Sea-ice prediction across timescales and the role of model complexity

PhD thesis

by

Lorenzo Zampieri

Department 1

Physics and Electrical Engineering

Universität Bremen

October 21, 2020

Abstract

In addition to observations and lab experiments, the scientific investigation of the Arctic and Antarctic sea ice is conducted through the employment of geophysical models. These models describe in a numerical framework the physical behavior of sea ice and its interactions with the atmosphere, ocean, and polar biogeochemical systems. Sea-ice models find application in the quantification of the past, present, and future sea-ice evolution, which becomes particularly relevant in the context of a warming climate system that causes the reduction of the Arctic sea ice cover. Because of the sea-ice decline, the navigation in the Arctic ocean increased substantially in the recent past, a trend that it is expected to continue in the next decades and that requires the formulation of reliable sea-ice predictions at various timescales. Sea-ice predictions can be delivered by modern forecast systems that feature dynamical sea-ice models. The simulation of sea ice is at the center of this thesis: A coupled climate model with a simple sea-ice component is used to quantify potential impacts of a geoengineering approach termed "Arctic Ice Management"; the skill of current operational subseasonal-to-seasonal sea-ice forecasts, based on global models with a varying degree of sea-ice model complexity, is evaluated; and, lastly, an unstructured-grid ocean model is equipped with state-of-the-art sea-ice thermodynamics to study the impact of sea-ice model complexity on model performance.

In chapter 2, I examine the potential of a geoengineering strategy to restore the Arctic sea ice and to mitigate the warming of the Arctic and global climate throughout the 21$^{\text{st}}$ century. The results, obtained with a fully coupled climate model, indicate that it is theoretically possible to delay the melting of the Arctic sea ice by $\sim$ 60 years, but that this does not reduce global warming. In chapters 3 and 4, I assess the skill of global operational

ensemble prediction systems in forecasting the evolution of the Arctic and Antarctic sea-ice edge position at subseasonal timescales. I find that some systems produce skillful forecasts more than 1.5 months ahead, but I also find evidence of substantial model biases and issues concerning data assimilation and model formulation.

Chapter 5 deals with the impact of sea-ice model complexity on model performance. I present a new formulation of the FESOM2 sea-ice/ocean model with a revised description of the sea-ice thermodynamics, including various parameterizations of physical processes at the subgrid-scale. The model formulation grants substantial modularity in terms of sea-ice physics and resolution. The new system is used for assessing the impact of the sea-ice model complexity on the FESOM2 performance in different atmosphere-forced setups with a specific parameter-tuning approach and a special focus on sea-ice related variables. The results evidence that a more sophisticated model formulation is beneficial for the model representation of the sea-ice concentration and snow thickness, while less relevant for sea-ice thickness and drift. I also highlight a dependence of the model performance on the atmospheric forcing product used as boundary condition.

In the final part of this thesis, I formulate recommendations for future developments in the field of sea-ice modeling, with particular emphasis on FESOM2 and, more generally, on the modeling infrastructure under development at the Alfred Wegener Institute.

Contents

1.   Introduction

The reader will forgive me if I begin this Ph.D. thesis with a brief digression, but I cannot resist acknowledging what, for a lot of reasons, made this project on sea-ice modeling possible and so exciting: the inorganic compound H_2O and its remarkable properties.

Every time an oxygen atom bonds covalently with two hydrogen atoms, a H_2O molecule is formed. This molecule is a fundamental brick for a lot of substances on our planet, and incidentally, it is the prime constituent of sea-ice. The electrons of this peculiar molecule cannot resist the charming attraction of the oxygen nucleus, and, consequently, they can be found more often in the oxygen vicinity than around the two hydrogen nucleus. The key to the existence of sea ice as we know it resides in this electronic imbalance, a behavior which in physical terms is described as polarity: the separation of positive and negative electrical charges within a molecule. The H_2O polarity allows these molecules to interact with each other, and the strength of the interactions varies depending on the thermal energy of the system, a quantity which is measured by temperature. For temperatures below freezing (T<0°C, at standard environmental pressure and without solutes), the H_2O molecules are less mobile and geometrically organized in space, leading to a solid state that we call ice. At higher temperatures (0°C<T<100°C), the molecular motion is more energetic and chaotic, leading to a fluid state which we call liquid water. Interestingly enough, H_2O is one of the few known substances for which the distance between molecules is slightly larger in the solid phase compared to the liquid phase. This implies that the ice is less dense than water and therefore fluctuates on it. As trivial as this behavior might appear to everyone who ever sipped a cold drink on a summer day, our oceans would look rather different if the H_2O molecules would interact in other ways, with thick sea ice filling the bottom of the ocean rather than covering the polar seas as a dynamic reflective blanket.

At this point, allow me to put aside my description of H_2O and to pause, just for a while, the suggestive speculations on sinking sea ice. I promise to return to them at the end of this chapter. Before this, I will describe the crucial role that sea ice plays in the climate system and for society, and I will present the efforts of the scientific community for modeling this remarkable material.

1.1 Sea ice in a changing climate system

Sea ice constitutes only a very small fraction of the total ice volume on our planet. However, the Arctic and Antarctic sea-ice systems combined occupy on average an area of 18×10^6 km^2, which corresponds to approximately 5% of the surface of the global oceans. Because of its wide-spread coverage, and due to the very different physical properties compared to the ocean and atmosphere, the sea-ice plays a role of prime importance in the climate system.

I previously condensed the description of sea-ice in the words *"dynamic reflective blanket"*. As a *blanket*, sea-ice acts as a thermal insulator, reducing the turbulent heat fluxes from the warmer ocean to the colder atmosphere during winter, and mitigating the warming of the ocean by absorbing heat during the polar summers (Notz, 2005). Sea-ice is overall a bright and thus *reflective* material, and even brighter is the snow that precipitates from the atmosphere and accumulates over it. For this reason, a large fraction of the incoming solar radiation in sea-ice covered regions is reflected back to space, contributing to keeping the polar atmosphere and ocean at relatively cool temperatures all year round (Perovich, 2003). The sea-ice system is not continuous and rigid, but it is constituted by a *dynamic* ensemble of floes with sizes that span over several orders of magnitude (Gherardi & Lagomarsino, 2015). Generally, the sea-ice cover is organized in larger floes during winter, when multiple floes freeze together and the sea ice becomes thicker, more compact, and less mobile. In contrast, it breaks into smaller heterogeneous floes during summer, when the ice warms and becomes weaker (Stern *et al.*, 2018). The floes that constitute the sea ice move in response to atmospheric winds, ocean currents, and interactions with other floes (Leppranta, 2009). Therefore, sea-ice is not only a mediator of the atmospheric-oceanic heat exchange, but also of the transfer of momentum from the atmosphere to the

ocean. Furthermore, the seasonal formation of new sea-ice and the consequent rejection of salty brine into the ocean destabilizes the water column and contributes to support the ocean thermohaline circulation on a global scale. In contrast, the melting of sea ice freshens the surface ocean and has a stabilizing effect on its vertical stratification (Rahmstorf, 1999).

Even though the fundamental physical processes that lead to sea-ice formation are very similar in the Arctic and in Antarctic regions, the resulting sea-ice differs substantially between the two hemispheres. The Arctic ocean is a rather closed environment, surrounded by the American and Eurasian continents, and with few localized exchange gates with other oceans. The configuration of the coastlines, in combination with the typical modes of atmospheric and oceanic circulation, favors the recirculation of the sea-ice into the Arctic basin and reduces its dispersion into warmer peripheral seas. This allows a substantial part of the sea-ice cover to survive multiple melting season and to form thicker multi-year ice which, nowadays, can be found mostly north of Greenland and off the Canadian Archipelago (Maslanik *et al.*, 2011). On the contrary, the Antarctic sea-ice forms in the Southern Ocean, a wide ocean basin that surrounds the Antarctic continent and that is in direct communication with warmer oceans at its northern boundaries. The seasonal cycle of the Antarctic sea-ice extent is wider than in the Arctic, with a winter expansion contained by the atmospheric jet stream and by the influence of the Antarctic Circumpolar Current, followed by a rapid melting phase that depletes most sea-ice cover. Nevertheless, some multi-year ice can be found in few protected locations at the end of the melting season (mostly in the Weddel Sea), while rigid and immobile patches of land-fast sea ice persist all year round along the Antarctic coastlines (Fraser *et al.*, 2020), attached to the ice shelves or to grounded icebergs.

The Arctic and Antarctic sea-ice systems respond very differently to the anthropogenic global warming that has affected our planet in the last decades. While the Antarctic sea-ice extent features a slightly increasing but not statistically significant trend since the beginning of satellite observational records in the late '70s (Parkinson, 2019), the Arctic sea ice exhibits a marked decline (Mueller *et al.*, 2018) in extent (Stroeve *et al.*, 2007), volume (Gascard *et al.*, 2019), and age (Comiso, 2012; Kwok, 2018). The rapid mutations that the sea-ice is facing, particularly in the Arctic, fostered the interest of the scientific community in studying this system. On one hand, this responds to an urgent need of understanding the impact of global warming on polar

regions, of quantifying feedback mechanisms that could amplify or reduce the current warming trends, and of preserving the fragile but rich ecosystem that thrives around the sea ice. On the other hand, the technological advances of our societies, in combination with the reduction of the sea-ice, open the polar regions to a more sustained human presence that is expected to grow in the near future (Larsen & Fondahl, 2015), demanding therefore an in-depth understanding of the sea-ice system for enabling a safe, sustainable, and proficuous development process of the polar regions.

1.2   Fifty years of sea-ice modeling

In addition to satellite observations, in-situ measurements, and lab experiments, the study of sea-ice is nowadays carried out through the employment of sophisticated mathematical models. The ultimate goal of these models is to simulate quantitatively the behavior of the sea ice, its evolution, and its interactions with the other components of the climate system.

When analyzing the evolution of sea ice models, I cannot refrain from connecting major developments in this field to fundamental advances in the field of sea-ice observations. The understanding of the heat conduction in sea ice, which led to the first mature sea-ice model of Maykut & Untersteiner (1971), would not have been possible without the data collected during the International Geophysical Year in 1957-58. Before the availability of the first satellite observations, the *'Arctic Ice Dynamics Joint Experiment'* (AIDJEX)—a series of experimental campaigns carried out during the '70s—laid the foundations for understanding the motion of sea ice in the Arctic, which contributed to the formulation of the first dynamical and thermodynamic model of sea ice by Hibler (1979), featuring a Viscous Plastic (VP) sea-ice rheology. The observations from the *'Surface Heat Budget of the Arctic'* (SHEBA) campaign (Uttal *et al.*, 2002) advanced our understanding of the interactions between sea ice and snow with radiation, resulting in a series of more physically consistent parameterizations added to the sea-ice models (e.g. Holland *et al.* (2012)). The *'Multidisciplinary drifting Observatory for the Study of Arctic Climate'* (MOSAiC)—arguably the largest Arctic expedition in history—can be considered the next step on this pathway, as this scientific initiative is explicitly designed to build a solid understanding of sea ice that will shape the future of sea-ice modeling. In this respect, future modeling improvements

are expected in the description of the snow layer on top of the sea ice, in the representation of biochemical processes, and finally in the exchange processes at the interfaces between the sea ice, ocean, and atmosphere.

New and better sea-ice observations have not only contributed to improving model performances, but they are responsible for continuously shaping the way we model the sea ice and how we analyze the model results. Right after the first sea-ice concentration satellite retrieval became available in the late '70s, Parkinson & Washington (1979) formulated a sea-ice model that allowed, for the first time, the coexistence of frozen sea ice and open water in the same grid cell. Such a formulation survives today in our models as the "sea-ice concentration" variable, and it is extremely successful because it allows easy comparison between observations and model data. Forty years afterward, many more examples of observational hybridization can be found among the variables of our sea-ice models. For example, we are now able to simulate the fraction of multi-year-ice, the sea-ice age, the sea-ice freeboard, etc. The adoption of these formulations in models was fostered by the development of compatible observations, mostly from remote sensing. In light of the success of sea-ice models and of their broad application, the aforementioned process has now also changed direction, with the requirements of sea-ice modeling and data assimilation communities explicitly taken into account when planning new satellite missions and observational campaigns.

The development of numerical sea-ice models has been driven not only by the growing quality of observations but also by the availability of computational resources for running the models and for analyzing the results. As described by Hunke *et al.* (2010), the first mature model of sea ice by Maykut & Untersteiner (1971) was too detailed and sophisticated for the computing capabilities of the time, and had to be simplified by Semtner (1976) to be employed over larger domains and for climate applications. The implementation of an energy-conserving multi-layer sea-ice model for pan-Arctic setups has only been established almost three decades later by Bitz & Lipscomb (1999). As the computational resources available to the scientific community grew steadily, pan-Arctic frontier sea-ice simulations feature, nowadays, very high spatial resolutions that reach the kilometer-scale (Menemenlis *et al.*, 2008; Wang *et al.*, 2020). At these resolutions, the continuum assumption, which is the foundation of classical sea-ice models, starts to fail, as the model resolution becomes comparable to the typical size of individual sea-ice floes. This does not necessarily imply that the use of the Hibler-type VP models

should be abandoned. In this respect, Ringeisen *et al.* (2019) show that the MITgcm model (Losch *et al.*, 2010) retains a remarkable consistency up to a resolution of 25m. Nevertheless, alternative sea-ice model formulations become attractive for certain types of applications that require high resolution. For example, Hunke *et al.* (2020c) calls for the use of Discrete Element Models (DEM; e.g. Herman (2016)) for simulating the anisotropic sea-ice pack at fine spatial scales (few hundred meters) and in small domains, a technique that would find its natural application into the field of operational short-term sea-ice forecasting.

In the past decades, sea-ice modeling played a key role in answering several scientific questions with a broad impact on our society. In this respect, the most remarkable achievement of sea-ice models is to have estimated correctly the declining trend of the Arctic sea-ice extent and volume, and to have demonstrated the causal link between this trend and the increase of greenhouse gases in the atmosphere (Notz & Stroeve, 2016). An interesting aspect that characterizes this finding is the multi-model framework from which it originated. The coordination and combination of different models have proven to be effective strategies for reducing model uncertainties and for increasing confidence in the model estimates, evidencing the importance of initiatives such as the Climate Model Intercomparison Project (CMIP) for formulating reliable climate projections.

At shorter timescales, sea-ice models are becoming more and more relevant for institutions and organizations that formulate operational environmental predictions (Jung *et al.*, 2016). On one side, this is motivated by the growing interest of stakeholders in a progressively ice-free and navigable Arctic Ocean, which opens new opportunities for the development of these regions, but which also emphasizes the need for predicting the evolution of the Arctic sea-ice to prevent hazards and to reduce the risks associated with the extreme polar environment (Emmerson & Lahn, 2012; Stephenson *et al.*, 2011). On the other side, improving the sea-ice representation in Earth System Models (ESM) is considered beneficial to increase our predictive capabilities of the ocean and atmosphere, with important relapses on human activities.

1.3 To each question its model

After having introduced the foundations on which modern sea-ice models rest, I continue this introductory chapter by presenting a rather drastic case study that, even if abstract and far from our classical sea-ice modelling experience, helps nonetheless to illustrate the background from which one of the arguments of this thesis arises.

Let us imagine, once more, a hypothetical world in which the H_2O molecules behave differently from what we are used to, causing the sea ice to sink into the ocean[1]. As the sea-ice density is represented by an arbitrary parameter in most of the models currently in use, we could legitimately think of studying this situation with one of these models. I embarked on such an exercise by increasing the sea-ice density in the Finite-volumE Sea ice-Ocean Model version 2 (FESOM2; Danilov *et al.* (2017)) sea-ice and ocean model to a value 10% larger than that of the ocean water. Interestingly, the model calculations do not fail, nor does the model produce any warning to flag an inappropriate density value. The outcome of the simulation is, not surprisingly, quite different from what the physical intuition would suggest: even if denser, the sea ice keeps fluctuating over the lighter ocean and does not sink into it. The major aspect that differs compared to a standard simulation is the thickness of the sea-ice, which is correctly reduced due to the higher density of the material. Does this mean that the FESOM2 model is wrong? Technically yes, in this situation FESOM2 produces nonphysical results. But, most importantly, we are trying to answer a rather strange, yet legitimate scientific question with an inappropriate tool that was not designed for this purpose. In fact, while allowing the coexistence of water and ice at the surface of the ocean, the model developers designed the sea-ice and ocean components of FESOM2 making the assumption—very reasonable before my attempt—of a complete separation of these two elements in the vertical column, with the sea-ice that, if present, floats over the ocean no matter the relative density of the two substances.

As described in Sec 1.2, both the complexity of sea-ice models and the number of problems and applications that they can address grew substantially over the past decades. Nevertheless, what has been demonstrated by the previous case study is that, despite the efforts made to adopt model for-

---

[1]Interestingly, D_2O in solid form or '*heavy ice*', a close relative of the standard ice, really sinks into H_2O (Maitra & Zare, 2016)

mulations simultaneously consistent with the laws of physics and with the observed sea-ice phenomenology, sea-ice models still return only a simplistic characterization of the real world and their performance is tightly connected to the research questions formulated by the model users. For this reason, each application that involves modelling should ideally adopt a specific and appropriate formulation of the model itself. In practice, this is rarely the case for sea-ice models, as the few available are shared across different scientific fields and employed for various applications. The approach adopted by the scientific community is surely efficient and not necessarily a bad one, provided to not forget that the level of complexity that might be required in some applications could become not justified, or even harmful, in other cases. I will revisit this argument in the concluding chapter of the thesis in light of the results presented in the continuation of this manuscript.

1.4 Scope and structure of the thesis

As emerges from the previous sections, the overarching theme of this Ph.D. dissertation is sea-ice modeling. In practice, this broad topic is here analyzed by following two main threads: the first concerns the investigations of sea ice through models, while the second regards the investigation of sea-ice models themselves. Although the majority of the studies here presented focus on the sea ice in the Arctic, I try to maintain a global perspective on the topic, keeping in mind that, as the laws of physics are the same in the two hemispheres, a good model should in principle be equally adequate in representing the Arctic and Antarctic sea ice.

The first thread of this thesis touches two of the scientific fields in which modern sea-ice models find their core area of application: climate projections at multi-decadal timescales on one side, and operational sea-ice predictions up to seasonal timescales on the other. As this thesis will demonstrate, the investigation of these problems is tackled with similar tools that nevertheless present some fundamental differences The approaches for the sea-ice investigation are, in fact, tightly connected to the timescale of the problem, and consequently to the scientific question and to the application requirements.

Climate projection studies typically deal with long timescales, and, in practice, this translates into a so-called boundary condition problem. The goal of these studies is understanding, in a statistical sense, the response

of sea ice to external forcings (i.e. boundary conditions), which can be, for example, the increase of greenhouse gases in the atmosphere. In this context, Chapter 2 presents a study in which the sea-ice component of a fully coupled climate model is adapted to investigate the impact on the climate system of a hypothetical geoengineering approach to counteract the decline of the Arctic sea ice. This geoengineering strategy, originally proposed by Desch *et al.* (2017) and called *'Arctic Ice Management'*, foresees a large number of winddriven pumps that shall spread seawater on the surface in winter to enhance ice growth, allowing more ice to survive the summer melt. We tested this hypothesis by modifying the surface exchange processes such that the physical effect of the pumps is simulated, and by performing century-long ensemble experiments that account for the increase of greenhouse gases in the atmosphere. The main scientific questions relative to this chapter (**Q1** and **Q2**) are summarized in the following box:

> **Q1** Can the actuation of the Arctic Ice Management strategy prevent the decline of the Arctic sea ice?
>
> **Q2** Can this approach mitigate the effects of global warming in the Arctic and beyond?

Chapter 2 has been published in the journal *'Earth's Future'* by Zampieri & Goessling (2019) under the title *'Sea Ice Targeted Geoengineering Can Delay Arctic Sea Ice Decline but not Global Warming'*.

