, 2018). However, this is a relatively low bar for models to reach. Other more elaborate consistency methods have been applied for other aspects of the climate system, e.g. Santer et al. (2008), and applied to Arctic sea ice by Stroeve et al. (2012). However the methodology of Santer et al. (2008) bases consistency assessments on trends rather than the internal variability independent of the trends, as is the goal here. Hence, we here use resampling and define consistency by comparing distributions, as it allows us to compare whether the resampled distributions overlap. Comparing distributions is a more stringent decision about consistency than comparing

single values for each ensemble member or observational dataset that would be available without resampling. Further augmentation to this binary classification is achieved by comparing SMILEs with four diverse observational datasets independently, adding the category of 'consistent within observational uncertainty'. We only use this three-category consistency classification rather than a significance or probability value (e.g., from a student t-test), as both the resampled average variability (μ_{mem}) and standard deviation of variability (σ_{mem}) are positively skewed across members. Nonetheless, we find that a 95% confidence interval is in fact similar to our consistency classification, but classifies fewer instances of inconsistency in the pan-Arctic than our method.

Applying this consistency metric to Arctic SIA, each SMILE realization or observational dataset has a different value of interannual variability for each of the 10,000 resamplings. These 10,000 resamplings from a single member or observational time series are approximately normally distributed and as such can be thought of as probability distribution functions (PDFs) (see Figure 2.2). The width of the PDFs show the distribution of the 10,000 resamplings, indicating the range of possible interannual variabilities (proportional to σ_{mem} and σ_{obs}). The location on the horizontal-axis indicates the average interannual variability (μ_{mem} and μ_{obs}). For models and observations to be considered 'consistent' in the following, we require their means (their position on the horizontal-axis in Figure 2.2) and their standard deviations (height on the vertical-axis) to overlap such that at least one member is greater than the lowest observational dataset and one member is lower than the highest observations for each σ and μ metric independently. Average SIC differences does not preclude a consistent classification as variability may be equal between a SMILE and observational datasets but about different means. However, due to the zero-bound nature of SIC, if a mean state differs so much that SMILE members have at least some sea ice where there is no sea ice in the observational datasets, we exclude those regions from the analysis rather than classifying them as inconsistent. We do this as the focus of our analysis is on assessing the realism of actual sea ice variability, so we only compare regions where there is variability in both models and observations.

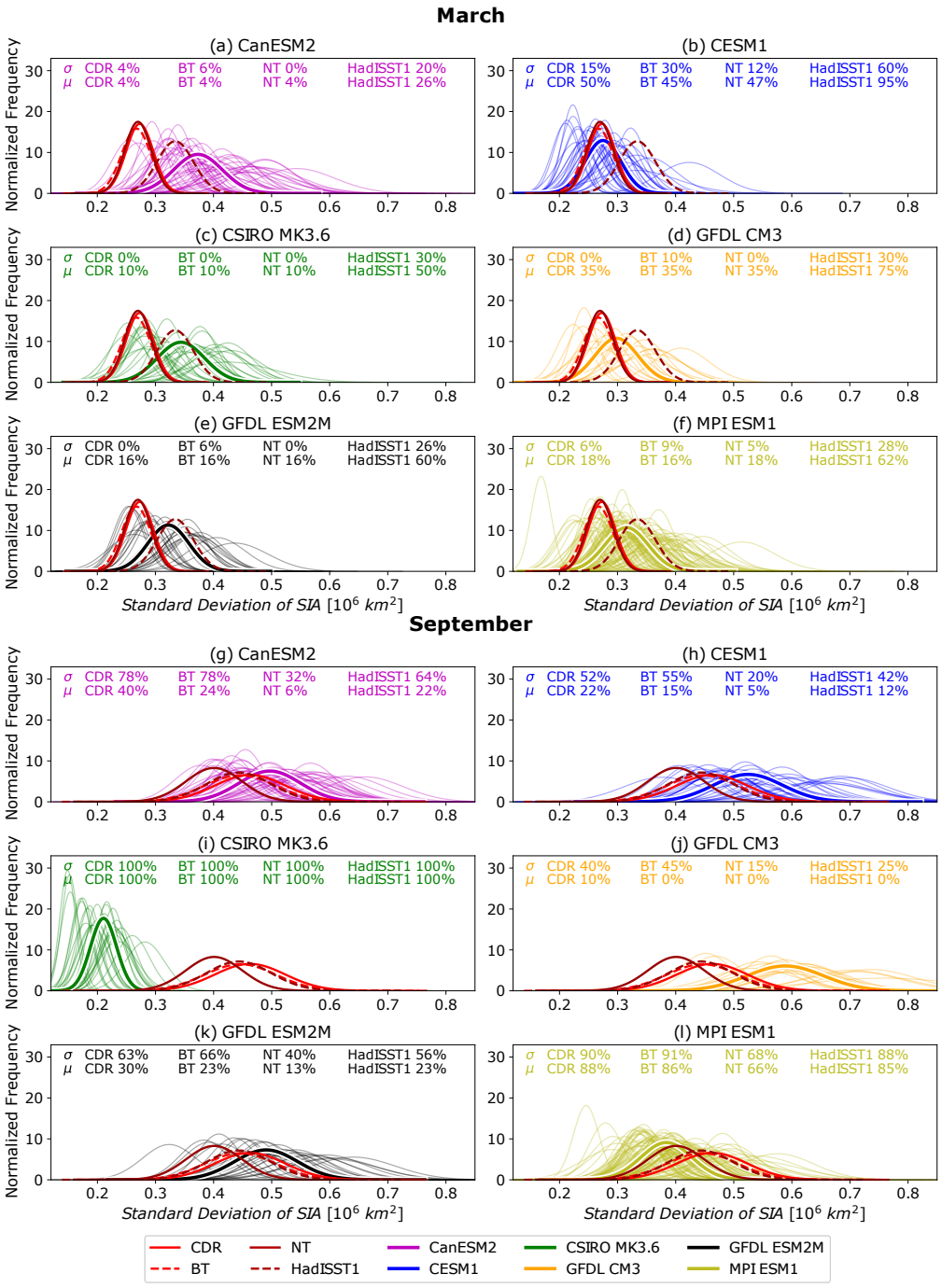

Figure 2.2: **Distribution of pan-Arctic SIA standard deviations across members, resamplings, and observations.** Probability distribution functions (PDFs) for detrended standard deviation of pan-Arctic SIA, for March (a-f) and September (h-l). PDFs are produced from the mean (μ) and standard deviation (σ) across the 10000 resamplings. Each individual resampled member (σ_{mem}) is plotted with a thin lines colored according to the legend, the average resampled member ($\bar{\sigma}_{mem}$) is colored similarly with a thick line, and the resampled observations (σ_{obs}) are in red for the four datasets according to the legend. Percentiles noted on the figure are the single values of σ_{obs} or μ_{obs} for the observational datasets relative to the distribution of σ_{mem} and μ_{mem} across members.

2.3.3 Ensemble size

We have included SMILEs with ensemble members as low as 20 in our analysis as the standard deviation between members (σ_{LE}), representing the full range of internal variability, increases only marginally beyond approximately 8-12 members, compared to the full range of 20-100 members (see Figure S2.2). This leads us to consider SMILEs of at least 12 members to generate enough diversity between realizations to capture most aspects of internal variability. The selection of a minimum number of members for SMILEs when assessing different time periods or other aspects of the climate system may require considerably more members (Milinski et al., 2020). With increasing ensemble size, the values of the minimum and maximum σ_{mem} diverge, making it easier for a SMILE to overlap with observations (see Figure S2.1). Our primary goal is to assess SMILEs' individual realism when compared with observations, using as much information from each model as is available. Hence, we present results without subsampling the members to a consistent ensemble size. However, as others may be interested in a direct comparison of the interannual variability in CMIP5 SMILEs, we provide subsampled results in the supplementary section, where consistency is standardized to 20 members, the size of the smallest SMILE, in the pan-Arctic (Figure S2.5) and spatially (Figure S2.6).

2.3.4 Detrending

The ensemble mean of a SMILE is considered a good representation of 'forced response' of the model to the changing climate (Frankcombe et al., 2018). However, observations only have one realization, and hence the observed forced trend must be computed from that single realization. Hence, in our analysis we use the individual members' trends over the period 1979-2020 as representation of the forced response, to enable the same methodology to be applied to observations and models, for direct comparisons. The SMILEs provide the perfect place to test the impact of this method: we find that linear detrending rather than removing the ensemble mean, results in only a marginal decrease in variability (8% reduction for σ_{mem} and 11% for σ_{LE}) yielding a very similar

ratio (see Figure S2.11).

Applying linear detrending largely removes low frequency variability. We reached this conclusion as detrending ensemble members and observations using a 2 year 5th order lowpass Butterworth filter (Roberts and Roberts, 1978), which explicitly removes low-frequency variability, obtains almost identical consistency results as with a simple linear trend (see Figure S2.7 in comparison to Figure 2.8). This lowpass filter removes variability on frequencies in excess of 2 years, the time period beyond which autocorrelation in the sea ice is negligible. Good agreement between the linear detrending and the lowpass filtered data suggests that both anomaly calculation methods effectively isolate interannual variability. The variability in our resampled anomalies of individual SMILE members (σ_{mem}) capture approximately three quarters of internal variability across SMILE realizations without resampling (σ_{LE}), as discussed further in section 2.4.3. This enables us to conclude that our detrending and resampling analysis primarily assesses interannual variability, and that this is the dominant timescale of internal variability for Arctic sea ice for the period 1979-2020.

In the spatial analysis we obtain a linear trend for each grid cell, using the same method of detrending as we did for the pan-Arctic. While we find some isolated incidences of grid cells where the linear SIC trend exceeds 100% or is lower than 0%, extremely small differences are found in consistency if a different detrending method is used, such as a 2 year low pass filter (see Figure S2.7) or trends capped to physical bounds (not shown). Hence, the detrending method does not affect the conclusions drawn from the analysis.

2.3.5 Time periods

The time period considered is the observational period 1979-2020, focused on the seasonal extremes of March and September for the spatial analysis. 1979-2020 is chosen for observations due to high quality spatial data from 1979 onward, which is particularly important for assessing interannual sea ice variability. We found that shifting the time period used from the models to

better match the observed mean sea ice state yielded negligible differences spatially and minimally affected pan-Arctic results for shifts of a few years to a decade. When matching the observed mean state required adjustments of many decades, the changes in the results were larger. However, in some instances a model did not have a time period where the mean state matched the observed mean state in the whole historical and future simulations. Furthermore, we want to assess the realism of the simulated interannual variability as simulated, to complement previous model assessments of trends and mean state which were done over the same periods in models and observations (e.g. Swart et al., 2015; Notz and SIMIP-Community, 2020). Hence, although internal variability has been shown to be sensitive to mean state (Goosse et al., 2009; Jahn et al., 2016; Olonscheck and Notz, 2017; Massonnet et al., 2018), and some models have more linear SIA and SIC declines than others over 1979-2020, we find the choice of exact period analyzed did not materially impact our results.

As the use of a 42-year time period is out of necessity, this raises the question of whether a 42-year period is sufficient for our analyses. To assess this question different time period lengths were assessed within the models (see Figure 2.3). Time periods longer than approximately 20 years yield similar ratios in $\sigma_{mem}/\sigma_{obs}$ ratios, which gives confidence in our results for this metric being representative of a broad range of time periods. Similarly, the ratio σ_{mem}/σ_{LE} changes rapidly for short time periods but becomes relatively stable for time periods of at least a few decades. To confirm this, we conducted a similar time period analysis for the period 1953-2020 using lowpass filtered SIA. This more clearly indicates the stabilization of the ratio σ_{mem} to σ_{LE}, at approximately 75%, independent of the length of the time period in excess of approximately 30 years (see Figure S2.3). Spatially, when we compare the shorter 32-year time periods 1979-2010 and 1989-2020 to the full time period of 1979-2020, we find that there are small consistency differences between the time periods for some regions, but these differences are not substantial enough for our main conclusions to be altered (see Figure S2.4).

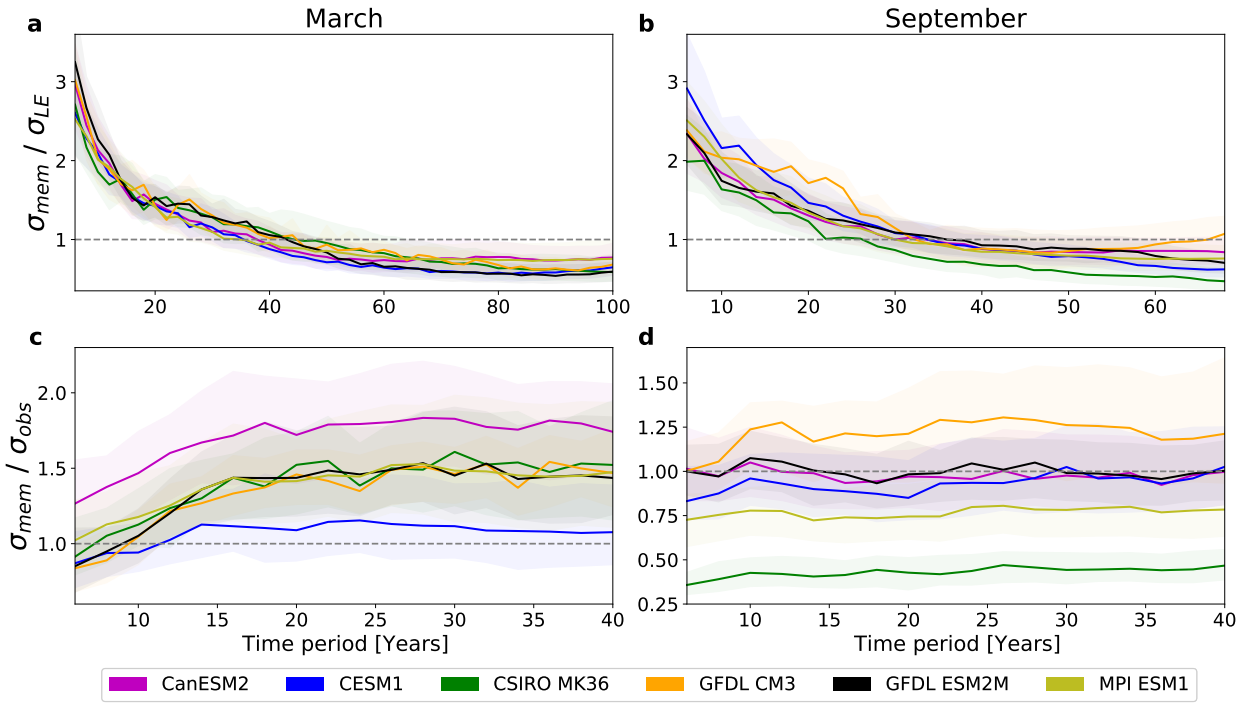

Figure 2.3: **Influence of length of time period on standard deviation of pan-Arctic SIA**. Standard deviation with respect to time for time periods between 6 years and the maximum length of a linear trend in SIA, bootstrapped 1000 times. Thick lines show the median ensemble member, shading shows ± 1 standard deviation. (a, b): the ratio of standard deviation across resamplings (σ_{mem}) to standard deviation across members (σ_{LE}) over a subset of the time periods 1965-2066 for March (a) and 1970-2040 for September (b). (c, d): the ratio of standard deviation across resamplings (σ_{mem}) to standard deviation across resampled observations (σ_{obs}) in the HadISST1 dataset for the period 1979-2020 in March (c) and September (d).

2.4 Results

2.4.1 Resampled variability in models and observations

Resampling the observations and SMILE models, we find that the variability of models is generally similar to observations, but with considerable seasonal and regional variability. Both the variability in models (σ_{mem}) and observations (σ_{obs}) show distinct seasonality in the pan-Arctic, peaking in the autumn with the exception of CSIRO MK3.6 (see Figure 2.4 and shown for average variability (μ) in Figure S2.8). In spring we find larger variation between different realizations of the same model than between model averages. This highlights the sensitivity of interannual variability to realization, and why we assess realism based on consistency rather than comparison between median SMILE member and observations (see section 2.3.2). The results of this consistency assessment are discussed further in section 2.4.2.

Observations have substantial uncertainties that impact the value of observational interannual variability (σ_{obs}). Hence, the choice of which dataset to use for comparison with models can affect whether observations fall within the large ensemble range, both for the pan-Arctic and spatially (see Figures 2.4 and 2.6). Furthermore, the uncertainties vary seasonally, with the largest relative uncertainty of pan-Arctic observational variability in the winter and spring (see Figure 2.5). Hence, it is easier for models to fall within the observational uncertainty in the winter and spring than in the summer and autumn. For most months we find the majority of the ensemble median variability ($\bar{\sigma}_{mem}$, grey bars in Figure 2.4) are similar or higher than observations (σ_{obs}, in red). However, as we do not know how typical observations are, we cannot use these differences to diagnose model biases.

Spatially, there is considerable difference in the locations of maximum variability between models and the observational datasets in September (see Figure 2.6). We find large magnitude differences throughout the ice covered region between different models and when comparing models with observations. Despite these large differences in the ensemble medians between models, we

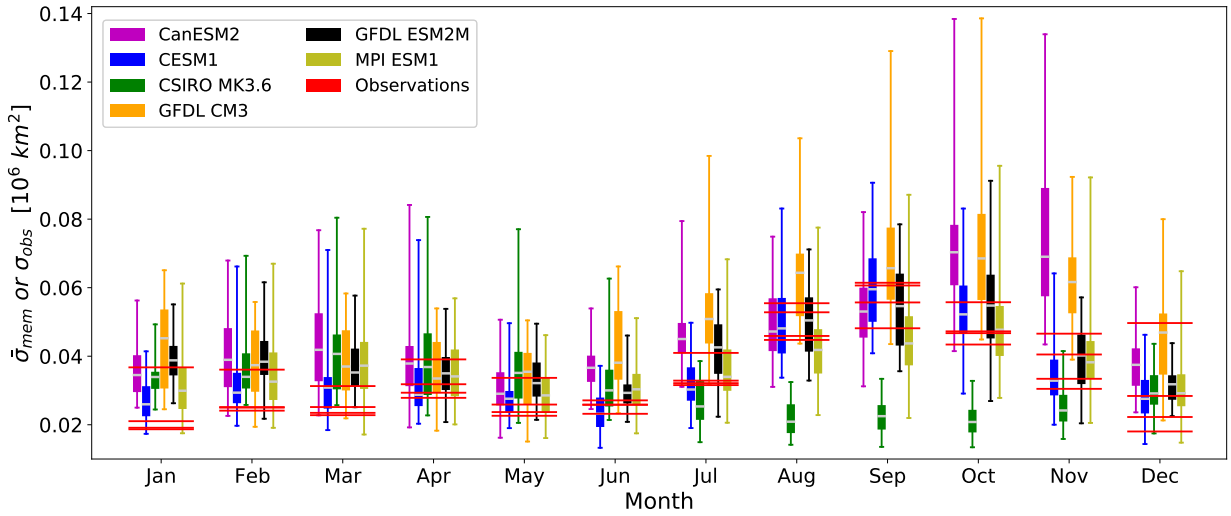

Figure 2.4: **Seasonality of resampled variability in members and observations for pan-Arctic sea ice area.** The distribution of Standard deviations (σ_{mem}) across members is shown for each model and month as a box and whisker charts, where whiskers show the full range of members, boxes show the interquartile range, and gray bars indicate the median member. Values of resampled variability in observations (σ_{obs}) are shown as horizontal lines for each of the four datasets.

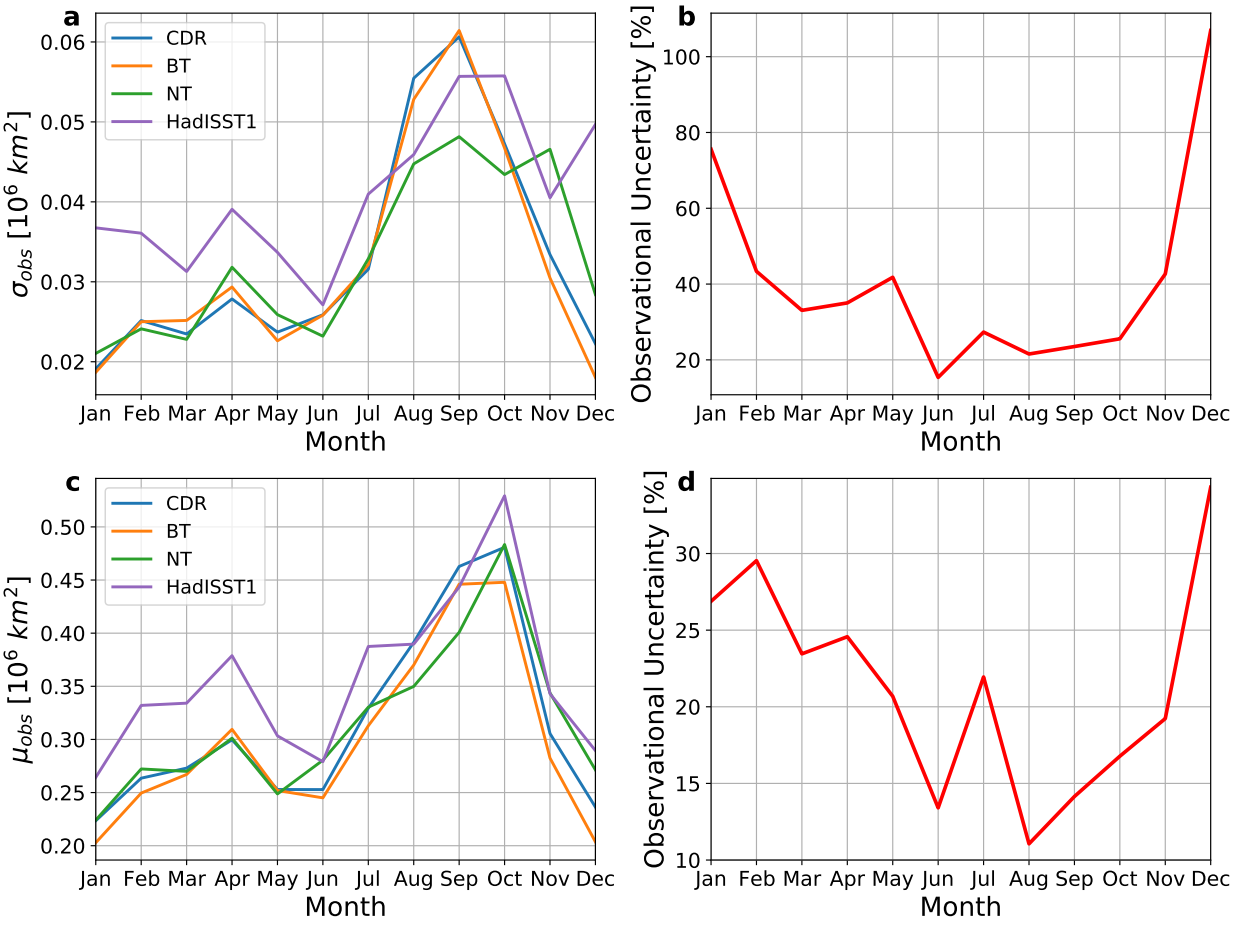

Figure 2.5: **Resampled variability of pan-Arctic sea ice area for the four observational datasets.** Absolute values shown in a) and c), and percentage uncertainty shown in b) and d) as calculated from the range of σ_{obs} divided by the mean of σ_{obs}.

find that the range between members for a given model is considerably larger in most instances. Again this draws attention to the difference of interannual variability between realizations. In comparison to September, the location and magnitude of highest variability in March is more similar between different models, with the range between members very large for the ice edge region (see Figure S2.9). Observational uncertainty is also highly variable between regions, for example NT exhibits much higher variability in the central Arctic in September than the other datasets (see Figure 2.6). When we combine both the spread of model simulations across realizations and the spread of interpretations of the observational record, we find broad agreement between models and observations. This is true both in the pan-Arctic and spatially in their representation of Arctic sea ice interannual variability.

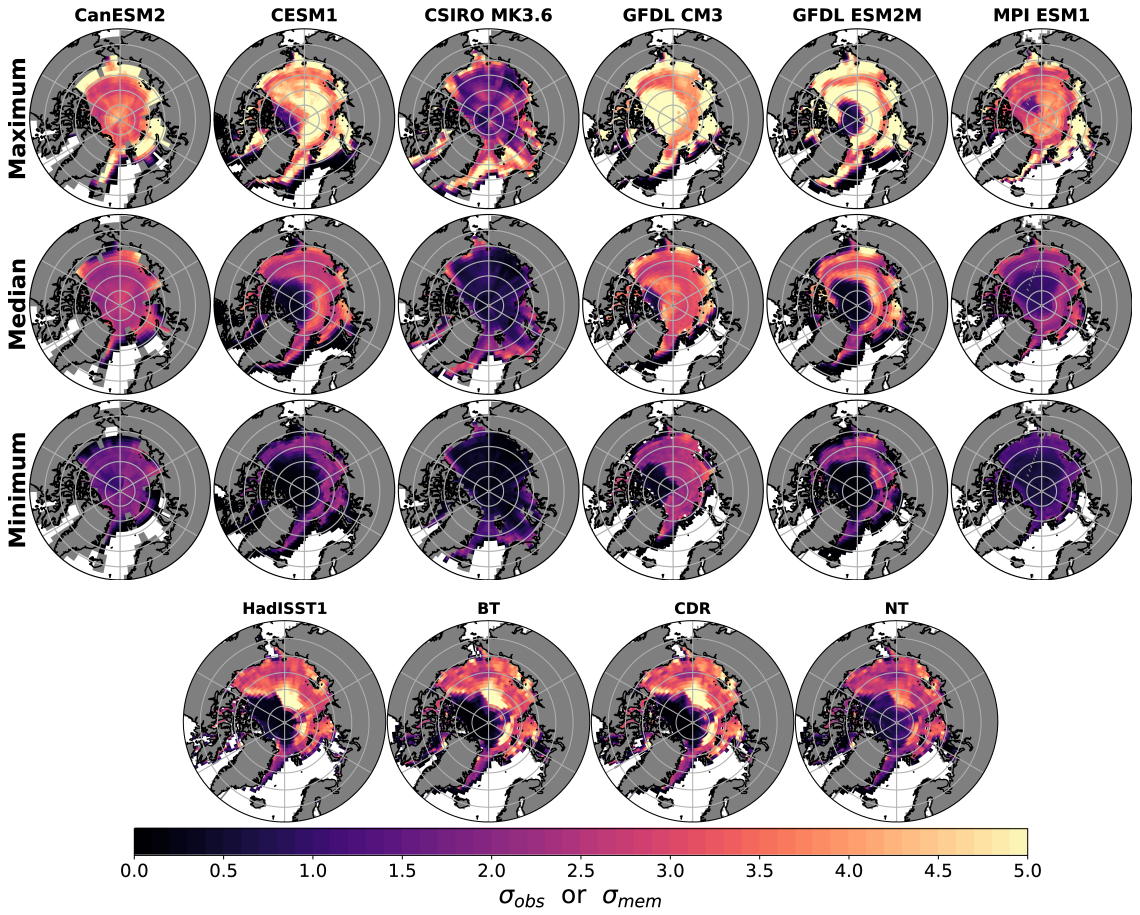

Figure 2.6: **Resampled modeled and observed variability of September sea ice concentration.** Standard deviation of resamplings for the six models (σ_{mem}) for the maximum, median and minimum member for each grid cell in rows 1-3. Standard deviation of resamplings for the four observational datasets (σ_{obs}) along the bottom row. The color bar applies to all subplots on this figure. The same analysis for March is shown in Figure S2.9.

2.4.2 Consistency of models and observations

When utilizing the range of observational datasets for the pan-Arctic, we find model consistency for a majority of the time (57%) across models and months (see Figure 2.7i). Models consistent within observational uncertainty accounts for 33% of months, far greater than the 10% of months identified as inconsistent. It is important to note that these proportions relate to the specific six models we analyzed, which capture the full spread of the CMIP5 sea ice simulations (Bonan et al., 2021). Nonetheless, the common pattern is for GCMs to be predominantly consistent within observational uncertainty. By our definition of consistency, all models except CSIRO MK3.6 and GFDL CM3 are consistent in September for all observational datasets. In the spring, when observational uncertainty is largest, we find all models are consistent within observational uncertainty and in April and May all models are consistent with all observational datasets. When looking across all months we find that only MPI ESM1 is unambiguously consistent with all observational datasets and CESM1 and GFDL ESM2M are consistent but not for all observational datasets. CanESM2, CSIRO MK3.6 and GFDL CM3 (the models with the largest mean-state bias) are the only models with inconsistent classifications beyond observational uncertainty. Our ability to more stringently assess realism by using the two metrics is demonstrated by CanESM2 and GFDL CM3 being considered consistent for all months for σ, but when also considering μ, we find both have two months with inconsistencies.

When considering consistency spatially, each grid cell can be considered to have a distribution of PDFs similar to Figure 2.2 and thus can be categorized in the same way. Consistency in σ and μ are highly correlated but with some differences, indicating the benefit of using both metrics (areas of light blue and light red in Figure 2.8). As noted earlier, we focus on the seasonal minimum and maximum sea ice area, September and March respectively and present a consistency classification only where both the model and observations exhibit non-zero sea ice.

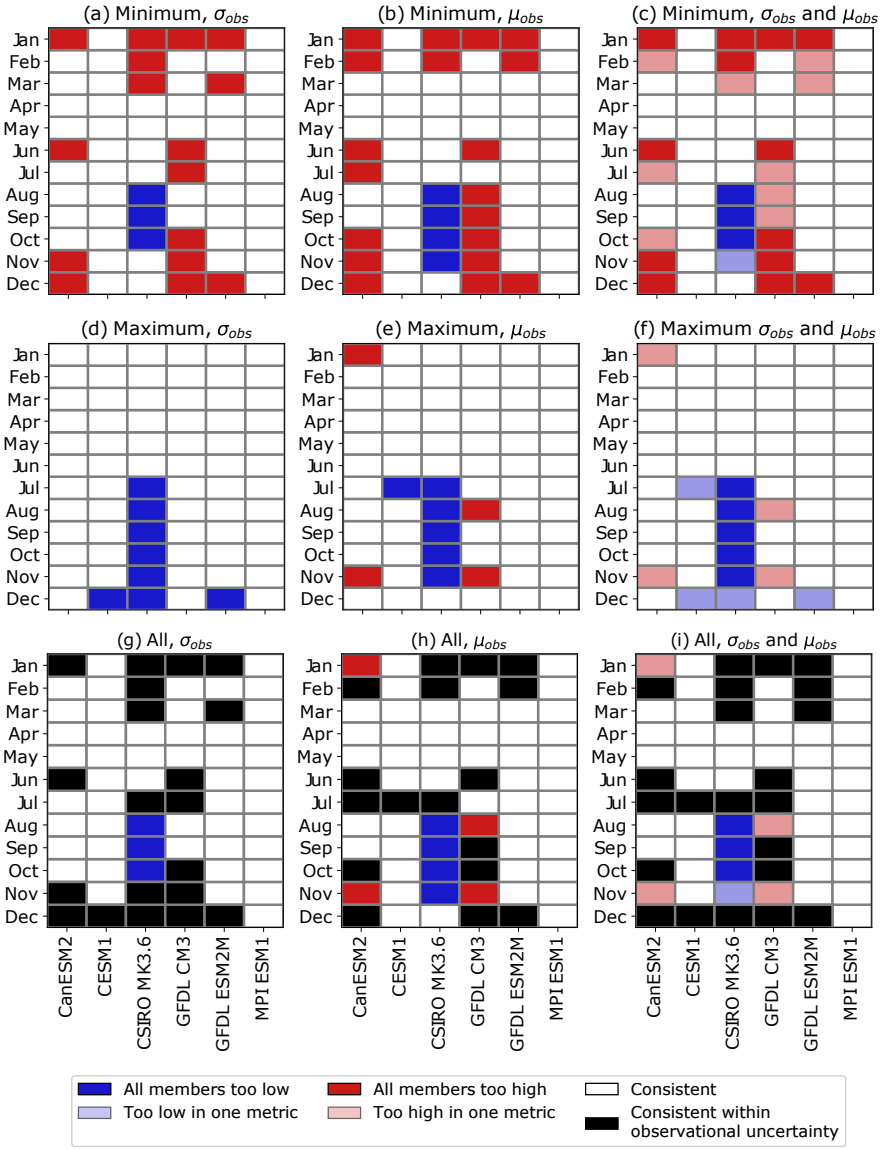

Figure 2.7: **Consistency between models and observations in pan-Arctic SIA**. White indicates consistency between models and all observational datasets, while reds and blues indicate inconsistency in at least one metric. Specifically, dark blue indicates the model is inconsistent with observations as all members are too low while dark red indicates inconsistency due to all members being too high. In the third column (c, f, i), where two metrics are combined, light blue means one of the metrics classifies the model as too low while the other metric is consistent, light red indicates that the model is too high in one metric but consistent in the other metric. There are no instances of too high and too low classifications for a given month by the different metrics. Finally, in the bottom row (g-i), where all observational products are combined, black indicates disagreement in classification between the observational datasets, indicating consistency within observational uncertainty.

Similarly to the pan-Arctic, we find no areas where the σ and μ metrics produce different signs of inconsistency. With the exception of CSIRO MK3.6, the shelf and marginal seas in September in all models are broadly consistent within observational uncertainty, with CESM1 and GFDL ESM2M performing the best. CSIRO MK3.6 shows the largest inconsistencies in March with underestimation of variability in the Barents Sea. All other models simulate consistent variability in the Barents Sea where atypically rapid SIC decline has occurred (Li et al., 2017). Both regions of too high variability and too low variability occur for MPI ESM1 in September, yet this model is consistent for September in the pan-Arctic, indicating these regions counteract each other for SIA. For March the models are more dissimilar than in September, with no regions of over or underestimation of interannual variability common to all models. Large portions of the central Arctic Ocean have very little observed and modeled variability in March, due to the 100% bounding of SIC. This means that small absolute biases in the modeled interannual variability can cause an inconsistent classification (see Figure S2.9). With our consistency classification we conclude that more models have greater realism of simulated interannual variability in September than in March. However, even well performing models in some regions in September or March generally do a poorer job in the other month, indicating that the skill of a certain model in simulating interannual variability is highly seasonally and regionally dependent.

2.4.3 Internal variability captured by resampling versus ensemble spread

Our best estimate of the full range of internal variability, on high and low frequency time scales, is through SMILEs, here we use the standard deviation between detrended members (σ_{LE}) to represent this. As we consider the resampled standard deviation of SMILE members and observations to be representative of interannual variability and not the full range of internal variability, we would expect the ratio σ_{mem}/σ_{LE} to be less than one. For all seasons, when looking at pan-Arctic SIA, interannual variability simulated by the median standard deviation across resamplings ($\bar{\sigma}_{mem}$) is less than the internal variability simulated by multiple realizations without resampling (σ_{LE}), an annual average of 75.9% across models (Figure 2.9). This ratio is robust irrespective

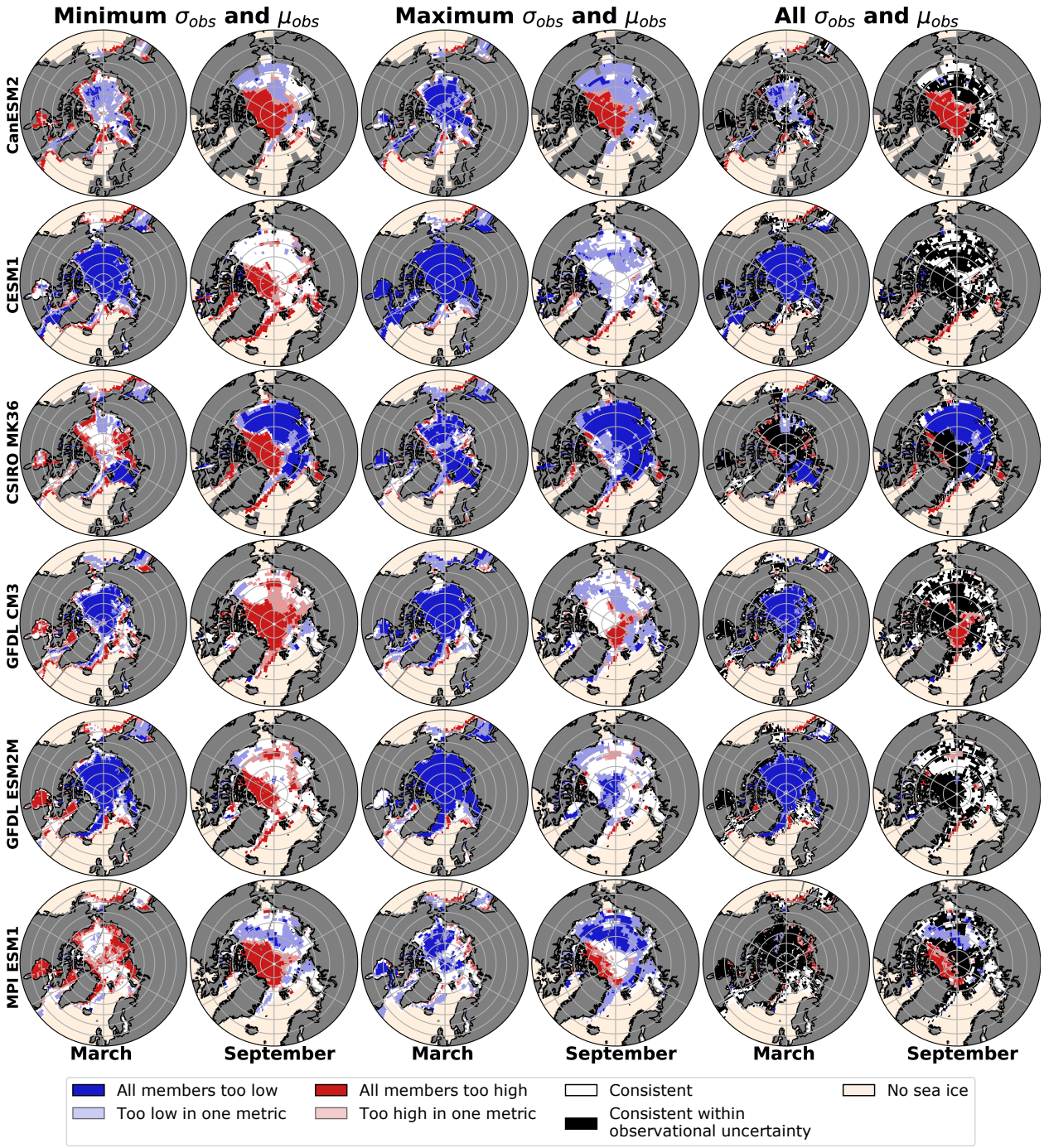

Figure 2.8: **Spatial consistency of interannual variability between large ensemble members and observations**. Members of the large ensembles which have at least one member overlapping with the variability of resampled observed SIC is shown in white, indicating consistency. Regions where the classification differs between the maximum and minimum observational datasets are shaded black indicating consistency within observational uncertainty. Areas without sea ice, either in the model or observations, are shaded beige. Shaded areas of red and blue indicate inconsistency in at least one metric, using the same color scheme as in Figure 2.7.

of detrending method with an average of 74.4% and 82.4% when the ensemble mean or a 2 year lowpass filter respectively is used for detrending (see Figure S2.11).