In contrast to climate projections, operational sea-ice predictions deal with much shorter timescales and attempt to predict, as accurately as possible, the trajectory between the present and the future sea-ice state. The knowledge of the initial sea-ice state (but also of the ocean and atmosphere initial conditions) is therefore crucial for a correct simulation of the sea-ice evolution, leading to a so-called initial condition problem. Because a perfect characterization of the initial state is in practice impossible, and because of limitation and biases in our models, an ensemble of forecasts is produced from slightly perturbed initial states, aiming to give indications on the range of possible future states of the variable of interest and not only the most likely one. In this context, Chapters 3 and 4 investigate the skill of several operational forecasting systems in predicting the evolution of the sea ice

edge at subseasonal to seasonal (S2S) timescales, both in the Arctic and in the Antarctic. The verification methodology employed for evaluating these forecasts is similar for both hemispheres, and it takes into account the probabilistic nature of these ensemble forecasts. Furthermore, this study goes beyond the classical sea-ice extent and area concepts by focusing instead on the sea-ice spatial distribution, which is a relevant piece of information for stakeholders and potential final forecast users. The main scientific questions relative to these chapters (**Q3** and **Q4**) are summarized in the following box:

Q3 How skillful are state-of-the-art operational forecast systems in predicting the evolution of the Arctic and Antarctic sea-ice edge at subseasonal timescales?

Q4 Which biases affect the S2S forecasting systems? And where do they originate from?

Chapter 3 has been published in the journal *'Geophysical Research Letters'* by Zampieri *et al.* (2018) under the title *'Bright Prospects for Arctic Sea Ice Prediction on Subseasonal Time Scales'*. Chapter 4 has been published in the same journal by Zampieri *et al.* (2019) under the title *'Predictability of Antarctic Sea Ice Edge on Subseasonal Time Scales'*.

Chapter 5 develops the second thread of this thesis, shifting the focus from the investigation of sea ice through modeling, to the investigation of sea-ice models. Specifically, this chapter explores the impact of sea-ice model complexity on the performances of an unstructured-mesh sea-ice model under different atmospheric forcings. For making such a study possible, I have first equipped the unstructured global sea-ice and ocean model FESOM2 with the single-column sea-ice model Icepack, a set of physical parameterizations that describe the sub-grid sea-ice processes not explicitly resolved in models. The update has substantially broadened the range of physical processes that can be represented by FESOM2. These new features are directly implemented on the unstructured FESOM2 mesh, and thereby benefit from the unique flexibility that comes with it in terms of spatial resolution. A subset of the parameter space of three model configurations with increasing complexity has been calibrated with an iterative Green's function optimization method.

This optimization creates the conditions to test fairly the impact of the model complexity on the sea-ice representation in the FESOM2 model. The main scientific questions relative to this chapter (**Q5** and **Q6**) are summarized in the following box:

> **Q5** Does a more complex and physically consistent formulation of a sea-ice model lead to better sea-ice simulations?
>
> **Q6** How does the impact of different atmospheric forcings on sea-ice simulations compare to the impact of model complexity?

Chapter 5 corresponds to a manuscript under review in the *'Journal on Advances in Modelling Earth Systems'* (at the time this thesis was written) under the title *'Impact of sea-ice model complexity on the performance of an unstructured sea-ice/ocean model under different atmospheric forcings'*.

Chapter 6 concludes this thesis and summarizes the main findings and conclusions of my research. The results from the main chapters will be placed in context with the overarching theme of the thesis and, based on these final considerations, I will give an outlook and recommendations for future developments in the field of sea-ice modeling.

Remarks Chapters 2, 3, 4, and 5 constitute either published or ready-for-submission papers, which have been written together with my co-authors. The contributions of each individual are detailed at the beginning of the respective chapters. I decided to retain the original format of the manuscripts, which generates small inconsistencies with the rest of this thesis regarding style, utilization of the first person plural, and abbreviations. There is a certain degree of redundancy also content-wise, as each paper is formulated to be independent from the others. I kindly ask the reader to be indulgent and overlook these imperfections.

2. Sea ice targeted geoengineering can delay Arctic sea ice decline but not global warming[1]

Abstract

To counteract global warming, a geoengineering approach that aims at intervening in the Arctic icealbedo feedback has been proposed. A large number of winddriven pumps shall spread seawater on the surface in winter to enhance ice growth, allowing more ice to survive the summer melt. We test this idea with a coupled climate model by modifying the surface exchange processes such that the physical effect of the pumps is simulated. Based on experiments with RCP 8.5 scenario forcing, we find that it is possible to keep the latesummer sea ice cover at the current extent for the next ∼60 years. The increased ice extent is accompanied by significant Arctic latesummer cooling by ∼1.3 K on average north of the polar circle (20212060). However, this cooling is not conveyed to lower latitudes. Moreover, the Arctic experiences substantial winter warming in regions with active pumps. The global annualmean nearsurface air temperature is reduced by only 0.02 K (20212060). Our results cast doubt on the potential of sea ice targeted geoengineering to mitigate climate change.

2.1 Introduction

The declining trend of the Arctic sea ice extent (Comiso, 2012; Kay *et al.*, 2011; Stroeve *et al.*, 2012; Lindsay & Schweiger, 2015), caused mainly by anthropogenic greenhouse-gas emissions (Notz & Stroeve, 2016), is expected to continue. Projections based on climate models foresee a largely ice-free Arctic ocean in late summer around the mid-21st century in the business-as-usual emission scenario (Collins *et al.*, 2013; Jahn, 2018; Niederdrenk & Notz, 2018; Notz & Stroeve, 2018). The replacement of the highly reflective ice cover by the dark ocean has been described as one of the most severe positive feedbacks in the climate system (Manabe & Stouffer, 1980) and contributes to the Arctic warming amplification (Pithan & Mauritsen, 2014).

The Paris Agreement stipulates the reduction of greenhouse-gas emissions to keep global warming well below 2°C (United Nations, 2015; Cornwall, 2015). However, even if all national commitments to curb emissions will be implemented, the 2°C target will likely be exceeded significantly (Rogelj *et al.*, 2016). The discussion around alternative approaches based on climate engineering—the anthropogenic large-scale modification of the Earth's climate to mitigate global warming (Keith, 2001; Bellamy *et al.*, 2017; Talberg *et al.*, 2018)—is highly controversial (Blackstock & Long, 2010; Hamilton, 2013; Givens, 2018). Nevertheless, with the prospect of insufficient emission reductions, the scientific examination of climate engineering strategies appears advisable.

Several climate engineering approaches that focus on the Arctic sea ice cover and the positive ice-albedo feedback have been proposed (Seitz, 2011; Cvijanovic *et al.*, 2015; Mengis *et al.*, 2016; Desch *et al.*, 2017; Field *et al.*, 2018). The Arctic Ice Management (AIM) strategy put forward in Desch *et al.* (2017) (D17 hereafter), which attracted the attention of the scientific community and the media alike (rated within the top 5% of all research output[2]), entails the large-scale employment of wind-driven pumps that spread seawater on the ice surface in the winter months. The sea ice and the snow that is accumulated over it are materials with low thermal conductivity compared to the ocean water. During the freezing season, even a thin layer of sea ice limits the heat flux from the warmer ocean to the cooler atmosphere considerably (Trodahl *et al.*, 2001), reducing the growth of additional sea ice.

---

[2]Altmetric Attention Score of Desch *et al.* (2017): `https://wiley.altmetric.com/details/72217339`

The AIM approach aims to bypass the thermally insulating effect of sea ice, allowing thereby more ice to grow thick enough during winter to withstand the summer melt (Fig. 2.1).
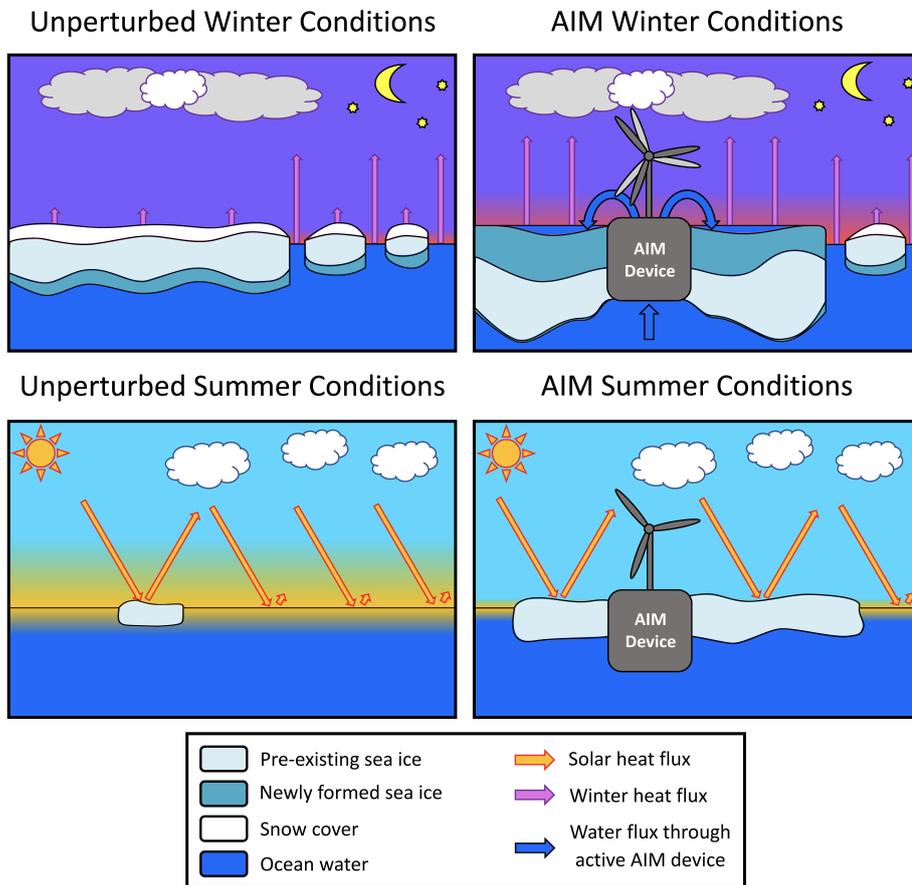
Figure 2.1: Idealized representation of the 21st century sea ice system with and without Arctic Ice Management (AIM). In unperturbed winter conditions (top left) the sea ice and snow act as insulator reducing the heat flux from the warmer ocean to the much colder atmosphere. The sea ice growth takes place mostly at the ice-ocean interface and is relatively slow (dark blue fraction of the ice floes). By summer (bottom left) most of the ice has melted, leading to an ice-free Arctic ocean in the second half of the century and amplifying the warming through the ice-albedo feedback (yellow fraction of ocean). In AIM conditions (top right) ocean water is pumped onto the ice, leading to larger heat flux and rapid ice growth at the surface. More ice withstands the summer melt (bottom right) and increases the surface albedo.

Based on simple thermodynamical arguments and observations from an ice mass balance buoy in the Beaufort Sea, D17 estimate that $\sim$1.4 m of seawater would need to be pumped onto the ice to generate $\sim$1.0 m of extra ice thickness over the course of one winter at a typical location in the

Arctic Ocean. They envisage the deployment of $\sim$10 million devices, each comprising a wind turbine, a pump, a water tank, and a delivery system that distributes the water over an area of $0.1\,\mathrm{km}^2$. D17 calculate that $1\,\mathrm{m}$ extra ice thickness would lead to a shift of the local melt date by $\sim$15 weeks ($3\,\mathrm{weeks}$ per $0.2\,\mathrm{metres}$). They argue that it might be possible to maintain a large part of the usually seasonal ice zone throughout the summer by appropriate annual repositioning and/or reseeding of the AIM array. Considering associated albedo changes, D17 calculate a global annual-mean short-wave radiative cooling by up to $0.14\,\mathrm{Wm}^{-2}$. This is about half of the estimate by Hudson (2011) for the global annual-mean forcing associated with a virtually ice-free Arctic summer ($0.3\,\mathrm{Wm}^{-2}$) and a significant fraction of the current anthropogenic radiative forcing by $\sim$$1\,\mathrm{Wm}^{-2}$.

Considering energy requirements, economical demands, as well as technical challenges, D17 conclude that such a major undertaking seems indeed feasible. However, the question is left open what the quantitative response of the Arctic as well as the global climate system would be. It is also unclear whether the local thermodynamic considerations can be scaled up to the whole Arctic. For example, the large-scale exposure of relatively warm ocean water is expected to generate positive near-surface temperature anomalies. Because the surface turbulent heat fluxes are proportional to the surface temperature gradient (Wallace & Hobbs, 2006; Serreze et $al.$, 2007), increased winter temperatures might induce a negative feedback that dampens the additional ice growth. Complex climate models that simulate the relevant physics, including the general circulations of the atmosphere and the ocean, can provide answers.

Here we use the Alfred Wegener Institute Climate Model (AWI-CM) (Sidorenko et $al.$, 2015; Danilov et $al.$, 2015a; Rackow et $al.$, 2016) to study the efficacy of Arctic Ice Management and the response of the climate system, in the Arctic and beyond. To this end we modify the parameterisation of the surface heat and mass fluxes in ice-covered ocean regions north of the polar circle ($\sim 66.5°\,\mathrm{N}$) such that the effect of the AIM devices is simulated. The modification is activated during the Arctic winter from October 21st to March 21st from 2020 onward. The strength of the modification is modulated with two parameters that affect the large-scale spatial extent and the local efficiency of the pumps. A sensitivity analysis with respect to these parameters is followed by a more detailed analysis based on ensemble simulations ($4\times$unperturbed and $4\times$AIM) with RCP 8.5 scenario forcing until

2100. Moreover, we analyse the effect of an abrupt suspension of AIM in 2030 to test its reversibility.

2.2 Results

2.2.1 Regulating the strength of AIM

The impact of AIM depends on the strength of the modification applied. In the real world, this would depend on the number and spatial distribution of the deployed AIM devices, as well as their efficiency in distributing the ocean water over the surrounding sea ice. To start with, we have performed a simulation where a liquid layer is maintained over the whole ice cover, allowing us to determine an upper bound for the impact of AIM on the ice and on the global climate. This extreme scenario should be regarded as an idealized case to test the response of the climate system to AIM. In this experiment the mean Arctic ice thickness increases almost linearly by $\sim 2.1\,\mathrm{m}$ per year from 2020 to 2030 (the historical 1850–2000 annual-mean value is $\sim 1.8\,\mathrm{m}$). Thereafter the thickness growth slows down until the mean thickness levels off around $65\,\mathrm{m}$ from 2080 onward, corresponding to a pan-Arctic ice volume of $\sim 900 \times 10^3\,\mathrm{km}^3$ (Fig. A.1; right). The ice extent attains values around $15 \times 10^6\,\mathrm{km}^2$ in late winter (February) and $13.5 \times 10^6\,\mathrm{km}^2$ in late summer (September) (Fig. A.1; left). This implies almost a doubling of the late-summer sea ice extent compared to historical conditions (1850–2000). The ice thickness and extent stop growing due to the gradual warming by increasing greenhouse-gas concentrations and, for the same reason, would start to decline beyond 2100 despite AIM.

The near-surface temperature response in this extreme case is profound: Averaged over 2021–2060 north of 66.5° N, the Arctic is colder by $\sim 5.2\,\mathrm{K}$ in September, compared to the 4-member ensemble of unperturbed runs without AIM, but warmer by $\sim 10.6\,\mathrm{K}$ in February when the pumps are active (Fig. A.2; top). The northern middle latitudes (30° N–60° N), however, are warmer by $0.5\,\mathrm{K}$–$1.0\,\mathrm{K}$ throughout the year. This implies that the radiative cooling from the increased albedo is not strong enough to (over-)compensate the effect from the direct Arctic winter warming which is transported to lower latitudes by atmospheric advection and persists there in the ocean mixed layer throughout the year. (The September warming of the northern middle latitudes tends to be present already after a single AIM season

(Fig. A.2; 2020), which is not compatible with the typical time scale for oceanic transport.) The results raise the question whether a more moderate implementation of AIM, where pumps are employed only where they are needed to make the ice thick enough to survive the summer melt, might be better suited to generate an overall cooling. A weaker AIM implementation is also more realistic given that it seems unlikely that the AIM devices would be able to maintain a closed cover of liquid water and also in regard to the number of devices required: The extreme case corresponds to more than 10 times the number of devices envisaged by D17.

We have thus introduced two parameters that affect the large-scale spatial extent and the local efficiency of the AIM devices: The Global Modulation Parameter (GMP) determines an ice thickness threshold beyond which the pumps are deactivated. Thereby the modification is active only in regions with relatively thin ice, where extra ice thickness can reduce the chances of the ice to melt completely over the course of the subsequent summer. In contrast, the Local Modulation Parameter (LMP) determines which fraction of the ice surface in model grid cells with active pumps is covered by water. The LMP represents the spatial density of AIM devices as well as their efficiency to maintain a liquid layer.

To explore the impact of the two parameters, we have conducted 9 simulations from 2020 to 2040 by combining 3 GMP values ($1\,\mathrm{m}$, $2\,\mathrm{m}$, and $3\,\mathrm{m}$) with 3 LMP values (25%, 50%, 75%) (Figs. 2.2 and A.3). Averaged over all 20 years, the March sea-ice extent falls short of the historic level by 1.1–$1.7{\times}10^6\,\mathrm{km}^2$ in any of these settings (Fig. 2.2). One reason for this low sensitivity is that the historical winter sea ice edge is located south of the southern bound of the AIM domain at 66.5° N, except in der Nordic Seas. Furthermore, this reflects that the winter ice edge is largely controlled by large-scale atmospheric (and oceanic) temperatures: if they never fall below the freezing point, no ice can grow, irrespective of AIM. In contrast, the September sea ice extent and the sea ice volume at any time of the year are strongly sensitive to the two parameters, with larger values of the parameters leading to larger extent and volume. The influence exerted by the thickness threshold (the GMP) is stronger than the one by the local density/efficiency (LMP). The LMP has only a minor influence for GMP $= 1\,\mathrm{m}$, where the impact of AIM is generally weak because $1\,\mathrm{m}$ ice thickness is typically not enough to withstand the summer melt. The influence of the LMP grows with increasing GMP.
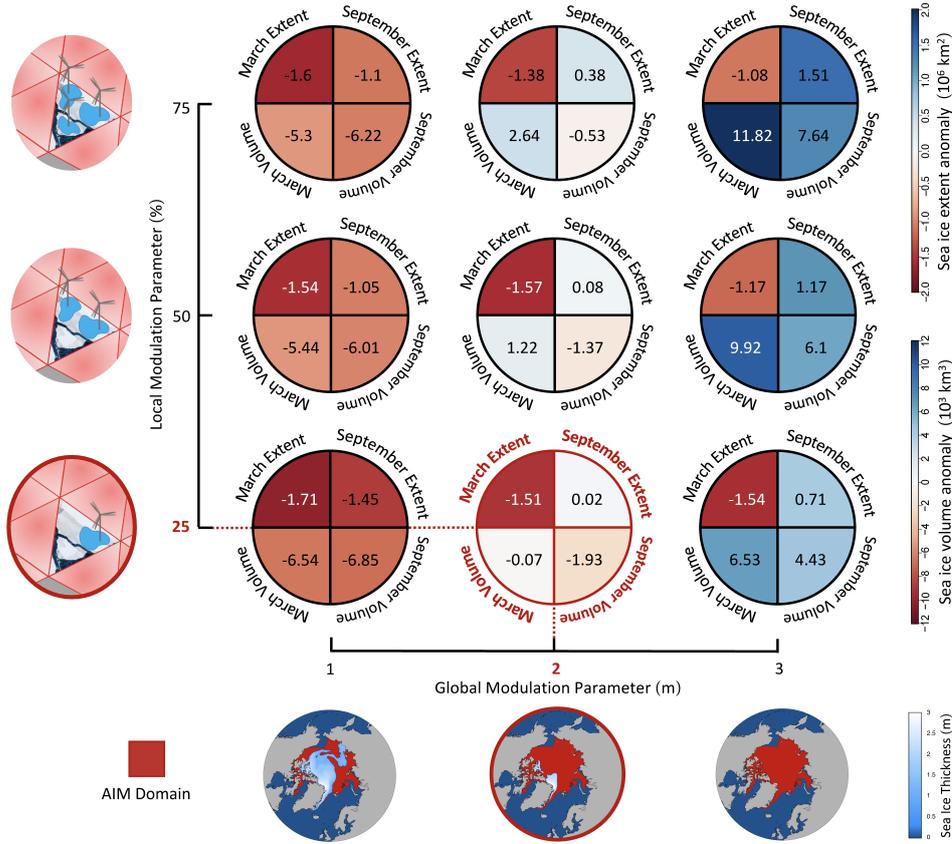
Figure 2.2: The nine pie charts show the sea ice extent and volume anomalies in nine sensitivity simulations (2020–2040) compared to historical conditions (1850–2000) for March and September. The numbers inside the pie charts provide the anomalies in $10^6\,km^2$ for the sea ice extent and in $10^3\,km^3$ for the sea ice volume. Each pie chart corresponds to one combination of the Global Modulation Parameter (GMP; increasing from left to right) and the Local Modulation Parameter (LMP; increasing from bottom to top). The LMP and GMP choice defines the active AIM domain (red area in the GMP maps) and therefore the strength of the AIM in the simulations. The combination GMP = 2m and LMP = 25% (marked in red) is used for the 21st century AIM simulations. AIM = Arctic Ice Management

For the 21st century simulations discussed in the following we have chosen GMP = 2 m and LMP = 25%. This setting restores the summer sea ice extent, which largely determines the ice-albedo feedback, quite accurately to historical levels (Fig. 2.2). Moreover, assuming that a single AIM device covers $\sim 0.1\,\mathrm{km}^2$, averaged over winter 2020–2040 this setting approximately corresponds to 10^6 active devices (Fig. 2.3; bottom), as envisaged by D17.
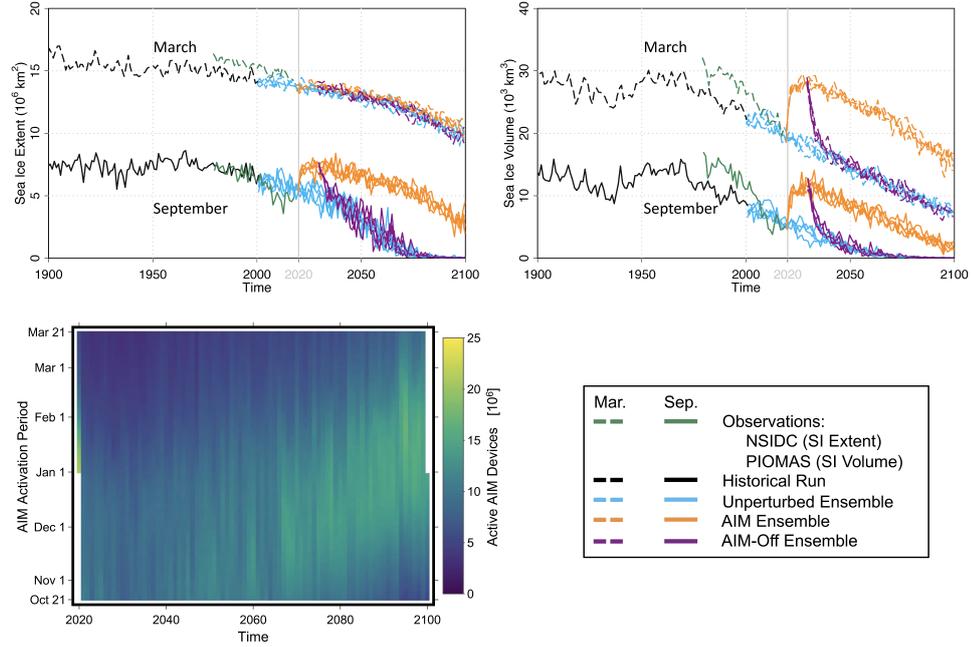
Figure 2.3: Top: Evolution of pan-Arctic sea ice extent (left) and volume (right) in March (upper curves) and September (lower curves). Bottom left: daily number of active devices as function of time of the year (vertical) and year (horizontal) in the AIM simulations (ensemble mean). NSIDC = National Snow and Ice Data Center, SI = Sea Ice, PIOMAS = Pan-Arctic Ice Ocean Modeling and Assimilation System, AIM = Arctic Ice Management.

2.2.2 Arctic sea ice in the 21st century with AIM

The unperturbed simulations without AIM coherently project a virtually ice-free Arctic ocean in late summer after 2060 (Fig. 2.3). The introduction of AIM in 2020 induces a strong and sudden perturbation of the sea ice state. At first a new quasi equilibrium close to historical conditions is reached within a few years. Compared to the unperturbed ensemble, the sea ice volume increases by ∼40% in March and ∼60% in September, and the September extent increases by ∼40%, whereas the March extent is again hardly affected. After the transition phase, however, the declining trend in sea ice volume is similar for both ensembles. For the month of March (September), the declining sea ice volume trend is $-163 \pm 2 \frac{km^3}{year}$ ($-121 \pm 3 \frac{km^3}{year}$) for the control ensemble mean and $-182 \pm 3 \frac{km^3}{year}$ ($-144 \pm 2 \frac{km^3}{year}$) for the AIM ensemble mean (Tab. A.1). Also the September sea ice extent shows a clear declining trend due to the greenhouse-gas induced warming: $-8.3 \pm 0.3 \times 10^4 \frac{km^2}{year}$ for the control ensemble mean and $-6.2 \pm 0.2 \times 10^4 \frac{km^2}{year}$ for the AIM ensemble mean (Tab. A.1). Based on the sea ice extent trends of the two ensemble means, a

virtually ice-free Arctic ocean (sea ice extent $< 1{\times}10^6 km^2$) occurs 66±6 years later with AIM. While the exact delay depends on the parameter considered, overall the Arctic sea ice decline is delayed by roughly 60 years through AIM.

Our approach entails that the number of active AIM devices continuously changes in response to the spatial ice thickness distribution, both between years and within a freezing season (Fig. 2.3, bottom-left). In early 2020, immediately after the AIM activation, the number of active devices (20×10^6) is particularly large because the sea ice thickness is below the thickness threshold (GMP = 2 m) in most places. Until around 2060, the area of ice less than 2 m thick and hence the number of active devices tends to decrease monotonically from 10×10^6 to $\leq 5 \times 10^6$ devices over the course of each freezing season. After 2060 the seasonal maximum is shifted gradually towards the end of the freezing season because the greenhouse-gas induced warming impedes the thickness growth, so that the seasonal ice area growth becomes faster than the seasonal growth of the ice area with thickness ≥2 m. Similarly, the seasonally-averaged number of active devices grows towards the end of the century because the ice area with thickness ≥2 m declines more rapidly than the total ice area.

If AIM would generate unanticipated detrimental effects of any kind, it would be important that the approach is reversible. To test this, we have branched off four additional simulations from the AIM ensemble in 2030 where AIM is turned off. The sea ice extent and volume return to the unperturbed trajectory within a transition period of less than 10 years (Fig. 2.3, purple curves). This is consistent with earlier findings that there is no tipping point associated with Arctic sea ice and the ice-albedo feedback (Tietsche *et al.*, 2011) and suggests that AIM is fully reversible. While this can be regarded as a beneficial property of AIM, it also implies that the array of devices would need to be maintained constantly to stay on a trajectory with delayed Arctic sea ice decline. The rapid loss of the response to geoengineering once it is discontinued seems to be common to geoengineering techniques trying to alter the Earth's albedo (McCusker *et al.*, 2014).

2.2.3 The climate impact of AIM

The increased surface albedo associated with the additional sea ice results in significantly more reflected solar radiation in the Arctic during summer (~5.0 Wm^{-2} at the top of the atmosphere in July north of the polar circle

for 2021–2060; $\sim$6.5 Wm^{-2} for 2061–2100; see also Fig. A.4). Averaged over the globe, the solar radiative forcing due to AIM amounts to $\sim$0.25 Wm^{-2} in July (for 2021–2060 as well as 2061–2100), but only $\sim$0.02 Wm^{-2} for 2021–2060 and $\sim$0.08 Wm^{-2} for 2061–2100 when averaged over the whole year. The latter corresponds to slightly more than half of the estimate by D17, and a quarter of the estimate by Hudson (2011) for a summer ice-free Arctic Ocean.

In the Arctic, AIM leads to a consistent late-summer cooling (Fig. 2.4; top-left; September). Averaged over the area north of the polar circle, September near-surface (2 m) temperatures are reduced by $\sim$1.3 K during the first half of the simulations (2021–2060) and by $\sim$1.4 K during the second half (2061–2100) compared to the unperturbed simulations. The Arctic winter response is more heterogeneous in both space and time (Fig. 2.4; top-left; February). In February, most areas of the Arctic Ocean are cooled by AIM during the first decades (average temperature anomaly over cooling regions is $\sim$-1.1 K), whereas some peripheral seas including the Baffin Bay area and the Kara Sea are subject to additional near-surface warming (average temperature anomaly over warming regions is $\sim$1 K); on average the Arctic is cooled by $\sim$0.3 K. Towards the end of the century the regions with AIM-induced warming expand further into the Arctic Ocean; the average Arctic cooling turns into a warming by $\sim$0.5 K. This adds to the 10.7 K of Arctic February warming in 2061–2100 relative to historical conditions in the unperturbed simulations.

The Arctic temperature response (Fig. 2.4; top-left) is caused mainly by four mechanisms:

1. In winter, the AIM devices maintain a layer of liquid water approximately at the freezing point on the ice surface. This leads to strongly enhanced surface heat fluxes and warm temperature anomalies in areas with active devices. This explains the February warming that expands gradually from the peripheral regions to the central Arctic. Fig. 2.4 shows a clear correspondence between regions with warm temperature anomalies (top-left; February) and the active AIM regions (bottom).

2. Some marginal thin-ice regions of the Arctic ice cover experience winter cooling instead of warming despite active AIM devices, simply because these regions are ice-free in the unperturbed simulations. These regions, including the northern Barents Sea, have gained ice through increased
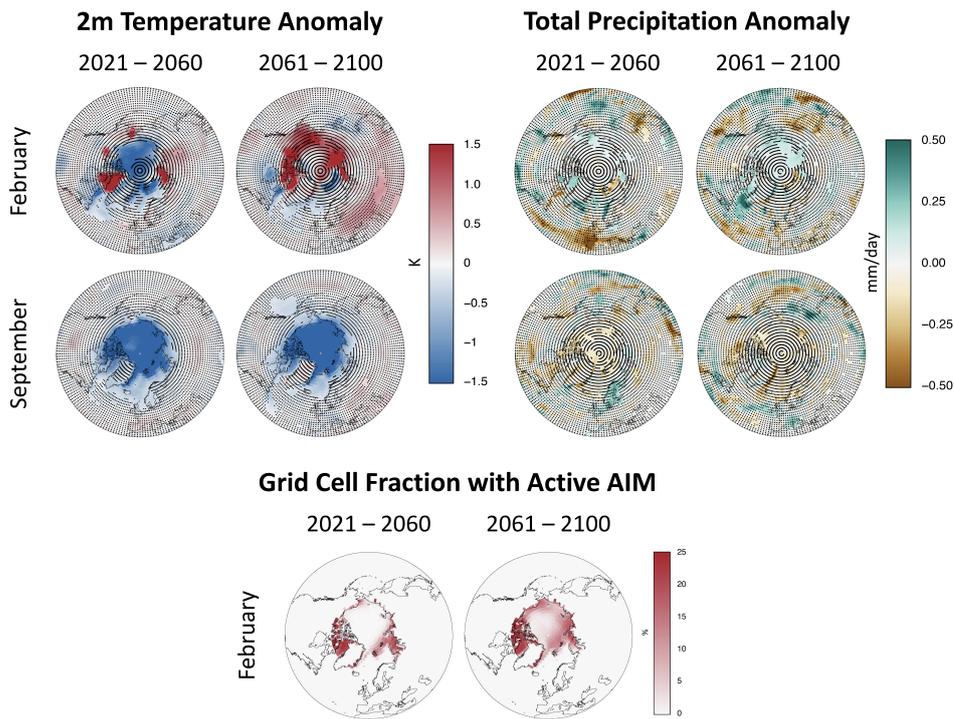
Figure 2.4: Top-left: Near-surface (2 m) temperature anomalies (AIM ensemble-mean minus unperturbed ensemble-mean) for the periods 2021–2060 and 2061–2100. Top-right: as before but for the total precipitation anomaly. Bottom: grid cell fraction with active AIM devices. Note that the 25% upper boundary is defined by the GMP. Stippling indicates local statistical non-significance of the anomaly at the 95% confidence level according to a two-tailed t-test.

advection from AIM-affected upstream regions.

3. Regions with ice thicker than 2 m that were previously subject to AIM encounter weaker winter heat fluxes from the ocean to the cold atmosphere due to the increased ice thickness compared to the unperturbed simulations. Such thick-ice regions without AIM activity thus experience cold temperature anomalies. This explains the February cooling in the central Arctic in 2021–2060.

4. In summer, the additional ice in the AIM ensemble has a direct cooling effect on the atmosphere and surface ocean by latent heat absorption associated with its melting, as well as an indirect cooling effect due to the increased surface albedo and accordingly reduced solar radiative heating. This explains the Arctic summer cooling.

While the impact of AIM on Arctic temperatures is substantial, lower-latitude regions are only weakly affected. The strongest influence outside

the Arctic is exerted on the northern North Atlantic (Fig. 2.4; top-left). In particular the Irminger Sea and the Labrador Sea are affected by enhanced ice export from the Arctic. The additional ice leads to a moderate cooling by up to ∼1 K that prevails year-round throughout the century, but the Atlantic Meridional Overturning Circulation (AMOC) is not sensitive to these changes (Fig. A.5). Outside the northern North Altlantic, the temperature response to AIM is weak and mostly not statistically significant. Some late-century anomalies, like the winter warming in central Eurasia and the summer cooling south of Alaska, appear to be locally significant (Fig. 2.4; top-left), but limited field significance for the middle and low latitudes as a whole suggests that these temperature anomalies might be spurious.

The annual-mean near-surface temperature response of the northern middle latitudes (30° N–60° N) to AIM is close to zero (∼-0.04 K and ∼-0.02 K in the first and second half of the simulations), with minor seasonality. This means that the middle-latitude warming obtained with the extreme-AIM experiment (Fig. A.2) can be prevented with a careful regulation of the interference. However, a significant cooling outside the Arctic (and northern North Atlantic) is still not accomplished. The annual global-mean near-surface warming of ∼1.9 K and ∼3.6 K in the first and second half of the unperturbed simulations is reduced by only ∼0.02 K and ∼0.05 K, despite the intervention in the Arctic ice-albedo feedback.

Finally, a large-scale interference with the climate system can in principle also affect other relevant aspects of climate besides radiation and temperature. The most obvious additional impact of AIM in our simulations is an enhancement of the hydrological cycle and precipitation in regions with warming and moistening due to active devices in winter (Fig. 2.4; top-right; February). We also find a drying across the Arctic Ocean in summer (Fig. 2.4; top-right; September), albeit less significant than the associated cooling. The precipitation response beyond the Arctic is weak, and small regions with locally significant anomalies again appear not to withstand field significance considerations. In general, the large-scale circulation does not respond coherently to AIM in our simulations, despite the modified meridional near-surface temperature gradient. We conclude that the impact of AIM on climate outside the Arctic (and the northern North Atlantic) is generally weak.

2.3 Discussion

This study involves a number of simplifying assumptions and approximations. The AIM implementation neglects peculiarities associated with ice formation at the surface instead of the bottom. Differences in the amount of salt rejected during the freezing and related to the flooding of the snow cover would have implications for the physical properties of the resulting sea ice, including its surface reflectivity and its mechanical behaviour. More generally, the use of a single climate model with necessarily simplified representations for all components of the physical climate system implies that our results are subject to uncertainty.

The difference between our estimate for the global annual-mean solar radiative forcing of AIM ($0.08\,\mathrm{Wm^{-2}}$ for 2061–2100) and the estimates by D17 ($0.14\,\mathrm{Wm^{-2}}$) and Hudson (2011) ($0.3\,\mathrm{Wm^{-2}}$) can have various reasons. The amount of clouds prevailing over the Arctic in summer for instance modulates the impact of changes in surface albedo. However, the Arctic summer cloud cover in our simulations amounts to about 80%, which is in line with the assumptions and observations used in D17 and Hudson (2011). We also do not find a response of the summer cloud cover that would be strong enough to explain the difference (Fig. A.4). Other relevant factors include the assumed or simulated ice surface albedo and how it develops when melt ponds form (which is treated by a diagnostic melt pond scheme in our model) as well as the assumed or simulated sea-ice area difference. In fact the latter might explain why D17 arrive at a higher estimate: They assume that the albedo change would occur over the entire area of the Arctic Ocean ($10^7\,\mathrm{km^2}$), whereas the ice extent anomalies in our simulations amount to roughly half of that area (depending on the year and time of the year; compare Fig. 3 for September).

Another element of uncertainty arises from the way we regulate the strength of AIM: Our implementation implicitly assumes that the deployment and relocation of devices is accomplished so efficiently that the evolving areas with ice thinner than $2\,\mathrm{m}$ are equipped with devices during the whole winter. Our estimate of $\sim 10 \times 10^6$ for the number of required devices, corresponding to the number suggested by D17, should thus be regarded as a lower bound.

Our work does not consider the economic and technical feasibility of the

construction, deployment, and maintenance of the enormous array of AIM devices that would be required. It also does not touch the political and societal dimension associated with such a planetary-scale intervention. Moreover, we do not attempt to provide precise estimates for the impacts of AIM on the climate system. This also holds for possible impacts on permafrost thaw and associated carbon emissions due to the summer cooling and winter warming in Arctic land regions. Rather, our results constitute a first assessment of the efficiency and impacts of AIM from a climate physics perspective. We find evidence that AIM can in principle delay the Arctic sea ice decline by several decades. Yet the cooling of lower latitudes, anticipated as a consequence of the intervention in the ice-albedo feedback, fails to materialise. These results cast doubt on the potential of sea ice targeted geoengineering as a meaningful contribution to mitigate climate change.

2.4 Methods

2.4.1 The AWI climate model

We use the Alfred Wegener Institute Climate Model (Sidorenko *et al.*, 2015; Rackow *et al.*, 2016, AWI-CM) which contributes to the Coupled Model Intercomparison Project Phase 6 (Eyring *et al.*, 2016, CMIP6). For the atmospheric model component ECHAM6 (Stevens *et al.*, 2013) we use the coarse-resolution version with $\sim 1.8°$ grid spacing. For the unstructured-mesh ocean and sea-ice model component FESOM-1.4 (Timmermann *et al.*, 2009) we use the "CORE2" mesh with a resolution of $\sim 25\,\mathrm{km}$ in the Arctic and $\sim 1.27 \times 10^5$ surface nodes globally. Details on the influence of the model resolution of the two model components can be found in Sein *et al.* (2018) and Rackow *et al.* (2019). The sea-ice model (Danilov *et al.*, 2015a) includes an elastic-viscous-plastic (EVP) rheology and a thermodynamical component based on (Parkinson & Washington, 1979), including a prognostic snow layer (Owens & Lemke, 1990). The heat, momentum and mass fluxes at the interface between the ocean (including the sea ice) and the atmosphere are computed within the atmospheric model and exchanged 6-hourly via the OASIS3-MCT coupler. The surface fluxes play a central role in this study because the implementation of AIM in AWI-CM is based on the modification of the surface exchange processes.

2.4.2 AIM implementation

Our implementation of AIM acts on the vertical fluxes of heat, mass and momentum across the ocean/ice-atmosphere interface. When AIM is active, it is assumed that a fraction of the sea ice defined by the Local Modulation Parameter (LMP) is covered by a thin but persistent water layer (PWL). The PWL has the same temperature as the sea surface and is thus close to the freezing point in regions with sea ice. The PWL is continuously restored by the AIM devices as soon as the water freezes or evaporates. The lower boundary of the atmosphere thus corresponds to an increased fraction of open water because the PWL masks the sea ice underneath. The latent and sensible heat fluxes, which represent the turbulent part of the surface heat budget, are calculated for a correspondingly altered open water fraction. Likewise, the surface thermal emissivity and the surface albedo are set to open water values for the PWL-covered part, even though the shortwave radiation plays a minor role during the Arctic winter. Since the PWL covers the sea ice and inhibits ice sublimation, only evaporation from the PWL is allowed in the AIM-affected part of the ice surface. Snow has a temperature of at most 0°C, whereas the PWL is close to the freezing point of salty sea water at ~ -1.8°C. Snow falling into the PWL is thus immediately added to the ice mass without latent heat changes, whereas snow falling into open water is assumed to melt and absorb latent heat.