This ratio of three quarters interannual variability and one quarter lower frequency variability also holds for different time period lengths, as discussed in section 2.3.5, and is relatively stable for a given 42 year time period sometime between 1950-1991 and 2050-2091 (see Figure S2.12). Hence, we expect interannual variability to remain the dominant portion of internal variability for the near future. The general underestimation of the resampled variability, compared with the benchmark of large ensemble spread, is in agreement with previous uses of this methodology on surface temperature, precipitation and sea level pressure (McKinnon et al., 2017; McKinnon and Deser, 2018). When considering the difference between σ_{LE} and σ_{mem} spatially, we find the largest underestimations along the ice edge but in general the signal in the pan-Arctic is replicated homogeneously across the Arctic (see Figure S2.10).

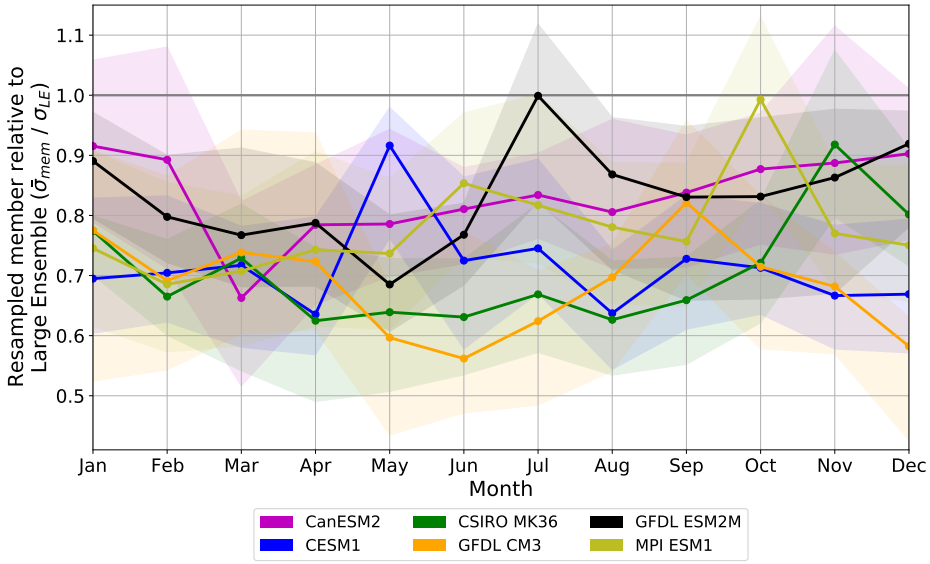

Figure 2.9: **Seasonality of the ratio of internal variability across SMILEs and interannual variability of resampled members for pan-Arctic sea ice area.** Lines show the ratio of the standard deviation of the median resampled member to the standard deviation across members without resampling ($\bar{\sigma}_{mem}$ to σ_{LE}), shading shows the interquartile range of the ratios for all members.

2.5 Discussion

Sea ice poses unique challenges in assessing internal variability: short time period of high quality observations, physical bounds of 0 to 100%, and changes in variability as mean state changes. Despite this, we were able to apply the synthetic ensemble method to Arctic sea ice as used in McKinnon et al. (2017) and McKinnon and Deser (2018) for temperature, precipitation and sea level pressure. Similarly to previous research, we found that resampling leads to an underestimation of the full range of internal variability captured by a large ensemble, both in the pan-Arctic (where $\bar{\sigma}_{mem} \approx 0.76\ \sigma_{LE}$, see Figure 2.9), and also locally across the Arctic Ocean (Figure S2.10). This agrees with the expectation that low frequency variability is not fully captured by the resampling (McKinnon and Deser, 2018). Hence our analysis primarily assesses the interannual component of internal variability. Interestingly, this proportion of three quarters of the internal variability being due to interannual variability matches closely with the 75% contribution from atmospheric temperature fluctuation to Arctic sea ice variability found by Olonscheck et al. (2019) via a 'decoupling' methodology. Both of these independent analyses hence suggest that Arctic sea ice interannual variability is largely unpredictable.

Our analysis assumes that a given anomaly is equally likely to have occurred in 1979 or 2020. This is a dependable assumption, despite the fact that it has been shown that variability increases as sea ice extent decreases (Goosse et al., 2009; Jahn et al., 2016; Olonscheck and Notz, 2017; Massonnet et al., 2018), as we showed that neither the length of the period considered (Figure 2.3) nor the period itself (Figure S2.12) substantially change the results. However, as the Arctic approaches seasonally ice-free conditions, an "equally likely" assumption will no longer be a valid approach. For example, it would not be appropriate to assume a September SIA negative anomaly of one million square km (as occurred in 2007) would be equally likely to occur when the mean state in September is practically zero in most models.

All of the SMILEs, except CSIRO MK3.6, capture the seasonal cycle of σ_{mem} and μ_{mem} with highest values in the summer. However, the magnitude of observational uncertainty also needs to be taken into account as it factors into how stringent consistent classifications are. Observational uncertainty is largest in the winter for the pan-Arctic (see Figure 2.5), therefore it is easier for models to be consistent during this part of the year. Spatially we find the largest differences in variability between observational datasets in the central Arctic during September (see Figure 2.6). Nevertheless, we still find that most models simulate too high variability in this region in September, and it is only the extreme variability of NT compared with the other observational datasets that allow a 'consistent within observational uncertainty' classification for most models (see Figure 2.8). Consensus regarding which observational dataset is the most realistic for these areas would be required before determining which models have the better representation of variability in the high SIC regions.

As we have shown that almost all models can simulate consistent members across seasons, we can say most of the SMILE models are realistic in their simulation of historical interannual variability. Realism of internal variability is a complementary assessment to the analysis of mean state, sensitivity to warming, and trends (Swart et al., 2015; Rosenblum and Eisenman, 2017; Winton, 2011; Niederdrenk and Notz, 2018; Davy and Outten, 2020). Some of these metrics are inter-related but each provide part of the picture for a full model assessment for Arctic sea ice. We show the CMIP5 models with inconsistent months or large regions of inconsistency are those with the largest mean state biases, but even these models are consistent for several months of the year in the pan-Arctic and for most regions in March and September. This suggests that avoiding mean state biases is important for correctly simulating the evolution of the Arctic sea ice cover (see Massonnet et al. (2018)), but models can have moderately large mean-state biases and still simulate realistic sea ice interannual variability. Furthermore, as we find that most CMIP5 SMILE models agree with observations in terms of their interannual variability for the pan-Arctic in September, the internal variability prediction uncertainty of an ice-free Arctic of over two decades from climate

models (Notz, 2015; Jahn et al., 2016) is likely realistic. However, no SMILE model performs well in all months and regions. But if one wishes to only focus on one season or region, one can find a CMIP5 SMILE model where the interannual variability is consistent with observations. This is true even for hotspots of internal variability such as the Barents Sea in winter and the shelf seas in summer (England et al., 2019; Bonan et al., 2021), showing the robustness of the consistency classification.

2.6 Conclusions

In this study, we showed that simulated interannual variability of CMIP5 large ensemble models is typically within observational uncertainty, by generating a synthetic ensemble of Arctic sea ice variability and using a binary classification of consistency that considers the full distribution of resamplings to aid the assessment of model realism. This analysis method considers approximately three quarters of Arctic sea ice internal variability, on the dominant interannual timescale for the period 1979-2020. Sea ice variability is another metric that augments the realism assessment of GCMs in the context of Arctic sea ice beyond the typical mean state and trend consistency and the assessment of sea ice sensitivity (Swart et al., 2015; Rosenblum and Eisenman, 2017; Winton, 2011; Niederdrenk and Notz, 2018; Davy and Outten, 2020).

We showed that all models are able to simulate the seasonal cycle of interannual variability with peaks in the summer, except CSIRO MK3.6 which has by far the largest mean state biases (see Table 2.4), caused by aerosol issues (Uotila et al., 2013). We demonstrate that all modeled interannual variability is within observational uncertainty, except for CanESM2 in January and November, GFDL CM3 in August and November, and CSIRO MK3.6 in August-October for the pan-Arctic. Except for areas of low absolute variability in the central Arctic Ocean, there are no inconsistencies that are common across all six models we assessed. Spatially, we find the models underestimate interannual variability for most regions in March, and in September most models

overestimate variability in the central Arctic. The marginal seas, which have high absolute variability, are generally realistically simulated, although our assessment is limited to where both models and observations have sea ice. No model simulated the spatial interannual variability in both March and September without inconsistencies, but most models simulated at least one of the two months realistically. CESM1 and GFDL ESM2M simulate September spatial variability very well, with very few areas of inconsistency, including the highly variable shelf seas. In March, MPI ESM1 performs best, with only the Siberian coast displaying too high variability.

In summary, in this first direct comparison of interannual variability between observations and models, we have shown that estimates of interannual variability from models are largely consistent with observations. However, model skill varies by month and region, highlighting that the best model to use for a study varies based on the context. To be able to assess the impact of the full range of internal variability, including the low frequency variability (McKinnon and Deser, 2018), first requires an improved understanding of the drivers of low-frequency variability on Arctic sea ice. Generally, the fact that the simulated interannual variability of most CMIP5 large ensembles agrees quite well with historical observations, especially in September, increases trust in the internal variability uncertainty of Arctic sea ice projections.

2.7 Acknowledgements

This work was supported by the National Science Foundation under grant number 1847398. We would also like to acknowledge high-performance computing on Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation.

2.8 Data Availability

All code required to replicate this study has been made open-access via Zenodo at https://www.doi.org/10.5281/zenodo.6687725. All data used in the analysis is already freely available. The CLIVAR Large Ensemble Archive can be obtained from the NCAR Climate Data Gateway (https://www.earthsystemgrid.org/dataset/ucar.cgd.ccsm4.CLIVAR_LE.html). The NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration (Version 4) is available from https://doi.org/10.7265/efmz-2t65. The Hadley Centre Sea Ice and Sea Surface Temperature dataset (HadISST) is available from https://www.metoffice.gov.uk/hadobs/hadisst/index.html.

Additionally, the summary statistics from the resampled synthetic ensemble (μ_{LE}, σ_{LE}, μ_{mem}, σ_{mem}, μ_{obs} and σ_{obs}), can be accessed from the Arctic Data Center at https://doi.org/10.18739/A2H98ZF3T.

2.9 Supplementary Material

Table S2.1: Replacement months for datasets with missing data for NSIDC or with discontinuities as for HadISST1.

Dataset	Discontinuity Month	Replacement Month
NSIDC CDR,BT,NT	1984-07	1985-07
NSIDC CDR,BT,NT	1987-12	1988-12
NSIDC CDR,BT,NT	1988-01	1989-01
HadISST1	2009-03	2007-03
HadISST1	2009-04	2008-04

Month	CanESM2		CESM1		CSIRO MK36		GFDL CM3		GFDL ESM2M		MPI ESM1		Model Mean	Model Mean Bias	NSIDC CDR
Jan	11.97	-1.42	13.53	0.13	14.88	1.49	11.90	-1.49	13.28	-0.11	12.01	-1.39	**12.93**	-0.46	**13.39**
Feb	13.36	-0.95	14.25	-0.06	15.69	1.38	12.91	-1.41	14.73	0.42	12.89	-1.42	**13.97**	-0.34	**14.31**
Mar	13.78	-0.65	14.39	-0.03	15.93	1.50	13.21	-1.22	15.06	0.64	13.11	-1.31	**14.25**	-0.18	**14.43**
Apr	12.99	-0.57	13.83	0.27	15.61	2.05	12.71	-0.85	14.25	0.69	12.56	-1.00	**13.66**	0.10	**13.56**
May	11.35	-0.64	12.30	0.30	14.88	2.88	11.41	-0.58	12.09	0.09	11.18	-0.81	**12.20**	0.21	**11.99**
Jun	8.69	-1.39	10.32	0.24	13.69	3.61	9.37	-0.71	9.53	-0.55	8.71	-1.37	**10.05**	-0.03	**10.08**
Jul	5.10	-2.62	7.93	0.21	11.50	3.78	6.57	-1.15	6.91	-0.81	5.84	-1.88	**7.31**	-0.41	**7.72**
Aug	2.75	-2.96	5.68	-0.03	10.11	4.41	4.44	-1.26	4.83	-0.88	4.01	-1.70	**5.30**	-0.40	**5.70**
Sep	2.44	-2.84	5.42	0.14	9.64	4.37	3.87	-1.40	4.63	-0.65	3.73	-1.55	**4.96**	-0.32	**5.28**
Oct	4.10	-3.03	7.25	0.12	9.77	2.64	5.37	-1.76	6.72	-0.41	5.44	-1.69	**6.44**	-0.69	**7.13**
Nov	7.04	-2.55	9.64	0.05	11.21	1.62	7.96	-1.63	9.47	-0.12	8.13	-1.45	**8.91**	-0.68	**9.59**
Dec	9.91	-1.87	11.82	0.04	13.32	1.54	10.18	-1.60	11.32	-0.46	10.26	-1.52	**11.13**	-0.65	**11.78**

Table S2.2: Mean SIA for the period 1979-2020 for all models and observations. SIA in 10^6 km^2, biases relative to CDR observations.

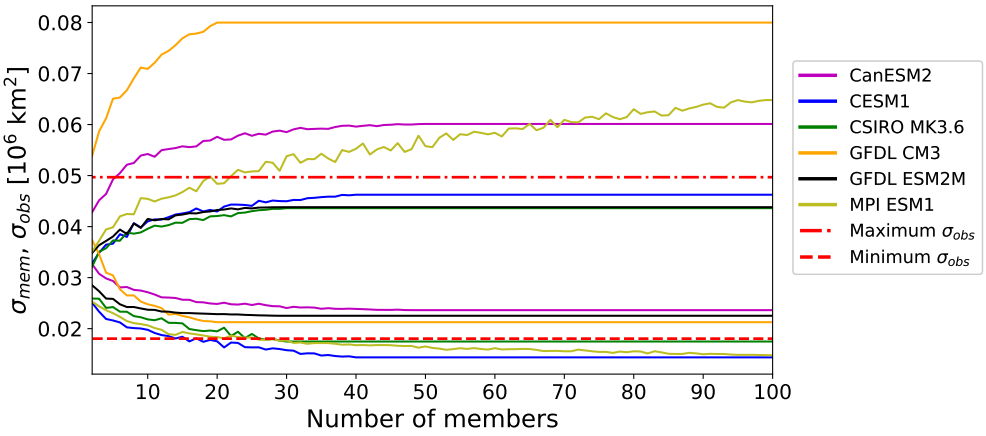

Figure S2.1: **Effect of ensemble size on maximum and minimum resampled standard deviations.** The maximum and minimum ensemble member for σ_{mem} when subsampled from 2-100 members 100 times is compared with the minimum and maximum values of σ_{obs} for December.

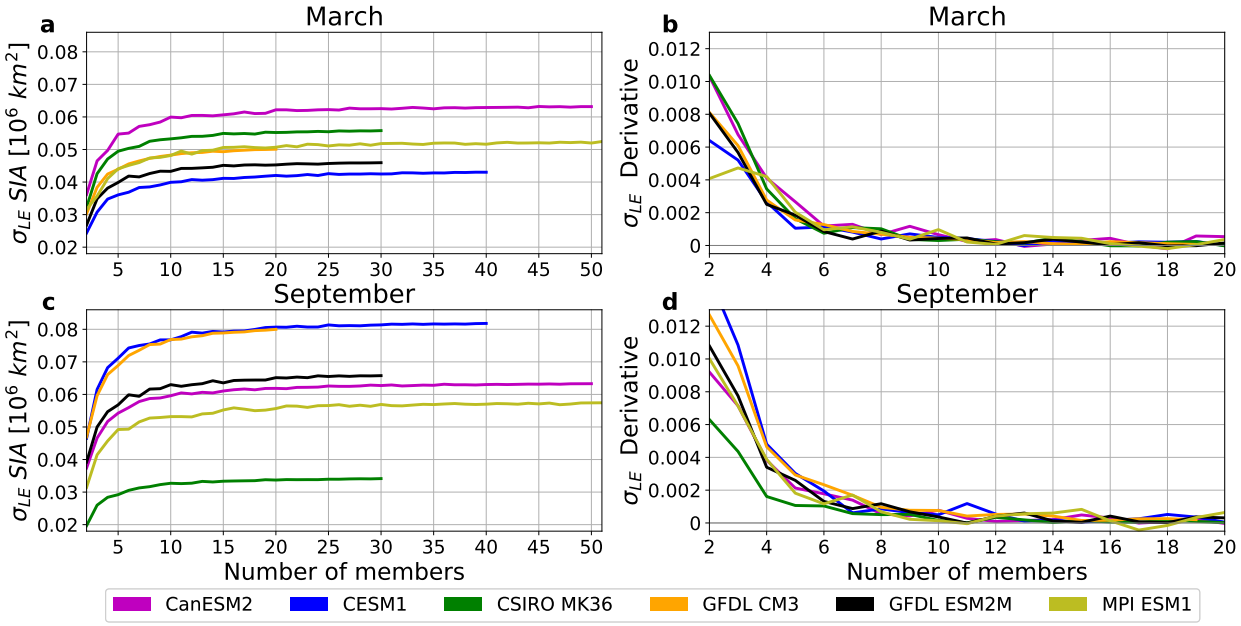

Figure S2.2: **Influence of ensemble size on standard deviation of pan-Arctic SIA**. Standard deviation with respect to time for a number of subsampled members between 2 and the lesser of 51 and the full ensemble for the time period 1979-2020. (a, c): average standard deviation across members (σ_{LE}) for 1000 bootstrapped members of a given number, for March (a) and September (c). (b, d): the derivative of the average standard deviation per member shown in (a) and (c).

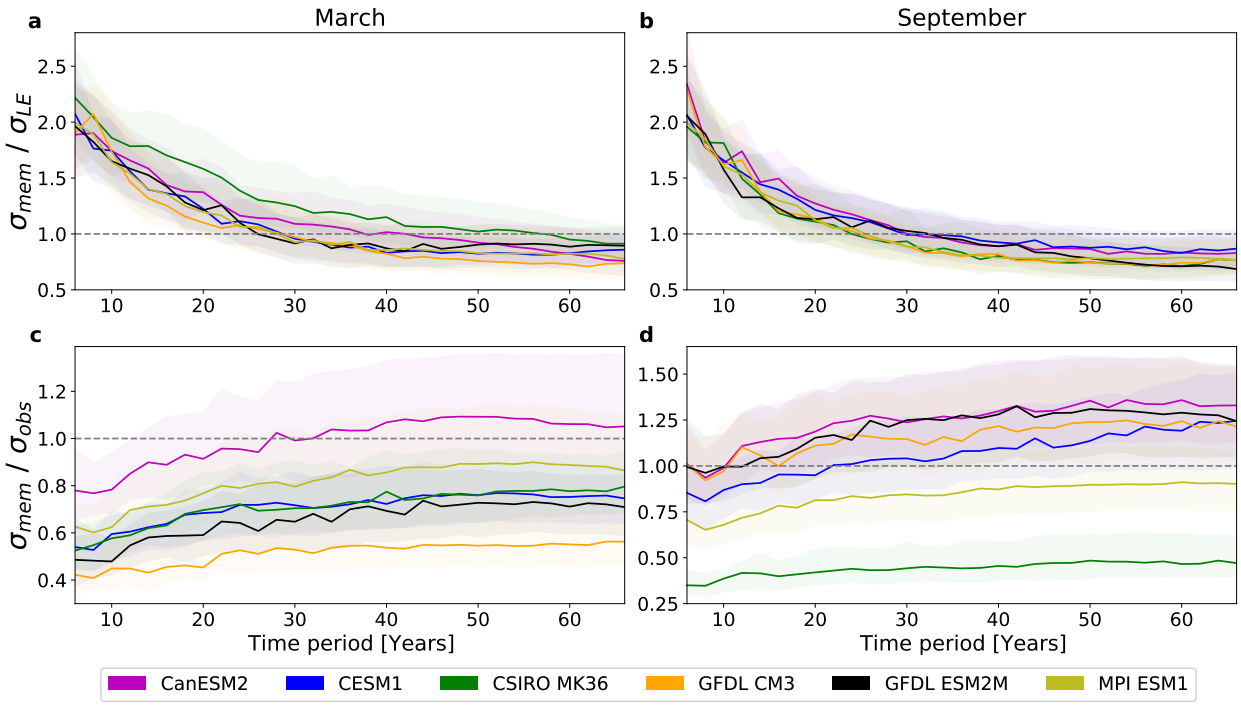

Figure S2.3: **Influence of length of time period on standard deviation of pan-Arctic SIA**. Standard deviation with respect to time for time periods between 6 and 66 years in SIA derived from anomalies relative to a 2 year lowpass Butterworth filter with 10 random start times for σ_{mem} and 1000 for σ_{LE} and σ_{obs}. Thick lines show the median ensemble member, shading shows $\pm$ 1 standard deviation. (a, b): the ratio of standard deviation across resamplings (σ_{mem}) to standard deviation across members (σ_{LE}) over a subset of the time periods for March (a) and September (b). (c, d): the ratio of standard deviation across resamplings (σ_{mem}) to standard deviation across resampled observations (σ_{obs}) in the HadISST1 dataset in March (c) and September (d).

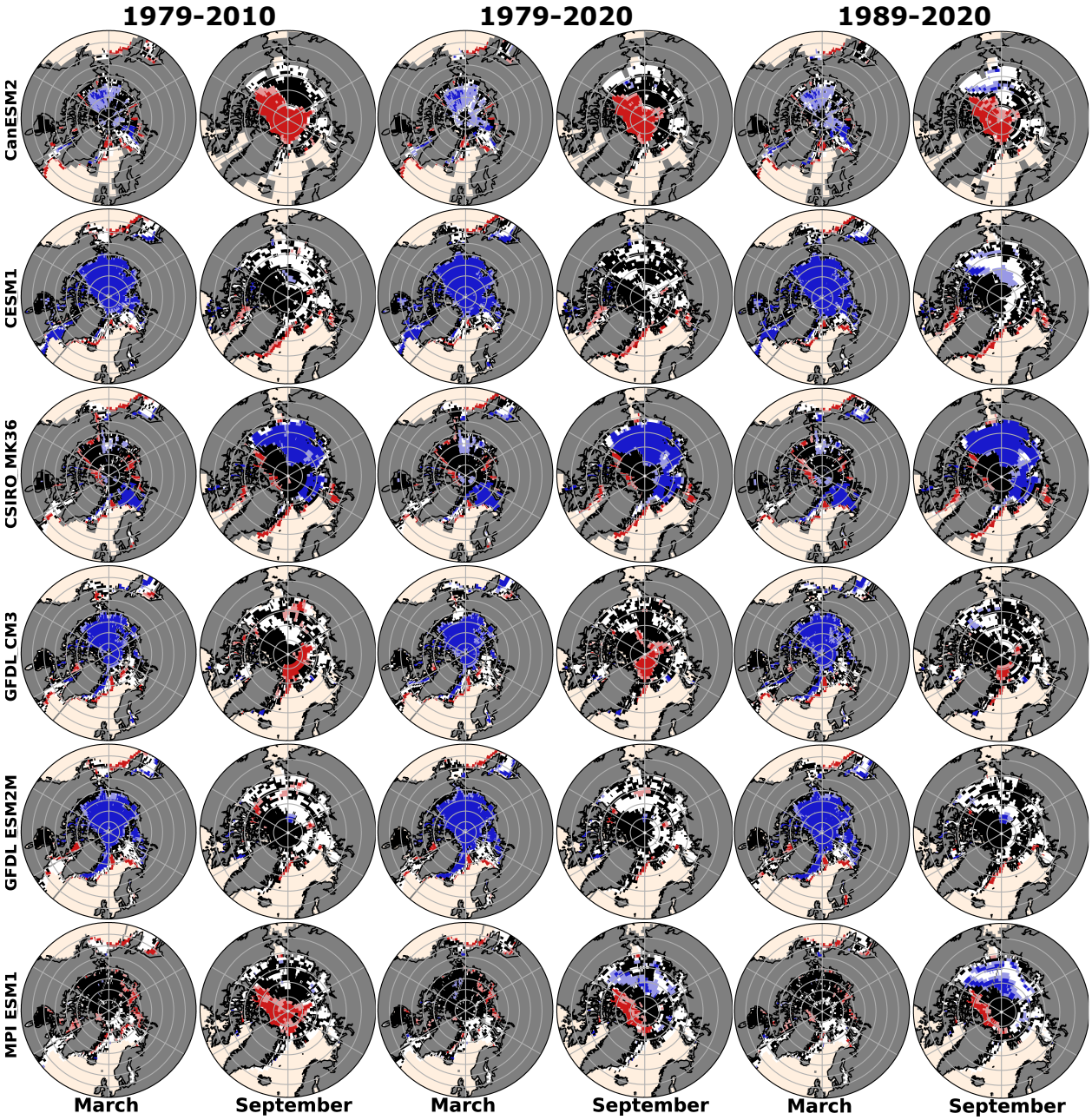

Figure S2.4: **Influence of time period on spatial consistency**. SIC is detrended using a linear trend. Same as 'All σ_{obs} and μ_{obs}' columns in Figure 2.8 for '1979-2020' columns. Model and observational data for 1979-2010 and 1989-2020 are shown in the columns labeled accordingly.

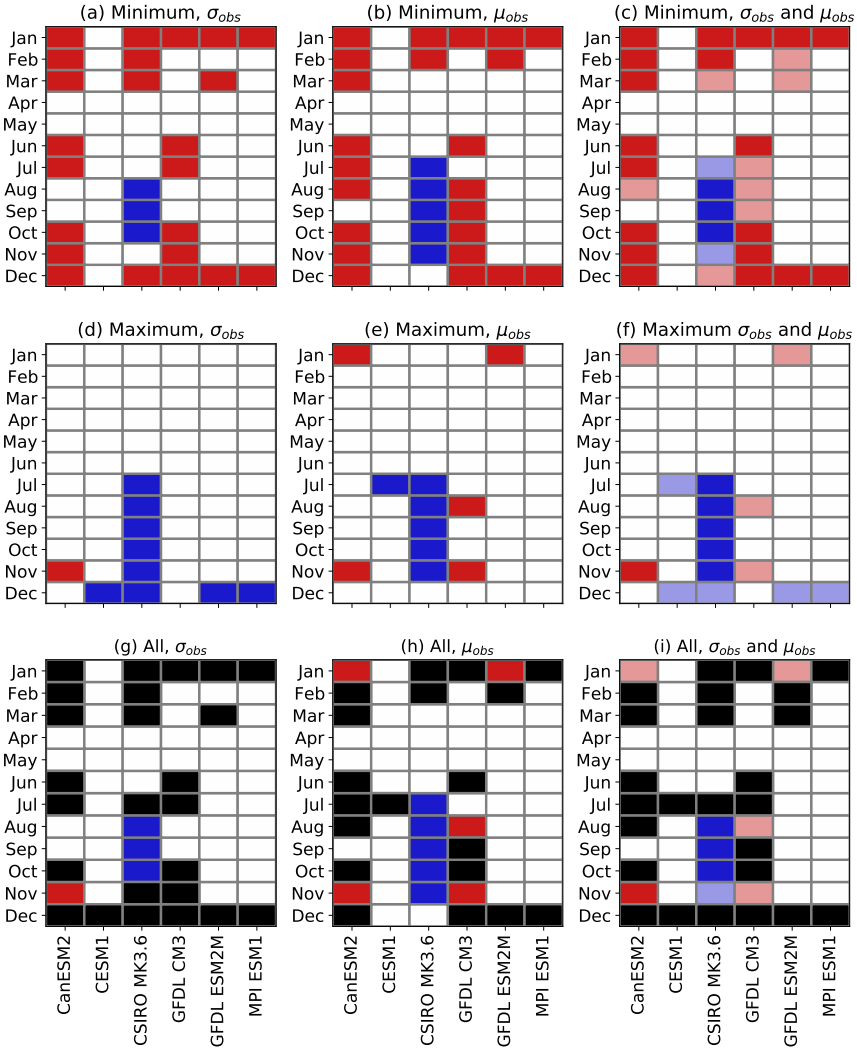

Figure S2.5: **Consistency between models and observations in pan-Arctic SIA, subsampled to 20 members**. Same as Figure 2.7, except now subsampled 1000 times to 20 members.

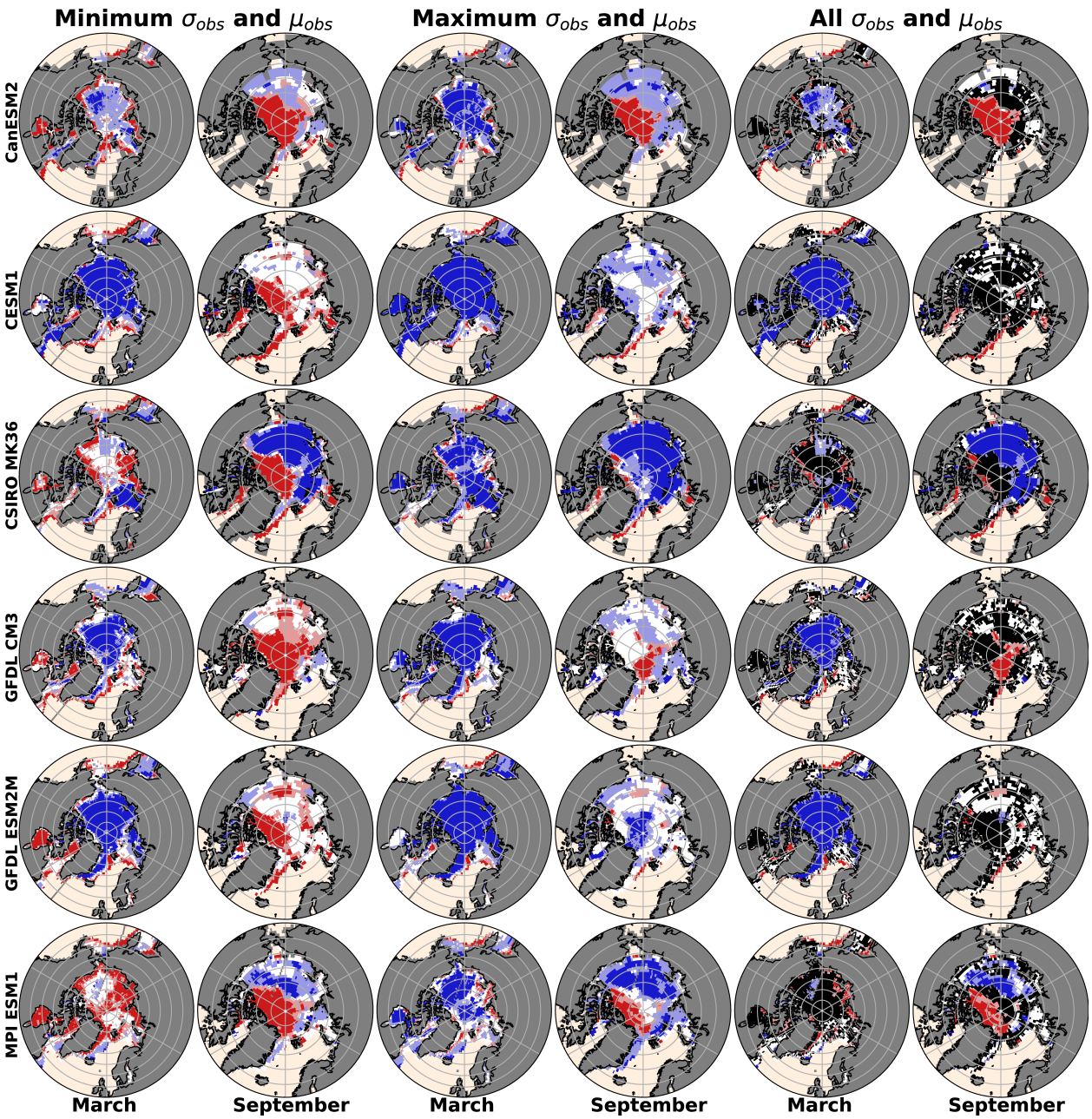

Figure S2.6: **Spatial consistency of internal variability between large ensemble members, subsampled to 20 members**. Same as Figure 2.8 except now members are subsampled 1000 times to a size of 20 members.

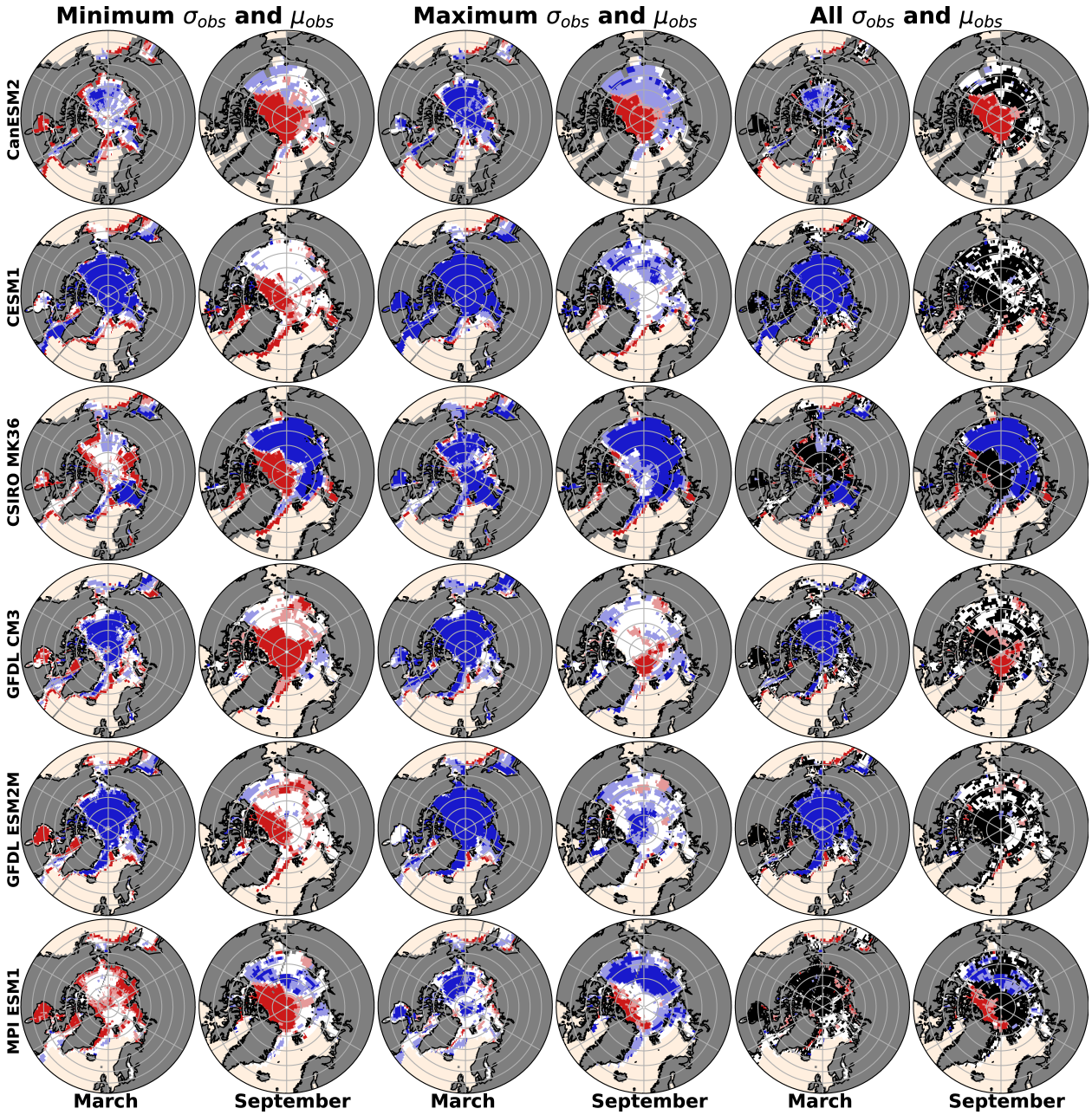

Figure S2.7: **Spatial consistency of internal variability between large ensemble members and observations using lowpass filtered data**. Same as Figure 2.8 except now SICs are detrended using a 2 year lowpass filter before resampling.

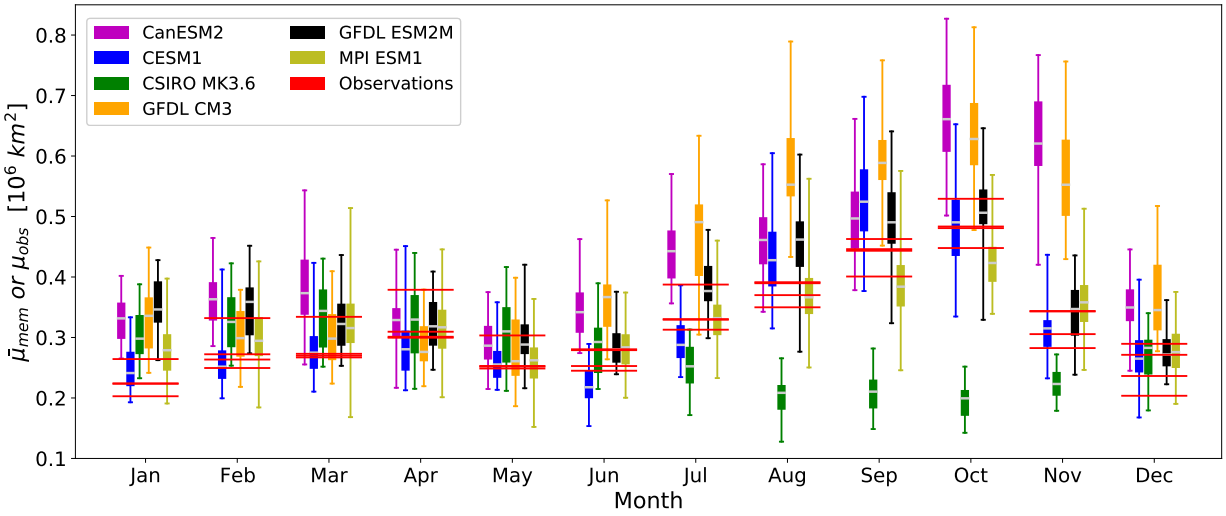

Figure S2.8: **Seasonality of average resampled variability in members and observations for pan-Arctic sea ice area.** Same as Figure 2.4, except now for μ instead of σ.