Formulated as a weighted average of the original fluxes over open water (w) and ice (i), the total heat flux H and the total mass flux M thus depend on the sea ice concentration A_i and the LMP as follows:

$$
\begin{aligned}
H = (1 - \mathrm{LMP} \cdot A_i)\left(Q_S^w + Q_L^w + Q_{LW}^w + Q_{SW}^w\right) + ... \\
... + (\mathrm{LMP} \cdot A_i)\left(Q_S^i + Q_L^i + Q_{LW}^i + Q_{SW}^i\right) + (1 - A_i)\left(P_{snow} \cdot L_f\right)
\end{aligned}
\tag{2.1}
$$

$$
M = (1 - \mathrm{LMP} \cdot A_i)\,E_{evap} + (\mathrm{LMP} \cdot A_i)\,E_{subl} + P_{snow} + P_{rain}
\tag{2.2}
$$

where A_i is the sea ice concentration, Q_{LW} and Q_{SW} are the net longwave and shortwave radiation, P_{rain} and P_{snow} are the liquid and solid precipitation, L_f is the latent heat of fusion of melting ice, Q_S is the sensible heat flux, Q_L is the latent heat flux, and E_{subl} and E_{evap} are the sublimation and evaporation fluxes. The momentum flux calculation remains unchanged.

This modified formulation is used from October 21$^{\text{st}}$ to March 21$^{\text{st}}$ (the Arctic freezing season) in grid cells north of the polar circle ($\sim$ 66.5°N) where the ice thickness is below the Global Modulation Parameter (GMP). No GMP is applied in the extreme AIM simulation.

Except for the described modifications, the sea ice physics remain the same as in the standard FESOM model. The sea ice model does not include a grounding scheme (i.e. no seabed stress is considered), and the ice thickness is not limited by the ocean depth, which has implications for the realism of our extreme AIM experiment in particular in shallow ocean regions.

2.4.3 Experimental setup

Our CMIP-type simulations are designed to test the response of the climate system to AIM in a progressively warming climate. After a 700-year spin-up simulation with constant CMIP6 pre-industrial (1850) forcing, we performed a single simulation until 1999 with transient CMIP6 historical forcing. In 2000, small perturbations were applied to the atmospheric model to generate a 4-members ensemble of simulations that continued until 2014 with CMIP6 historical forcing. Since the new CMIP6 scenario forcings (O'Neill *et al.*, 2016) were not yet available at the time, we used CMIP5 scenario forcing from 2015 onward, accepting a minor discontinuity in the forcing. RCP 8.5 corresponds to the "business-as-usual" scenario where no substantial efforts are implemented to curb greenhouse-gas emissions. The 4 unperturbed simulations were conducted until 2100.

In 2020 a total of 13 simulations was branched off from the unperturbed simulations:

- 1 simulation with extreme AIM, that is, with LMP=100% and no GMP applied, until 2100,

- 9 sensitivity simulations combining 3 GMP values (1 m, 2 m, and 3 m) with 3 LMP values (25%, 50%, 75%) until 2040, one of which (GMP=2 m, LMP=25%) is extended to 2100, and

- 3 additional simulations with GMP=2 m and LMP=25% until 2100, with each member of the resulting 4-member ensemble initialised from one of the 4-member unperturbed ensemble.

2.4.4   Simulated versus observed historical sea ice state

A realistic simulated sea ice state is an important prerequisite for a meaning-ful quantitative assessment of AIM. The Arctic sea ice extent simulated for the period 1979–2017 is in overall agreement with observations in terms of mean value, trend, and inter-annual variability (Fig. 2.3; top-left), although the model seems to slightly underestimate the March sea ice extent and fails to simulate years with particularly low sea ice extent as they occurred in 2007 and 2012. The AWI-CM slightly underestimates the Arctic sea ice volume compared to PIOMAS (Schweiger *et al.*, 2011) during the period 1979–2005. The more recent volume values are better represented. Nevertheless, the model captures the declining sea ice volume trend (Fig. 2.3; top-right). The spatial thickness distribution are also realistically simulated, with thicker ice north of Greenland and the Canadian Archipelago compared to the rest of the Arctic (Fig. A.6). The modelled sea ice thickness can be visually com-pared to sea ice thickness satellite retrievals and reanalysis products (Wang *et al.*, 2016; Ricker *et al.*, 2017).

3. Bright prospects for Arctic sea ice prediction on subseasonal time scales[1]

Abstract

With retreating sea ice and increasing human activities in the Arctic come a growing need for reliable sea ice forecasts up to months ahead. We exploit the subseasonaltoseasonal prediction database and provide the first thorough assessment of the skill of operational forecast systems in predicting the location of the Arctic sea ice edge on these time scales. We find large differences in skill between the systems, with some showing a lack of predictive skill even at short weather time scales and the best producing skillful forecasts more than 1.5 months ahead. This highlights that the area of subseasonal prediction in the Arctic is in an early stage but also that the prospects are bright, especially for late summer forecasts. To fully exploit this potential, it is argued that it will be imperative to reduce systematic model errors and develop advanced data assimilation capacity.

3.1 Introduction

The observed rapid retreat of Arctic sea ice and the prospect of a virtually ice-free Arctic Ocean in late summer by the middle of this century (Collins *et al.*, 2013; Wang & Overland, 2009; Overland & Wang, 2013; Stroeve *et al.*, 2007) have fueled socio-economic interests in the region (Emmerson & Lahn, 2012; Stephenson *et al.*, 2011). As a consequence there is a growing demand for reliable predictions of Arctic weather and sea ice across a wide range of time scales, to reduce the risks that come with enhanced activities in the high north (Jung *et al.*, 2016).

Much of what is known about the skill of existing systems in predicting Arctic sea ice is based on the Sea Ice Outlook (SIO) (Stroeve *et al.*, 2014)—an effort of the international research community that since 2008 has been aiming to build and evaluate seasonal sea ice prediction capabilities. So far, SIO dynamical predictions have shown limited skill, with simple statistical forecasts being of comparable quality (Stroeve *et al.*, 2014; Blanchard-Wrigglesworth *et al.*, 2017). On the other hand, perfect-model studies suggest significant potential predictability at seasonal time scales (Tietsche *et al.*, 2014; Goessling *et al.*, 2016a; Guemas *et al.*, 2016), indicating that there is scope for major improvements. On much shorter weather time scales (up to ∼10 days ahead) high-resolution forecast systems are increasingly being used by operational ice services (Carrieres *et al.*, 2017; World Meteorological Organization, 2017), and recently research has started into exploring the predictability of sea ice on these shorter time scales (e.g. Mohammadi-Aragh *et al.* (2018)).

The potential for skillful predictions of Arctic sea ice on subseasonal-to-seasonal time scales has improved considerably through recent developments. Recognising the urgent need for a better representation of the sea ice-ocean system, forecast centres are moving towards using fully coupled models (Smith *et al.*, 2015). This also holds for shorter weather time scales, where features such as the location of the sea ice edge can feed back significantly to the atmosphere, thereby influencing the further evolution of the coupled system (Jung *et al.*, 2016). This development towards using coupled models is reflected by the fact that six out of eleven forecast systems contributing to the recently established Subseasonal to Seasonal (S2S) Prediction database (Vitart *et al.*, 2012, 2016) include dynamical sea ice compo-

nents. These dynamical models replace relatively crude schemes where the sea ice state is simply persisted from its initial state and/or relaxed towards climatological conditions. In fact, the S2S database constitutes an unprecedented opportunity for a thorough assessment of state-of-the-art operational predictions of Arctic sea ice on subseasonal timescales. Numerous reforecasts are available for each of the contributing systems, which is critical for making robust statements about the skill and the associated uncertainties. Furthermore, the forecasts cover the whole annual cycle, allowing to determine seasonal variations in skill. To our knowledge, this study represents the first assessment of these systems in the Arctic, showing that the field of subseasonal prediction of Arctic sea ice is in a early stage, but also highlighting that prospects for skillful predictions are bright.

3.2 Data

The ensemble forecasts analysed here have been obtained from the database of the Subseasonal-to-Seasonal Prediction (S2S) project. Here we consider only those six systems that include a sea ice model coupled to an atmospheric and ocean model, thereby producing actual dynamical sea ice forecasts. The only exception is the older ECMWF forecast system (ECMWF Pres.) where the sea ice state is persisted for the first 15 days of the forecast and then relaxed towards climatology. Archiving of real-time ensemble forecasts in the S2S database started in January 2015 only. However, corresponding reforecasts are available approximately for the previous two decades. The S2S forecast systems exhibit different forecast lengths, initialisation frequencies, ensemble sizes, data assimilation methods and model physics (Tab. B.1). Despite their differences, however, some forecast centers also share some of the same model components, typically the ocean or sea ice model, including the extreme case of UKMO and KMA which share the same forecasting system altogether. Differences in ensemble size and initialisation frequency exist between real-time forecasts and the corresponding reforecasts. The initialization strategy also varies among the systems: some feature a balanced assimilation among sea ice,ocean and atmospheric components (EMCWF, UKMO, KMA, NCEP), in contrast MF and CMA adopt a two tier initialization strategy. To ensure a sufficiently large sample size, while allowing comparability between the systems, our analysis is focused on the common

reforecast period 1999–2010. The sea ice concentration fields from the S2S database are provided on a 1.5°×1.5° longitude-latitude grid, although the sea ice models run are at higher resolution (from 0.25°to 1°).

The verification is carried out against daily sea ice concentration data from passive microwave (PMW) satellite measurements. As for the forecast data, we use the 15% sea ice concentration contour to determine the location of the ice edge. The main observational product used here is the Global sea ice Concentration data record (OSI-SAF, 2016). Discrepancies between true and observed ice edge locations are mainly caused by the presence of summer melt ponds over the sea ice. These are interpreted as open water by PMW sensors (Kwok, 2002; Notz, 2014) and cause a northward shift of the ice edge (Comiso & Nishio, 2008). However, since most of the forecast centers also assimilate PMW measurements, we expect this systematic error to be propagated also to the forecasts and to have a limited impact on our analysis.

3.3   Methods

We apply the recently introduced Spatial Probability Score (SPS; Goessling & Jung (2018)) as verification metric, which can be regarded as the extension of the Integrated Ice Edge Error (IIEE; Goessling *et al.* (2016a)) to probabilistic ice edge forecasts. These metrics are specifically designed to capture the accuracy of the forecasted ice edge and to overcome the limitations of more widely used metrics such as the difference in pan-Arctic sea ice extent or area. The latter only evaluate the total extent of the ice cover, but fail to provide useful information about its spatial distribution. In contrast, the SPS and the IIEE account not only for differences in total sea ice extent but also for ice that is forecast at a wrong location.

The decomposition of the IIEE for the ensemble-median ice edge into Overestimation (O) and Underestimation (U) or, alternatively, Absolute Extent Error (AEE) and Misplacement Error (ME) (Goessling *et al.*, 2016a), adds information to the SPS and provides insights into the origin of forecast errors. O is the spatial integral of all areas where the forecast sea ice concentration is above 15% but the observed sea ice concentration is below 15%; U is the spatial integral of all areas where the forecast sea ice concentration is below 15% but the observed sea ice concentration is above 15%. The AEE

component represents the total difference in sea ice extent between forecast and observation, while the ME component accounts for sea ice that is forecast at a wrong location. A more extensive description of the verification metrics can be found in Sec. B.2.

The computation of verification scores is conducted on a per-grid-cell basis. Therefore it is necessary to remap either the forecast data or the observations (or both) to a common grid and to investigate the impact of the forecasts and observation resolution on our results. In the analysis, the observational data were remapped by first-order conservative remapping to the relatively coarse-resolution forecast data. Further details on the role of resolution in observations and forecasts can be found in Sec. B.3. Only grid cells that are classified as ocean (including sea ice) in all models and in the observations were used (see the resulting land-mask in Fig. B.4. Employing a common conservative land-mask guarantees an unbiased comparison of the skill of different forecast systems.

A meaningful assessment of the forecast skill requires the introduction of observation-based benchmarks based on the same metric employed for measuring the forecast error. If the forecast error is lower than that of a benchmark, the dynamical forecasting system has some predictive skill. Otherwise, the observational record can be used to build a better forecast. We have followed two strategies to construct a meaningful benchmark. Firstly, we defined a climatological benchmark forecast as the 10-member ensemble of states observed at the same time of the year during those 10 years preceding the respective forecast target time. Secondly, we defined a persistence benchmark based on the observed sea ice conditions one month before the forecast target time (Blanchard-Wrigglesworth *et al.*, 2010). The climatological benchmark is more restrictive than the persistence benchmark for most of the year (see Sec. B.4 and Fig. B.1) and is therefore used to assess the skills of the S2S systems.

3.4 Results

3.4.1 Annual-mean sea ice forecast skill

The annual-mean skill of different forecasts in predicting the Arctic sea ice edge can be inferred from Fig. 3.1. The most striking feature is that the forecast skill varies substantially across the different systems. Compared to

the climatological benchmark, the CMA and MF systems do not show any predictive skill, even at initialization time. On the other hand, the ECMWF system shows predictive skill all the way to a lead time of 45 days. The other systems (KMA, NCEP and UKMO) are comparable to ECMWF for short lead times; the error growth is larger, however, leading to a faster loss of predictive skill.

The wide range of error growth rates among the different models is in stark contrast to what can be found for predictions of atmospheric fields, which are much more similar in terms of skill (Jung & Matsueda, 2016). This highlights the fact that the field of sea ice prediction with weather and climate models is still in its infancy.
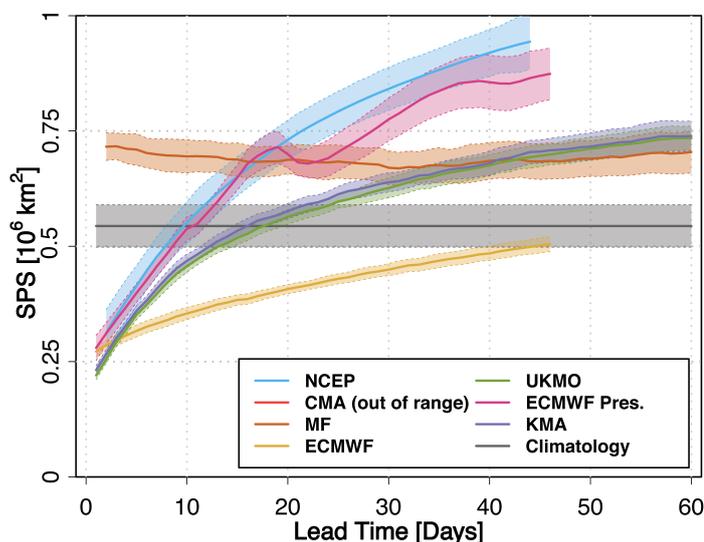

Figure 3.1: Annual-mean skill in terms of the Spatial Probability Score (SPS) of the different forecast systems (colored-solid lines) and the climatological benchmark forecast (gray-solid line) in predicting the Arctic sea ice edge as a function of lead time. Results have been averaged over the common reforecast period 1999–2010. Predictions with SPS values smaller than the climatological value ($\approx 0.55 \cdot 10^6 \, km^2$) can be considered skillful. The shading and dashed lines indicate $\sim 95\%$ confidence intervals, based on standard errors obtained from the twelve individual annual means. Note that the CMA forecast system is not depicted given that its large errors lie outside of the range shown. ECMWF Pres. is based on the predecessor ECMWF system, the main difference being that sea ice was not simulated dynamically but prescribed based on a combination of persistence and climatology. SPS = Spatial Probability Score; S2S = Subseasonal to Seasonal; NCEP = National Centers for Environmental Prediction; CMA = China Meteorological Administration; MF = Météo-France; ECMWF = European Centre for Medium-Range Weather Forecasts; UKMO = UK Met Office; KMA = Korea Meteorological Administration.

Although the skill of ECMWF, KMA, NCEP and UKMO at initial time is much better than that of MF and CMA, initial errors are still quite large (half the values of the climatological benchmark; Fig. 3.1). Given that, based on satellite data, the sea ice conditions should be reasonably well known at the time of the initialization, the large initial errors suggest that there is still substantial scope for improving the data assimilation procedure and thereby the prediction skill of subseasonal forecast systems.

The skill of the UKMO and KMA systems is almost identical (Fig. 3.1) because of the same system shared. However, given that they represent independent forecast realizations (ensemble members) of the chaotic climate system, their agreement demonstrates that the data available in the S2S database allow to draw robust conclusions about the skill of sea ice forecasts. Furthermore, noting that UKMO ensemble size is larger than KMA (Tab. B.1), the slightly higher skill of UKMO compared to KMA suggests that ensemble size matters to improve sea ice edge predictions.

3.4.2 Seasonal variations in forecast skill and origins of error

The results discussed so far were based on annually-averaged values. However, since high latitudes experience very different physical conditions at different times of the year, it appears likely that the predictability of Arctic sea ice is seasonally dependent. In this section, this seasonality will be further explored.

Despite the specific biases affecting each system, a general feature of the SPS, including the climatological benchmark, is a pronounced seasonal cycle with two peaks at the end of the winter and summer seasons (Fig. 3.2). This pattern can be explained by a corresponding seasonality of the ice edge length, which reaches its maxima in late winter and in summer. In general, a longer edge simply implies on average a larger area where forecast and observations can disagree.

The ECMWF system achieves the largest skill in late summer, when actual predictions remain for all the lead times much better than climatological forecasts, which exhibits particularly low skill in this period (Fig. 3.2, top left). A possible explanation for this is that around September the uncertainty in the ice-edge location is the largest due to higher mobility of the ice. However, the ECMWF forecast system is able to capture a relatively
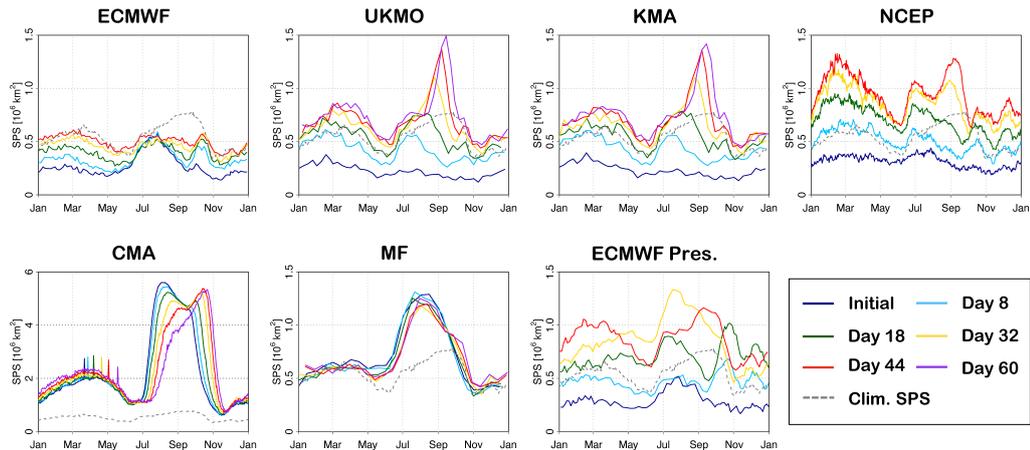
Figure 3.2: Skill in terms of the SPS of individual forecast systems in predicting the Arctic sea ice edge as a function of the time of the year (target date) and for six different lead times (see legend). Results have been averaged over the common reforecast period 1999–2010. Note that Day 60 is missing for NCEP and ECMWF (both versions) due to their shorter lead time ranges, and that *Initial Time* corresponds to Day 1 for all systems except NCEP and MF where it corresponds to Day 2 for technical reasons. SPS = Spatial Probability Score; S2S = Subseasonal to Seasonal; NCEP = National Centers for Environmental Prediction; CMA = China Meteorological Administration; MF = Météo-France; ECMWF = European Centre for Medium-Range Weather Forecasts; UKMO = UK Met Office; KMA = Korea Meteorological Administration.

large fraction of that variability and therefore the forecast error is not larger around September than at other times of the year. Lower relative skill is found from October through July; during this time of the year only short-term forecasts out to ∼18 days achieve meaningful skill compared to the climatological benchmark.

The error components provide further insights into the performance of the ECMWF forecast system. An evident feature is a peak in SPS in July for short lead times (Initial, Day 8 and Day 18) (Fig. 3.2, ECMWF). This reflects a less accurate initialization of the ice edge compared to the rest of the year. The O,U error decomposition (Fig. B.2) reveals that the peak is associated with a development of a substantial model bias: The initial position of the ice edge is systematically underestimated (O ≈ 0% and U ≈ 100%) from July to October.

Interestingly, the forecasts less accurately initialized in July produce comparably skillful long-range (day 45) predictions for late summer, with an approximate balance between O and U (O ≈ 40% and U ≈ 60%, Fig. B.2) and the ME dominating over the AEE (ME ≈ 70% and AEE ≈ 30%, Fig. B.3). A

possible reason for this apparent contradiction is that the skill in late September, which marks the beginning of the freezing season, is related to sources of predictability residing in components of the climate system other than the sea ice. For example, the heat content stored in the surface ocean could influence the sea ice edge position in the early freezing season (Blanchard-Wrigglesworth *et al.* (2010); Sec. B.4). The underestimation of the initial ice edge in the ECMWF system continues until late September, affecting the forecasts at longer lead times in October. The striking transition at the beginning of the freezing season, when the underestimation and the absolute extent error components start to dominate, hints at a delayed onset of the ice growth season in the ECMWF system.

A similar seasonal cycle as for ECMWF can be found for UKMO, KMA and NCEP, at least for forecasts out to 8–18 days, which show still some skill. For longer lead times (beyond day 18), UKMO and KMA show a rapid error growth in August and September. The decomposition of the forecast error reveals that this deterioration of skill is associated with the development of a substantial model bias that is reflected by an underestimation of the integrated Arctic sea ice extent ($O \approx 10\%$ and $U \approx 90\%$, Fig. B.2, KMA and UKMO). The NCEP system exhibit notable differences in how the initially similar imbalances evolve with lead time (Fig. B.2, NCEP). In particular, the dominance of overestimation in January and February increases, and an initially balanced state in August and September turns overestimation-dominated with lead time, pointing to positive model biases for sea ice extent during these months. In contrast, a rapid transition from overestimation-dominated to underestimation-dominated errors around the end of September hints at a delayed onset of the ice growth season in the model, similar as in the ECMWF system.

The CMA system, which is outperformed by the climatological benchmark for all lead times and times of the year, exhibits particularly large errors from August to October (Fig. 3.2, CMA). From July to September the skill decreases (i.e., the SPS increases) with lead time, implying that very large initial errors during this part of the year are amended over the course of the forecast model integration towards a less unrealistic state. Furthermore, the CMA system considerably overestimates the Arctic sea ice extent from November to June, and underestimates the extent even more strongly from July to October (Fig. B.2, CMA). Moreover, the CMA system features a series of negative SPS spikes in spring; the cause of these can be tracked

down to a single forecast bust associated with an erroneous initialisation on 25 March 2007.

The MF system is approximately as skillful as the climatological benchmark from October to April, with only a weak dependence on lead time (Fig. 3.2, MF). During the melting season from May to September, however, the MF system is less skillful and exhibits large initial errors that are slightly amended with growing lead time. Errors in long-term prediction in September are dominated by an underestimation of the pan-Arctic sea ice cover, whereas biases play a minor role in the MF system at other times of the year. This suggests that a more accurate initialisation of the MF system might already be sufficient to improve ice-edge forecasts of this system considerably.

3.4.3 The benefit of using a more realistic representation of sea ice and ocean

ECMWF updated its operational forecast system in November 2016. Until then, sea ice conditions were determined based on the persistence of the initial conditions for the first 15 forecast days, followed by a relaxation towards average sea ice conditions observed during the five years preceding the forecast target time (ECMWF Pres.). The change to a more advanced approach, in which sea ice dynamics and thermodynamics are explicitly represented by a sea ice model, provides a unique opportunity to study the impact of this critical development of the forecast system. Note that the system update also included an increase of the ocean model resolution from 1°to 0.25°. For our assessment we exploit the fact that reforecasts for 1999–2010 are available for both versions of the ECMWF system. Figure B.4 illustrates recent forecasts from the two ECMWF system versions in comparison with the observed sea ice edge derived from different passive-microwave products (OSI-SAF, 2016; Spreen *et al.*, 2008).

The accuracy of the ice-edge location in the initial conditions is similar for the two versions of the ECMWF system; with increasing lead time, however, the version with explicit sea ice physics included quickly outperforms the older version with simple sea ice treatment (Figs. 3.1 and 3.2). This highlights that investments in forecast system development can lead to major advances in predictive skill.

Not surprisingly, using persistence, even for short lead times, leads to

an overestimation of sea ice during the melting season from April to August and an underestimation during the growing season from October to February (Fig. B.2, ECMWF Pres., dark and light blue lines). Around day 18 of the forecasts, the older version of the ECMWF system exhibits an intermittent increase in skill that is a result of the gradual transition from initial-state persistence towards average conditions of previous years (Fig. 3.1). In fact, the temporary decrease of the SPS from day 19 to day 22 suggests that the older version could have benefited from an earlier transition towards climatological sea ice fields.

3.4.4 Case study: the summer of 2007

Some of our main results can be further illustrated by considering subseasonal sea ice forecasts for the exceptional summer of September 2007, which was the first in a series of summers with anomalously low Arctic sea ice extent. Not surprisingly, the climatological forecast clearly overestimates the ice extent in large parts of the Arctic (Fig. 3.3). The ECMWF system clearly captures the observed sea ice edge in its 30-days forecast. The ECMWF ensemble spread appears reasonable, with probabilities transitioning smoothly from 0 to 1 along the observed ice edge. This indicates that the ensemble is reliable, that is, neither under- nor over-dispersive. In contrast, the NCEP forecast, although clearly more skillful than the climatology, is overconfident regarding the ice edge location, with probabilities transitioning sharply from 0 to 1 in disagreement with observed ice edge. The UKMO and KMA systems produce very similar forecasts, including a region at about 170°W where the amount of sea ice is strongly underestimated, also confirming the similarity of the systems. The CMA model is a clear outlier in the sense that initialization and model errors lead to the complete absence of Arctic sea ice during this time of the year. The MF forecast is characterised mostly by overestimation of the ice extent in the Siberian sector, combined with an underestimation along eastern Greenland. This misplacement suggests that the MF system does not capture the particularly high sea ice transport trough Fram Strait which occurred in summer 2007. In this specific year, the persistence benchmark provides a better representation of the September ice edge than other empirical schemes based on the climatological sea ice state (ECMWF Pres. and the climatological benchmark forecast). This suggests that the use of the climatological benchmark has particularly pronounced drawbacks in unusual

years such as 2007, which are more common in a rapidly changing climate.
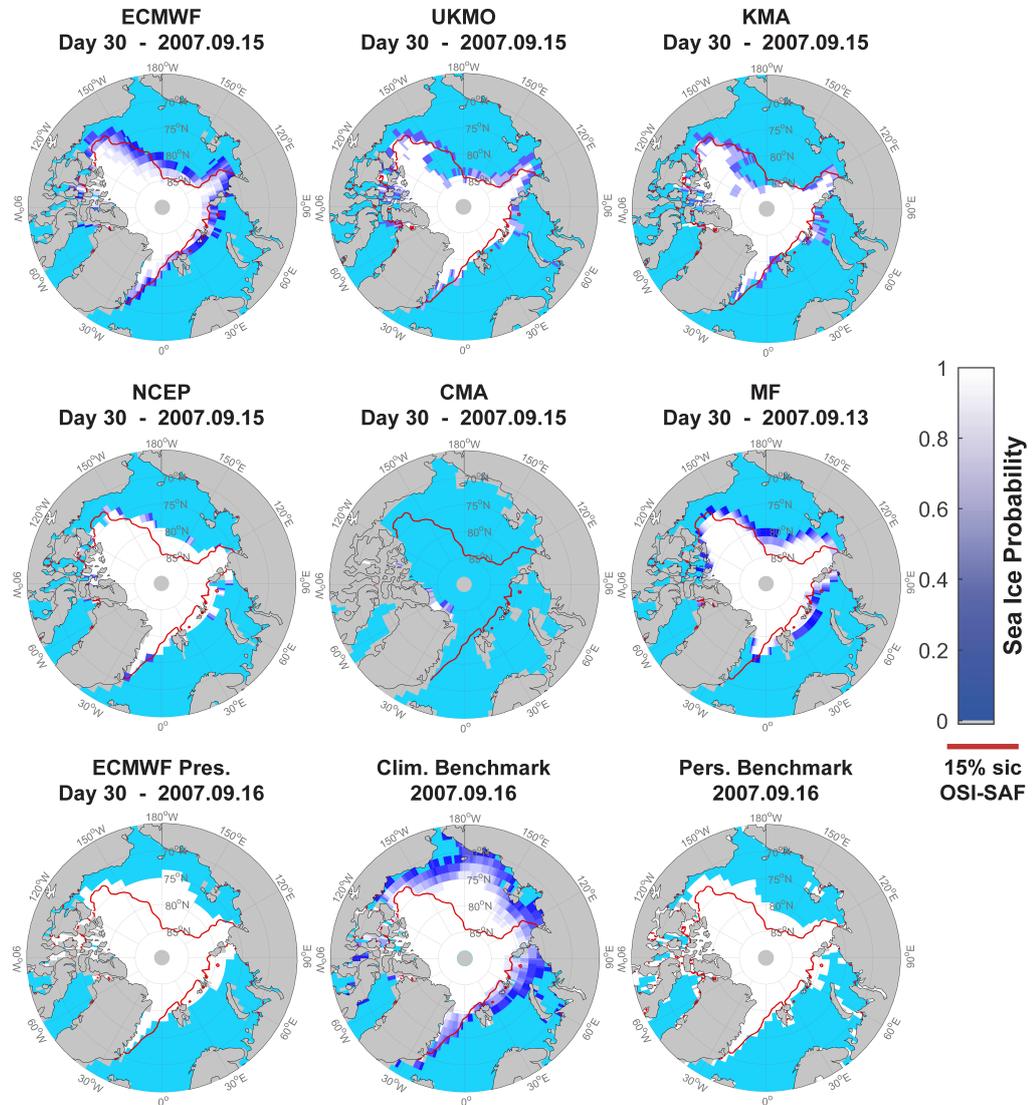
Figure 3.3: 30-day forecasts for 15 September 2007 of the sea ice probability (probability that sea ice concentration exceeds 15%) as obtained from different forecast systems and from climatological and persistence benchmarks. The observed sea ice edge (15% contour of OSI-SAF sea ice concentration) is also shown (red contour). ECMWF = European Centre for Medium-Range Weather Forecasts; UKMO = UK Met Office; KMA = Korea Meteorological Administration; NCEP = National Centers for Environmental Prediction; CMA = China Meteorological Administration; MF = Météo-France.

3.5 Discussion

This paper provides the first overview of the subseasonal skill of state-of-the-art coupled forecast systems in predicting the sea ice edge in the Arctic. By exploiting the recently established S2S database, we find a surprisingly large range of skills with some of the systems showing no skill at all, even at short weather time scales, and the best system producing skillful forecasts up to 45 days in advance. The fact that prediction skill is largest in late summer suggests that useful long-range forecasts can be provided to stakeholders during a time of the year when marine operations peak.

Our analysis of error components has revealed that seasonally dependent model biases play a critical role. This calls for dedicated efforts to improve the realism of coupled models in the Arctic, with the ultimate aim of reducing systematic model errors. Bias correction could be a means to improve real-time forecasts. In fact, a method specifically designed to bias-correct ice-edge forecasts has been recently proposed (Director *et al.*, 2017), and the reforecasts needed for bias correction are available in the S2S database. However, the size of the biases in some of the models, which are comparable in size or even larger than the anomalies one would like to predict, suggests that non-linearity may be an issue.

The large differences in the accuracy of the initial conditions for sea ice between the systems is related to the way how the forecasts are initialized, that is, the way how observations are assimilated into the coupled models. A major difference between the CMA and MF systems and the other (more skillful) systems is that the former two systems do not directly assimilate any sea ice observations into their models, unlike the other systems that assimilate sea ice concentration. In principle, one could have expected to see some skill also for the CMA and MF systems because (i) they do assimilate other ocean variables that affect the sea ice, in particular sea-surface temperature (SST), and (ii) the evolution of the atmosphere, which largely drives sea ice anomalies, is constrained through the assimilation of atmospheric observations. However, our results indicate that these aspects are not sufficient to generate realistic sea ice initial states, and that direct assimilation of sea ice observations is required.

Even the systems with a more accurate initialisation of sea ice (ECMWF, UKMO, KMA, and NCEP) exhibit considerable ice-edge initial errors that

amount to about half of the error of the climatological benchmark. This agrees well with the assessments of the Arctic sea ice cover in reanalyses by Chevallier *et al.* (2016) and Uotila *et al.* (2018), who found a substantial spread in the sea ice edge position between reanalyses, particularly in late summer. Several mechanisms could contribute to the initial error: one is that adjustments of sea ice concentrations based on other assimilated variables (in particular SST) to obtain more consistent states introduce inaccuracies in the ice-edge location. Constraints related to delays in the availability of observational sea ice products might also contribute to the initial errors, although it is not obvious whether such constraints applying to real-time operations are also an issue for the reforecasts.

We conclude that the accuracy of sea ice initial states needs further research and will be critical to advance the field of Arctic sea ice forecasting on subseasonal time scales. While for short-range summer predictions (below 10 days) or subseasonal winter predictions a correct initialization of the sea ice concentration field might be sufficient to achieve skillful forecasts of the ice edge, for longer timescales the role of the sea ice thickness initialization will be crucial, especially during the melting season. In this regard, new satellite observational products have the potential to improve sea ice initial conditions considerably. Of particular interest are, for example, sea ice thickness observations from multiple instruments, with a proven potential to help constrain sea ice initial states (Mu *et al.*, 2017; Day *et al.*, 2014).

The sea ice prediction is a central element of major international efforts such as the Polar Prediction Project along with its flagship activity, the Year of Polar Prediction (Jung *et al.*, 2016), suggesting that there is an opportunity for resource mobilization and international coordination that promises imminent progress. This factors, and the already achieved progress documented by our analysis, indicate that the prospects for subseasonal prediction of Arctic sea ice are bright.

4. Predictability of Antarctic sea ice edge on subseasonal time scales[1]

Abstract

Coupled subseasonal forecast systems with dynamical sea ice have the potential of providing important predictive information in polar regions. Here, we evaluate the ability of operational ensemble prediction systems to predict the location of the sea ice edge in Antarctica. Compared to the Arctic, Antarctica shows on average a 30% lower skill, with only one system remaining more skillful than a climatological benchmark up to ∼30 days ahead. Skill tends to be highest in the west Antarctic sector during the early freezing season. Most of the systems tend to overestimate the sea ice edge extent and fail to capture the onset of the melting season. All the forecast systems exhibit large initial errors. We conclude that subseasonal sea ice predictions could provide marginal support for decisionmaking only in selected seasons and regions of the Southern Ocean. However, major progress is possible through investments in model development, forecast initialization and calibration.

4.1 Introduction

Reliable predictions of the sea ice edge location are becoming increasingly important to ensure the safety of human activities at both poles. Furthermore, providing skillful predictions has been recognized as an important scientific challenge that will need to be addressed in the coming years (Alley *et al.*, 2019). Previous efforts of the research community have focused mostly on the Arctic, partly due to the higher economic interests that are at stake and due to its proximity to highly-populated regions. While the number of stakeholders that requires sea ice predictions in the Arctic is relatively large and ranges from shipping companies to tourism (Stephenson *et al.*, 2011; Emmerson & Lahn, 2012), Antarctic sea ice predictions in the past were relevant mostly for logistical aspects related to research activities. However, in recent years the tourism industry is flourishing also around Antarctica (Eijgelaar *et al.*, 2010), and the presence of the fishing industry in the Southern Ocean is also expected to increase (Cheung *et al.*, 2010; Smetacek & Nicol, 2015), calling for reliable Antarctic sea ice forecasts to manage the risks that come with enhanced activities.

Sea ice forecasting is not only relevant at short "weather" timescales (forecasts up to 10 days ahead), but also at subseasonal and seasonal timescales (forecasts from weeks to months ahead). The work by Chen & Yuan (2004) is one of the first attempts at providing seasonal predictions of the Antarctic sea ice cover using a statistical approach. Holland *et al.* (2013) evaluate the mechanisms of Antarctic sea ice predictability. More recently, Ordoñez *et al.* (2018) compared sea ice predictability between the Arctic and Antarctic. Both these studies are based on climate models as research tools. The systematic investigation of operational sea ice prediction systems, with the assimilation of the observed sea ice state and possibly ensemble-based, is still at a very early stage.

While the Sea Ice Outlook (Stroeve *et al.*, 2014; Blanchard-Wrigglesworth *et al.*, 2017) has established a framework to build and evaluate Arctic late-summer sea ice prediction capabilities in 2008, a similar exercise for the Antarctic region, targeting the February sea ice minimum (SIPN South— 2017–2019), has been initiated only very recently (Massonnet *et al.*, 2018, 2019), that is almost ten years later. In fact, the international scientific community has recognized the need to advance the field of sea ice prediction

at both poles simultaneously (Jung *et al.*, 2016; Goessling *et al.*, 2016c). In this sense, the present study contributes to closing an important knowledge gap.

The recently established database of the Subseasonal to Seasonal (S2S) Prediction Project (Vitart *et al.*, 2012, 2016) has proven to be valuable for evaluating the predictive skill of operational S2S ensemble forecast systems in the Arctic (Zampieri *et al.*, 2018; Wayand *et al.*, 2019). The availability of comprehensive sets of both reforecasts and real-time forecasts allows for a robust assessment of the forecast skill over a relatively long time period (> 10 years), covering the whole seasonal cycle. Here, we extend the analysis by Zampieri *et al.* (2018) for the Arctic, to Antarctica, addressing the two following guiding questions:

- Are fully coupled forecasting systems in the Antarctic better than observation-based benchmark forecasts in predicting the sea ice edge?

- Does the predictive skill of dynamical forecast systems differ between the two hemispheres?

Thereby, the goal is to establish a reference against which future progress in Antarctic sea ice prediction can be quantified. To our knowledge, this study is the first assessment of the S2S forecast systems in the Antarctic, especially when it comes to focusing on the sea ice edge position, which is a crucial variable for navigation and for planning human activities in the Southern Ocean.

4.2 Data and Methods

The sea ice forecasts are verified against observations using a verification metric suitable for quantifying the accuracy of the sea ice edge location. The resulting forecast error is compared to that of observation-based benchmark forecasts to assess the predictive skills of the forecast systems and to understand associated shortcomings and model biases. This section briefly describes the main features of forecasts, observations, verification metrics and benchmark forecasts used in this study. A more detailed description of the methods, forecasts and observations can be found in the work of Zampieri *et al.* (2018), including its supplements.

4.2.1 Forecasts and observations

The ensemble sea ice forecasts considered here belong to the S2S Database (Vitart *et al.*, 2016), which provides sea ice concentration as a standard output variable. Here we focus on the six forecasting systems that employ a dynamical sea ice model in their coupled model: the National Centers for Environmental Prediction (NCEP), China Meteorological Administration (CMA), Météo-France (MF), European Centre for Medium-Range Weather Forecasts (ECMWF), UK Met Office (UKMO) and the Korea Meteorological Administration (KMA) forecast systems. Additionally, we also consider the old version of the ECMWF forecast system in which the sea ice concentration was prescribed based on combining initial sea ice fields with relaxation towards climatological fields (ECMWF Pres.), a method that could be described as *damped persistence*. The technical features of these forecast systems are quite diverse: they differ in terms of initialization frequency (from daily to monthly), ensemble size (from 3 to 15 ensemble members), forecast length (from 44 to 60 days) and assimilation strategy. Only some of the systems directly assimilate sea ice concentration from observations and none of them assimilates sea ice thickness. Here, we consider the raw forecast data without calibration (bias/drift correction). The S2S Model Description[2] includes a detailed description of the S2S forecast systems.