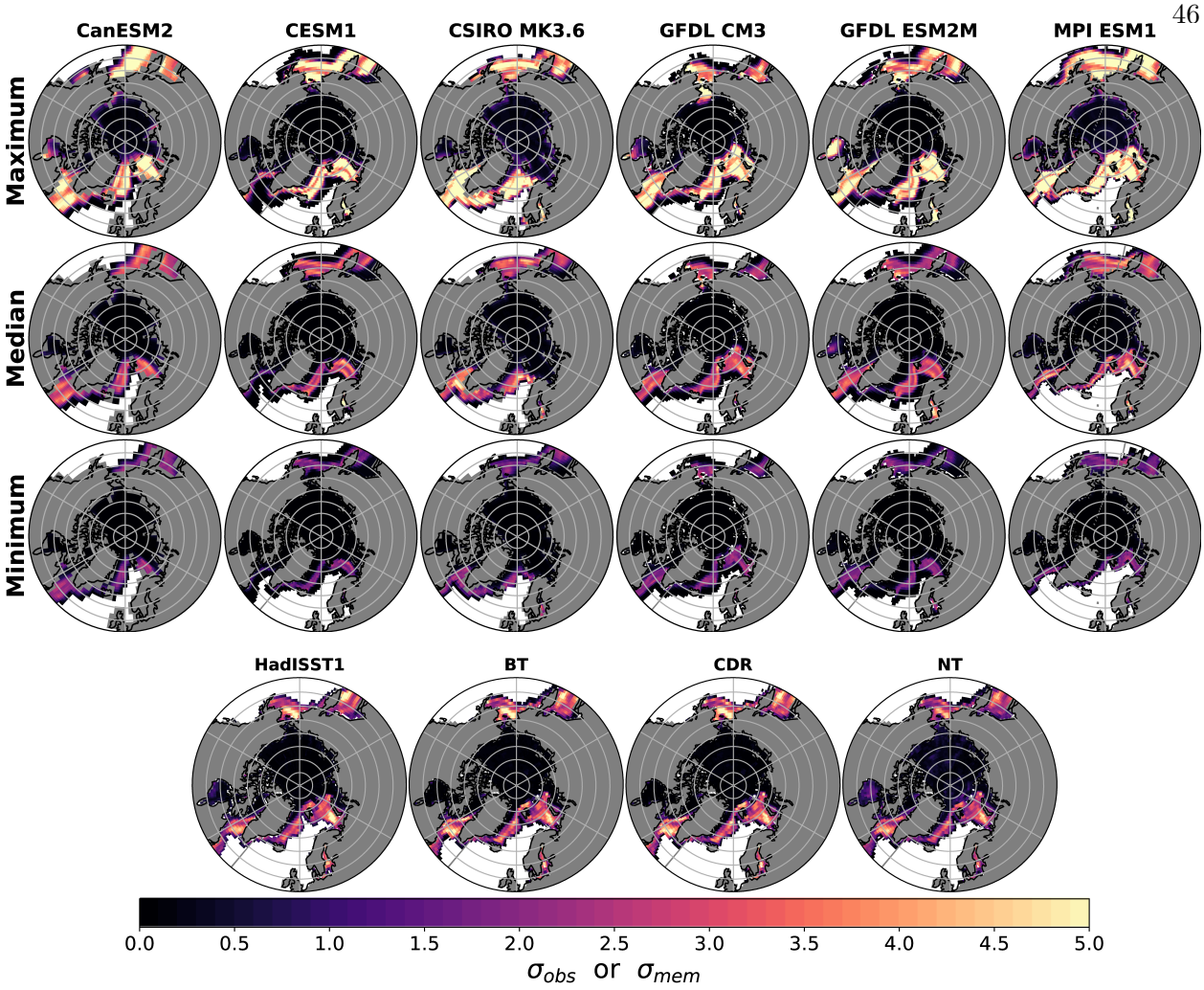

Figure S2.9: **Resampled modeled and observed variability of March SIC.** Same as Figure 2.6, but now for March.

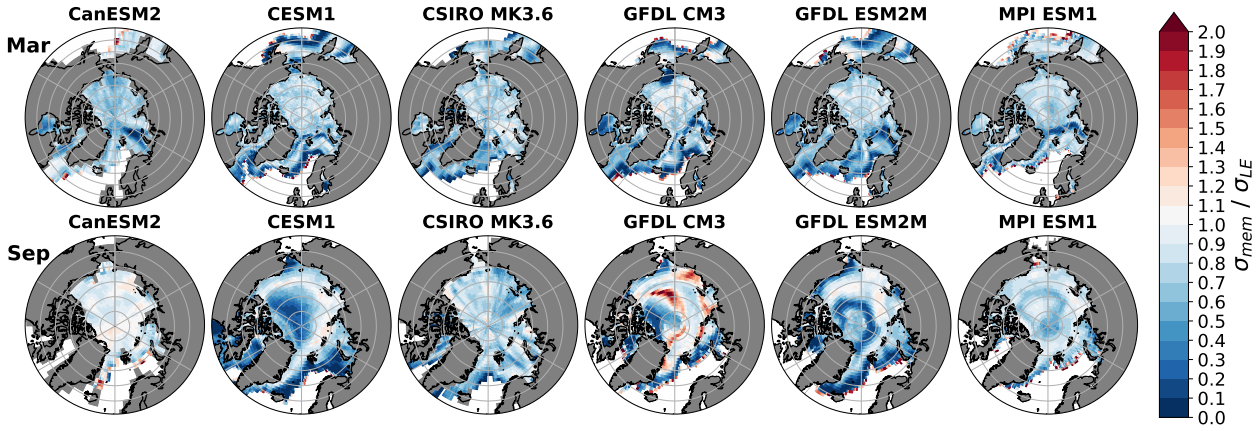

Figure S2.10: **Spatial distribution of the ratio of internal variability in large ensembles and resampled members.** The ratio $\bar{\sigma}_{mem}$ to σ_{LE} for March and September in all models represents the proportion of large ensemble variability captured via the resampling technique.

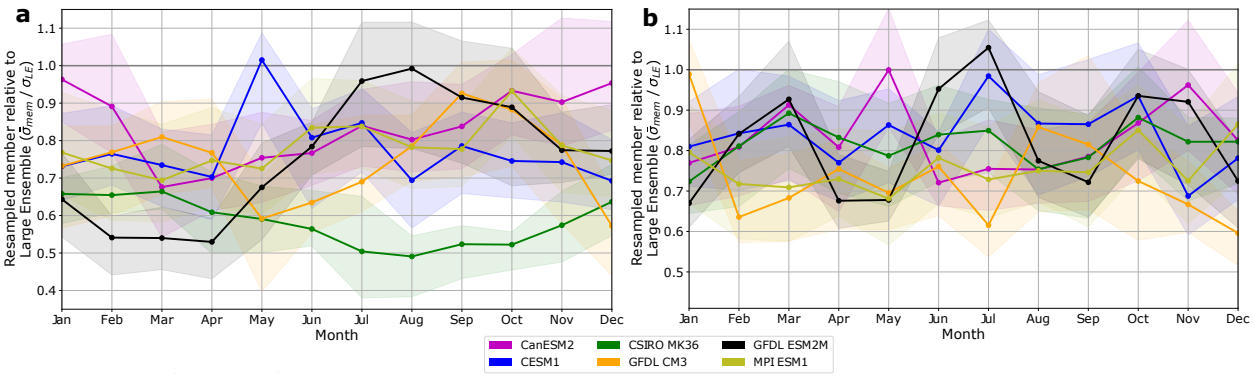

Figure S2.11: **Seasonality of the ratio of internal variability across SMILEs and inter-annual variability of resampled members for pan-Arctic sea ice area, using ensemble mean detrended data and lowpass filtered data.** Same as Figure 2.9, except now SIA anomalies were detrended relative to the linear trend of the ensemble mean for (a) and relative to a 2 year lowpass filter for (b).

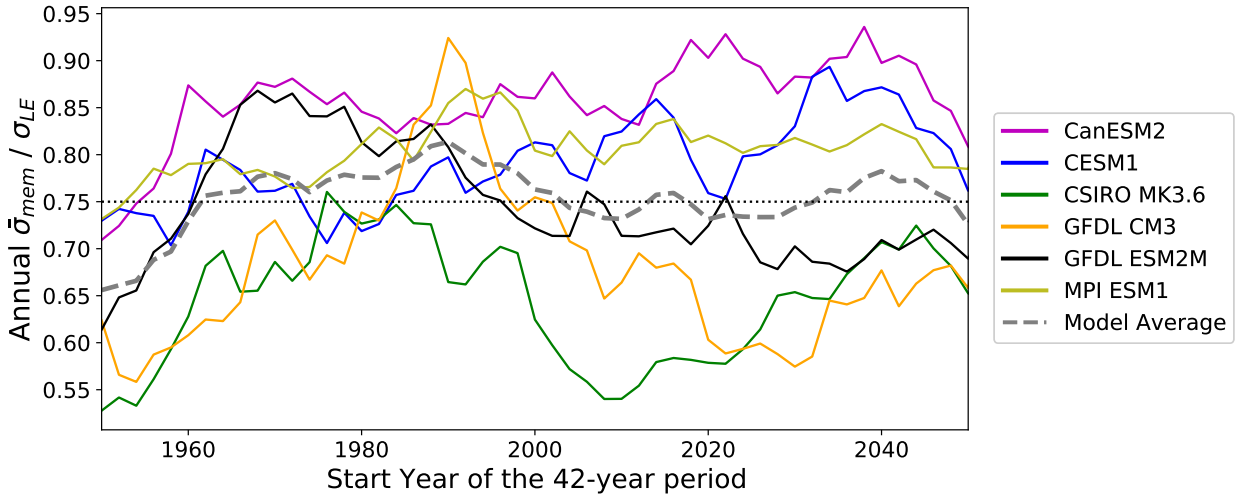

Figure S2.12: **Effect of starting year on $\bar{\sigma}_{mem}$ / σ_{LE} ratio for a 42-year time period.** The ratio of $\bar{\sigma}_{mem}$ to σ_{LE} is calculated for 42-year time periods between 1950-1991 and 2050-2091 for the annual mean. $\bar{\sigma}_{mem}$ is calculated from 1000 resamplings.

CHAPTER 3

LARGE-SCALE CLIMATE MODES DRIVE LOW-FREQUENCY REGIONAL ARCTIC SEA ICE VARIABILITY

Preface

This chapter is a manuscript being revised for Journal of Climate (manuscript number: JCLI-D-23-0326). The authors are Wyburn-Powell, C. and Jahn, A.

Abstract

Summer Arctic sea ice is declining rapidly but with superimposed variability on multiple timescales that introduces large uncertainties into projections of future sea ice loss. To better understand what drives at least part of this variability, we show how a simple linear model can link dominant modes of climate variability to low-frequency regional Arctic sea ice concentration (SIC) anomalies. Focusing on September, we find skillful projections from global climate models (GCMs) from the Coupled Model Intercomparison Project Phase 6 (CMIP6) at lead times of 4-20 years, with up to 60% of observed low-frequency variability explained at a 5-year lead time. The dominant driver of low-frequency SIC variability is the Interdecadal Pacific Oscillation (IPO) which is positively correlated with SIC anomalies in all regions up to a lead time of 15 years, but with large uncertainty between GCMs and internal variability realization. The Niño 3.4 Index and Atlantic Multidecadal Oscillation have better agreement between GCMs of being positively and negatively related respectively with low-frequency SIC anomalies for at least 10-year lead times. The large variation between GCMs and between members within large ensembles indicate the diverse simulation of

teleconnections between the tropics and Arctic sea ice, and the dependence on initial climate state. Further, the influence of the Niño 3.4 Index was found to be sensitive to the background climate. Our results suggest that, based on the 2022 phases of dominant climate variability modes, enhanced loss of sea ice area across the Arctic is likely during the next decade.

Significance Statement

The purpose of this study is to better understand the drivers of low-frequency variability of Arctic sea ice. Teasing out the complicated relationships within the climate system takes a large number of examples. Here we use 42 of the latest generation of global climate models to construct a simple linear model based on dominant named climate features to predict regional low-frequency sea ice anomalies at a lead time of 2-20 years. In 2022, these modes of variability happen to be in the phases most conducive to low Arctic sea ice concentration anomalies. Given the context of the longer-term trend of sea ice loss due to global warming, our results suggest accelerated Arctic sea ice loss in the next decade.

3.1 Introduction

Over the past four decades, summer Arctic sea ice has rapidly declined and is projected to continue to decline in the future (Wang and Overland, 2012; Notz and Stroeve, 2016; Sigmond et al., 2018). However, large variability on multiple timescales is superimposed on this declining trend, which can lead to 10-20 year periods of accelerated sea ice loss but also to a period of over a decade of no sea ice loss (Kay et al., 2011; Swart et al., 2015). Hence, it is not unexpected that no new record low September sea ice area has occurred since 2012 (Francis and Wu, 2020), in particular as September internal variability is currently elevated due to the decrease in the thickness and mean sea ice state (Goosse et al., 2009; Eisenman, 2010; Jahn, 2018; Mioduszewski et al., 2019). The shelf seas have been the focus of the observed decline as well as of the impact of internal variability, with lower average sea ice concentration and thinner ice making the area a

hotspot of internal variability over the past few decades (Lindsay and Zhang, 2006; England et al., 2019; VanAchter et al., 2020; Årthun et al., 2021). The shelf seas are also coincident with areas of interest for shipping (Eguíluz et al., 2016; Melia et al., 2017), natural resource exploration (Petrick et al., 2017), and ecological changes (Kovacs et al., 2011). However, the current characteristics of variability are likely transitory as the shelf seas in the next few decades will become more reliably ice-free throughout the summer (Barnhart et al., 2016; Crawford et al., 2021), ending the dominant role of internal variability in projection uncertainty for this region (Bonan et al., 2021).

The internal variability of Arctic sea ice acts on multiple timescales and has therefore been challenging to cleanly separate from the forced response (Stroeve et al., 2007; Kay et al., 2011; Swart et al., 2015; Dörr et al., 2023). High-frequency drivers such as atmospheric temperature and wind anomalies are generally considered dominant over lower-frequency drivers (Ding et al., 2019; Olonscheck et al., 2019; Roach and Blanchard-Wrigglesworth, 2022), but separating the drivers is difficult due to large spatial and temporal heterogeneity in variability (Onarheim et al., 2018). By defining low-frequency variability as periods of at least 2 years, approximately one quarter of September pan-Arctic internal variability can be accounted for by low-frequency variability in a sample of global climate models (GCMs) (Wyburn-Powell et al., 2022). Although low-frequency variability is only a small component of internal variability, it promises some longer term predictability, as the influence of initial conditions and high-frequency drivers of variability decay rapidly beyond the current season (Blanchard-Wrigglesworth et al., 2011; Bonan et al., 2019; Bushuk et al., 2019), and have been shown to be useful to a maximum of two-year lead time (Day et al., 2014; Yeager et al., 2015; Bushuk and Giannakis, 2017; Holland et al., 2019; Gregory et al., 2021; Wang et al., 2021).

There is some prospect of summer Arctic sea ice predictability at lead times greater than 2 years due to ocean heat transports (Zhang and Wallace, 2015; Docquier et al., 2021) and climate modes of variability (Guemas et al., 2016). However, results so far seem to be model dependent (Tietsche et al., 2014; Blanchard-Wrigglesworth and Bushuk, 2019), and our current length of ob-

servations is likely too short to verify such relationships (Bonan and Blanchard-Wrigglesworth, 2020; Karami et al., 2023). Despite these challenges, extra-tropical modes of sea level pressure variability have been suggested to directly affect Arctic sea ice variability, but so far only with strong evidence on high-frequency timescales (Ukita et al., 2007; Serreze et al., 2007; L'Heureux et al., 2008; Zhang et al., 2019; Liu et al., 2021). Tropical teleconnections have also been identified as influencing Arctic sea ice loss, primarily associated with Pacific sea surface temperatures (SSTs) (Hu et al., 2016; Li et al., 2018a; Screen and Deser, 2019; Ding et al., 2019; Kim et al., 2020; Clancy et al., 2021; Jeong et al., 2022b; Simon et al., 2022), but also with Atlantic variability (Day et al., 2012; Miles et al., 2014; Meehl et al., 2018; Li et al., 2018b; Karami et al., 2023). Rossby wave trains are the primary mechanism linking tropical Pacific SST anomalies to the Arctic (Yuan et al., 2018). These Rossby waves propagate from the tropics to the Arctic in the order of two weeks (Alexander et al., 2002), but can have seasonal Arctic sea ice effects due to persistent positive geopotential height anomalies and associated subsidence and diabatic warming leading to reduced sea ice cover (Baxter et al., 2019; Hofsteenge et al., 2022). These insights into drivers of variability show promise, but skillful regional sea ice predictions combining multiple modes of variability at lower-frequency timescales has so far been elusive.

Assessing drivers of low-frequency variability in the climate system is difficult to do without large quantities of consistent data, such as that available from single model initial-condition large ensembles (Deser et al., 2020; Milinski et al., 2020). This requirement for assessing drivers of low-frequency Arctic sea ice variability stems from a multitude of drivers likely interacting on heterogeneous spatial and temporal scales to cause this variability (Zhang et al., 2020). This has, so far, lead to a lack of consensus of many of the drivers at time periods in excess of 2 years, especially as GCMs and observations have been shown to represent these relationships differently. We therefore leverage all available GCMs from the Coupled Model Intercomparison Project Phase 6 (CMIP6) archive to investigate model consensus of these low-frequency relationships. Additionally, we do not prescribe the nature of any of these relationships such as linearity

and independence, and perform a detailed regional analysis as well as assess multiple lead times. To enable interpretation of these potentially complex relationships in the climate system we use machine learning which has been used successfully before to explain patterns of surface climate variability (e.g. Barnes et al., 2019; Labe and Barnes, 2022). With this coherent approach to determine the drivers of low-frequency Arctic sea ice variability on multiple timescales and locations, we determine the modes of variability which are simulated to have the largest impact and use the resulting model to make predictions of low frequency SIC variability over the next decade.

3.2 Methods

3.2.1 Data sources

In order to gather sufficient data of both climate modes of variability and associated sea ice concentrations, we use 42 GCMs with historical CMIP6 forcing (O'neill et al., 2016). These GCMs are those for which both monthly sea ice concentration is available and the full suite of climate mode data has been processed using the Climate Variability Diagnostics Package (CVDP) (Phillips et al., 2014). In total we use 609 realizations, from 42 GCMs and 23 modeling centers; a full list can be found in Table 3.1. In using the full suite of CMIP6 GCMs we can get a consensus of low-frequency drivers of Arctic sea ice variability, as individual GCMs have biases in their simulation of teleconnections (Dalelane et al., 2023), but some systematic biases pervasive across CMIP5 are improved in CMIP6 (Fasullo et al., 2020).

Alternatives to the historical simulations which could provide a similarly large quantity of data include future scenarios or pre-industrial control simulations. However, as the mean-state and variability of the Arctic sea ice (VanAchter et al., 2020; Årthun et al., 2021) and some aspects of the rest of the climate system such as El Niño Southern Oscillation (ENSO) (Brown et al., 2020) or AMOC (Weijer et al., 2020) differ from present conditions, this approach would be less appro-

priate to analyze near-contemporaneous variability. Despite differences in mean state, we do utilize pre-industrial control simulations to assess the validity of our detrending methodologies, but not make projections, as detailed in section 3.2.4.

Within the historical period we use the 95-year time period 1920-2014 for sea ice concentration (SIC), which we average over regions of the Arctic as defined by the National Snow and Ice Data Center (NSIDC) Multisensor Analyzed Sea Ice Extent - Northern Hemisphere (Fetterer et al., 2010) (see Figure 3.1d). These seven regions cover the vast majority of the sea ice found during the summer, although we do exclude the Canadian Arctic Archipelago due to complex coastal zones which are typically poorly represented in GCMs (Long et al., 2021). We linearly detrend the average SIC for each region and then apply a 2-year lowpass filter to exclude the high-frequency interannual variability and leave only the low-frequency anomalies (see Figure 3.1a-c). This lowpass filtered regional sea ice concentration data becomes the predictands in our regression analysis.

Table 3.1: Global climate model output used in this analysis

Modeling Center	GCM Name	Members	Citation
CSIRO-ARCCSS	ACCESS-CM2	5	Dix et al., 2019
CSIRO	ACCESS-ESM1.5	40	Ziehn et al., 2019
BCC	BCC-CSM2-MR	3	Wu et al., 2018
BCC	BCC-ESM1	3	Zhang et al., 2018
CAMS	CAMS-CSM1.0	3	Rong, 2019
NCAR	CESM2-FV2	3	Danabasoglu, 2019a
NCAR	CESM2-LENS	50	Danabasoglu, 2019b
NCAR	CESM2-WACCM	3	Danabasoglu, 2019d
NCAR	CESM2-WACCM-FV2	3	Danabasoglu, 2019c
THU	CIESM	3	Huang, 2019
CMCC	CMCC-CM2-SR5	11	Lovato and Peano, 2020
CNRM-CERFACS	CNRM-CM6-1	21	Voldoire, 2018
CNRM-CERFACS	CNRM-ESM2-1	6	Seferian, 2018
CCCma	CanESM5	65	Swart et al., 2019b
CCCma	CanESM5-CanOE	3	Swart et al., 2019a
E3SM-Project	E3SM1.0	4	Bader et al., 2019
EC-Earth-Consortium	EC-Earth3	23	EC-Earth-Consortium, 2019a
EC-Earth-Consortium	EC-Earth3-CC	10	EC-Earth-Consortium, 2021
EC-Earth-Consortium	EC-Earth3-Veg	7	EC-Earth-Consortium, 2019b
EC-Earth-Consortium	EC-Earth3-Veg-LR	3	EC-Earth-Consortium, 2020
FIO-QLNM	FIO-ESM2.0	3	Song et al., 2019
NOAA-GFDL	GFDL-ESM4	3	Krasting et al., 2018
NASA-GISS	GISS-E2-1-G	46	NASA Goddard Institute for Space Studies, 2018
NASA-GISS	GISS-E2-1-H	25	NASA Goddard Institute for Space Studies, 2019b
NASA-GISS	GISS-E2-2-G	11	NASA Goddard Institute for Space Studies, 2019a
NASA-GISS	GISS-E2-2-H	5	NASA Goddard Institute for Space Studies, 2019c
MOHC	HadGEM3-GC31-LL	5	Ridley et al., 2019a
MOHC	HadGEM3-GC31-MM	4	Ridley et al., 2019b
INM	INM-CM5-0	10	Volodin et al., 2019
IPSL	IPSL-CM6A-LR	32	Boucher et al., 2018
MIROC	MIROC-ES2H	3	Watanabe et al., 2021
MIROC	MIROC-ES2L	31	Hajima et al., 2019
MIROC	MIROC6	50	Tatebe and Watanabe, 2018
HAMMOZ-Consortium	MPI-ESM1.2-HAM	3	Neubauer et al., 2019
MPI-M	MPI-ESM1.2-HR	10	Schupfner et al., 2019
MPI-M	MPI-ESM1.2-LR	30	Wieners et al., 2019
MRI	MRI-ESM2.0	12	Yukimoto et al., 2019
NUIST	NESM3	5	Cao and Wang, 2019
NCC	NorCPM1	30	Bethke et al., 2019
NCC	NorESM2-LM	3	Seland et al., 2019
NCC	NorESM2-MM	3	Bentsen et al., 2019
MOHC	UKESM1.0-LL	16	Tang et al., 2019

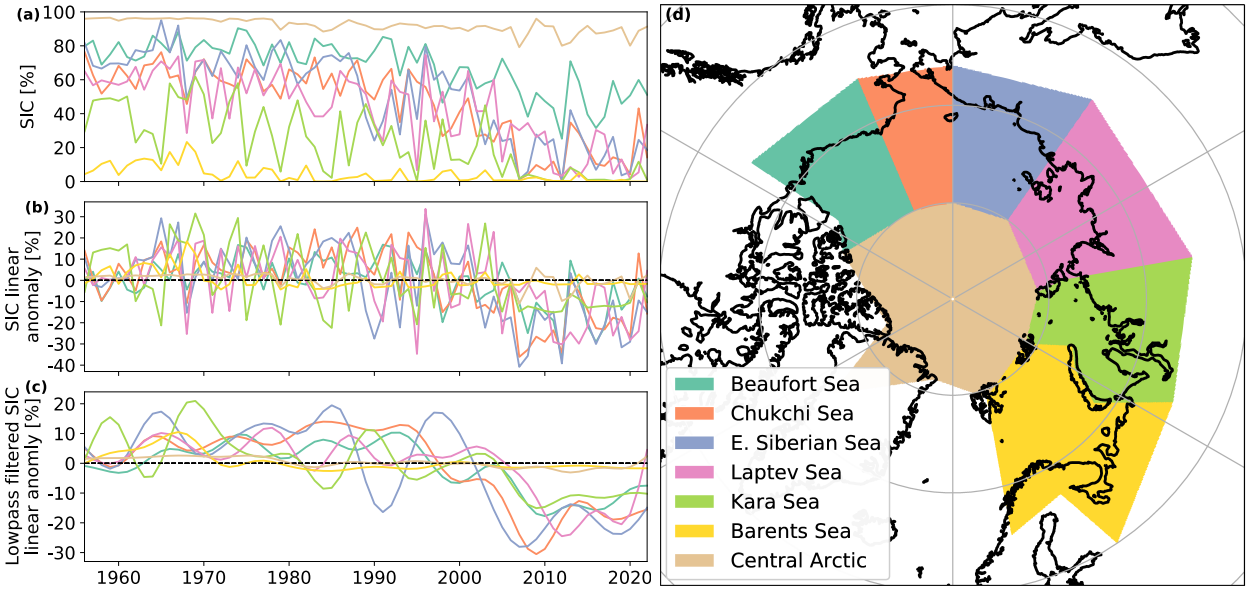

Figure 3.1: **Observed September sea ice concentrations for the seven Arctic regions used in this analysis.** The observational HadISST1 sea ice concentration data shown for (a) the regional average, (b) the linearly detrended version of (a), and (c) a 2-year lowpass filter applied on (b). What is shown in (c) is the data used in the analysis presented here. The outline of the different regions considered are shown in (d) and defined as for the National Snow and Ice Data Center (NSIDC) Multisensor Analyzed Sea Ice Extent - Northern Hemisphere (MASIE-NH) dataset (Fetterer et al., 2010).

We use nine variables from the CVDP to assess their influence on regional SIC anomalies in our regression analysis. Below we have included a brief description of these modes of variability, we have also included a citation of a relevant article using the same index. These climate modes of variability aim to capture different aspects of variability within the climate system, although some of these do overlap in spatial or temporal domains, and thus should not be considered independent. We obtain seasonal values for all variability modes which are then linearly detrended over the period 1920-2014 and standardized (if not already in such a format). As we lag the SIC data between 2 and 20 years from the CVDP data, only the latest 74 of the 95 year time period is used for a given lag time (1941-2014). When we present the linear effects of each mode of variability, we only use one seasonal value for the climate modes listed below (see section 3.2.2 for selection of the season):

- AMO: Atlantic Multidecadal Oscillation, winter - DJF. The area-weighted SST anomalies in the north Atlantic basin (0-60°N, 80°W-0°E), which is thought to have a period of approximately 60-80 years (Trenberth and Shea, 2006).

- NAO: North Atlantic Oscillation, winter - DJF. The leading principal component of the Atlantic (20-80°N, 90°W-40°E) seasonal average sea level pressure anomalies. Positive phase indicates a relatively enhanced Azores high and deepened Icelandic low (Hurrell and Deser, 2009). The NAO may have some small decadal predictability, such as from the AMO, but is dominated by large interannual variability (Klavans et al., 2021).

- ATN: Atlantic Niño, spring - MAM. The area-averaged tropical Atlantic SST anomalies (3°S-3°N, 20°W-0°E), with a similar periodicity to the Pacific El Niño/La Niña phases (Zebiak, 1993).

- NINO34: Niño 3.4 Index, winter - DJF. 5-month running mean SST anomalies in the equatorial Pacific (5°N–5°S, 120°–170°W). Values continuously in excess of +0.4°C for 6 months indicate El Niño conditions, below -0.4°C indicates La Niña (Trenberth, 1997). Such oscillations between positive and negative states occur approximately every 2-7 years

in the observational record.

- PDO: Pacific Decadal Oscillation, spring - MAM. The leading principal component of north Pacific SST anomalies (20-70°N, 110°E-100°W). Positive phases are associated with positive SST anomalies in the eastern Pacific and negative SST anomalies in the western and central Pacific (Mantua et al., 1997). The PDO is thought to have a periodicity of approximately 50-70 years over the last 200 years (MacDonald and Case, 2005).

- NPO: North Pacific Oscillation, spring - MAM. The second principal component of seasonal sea level pressures over the north Pacific and North American continent (20-85°N, 120°E-120°W) (Phillips et al., 2014). A positive phase is indicative of a deepened Aleutian low and enhanced sea level pressure in the region of 20-40°N as per Rogers (1981) who defined the NPO based on geopotential height. A given phase usually persists on the order of a week.

- PNA: Pacific/North American Teleconnection, spring - MAM. The leading principal component of seasonal sea level pressures over the north Pacific and North American continent (20-85°N, 120°E-120°W) (Phillips et al., 2014). A positive phase is similar to the NPO with a deepened Aleutian low, but this mode of variability is more extensive, also including enhanced pressure over western Canada, see Leathers et al. (1991) who used geopotential height anomalies.

- IPO: Interdecadal Pacific Oscillation, spring - MAM. The leading principal component of 13-year lowpass filtered Pacific (40°S-60°N, 110°E-70°W) area-weighted SST anomalies. In its positive phase SST anomalies in the equatorial Pacific are positive with the western extra-tropical Pacific in both hemispheres experiencing cooler SST anomalies (Meehl et al., 2013). The period and symmetry of the IPO is thought to have varied considerably over time, but over the observational period it has been shown to change phase approximately every 20-30 years (Vance et al., 2022).

In addition to these modes of variability, we also include the summer (JJA) global average surface temperature (TAS), as motivated in section 3.2.4.

Several additional modes of variability were also available from the CVDP but were not included in the final analysis. The modes investigated but not used are as follows: the Indian Ocean Dipole, the Atlantic Meridional Mode, the Southern Annular Mode, the North Pacific Index. All of these modes of variability had no measurable effect on the regression model. Furthermore, including the Northern Annular Mode led to over-fitting with the highly related NAO.

To compare model results to observations, we use SIC from the Hadley Centre Sea Ice and Sea Surface Temperature data set (HadISST1) (Rayner et al., 2003) for the period 1956-2022. We use the HadISST1 SIC record before the beginning of the satellite era in 1978 to enable longer analyses in our correlation analysis in section 3.3.5. We start using the HadISST1 SIC data in 1956, as variability is degraded substantially before 1956 due to interpolations during winter (Rayner et al., 2003). However, when calculating linear trends for detrending, we use SIC data for 1920-2014 in order to be consistent with the GCMs. This is possible due to moderate confidence in the mean state for 1920-1955 despite the increased uncertainty in the interannual sea ice variability for that period. The HadISST1 data, similarly to the SIC in the GCMs, is divided into regions, linearly detrended and interannual variability is removed with a 2-year lowpass filter. For observed climate variability data we also obtain these from the CVDP where we use the HadISST1 dataset to calculate sea surface temperature-derived variables, the NCEP-NCAR record for sea level pressures (Kalnay et al., 1996), and GISTEMP version 4 for global surface temperatures (Lenssen et al., 2019). Similarly to the CVDP output variables for the GCMs, we apply a linear detrending and standardization to the variables not already in this format.

3.2.2 Machine Learning Methods

To determine the relationship between the climate variability modes and the lagged effects on regional Arctic SIC gain and loss, we use machine learning. Specifically we use neural networks which excel at finding relationships within large data sets (e.g. Diffenbaugh and Barnes, 2023). At its simplest, the neural networks used here are multiple linear regression, but we can also account for non-linear relationships and covariance by using more advanced neural network configurations. In order to constrain the potentially complicated relationships between climate modes and subsequent SIC changes, we require large quantities of data to train, validate and test our neural networks. We therefore utilize three data sets as listed below, which fulfill different purposes:

- 12 LEs, individual CMIP6 GCM large ensembles of at least 20 members.
- MMLE 3+, all CMIP6 GCMs (42) with at least 3 members.
- MMLE 30+, all CMIP6 GCMs (8) with at least 30 members.

To determine the climate mode relationships with Arctic sea ice within an individual GCM we require at least 20 members to provide sufficient data. This means we can train a neural network separately on 12 of the 42 GCMs, referred to as LEs. To get a consensus across the 42 CMIP6 GCMs and weight them equally, we train a neural network on the 1st members of all 42 GCMs, validate on the 2nd members and test on the 3rd members (the MMLE 3+). Finally, we also train a neural network on the first 23 member of 8 GCMs with sufficiently large ensembles, this allows us to see whether maximizing the available data increases predictive skill (MMLE 30+).

For all LEs, MMLE3+ and MMLE30+ we use a single seasonal time series from 8 climate modes and TAS to predict lagged sea ice anomalies at one lead time, one region, and one sea ice anomaly month at a time. Allowing any patterns between the lags, region or sea ice anomaly months to be discovered rather than prescribed. The SIC anomalies are in % points for consis-

tency across regions. Hence, when comparing the influence of modes of variability in aggregate, the % point change should be scaled by the variability of that region (as is done for Figure 3.8). The use of % SIC deviation from the trend has identical meaning to using sea ice area and is not sensitive to the mean state, other than the 0-100% bounds capping anomalies. The neural networks have no knowledge of the initial sea ice state, but as the memory for the summer at lead times in excess of 1 year is considered negligible (Giesse et al., 2021), this omission is considered unimportant at the timescales we consider. Further, including initial sea ice state as a predictand would add complexity to our methods which would be difficult to constrain without additional data.

We utilize four configurations of machine learning model to test whether nonlinearities and covariance between the climate modes is required to make skillful predictions of Arctic sea ice anomalies. We achieve this by constructing four models listed below differing in their linear or nonlinear relationships (activation functions) and whether they take into account climate mode covariance (presence or absence of hidden layers). Model 1 has independent linear relationships between the climate modes and sea ice anomalies, and hence is effectively multiple linear regression. Model 2 is the same as model 1 but permits nonlinear relationships. Model 3 uses only linear relationships but can take advantage of covariance between climate modes, such as a positive phase of the IPO and a positive phase of the PDO having a different combined effect than the individual effect of those modes. Model 4 is the most complicated, allowing both nonlinear relationships and also covariance between the modes of variability. For further details on the machine learning models see Supplementary section S1.

3.2.3 Assessing Predictive Skill

The threshold for our machine learning model to be useful at a given lag time is defined as when its Pearson correlation coefficient for the validation data exceeds that obtained from persistence. The persistence correlation coefficient in this instance is calculated from the 2-year lowpass

filtered regional SIC anomalies lagged between 2 and 20 years, the same lag times as used for our regression analysis. When using the correlation coefficient, it is important to note that, especially at longer lag times, there may be a high correlation between the linear model output and the validation data, but this skill may be present with a smaller amplitude than for the validation data. Further, for regions that are close to zero or 100% SIC, we are trying to predict very small variations in SIC. Hence we could have poor predictability in these regions but still have small errors in absolute terms.

As we do not have sufficiently long periods of observations, we cannot train a separate machine learning model on the observations. Instead, by pooling several regions and SIC anomaly months, we calculate the proportion of positively and negatively correlated modes of variability with the most extreme 10% of SIC positive and negative anomalies. This is not a way of verifying the GCM predictive models per se, rather it shows the range of correlations present within a large ensemble and allows observation to be placed alongside that range. Observations would be expected to typically fall within the large ensemble distribution, but as we do not know how atypical our one realization of reality is, we cannot ascribe meaning to differences from the ensemble mean (Notz, 2015). Similarly, when in section 3.3.5 we provide predictions of past and future regional SIC anomalies, good agreement to observations does not explicitly validate our results.

3.2.4 Sensitivities to time period and forcing

We use a linear detrending for both the SIC and the CVDP variables over the period 1920-2014 as this is a simple process to understand and does not make specific assumptions about the time period in question. This is not perfect as during that period the radiative forcing as well as the observed and modeled sea ice decline were not entirely linear (see Figure 1 from Mcbride et al. (2021) for global temperature). This means that some of the very low-frequency variability of the forced response is incorporated into the anomalies of SIC and CVDP variables, rather than being removed by detrending. Therefore, some predictability is due to the shape of the forced response,

primarily represented by our input variable of global average surface temperature (TAS), and likely, to a small extent, the SST-derived variables of NINO34, PDO, ATN, AMO, and the IPO. As the simple linear model used in our results considers each variable independently, we can consider TAS similarly to a residual term in the model which does not affect the conclusions we draw about other modes of variability.

To verify that our results from the period 1920-2014 are robust to different forcing conditions, we compare results with a more linear forcing scenario for the historical period 1970-2014 and a constant pre-industrial forcing scenario. For the 1970-2014 time period the global surface temperature and sea ice area trends are both highly linear (Notz and Stroeve, 2016; Mcbride et al., 2021). Consequently, for 1970-2014 we find that the linear response to TAS in our models is far smaller than in 1920-2014 (see Figure S3.1, compared with Figure 3.4). The 1970-2014 time period, after accounting for lags, only uses 24 years of data (compared with 74 for 1920-2014) and hence the linear response is much more noisy than for 1920-2014. Therefore, although we get a broadly similar linear responses for each climate mode, the low skill relative to persistence means we cannot use this shorter time period, despite the more linear variables and more similar mean state to the present day.

Pre-industrial control runs (of which 35 GCMs are available to each provide 222 training years) use constant 1850 radiative forcing and hence TAS trends are near zero over a 74-year time periods. Despite the different mean state and variability, we still find very similar linear coefficients to the 1920-2014 time period, but with a smaller influence of TAS (see Figure S3.2 compared with Figure 3.4). However, the pre-industrial control results provide much smaller linear responses, likely due to the 1850 mean-state exhibiting less variability than the 21st century, primarily due to thicker Arctic sea ice (Kwok and Rothrock, 2009). Despite the pre-industrial control climate being too different to present day to make projections, the similar results to the 1920-2014 period implies that the relationships are inherent to the climate system, not artifacts of the detrending

methodology, with the possible exception of NINO34 as discussed in section 3.4. We therefore use the 1920-2014 time period, despite the TAS and SIC nonlinearity, as it both captures similar SIC mean state and variability to the present day, and enables the use of sufficient training data.

3.3 Results

3.3.1 A simple linear model captures drivers of low-frequency variability

Predictions of regional low-frequency Arctic sea ice concentration anomalies can be produced from climate modes of variability using a linear model, which are skillful when compared with persistence. In general, we find that the simple linear variant of the machine learning models (model 1) produces the highest predictive skill of the four models across GCMs, regions and seasons. When validating our linear model we find it generally exceeds the skill from persistence for lead times beyond approximately 4 years, but is dependent on the GCM (see Figure 3.2 for the Chukchi Sea in September). The highest predictive skill is found at approximately a 5-year lead time when the r^2 value of persistence has decayed close to zero while the r^2 value of the linear model declines more slowly with lead time. This temporal pattern of persistence, as well as the superiority of the linear model, is found across regions and months with nonzero skill (see section 3.3.2).

The simple linear model with no hidden layers (model 1) and the linear neural network allowing climate mode covariance (model 3) are nearly identical in their performance across different LEs and MMLEs (see Figure 3.2). The high performance of models 1 and 3 imply that nonlinearities are not required to produce a skillful predictive model. The simple nonlinear model 2 consistently performs poorly, with model 4 performing erratically for small training data but can exceed the skill of other models for short lead-times and for the largest LEs and MMLEs. As model 4 includes the effect of covariance of climate modes and nonlinearities, this complex relationship between climate modes and sea ice anomalies is shown to only provide a modest benefit to predic-

tions. Subsequently, we therefore only utilize model 1, the simple linear model, to clearly determine the independent linear effect of each climate mode of variability. However, with additional data, the likely interdependent and nonlinear relationships may be able to be detected robustly to allow greater generalization and produce better predictions.

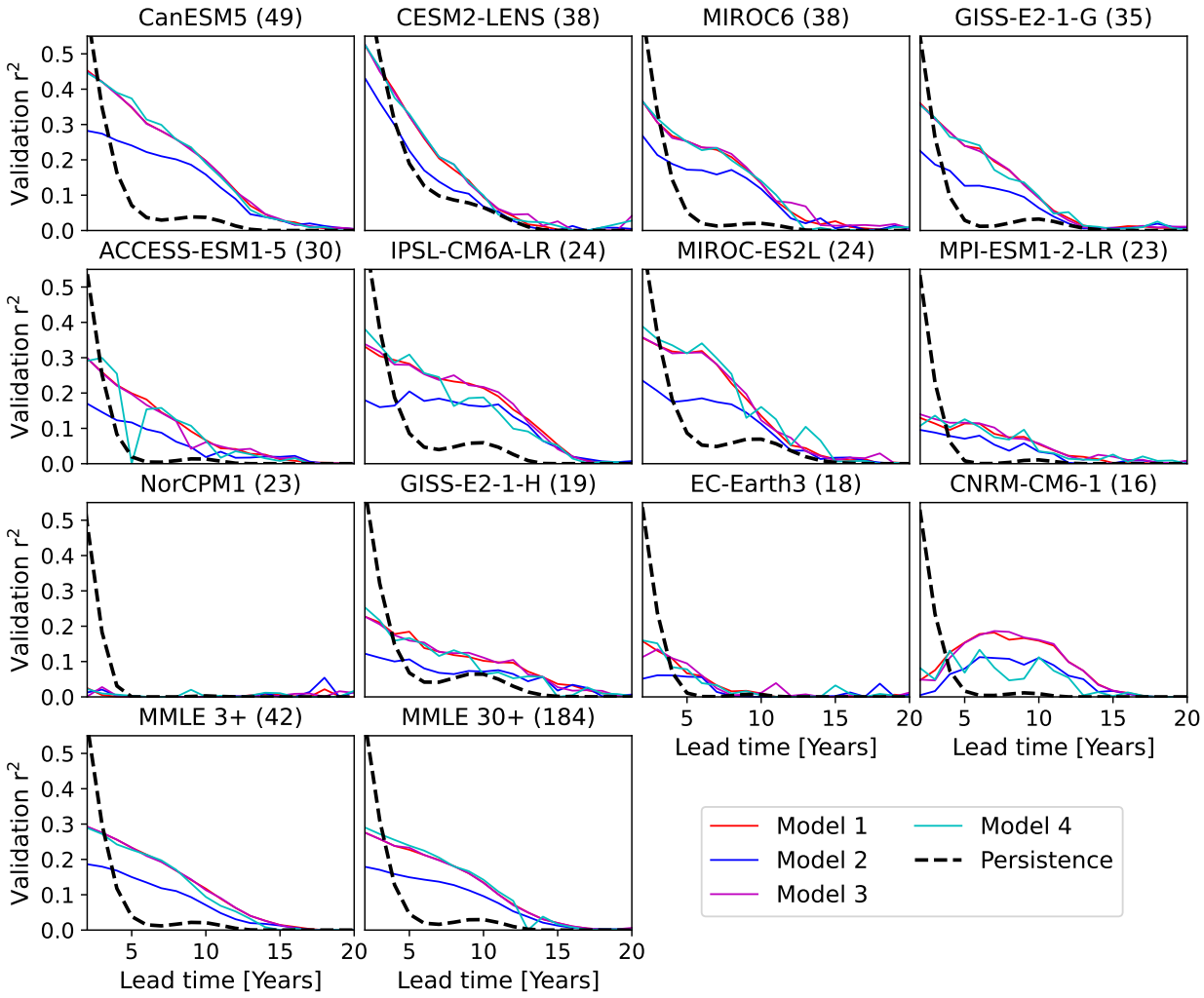

Figure 3.2: **The effect of machine learning model complexity on predictive skill.** Pearson correlation coefficients in the Chukchi Sea in September for the validation data for four machine learning models as shown for the 12 LEs and 2 MMLE datasets. Model 1 refers to the simple linear model (red), model 2 to the simple nonlinear model (blue), and Model 3 and Model 4 to the fully-connected 9-3-3-1 neural network with linear (purple) and nonlinear (cyan) activation functions, respectively. The black dashed line indicates the average persistence for that lag time for the GCM or GCMs used. Where the model validation r^2 values exceed persistence the model has predictive skill. Numbers in parentheses indicate the number of ensemble members used in training.

3.3.2 Hotspots of low-frequency variability predictive skill

The summer and autumn marginal seas are generally able to produce the highest skill at a 5-year lead time, however the predictive skill varies considerably between GCM. Based on the MMLE 3+, which takes into account the full suite of CMIP6 GCMs with at least 3 ensemble members, the pattern of highest predictability is found in the Beaufort Sea in September, with decaying skill for regions further from the Pacific and for months more distant from September (Figure 3.3). The MMLE 3+ model is unable to produce high predictive skill in the Barents Sea for any season likely due to frequently near zero SIC, and the Kara sea appears to have distinct peaks of predictive skill in July and late autumn.

For models using individual GCMs, the temporal and regional patterns of predictive skill are often noisy for neighboring regions and months, unlike the clearer MMLE models. The relatively high predictive skill values of the LEs typically exceed that of the MMLE 3+ for the best regions, but with less coherence between regions and months. Selecting the LE with the highest skill for a region and month may be appropriate, but each LE's specific spatial and temporal limitations should be taken into account. The MMLE 3+ has lower predictive skill than the best LEs, but is influenced by all 42 CMIP6 GCMs. Therefore, the relatively higher predictive skill in the MMLE 3+ should be seen as less sensitive to individual GCM biases as it is representative of the general agreement between all GCMs. Some LEs such as CanESM5 and ACCESS-ESM1-5 exhibit unusual patterns of high predictability in the Kara and Chukchi Seas in the winter. Other LEs such as CESM2-LENS, GISS-E2-1-H and MIROC-ES2L have particular regions which are far more predictable than others. For example, the CESM2-LENS simulates high persistence for the Chukchi Sea but not for the Beaufort Sea (see Figure S3.4 for 5-year persistence, and section 3.3.6 for a CESM2 bias discussion) which causes the large disparity in predictive skill between these two regions. As September is of particular interest as the typical minimum annual pan-Arctic sea ice cover, and relatively high validation r^2 values occur across regions for September in the MMLE 3+, this is our focus in subsequent analyses.

3.3.3 Linear drivers of regional sea ice anomalies

Using a linear model trained on 42 CMIP6 GCMs (the MMLE 3+ model), we can establish the consensus across GCMs for the independent effect of each mode of variability on regional September SIC anomalies. The lead times where the MMLE 3+ model has no predictive skill is before a 4-year lead time for all regions except the Central Arctic where it is not until a 5-year lag time that the validation r^2 exceeds persistence (see the dotted lines in Figure 3.4). The most important mode of variability is the IPO, which is strongly positively correlated with the SIC in all regions, especially in the East Siberian and Beaufort Seas (Figure 3.4). The IPO decays in influence over time, reaching near zero influence on SIC at approximately a 15-year lead time. The global average surface temperature (TAS) also has a very large coefficients, but as this is not a mode of variability and is considered to integrate modes of variability not represented (see section 3.2.4 for a more detailed explanation), we do not discuss in detail the influence of TAS further.

Aside from the large influence of the IPO, the Niño 3.4 index (NINO34) and the Atlantic Multidecadal Oscillation (AMO) both display a very consistent sign of influence which decays with time. The NINO34 and AMO both have smaller influences than the dominant IPO (approximately one third and one quarter, respectively) for a given one standard deviation anomaly in each mode of variability. Like the IPO and TAS, the influence of the AMO and NINO34 decays relatively monotonically with time. As the skill of persistence also declines nearly monotonically, and the IPO, TAS, NINO34 and AMO all display low-frequency variability, this increases confidence in the validity of these relationships found in the MMLE 3+. The low-frequency oscillations of the other sea surface temperature-derived indices of the Pacific Decadal Oscillation (PDO), and to a lesser extent the Atlantic Niño (ATN), implies the potential for longer-term predictability as with the IPO, TAS, NINO34 and AMO. However the influence of these modes is small at most time periods and does not display a monotonic decline with time. This suggests these two modes are not highly

important in driving low-frequency Arctic sea ice variability, but consistency or lack thereof between LEs (see section 3.3.4) may clarify whether the relationships in the MMLE 3+ are small and independently consistent in magnitude between GCMs, or small due to disagreement between GCMs.

The modes of variability based on sea level pressure patterns are generally a small influence on low-frequency variability of Arctic sea ice. The Pacific/North American Teleconnection (PNA) and North Pacific Oscillation (NPO) do have some coherent regional effects but the switch in sign of influence over time may be indicative of the expectation of a change in the mode itself rather than the effect of the initial sign of the mode. Further, the NPO and PNA are influenced by longer-lived modes of variability in the Pacific (Furtado et al., 2012), potentially meaning these modes are not independent. The North Atlantic Oscillation (NAO) is less erratic than the NPO and PNA with a general negative effect on SIC anomalies but is very small in magnitude and is shown to affect SIC anomalies minimally.

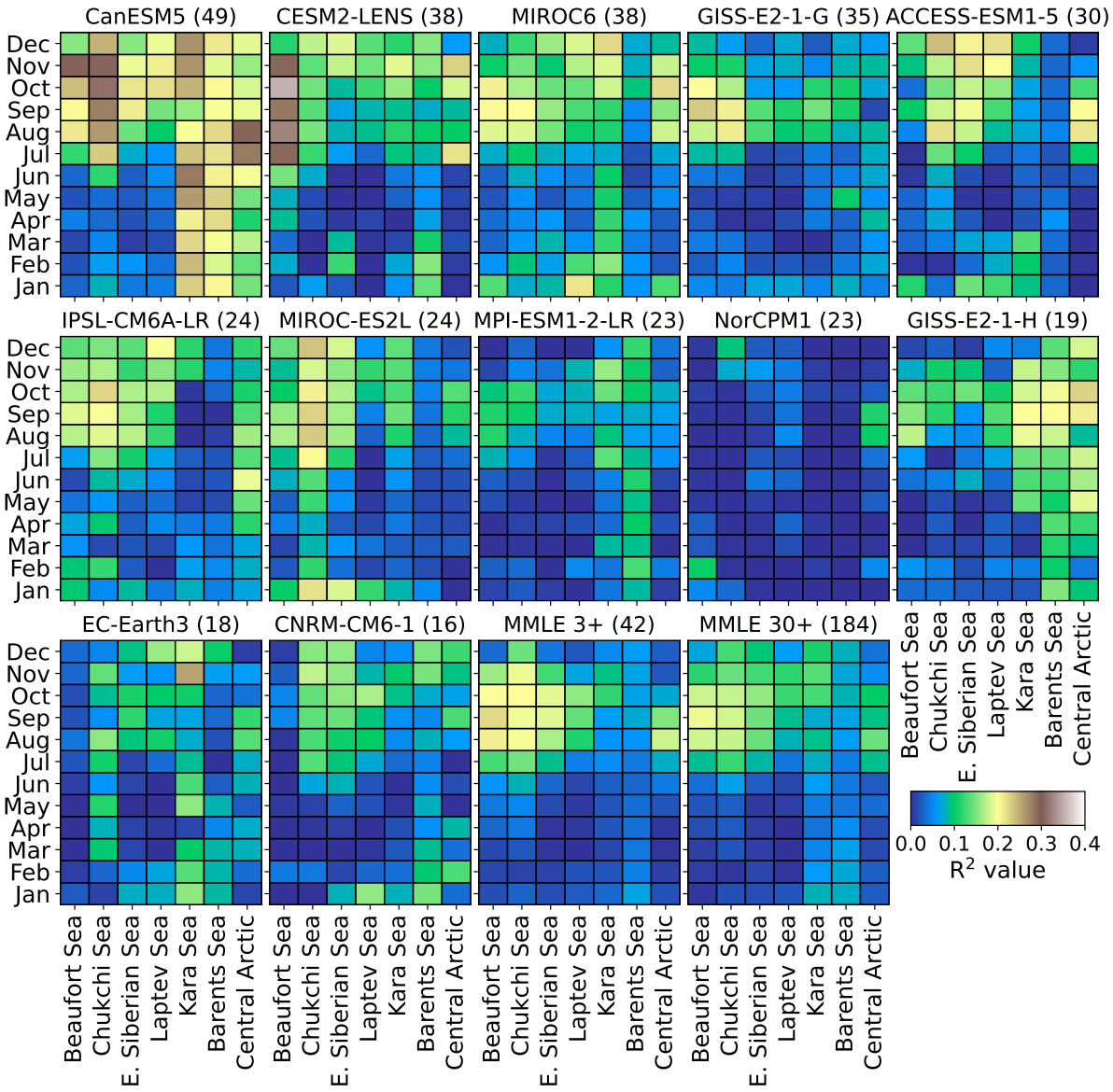

Figure 3.3: **5-year lagged predictive skill for multiple global climate models and the CMIP6 multi-model ensembles.** Pearson correlation coefficients are shown for the validation data minus persistence at a 5-year lag time between the input climate modes and sea ice concentration anomalies. Persistence is removed to indicate the regions and months for each LE or MMLE where predictive skill is high, rather than where explained variability is high. Numbers in parentheses indicate the total number of ensemble members used for training.

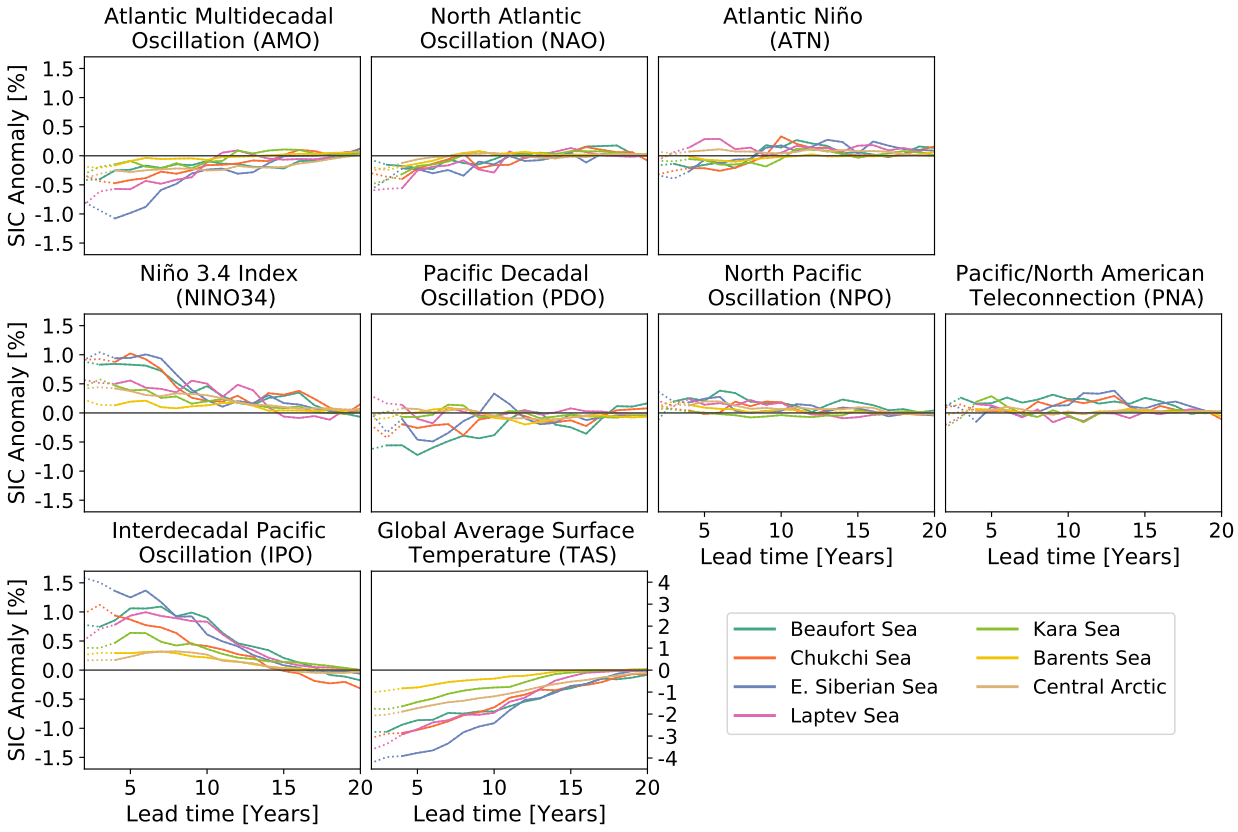

Figure 3.4: **Linear drivers of September regional sea ice concentration anomalies.** Linear response of a +1 standard deviation anomaly of each of the 8 climate modes and global average surface temperature on sea ice concentration anomalies in each of the seven Arctic regions. Positive SIC anomaly values indicate a positive SIC anomaly results from the +1 standard deviation anomaly in the climate mode of variability. Solid lines indicate that the validation r^2 value exceeds persistence for a given region and lead time, dashed lines indicate where there is no predictive skill beyond persistence. Non-zero predictive skill occurs for 4- to 20- year lead times for all regions except for the Beaufort Sea which has some predictive skill for a wider range of 3- to 20-year lead times.

3.3.4 Low-frequency driver representation across global climate models

Comparing the independent results from 12 LEs aids our interpretation of the linear drivers of SIC anomalies captured in the MMLE 3+. We do this by comparing the datasets for both the medium-term for lead times of 4-9 years (Figure 3.5). Although the LE analysis only includes 12 of the 42 GCMs that went into the MMLE 3+ linear model, we can get a sense of the consistency between the CMIP6-suite of GCMs. This informs our interpretation of the two dominant modes of variability, namely the IPO and NINO34 with the LEs varying considerably for both modes of variability during both periods. Although the influence of the IPO and NINO34 are seen to gradually decrease over time for the MMLE 3+, the individual LEs show large magnitudes of influence on SIC for both time periods and the sign is inconsistent between LEs. We find little consensus across GCMs on the sign of influence of the IPO across the 12 LEs. However, when we include these same 12 GCMs and 30 others in the MMLE 3+, a more positive signal emerges. This suggests either the additional 30 GCMs used in the MMLE 3+ have stronger positive linear relationships, and/or that by chance the first members used in the MMLE 3+ have a disproportionately strong positive relationship compared to the many members used for training in the LEs for a given GCM.

For the NINO34 there appears more consistency across the full CMIP6-suite of models with similarities between the collection of 12 LEs and both MMLEs. This again highlights the importance of taking a multi-model approach for the detection of low-frequency variability as two GCMs selected at random may produce vastly different results. Further, although we use large ensembles, the teleconnections between the tropics and Arctic may vary considerably between realization within a large ensemble. Without pooling multiple GCMs and members we may not be able to capture the full possible range of tropical-Arctic linkages which could be present over a 74-year time period.

The AMO has reasonably good agreement between the LEs with almost all indicating negative influence on regional SIC in the medium-term. The PDO in the MMLE 3+ has near zero

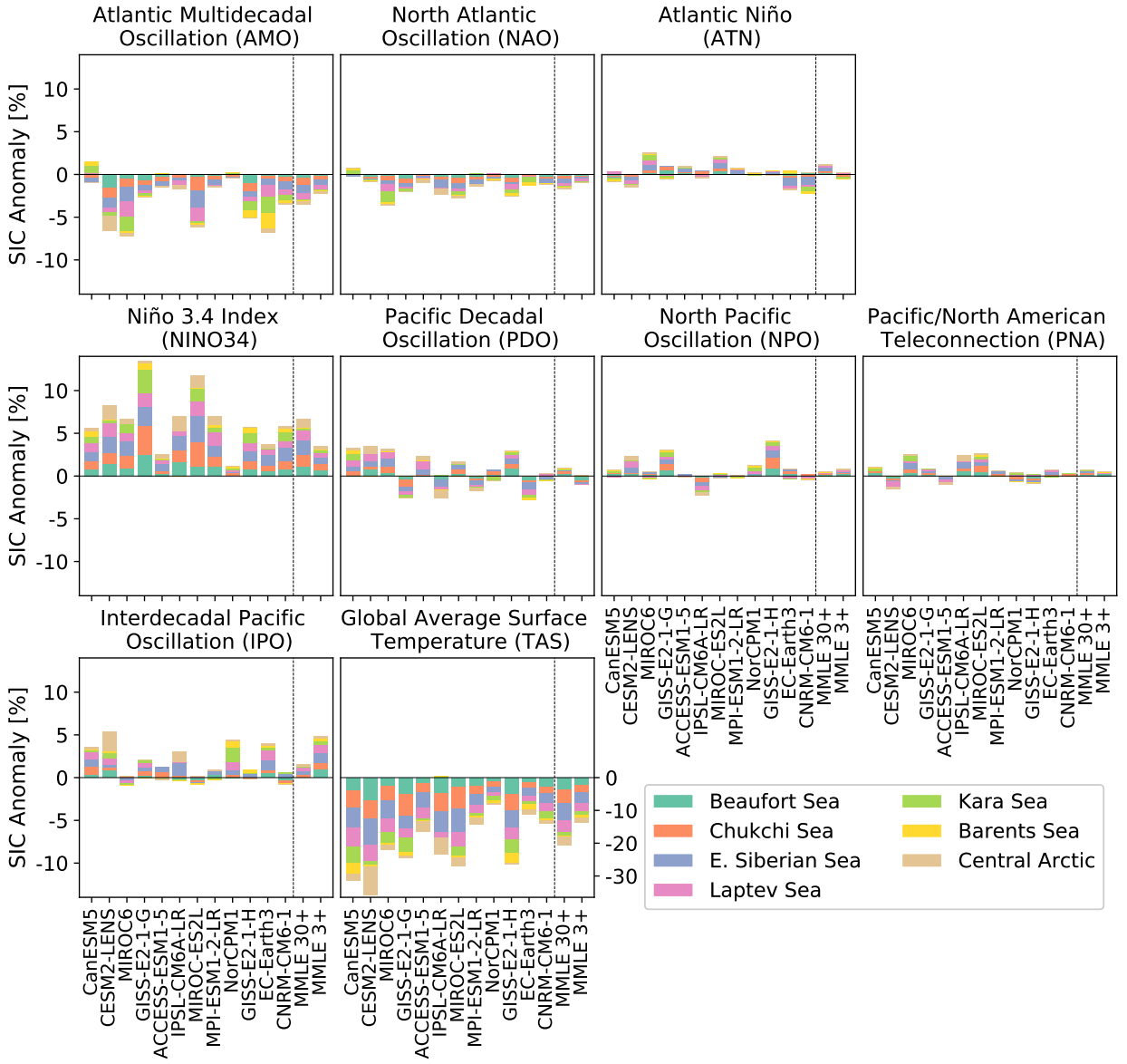

Figure 3.5: **The linear effect on regional SIC for 12 large ensembles and the two multi-model large ensembles.** Linear response in September sea ice concentration for a +1 standard deviation anomaly of each climate mode, as in Figure 3.4, but averaged over two distinct lead times. Bars are the linear response averaged over 4 to 9-year lead times. Agreement within the CMIP6-suite of GCMs is high where bars are similar in magnitude and sign. Note the different y-axis scale for the global average surface temperature.

influence, across all 12 LEs we can see that none indicate the PDO as being particularly influential, with disagreement in sign reducing the overall effect for the MMLE 3+. For the other modes of variability we find that almost all of the LEs coefficients are small in magnitude and without overwhelming agreement on sign. This allows us to interpret the MMLE 3+ near zero coefficients as being representative of both the lack of consensus across CMIP6 GCMs and no strong relationships being found in any of the LEs.

The average magnitude of influence across all modes of variability differs considerably between individual LEs. For example CESM2-LENS often produces the largest magnitudes for a given mode and NorCPM1 the smallest. Such systematic differences may occur due to differences in the mean state and magnitude of variability by GCM. This may well be the case considering the SIC anomaly is recorded in percentage points and CEMS2-LENS has a low biased summer mean-state (DuVivier et al., 2020) and consequently large variability. Conversely, NorCPM1 has been noted as having a high biased sea ice thickness (Bethke et al., 2021), which may explain why NorCPM1 is an outlier for small low-frequency SIC variability. Again, this indicates care must be taken to understand the effect of limitations to the results from individual LEs. Although many of the CMIP6-suite GCMs are related (Knutti et al., 2013), and their biases may not average out, taking the results from the MMLE 3+ can reduce the risk of extreme outliers.

When testing our MMLE 3+ model on unseen members from the 42 GCMs, we find large variation between GCMs and ensemble members (see Figure 3.6). This limits our ability to determine which GCMs are most like the CMIP6 consensus if they have small ensemble sizes which cannot populate the full range of potential values (Notz, 2015). Observational comparison with a similar time period will therefore be also difficult as observations could be expected, due to internal variability, to fall somewhere between 0 and 0.5 r^2 if internal variability in the actual climate system behaves similarly to the the range of ensemble members in a large ensemble such as CanESM5.

Despite the ensemble member disagreement, the MMLE 3+ model appears to be well generalized to multiple GCMs as the test r^2 values appear very similar if a linear model is trained on all 42 GCMs as for the MMLE 3+ (blue circles in Figure 3.6) or only on other members from the same GCM as for the LE (red triangles). CESM2-LENS has a wide distribution of test r^2 values between ensemble members, with larger variations between the micro-perturbations (atmospheric state), than between ensemble members with different ocean states (macro-perturbations) (see Figure S3.3), as also found for pan- Arctic sea ice volume variability (Kay et al., 2022). This indicates that for a 74-year time period, the specific manifestation of the relationships between climate variability modes and regional Arctic SIC anomalies can be highly dependent on the initial climate state.

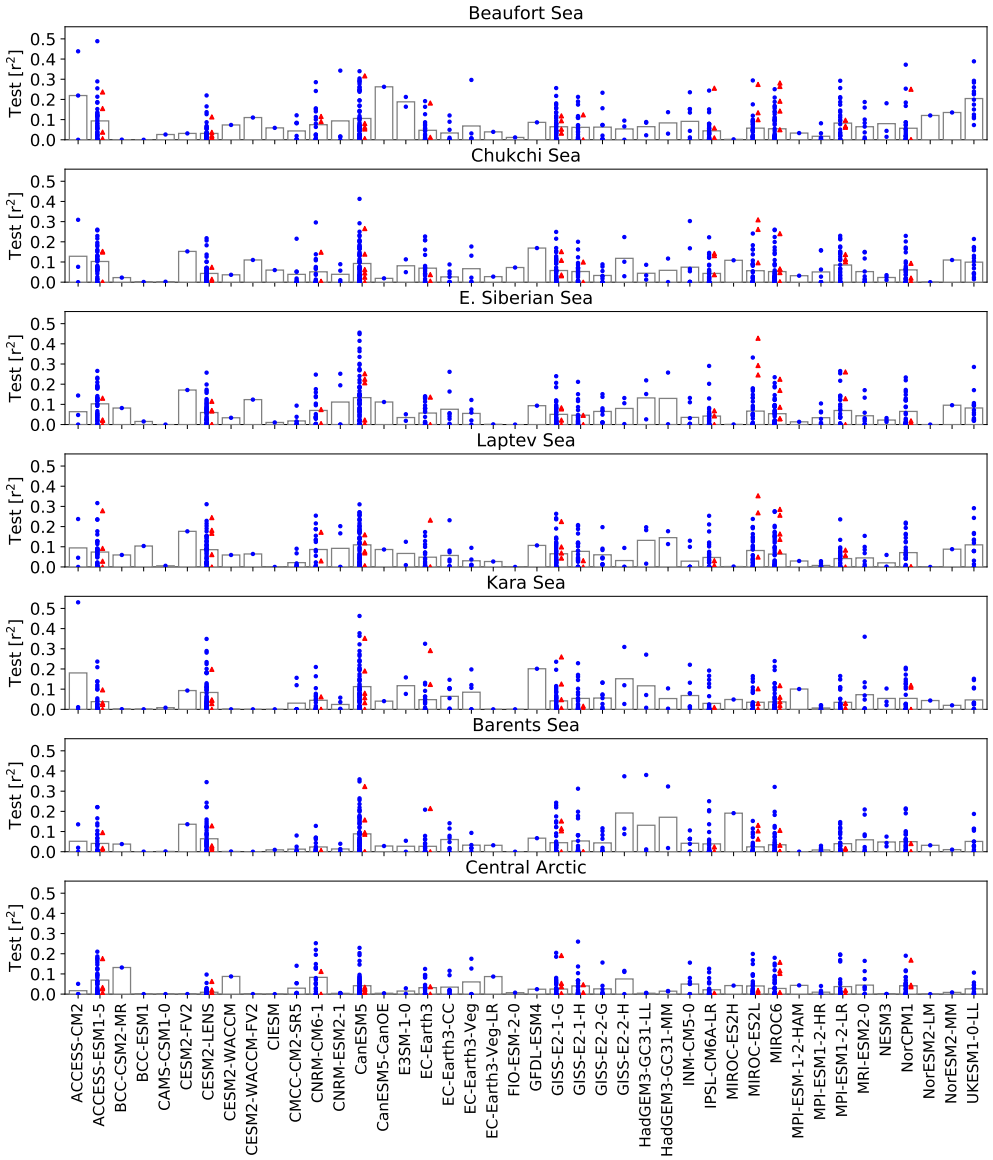

Figure 3.6: **September r^2 values for the test ensemble members from either the multi-model large ensemble (3+, blue) or the 12 single GCM large ensembles (red).** The performance of the test members (third and later ensemble members) for the 42 GCMs included in the MMLE 3+ model are shown as blue circles, ensemble mean values are indicated by gray bars. The red triangles indicate the performance of the test members for the individually trained linear models for each of the 12 LEs, where 10% of the LE members were reserved for testing against the linear model trained and validated on the first 75% and 15% of members from each GCM. Where the red triangles and blue circles for a given GCM have a similar distribution, the MMLE 3+ is equally good at capturing the relationships between climate modes and SIC as the LE, indicating the MMLE 3+ is well generalized. The r^2 values are for a 5-year lead time minus persistence.

3.3.5 Observational comparisons

Correlations between the climate modes and extreme SIC anomalies show observations broadly fall within what is simulated for the LEs, but validation is difficult due to the large differences between realizations. In order to directly compare observations with ensemble members, we compute the correlation between the 6 most extreme regional SIC anomaly years in the period 1956-2022 and correlate whether each mode of variability was in a positive or negative phase. To make a more representative sample, we pool the seven regions (except the Barents Sea where summer variability is near zero), averaged over a 4- to 9-year lead time. However, the correlations should not be seen as comparable to the linear model as the correlations are binary, unlike the abilities of the linear model to apply lower weights to less important climate modes. Observations fall within the ensemble spread for all GCMs for all modes of variability except for the AMO which falls outside of only the NorCPM1, and the ATN and NPO which are outside multiple GCM ensemble ranges (see Figure 3.7). This suggests that the observed correlations between most climate modes of variability and SIC anomalies is consistent with the CMIP6 large ensembles, within internal variability uncertainty. The far stronger correlation of observations for the ATN and NPO may mean in our one realization of reality these modes of variability have played a larger role than has been simulated in many climate models. Again, the large spread between realizations within a large ensemble highlights the extremely large range that observations would be expected to fall within (particularly for the IPO), and hence the difficulty of validating the simulated low-frequency drivers with observations.

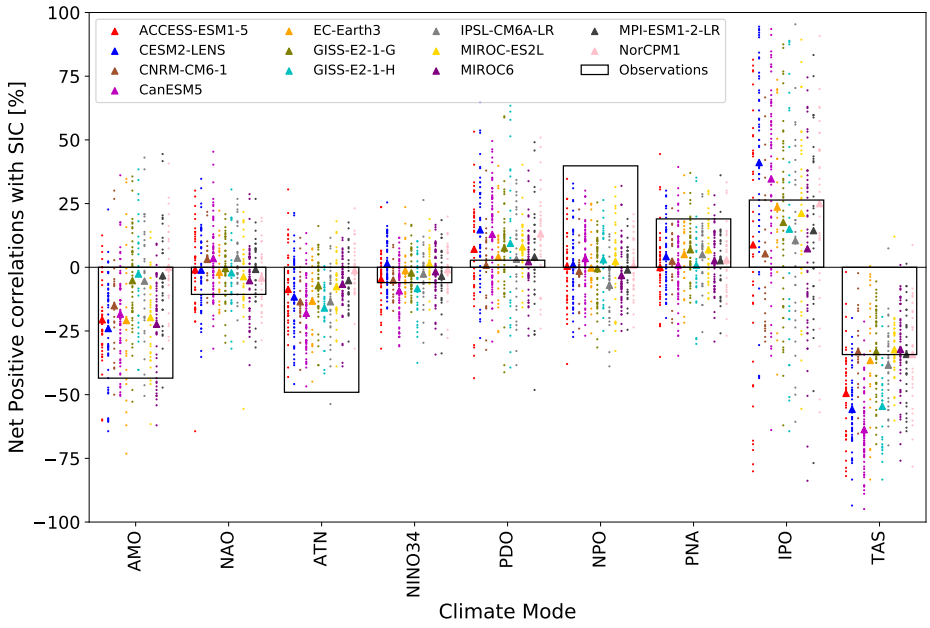

Figure 3.7: **Correlations between ensemble members and observations between modes of variability and extreme SIC anomaly events.** The 6 most extreme SIC positive and negative anomalies are found for each ensemble member and September observations over the period 1956-2014. For a lead time of 4-9 years the positive and negative correlations with each mode of variability is summed. These data are the average for the Beaufort, Chukchi, East Siberian, Kara and Laptev Seas and the Central Arctic. Each colored dot indicates the correlations for a single ensemble member, with the same colored triangle indicating the ensemble mean. The observed value for each variable is shown with a black hollow bar. When observations lie within a given GCM ensemble member distribution, the correlation in the observations is consistent with that simulated in the GCM.

3.3.6 Future projections

Our limited time period of observations may not be representative of a typical climate realization and therefore may arbitrarily match well or poorly to a specific machine learning model trained on GCMs. However, validation of our LE and MMLE 3+ models against the period 1956-2022 may have some implications for how well we can expect projections over the next 4-20 years to hold up. The r^2 values of the MMLE 3+ validated against the observations (Figure 3.8 prediction columns) is similar to that of the MMLE 3+ validated against the second large ensemble members (Figure 3.3). The MMLE 3+ and the best LEs when used for hindcasting SIC anomalies from observed climate modes, often achieve r^2 values of between 0.2-0.3 above persistence, but is highly regionally dependent. As the MMLE 3+ typically has the highest or near highest validation skill against the observations, we use these for future projections in the following.

For all regions of the Arctic, our linear model predicts below trend sea ice concentrations over the coming decade. The seven regions have different time evolutions of the projected SIC anomalies, however all regions for the MMLE 3+ projections show accelerated SIC loss due to low-frequency variability over the 20 years following 2022 (see Figure 3.8). Taking the pan-Arctic as a whole, the predicted negative anomaly from the linear trend is the largest anomaly at a 5-year lead time during the period 1956-2022. Therefore, our MMLE 3+ model predicts current climate modes as being particularly conducive to a large low-frequency SIC anomaly. This is fairly consistent across LEs, with the only large outlier being the CESM2-LENS which predicts an extreme accelerated loss due to being a large outlier in Central Arctic projections. This outlier is likely due to thin biased ice as discussed in section 3.3.4. Comparing the persistence of CESM2-LENS with CESM2-lessmelt runs which have thicker sea ice (Kay et al., 2022), the lessmelt CESM2 variant is more in line with the persistence in other GCMs (see Figure S3.4). This indicates the low thickness bias likely caused the enhanced simulated variability outlier.

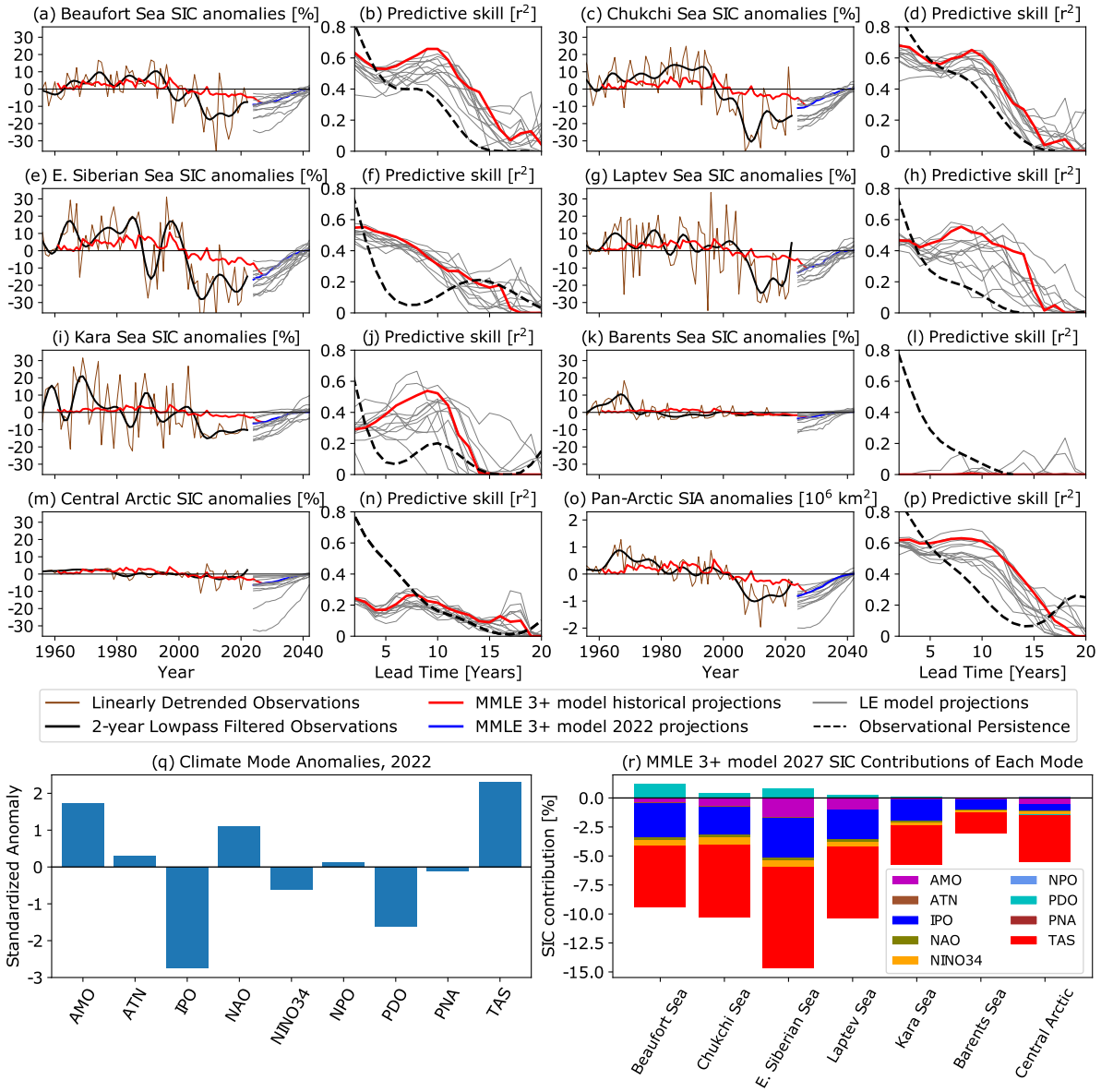

Figure 3.8: **Linear model projections of SIC anomalies based on observed climate modes.** The projection subplots a,c,e,g,i,k,m,o show the observed 1956-2022 regional or pan-Arctic SIC anomalies (brown), the 2-year lowpass filtered anomalies (black), the MML3+ linear model historical hindcasts on a 5-year lead time (red), and the future projections based on the climate mode anomalies observed in 2022 using the MMLE 3+ (blue) and individual LEs (grey). The prediction skill subplots b,d,f,h,j,l,n,p show the observed persistence in dashed back lines while the MMLE 3+ and LE hindcast performances for 1976-2022 at 2- to 20- year lead times are shown in red and gray respectively. The subplot q depicts the observed climate mode anomalies for the year 2022. Subplot r shows the MMLE 3+ contribution to the projected anomalies in 2027 based on 2022 data of each of the modes of variability.