The observations used to verify the forecasts are daily sea ice concentration fields retrieved from passive-microwave satellite measurements (OSI-450 – OSI-SAF (2016); Lavergne *et al.* (2019)). The sea ice edge has been defined as the 15% sea ice concentration contour line for both the forecast ensemble members and the observations. The verification results are averaged over a 12-years reforecast period (1999–2010) common to all of the S2S forecast systems. All the analyses have been conducted with the sea ice observation fields interpolated to the $1.5° \times 1.5°$ grid on which the S2S forecasts are provided. A common conservative land-sea mask has been obtained by combining the land-sea masks of all the models and observations based on the following criteria: if a grid cell is classified as land in one forecast system or in the observations, such classification is extended to all the other forecast systems, thus excluding that grid cell from all the analyses. The verification has been constrained to this land mask to allow a fair comparison between the different systems.

---

[2]`https://software.ecmwf.int/wiki/display/S2S/Models`

4.2.2 Verification metrics

The basic verification metric employed in this study is the Spatial Probability Score (SPS) (Goessling & Jung, 2018), which is defined as follows:

$$\text{SPS} = \int_A (P_f(x) - P_o(x))^2 \, \mathrm{d}A. \tag{4.1}$$

P_f and P_o are the local sea ice probabilities (SIP: the ensemble-based probability of sea ice concentration being above a certain threshold—here 15% if not differently stated) of respectively forecast and observation at location x. A property of the SPS that makes this metric suitable for verifying ensemble forecasts is its ability to deal directly with probabilities, which allows avoiding degrading probabilistic forecasts to deterministic ones. Since the sea ice observations considered here are deterministic and not probabilistic, their SIP simply consists of binary fields with 0 (no ice) and 1 (ice-covered cell). A is the integration domain, which is the northern hemisphere for the Arctic forecasts and the southern hemisphere for Antarctic forecasts.

Unlike the pan-Arctic sea ice extent, which measures only the total sea ice coverage, the SPS is designed to capture the accuracy of the sea ice spatial distribution and thus that of the sea ice edge location. Furthermore, the SPS can be decomposed into an Overestimation component (0 – SPS fraction caused by a local overestimation of the ice edge extent) and an Underestimation component (U – SPS fraction caused by a local underestimation of the ice edge extent), which provide additional insight into the type of the forecast error (Goessling *et al.*, 2016a; Zampieri *et al.*, 2018). Finally, the SPS can be also normalized (Norm. SPS) if divided by the length of the sea ice edge (Goessling *et al.*, 2016a; Melsom *et al.*, 2019; Palerme *et al.*, 2019). The Norm. SPS provides an estimate of the average distance between the (probabilistic) forecast edge and the (deterministic) observed edge. An advantage of this version of the metric is that it is easily understandable by potential forecast users. In this study, the length of the observed climatological sea ice edge, defined as the median of the climatological SIP (Fig. C.1), is used as normalization factor to assess longitudinal variations in Antarctic sea ice forecast skill (Sec. 4.3.3).

4.2.3 Benchmark forecasts

The predictive skill assessment of the forecast systems is based on the following approach: if for a given lead time the forecast SPS is lower than the SPS of some observational-based benchmarks, we consider this system to have predictive skill for that lead time. We employ two benchmark forecasts as a reference to assess the predictive skills of the S2S forecast systems: *1.* a probabilistic climatological forecast (CLIM) based on the observed sea ice conditions of the 10 years previous to the forecast target time at the same time of the year and *2.* a deterministic persistence forecast (PERS) based on the observed sea ice state at the forecast initial time.

4.3 Results

4.3.1 Comparison of the annual-mean forecast skills at the two poles

The annual-mean forecast skills in predicting the Arctic and Antarctic sea ice edge location are shown in Fig. 4.1 in terms of the SPS. In the following, we first focus on the Antarctic and then compare the predictive skills in the two hemispheres.

The ECMWF system (yellow line) is overall the most skillful system when it comes to predicting the Antarctic ice edge location. The system outperforms the CLIM and PERS benchmark forecasts from about day 5 to day ~30. The UKMO and KMA forecast systems (green and purple lines), which share the same model configuration, exhibit virtually identical results and show marginal predictive skill from day 8 to day 15. The old version of the ECMWF forecast system (ECMWF Pres. – magenta line) is less skillful than the benchmarks at all lead times and is characterized by a non-monotonic growth of the forecast error. The non-monotonicity is caused by the blending of different observations: first, the initial sea ice conditions are persisted up to day 15 of the forecast, and afterwards, the sea ice concentration is relaxed towards the climatological state based on the observations of the 5 years before the forecast target date.

The NCEP forecast system (light blue line) shows a rapid growth of the forecast error and has on average no predictive skill over the benchmarks. The wide uncertainty band is the result of large inter-annual variability of
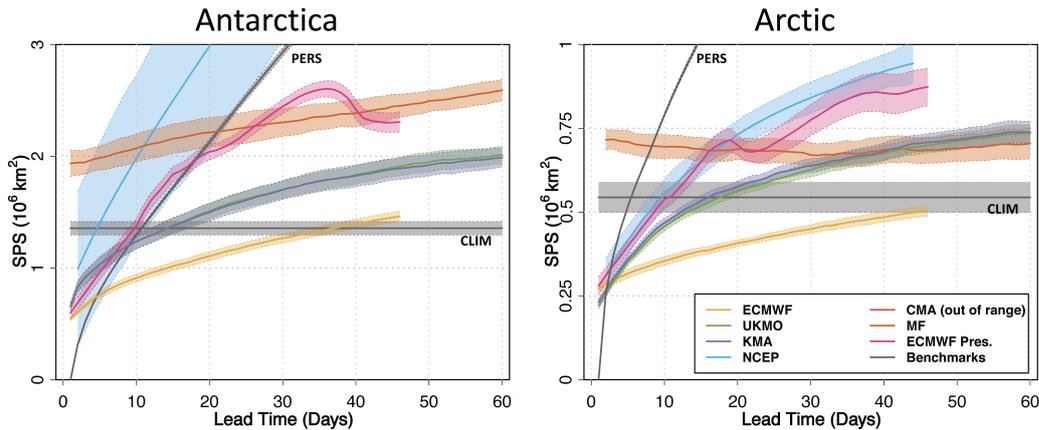
Figure 4.1: Annual-mean forecast skill in predicting the sea ice edge location in terms of the SPS of the different S2S systems (colored-solid lines), the climatological benchmark (constant gray-solid line) and the persistence benchmark (growing gray-solid line) as a function of forecast lead time for the Antarctic (left) and Arctic (right) regions. Note the different scales for the SPS. The averaging is performed over the common 12-years reforecast period (1999–2011). The shading and dashed lines indicate ∼95% confidence intervals, based on standard errors obtained from the twelve individual annual means. SPS = Spatial Probability Score; S2S = Sub-seasonal to Seasonal; NCEP = National Centers for Environmental Prediction; CMA = China Meteorological Administration; MF = Météo-France; ECMWF = European Centre for Medium-Range Weather Forecasts; UKMO = UK Met Office; KMA = Korea Meteorological Administration.

the NCEP forecast error. The MF forecast system exhibits an error 30% larger than CLIM already at initial time, growing further with lead time. Finally, the CMA forecast system (not visible in Fig. 4.1 because out of range for all lead times) is affected by strong biases related to the lack of assimilation of sea ice observations as well as to significant model biases in the polar regions. In the Antarctic, the ice edge extent is almost always and everywhere underestimated (Fig. 4.3), pointing to a wide-spread warm bias in the CMA system.

The results indicate some similarities between the two hemispheres. Firstly, the model ranking in the Antarctic is comparable to that in the Arctic. The only exception is the NCEP forecast system, which shows a degradation of its predictive skill in the Southern Ocean relative to the skills of the other systems and benchmarks. With the exception of April and May, the NCEP sea ice edge extent tends to be overestimated in most places (Fig. 4.3), pointing to a prevailing cold bias. Since the same sea ice model physics are implemented for both hemispheres, our results suggest that the NCEP forecast

system would benefit from a more careful tuning of its parameters to match better the observed state in the Southern Ocean. A second feature common to the two hemispheres is the large initial error, which amounts to $\sim 50\%$ of the CLIM error in the decently initialized systems (ECMWF, UKMO, KMA). As described in Zampieri *et al.* (2018), the initial error can have multiple sources, such as the adjustment of the sea ice edge to the sea surface temperature during the data assimilation, employment of different sea ice observations in the assimilation and verification phases and finally interpolation errors due to the regridding of the model and observational data to the coarse S2S grid. Understanding the relative contributions of different sources to the total initial error is challenging and beyond the scope of the present study.

Selected forecasts users might be interested in the verification of different sea ice concentration contours rather than the usual 15% threshold that defines the ice edge. Fig. C.2 shows a moderate error reduction when considering a higher threshold (50%), both for the forecast systems (only ECMWF is displayed) and for the climatological benchmark. This leads to a slight increase of the predictive skill at longer lead times (the forecast loses predictive skills at day 39 instead of day 37) that could be explained by a reduced sensitivity of the compact ice to weather events. Moreover, we observe a substantial reduction of the initial error ($\sim 40\%$), suggesting that this error is in part caused by a misrepresentation of dispersed sea ice in the marginal ice zone.

Finally, an obvious difference between the annual-mean forecast errors in the two hemispheres is their overall magnitude. The Antarctic SPS is on average $\times 2.6$ larger than the Arctic SPS. This difference is in part explained by the fact that the Antarctic sea ice edge is on average $\times 1.8$ longer than the Arctic one (Fig. C.1). If one assumes errors in terms of ice edge distance to be regionally independent, then the forecast SPS would tend to be proportional to the length of the edge. However, under this assumption, the sea ice edge length difference can explain only $\sim 70\%$ of the hemispheric SPS discrepancy, while the remaining $\sim 30\%$ reflects increased errors in terms of ice-edge distance in the Antarctic. A way to account for variations in ice edge length explicitly is to normalize the SPS with the ice edge length; such an approach is taken in Sec. 4.3.3.

4.3.2 Seasonality and components of the Antarctic forecast error

One of the strengths of the S2S Database is the availability of forecasts all year round for a period of time longer than a decade. This allows us to assess seasonal variations of the forecast error.
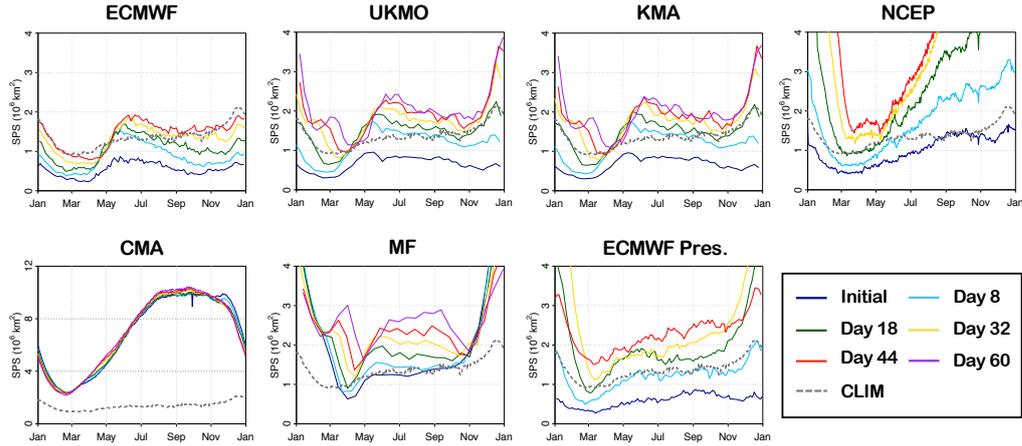

Figure 4.2: Skill in predicting the Antarctic sea ice edge location in terms of the SPS of seven individual S2S forecast systems. The results have been averaged over the common 12-years reforecast period (1999–2011). The SPS is displayed for six different lead times (see legend) as a function of the target date (expressed in days of the year). The different resolution with respect to the target date reflects differences in the initialization frequency of the reforecasts. Note the different SPS scale adopted for the CMA forecast system. SPS = Spatial Probability Score; S2S = Subseasonal to Seasonal; NCEP = National Centers for Environmental Prediction; CMA = China Meteorological Administration; MF = Météo-France; ECMWF = European Centre for Medium-Range Weather Forecasts; UKMO = UK Met Office; KMA = Korea Meteorological Administration.

The CLIM benchmark forecast exhibits seasonal variations of the SPS that correlate well to the length of the sea ice edge (Fig. 4.2, dashed curves; compare with Fig. C.1). The SPS reaches its minimum value in March, immediately after the annual sea ice extent minimum and when the sea ice edge is the shortest. The CLIM SPS slowly grows during the following months as the ice edge becomes longer and stretches further to the north. The CLIM SPS maximum is finally reached during the melting season in November and December when the Antarctic sea ice edge is the longest.

In general, the S2S forecast systems exhibit similar seasonal variations as the CLIM benchmark, in particular at the initial time. The only exception is CMA, which, as already mentioned, is affected by strong model and data
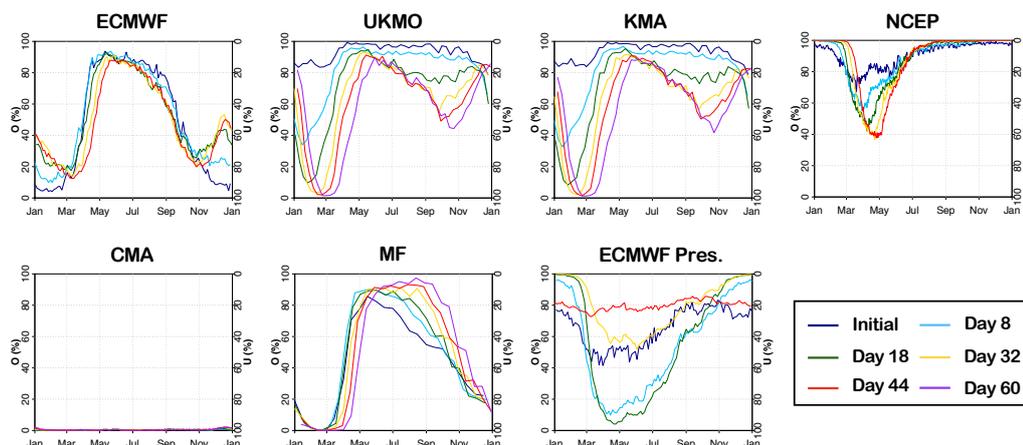
Figure 4.3: Relative contributions to the Integrated Ice Edge Error of the ensemble-median ice edge from Overestimation(O) versus Underestimation (U) of individual S2S systems as a function of the time of the year (target date) and for six different lead times (see legend). Results have been averaged over the common reforecast period 1999–2010. SPS = Spatial Probability Score; S2S = Subseasonal to Seasonal; NCEP = National Centers for Environmental Prediction; CMA = China Meteorological Administration; MF = Météo-France; ECMWF = European Centre for Medium-Range Weather Forecasts; UKMO = UK Met Office; KMA = Korea Meteorological Administration.

assimilation related biases that we do not further discuss. The ECMWF seasonality is in line with the CLIM benchmark, with the forecast error approaching the climatological error with increasing lead time. Only during the second half of the freezing season (May to August) the forecast errors at longer lead times significantly exceed the CLIM error due to an overall overestimation of the sea ice edge extent (Fig. 4.3 – ECMWF). The UKMO and KMA systems show a similar freezing-season bias, also linked to an overestimation of the ice edge extent. These two systems exhibit an additional degradation of the predictive skills during the melting season (December and January, Fig. 4.2) for lead times longer than 18 days. This suggests that the two systems have difficulties transitioning into the sea ice melting regime when initialized during a maximum-extent phase. The NCEP forecast system is characterized by a similar bias that is largest during the melting season. Specifically, NCEP strongly overestimates the ice edge extent during most of the year, except in the first two months of the freezing season (March to May – Fig. 4.3).

4.3.3 Regional skill in terms of ice edge distance

Fig. 4.4 displays the longitudinal variation of the forecast and CLIM benchmark errors in terms of the Norm. SPS. In agreement with our previous findings, only the ECMWF forecast system is still partially skillful after one forecast month. The forecast error exceeds the error of the climatological benchmark after 32 forecast days in the east Antarctic sector (from 80°E to 170°E; Fig. 4.4) and even earlier in the Haakon VII Sea. However, the system is skillful up to day 44 in some portions of the West Antarctic sector (Ross, Amundsen and Weddell Seas), where the Norm. SPS remains up to 40 km lower compared to CLIM. The other forecast systems lose their predictive skill much faster and none of them is skillful at the monthly range in any location around Antarctica (Fig. 4.4). The very similar UKMO and KMA systems are on average skillful up to day 18 (green lines lower than CLIM), whereas the remaining systems lose their predictive skill before day 8 (ECMWF Pres. and NCEP) or are not even skillful at initial time (MF and CMA).

The skill in predicting the sea ice edge location differs substantially among the S2S forecast systems. However, the analysis of the annual-mean longitudinal variation of the forecast error reveals also some features common to multiple systems. The forecasts are overall less skillful (relative to the climatological benchmark) in the eastern Antarctic [0°E;180°E] than in the western Antarctic [−180°E;0°E]. This does not necessarily imply that the models are particularly good at capturing the evolution of the sea ice edge in the West Antarctic regions, but rather that the climatological forecasts are more accurate in the eastern sectors because of a lower sea ice edge variability. Both CLIM (Fig. 4.4; gray-dashed line) and the S2S forecasts (coloured lines) exhibit larger errors in terms of ice edge distance (Norm. SPS) in the Ross and Weddell Seas, suggesting that formulating accurate subseasonal sea ice edge predictions in these regions is challenging because of the high complexity and variability of the local climate system. Our results agree with Massonnet *et al.* (2018) who find large sea ice area prediction uncertainties in the Weddell and Ross Seas for late summer.
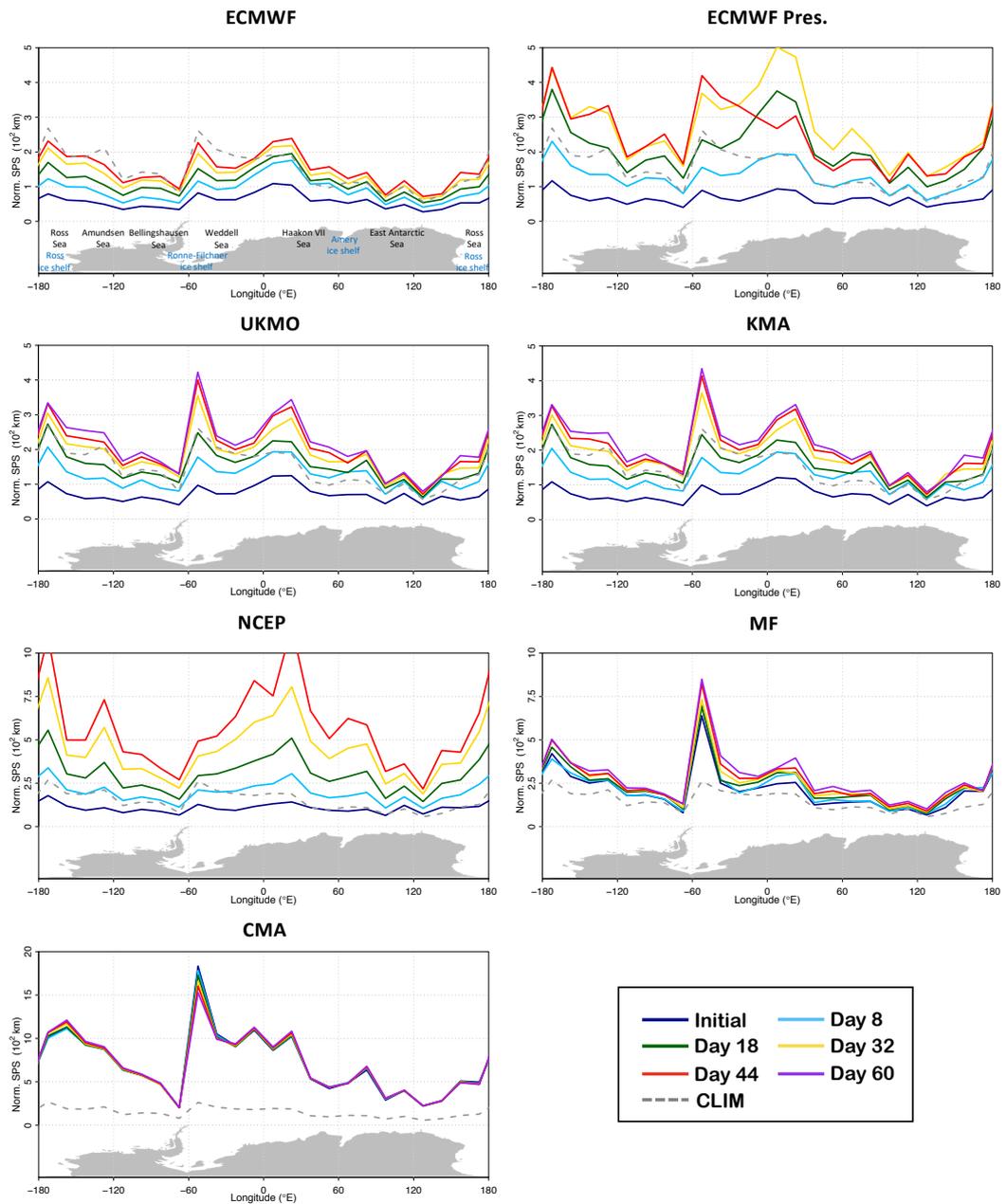
Figure 4.4: Annual-mean sea ice edge forecast error in terms of the Norm. SPS of seven individual S2S forecast systems and of the CLIM benchmark as a function of longitude. The results are averaged over the common 12-years reforecast period (1999–2010) and displayed for 6 lead times (see legend). The longitude domain $[-180°\text{E}, 180°\text{E}]$ is divided into 24 equally-spaced bins. Note the different Norm. SPS scales adopted for the forecast systems. Geographical names of the main oceanic sectors and ice shelves are indicated in respectively black and blue in the upper-left plot. Norm. SPS = Normalized Spatial Probability Score; NCEP = National Centers for Environmental Prediction; CMA = China Meteorological Administration; MF = Météo-France; ECMWF = European Centre for Medium-Range Weather Forecasts; UKMO = UK Met Office; KMA = Korea Meteorological Administration.