The contributions to this predicted accelerated SIC loss throughout the Arctic in the coming decade is dominated by the large anomalies in 2022 of a negative IPO and strongly positive AMO, alongside a moderately negative NINO34 value. Furthermore, the above trend surface temperature warming in 2022 is also modeled as being a large contribution in the year 2027 (see Figure 3.8q,r). Only the negative phase of the PDO in 2022 is expected to counter the accelerated sea ice loss by leading to positive SIC anomalies in the Pacific sector. The remaining modes of variability are either in near neutral phase in 2022 or have small influences on the linear model and hence do not feature as contributing to future anomalies. The alignment of modes of variability phases in 2022 combine to simulate a negative anomaly to the linear trend larger than any anomaly predicted during the period 1956-2021. Even if for example there were a sudden switch in phase of the NINO 3.4 index from -0.8 in 2022 to +2.0 standard deviation, the SIC % point influence in the East Siberian sea would change the projected 2027 contribution from -0.78% to 1.89%. The more dominant modes of variability, namely the IPO and AMO as well as the TAS influence, transition phase more slowly, which would therefore reduce the sensitivity of projections based on a single year.

3.4 Discussion

Skillful projections of regional Arctic SIC anomalies are able to be produced using simple linear relationships. This has allowed us to identify the individual impact of the dominant modes of variability, namely the IPO, NINO34 and the AMO. However, covariance between the climate modes of variability and nonlinearities in their effect on sea ice are likely to exist (e.g. Heo et al., 2021). This is also supported by the fact that many of the climate modes of variability are not spatially or temporally independent of one another, such as the PDO and NINO34 (Chen and Wallace, 2016). The lack of improvement in skill when we included nonlinearities or covariance in our analysis may be a result of a lack of data (see section 3.3.1), despite the quantity and quality of GCM data available from CMIP6 simulations being unprecedented (Davy and Outten, 2020). The multimodal ensemble approach to learning the drivers of Arctic sea ice variability did not de-

grade our skill when compared with training our linear model on a single GCM (see section 3.3.4). This shows that a generalized model can be obtained from a variety of GCMs differing in model physics, model biases and ocean states. However, large differences in validation between realizations indicates the extent to which the linear relationships themselves can differ due to internal variability.

Previous studies have primarily focused on seasonal or interannual timescales of variability, with the notable exceptions of the IPO and AMO which have been considered on decadal timescales. As these modes of variability persist in one phase for several years to decades it is unsurprising for their influence on sea ice to also persist with a slow near-monotonic decay with time. We found the IPO to be the most influential mode of variability on all lead times between 4 and 20 years, positively correlated with SIC anomalies in all regions. In previous research the IPO's influence on Arctic sea ice has not been featured, except as found by Screen and Deser (2019) for the CMIP5 GCM CESM1. The transition between the negative and positive IPO phase in CESM1 was found to be associated with a strengthened Aleutian Low which enhances poleward heat and moisture transports, facilitating enhanced Arctic sea ice loss. The disagreement in sign and longevity of the IPO's influence on SIC between the CMIP6 LEs follows on from research that CMIP5 GCMs generally poorly simulate the extratropical effects of the IPO (Baxter et al., 2019; Ding et al., 2019; Topál et al., 2020). In addition to the lack of consensus between GCMs broadly, the correlation appears highly sensitive to realization (see Figure 3.7). Previous studies have suggested the variability in the strength of the Pacific-Arctic (PARC) teleconnection may cause the linkages between Pacific SST anomalies and Arctic sea ice to vary substantially and even change sign depending on initial climate state (Bonan and Blanchard-Wrigglesworth, 2020). Additional focus on this mode with a wider variety of modeling applications appears needed and is particularly pressing given the strong current negative phase (see Figure 3.8q).

The AMO was found in our MMLE 3+ linear model to be negatively correlated with all regions of the Arctic sea ice, which shows good agreement with previous studies (e.g. Day et al.,

2012; Miles et al., 2014; Li et al., 2018b) for the pan-Arctic or Atlantic sector on decadal timescales. The suggested physical mechanism for a positive AMO leading sea ice loss is the enhanced winter atmospheric heat transport into the Atlantic sector of the Arctic (Day et al., 2012). However, the AMO itself may have a forced component (Murphy et al., 2021; Cai et al., 2021; Klavans et al., 2022), and its oscillatory timescale varies considerably between GCMs (Lee et al., 2021), potentially limiting the use of the AMO as an independent variable. Despite this, the pre-industrial control simulations (see Figure S3.2) match well with the MMLE 3+ for 1920-2014 for the IPO and AMO, suggesting that forcing context is not highly important for these modes. A much smaller influence of the NINO34 in the pre-industrial control simulations may suggest sensitivity to climate state. Similarly the selection of the dominant season based on the LEs in the period 1920-2014 could also cause an array of responses in different climate conditions and for different GCMs if they simulate other seasons as being most influential.

El Niño and La Niña have been shown to be influential on Arctic sea ice and generally suggest that NINO34 is positively correlated with SIC except for the Beaufort Sea (e.g. Clancy et al., 2021; Hu et al., 2016; Jeong et al., 2022b). However, the lead times considered previously were shorter than our 4- to 20-year timescale, making our positively correlated influence hard to directly compare with previous research. Furthermore, previous literature on shorter timescales have noted the importance of the type of El Niño regime (Jeong et al., 2022a; Lee et al., 2023) and the likely nonlinear and asymmetrical climate response from NINO34 variations (Hoerling et al., 1997). Limitations to the detection of lagged influence from ENSO could be derived from the fact that CMIP6 GCMs still fail to accurately simulate the periodicity and phase-locking characteristics of ENSO (Hou and Tang, 2022). Nonetheless, El Niño favoring a positive Arctic Oscillation and a deepened Aleutian Low via Rossby wave trains, which in turn promote enhanced sea ice export, have been shown to exist in a number of CMIP5 and CMIP6 GCMs (Clancy et al., 2021; Lee et al., 2023). What remains elusive is the persistence of the influence of the NINO34 on sea ice across GCMs up to approximately 14 years (Figure 3.4). This is surprising due to the transition between phases

often occurring interannually, which would therefore require persistent lagged teleconnection pathway to the Arctic which has not been suggested. During 1970-2014 when radiative forcing increase was more linear compared with 1920-2014, we found NINO34 still has a positive correlation on sea ice, but only for approximately 5 years (Figure S3.1). For pre-industrial control simulations, a constant near zero relationship was found (Figure S3.2). It would therefore appear that for periods when nonlinearities in radiative forcing is more effectively removed by detrending or by constant forcing, the lagged NINO34 influence sea ice is substantially reduced. As no physical mechanism for a lagged sea ice response to a given phase of NINO34 is suggested, the influence of NINO34 encapsulating residuals in the model or nonlinearities in warming must be considered.

The PDO was previously not found to be highly important for Arctic sea ice by itself (Zhang et al., 2020; Karami et al., 2023) and its influence may also have changed over time (Bonan and Blanchard-Wrigglesworth, 2020; Kim et al., 2020). Similarly we also only found a small influence of the PDO over our long time period of 1920-2014. The SST pattern of the PDO is related to the Pacific SST climate modes of the NINO34 and IPO, which each have similar teleconnection mechanisms via Rossby wave train formation (Yuan et al., 2018). Further, the mechanisms of an oceanic 'tunnel' and atmospheric 'bridge' may be common between the IPO and PDO, leading to the similar characteristics of persistence, Arctic influence, and modulation by ENSO (Liu and Alexander, 2007; Henley et al., 2017). Therefore it is unsurprising for a subset of tropical Pacific modes of variability to dominate due to their similarity. The ATN is the only negligible mode of variability derived from SSTs, but has not previously been identified as specific driver of Arctic sea ice variability. However the tropical Atlantic was been suggested to influence Arctic sea ice (Meehl et al., 2018). Thus, the unimportant nature of the ATN does not preclude other aspects of tropical Atlantic being important.

Consistent with the previous lack of evidence of influence beyond interannual timescales, the sea level pressure-derived modes of variability (the NAO, NPO and PNA) were found to have negli-

gible effect at lead times of 4-20 years. Previous research has shown the effect of the NAO to decay to zero after approximately 2 years (Ukita et al., 2007), this timescale and the regional correlations (e.g. Serreze et al., 2007; Döscher et al., 2010) align with our findings given the smoothing inherent in our lowpass filtered data. This provides confidence in our linear model's ability to capture higher frequency variability but dismiss low-frequency influence from these modes of variability.

3.5 Conclusions

We have shown that low-frequency variability of regional Arctic sea ice concentration can be modeled using linear drivers consisting of climate modes of variability. We achieve predictions superior to persistence for most regions for a lead time of 4-20 years and find that the climate modes of variability can be considered independently without reducing skill. By comparing the linear responses between twelve large ensembles from CMIP6 and a multi-model large ensemble comprising of 42 GCMs, we find where there is consensus of the dominant linear drivers of low-frequency sea ice variability except in the case of the Interdecadal Pacific Oscillation (IPO). In the pan-Arctic we are able to explain up to 60% of observed low-frequency sea ice concentration variability at lead times of 5 years. However, the ability of a GCM or a multi-model large ensemble to predict unseen ensemble members or observations can vary wildly depending on the realization of internal variability; between 0-0.46 r^2 in the case of CanESM5, the largest single GCM ensemble. Hence, this both complicates the analysis of small samples of GCMs and the application and verification of these relationships with our single realization of observations.

The most important modes of variability we found were the IPO, Nino 3.4 Index (NINO34) and the Atlantic Mutidecadal Oscillation (AMO). The multi-model large ensemble linear model showed the IPO to have a strong positive correlation with this being most pronounced in the East Siberian, Beaufort and Laptev Seas at lead times of up to 14 years. Although this large magnitude of influence of the IPO was found across GCMs, the sign and regional influence was especially

dependent on the GCM used and the specific realization of internal variability. NINO34 was found to be positively correlated with SIC anomalies in all regions, particularly in the Pacific sector. This correlation was more consistent between GCMs, with disagreement growing at longer lead-times. The AMO was the only other mode of variability considered important for long periods of time, being modeled as highly negatively correlated with SIC across all regions for up to approximately 10 years. However, the agreement across CMIP6 GCMs for the AMO was less consistent than NINO34. The persistence of influence on sea ice from the IPO and AMO is not surprising due to their similar phase persistence. NINO34 in contrast can change phase interannually, suggesting the influence found here may be an artifact of forcing conditions, as no physical mechanism for a delayed Arctic influence has been suggested.

When using our linear model to make predictions, we find a near 'perfect storm' of modes of variability in the year 2021/2022 to induce an acceleration to the sea ice loss trend over the next decade. The primary influences of this projected acceleration of low-frequency variability driven sea ice loss are an above trend global average surface temperature warming, a negative IPO, La Niña conditions, and a positive AMO. For the pan-Arctic, the projected low-frequency deviation from the long-term trend due to current climate mode phase configurations is expected to be the largest since at least 1956. While the transition between La Niña and El Niño can occur rapidly, the fact that our strong negative predictions are primarily due to the slowly changing climate modes of the AMO and IPO imply robustness of this prediction to interannual variability. Of course, the sea ice anomalies that will actually be observed are still dominated by interannual variability, which makes up roughly three quarters of the total variability. Thus, while we cannot say with confidence that a new record low September extent will occur over the next decade, the modeled low-frequency variability suggests that extreme low SIC values will be more likely over the coming decade, with low-frequency variability likely to enhance the long-term negative trend.

3.6 Acknowledgements

This work was supported by the National Science Foundation under grant number 1847398. A. Jahn's contribution was also supported by an Alexander von Humboldt fellowship. We would also like to acknowledge high-performance computing on Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. Further, we thank Adam Philips for processing updates to the Climate Variability Diagnostics Package output used in this analysis. We also thank Elizabeth Barnes and David Hall for useful conversations in the early part of the project.

3.7 Data Availability

The data used in this study is freely available at https://esgf-node.llnl.gov/projects/ cmip6/ for the CMIP6 global climate model data except for the CESM2-LENS, which is available at https://www.cesm.ucar.edu/community-projects/lens2/data-sets. The CVDP data is available at https://www.cesm.ucar.edu/ projects/cvdp/data-repository. The observational sea ice concentration data is freely available from https://nsidc.org/data/g02202/versions/4. Upon publication, all code required to replicate this study will be made open-access at Zenodo, currently it is available at https://github.com/chrisrwp/low-frequency-variability. Additionally the linear model coefficients and the observational CVDP output data processed for this analysis will be made available at the Arctic Data Center upon acceptance of this manuscript.

3.8 Supplementary Material

3.8.1 Machine Learning Method Details

3.8.1.1 Splitting for training, validation and testing

For each LE we divide the members into training, validation, and testing sets with 75%, 15%, 10% of members respectively. For the MMLE 3+, we then use the first member for the training data set, the second member for the validation set, and leave the third and any other members for testing. For the MMLE 30+, we pool the first 23 members from all 8 GCMs for training, we use the next 4 members for validation, and the final 3 or more members for testing. As we use 74 years of data for each ensemble member (1920-2014) the smallest LE uses 74 years with 21 ensemble members, yielding an effective 1554 years for training - far in excess of observations and typically longer than pre-industrial control runs from any individual GCM. The MMLE3+ has 2294 years of training data and the MMLE 30+ maximizes the number of training years at 13,320, allowing us to determine whether substantially increasing the training data provides any gain in predictive skill.

3.8.1.2 Neural network configurations

All of the four machine learning models (see section 3.2.2 for a physical explanation of their utility) use a fully-connected neural network with the same L1 loss function to encourage sparseness and an Adam optimizer for suitability to the four diverse models. We selected models 1 and 2 to have no bias term, which ensures the zero values of the standardized input variables (a neutral phase of the climate mode) predicts a zero SIC anomaly. With these four machine learning models, as detailed below, we can separate the effect of linear/nonlinear activation functions from the effect of additional neural network layers which allows one climate variable to interact with another:

- Model 1 - Model layers: 9-1 with linear activation functions and no bias.
- Model 2 - Model layers: 9-1 with nonlinear (ReLU) activation functions and no bias.

- Model 3 - Model layers: 9-3-3-1 with linear activation functions with bias.
- Model 4 - Model layers: 9-3-3-1 with nonlinear (ReLU) activation functions with bias.

By comparing the predictive skill of model 1 versus 2 and model 3 versus 4 we can identify the effect of increasing the model complexity from a linear to nonlinear activation functions. This is because the only difference between those two groups is the activation function, analogous to linear or nonlinear relationships between the climate modes of variability and SIC. Then, by comparing the predictive skill of models 1 versus 3 as well as model 2 versus 4, we can determine the difference in allowing the climate modes of variability to be independent of one another. This independence is facilitated in the simple models 1 and 2 where each of the 9 neurons in the input layer connects directly with the output layer. Models 3 and 4 which take into account covariance of different climate modes, this is achieved by connecting the input layer to two hidden intermediate layers of 3 fully-connected neurons before reaching the output layer.

3.8.1.3 Determining dominant seasons

We first compute our machine learning models on four seasonal values for each of our 9 input variables for the LE datasets and find the most dominant season averaged over all 7 regions of the Arctic for September. This is done by using model 1 with model layers of 36-1 and a linear activation function, and then selecting the largest seasonal input. We considered this reduction in seasons necessary due to the slowest changing climate modes such as the IPO having very similar seasonal values, resulting in over-fitting. Using all 4 seasons for some climate modes and a single season for one would not allow a fair comparison between climate modes, while using only one season rather than four does not substantially decrease the predictive skill. However, there is a limitation of this approach, whereby the dominant season for a given climate mode may differ between the GCMs, regions, and SIC anomaly months. Further, 'best' seasons for some climate modes of variability have little physical meaning as no season has much influence on Arctic sea ice.

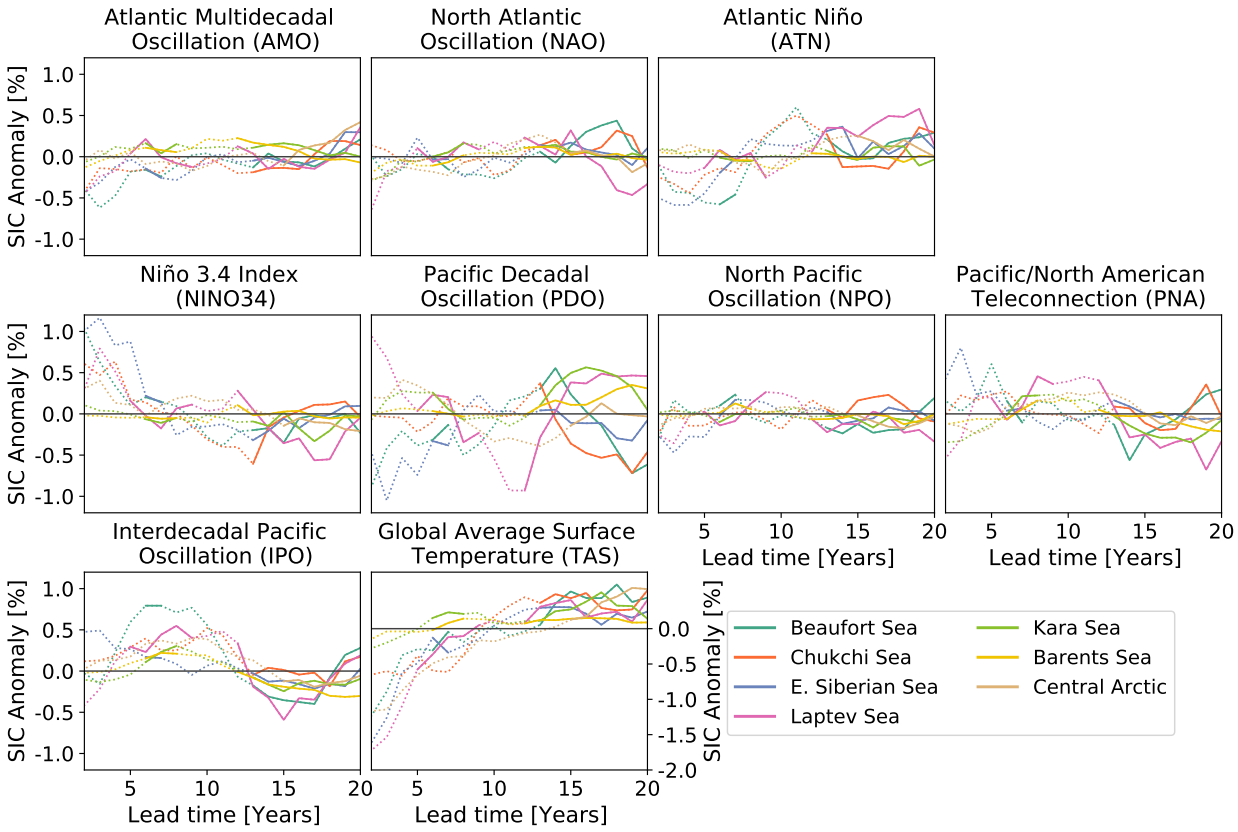

Figure S3.1: **Linear drivers of regional sea ice concentration anomalies for a reduced time period.** Same as Figure 3.4, except for the reduced time period of 1970-2014 instead of 1920-2014. By comparing this figure with Figure 3.4, we can see that the modes of variability have a similar influence as for the 1920-2014 time period, although the results are far more noisy and predictive skill does not exceed persistence for as much of the lead times as for the period 1920-2014.

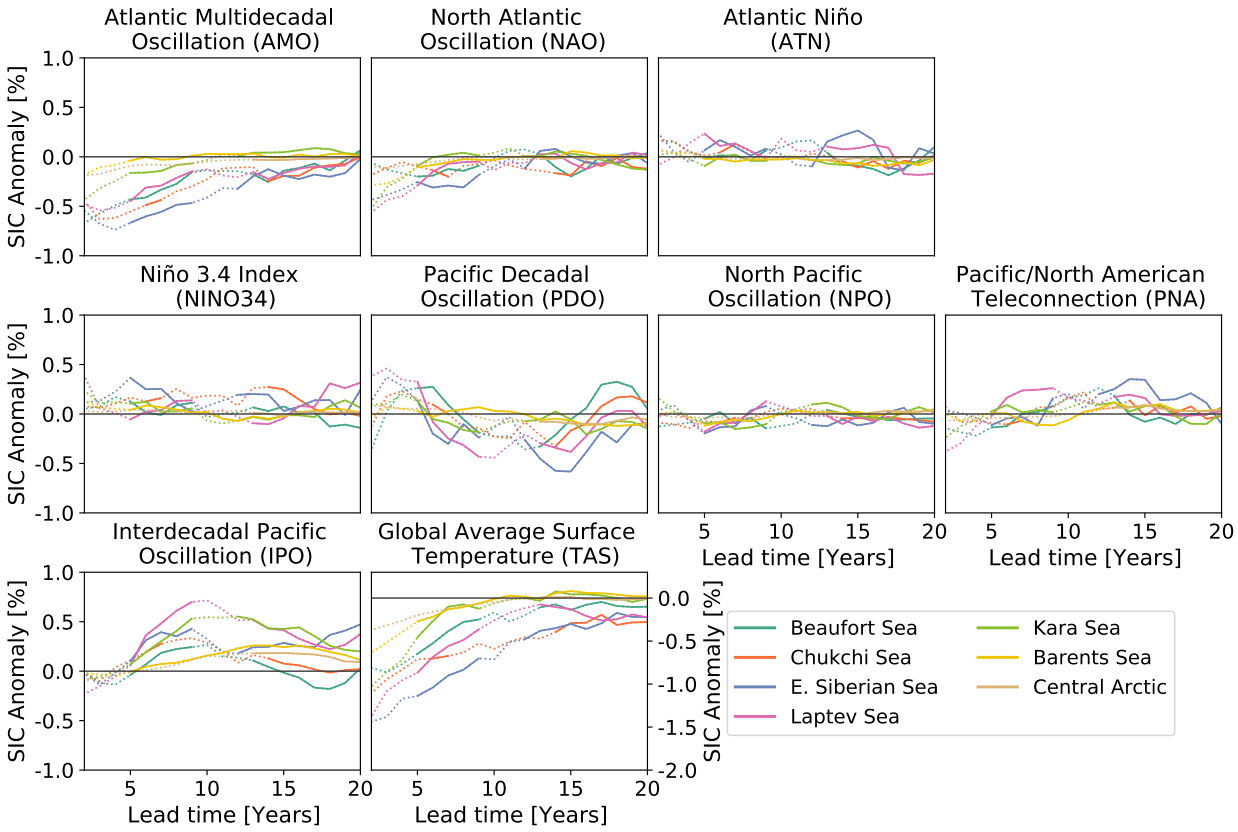

Figure S3.2: **Linear drivers of regional sea ice concentration anomalies for pre-industrial control runs.** Same as Figure 3.4 and S3.1, except here using the 1850 control simulations instead of the period 1920-2014 in the historical simulations. As for Figure S3.1, the influence of the climate variability modes are very similar as for the period 1920-2014 (Figure 3.4), but the coefficients are smaller, likely due to the lower variability in the pre-industrial mean state. Instead of different ensemble members, the available 35 GCMs are each split into several members of 74 year length each, with the first 222 years used for training and the following 74 years for validation.

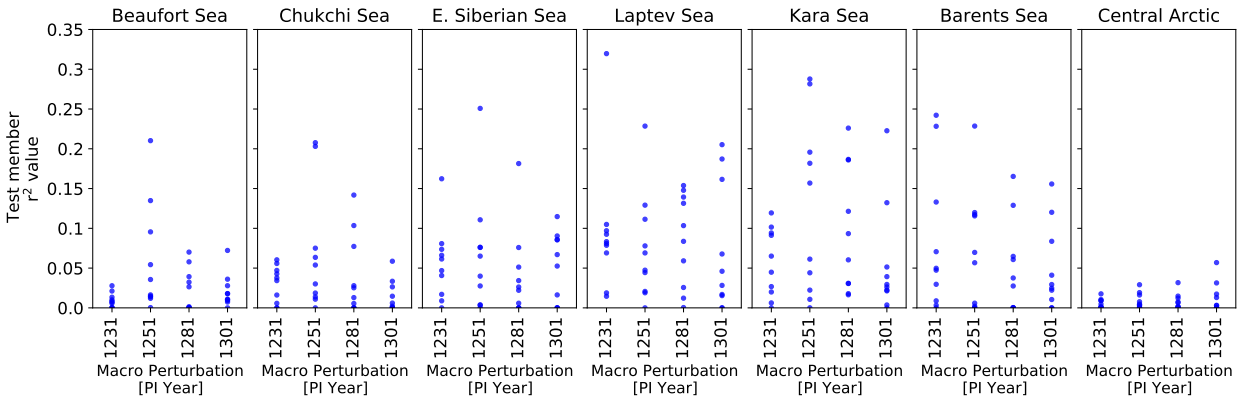

Figure S3.3: **Influence of macro versus micro initializations in the CESM2-LENS on September test member r^2 values**. Of the 48 test members from the CESM2-LENS (see Figure 3.6), 12 are created through macro initializations by choosing different start years from the pre-industrial simulation, and hence differ in their ocean and atmospheric state. Of those 12, four (here shown on the x-axis by branch year) have 9 additional ensemble members branched from them, which all only differ slightly in their atmospheric state due to small atmospheric perturbations, i.e., referred to as micro initializations. Here we show these latter 40 simulations (blue circles), to assess whether macro or micro initializations dominate the possible r^2 values (with persistence removed). As the four distributions of 10 realizations for each macro initialization are very similar, this shows that the ocean state (macro perturbation) can influence the predictive skill, but generally does not narrow the potential range of r^2 values which can occur due to micro perturbation.

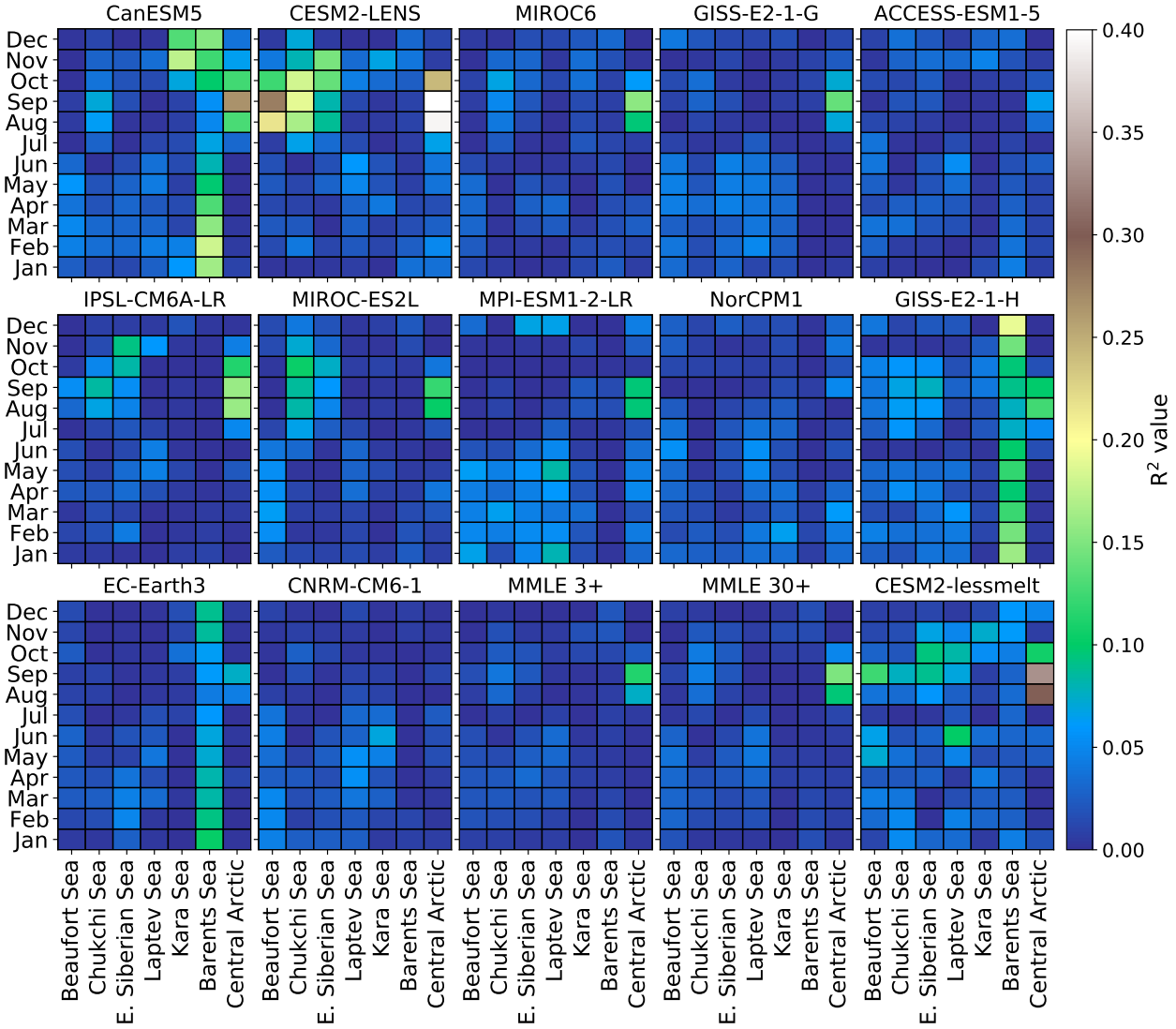

Figure S3.4: **Persistence r^2 values for LEs, MMLEs, and CESM2-lessmelt at a 5-year lag time.** This figure shows the persistence r2 value that was subtracted from the absolute value of the validation r^2 in Figure 3.3. Additionally the CESM2-lessmelt persistence is shown for comparison with CESM2-LENS. CESM2-lessmelt has a thicker sea ice mean state than CESM2-LENS and, as shown in this figure, has a smaller persistence validation r^2 value, although this value is still an outlier compared with the other GCMs.

CHAPTER 4

FOLLOWING AN INITIALLY NEGATIVE PHASE OF THE INTERDECADAL PACIFIC OSCILLATION 75% OF LARGE ENSEMBLES SHOW AN EARLIER ICE-FREE ARCTIC OCEAN

Preface

This chapter is a manuscript submitted to Geophysical Research Letters (manuscript number: 2024GL109158). The authors are Wyburn-Powell, C. and Jahn, A.

Abstract

The Arctic is transitioning towards a seasonally ice-free state over the next few decades, with internal variability determining precisely when. The phase of the Interdecadal Pacific Oscillation (IPO) has been suggested as a way to reduce internal variability uncertainty. We show that in a selection of 8 large ensembles from the Coupled Model Intercomparison Project - Phase 5 (CMIP5) and CMIP6, ice-free conditions occur earlier in 75% of these ensembles for an initially negative IPO phase. However, the effect of the initial IPO phase is modulated by the Pacific-Arctic Teleconnection (PARC) in a majority of models, but with little consensus on sign and magnitude. An initially negative PARC phase is associated with reduced sea ice concentrations preceding ice-free conditions for all LEs except the CanESM models. For the negative initial IPO and PARC teleconnection that occurred in 2023 observations, we find 75% of LEs simulate an earlier than average ice-free Arctic.

Plain Language Summary

The Arctic Ocean is expected to become effectively free of sea ice in September consistently during the mid-21st century. There is a large amount of uncertainty on precisely when this will occur, due both to uncertainty in our climate models, as well as the chaotic nature of the climate system, called internal variability. Some of this internal variability can be explained by long-lasting variations in the Pacific Ocean and atmosphere. We find that 75% of climate models simulate an ice-free Arctic occurs earlier when the Pacific Ocean is in a decadal climate state similar to what we observe today, namely a negative phase of the Interdecadal Pacific Oscillation (IPO). How the Pacific Ocean affects the Arctic atmosphere also changes on decadal timescales, described by the phase of the Pacific-Arctic teleconnection (PARC). When matching both the IPO and PARC teleconnection phase to current observations, an earlier ice-free Arctic is still simulated by 75% of models. This suggests that due to the current state of Pacific climate variability, the Arctic will be effectively free of sea ice earlier than the ensemble-mean prediction.

4.1 Introduction

Arctic sea ice is projected to rapidly decline over the next few decades, likely becoming seasonally ice-free by mid-century, independent of emissions scenario (Notz and SIMIP-Community, 2020; Årthun et al., 2021; Jahn et al., 2024). Internal climate variability introduces large variations to the the sea ice trajectory over interannual to decadal timescales (Kay et al., 2011; Wettstein and Deser, 2014; Swart et al., 2015; Bonan et al., 2021). Single-model initial condition large ensembles (LEs) can isolate the contribution of internal variability (Fasullo et al., 2020). The timing of an ice-free Arctic can vary by over two decades based solely on the climate realization (Notz, 2015; Jahn et al., 2016), with this simulated range of internal variability differing somewhat between

LEs and ice-free definitions (Holland and Hunke, 2022; Jahn et al., 2024). Thus, understanding the predictability and magnitude of internal variability at decadal timescales, can help to reduce overall projection uncertainty of consistently ice-free conditions.

The Interdecadal Pacific Oscillation (IPO) has been suggested as an important driver of Arctic sea ice variability, influencing the time of an ice-free Arctic by 9 years in the Community Earth System Model version 1 (CESM1) (Screen and Deser, 2019). The transition from a negative to positive phase of the IPO was associated with a persistent strengthening of the Aleutian Low in CESM1, favoring enhanced poleward heat and moisture transports, facilitating faster Arctic sea ice loss. However, whether the IPO influences ice-free conditions in the Arctic in other CMIP5 and CMIP6 LEs is so far unknown. The non-stationary Pacific-Arctic (PARC) teleconnection has also been suggested to substantially modulate the strength and sign of tropical Pacific SSTs' effect on Arctic geopotential heights and sea ice (Ding et al., 2014; Baxter et al., 2019; Bonan and Blanchard-Wrigglesworth, 2020). However, the ability of CESM1, and CMIP5 models more broadly, to simulate tropical-Arctic teleconnections similar to observations has been called into question (Baxter et al., 2019; Blanchard-Wrigglesworth and Ding, 2019; Topál et al., 2020). In CMIP6 models, earlier ensemble mean ice-free conditions were simulated for the models with a strong Pacific-Arctic connection (Jahfer et al., 2024).

With the recent advent of multiple very large ensembles (Maher et al., 2019), isolating the independent and combined effects of the IPO and PARC is now possible in several models. Thus, we here show that a negative initial IPO indeed leads to earlier ice-free conditions in 6 out of the 8 LEs. Furthermore, we find that the phase of the PARC teleconnection modulates the effect of the IPO on Arctic sea ice. However, differences between model families emerge regarding how the PARC phase affects the timing of an ice-free Arctic under different initial IPO phases. Thus, to reduce the projection uncertainty of a consistently ice-free Arctic, it needs to be determined which, if any, model family represents the Pacific-Arctic teleconnection most realistically.

Table 4.1: List of Global Climate Models used, listed with the number of ensemble members that have the data necessary for calculating all pertinent analysis metrics.

Modeling Center	Model Name	Members	CMIP era	Citation
CSIRO	ACCESS-ESM1-5	40	CMIP6	**?**
CCCma	CanESM5	49	CMIP6	**?**
NCAR	CESM2-LENS	50	CMIP6	Danabasoglu et al. (2020)
SMHI	EC-Earth3	50	CMIP6	Wyser et al. (2021)
MIROC	MIROC6	50	CMIP6	Shiogama et al. (2023)
MPI	MPI-ESM1-2-LR	30	CMIP6	**?**
CCCma	CanESM2	50	CMIP5	Kirchmeier-Young et al. (2017)
NCAR	CESM1	40	CMIP5	Kay et al. (2015)

4.2 Data and Methods

We use eight climate models which provide large ensembles for our analyses, see Table 4.1. The large ensembles selected are those from CMIP5 and CMIP6 which provide at least 30 ensemble members with output sufficient to obtain sea ice concentration, sea ice extent (SIE), the IPO phase, and the PARC teleconnection over both the historical and future scenario (RCP 8.5 or SSP3-7.0). For observations, we use the National Snow and Ice Data Center (NSIDC) sea ice index (Fetterer et al., 2017) as observed SIE. Sea surface temperatures (SSTs), used for calculating the IPO, are from HadISST (Rayner et al., 2003). Geopotential height reanalysis, used for calculating the PARC teleconnection, are taken from ERA-5 (Hersbach et al., 2023).

We assess projections of consistently ice-free conditions, by defining the ice-free year as the first year an 11-year centered mean September SIE drops below 1 million square km, following Screen and Deser (2019). When calculating the IPO, we use the Henley et al. (2015) definition, using a tri-polar index of SST anomalies. The PARC teleconnection is calculated over a 40-year period as the correlation between the summer (JJA) Pacific SST anomalies in the region (5-20°N and 180-120°W), and the 200 hPa geopotential height anomalies in the Pacific sector of the Arctic (70-90°N and 180-90°W), following Bonan and Blanchard-Wrigglesworth (2020). The current observational mean state SIE was 4.7 million km^2 for the 11-year running average in both 2017 (as

for Screen and Deser (2019)) and for 2023 (Fetterer et al., 2017).

For our initial analysis of the influence of the IPO alone (Figure 4.1) we follow the methodology of Screen and Deser (2019). This involves selecting all members starting at 4.7 million km^2 and classifying their IPO phase based on 0.9 standard deviations from the mean. For subsequent analyses, to better align the IPO phase, incorporate the PARC teleconnection and pool multiple trajectories with different initial SIE values we diverge from the method of Screen and Deser (2019). In this case we classify the IPO phase based on persistence above/below zero for 10 years. This is done to better align the subsequent transitions of the IPO to the alternate phase. Furthermore, this definition aligns with the current weak but persistent negative IPO phase. The PARC is classified as positive or negative based on the sign of the correlation over the preceding 40 years, following Bonan and Blanchard-Wrigglesworth (2020). Further, the pooling of multiple trajectories is achieved by selecting 11 trajectories at 4.2-5.2 million km^2 at 0.1 increments. If a specific IPO and PARC phase combination does not occur in any member for a given initial SIE, (as is the case for 6% of trajectories) the ensemble mean is used instead, to facilitate averaging across the initial trajectories. For MPI-ESM1-2-LR there is no ensemble member where both the positive phase of the IPO and negative phase of the PARC exists (see Table S4.1), and hence MPI-ESM1-2-LR is removed from the analysis in section 4.3.2.

Statistical significance is calculated at the 95% level using a student's t-test. For significance regarding just the IPO or PARC teleconnection's influence, the difference between the two respective initial states is computed. For the smaller samples considering the simultaneous IPO and PARC states, the difference from the mean is computed for significance.

4.3 Results

4.3.1 The Influence of the IPO on Ice-free Conditions

The initial phase of the IPO is only statistically significant in determining the trajectories of September SIE for a minority (37.5%) of LEs. Following the methodology by Screen and Deser (2019) for CESM1, significant differences in SIE are also found for MPI-ESM1-2-LR and EC-Earth3 (see Figure 4.1). For both CESM1 and MPI-ESM1-2-LR the initially negative phase of the IPO is seen to produce lower SIE when compared with trajectories which start with an initially positive phase of the IPO. EC-Earth3, in contrast, simulates lower SIE trajectories when the initial positive phase of the IPO occurs. For the LEs that do not show a statistical significant divergence in SIE trajectories, the initially negative phase of the IPO generally produces lower SIE (see Figure 4.1). When ice-free conditions approach, divergence between the trajectories is seen to converge for these LEs. Overall, there is little consensus between the LEs on the predictive skill of the IPO alone in determining SIE trajectories towards an ice-free Arctic, when following the methodology of Screen and Deser (2019).

Broadening the analysis of the IPO's influence from one initial mean state to the average of 11 mean states, and looking at states with persistent IPO phases, significant trajectory divergence now occurs for 75% of LEs. The LEs which still do not show any significant divergence between the two SIE trajectories with positive and negative IPO phases are CanESM2 and CanESM5 (see Figure 4.2A). However, even for the LEs that now show significant divergence between the two SIE trajectories, the periods of significant difference occur at different times. For some (CESM2-LENS and MPI-ESM1-2-LR), the significance occurs during the first decade, while for the majority (CESM1, EC-Earth3, ACCESS-ESM1-5, MIRCO6), the significant difference occurs in the second decade. This shows that while the IPO has an effect on Arctic SIE, its utility for narrowing the prediction uncertainty for an ice-free Arctic depends on the number of decades it takes to go from 4.2-5.2 million km^2 to 1 million km^2 in the different models. Although trajectory divergence occurs

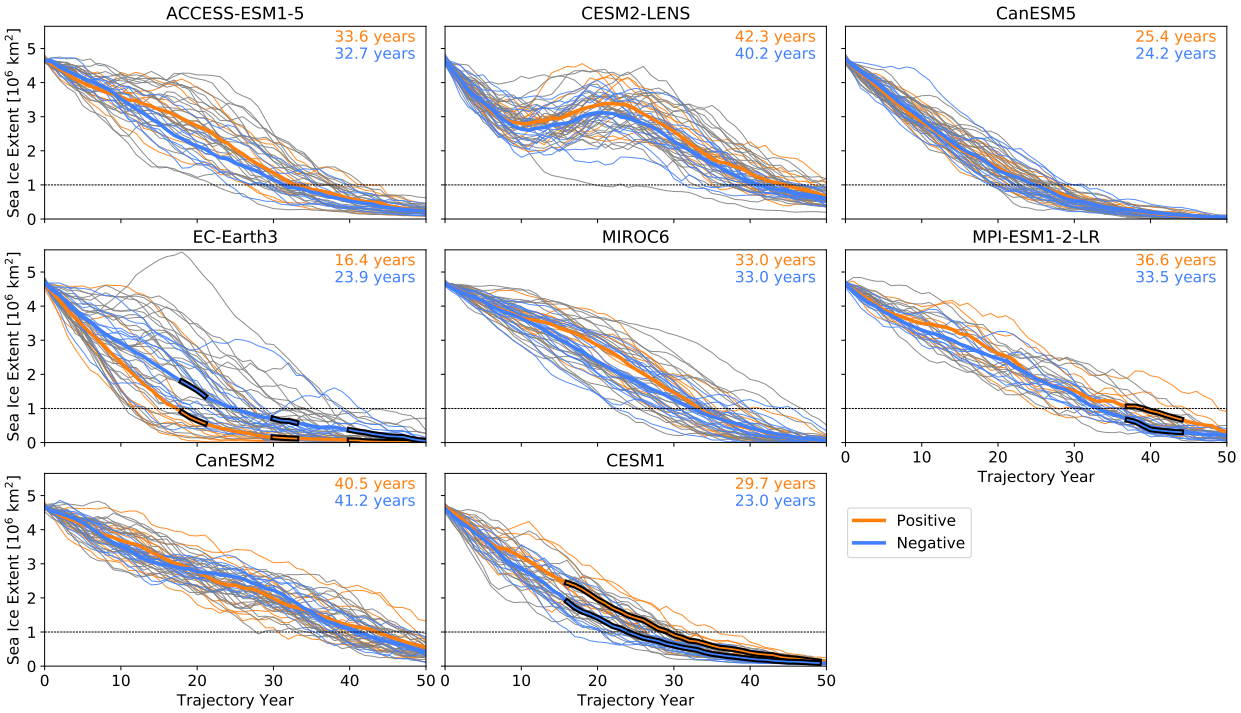

Figure 4.1: **SIE trajectories starting at 4.7 million square kilometers, categorized based on initial IPO phase with a threshold of 0.9 σ, following Screen and Deser (2019).** The thin orange lines indicate trajectories where the initial IPO exceeds 0.9 σ and the bold orange line is the average of the positive IPO trajectories. Similarly the blue lines are for initial IPO values of -0.9 σ and lower, grey indicates trajectories with an initially neutral IPO phase. The numbers in the top right of each subplot indicate the average trajectory year before the first ice-free conditions for the average trajectory starting with a positive or negative IPO phase. Periods where the IPO phase trajectory averages are significantly different from one another are bounded in black.

at different times, the period 17-22 years coincides with when the initial IPO phase has changed into the opposite or neutral phase (see Figure S4.2). Contrasting this period to trajectories over the period 2-7 years, can thus lead us to assess the impact of the IPO phase transition.

During the period 17-22 years of the trajectories, there is clear spatial evidence of the influence of initial IPO phase across the LEs. For all of the LEs there are noticeably higher SIC for trajectories with an initially positive phase of the IPO compared with trajectories starting with a negative phase of the IPO. Such a large difference is not seen in the 2-7 year time period, indicating the initial phase of the IPO does not greatly impact the first decade of trajectories. Instead, it is the change in IPO phase that impacts the sea ice cover most strongly, in agreement with Screen and Deser (2019). Spatially, the SIC anomalies during the 17-22 year time period vary between the LEs, but for all LEs, except CanESM2 and CanESM5 which do not show any significant divergence, there is symmetry between the initially positive and negative IPO trajectory averages. This indicates that both phases of the IPO act upon Arctic sea ice. The influence of the IPO at time periods of 17-22 years gives an insight into the strong effect that in particular changes in the IPO can have on Arctic sea ice.

For 75% of the LEs, an initially negative phase of the IPO leads to earlier ice-free conditions relative to the ensemble mean. However, it is only CESM1 and CESM2-LENS for which the initially negative phase of the IPO is significantly earlier than the initially positive phase (see Figure 4.3A). MIROC6 and EC-Earth3 show the opposite influence of the initially negative IPO, with the difference between the IPO phases significant only for EC-Earth3. Overall the majority of the LEs (75%) agree with CESM1, but CESM1 and CESM2-LENS are particularly sensitive to the effects of the IPO on ice-free conditions. Trajectories selected for either initial phase of the IPO tend to lead to earlier ice-free conditions than the ensemble mean which considers all IPO states, which is particularly apparent for CanESM2. In the following, section 4.3.2, we examine the combined and independent effect of the PARC teleconnection to further assess how the Arctic sea ice is affected by conditions in the Pacific.

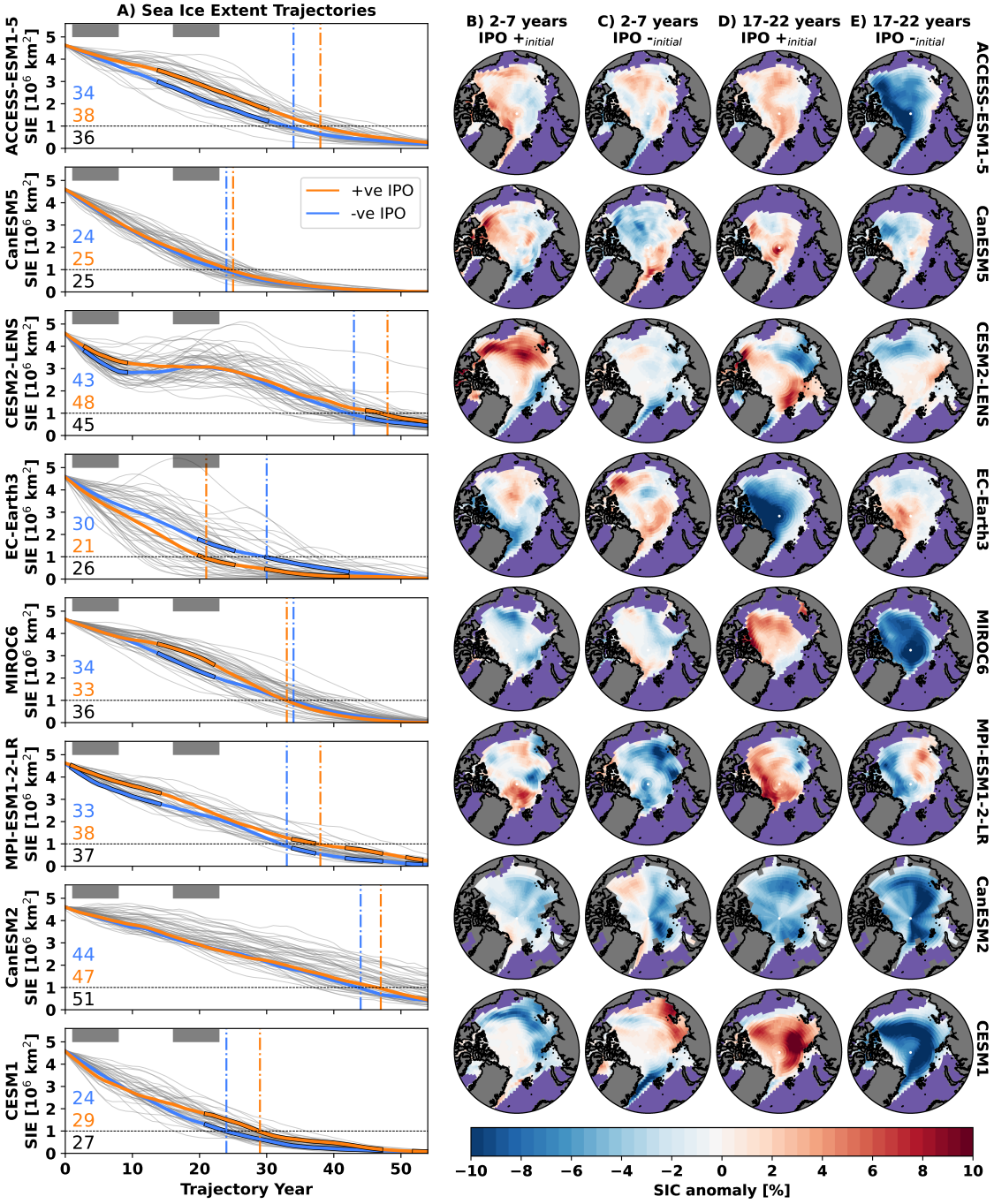

Figure 4.2: **Sea ice trajectories categorized by initial IPO phase.** A) The SIE trajectories from 11 initial SIE states of 4.2-5.2 million km^2 selected based on persistence of the IPO in a positive or negative phase for at least 10 years. When the two trajectories are statistically different from one another they are bounded in black. Light grey lines indicate the individual ensemble members. B), C) Sea ice concentration anomalies in years 2 to 7 of the trajectories relative to the ensemble mean, for initially positive (B) and negative (C) IPO phases. D), E) is the same as B), C) except for the averaging period of 17 to 22 years. Purple shaded regions indicates open ocean where no ensemble member exceeds 15% concentration over the time period.

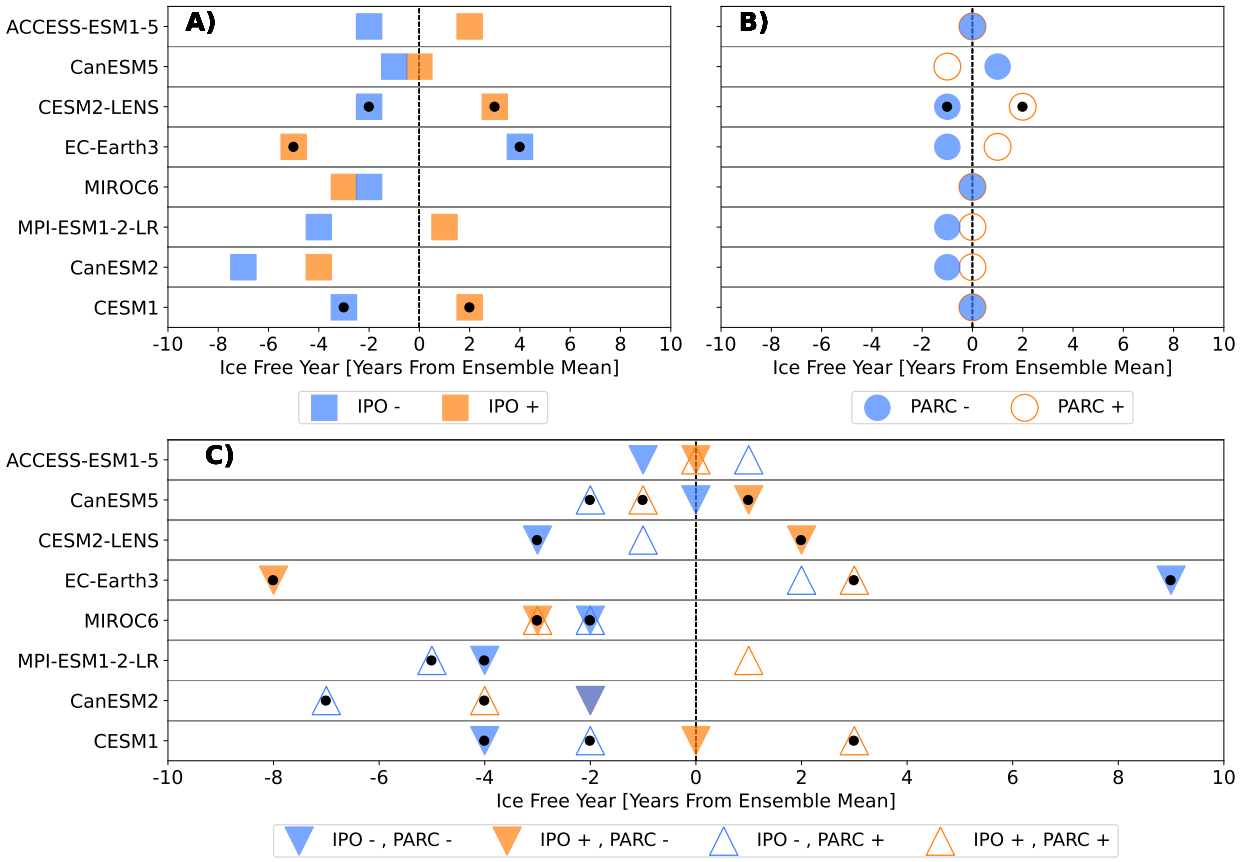

Figure 4.3: **Ice-free years based on initial IPO phase and PARC teleconnection state, relative to the ensemble mean.** A) Average ice-free years selecting ensemble members based on the initial IPO phase. B) as A) with ice-free years categorized by PARC phase. For both A) and B) significance is denoted by a black dot, representing the two trajectories are significantly different from one another. C) Ice-free years selected by both the initial IPO phase and the PARC teleconnection state, significance is denoted by black dots, indicating the ice-free year is outside of the 95% confidence interval of the ensemble distribution. Note there is no instance of the IPO positive and PARC negative for MPI-ESM1-2-LR, hence that category is excluded from C).

4.3.2 Modulation by the PARC teleconnection

The impact of both the IPO and the PARC teleconnection phases on ice-free years is greater than the sum of each of their independent effects. For half of the LEs a negative phase of the PARC teleconnection alone leads to earlier ice-free years, with this difference being significant only for CESM2-LENS (see Figure 4.3B). This independent influence of the PARC teleconnection on ice-free years is small in comparison to the impact of the initial phase of the IPO (with a maximum effect of 3 years versus 9 years, respectively). When the combined influence of the IPO and PARC teleconnection are considered, in 75% of LEs, a wider range of ice-free years are found (see Figure 4.3C). While no difference in ice-free years occur in CESM1 dependent only on the PARC teleconnection phase, the modulation of the IPO by the PARC increases the range by 2 years. Similarly to CESM1, a negative PARC leads to earlier ice-free conditions in CESM2-LENS. In contrast, for CanESM2 and CanESM5 it is the modulation of the initial negative IPO by a positive PARC that leads to earlier ice-free conditions. EC-Earth3 behaves still differently, with the negative PARC amplifying the signal from both phases of the IPO. The three remaining LEs (ACCESS-ESM1-5, MIROC6 and MPI-ESM1-2-LR) show little sensitivity to the independent effect of the PARC teleconnection or its modulation of the initial IPO phase in regards to ice-free year. Thus, there is inconsistent modulation of IPO's effect by the PARC on ice-free timing across models.

In contrast to the lack of consensus of the PARC teleconnection's effect on ice-free timing, agreement exists on the effect of SIC anomalies in the lead-up to ice-free conditions. The PARC teleconnection modulation, where a positive PARC increases SIC and the negative phase decreases SIC, is common when approaching ice-free conditions, aside from the CanESM LEs (see Figure 4.4). This agreement between the LEs aside from CanESM2 and CanESM5 on the PARC influence is apparent despite the low consensus of the IPO's spatial effect when approaching ice-free conditions. Notably EC-Earth3 displays the common PARC modulation of positive SIC anomalies with a positive phase, despite the IPO effect in EC-Earth3 being opposite to the majority of the

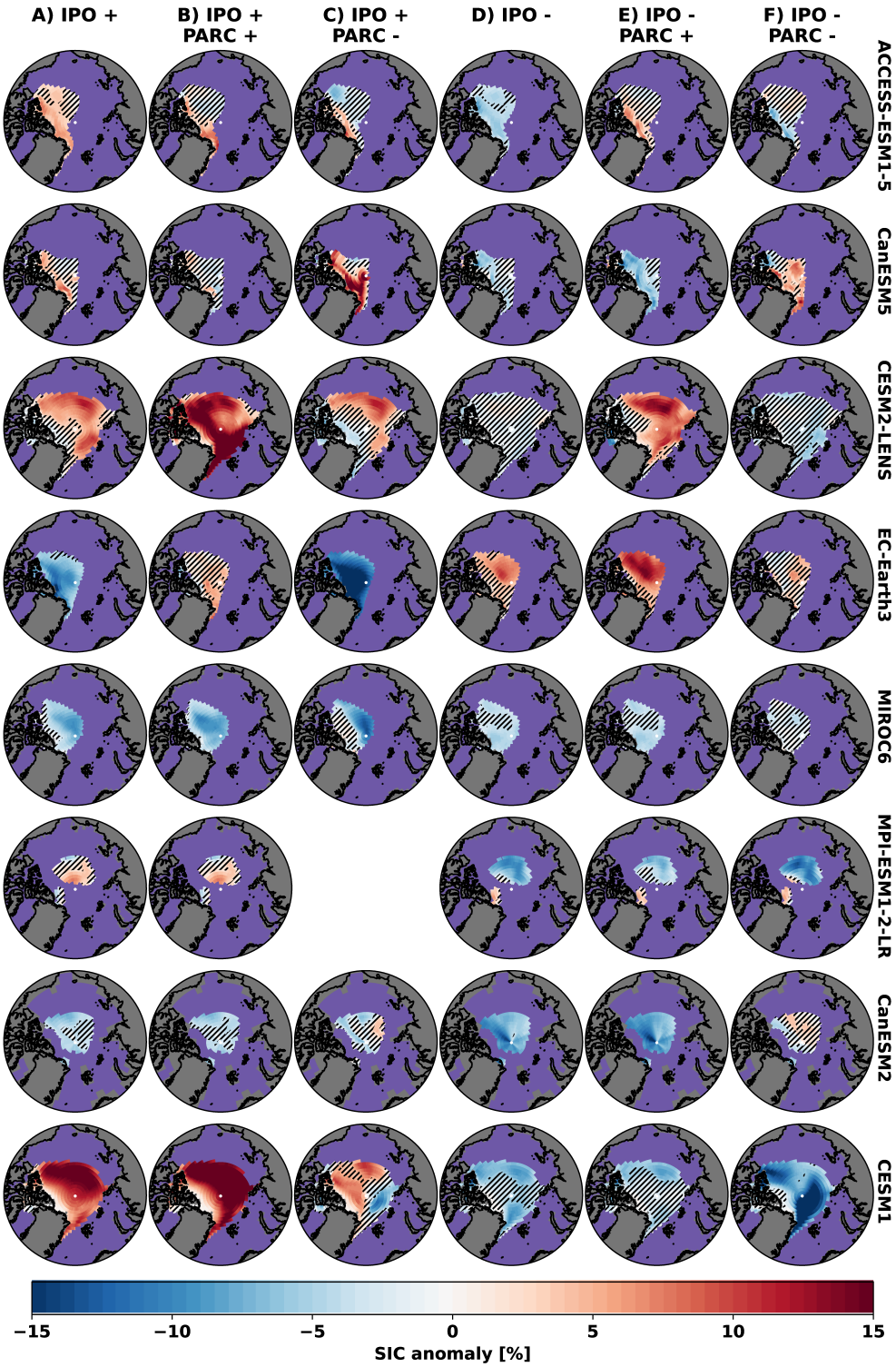

Figure 4.4: **Sea ice concentration anomalies relative to the ensemble mean for the 5 years leading up to the average trajectory year of ice-free conditions.** Black lines indicate the difference is not significant from the ensemble mean at the 95% level. Purple shaded regions indicates open ocean where the ensemble mean does not exceed 15% concentration 5 years before ice-free conditions. Note MPI-ESM1-2-LR has no ensemble members where an initially positive phase of the IPO is coincident with the positive phase of the PARC teleconnection, leading to this subplot being missing from the figure.

LEs. The initially positive phase of the PARC is also regionally associated with increased SIC in the Greenland Sea, which is especially notable for CESM2-LENS and EC-Earth3. This region is still shown to be modulated by the PARC teleconnection phase in CanESM5, despite the opposite SIC sign to the other LEs sensitive to PARC modulation. For all LEs, considering the PARC modulation of the IPO signal leads to larger SIC anomalies than considering the IPO on its own. The most sensitive LEs to the initial IPO phase (CESM1, CESM2-LENS, EC-Earth3) are also the most sensitive to modulation by the PARC teleconnection. These three LEs also have the smallest open water area for this time period, and hence simulate large SIC internal variability.

4.4 Discussion

It is only recently that a collection of LEs with sufficiently numerous ensemble members has been available to disentangle the independent and combined effects of the IPO and PARC teleconnection on Arctic sea ice trajectories. The 8 LEs considered here also cover a wide range of modeling center genealogies (Kuma et al., 2023), representing few model component similarities aside from the CMIP5/6 versions of CanESM and CESM. Within this broad data set, 75% of the LEs simulate the impact of the initial phase of the IPO in the same direction as, but generally less sensitively than, CESM1 as found by (Screen and Deser, 2019). EC-Earth3 is the major outlier of the 8 LEs analyzed, simulating a much delayed ice-free Arctic for trajectories with an initially negative phase of the IPO. While the reason for this difference is not known, we find that based on all available CMIP6 pre-industral control simulations, EC-Earth3 has the largest September sea ice area variability of all (see table S3 of Notz and SIMIP-Community (2020)). Further, EC-Earth3 is an outlier in extreme low-frequency variability in the future simulations, as represented by the largest range of ice-free years in the 11-year running mean (see Figure S4.1). Excluding EC-Earth3, all LEs analyzed simulate an earlier ice-free Arctic compared with the ensemble mean for trajectories starting with an initially negative phase of the IPO, albeit not all significantly earlier.

The strongly negative current observations of the PARC teleconnection are at the extreme end of that simulated by the large ensembles (see Figure S4.4). This raises the question of whether observations are atypical or whether models cannot represent the climate variability necessary for accurate tropical-Arctic teleconnections, as has been suggested for CESM1 (Blanchard-Wrigglesworth and Ding, 2019). As the effect of the PARC teleconnection in modulating the IPO signal is relatively small and inconsistent compared with the IPO signal alone (aside from EC-Earth3) the IPO alone appears the best indication of the timing of ice-free conditions. Nonetheless, when the PARC teleconnection is initially in a negative phase, regions of lower SIC occur at multi-decadal timescales in all LEs except for CanESM2 and CanESM5, similar to recent observations (Baxter et al., 2019). As the CanESM LEs do not show significant influence of the IPO on SIE trajectories at any time period, the PARC is less likely to be acting mechanistically to modulate the IPO signal. We found non-stationarity of the PARC teleconnection over the period 1953-2033 for CMIP5 and CMIP6 LEs (see Figure S4.4) which agrees with the findings of Bonan and Blanchard-Wrigglesworth (2020) for CMIP5 pre-industrial control simulations. However, despite the non-stationarity of the PARC, it is simulated to evolve sufficiently slowly with time such that the signal is evident both at the 17-22 year time period (see Figure S4.3) and in the subsequent lead-up to ice-free conditions (see Figure 4.4).

4.5 Conclusions

The transition of the Arctic Ocean to becoming seasonally consistently ice-free is highly dependent on internal variability, with a range of 11-25 years between LEs (see Figure S4.1). To reduce this large internal variability uncertainty, we here focus on the IPO, a decadal to interdecadal mode of variability that has been suggested to allow a narrowing of the prediction uncertainty (Screen and Deser, 2019), based on the CESM1. By repeating this analysis for 7 additional LEs, we find that it is only CESM1 that shows a statistically significant difference of SIE trajectories based on the initial IPO state at the time when ice-free conditions are reached.

Increasing the number of cases considered, by including different starting SIE and persistently negative and positive IPO phases, we find that different IPO phases show significant differences between SIE trajectories over the next 55 years in all LEs except for CanESM2 and CanESM5. Looking specifically at ice-free conditions, we show that a negative initial phase of the IPO brings about an earlier than average ice-free Arctic for 75% of the large ensembles. In the second and third decade of SIE trajectories, when the initial IPO has reversed sign from its initial state, the influence of the IPO on Arctic sea ice is typically the greatest. The IPO is therefore most influential in terms of affecting the timing of ice-free conditions for the LEs that have both high sensitivity to the initial phase of the IPO and simulate a rapid decline towards ice-free conditions, notably CESM1 and EC-Earth3. However, EC-Earth3 is the outlier model in simulating a vastly delayed ice-free Arctic for the case of an initially negative IPO.

The modulation of the IPO's effect on ice-free conditions by the PARC is relatively small and inconsistent across the LEs compared with the IPO signal alone. However, during the transition to an ice-free Arctic, the positive (negative) PARC is associated with positive (negative) spatial SIC anomalies, except for the CanESM LEs. More work is needed to determine which, if any, model family represents the Pacific-Arctic teleconnection most realistically.

Based on an initial state of a negative IPO, as is the case today, 75% of the LEs show earlier than average consistently ice-free conditions.

4.6 Open Research

The climate model data is available from https://aims2.llnl.gov/search/cmip5/ for CMIP5 https://aims2.llnl.gov/search/cmip6/ for CMIP6. The observational sea ice index dataset is available from NSIDC at https://doi.org/10.7265/N5K072F8. Sea surface temperature dataset HadISST is available from https://doi.org/10.5065/XMYE-AN84. The geopotential height data from the ERA-5 reanalysis can be found from ECMWF at https://doi.org/10.24381/cds.f17050d7. Data archiving is underway with the code used to compute the analyses to be archived with Zenodo, currently it is available from GitHub https://github.com/chrisrwp/ice-free. The datasets used to reproduce this material will be archived at the Arctic Data Center.

4.7 Acknowledgments

This work was supported by the National Science Foundation under grant number 1847398. We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. We also acknowledge the US CLIVAR Working Group on Large Ensembles and the modeling centers that contributed to the CLIVAR Large Ensemble project. We would also like to acknowledge high-performance computing on Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. We also acknowledge helpful feedback from J.E. Kay, M.M. Holland, D.A. Bailey and K.B. Karnauskas.

4.8 Supplementary Material

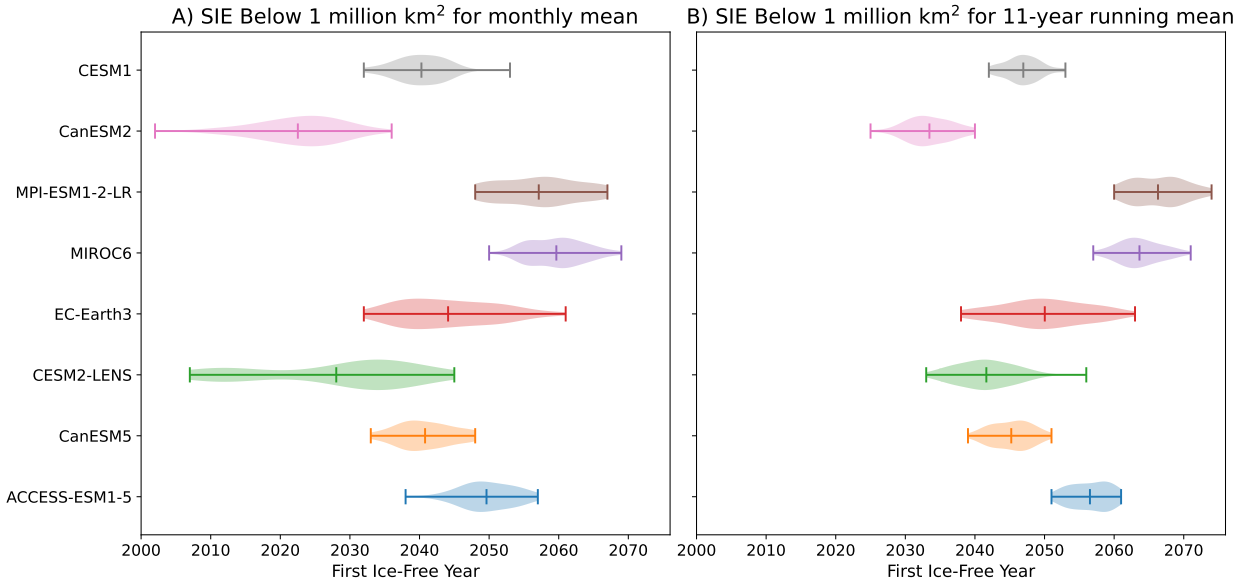

Figure S4.1: **Ice-free timing for September in the different LEs.** The range of first ice-free conditions simulated between different ensemble members from the large ensembles of interest based on two averaging periods.

Table S4.1: List of total number of ensembles used for each category of IPO and PARC teleconnection categories.

LE Name	All	IPO+	IPO-	PARC+	PARC-	IPO+ PARC+	IPO+ PARC-	IPO- PARC+	IPO- PARC-
ACCESS-ESM1-5	40	19	12	15	13	9	3	3	5
CanESM5	49	13	13	43	17	10	4	10	6
CESM2-LENS	50	10	16	26	31	3	7	4	12
EC-Earth3	50	13	17	37	28	8	8	10	8
MIROC6	50	14	18	47	17	13	5	17	1
MPI-ESM1-2-LR	30	7	7	25	10	7	0	4	3
CanESM2	50	21	13	42	14	17	4	10	4
CESM1	40	7	12	31	17	5	3	8	4

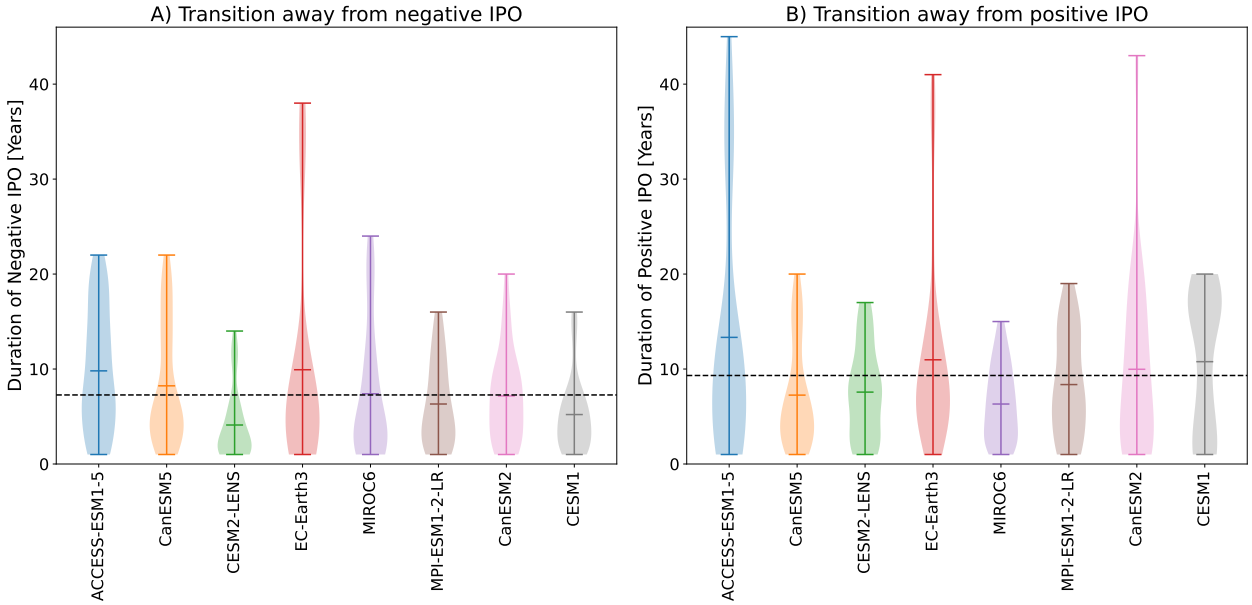

Figure S4.2: **Transition time from an initial IPO phase, to neutral or the opposite phase, following the initial IPO phase persistence of at least 10 years.** Data taken from each of the ensemble member trajectories starting for the 11 mean states of 4.2-5.2 million km^2.

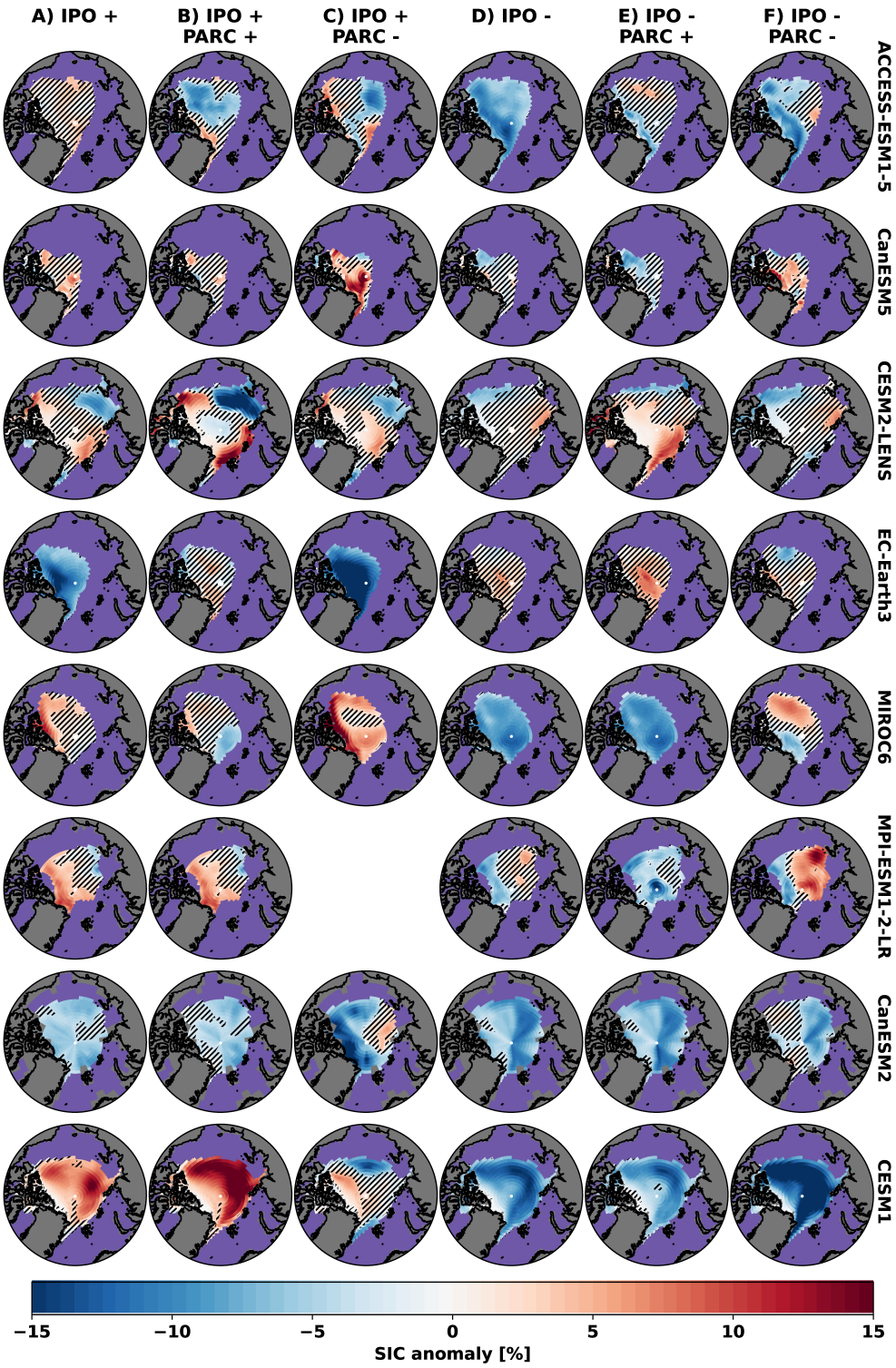

Figure S4.3: **Sea ice concentration anomalies in years 17 to 22 of the trajectories relative to the ensemble mean.** For initial IPO phases A), D) and additional PARC phase categorization for B), C), E), F). Significance shown by the black lines is relative to the ensemble mean.