A further error peak can be observed in the west Haakon VII Sea (0°E to 40°E). Unlike the previous error peaks in the Ross and Weddell Seas (featured both in the CLIM benchmarks and the S2S forecasts), the west Haakon VII Sea error peak is more pronounced for the forecast systems (ECMWF, ECMWF Pres., UKMO, KMA and NCEP) than for the CLIM benchmark. The NCEP system displays a particularly fast error growth with lead time in this region. In contrast, in the more skillful systems (ECMWF, UKMO and KMA) this regional error peak appears to be caused mainly by accordingly large initial errors ($\geq 100\,\mathrm{km}$). More generally, the Antarctic average initial error in these systems is considerable ($\geq\sim 70\,\mathrm{km}$), suggesting again that investments into the sea ice initialization procedure appear promising to enhance predictive capacity.

4.4 Discussion

This study provides the first thorough assessment of the skill of current operational ensemble forecasting systems in predicting the location of the Antarctic sea ice edge on subseasonal timescales. We find that only one of the considered forecast systems outperforms two benchmarks (persistence and climatology) for a wide range of lead times, namely from about 5–30 days. On average, the other systems perform worse than either persistence or climatology at any lead time considered here. The forecasts are in general more skillful in the west Antarctic sector than in the east Antarctic sector, where the climatological benchmark forecast provides a more accurate estimate of the sea ice edge location. In particular, the ECMWF forecast system outperforms the climatological benchmark forecast in the Ross, Amundsen and Weddell Seas, where predictive skill up to 44 days into the forecast is found.

We identify two types of errors that are common to several forecast systems: (i) a "freezing-season bias" that affects ECMWF, UKMO, KMA and MF and (ii) a "melting-transition bias" that affects UKMO, KMA and NCEP (Balan-Sarojini *et al.*, 2019; Blockley & Peterson, 2018). Both are caused by a systematic overestimation of the sea ice edge location (i.e. predicted to be too northward). While the first bias can be explained by a misrepresentation of thermodynamical processes in the coupled models, with the oceanic surface cooling and freezing too rapidly, the second bias could be linked to an initial overestimation of the sea ice thickness, which would delay the melting

onset and thus the ice edge retreat in spring. At the moment we are not able to test this last hypothesis because the S2S Database does not include sea ice thickness as a standard output variable.

The hemispheric comparison reveals that differences between the Arctic and Antarctic cannot be explained by differences in the sea ice edge length. This holds not only for the S2S forecast systems but also for the climatological benchmark forecast, suggesting that larger model biases in the Southern Ocean are not the major cause for this difference, but rather that this is due to an intrinsic property of the Antarctic climate system. The Antarctic forecast skill degradation points to a higher variability of the Antarctic sea ice edge at subseasonal timescales compared to the Arctic. Similar differences in skill between the hemispheres have been found for atmospheric predictions in polar regions and beyond (Jung & Matsueda, 2016; Bauer *et al.*, 2015).

Given the relatively large forecast errors—ranging from 50 km to 250 km even for the best forecast systems—sea ice edge forecasts with state-of-the-art operational systems need to be used carefully. However, there might be some useful applications already. One example relates to the medium-term planning of ship tracks to optimize the provision of research stations in the Antarctic continent during the brief Antarctic summer and at the beginning of the freezing season. Furthermore, the probabilistic nature of the S2S forecasts could be beneficial for identifying the possibility of extreme sea ice conditions.

Our results suggest that current sea ice edge forecast capabilities for the Southern Hemisphere are lagging behind those for the Northern Hemisphere. Nevertheless, we anticipate that major improvements in forecast models and initialization techniques, together with further in-situ observations to better understand the physical processes at the atmosphere-sea ice-ocean interfaces, will render Antarctic sea ice forecasts a valuable resource for guiding operational decision-making in the Southern Ocean.

5. Impact of sea-ice model complexity on the performance of an unstructured sea-ice/ocean model under different atmospheric forcings[1]

Abstract

We have equipped the unstructured-mesh global sea-ice and ocean model FESOM2 with a set of physical parameterizations derived from the single-column sea-ice model Icepack. The update has substantially broadened the range of physical processes that can be represented by the model. The new features are directly implemented on the unstructured FESOM2 mesh, and thereby benefit from the flexibility that comes with it in terms of spatial resolution. A subset of the parameter space of three model configurations, with increasing complexity, has been calibrated with an iterative Green's function optimization method to test fairly the impact of the model update on the sea-ice representation. Furthermore, to explore the sensitivity of the results to different atmospheric forcings, each model configuration was calibrated separately for the NCEP-CFSR/CFSv2 and ERA5 forcings. The results suggest that a complex model formulation leads to a better agreement between modeled and the observed sea-ice concentration and snow thickness, while differences are smaller for sea-ice thickness and drift speed. However, the choice of the atmospheric forcing also impacts the agreement of FESOM2 simulations and observations, with NCEP-CFSR/CFSv2 being particularly beneficial for the simulated sea-ice concentration and ERA5 for sea-ice drift speed. In this respect, our results indicate that the parameter calibration can better compensate for differences among atmospheric forcings in a simpler model (i.e. sea-ice has no heat capacity) than in more energy consistent formulations with a prognostic ice thickness distribution.

5.1 Introduction

Sea-ice is a key component of the climate system (Dieckmann & Hellmer, 2010) and it plays a central role as a physical regulator of the energy exchange between atmosphere and ocean in polar regions (Döscher *et al.*, 2014). Furthermore, sea-ice represents by itself a platform where large ecosystems thrive (Spindler, 1994), and it is a fundamental element in the lives of coastal human communities in the Arctic (Cooley *et al.*, 2020). Because of the strong and rapid transformations that sea-ice has faced in recent years due to global warming (particularly in the Arctic; Notz & Stroeve (2016)), there is an urgent need to better understand and being able to quantify the physical and biogeochemical mechanisms regulating the sea-ice system, to inform decision-makers and various stakeholders. Reliable dynamical sea-ice models are fundamental tools for accurately predicting the evolution of sea ice at multiple timescales, from days to centuries into the future.

In the past decades, there has been a constant development of more complex and physically consistent sea-ice model formulations, summarized by Hunke *et al.* (2010) and Notz (2012), and of which we give a brief overview in Sec. 5.2.2. At the same time, the resolution of sea-ice and ocean models has increased due to the growing availability of computational resources, and so has the resolution and quality of the atmospheric reanalyses used to force the models. These developments, together with the growing availability of more accurate sea-ice observations to constrain our models, have lead to better sea-ice simulations. Multiple studies attribute a relevant role in improving the sea-ice model performances to more realistic model formulations (Vancoppenolle *et al.* (2009); Massonnet *et al.* (2011); Flocco *et al.* (2012); Roach *et al.* (2018b), among others). However, in the framework of the Coupled Model Intercomparison Project (CMIP), the SIMIP Community (2020) (Sea Ice Model Intercomparison Prohect) shows that it is unclear to what degree differences between CMIP6, CMIP5, and CMIP3 sea-ice simulations are caused by better model physics versus other changes in the forcing. In the field of subseasonal and seasonal sea-ice forecasting, simple dynamical models exhibit predictive skills comparable or even better than those of more complex forecast systems (Zampieri *et al.*, 2018, 2019), suggesting that the yeartoyear variability, the skill of the atmospheric models, and the quality of initial conditions dominate the variation in ensemble prediction success

(Stroeve *et al.*, 2014). In conclusion, to what extent the model complexity impacts the quality of sea-ice simulations remains an open question always evolving with our models (Blockley *et al.*, 2020).

A key aspect to examine when assessing the relative performances of multiple model formulations is whether these are all appropriately tuned (Miller *et al.*, 2006). Because of an interdependency of model parameters and a lack of comprehensive ice and snow observations, the model parameters are in general underconstrained (Urrego-Blanco *et al.*, 2016), and their systematic calibration can substantially impact the quality of the simulations (Turner *et al.*, 2013b; Massonnet *et al.*, 2014; Ungermann *et al.*, 2017; Sumata *et al.*, 2019a; Roach *et al.*, 2018a). Furthermore, acknowledging the substantial differences between the reanalysis products used to force the sea-ice models in stand-alone setups (Batrak & Müller, 2019), we argue that the same model configuration should be also optimized separately for different forcing conditions. As shown by Miller *et al.* (2007), the behavior of a specific model formulation can change substantially based on the forcing used.

Most of the relevant sea-ice parameterizations and modeling strategies developed over the years have been collected by the scientific community and integrated into sophisticated sea-ice models, the most advanced and complete of which is arguably CICE, (Hunke *et al.*, 2020a). The CICE model is distributed in combination with the Icepack column-physics package (Hunke *et al.*, 2020b) – a collection of physical parameterizations that account for thermodynamic and mechanic sub-grid processes not explicitly resolved by the models. Because of its modularity, Icepack can be conveniently implemented in ocean and sea-ice models other than CICE. In this regard, this study presents a new version of the Finite-volumE Sea ice-Ocean Model version 2 (FESOM2; Danilov *et al.* (2017)) that exploits the capabilities of the Icepack column physics package. As we describe in Sec. 5.2.2, the development of the FESOM2 sea-ice component has been mostly focused on dynamical aspects, while the adopted sub-grid sea-ice parameterizations were quite simple and outdated if compared to those implemented in other sea-ice models. This resulted in a partially inconsistent physical formulation of the standard FESOM2 model, caused for example by the missing representation of the sea-ice internal energy. The inclusion of Icepack in FESOM2 has substantially broadened the range of sea-ice physical processes that can be simulated by the FESOM2 model, making it an ideal tool for answering the scientific questions posed below.

Based on the new FESOM2-Icepack implementation, we designed a set of experiments to assess the impact of the sea-ice model complexity on the quality of the sea-ice simulations. Ten parameters from three distinct model setups are optimized with a semi-automated calibration technique and compared to different types of sea-ice and snow observations. Because we deal with a standalone ocean and sea-ice model (i.e. no coupling to an atmospheric model) the calibration process is conducted separately for two different atmospheric reanalysis products used to force FESOM2. Based on the outcome of the calibration and the resulting model performance, we try to address the following questions:

1. Does a more complex and physically consistent formulation of the sea-ice model lead to better sea-ice simulations given the resolution, coverage and uncertainty of satellite Earth Observations (EO) of the sea-ice available today?

2. How does the impact of different atmospheric forcings on the sea-ice model performance relate to the impact of model complexity?

3. Which sea-ice formulation can be calibrated more effectively?

The remainder of this paper is organized as follows: the method section presents the standard (Sec. 5.2.1) and Icepack (Sec. 5.2.2) FESOM2 formulations, followed by the theoretical description of the Green's function approach for the calibration of the model parameter space (Sec. 5.2.3). We then describe the experimental setups employed in the study and we present the practical implementation of the calibration technique (Sec. 5.2.4), as well as the observations used for constraining the parameter space and for validating the model results (Sec. 5.2.5). The results section (Sec. 5.3) describes the impact of the parameter optimization on the model performances in terms of cost function reduction. Furthermore, we explore the discrepancies of the various optimized model configurations by comparing the simulated sea-ice and snow state to different types of observations, and by linking this to differences in the optimized model parameters. Finally, the computational performances of three model setups is analyzed for assessing the sustainability of more sophisticated, and thus computationally more demanding, sea-ice setups for diverse modeling applications (Sec. 5.4.3).

5.2 Methods

5.2.1 Standard sea-ice formulation in FESOM2

Danilov *et al.* (2015b) describes in detail the numerical implementation of
the Finite Element Sea-Ice Model (FESIM), which is the standard sea-ice
component of FESOM2. Three alternative algorithms are available for solv-
ing the sea-ice momentum equation: a classical elastic-viscous-plastic (EVP)
approach coded following Hunke & Dukowicz (1997) plus two modified ver-
sions of the EVP solver: the modified EVP (mEVP; Kimmritz *et al.* (2015)),
and the adaptive EVP (aEVP; Kimmritz *et al.* (2016)). Three sea-ice tracers
are advected based on a finite element (FE) flux corrected transport (FCT)
scheme (Lhner *et al.*, 1987): the sea-ice area fraction a_i, and the sea-ice and
snow volumes per unit area, v_i and v_s. The thermodynamic evolution of sea
ice is described by a simple 0-layer model (i.e. the sea-ice and snow layers
have no heat capacity) that follows Parkinson & Washington (1979). The
interaction between the radiation and sea ice is mediated by four constant
albedo values (dry ice, wet (melting) ice, dry snow, and wet (melting) snow)
that respond to changes in the atmospheric near-surface temperature, thus
including an implicit description of the radiative effect of melt ponds during
the melting season. No incoming shortwave radiation penetrates through the
snow and sea-ice layers.

5.2.2 Icepack implementation in FESOM2

Icepack (Hunke *et al.*, 2020b) – the column physics package of the sea-ice
model CICE – is a collection of physical parameterizations that account for
thermodynamic and mechanic sub-grid processes not explicitly resolved by
the hosting sea-ice model. The modular implementation of Icepack allows
the users to vary substantially the complexity of the sea-ice model, with
the possibility of choosing between several schemes and a broad set of ac-
tive and passive tracers that describe the sea-ice state. Similarly to FESIM,
Icepack can make use of a simple 0-layer sea-ice and snow thermodynam-
ics scheme (Semtner, 1976). However, two more sophisticated and energy
consistent multi-layer thermodynamics formulations, taking into account the
sea-ice enthalpy and salinity, are also available: the Bitz & Lipscomb (1999)
thermodynamics (BL99 hereafter), which assumes a temporally constant sea-

ice salinity profile, and the "mushy layer" implementation, with a prognostic sea-ice salinity description (Turner *et al.*, 2013a). To account for the sea-ice thickness variations typically observed at sub-grid scales, Icepack discretizes the sea-ice cover in multiple classes, each representative of a sea-ice thickness range, and describes prognostically the evolution of the Ice Thickness Distribution (ITD) in time and space (Bitz *et al.*, 2001). The processes leading to changes in the ITD are sea-ice growth and melt, snow-ice formation (flooding), and mechanical redistribution (i.e. sea-ice ridging and rafting due to dynamical deformation; Lipscomb *et al.* (2007)). In terms of the interaction between sea ice and radiation, Icepack includes two more sophisticated parameterizations in addition to a simple albedo scheme similar to that of FESIM. In the "CCSM3" formulation, the surface albedo depends on the sea-ice and snow thickness and temperature, and it is defined separately for the visible and infrared portion of the spectrum. The main difference between this and the constant albedo approach is a reduction of the surface reflectivity for thin sea-ice or snow. The even more sophisticated "Delta-Eddington" formulation exploits the inherent optical properties of snow and sea ice for solving the radiation budget (Holland *et al.*, 2012), and it can be combined with three explicit prognostic melt pond schemes (Holland *et al.*, 2012; Flocco *et al.*, 2010; Hunke *et al.*, 2013). Finally, the Icepack radiation implementation allows the penetration of part of the incoming shortwave radiation through snow and sea ice, leading to additional energy absorption in the water column below the sea ice.

Icepack v1.2.1 has been implemented in FESOM2 and can now be used as an alternative to the standard FESIM thermodynamic module. As the standard FESIM implementation, the Icepack column-physics subroutines run every ocean time step. All the Icepack variables are defined directly on the FESOM2 mesh, ensuring an optimal consistency between the ocean and the sea-ice components of the model. The inclusion of Icepack in FESOM2 required a revision of the calling sequence within the sea-ice model (Fig. 5.1), which now follows that of the CICE model (Hunke *et al.*, 2020a). The coefficients mediating the momentum and heat exchanges between atmosphere and ice, previously constant in FESIM, have been updated and are now computed iteratively based on the stability of the atmospheric near-surface layer (Jordan *et al.*, 1999). The solution of the momentum equation for computing the sea-ice velocity does not change when running in FESOM2-Icepack configuration. Two alternative formulations of the sea-ice strength P are

available in Icepack and can be used in the EVP solver:
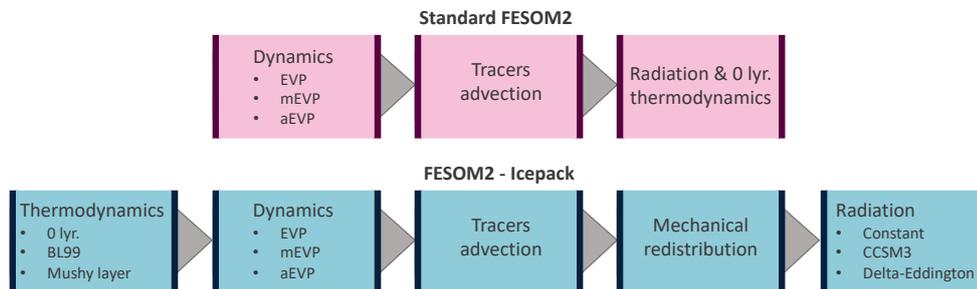
Figure 5.1: Schematic describing the calling sequences of the Standard FESOM2
and FESOM2-Icepack implementations.

$$\text{Hibler (1979):} \quad P = P^* v e^{-C^*(1-a_i)} \tag{5.1}$$

$$\text{Rothrock (1975):} \quad P = C_p C_f \int_0^\infty h^2 \omega_r(h) dh \tag{5.2}$$

where $h = v/a$ is the ice thickness, P^*, C^*, and C_f are empirical parameters,
$C_p = \rho_i(\rho_w - \rho_i)g/(2\rho_w)$ is a combination of the gravitational acceleration
and the densities of ice and water, and $\omega_r(h)$ is a function that represents
the effective sea-ice volume change for each thickness class due to mechanical
redistribution processes. In this study, the Hibler (1979) approach (H79
hereafter) is adopted for all model setups instead of the Rothrock (1975)
approach (R75 hereafter). The reasoning behind this choice will be discussed
in Sec. 5.2.4.