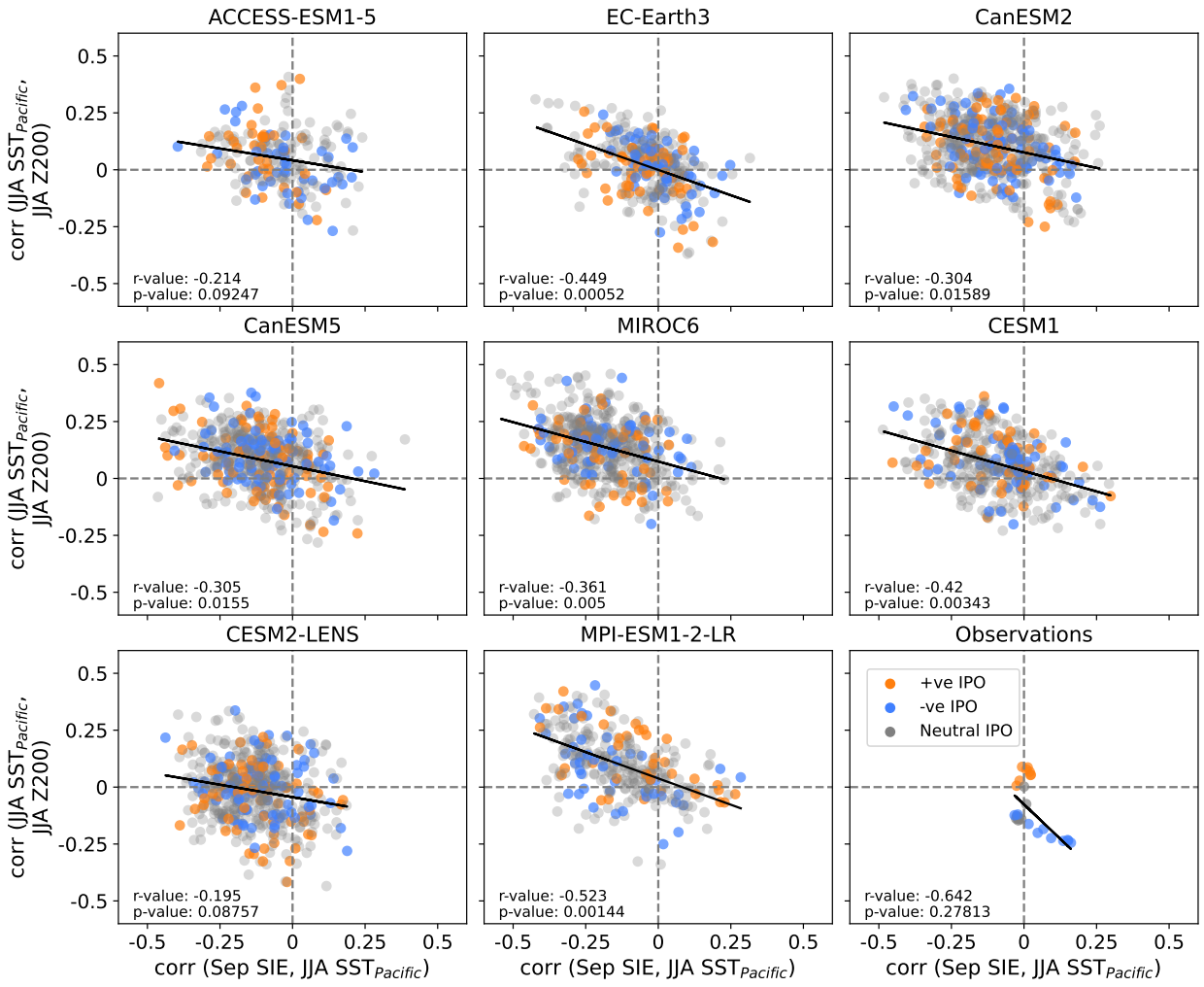

Figure S4.4: **PARC teleconnection and IPO, SIE correlations.** Subplots show the correlation between the IPO and the JJA Arctic geopotential height anomalies on the y-axis, and the correlations between the IPO and Arctic SIE on the x-axis. The correlations are calculated for each ensemble member for 40-year time periods at 5 year increment for each ensemble member. The black line indicates the linear least-squares fit for all correlations, with associated statistics printed on the subplots. The colors of the scatter plots relate to the phase of the IPO for the last 10 years of the 40-year period, to align with the IPO phase used for classifying trajectories.

CHAPTER 5

CONCLUSION

5.1 Summary of Major Findings

Climate model projections of Arctic sea ice are crucial tools for understanding how the Arctic responds to increasing global temperatures. However, evaluation of forced sea ice changes in climate models is complicated by the internal variability of the climate system, model differences and scenario uncertainty. In this dissertation I explored the realism of interannual variability of Arctic sea ice and diagnosed drivers of low-frequency variability. This allows us to use climate models to better understand the internal variability of Arctic sea ice, and to begin to reduce internal variability uncertainty of the future evolution of Arctic sea ice.

In Chapter 2, I assessed climate model realism in the simulation of interannual variability of Arctic sea ice. Variability which acts on timescales shorter than 2 years was found to account for approximately three quarters of total internal variability. This interannual variability was generally modeled within observational uncertainty for a collection of 6 large ensembles from the CMIP5 archive. The large ensembles generally underestimate interannual variability in the central Arctic in September and overestimate variability in the same region in March. No model was able to simulate realistic interannual variability across the Arctic for all seasons, hence it is important to select the most accurate model for a given task. Particularly challenging to accurately model are the coastal regions where the approximately 1-3° resolution of the GCMs is an impediment to capturing the local processes, limiting meaningful observational comparisons.

For many spatial and temporal domains, particularly in the central Arctic and during the winter months, observational uncertainty is so large that model consistency with one observational record is an easy achievement. Consequently, better agreement on the interpretation of satellite records is required to further diagnose and characterize model biases in interannual variability. However, the general realism of these models increases confidence of their ability to simulate future interannual variability, especially for the annual minimum in September when observations are most consistent.

In Chapter 3, I identified and characterized the dominant drivers of regional Arctic sea ice variability on multi-year to decadal timescales. I showed that predictive skill of up to 46% above persistence can be achieved using a simple linear model, using inputs solely of climate modes of variability. Leveraging 42 GCMs from the CMIP6 archive, four key climate drivers were found to dominate predictions. These drivers are the interdecadal Pacific oscillation (IPO), the Nino 3.4 Index, the Atlantic Mutidecadal Oscillation (AMO), and global surface temperatures. Although the IPO was found to be the most influential climate mode of variability, particularly in the East Siberian, Beaufort and Laptev Seas, there is large disagreement between the GCMs on the sign and magnitude of its influence. The Nino 3.4 index and AMO were more consistently identified by the collection of CMIP6 GCMs as being highly influential. Using our linear model trained and validated on CMIP6 data to then test on historical observational data, our predictions reach 0.3 r^2 above persistence for the pan-Arctic at a 10-year lead time. Using our linear model to make predictions, I found a near 'perfect storm' of the current alignment of climate modes of variability to induce an acceleration to the sea ice loss trend over the next decade. In fact, our model predicts a larger pan-Arctic anomaly in low-frequency variability than at any time since 1956.

In Chapter 4, I determined the importance of two specific climate modes of variability in driving multi-decadal variability of Arctic sea ice. I showed that the IPO and PARC are influential in a number of CMIP5 and CMIP6 climate models. The initially negative phase of the IPO was shown to lead to an earlier ice-free Arctic than average for 75% of the large

ensembles. The largest influence from the IPO occurs following the transition of the IPO from one phase to another, approximately two decades from the start of the trajectory. From this we can deduce that it is the transition of the IPO away from a negative phase which leads to earlier ice-free conditions than average in all large ensembles aside from EC-Earth3, which has the opposite behavior. I further found that the Pacific-Arctic (PARC) teleconnection can also be highly influential in affecting the timing of ice-free conditions. However, there is a large range of sensitivities to both the IPO and PARC teleconnection across the CMIP5 and CMIP6 large ensembles. Despite this, the strong influence of the IPO and the PARC teleconnection leads to ice-free conditions occurring earlier than the ensemble mean when both climate modes of variability are in their negative phases, which is the state of current observations.

Taken together, this work has allowed us to address the overarching scientific question of this dissertation: "How can we best understand internal variability of Arctic sea ice both in models and observations in order to reduce our projection uncertainty of future sea ice conditions?" This has been achieved by comparing models and observations, and examining large-scale climate interconnections within numerous historical and future model simulations. We have shown climate model large ensembles are able to realistically simulate historical interannual variability of Arctic sea ice. This increases our confidence of short-term variability of future simulations, while highlighting spatial and temporal model biases important for applications such as the frequency of shipping route navigability. We have also shown that using the initial state of named climate modes of variability can reduce internal variability uncertainty in decadal trends of Arctic sea ice. With the knowledge of the current phases of the identified important climate modes of variability, we project that over the next decade there will be accelerated loss of Arctic sea ice. Further, on multi-decadal timescales, we have shown that the IPO and PARC teleconnection in their current negative phases will likely lead to an earlier ice-free Arctic than predicted based on the ensemble mean.

5.2 Future Work

The spatial and temporal characteristics of interannual variability are evolving rapidly with the thinning and shrinking of sea ice extent. In Chapter 2 I found that CMIP5 large ensembles were generally able to simulate interannual variability within observational uncertainty. However, areas where large biases exist are in the central Arctic, which is increasingly dominating the summer Arctic sea ice cover. The MOSAiC expedition and recent high-resolution satellite measurements have advanced our understanding of melt ponds and other high latitude summer processes (e.g. Nicolaus et al., 2022; Dawson and Landy, 2023). This is enabling process-based model assessments (e.g. Webster et al., 2022), which will likely help inform the cause of large-scale concentration model biases. Budget analyses e.g. Keen et al. (2021) can inform such biases as well, but longer high quality thickness records must first be obtained in order to make meaningful observational comparisons. This leads to the research question of whether reducing observational uncertainty will facilitate better identification of model biases in interannual variability of Arctic sea ice. A concerted effort is being made to produce high quality climate data record (CDR) products across observational records, with the sea ice concentration CDR already proving to be highly value as the primary record for model assessments. Expansion of these CDRs to include thickness data, along with explicit inclusion of uncertainty quantification will help standardize future model evaluations.

Decadal variability in summer Arctic sea ice has many drivers which are likely to have nonlinear effects and have interactions with each other. In Chapter 3 I found that a simple liner model was sufficient to produce skillfull predictions of regional Arctic sea ice concentration anomalies. However, when using the largest number of ensemble members, the more complicated neural networks, which incorporated nonlinearies and covariance, were able to slightly exceed the linear model's skill. This leads to the research question of whether we can incorporate additional ensemble members in our machine learning model to capture these nonlinear effects of drivers of Arctic sea ice variability. We could address this by leveraging more sophisticated machine learning models to explicitly take

into account additional variables such as the initial state of the sea ice, and temporal patterns. Specifically, a long short-term memory (LSTM) machine learning architecture can explicitly learn the time decay component of each variable. Deploying a methodology like this would allow short-term variables to be included in the model, such as the initial concentration and thickness of the sea ice, without over-fitting. This may produce a more generalized model which could be trained on many different time periods, utilizing all historical and future simulations. The result could be a single equation with linear and nonlinear terms for each driver, which would also include time decay terms.

Multidecadal variability of summer Arctic sea ice is expected to greatly impact the timing of the Arctic becoming seasonally ice free. In Chapter 4, I found that an ice-free Arctic is more likely to occur earlier when trajectories start at present-day sea ice extents during a negative IPO and negative PARC teleconnection phase. However, as these climate modes of variability change so slowly, even with large ensembles there are few examples to from which to draw these conclusions. In the large ensembles analyzed I found a large difference in the sensitivity of the IPO and PARC teleconnections between the CESM and CanESM model families. Large differences across the two model families exist, in their Pacific SST patterns as well as their sea ice mean state and thickness. The number of differences makes it difficult to determine the reasons for the far weaker sea ice response to the IPO sign and opposite response to the PARC teleconnection in the two model families. Detection of the transmission mechanism between the tropics and Arctic has been successfully achieved using a nudging method in CESM1 (e.g. Ding et al., 2017; Blanchard-Wrigglesworth and Ding, 2019; Ding et al., 2022). A similar experimental setup for CanESM5 would help to determine what drives the differences between the two model families. Further, commonalities and differences between the observations and each model family would further inform what real-word sea ice sensitivity to the IPO and PARC we should expect.

BIBLIOGRAPHY

Alexander, M. A., I. Bladé, M. Newman, J. R. Lanzante, N.-C. Lau, and J. D. Scott, 2002: The Atmospheric Bridge: The Influence of ENSO Teleconnections on Air–Sea Interaction over the Global Oceans. *Journal of Climate*, **15**, 2205 – 2231, doi:10.1175/1520-0442(2002)015¡2205:TABTIO¿2.0.CO;2.

Bader, D. C., R. Leung, M. Taylor, and R. B. McCoy, 2019: E3SM-Project E3SM1.0 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.4497.

Barnes, E. A., J. W. Hurrell, I. Ebert-Uphoff, C. Anderson, and D. Anderson, 2019: Viewing Forced Climate Patterns Through an AI Lens. *Geophysical Research Letters*, **46**, 13389–13398, doi:10.1029/2019GL084944.

Barnhart, K. R., C. R. Miller, I. Overeem, and J. E. Kay, 2016: Mapping the future expansion of Arctic open water. *Nature Climate Change*, **6**, 280–285, doi:10.1038/nclimate2848.

Baxter, I., Q. Ding, A. Schweiger, M. L'Heureux, S. Baxter, T. Wang, Q. Zhang, K. Harnos, B. Markle, D. Topal, and J. Lu, 2019: How tropical Pacific surface cooling contributed to accelerated sea ice melt from 2007 to 2012 as ice is thinned by anthropogenic forcing. *Journal of Climate*, **32**, 8583–8602, doi:10.1175/JCLI-D-18-0783.1.

Bentsen, M., D. J. L. Oliviè, Øyvind Seland, T. Toniazzo, A. Gjermundsen, L. S. Graff, J. B. Debernard, A. K. Gupta, Y. He, A. Kirkevag, J. Schwinger, J. Tjiputra, K. S. Aas, I. Bethke, Y. Fan, J. Griesfeller, A. Grini, C. Guo, M. Ilicak, I. H. H. Karset, O. A. Landgren, J. Liakka, K. O. Moseid, A. Nummelin, C. Spensberger, H. Tang, Z. Zhang, C. Heinze, T. Iversen, and M. Schulz, 2019: NCC NorESM2-MM model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.8040.

Bethke, I., Y. Wang, F. Counillon, N. Keenlyside, M. Kimmritz, F. Fransner, A. Samuelsen, H. Langehaug, L. Svendsen, P. G. Chiu, L. Passos, M. Bentsen, C. Guo, A. Gupta, J. Tjiputra, A. Kirkeväg, D. Olivié, Øyvind Seland, J. S. Vägane, Y. Fan, and T. Eldevik, 2021: NorCPM1 and its contribution to CMIP6 DCPP. *Geoscientific Model Development*, **14**, 7073–7116, doi:10.5194/gmd-14-7073-2021.

Bethke, I., Y. Wang, F. Counillon, M. Kimmritz, F. Fransner, A. Samuelsen, H. R. Langehaug, P.-G. Chiu, M. Bentsen, C. Guo, J. Tjiputra, A. Kirkevag, D. J. L. Oliviè, Øyvind Seland, Y. Fan, P. Lawrence, T. Eldevik, and N. Keenlyside, 2019: NCC NorCPM1 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.10894.

Blanchard-Wrigglesworth, E., K. C. Armour, C. M. Bitz, and E. Deweaver, 2011: Persistence and inherent predictability of Arctic sea ice in a GCM ensemble and observations. *Journal of Climate*, **24**, 231–250, doi:10.1175/2010JCLI3775.1.

Blanchard-Wrigglesworth, E. and M. Bushuk, 2019: Robustness of Arctic sea-ice predictability in GCMs. *Climate Dynamics*, **52**, 5555–5566, doi:10.1007/s00382-018-4461-3.

Blanchard-Wrigglesworth, E. and Q. Ding, 2019: Tropical and Midlatitude Impact on Seasonal Polar Predictability in the Community Earth System Model. *Journal of Climate*, **32**, 5997–6014, doi:10.1175/JCLI-D-19-0088.1.

Boeke, R. C. and P. C. Taylor, 2018: Seasonal energy exchange in sea ice retreat regions contributes to differences in projected Arctic warming. *Nature Communications*, **9**, 1–14, doi:10.1038/s41467-018-07061-9.

Bonan, D. B. and E. Blanchard-Wrigglesworth, 2020: Nonstationary Teleconnection Between the Pacific Ocean and Arctic Sea Ice. *Geophysical Research Letters*, **47**, e2019GL085666, doi:10.1029/2019GL085666.

Bonan, D. B., M. Bushuk, and M. Winton, 2019: A Spring Barrier for Regional Predictions of Summer Arctic Sea Ice. *Geophysical Research Letters*, **46**, 5937–5947, doi:10.1029/2019GL082947.

Bonan, D. B., F. Lehner, and M. M. Holland, 2021: Partitioning uncertainty in projections of Arctic sea ice. *Environmental Research Letters*, **16**, doi:10.1088/1748-9326/abe0ec.

Boucher, O., S. Denvil, G. Levavasseur, A. Cozic, A. Caubel, M.-A. Foujols, Y. Meurdesoif, P. Cadule, M. Devilliers, J. Ghattas, N. Lebas, T. Lurton, L. Mellul, I. Musat, J. Mignot, and F. Cheruy, 2018: IPSL IPSL-CM6A-LR model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.5195.

Brennan, M. K. and G. J. Hakim, 2022: Reconstructing Arctic Sea Ice over the Common Era Using Data Assimilation. *Journal of Climate*, **35**, 1231–1247, doi:10.1175/JCLI-D-21-0099.1.

Brown, J. R., C. M. Brierley, S. I. An, M. V. Guarino, S. Stevenson, C. J. Williams, Q. Zhang, A. Zhao, A. Abe-Ouchi, P. Braconnot, E. C. Brady, D. Chandan, R. D'Agostino, C. Guo, A. N. Legrande, G. Lohmann, P. A. Morozova, R. Ohgaito, R. O'Ishi, B. L. Otto-Bliesner, W. R. Peltier, X. Shi, L. Sime, E. M. Volodin, Z. Zhang, and W. Zheng, 2020: Comparison of past and future simulations of ENSO in CMIP5/PMIP3 and CMIP6/PMIP4 models. *Climate of the Past*, **16**, 1777–1805, doi:10.5194/cp-16-1777-2020.

Bushuk, M. and D. Giannakis, 2017: The seasonality and interannual variability of Arctic Sea ice reemergence. *Journal of Climate*, **30**, 4657–4676, doi:10.1175/JCLI-D-16-0549.1.

Bushuk, M., R. Msadek, M. Winton, G. Vecchi, X. Yang, A. Rosati, and R. Gudgel, 2019: Regional Arctic sea–ice prediction: potential versus operational seasonal forecast skill. *Climate Dynamics*, **52**, 2721–2743, doi:10.1007/S00382-018-4288-Y.

Cai, Q., D. Beletsky, J. Wang, and R. Lei, 2021: Interannual and Decadal Variability of Arctic Summer Sea Ice Associated with Atmospheric Teleconnection Patterns during 1850-2017. *Journal of Climate*, **34**, 9931–9955, doi:10.1175/JCLI-D-20-0330.1.

Cao, J. and B. Wang, 2019: NUIST NESMv3 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.8769.

Capotondi, A., S. McGregor, M. J. McPhaden, S. Cravatte, N. J. Holbrook, Y. Imada, S. C. Sanchez, J. Sprintall, M. F. Stuecker, C. C. Ummenhofer, M. Zeller, R. Farneti, G. Graffino, S. Hu, K. B. Karnauskas, Y. Kosaka, F. Kucharski, M. Mayer, B. Qiu, A. Santoso, A. S. Taschetto, F. Wang, X. Zhang, R. M. Holmes, J. J. Luo, N. Maher, C. Martinez-Villalobos, G. A. Meehl, R. Naha, N. Schneider, S. Stevenson, A. Sullivan, P. van Rensch, and T. Xu, 2023: Mechanisms of tropical pacific decadal variability. *Nature Reviews Earth and Environment*, **4**, 754–769, doi:10.1038/s43017-023-00486-x.

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

Chan, D., A. Cobb, L. R. Zeppetello, D. S. Battisti, and P. Huybers, 2020: Summertime temperature variability increases with local warming in midlatitude regions. *Geophysical Research Letters*, **47**, 1–7, doi:10.1029/2020GL087624.

Chemke, R. and L. M. Polvani, 2020: Using Multiple Large Ensembles to Elucidate the Discrepancy Between the 1979–2019 Modeled and Observed Antarctic Sea Ice Trends. *Geophysical Research Letters*, **47**, doi:10.1029/2020GL088339.

Chen, J., S. Kang, C. Chen, Q. You, W. Du, M. Xu, X. Zhong, W. Zhang, and J. Chen, 2020: Changes in sea ice and future accessibility along the Arctic Northeast Passage. *Global and Planetary Change*, **195**, 103319, doi:10.1016/j.gloplacha.2020.103319.

Chen, X. and J. M. Wallace, 2016: Orthogonal PDO and ENSO Indices. *Journal of Climate*, **29**, 3883–3892, doi:10.1175/JCLI-D-15-0684.1.

Christensen, M. and A. E. Nilsson, 2017: Arctic sea ice and the communication of climate change. *Popular Communication*, **15**, 249–268, doi:10.1080/15405702.2017.1376064.

Clancy, R., C. Bitz, and E. Blanchard-Wrigglesworth, 2021: The Influence of ENSO on Arctic Sea Ice in Large Ensembles and Observations. *Journal of Climate*, **34**, 9585–9604, doi:10.1175/JCLI-D-20-0958.1.

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

Comiso, J. C., 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, doi:10.1002/2017JC012768.

Crawford, A., J. Stroeve, A. Smith, and A. Jahn, 2021: Arctic open-water periods are projected to lengthen dramatically by 2100. *Communications Earth and Environment*, **2**, doi:10.1038/s43247-021-00183-x.

Dai, A. and C. E. Bloecker, 2019: Impacts of internal variability on temperature and precipitation trends in large ensemble simulations by two climate models. *Climate Dynamics*, **52**, 307, doi:10.1007/s00382-018-4204-5.

Dalelane, C., K. Winderlich, and A. Walter, 2023: Evaluation of global teleconnections in CMIP6 climate projections using complex networks. *Earth System Dynamics*, **14**, 17–37, doi:10.5194/esd-14-17-2023.

Danabasoglu, G., 2019a: NCAR CESM2-FV2 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.11297.

— 2019b: NCAR CESM2 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.7627.

— 2019c: NCAR CESM2-WACCM-FV2 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.11298.

— 2019d: NCAR CESM2-WACCM model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.10071.

Danabasoglu, G., J. F. Lamarque, J. Bacmeister, D. A. Bailey, A. K. DuVivier, J. Edwards, L. K. Emmons, J. Fasullo, R. Garcia, A. Gettelman, C. Hannay, M. M. Holland, W. G. Large, P. H. Lauritzen, D. M. Lawrence, J. T. Lenaerts, K. Lindsay, W. H. Lipscomb, M. J. Mills, R. Neale, K. W. Oleson, B. Otto-Bliesner, A. S. Phillips, W. Sacks, S. Tilmes, L. van Kampenhout, M. Vertenstein, A. Bertini, J. Dennis, C. Deser, C. Fischer, B. Fox-Kemper, J. E. Kay, D. Kinnison, P. J. Kushner, V. E. Larson, M. C. Long, S. Mickelson, J. K. Moore, E. Nienhouse, L. Polvani, P. J. Rasch, and W. G. Strand, 2020: The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, **12**, e2019MS001916, doi:10.1029/2019MS001916.

Davy, R. and S. Outten, 2020: The Arctic surface climate in CMIP6: Status and developments since CMIP5. *Journal of Climate*, **33**, 8047–8068, doi:10.1175/JCLI-D-19-0990.1.

Dawson, G. J. and J. C. Landy, 2023: Comparing elevation and backscatter retrievals from CryoSat-2 and ICESat-2 over Arctic summer sea ice. *Cryosphere*, **17**, 4165–4178, doi:10.5194/tc-17-4165-2023.

Day, J. J., J. C. Hargreaves, J. D. Annan, and A. Abe-Ouchi, 2012: Sources of multi-decadal variability in Arctic sea ice extent. *Environmental Research Letters*, **7**, doi:10.1088/1748-9326/7/3/034011.

Day, J. J., E. Hawkins, and S. Tietsche, 2014: Will Arctic sea ice thickness initialization improve seasonal forecast skill? *Geophysical Research Letters*, **41**, 7566–7575, doi:10.1002/2014GL061694.

Derepentigny, P., A. Jahn, M. M. Holland, J. E. Kay, J. Fasullo, J.-F. Lamarque, S. Tilmes, C. Hannay, M. J. Mills, D. A. Bailey, and A. P. Barrett, 2022: Enhanced simulated early 21st century Arctic sea ice loss due to CMIP6 biomass burning emissions. *Sci. Adv*, **8**, 2405, doi:10.1126/sciadv.abo2405.

Deser, C., F. Lehner, K. B. Rodgers, T. Ault, T. L. Delworth, P. N. DiNezio, A. Fiore, C. Frankignoul, J. C. Fyfe, D. E. Horton, J. E. Kay, R. Knutti, N. S. Lovenduski, J. Marotzke, K. A. McKinnon, S. Minobe, J. Randerson, J. A. Screen, I. R. Simpson, and M. Ting, 2020: Insights from Earth system model initial-condition large ensembles and future prospects. *Nature Climate Change*, **10**, 277–286, doi:10.1038/s41558-020-0731-2.

Diffenbaugh, N. S. and E. A. Barnes, 2023: Data-driven predictions of the time remaining until critical global warming thresholds are reached. *Proceedings of the National Academy of Sciences of the United States of America*, **120**, doi:10.1073/pnas.2207183120.

Ding, Q., A. Schweiger, and I. Baxter, 2022: Nudging Observed Winds in the Arctic to Quantify Associated Sea Ice Loss from 1979 to 2020. *Journal of Climate*, **35**, 6797–6813, doi:10.1175/JCLI-D-21-0893.1.

Ding, Q., A. Schweiger, M. L'Heureux, D. S. Battisti, S. Po-Chedley, N. C. Johnson, E. Blanchard-Wrigglesworth, K. Harnos, Q. Zhang, R. Eastman, and E. J. Steig, 2017: Influence of high-latitude atmospheric circulation changes on summertime Arctic sea ice. *Nature Climate Change*, **7**, 289–295, doi:10.1038/nclimate3241.

Ding, Q., A. Schweiger, M. L'Heureux, E. J. Steig, D. S. Battisti, N. C. Johnson, E. Blanchard-Wrigglesworth, S. Po-Chedley, Q. Zhang, K. Harnos, M. Bushuk, B. Markle, and I. Baxter, 2019: Fingerprints of internal drivers of Arctic sea ice loss in observations and model simulations. *Nature Geoscience*, **12**, 28–33, doi:10.1038/s41561-018-0256-8.

Ding, Q., J. M. Wallace, D. S. Battisti, E. J. Steig, A. J. Gallant, H. J. Kim, and L. Geng, 2014: Tropical forcing of the recent rapid Arctic warming in northeastern Canada and Greenland. *Nature*, **509**, 209–212, doi:10.1038/nature13260.

Dix, M., D. Bi, P. Dobrohotoff, R. Fiedler, I. Harman, R. Law, C. Mackallah, S. Marsland, S. O'Farrell, H. Rashid, J. Srbinovsky, A. Sullivan, C. Trenham, P. Vohralik, I. Watterson, G. Williams, M. Woodhouse, R. Bodman, F. B. Dias, C. M. Domingues, N. Hannah, A. Heerdegen, A. Savita, S. Wales, C. Allen, K. Druken, B. Evans, C. Richards, S. M. Ridzwan, D. Roberts, J. Smillie, K. Snow, M. Ward, and R. Yang, 2019: CSIRO-ARCCSS ACCESS-CM2 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.4271.

Docquier, D., T. Koenigk, R. Fuentes-Franco, M. P. Karami, and Y. Ruprich-Robert, 2021: Impact of ocean heat transport on the Arctic sea-ice decline: a model study with EC-Earth3. *Climate Dynamics*, **56**, 1407–1432, doi:10.1007/s00382-020-05540-8.

Döscher, R., K. Wyser, H. E. Meier, M. Qian, and R. Redler, 2010: Quantifying Arctic contributions to climate predictability in a regional coupled ocean-ice-atmosphere model. *Climate Dynamics*, **34**, 1157–1176, doi:10.1007/s00382-009-0567-y.

DuVivier, A. K., M. M. Holland, J. E. Kay, S. Tilmes, A. Gettelman, and D. A. Bailey, 2020: Arctic and Antarctic Sea Ice Mean State in the Community Earth System Model Version 2 and the Influence of Atmospheric Chemistry. *Journal of Geophysical Research: Oceans*, **125**, doi:10.1029/2019JC015934.

Dörr, J., D. B. Bonan, M. Årthun, L. Svendsen, and R. C. J. Wills, 2023: Forced and internal components of observed Arctic sea-ice changes. *The Cryosphere*, **17**, 4133–4153, doi:10.5194/tc-17-4133-2023.

Dörr, J., M. Årthun, T. Eldevik, and E. Madonna, 2021: Mechanisms of regional winter sea-ice variability in a warming Arctic. *Journal of Climate*, **34**, 8635–8653, doi:10.1175/JCLI-D-21-0149.1.

EC-Earth-Consortium, 2019a: EC-Earth-Consortium EC-Earth3 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.4700.

— 2019b: EC-Earth-Consortium EC-Earth3-Veg model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.4706.

— 2020: EC-Earth-Consortium EC-Earth3-Veg-LR model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.4707.

— 2021: EC-Earth-Consortium EC-Earth-3-CC model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.4702.

Eguíluz, V. M., J. Fernández-Gracia, X. Irigoien, and C. M. Duarte, 2016: A quantitative assessment of Arctic shipping in 2010–2014. *Scientific Reports 2016 6:1*, **6**, 1–6, doi:10.1038/srep30682.

Eisenman, I., 2010: Geographic muting of changes in the Arctic sea ice cover. *Geophysical Research Letters*, **37**, doi:10.1029/2010GL043741.

Elsworth, G. W., N. S. Lovenduski, and K. A. McKinnon, 2021: Alternate history: A synthetic ensemble of ocean chlorophyll concentrations. *Global Biogeochemical Cycles*, **35**, 1–13, doi:10.1029/2020GB006924.

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

England, M. R., 2021: Are Multi-Decadal Fluctuations in Arctic and Antarctic Surface Temperatures a Forced Response to Anthropogenic Emissions or Part of Internal Climate Variability? *Geophysical Research Letters*, **48**, e2020GL090631, doi:10.1029/2020GL090631.

England, M. R., I. Eisenman, N. J. Lutsko, and T. J. Wagner, 2021: The Recent Emergence of Arctic Amplification. *Geophysical Research Letters*, **48**, 1–10, doi:10.1029/2021GL094086.

Fasullo, J. T., A. S. Phillips, and C. Deser, 2020: Evaluation of Leading Modes of Climate Variability in the CMIP Archives. *Journal of Climate*, **33**, 5527–5545, doi:10.1175/jcli-d-19-1024.1.

Fetterer, F., K. Knowles, W. N. Meier, M. Savoie, and A. K. Windnagel., 2017: Sea ice index, version 3.0. *National Snow and Ice Data Center (NSIDC)*, doi:10.7265/N5K072F8.

Fetterer, F., M. Savoie, S. Helfrich, and P. Clemente-Colón, 2010: Multisensor Analyzed Sea Ice Extent - Northern Hemisphere (MASIE-NH), Version 1. *U.S. National Ice Center and National Snow and Ice Data Center*, doi:10.7265/N5GT5K3K.

Francis, J. A. and B. Wu, 2020: Why has no new record-minimum Arctic sea-ice extent occurred since September 2012? *Environmental Research Letters*, **15**, doi:10.1088/1748-9326/abc047.

Frankcombe, L. M., M. H. England, J. B. Kajtar, M. E. Mann, and B. A. Steinman, 2018: On the choice of ensemble mean for estimating the forced signal in the presence of internal variability. *Journal of Climate*, **31**, 5681–5693, doi:10.1175/JCLI-D-17-0662.1.

Furtado, J. C., E. D. Lorenzo, B. T. Anderson, and N. Schneider, 2012: Linkages between the North Pacific Oscillation and central tropical Pacific SSTs at low frequencies. *Climate Dynamics*, **39**, 2833–2846, doi:10.1007/s00382-011-1245-4.

Giesse, C., D. Notz, and J. Baehr, 2021: On the Origin of Discrepancies Between Observed and Simulated Memory of Arctic Sea Ice. *Geophysical Research Letters*, **48**, doi:10.1029/2020GL091784.

Goosse, H., O. Arzel, C. M. Bitz, A. D. Montety, and M. Vancoppenolle, 2009: Increased variability of the Arctic summer ice extent in a warmer climate. *Geophysical Research Letters*, **36**, 1–5, doi:10.1029/2009GL040546.

Gregory, W., J. Stroeve, and M. Tsamados, 2021: Network connectivity between the winter Arctic Oscillation and summer sea ice in CMIP6 models and observations. *The Cryosphere*, 1653–1673.

Guemas, V., E. Blanchard-Wrigglesworth, M. Chevallier, J. J. Day, M. Déqué, F. J. Doblas-Reyes, N. S. Fučkar, A. Germe, E. Hawkins, S. Keeley, T. Koenigk, D. S. y Mélia, and S. Tietsche, 2016: A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales. *Quarterly Journal of the Royal Meteorological Society*, **142**, 546–561, doi:10.1002/qj.2401.

Hajima, T., M. Abe, O. Arakawa, T. Suzuki, Y. Komuro, T. Ogura, K. Ogochi, M. Watanabe, A. Yamamoto, H. Tatebe, M. A. Noguchi, R. Ohgaito, A. Ito, D. Yamazaki, A. Ito, K. Takata, S. Watanabe, M. Kawamiya, and K. Tachiiri, 2019: MIROC MIROC-ES2L model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.5602.

Hawkins, E. and R. Sutton, 2009: The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, **90**, 1095–1107, doi:10.1175/2009BAMS2607.1.

Henley, B. J., J. Gergis, D. J. Karoly, S. Power, J. Kennedy, and C. K. Folland, 2015: A Tripole Index for the Interdecadal Pacific Oscillation. *Climate Dynamics*, **45**, 3077–3090, doi:10.1007/s00382-015-2525-1.

Henley, B. J., G. Meehl, S. B. Power, C. K. Folland, A. D. King, J. N. Brown, D. J. Karoly, F. Delage, A. J. Gallant, M. Freund, and R. Neukom, 2017: Spatial and temporal agreement in climate model simulations of the Interdecadal Pacific Oscillation. *Environmental Research Letters*, **12**, doi:10.1088/1748-9326/aa5cc8.

Heo, E. S., M. K. Sung, S. I. An, and Y. M. Yang, 2021: Decadal phase shift of summertime Arctic dipole pattern and its nonlinear effect on sea ice extent. *International Journal of Climatology*, **41**, 4732–4742, doi:10.1002/joc.7097.

Hersbach, H., P. Bell B. Berrisford, G. Biavati, A. Horányi, J. Muñoz Sabater, J. Nicolas, C. Peubey, R. Radu, I. Rozum, D. Schepers, A. Simmons, C. Soci, D. Dee, and J.-N. Thépaut, 2023: Era5 monthly averaged data on single levels from 1940 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*, doi:10.24381/cds.adbb2d47.

Hoerling, M. P., A. Kumar, and M. Zhong, 1997: El Niño, La Niña, and the Nonlinearity of Their Teleconnections. *Journal of Climate*, **10**, 1769–1786, doi:https://doi.org/10.1175/1520-0442(1997)010¡1769:ENOLNA¿2.0.CO;2.

Hofsteenge, M. G., R. G. Graversen, J. H. Rydsaa, and Z. Rey, 2022: The impact of atmospheric Rossby waves and cyclones on the Arctic sea ice variability. *Climate Dynamics*, **59**, 579–594, doi:10.1007/s00382-022-06145-z.

Holland, B. M. M. and E. C. Hunke, 2022: A Review of Arctic Sea Ice Climate Predictability in Large-Scale Earth System Models. *Oceanography*, **35**, 20–27, doi:10.5670/oceanog.2022.113.

Holland, M. M., L. Landrum, D. Bailey, and S. Vavrus, 2019: Changing seasonal predictability of Arctic summer sea ice area in a warming climate. *Journal of Climate*, **32**, 4963–4979, doi:10.1175/jcli-d-19-0034.1.

Hou, M. and Y. Tang, 2022: Recent progress in simulating two types of ENSO – from CMIP5 to CMIP6. *Frontiers in Marine Science*, **9**, doi:10.3389/fmars.2022.986780.

Hu, C., S. Yang, Q. Wu, Z. Li, J. Chen, K. Deng, T. Zhang, and C. Zhang, 2016: Shifting El Niño inhibits summer Arctic warming and Arctic sea-ice melting over the Canada Basin. *Nature Communications*, **7**, 1–9, doi:10.1038/ncomms11721.

Hu, K., G. Huang, and S. P. Xie, 2019: Assessing the internal variability in multi-decadal trends of summer surface air temperature over East Asia with a large ensemble of GCM simulations. *Climate Dynamics*, **52**, 6229–6242, doi:10.1007/s00382-018-4503-x.

Huang, W., 2019: THU CIESM model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.8843.

Hurrell, J. W. and C. Deser, 2009: North Atlantic climate variability: The role of the North Atlantic Oscillation. *Journal of Marine Systems*, **78**, 28–41, doi:10.1016/j.jmarsys.2008.11.026. http://dx.doi.org/10.1016/j.jmarsys.2008.11.026

Jahfer, S., K.-J. Ha, E.-S. Chung, C. L. E. Franzke, and S. Sharma, 2024: Unveiling the role of tropical Pacific on the emergence of ice-free Arctic projections. *Environmental Research Letters*, **19**, 044033, doi:10.1088/1748-9326/ad3141.

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

Jahn, A., M. Holland, and J. Kay, 2024: Projections of an ice-free Arctic Ocean. *Nature Reviews Earth and Environment*, **5**, 164–176, doi:10.1038/s43017-023-00515-9.

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, doi:10.1002/2016GL070067.

Jeffrey, S., L. Rotstayn, M. Collier, S. Dravitzki, C. Hamalainen, C. Moeseneder, K. Wong, and J. Syktus, 2013: Australia's CMIP5 submission using the CSIRO-Mk3.6 model. *Australian Meteorological and Oceanographic Journal*, **63**, 1–13, doi:10.22499/2.6301.001.

Jeong, H., H.-S. Park, M. F. Stuecker, and S.-W. Yeh, 2022a: Distinct impacts of major El Niño events on Arctic temperatures due to differences in eastern tropical Pacific sea surface temperatures. *Sci. Adv*, **8**, 8278, doi:10.1126/sciadv.abl8278.

Jeong, H., H. S. Park, M. F. Stuecker, and S. W. Yeh, 2022b: Record Low Arctic Sea Ice Extent in 2012 Linked to Two-Year La Niña-Driven Sea Surface Temperature Pattern. *Geophysical Research Letters*, **49**, doi:10.1029/2022GL098385.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, R. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, R. Jenne, and D. Joseph, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, **77**, 437–471, doi:10.1175/1520-0477(1996)077¡0437:TNYRP¿2.0.CO;2.

Karami, M. P., T. Koenigk, and B. Tremblay, 2023: Variability modes of September Arctic sea ice: drivers and their contributions to sea ice trend and extremes. *Environmental Research: Climate*, **2**, 025005, doi:10.1088/2752-5295/accbe3.

Kashiwase, H., K. I. Ohshima, S. Nihashi, and H. Eicken, 2017: Evidence for ice-ocean albedo feedback in the Arctic Ocean shifting to a seasonal ice zone. *Scientific Reports 2017 7:1*, **7**, 1–10, doi:10.1038/s41598-017-08467-z.

Kay, J. E., P. DeRepentigny, M. M. Holland, D. A. Bailey, A. K. DuVivier, E. Blanchard-Wrigglesworth, C. Deser, A. Jahn, H. Singh, M. M. Smith, M. A. Webster, J. Edwards, S. S. Lee, K. B. Rodgers, and N. Rosenbloom, 2022: Less Surface Sea Ice Melt in the CESM2 Improves Arctic Sea Ice Simulation With Minimal Non-Polar Climate Impacts. *Journal of Advances in Modeling Earth Systems*, **14**, doi:10.1029/2021MS002679.

Kay, J. E., C. Deser, A. Phillips, A. Mai, C. Hannay, G. Strand, J. M. Arblaster, S. C. Bates, G. Danabasoglu, J. Edwards, M. Holland, P. Kushner, J. F. Lamarque, D. Lawrence, K. Lindsay, A. Middleton, E. Munoz, R. Neale, K. Oleson, L. Polvani, and M. Vertenstein, 2015: The community earth system model (CESM) large ensemble project : A community resource for studying climate change in the presence of internal climate variability. *Bulletin of the American Meteorological Society*, **96**, 1333–1349, doi:10.1175/BAMS-D-13-00255.1.

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**, 2–7, doi:10.1029/2011GL048008.

Keen, A., E. Blockley, D. A. Bailey, J. Boldingh Debernard, M. Bushuk, S. Delhaye, D. Docquier, D. Feltham, F. Massonnet, S. O'Farrell, L. Ponsoni, J. M. Rodriguez, D. Schroeder, N. Swart, T. Toyoda, H. Tsujino, M. Vancoppenolle, and K. Wyser, 2021: An inter-comparison of the mass budget of the Arctic sea ice in CMIP6 models. *Cryosphere*, **15**, 951–982, doi:10.5194/tc-15-951-2021. https://doi.org/10.5194/tc-15-951-2021

Kern, S., T. Lavergne, D. Notz, L. T. Pedersen, R. T. Tonboe, R. Saldo, and A. M. Sørensen, 2019: Satellite passive microwave sea-ice concentration data set intercomparison: Closed ice and ship-based observations. *Cryosphere*, **13**, 3261–3307, doi:10.5194/TC-13-3261-2019.

Kim, H., S. W. Yeh, S. I. An, and S. Y. Song, 2020: Changes in the role of Pacific decadal oscillation on sea ice extent variability across the mid-1990s. *Scientific Reports*, **10**, doi:10.1038/s41598-020-74260-0.

Kirchmeier-Young, M. C., F. W. Zwiers, and N. P. Gillett, 2017: Attribution of extreme events in Arctic Sea ice extent. *Journal of Climate*, **30**, 553–571, doi:10.1175/JCLI-D-16-0412.1.

Klavans, J. M., M. A. Cane, A. C. Clement, and L. N. Murphy, 2021: NAO predictability from external forcing in the late 20th century. *npj Climate and Atmospheric Science*, **4**, 1–8, doi:10.1038/s41612-021-00177-8.

Klavans, J. M., A. C. Clement, M. A. Cane, and L. N. Murphy, 2022: The Evolving Role of External Forcing in North Atlantic SST Variability over the Last Millennium. *Journal of Climate*, **35**, 2741–2754, doi:10.1175/JCLI-D-21-0338.1.

Knutti, R., D. Masson, and A. Gettelman, 2013: Climate model genealogy: Generation CMIP5 and how we got there. *Geophysical Research Letters*, **40**, 1194–1199, doi:10.1002/grl.50256.

Kovacs, K. M., C. Lydersen, J. E. Overland, and S. E. Moore, 2011: Impacts of changing sea-ice conditions on Arctic marine mammals. *Marine Biodiversity*, **41**, 181–194, doi:10.1007/s12526-010-0061-0.

Krasting, J. P., J. G. John, C. Blanton, C. McHugh, S. Nikonov, A. Radhakrishnan, K. Rand, N. T. Zadeh, V. Balaji, J. Durachta, C. Dupuis, R. Menzel, T. Robinson, S. Underwood, H. Vahlenkamp, K. A. Dunne, P. P. G. Gauthier, P. Ginoux, S. M. Griffies, R. Hallberg, M. Harrison, W. Hurlin, S. Malyshev, V. Naik, F. Paulot, D. J. Paynter, J. Ploshay, B. G. Reichl, D. M. Schwarzkopf, C. J. Seman, L. Silvers, B. Wyman, Y. Zeng, A. Adcroft, J. P. Dunne, R. Dussin, H. Guo, J. He, I. M. Held, L. W. Horowitz, P. Lin, P. C. D. Milly, E. Shevliakova, C. Stock, M. Winton, A. T. Wittenberg, Y. Xie, and M. Zhao, 2018: NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.8597.

Kuma, P., F. A. Bender, and A. R. Jönsson, 2023: Climate model code genealogy and its relation to climate feedbacks and sensitivity. *Journal of Advances in Modeling Earth Systems*, **15**, doi:10.1029/2022MS003588.

Kwok, R. and D. A. Rothrock, 2009: Decline in Arctic sea ice thickness from submarine and ICESat records: 1958-2008. *Geophysical Research Letters*, **36**, 1–5, doi:10.1029/2009GL039035.

Labe, Z. M. and E. A. Barnes, 2022: Comparison of Climate Model Large Ensembles With Observations in the Arctic Using Simple Neural Networks. *Earth and Space Science*, **9**, e2022EA002348, doi:10.1029/2022EA002348.

Landrum, L. and M. M. Holland, 2020: Extremes become routine in an emerging new Arctic. *Nature Climate Change*, **10**, 1108–1115, doi:10.1038/s41558-020-0892-z.

Leathers, D. J., B. Yarnal, and M. A. Palecki, 1991: The Pacific/North American Teleconnection Pattern and United States Climate. Part I: Regional Temperature and Precipitation Associations. *Journal of Climate*, **4**, 517 – 528, doi:10.1175/1520-0442(1991)004¡0517:TPATPA¿2.0.CO;2.

Lee, J., K. R. Sperber, P. J. Gleckler, K. E. Taylor, , Céline, and J. W. Bonfils, 2021: Benchmarking Performance Changes in the Simulation of Extratropical Modes of Variability across CMIP Generations. *Journal of Climate*, **34**, doi:10.1175/JCLI-D-20.

Lee, S., H. S. Park, S. Y. Song, and S. W. Yeh, 2023: Distinct impacts of two types of El Niño events on northern winter high-latitude temperatures simulated by CMIP6 climate models. *Environmental Research Letters*, **18**, doi:10.1088/1748-9326/acbce9.

Lehner, F., C. Deser, N. Maher, J. Marotzke, E. M. Fischer, L. Brunner, R. Knutti, and E. Hawkins, 2020: Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. *Earth Syst. Dynam*, **11**, 491–508, doi:10.5194/esd-11-491-2020.

Lenssen, N. J. L., G. A. Schmidt, J. E. Hansen, M. J. Menne, A. Persin, R. Ruedy, and D. Zyss, 2019: Improvements in the GISTEMP Uncertainty Model. *Journal of Geophysical Research: Atmospheres*, **124**, 6307–6326, doi:10.1029/2018JD029522.

L'Heureux, M. L., A. Kumar, G. D. Bell, M. S. Halpert, and R. W. Higgins, 2008: Role of the Pacific-North American (PNA) pattern in the 2007 Arctic sea ice decline. *Geophysical Research Letters*, **35**, doi:10.1029/2008GL035205.

Li, C., D. Notz, S. Tietsche, and J. Marotzke, 2013: The transient versus the equilibrium response of sea ice to global warming. *Journal of Climate*, **26**, 5624–5636, doi:10.1175/JCLI-D-12-00492.1.

Li, D., R. Zhang, and T. Knutson, 2018a: Comparison of Mechanisms for Low-Frequency Variability of Summer Arctic Sea Ice in Three Coupled Models. *Journal of Climate*, **31**, 1205–1226, doi:10.1175/JCLI-D-16-0617.1.

Li, D., R. Zhang, and T. R. Knutson, 2017: On the discrepancy between observed and CMIP5 multi-model simulated Barents Sea winter sea ice decline. *Nature Communications*, **8**, doi:10.1038/ncomms14991.

Li, F., Y. J. Orsolini, H. Wang, Y. Gao, and S. He, 2018b: Atlantic Multidecadal Oscillation Modulates the Impacts of Arctic Sea Ice Decline. *Geophysical Research Letters*, **45**, 2497–2506, doi:10.1002/2017GL076210.

Lim, S. M., C. M. Payne, G. L. van Dijken, and K. R. Arrigo, 2022: Increases in Arctic sea ice algal habitat, 1985-2018. *Elementa*, **10**, doi:10.1525/elementa.2022.00008.

Lindsay, R. W. and J. Zhang, 2006: Arctic Ocean ice thickness: Modes of variability and the best locations from which to monitor them. *Journal of Physical Oceanography*, **36**, 496–506, doi:10.1175/JPO2861.1.

Liu, J., J. A. Curry, and Y. Hu, 2004: Recent Arctic sea ice variability: Connections to the Arctic Oscillation and the ENSO. *Geophysical Research Letters*, **31**, doi:10.1029/2004GL019858.

Liu, Z. and M. Alexander, 2007: Atmospheric bridge, oceanic tunnel, and global climatic teleconnections. *Reviews of Geophysics*, **45**, doi:10.1029/2005RG000172.

Liu, Z., C. Risi, F. Codron, X. He, C. J. Poulsen, Z. Wei, D. Chen, S. Li, and G. J. Bowen, 2021: Acceleration of western Arctic sea ice loss linked to the Pacific North American pattern. *Nature Communications*, **12**, doi:10.1038/s41467-021-21830-z.

Long, M., L. Zhang, S. Hu, and S. Qian, 2021: Multi-aspect assessment of CMIP6 models for Arctic sea ice simulation. *Journal of Climate*, **34**, 1515–1529, doi:10.1175/JCLI-D-20-0522.1.

Lovato, T. and D. Peano, 2020: CMCC CMCC-CM2-SR5 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.3825.

MacDonald, G. M. and R. A. Case, 2005: Variations in the Pacific Decadal Oscillation over the past millennium. *Geophysical Research Letters*, **32**, 1–4, doi:10.1029/2005GL022478.

Maher, N., F. Lehner, and J. Marotzke, 2020: Quantifying the role of internal variability in the temperature we expect to observe in the coming decades. *Environmental Research Letters*, **15**, 54014, doi:10.1088/1748-9326/ab7d02.

Maher, N., S. Milinski, L. Suarez-Gutierrez, M. Botzet, M. Dobrynin, L. Kornblueh, J. Kröger, Y. Takano, R. Ghosh, C. Hedemann, C. Li, H. Li, E. Manzini, D. Notz, D. Putrasahan, L. Boysen, M. Claussen, T. Ilyina, D. Olonscheck, T. Raddatz, B. Stevens, and J. Marotzke, 2019: The Max Planck Institute Grand Ensemble: Enabling the Exploration of Climate System Variability. *Journal of Advances in Modeling Earth Systems*, **11**, 2050–2069, doi:10.1029/2019MS001639.

Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis, 1997: A Pacific Interdecadal Climate Oscillation with Impacts on Salmon Production. *Bulletin of the American Meteorological Society*, **78**, 1069–1080, doi:10.1175/1520-0477(1997)078¡1069:APICOW¿2.0.CO;2.

Massonnet, F., M. Vancoppenolle, H. Goosse, D. Docquier, T. Fichefet, and E. Blanchard-Wrigglesworth, 2018: Arctic sea-ice change tied to its mean state through thermodynamic processes. *Nature Climate Change*, **8**, 599–603, doi:10.1038/s41558-018-0204-z.

Mcbride, L. A., A. P. Hope, T. P. Canty, B. F. Bennett, W. R. Tribett, and R. J. Salawitch, 2021: Comparison of CMIP6 historical climate simulations and future projected warming to an empirical model of global climate. *Earth System Dynamics*, **12**, 545–579, doi:10.5194/esd-12-545-2021.

McKinnon, K. A. and C. Deser, 2018: Internal variability and regional climate trends in an observational large ensemble. *Journal of Climate*, **31**, 6783–6802, doi:10.1175/JCLI-D-17-0901.1.

— 2021: The Inherent Uncertainty of Precipitation Variability, Trends, and Extremes due to Internal Variability, with Implications for Western U.S. Water Resources. *Journal of Climate*, **34**, 9605–9622, doi:10.1175/JCLI-D-21-0251.1.

McKinnon, K. A., A. Poppick, E. Dunn-Sigouin, and C. Deser, 2017: An "observational large ensemble" to compare observed and modeled temperature trend uncertainty due to internal variability. *Journal of Climate*, **30**, 7585–7598, doi:10.1175/JCLI-D-16-0905.1.

Meehl, G. A., C. T. Chung, J. M. Arblaster, M. M. Holland, and C. M. Bitz, 2018: Tropical Decadal Variability and the Rate of Arctic Sea Ice Decrease. *Geophysical Research Letters*, **45**, 11,326–11,333, doi:10.1029/2018GL079989.

Meehl, G. A., A. Hu, J. M. Arblaster, J. Fasullo, and K. E. Trenberth, 2013: Externally forced and internally generated decadal climate variability associated with the Interdecadal Pacific Oscillation. *Journal of Climate*, **26**, 7298–7310, doi:10.1175/JCLI-D-12-00548.1.

Meier, W. N., F. Fetterer, A. Windnagel, and J. Stewart, 2021: NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 4. doi:10.7265/efmz-2t65.

Meier, W. N. and J. S. Stewart, 2019: Assessing uncertainties in sea ice extent climate indicators. *Environmental Research Letters*, **14**, doi:10.1088/1748-9326/AAF52C.

Melia, N., K. Haines, E. Hawkins, and J. J. Day, 2017: Towards seasonal Arctic shipping route predictions. *Environmental Research Letters*, **12**, 084005, doi:10.1088/1748-9326/AA7A60.

Miles, M. W., D. V. Divine, T. Furevik, E. Jansen, M. Moros, and A. E. Ogilvie, 2014: A signal of persistent Atlantic multidecadal variability in Arctic sea ice. *Geophysical Research Letters*, **41**, 463–469, doi:10.1002/2013GL058084.

Milinski, S., N. Maher, and D. Olonscheck, 2020: How large does a large ensemble need to be? *Earth Syst. Dynam*, **11**, 885–901, doi:10.5194/esd-11-885-2020.

Mioduszewski, J. R., S. Vavrus, M. Wang, M. Holland, and L. Landrum, 2019: Past and future interannual variability in Arctic sea ice in coupled climate models. *Cryosphere*, **13**, 113–124, doi:10.5194/tc-13-113-2019.

Murphy, L. N., J. M. Klavans, A. C. Clement, and M. A. Cane, 2021: Investigating the Roles of External Forcing and Ocean Circulation on the Atlantic Multidecadal SST Variability in a Large Ensemble Climate Model Hierarchy. *Journal of Climate*, 1–51, doi:10.1175/JCLI-D-20-0167.1.

NASA Goddard Institute for Space Studies, 2018: NASA-GISS GISS-E2.1G model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.7127.

— 2019a: NASA-GISS GISS-E2-2-G model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.7129.

— 2019b: NASA-GISS GISS-E2.1H model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.7128.

— 2019c: NASA-GISS GISS-E2.2H model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.15871.

Neubauer, D., S. Ferrachat, C. S.-L. Drian, J. Stoll, D. S. Folini, I. Tegen, K.-H. Wieners, T. Mauritsen, I. Stemmler, S. Barthel, I. Bey, N. Daskalakis, B. Heinold, H. Kokkola, D. Partridge, S. Rast, H. Schmidt, N. Schutgens, T. Stanelle, P. Stier, D. Watson-Parris, and U. Lohmann, 2019: HAMMOZ-Consortium MPI-ESM1.2-HAM model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.5016.

Nicolaus, M., D. K. Perovich, G. Spreen, M. A. Granskog, L. von Albedyll, M. Angelopoulos, P. Anhaus, S. Arndt, H. J. Belter, V. Bessonov, G. Birnbaum, J. Brauchle, R. Calmer, E. Cardellach, B. Cheng, D. Clemens-Sewall, R. Dadic, E. Damm, G. de Boer, O. Demir, K. Dethloff, D. V. Divine, A. A. Fong, S. Fons, M. M. Frey, N. Fuchs, C. Gabarró, S. Gerland, H. F. Goessling, R. Gradinger, J. Haapala, C. Haas, J. Hamilton, H. R. Hannula, S. Hendricks, A. Herber, C. Heuzé, M. Hoppmann, K. V. Høyland, M. Huntemann, J. K. Hutchings, B. Hwang, P. Itkin, H. W. Jacobi, M. Jaggi, A. Jutila, L. Kaleschke, C. Katlein, N. Kolabutin, D. Krampe, S. S. Kristensen, T. Krumpen, N. Kurtz, A. Lampert, B. A. Lange, R. Lei, B. Light, F. Linhardt, G. E. Liston, B. Loose, A. R. Macfarlane, M. Mahmud, I. O. Matero, S. Maus, A. Morgenstern, R. Naderpour, V. Nandan, A. Niubom, M. Oggier, N. Oppelt, F. Pätzold, C. Perron, T. Petrovsky, R. Pirazzini, C. Polashenski, B. Rabe, I. A. Raphael, J. Regnery, M. Rex, R. Ricker, K. Riemann-Campe, A. Rinke, J. Rohde, E. Salganik, R. K. Scharien, M. Schiller, M. Schneebeli, M. Semmling, E. Shimanchuk, M. D. Shupe, M. M. Smith, V. Smolyanitsky, V. Sokolov, T. Stanton, J. Stroeve, L. Thielke, A. Timofeeva, R. T. Tonboe, A. Tavri, M. Tsamados, D. N. Wagner, D. Watkins, M. Webster, and M. Wendisch, 2022: Overview of the MOSAiC expedition: Snow and sea ice. *Elementa*, **10**, doi:10.1525/elementa.2021.000046.

Niederdrenk, A. L. and D. Notz, 2018: Arctic Sea Ice in a 1.5°C Warmer World. *Geophysical Research Letters*, **45**, 1963–1971, doi:10.1002/2017GL076159.

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

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

Notz, D. and J. Marotzke, 2012: Observations reveal external driver for Arctic sea-ice retreat. *Geophysical Research Letters*, **39**, doi:10.1029/2012GL051094.

Notz, D. and SIMIP-Community, 2020: Arctic Sea Ice in CMIP6. *Geophysical Research Letters*, **47**, doi:10.1029/2019GL086749.

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

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

Ogi, M. and J. M. Wallace, 2007: Summer minimum Arctic sea ice extent and the associated summer atmospheric circulation. *Geophysical Research Letters*, **34**, 1–4, doi:10.1029/2007GL029897.

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

Olonscheck, D. and D. Notz, 2017: Consistently estimating internal climate variability from climate model simulations. *Journal of Climate*, **30**, 9555–9573, doi:10.1175/JCLI-D-16-0428.1.

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

O'neill, B. C., C. Tebaldi, D. P. V. Vuuren, V. Eyring, P. Friedlingstein, G. Hurtt, R. Knutti, E. Kriegler, J.-F. Lamarque, J. Lowe, G. A. Meehl, R. Moss, K. Riahi, and B. M. Sanderson, 2016: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geosci. Model Dev*, **9**, 3461–3482, doi:10.5194/gmd-9-3461-2016.

Petrick, S., K. Riemann-Campe, S. Hoog, C. Growitsch, H. Schwind, R. Gerdes, and K. Rehdanz, 2017: Climate change, future Arctic Sea ice, and the competitiveness of European Arctic offshore oil and gas production on world markets. *Ambio*, **46**, 410–422, doi:10.1007/s13280-017-0957-z.

Phillips, A., C. Deser, and J. Fasullo, 2014: A New Tool for Evaluating Modes of Variability in Climate Models. *Eos*, **95**, 453–455, doi:10.1002/2014EO490002.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres*, **108**, 4407, doi:10.1029/2002jd002670.

Ridley, J., M. Menary, T. Kuhlbrodt, M. Andrews, and T. Andrews, 2019a: MOHC HadGEM3-GC31-LL model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.6109.

— 2019b: MOHC HadGEM3-GC31-MM model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.6112.

Roach, L. A. and E. Blanchard-Wrigglesworth, 2022: Observed Winds Crucial for September Arctic Sea Ice Loss. *Geophysical Research Letters*, **49**, 1–10, doi:10.1029/2022GL097884.

Roberts, J. and T. D. Roberts, 1978: Use of the Butterworth low-pass filter for oceanographic data. **83**, 5510–5514, doi:10.1029/JC083iC11p05510.

Rodgers, K. B., J. Lin, and T. L. Frölicher, 2015: Emergence of multiple ocean ecosystem drivers in a large ensemble suite with an Earth system model. *Biogeosciences*, **12**, 3301–3320, doi:10.5194/bg-12-3301-2015.

Rogers, J. C., 1981: The North Pacific Oscillation. *Journal of Climatology*, **1**, 39–57, doi:10.1002/joc.3370010106.

Rong, X., 2019: CAMS CAMS-CSM1.0 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.9754.

Rosenblum, E. and I. Eisenman, 2017: Sea ice trends in climate models only accurate in runs with biased global warming. *Journal of Climate*, **30**, 6265–6278, doi:10.1175/JCLI-D-16-0455.1.

Santer, B. D., P. W. Thorne, L. Haimberger, K. E. Taylor, T. M. Wigley, J. R. Lanzante, S. Solomon, M. Free, P. J. Gleckler, P. D. Jones, T. R. Karl, S. A. Klein, C. Mears, D. Nychka, G. A. Schmidt, S. C. Sherwood, and F. J. Wentz, 2008: Consistency of modelled and observed temperature trends in the tropical troposphere. *International Journal of Climatology*, **28**, 1703–1722, doi:10.1002/JOC.1756.

Schupfner, M., K.-H. Wieners, F. Wachsmann, C. Steger, M. Bittner, J. Jungclaus, B. Früh, K. Pankatz, M. Giorgetta, C. Reick, S. Legutke, M. Esch, V. Gayler, H. Haak, P. de Vrese, T. Raddatz, T. Mauritsen, J.-S. von Storch, J. Behrens, V. Brovkin, M. Claussen, T. Crueger, I. Fast, S. Fiedler, S. Hagemann, C. Hohenegger, T. Jahns, S. Kloster, S. Kinne, G. Lasslop, L. Kornblueh, J. Marotzke, D. Matei, K. Meraner, U. Mikolajewicz, K. Modali, W. Müller, J. Nabel, D. Notz, K. P. von Gehlen, R. Pincus, H. Pohlmann, J. Pongratz, S. Rast, H. Schmidt, R. Schnur, U. Schulzweida, K. Six, B. Stevens, A. Voigt, and E. Roeckner, 2019: DKRZ MPI-ESM1.2-HR model output prepared for CMIP6 historical. doi:10.22033/ESGF/CMIP6.6594.

Schweiger, A., R. Lindsay, J. Zhang, M. Steele, H. Stern, R. Kwok, C. . Schweiger, R. Lindsay, J. Zhang, M. Steele, H. Stern, and R. Kwok, 2011: Uncertainty in modeled Arctic sea ice volume. *J. Geophys. Res*, **116**, 0–06, doi:10.1029/2011JC007084.

Screen, J. A. and C. Deser, 2019: Pacific Ocean Variability Influences the Time of Emergence of a Seasonally Ice-Free Arctic Ocean. *Geophysical Research Letters*, **46**, 2222–2231, doi:10.1029/2018GL081393.

Seferian, R., 2018: CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.4068.

Seland, , M. Bentsen, D. J. L. Oliviè, T. Toniazzo, A. Gjermundsen, L. S. Graff, J. B. Debernard, A. K. Gupta, Y. He, A. Kirkevåg, J. Schwinger, J. Tjiputra, K. S. Aas, I. Bethke, Y. Fan, J. Griesfeller, A. Grini, C. Guo, M. Ilicak, I. H. H. Karset, O. A. Landgren, J. Liakka, K. O. Moseid, A. Nummelin, C. Spensberger, H. Tang, Z. Zhang, C. Heinze, T. Iversen, and M. Schulz, 2019: NCC NorESM2-LM model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.8036.

Serreze, M. C., M. M. Holland, and J. Stroeve, 2007: Perspectives on the Arctic's Shrinking Sea-Ice Cover. *Science*, **315**, 1533–1536, doi:10.1126/science.1139426.

Shiogama, H., H. Tatebe, M. Hayashi, M. Abe, M. Arai, H. Koyama, Y. Imada, Y. Kosaka, T. Ogura, and M. Watanabe, 2023: MIROC6 Large Ensemble (MIROC6-LE): experimental

design and initial analyses. *Earth System Dynamics*, **14**, 1107–1124, doi:10.5194/esd-14-1107-2023.

Sigmond, M., J. C. Fyfe, and N. C. Swart, 2018: Ice-free Arctic projections under the Paris Agreement. *Nature Climate Change*, **8**, 404–408, doi:10.1038/s41558-018-0124-y.

Simon, A., G. Gastineau, C. Frankignoul, V. Lapin, and P. Ortega, 2022: Pacific Decadal Oscillation modulates the Arctic sea-ice loss influence on the midlatitude atmospheric circulation in winter. *Weather and Climate Dynamics*, **3**, 845–861, doi:10.5194/wcd-3-845-2022.

Smith, D. M., R. Eade, M. B. Andrews, H. Ayres, A. Clark, S. Chripko, C. Deser, N. J. Dunstone, J. García-Serrano, G. Gastineau, L. S. Graff, S. C. Hardiman, B. He, L. Hermanson, T. Jung, J. Knight, X. Levine, G. Magnusdottir, E. Manzini, D. Matei, M. Mori, R. Msadek, P. Ortega, Y. Peings, A. A. Scaife, J. A. Screen, M. Seabrook, T. Semmler, M. Sigmond, J. Streffing, L. Sun, and A. Walsh, 2022: Robust but weak winter atmospheric circulation response to future Arctic sea ice loss. *Nature Communications*, **13**, doi:10.1038/s41467-022-28283-y.

Smith, L. C. and S. R. Stephenson, 2013: New Trans-Arctic shipping routes navigable by mid-century. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, E1191–E1195, doi:10.1073/PNAS.1214212110.

Song, Z., F. Qiao, Y. Bao, Q. Shu, Y. Song, and X. Yang, 2019: FIO-QLNM FIO-ESM2.0 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.9199.

Stroeve, J., M. M. Holland, W. Meier, T. Scambos, and M. Serreze, 2007: Arctic sea ice decline: Faster than forecast. *Geophysical Research Letters*, **34**, 9501, doi:10.1029/2007GL029703.

Stroeve, J. C., V. Kattsov, A. Barrett, M. Serreze, T. Pavlova, M. Holland, and W. N. Meier, 2012: Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations. *Geophysical Research Letters*, **39**, L16502, doi:10.1029/2012GL052676.

Suarez-Gutierrez, L., S. Milinski, and N. Maher, 2021: Exploiting large ensembles for a better yet simpler climate model evaluation. *Climate Dynamics*, **57**, 2557–2580, doi:https://doi.org/10.1007/s00382-021-05821-w.

Sun, L., M. Alexander, and C. Deser, 2018: Evolution of the Global Coupled Climate Response to Arctic Sea Ice Loss during 1990–2090 and Its Contribution to Climate Change. *Journal of Climate*, **31**, 7823–7843, doi:10.1175/JCLI-D-18-0134.1.

Swart, N. C., J. N. S. Cole, V. V. Kharin, M. Lazare, J. F. Scinocca, N. P. Gillett, J. Anstey, V. Arora, J. R. Christian, Y. Jiao, W. G. Lee, F. Majaess, O. A. Saenko, C. Seiler, C. Seinen, A. Shao, L. Solheim, K. von Salzen, D. Yang, B. Winter, and M. Sigmond, 2019a: CCCma CanESM5-CanOE model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.10260.

— 2019b: CCCma CanESM5 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.3610.

Swart, N. C., J. C. Fyfe, E. Hawkins, J. E. Kay, and A. Jahn, 2015: Influence of internal variability on Arctic sea-ice trends. *Nature Climate Change*, **5**, 86–89, doi:10.1038/nclimate2483.

Sévellec, F., A. V. Fedorov, and W. Liu, 2017: Arctic sea-ice decline weakens the Atlantic Meridional Overturning Circulation. *Nature Climate Change*, **7**, 604–610, doi:10.1038/nclimate3353.

Tang, Y., S. Rumbold, R. Ellis, D. Kelley, J. Mulcahy, A. Sellar, J. Walton, and C. Jones, 2019: MOHC UKESM1.0-LL model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.6113.

Tatebe, H. and M. Watanabe, 2018: MIROC MIROC6 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.5603.

Thompson, D. W., E. A. Barnes, C. Deser, W. E. Foust, and A. S. Phillips, 2015: Quantifying the role of internal climate variability in future climate trends. *Journal of Climate*, **28**, 6443–6456, doi:10.1175/JCLI-D-14-00830.1.

Tietsche, S., J. J. Day, V. Guemas, W. J. Hurlin, S. P. Keeley, D. Matei, R. Msadek, M. Collins, and E. Hawkins, 2014: Seasonal to interannual Arctic sea ice predictability in current global climate models. *Geophysical Research Letters*, **41**, 1035–1043, doi:10.1002/2013GL058755.

Topál, D., Q. Ding, J. Mitchell, I. Baxter, M. Herein, T. Haszpra, R. Luo, and Q. Li, 2020: An internal atmospheric process determining summertime Arctic sea ice melting in the next three decades: Lessons learned from five large ensembles and multiple CMIP5 climate simulations. *Journal of Climate*, **33**, 7431–7454, doi:10.1175/JCLI-D-19-0803.1.

Trenberth, K. E., 1997: The Definition of El Niño. *Bulletin of the American Meteorological Society*, **78**, 2771–2778, doi:10.1175/1520-0477(1997)078¡2771:TDOENO¿2.0.CO;2.

Trenberth, K. E. and D. J. Shea, 2006: Atlantic hurricanes and natural variability in 2005. *Geophysical Research Letters*, **33**, doi:10.1029/2006GL026894.

Ukita, J., M. Honda, H. Nakamura, Y. Tachibana, D. J. Cavalieri, C. L. Parkinson, H. Koide, and K. Yamamoto, 2007: Northern Hemisphere sea ice variability: Lag structure and its implications. *Tellus, Series A: Dynamic Meteorology and Oceanography*, **59**, 261–272, doi:10.1111/j.1600-0870.2006.00223.x.

Uotila, P., S. O'farrell, S. J. Marsland, and D. Bi, 2013: The sea-ice performance of the Australian climate models participating in the CMIP5. *Australian Meteorological and Oceanographic Journal*, **63**, 121–143, doi:10.22499/2.6301.008.

VanAchter, G., L. Ponsoni, F. Massonnet, T. Fichefet, and V. Legat, 2020: Arctic sea ice thickness internal variability and its changes under historical and anthropogenic forcing. *Cryosphere*, **14**, 3479–3486, doi:10.5194/tc-14-3479-2020.

Vance, T. R., A. S. Kiem, L. M. Jong, J. L. Roberts, C. T. Plummer, A. D. Moy, M. A. Curran, and T. D. van Ommen, 2022: Pacific decadal variability over the last 2000 years and implications for climatic risk. *Communications Earth and Environment*, **3**, doi:10.1038/s43247-022-00359-z.

Voldoire, A., 2018: CNRM-CERFACS CNRM-CM6-1 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.4066.

Volodin, E., E. Mortikov, A. Gritsun, V. Lykossov, V. Galin, N. Diansky, A. Gusev, S. Kostrykin, N. Iakovlev, A. Shestakova, and S. Emelina, 2019: INM INM-CM5-0 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.5070.

Wang, M. and J. E. Overland, 2012: A sea ice free summer Arctic within 30 years: An update from CMIP5 models. *Geophysical Research Letters*, **39**, 6–11, doi:10.1029/2012GL052868.

Wang, Q., S. Danilov, L. Mu, D. Sidorenko, and C. Wekerle, 2021: Lasting impact of winds on Arctic sea ice through the ocean's memory. *Cryosphere*, **15**, 4703–4725, doi:10.5194/tc-15-4703-2021.

Watanabe, S., T. Hajima, K. Sudo, M. Abe, O. Arakawa, K. Ogochi, T. Arakawa, H. Tatebe, A. Ito, A. Ito, Y. Komuro, T. Nitta, M. A. Noguchi, T. Ogura, R. Ohgaito, M. Sekiguchi, T. Suzuki, K. Tachiiri, K. Takata, T. Takemura, M. Watanabe, A. Yamamoto, D. Yamazaki, K. Yoshimura, and M. Kawamiya, 2021: MIROC MIROC-ES2H model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.5601.

Webster, M. A., M. Holland, N. C. Wright, S. Hendricks, N. Hutter, P. Itkin, B. Light, F. Linhardt, D. K. Perovich, I. A. Raphael, M. M. Smith, L. Von Albedyll, and J. Zhang, 2022: Spatiotemporal evolution of melt ponds on Arctic sea ice: MOSAiC observations and model results. *Elementa*, **10**, 1–22, doi:10.1525/elementa.2021.000072.

Weijer, W., W. Cheng, O. A. Garuba, A. Hu, and B. T. Nadiga, 2020: CMIP6 Models Predict Significant 21st Century Decline of the Atlantic Meridional Overturning Circulation. *Geophysical Research Letters*, **47**, doi:10.1029/2019GL086075.

Werb, B. E. and D. L. Rudnick, 2023: Remarkable Changes in the Dominant Modes of North Pacific Sea Surface Temperature. *Geophysical Research Letters*, **50**, 1–8, doi:10.1029/2022GL101078.

Wettstein, J. J. and C. Deser, 2014: Internal variability in projections of twenty-first-century Arctic sea ice loss: Role of the large-scale atmospheric circulation. *Journal of Climate*, **27**, 527–550, doi:10.1175/JCLI-D-12-00839.1.

Wieners, K.-H., M. Giorgetta, J. Jungclaus, C. Reick, M. Esch, M. Bittner, V. Gayler, H. Haak, P. de Vrese, T. Raddatz, T. Mauritsen, J.-S. von Storch, J. Behrens, V. Brovkin, M. Claussen, T. Crueger, I. Fast, S. Fiedler, S. Hagemann, C. Hohenegger, T. Jahns, S. Kloster, S. Kinne, G. Lasslop, L. Kornblueh, J. Marotzke, D. Matei, K. Meraner, U. Mikolajewicz, K. Modali, W. Müller, J. Nabel, D. Notz, K. P. von Gehlen, R. Pincus, H. Pohlmann, J. Pongratz, S. Rast, H. Schmidt, R. Schnur, U. Schulzweida, K. Six, B. Stevens, A. Voigt, and E. Roeckner, 2019: MPI-M MPIESM1.2-LR model output prepared for CMIP6 historical. doi:10.22033/ESGF/CMIP6.6595.

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

Winton, M., M. Bushuk, Y. Zhang, B. Hurlin, L. Jia, N. C. Johnson, and F. Lu, 2022: Prospects for Seasonal Prediction of Summertime Trans-Arctic Sea Ice Path. *Journal of Climate*, **35**, 4253–4263, doi:10.1175/JCLI-D-21-0634.1.

Wu, T., M. Chu, M. Dong, Y. Fang, W. Jie, J. Li, W. Li, Q. Liu, X. Shi, X. Xin, J. Yan, F. Zhang, J. Zhang, L. Zhang, and Y. Zhang, 2018: BCC BCC-CSM2MR model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.2948.

Wyburn-Powell, C., A. Jahn, and M. England, 2022: Modeled Interannual Variability of Arctic Sea Ice Cover is Within Observational Uncertainty. *Journal of Climate*, **35**, 6827–6842, doi:10.1175/jcli-d-21-0958.1.

Wyser, K., T. Koenigk, U. Fladrich, R. Fuentes-Franco, M. P. Karami, and T. Kruschke, 2021: The SMHI Large Ensemble (SMHI-LENS) with EC-Earth3.3.1. *Geoscientific Model Development*, **14**, 4781–4796, doi:10.5194/gmd-14-4781-2021.

Yeager, S. G., A. R. Karspeck, and G. Danabasoglu, 2015: Predicted slowdown in the rate of Atlantic sea ice loss. *Geophysical Research Letters*, **42**, 10704–10713, doi:10.1002/2015GL065364.

Yuan, X., M. R. Kaplan, and M. A. Cane, 2018: The interconnected global climate system-a review of tropical-polar teleconnections. *Journal of Climate*, **31**, 5765–5792, doi:10.1175/JCLI-D-16-0637.1.

Yukimoto, S., T. Koshiro, H. Kawai, N. Oshima, K. Yoshida, S. Urakawa, H. Tsujino, M. Deushi, T. Tanaka, M. Hosaka, H. Yoshimura, E. Shindo, R. Mizuta, M. Ishii, A. Obata, and Y. Adachi, 2019: MRI MRI-ESM2.0 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.6842.

Zebiak, S. E., 1993: Air–sea interaction in the equatorial atlantic region. *Journal of Climate*, **6**, 1567–1586, doi:10.1175/1520-0442(1993)006¡1567:AIITEA¿2.0.CO;2.

Zhang, J., T. Wu, X. Shi, F. Zhang, J. Li, M. Chu, Q. Liu, J. Yan, Q. Ma, and M. Wei, 2018: BCC BCC-ESM1 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.2949.

Zhang, M., W. Perrie, and Z. Long, 2019: Springtime North Pacific Oscillation and summer sea ice in the Beaufort sea. *Climate Dynamics*, **53**, 671–686, doi:10.1007/s00382-019-04627-1.

Zhang, R. and J. M. Wallace, 2015: Mechanisms for low-frequency variability of summer Arctic sea ice extent. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, doi:10.1073/pnas.1422296112.

Zhang, S., T. Y. Gan, and A. B. Bush, 2020: Variability of Arctic sea ice based on quantile regression and the teleconnection with large-scale climate patterns. *Journal of Climate*, **33**, 4009–4025, doi:10.1175/JCLI-D-19-0375.1.

Ziehn, T., M. Chamberlain, A. Lenton, R. Law, R. Bodman, M. Dix, Y. Wang, P. Dobrohotoff, J. Srbinovsky, L. Stevens, P. Vohralik, C. Mackallah, A. Sullivan, S. O'Farrell, and K. Druken, 2019: CSIRO ACCESS-ESM1.5 model output prepared for CMIP6 CMIP historical. doi:10.22033/ESGF/CMIP6.4272.

Årthun, M., I. H. Onarheim, J. Dörr, and T. Eldevik, 2021: The Seasonal and Regional Transition to an Ice-Free Arctic. *Geophysical Research Letters*, **48**, e2020GL090825, doi:10.1029/2020GL090825.