In the FESOM2 implementation of Icepack, each tracer is advected sep-
arately using the FE-FCT scheme by Lhner *et al.* (1987) as described in
Kuzmin (2009). The tracer advection is based on the conservation equation

$$\partial_t T + \nabla \cdot (T\mathbf{v}) = 0 , \tag{5.3}$$

where T is a generic advected tracer with no dependencies and $\mathbf{v}$ is the sea-
ice velocity that solves the momentum equation. If a tracer T_2 depends on
another tracer T_1, the advected quantity that satisfies Eq. 5.3 is $T = T_1 T_2$.
This concept can be generalized for a tracer with more than one dependency.
Icepack comes with a vast set of required and optional tracers. As for the
standard FESIM, a_i, v_i, and v_s are required tracers. However, in Icepack
these three variables are defined separately for each ice thickness class. The
skin temperature of the sea-ice, or in the presence of snow of the snow, T_s

is also defined separately for each thickness class and depends on a_i for the advection. If the BL99 or mushy thermodynamics are used, the enthalpy of sea-ice and snow layers (q_i,q_s), and the sea-ice salinity s_i become also required tracers and depend on v_i or v_s. Several more tracers are available (melt pond fraction and depth, sea-ice age, first-year ice fraction, level ice fraction and volume, etc.) depending on the chosen setup of the model. All these tracers are implemented in the FESOM2-Icepack model.

5.2.3   Green's function approach for the optimization of model parameters

The Green's function approach is a simple, yet powerful method that, given some observations, can be used for the calibration of the parameter space of general circulation models (Stammer & Wunsch, 1996; Menemenlis & Wunsch, 1997; Menemenlis *et al.*, 2005; Nguyen *et al.*, 2011; Ungermann *et al.*, 2017). The practical realization of one iteration of this method requires to compute an ensemble of n sensitivity simulations by perturbing separately each one of the n parameters that we choose to optimize. The Green's functions of these sensitivity simulations are then combined through discrete inverse theory for constructing an optimal linear solution that minimizes the difference between the model state and the observations, and which corresponds to a set of optimal parameter perturbations. Menemenlis *et al.* (2005) and Ungermann *et al.* (2017) provide an extensive mathematical derivation of the method. Here, we limit our description to a few important points.

Given a vector of m observations $\mathbf{y}$ and their measurement uncertainties $\boldsymbol{\sigma}$, the relationship between the observations and a model operator G can be expressed as

$$\mathbf{y} = G(\boldsymbol{\nu}) + \boldsymbol{\epsilon} \; , \qquad (5.4)$$

where $\boldsymbol{\nu}$ contains a generic set of n parameter perturbations around a reference state $\boldsymbol{\nu}_0$, and $\boldsymbol{\epsilon}$ represents the discrepancy between the observations and the model results. The optimal set of parameters $\boldsymbol{\nu}_{opt}$ can be obtained by minimizing a quadratic cost function

$$F = \boldsymbol{\epsilon}^T \mathbf{R} \boldsymbol{\epsilon} \; , \qquad (5.5)$$

where $\mathbf{R}$, the covariance matrix of $\boldsymbol{\epsilon}$, is assumed to be a simple diagonal

matrix with elements $R_{ij} = (\sigma_i)^{-2}$ (with $i, j = 1 \ldots m$), meaning that observation errors are considered independent. In this study, each element of $\mathbf{R}$ is further divided by the total number of observations of its corresponding observation type. In this way, the same weight is given to each observational type employed in the optimization. We assume now that a linearization of the system holds, and that the model operator G can be represented by a matrix $\mathbf{G}$, so that the misfit between observations and the control simulation (for which $\boldsymbol{\nu} = 0$) can be expressed as

$$\Delta \mathbf{y} = \mathbf{y} - G(0) = \mathbf{G}\,\boldsymbol{\nu} + \boldsymbol{\epsilon} \;. \tag{5.6}$$

In practice, $\mathbf{G}$ is an $m \times n$ matrix constructed by combining the Green's function for each of the parameter perturbations $\boldsymbol{\nu} = (\nu_1 \ldots \nu_n)$. Specifically, $\boldsymbol{g}_j$—the jth-column of the matrix $\mathbf{G}$—is

$$\boldsymbol{g}_j = \frac{G(\boldsymbol{\nu}_j) - G(0)}{\nu_j} \;, \tag{5.7}$$

where $G(\boldsymbol{\nu}_j)$ is the sensitivity simulation where only the parameter ν_j is perturbed. The set of optimal parameters that minimizes the cost function is given by

$$\boldsymbol{\nu}_{opt} = \boldsymbol{\nu}_0 + (\mathbf{G^T R\,G})^{-1}\mathbf{G^T R}\,\Delta \mathbf{y} \;. \tag{5.8}$$

Even if the Green's function approach is a robust method for tuning the model effectively, there is no guarantee that the estimated optimal parameters lead to a model state that corresponds to a global minimum of the cost function, in particular if the cost function is not a "well-behaved" function as in the case of sea-ice observations. In this respect, the results by Sumata *et al.* (2013) shows that a stochastic optimization method is more appropriate for finding a global minimum of the cost function than gradient descent methods as the Green's function approach (Figs. 4 and 5 of Sumata *et al.* (2013) reveal the heterogeneity of the sea-ice concentration cost function). In the context of this study, where the model optimization is performed for three model configurations each forced with two sets of atmospheric boundary conditions, the Green's function approach has been chosen because it provides a balance between the effectiveness of the method, simplicity of implementation, and associated computational costs.

5.2.4 Model simulations

All model simulations are run on a global mesh with 1.27×10^5 surface nodes and 46 vertical levels. This unstructured mesh has approximately a 1° resolution over most of the domain, but it is refined along the coastlines, in the equatorial regions, and north of 50°N, where the resolution reaches ∼25km (see Fig. 4a in Sein *et al.* (2016) for more details on the mesh). The atmospheric boundary conditions used to force the FESOM2 model are derived from two reanalysis products: the European Centre for Medium-Range Weather Forecasts Reanalysis, 5th Generation (ERA5) global reanalysis (Hersbach *et al.*, 2020) and the NCEP Climate Forecast System (NCEP hereafter; Saha *et al.* (2010, 2014)). The fields used to force the model are the 2-m air temperature and specific humidity, the 10-m wind velocity, the downward longwave and shortwave radiation, and the liquid and solid precipitations. The ocean component of the FESOM2 model is initialized in 1980 from the PHC3 ocean climatology (Steele *et al.*, 2001). A sea-ice thickness of 2m is set at initial time in regions with sea surface temperature below the freezing temperature of sea water of typical salinity for the inner Arctic surface ocean.

The Green's function approach for parameter optimization is applied to three different model setups of increasing complexity:

C1 Low-complexity configuration corresponding to the standard FESIM implementation within FESOM2, as described in Sec. 5.2.1.

C2 Medium-complexity configuration based on the FESOM2-Icepack implementation described in Sec. 5.2.2. This configuration features an ITD with 5 thickness classes, the BL99 thermodynamics (4 sea-ice layers and 1 snow layer), and the CCSM3 radiation scheme.

C3 High-complexity configuration based on the FESOM2-Icepack implementation. Like C2, C3 features an ITD with 5 thickness classes and the BL99 thermodynamics. The CCSM3 radiation is replaced by the Delta-Eddington scheme, and the melt ponds are prognostically described with the CESM parameterizations (Holland *et al.*, 2012).

Each configuration is optimized twice, once for each atmospheric forcing employed: ERA5 (suffix "E" hereafter) and NCEP (suffix "N" hereafter). This leads to a total of 6 optimal parameter sets, each one optimized by
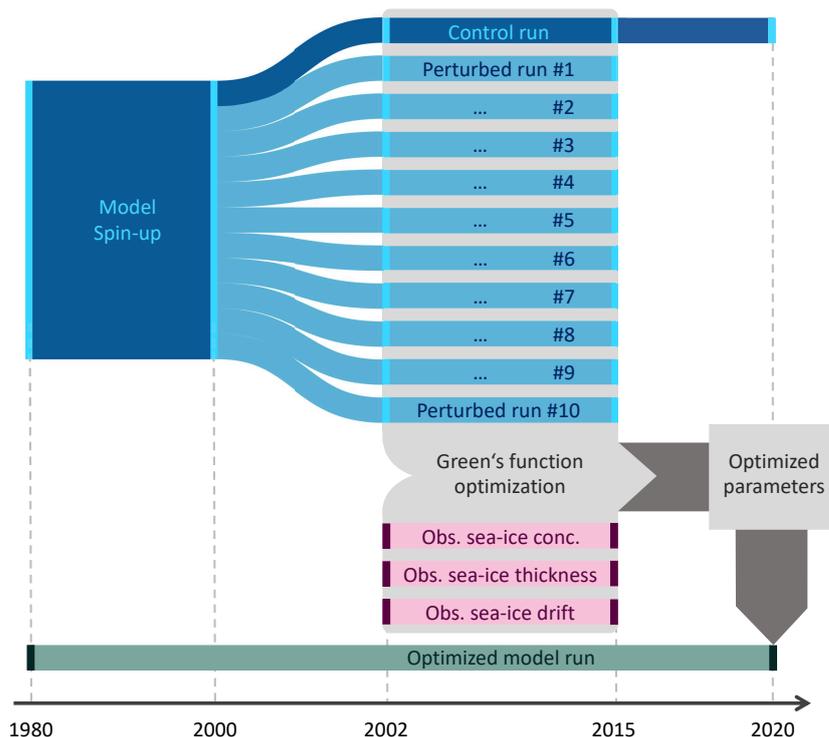
Figure 5.2: Schematic of one iteration of the Green's function approach for parameter optimization as employed in our study for each configuration. When the second iteration is performed, the optimized model run computed at the end of the first iteration serves as control run for the second one.

performing two iterations of the Green's function method. A schematic of the Green's function optimization procedure is displayed in Fig. 5.2. Each configuration undergoes a 20-year spin-up (1980–2000) to guarantee a realistic state of the modelled upper ocean and of the sea-ice cover in (quasi-)equilibrium with the chosen atmospheric forcing product and the individual parameter set. The model optimization window is limited to the 14 years period 2002–2015, i.e. the cost function is evaluated in this period.

The R75 formulation of the sea-ice strength is arguably more physically consistent than the H79 formulation, as it includes information about the ITD in each grid cell and it considers potential energy changes associated with the redistribution. However, Ungermann *et al.* (2017) show that the H79 approach leads to a better fit between model data and observations when properly tuned. In addition, the R75 sea-ice strength is much more non-linear then H79 one. For these reasons, and for being able to compare the C1 setup (no ITD; only H79 available) to the C2 and C3 setups (with ITD; both H79 and R75 available), all the simulations here presented employ

the H79 sea-ice strength formulation.

Because the finite availability of computational resources limits in practice the number of parameters that can be optimized with the Green's function approach (a separate sensitivity run is needed for each parameter one intends to optimize), the parameters have been chosen based on their ability to influence the sea-ice state of the model, as described in previous studies (Massonnet *et al.*, 2014; Urrego-Blanco *et al.*, 2016; Ungermann *et al.*, 2017; Sumata *et al.*, 2019a). In total, 10 model parameters are optimized for each of the three model setups (Tab. 5.1). The chosen parameters act on various sea-ice parameterizations: thermodynamics, dynamics, radiation, and mechanical redistribution. Some are common to all three configurations (α_O, k_S, P^*, C^*, and c_{IO}), while others are specific to the formulation of each setup. Note that δ_P has been classified as radiation parameter (Tab. 5.1a) because the chosen melt pond scheme describes only the radiation effects of melt ponds. The ice-atmosphere drag coefficient c_{IA} has not been optimized following the results of Massonnet *et al.* (2014), which show that optimizing the atmospheric drag is not necessary if P^* and c_{IO} are already optimized.

5.2.5 Observational products

The Green's function optimization method is based on three types of monthly averaged satellite observations and their uncertainties: sea-ice concentration, thickness, and drift (Fig. 5.2). We employ the OSI SAF Global Sea Ice Concentration Climate Data Record v2.0 (EUMETSAT Ocean and Sea Ice Satellite Application Facility, 2017) for the period 2002–2015. The retrieval of this product is based on passive microwave data from the SSM/I (Special Sensor Microwave/Imager) and SSMIS (Special Sensor Microwave Imager/-Sounder) sensors (Lavergne *et al.*, 2019). The data are distributed on a polar stereographic 25km resolution grid, which is approximately the same resolution of our model in the Arctic.

Two complementary sea-ice thickness datasets are considered during the freezing season (October to April): the monthly northern hemisphere sea-ice thickness from Envisat (2002–2010; Hendricks *et al.* (2018b)) and from CryoSat-2 (2011–2015; Hendricks *et al.* (2018a)). The merged CryoSat-2/SMOS sea-ice thickness product has not been considered for the parameter optimization because we decided to prioritize the optimization of thick sea-ice regions over the marginal ice zone. The evolution of the thin ice cover

(a) Optimized parameters in C1

Ocean albedo	α_O	Therm. conductivity snow	k_S
Dry sea-ice albedo	α_{Id}	H79 ice strength const.	P^*
Wet sea-ice albedo	α_{Iw}	H79 ice strength const.	C^*
Dry snow albedo	α_{Sd}	Ice-ocean drag	c_{IO}
Wet snow albedo	α_{Sw}	Lead closing param.	H_0

(b) Optimized parameters in C2

Ocean albedo	α_O	Therm. conductivity snow	k_S
Visible sea-ice albedo	α_{Iv}	H79 ice strength const.	P^*
Infrared sea-ice albedo	α_{Ii}	H79 ice strength const.	C^*
Visible snow albedo	α_{Sv}	Ice-ocean drag	c_{IO}
Infrared snow albedo	α_{Si}	Redistribution ridged ice	μ

(c) Optimized parameters in C3

Ocean albedo	α_O	Therm. conductivity snow	k_S
Sigma coeff. for ice albedo	R_I	H79 ice strength const.	P^*
Sigma coeff. for snow albedo	R_S	H79 ice strength const.	C^*
Sigma coeff. for pond albedo	R_P	Ice-ocean drag	c_{IO}
Melt pond shape	δ_P	Redistribution ridged ice	μ

Parameter types

Radiation	Sea-ice thermodynamics
Sea-ice thickness / ITD	Sea-ice dynamics

Table 5.1: Model parameters optimized for each of the three model configurations C1, C2, and C3. The division of the model parameters in four groups reflects the sea-ice model aspect regulated by the parameters. These groups are defined and color-coded as follows: radiation=blue, sea-ice thickness / ITD = gray, sea-ice thermodynamics=green, and sea-ice thermodynamics=red.

is implicitly constrained by the parallel employment of sea-ice concentration observations during the optimization, which compensates, at least to some extent, for the exclusion of the SMOS observations from the optimization.

Following Sumata *et al.* (2019a), sea-ice drift data covering the whole seasonal cycle are obtained by combining three different pan-Arctic low-resolution products: the OSI-405 (Lavergne *et al.*, 2010), the sea-ice motion estimate by Kimura *et al.* (2013), and the Polar Pathfinder Daily 25 km EASE-Grid Sea Ice Motion Vectors, Version 2 (NSIDC Drift hereafter; Tschudi *et al.* (2010); Fowler *et al.* (2013)). OSI-405 is the drift product with the smallest observational uncertainties (Sumata *et al.*, 2014) and therefore, when possible, it is preferred to the others. The estimates by Kimura *et al.* (2013) are used in summer because the OSI-405 temporal coverage is limited to the winter months. The NSIDC Drift data are used to cover a gap left by the other two products during part of 2011 and 2012.

Additionally, the model simulations are compared to other types of sea-ice observations than those employed for the Green's function optimization. As for the northern hemisphere, the southern hemisphere sea-ice concentration is taken from the OSI SAF Global Sea Ice Concentration Climate Data Record v2.0. Starting from 2016, we use the operational extension of the OSI-450, denominated OSI-430-b, for both hemispheres (EUMETSAT Ocean and Sea Ice Satellite Application Facility, 2019). The retrieval of snow depth on top of the sea ice is based on an empirical algorithm that uses passive microwave satellite observations from the AMSR-E (Advanced Microwave Scanning Radiometer; Rostosky *et al.* (2019b)) and AMSR-2 (Rostosky *et al.*, 2019a) sensors, as described by Rostosky *et al.* (2018).

5.2.6   Cost Function

The optimization of the model parameter space leads to modifications of the sea-ice state and, consequently, to a variation of the cost function measuring the mismatch between model results and observations. Studying the cost function represents therefore a useful approach to assess changes in model performance. Before presenting the main findings of our study, we clarify some aspects related to the cost function formulation and interpretation. From a mathematical viewpoint, the cost function F (Eq. 5.9) employed in the assessment of the model performances is the same quadratic cost function that is minimized during the Green's function parameter optimization

(Eq. 5.5):

$$F = \frac{1}{N_o} \sum_{i=1}^{N_o} \frac{(y_i - x_i)^2}{\sigma_i^2} \ , \tag{5.9}$$

where y_i is a single observation with standard deviation σ_i, x_i is the corresponding model value, and N_o the total number of observations. In the context of model performance evaluation, F is computed separately for each observation type at different stages of the parameter optimization procedure (before optimization, after one iteration, and lastly after the second iteration). Assuming that the observations represent accurately the "true" state of the sea-ice cover, a change in cost function (c_F) can indicate an improvement ($c_F < 0$) or degradation ($c_F > 0$) of the model performance. Note that, due to the quadratic nature of the cost function, F= 4 indicates that, on average, the mismatch between model results and observations is equal to 2 ($= \sqrt{4}$) standard deviations of the observations.

Although the initial parameter values of different model setups before the optimization has been made as homogeneous as possible, the pre-optimization cost function values differ inevitably for each model configuration (Fig. 5.3). This behavior depends on multiple factors:

1. The intrinsic ability of a specific model formulation to reproduce the observed state.

2. The quality of the employed atmospheric forcing and its compatibility with each model formulation.

3. The "distance" of each pre-optimization parameter set from the optimized one (i.e. how well the model parameters are manually tuned already).

The relative contribution of these factors is difficult to quantify and can change substantially depending on the variable of interest (e.g. sea-ice concentration, thickness, etc.). An obvious consequence of point 3 is that a configuration far from its optimal state can be optimized more effectively than a configuration closer to it. For being able to evaluate more reasonably a property that we call the model "flexibility"—the extent to which a model configuration can be optimized for a variable—we propose a normalized ver-

sion of c_F for each of the model variables and observations considered:

$$\hat{c}_F = \frac{\sqrt{F_f} - \sqrt{F_i}}{\sqrt{F_i}} \cdot \sqrt{\frac{\min\{F_i^{C1\text{-}E}, \ldots, F_i^{C3\text{-}N}\}}{F_i}} \ , \qquad (5.10)$$

where F_i and F_f are the cost function values respectively before and after the Green's function parameter optimization. The square-roots in Eq. 5.10 are introduced as compensation for the quadratic nature of the cost function. In practice, the normalized formulation $\hat{c}_F$ (Fig. 5.3; gray percentages) has the effect of reducing the cost function change in those configurations that start further away from the optimal state before the optimization, providing a suitable metric for assessing the flexibility of the model configurations.

5.3 Results

5.3.1 Sea-ice concentration and position of the ice edge

The Green's function parameter optimization improves the model representation of the sea-ice concentration for each of the six configurations considered (Fig. 5.3; top-left). The C3 setup shows better performances than C1 and C2 both under ERA5 and NCEP atmospheric forcing, suggesting that a more complex formulation of the sea-ice model is beneficial for accurately simulating this variable. In the Icepack setups C2 and C3, the employment of the NCEP forcing leads to better results than ERA5 in terms of the absolute values of the cost function. In contrast, the cost function values of the optimized C1 configurations are comparable under ERA5 and NCEP forcing. Overall, the C1 setup shows higher flexibility, and it is capable of compensating more effectively for differences in boundary conditions.

Simulating correctly the sea-ice edge position is a requirement for every modern sea-ice model. Because the definition of the ice edge position is based on the sea-ice concentration, one might expect the parameter calibration technique based on sea-ice concentration observations to also improve the representation of this feature. This assumption is reasonable, with one caveat: the observational uncertainties of the sea-ice concentration are largest in the vicinity of the ice edge, slightly reducing the weight of these key regions on the total cost function and prioritizing the optimization of pack ice locations, where however the agreement between model and observations is generally already good. Here we analyze the correctness of the sea-ice edge
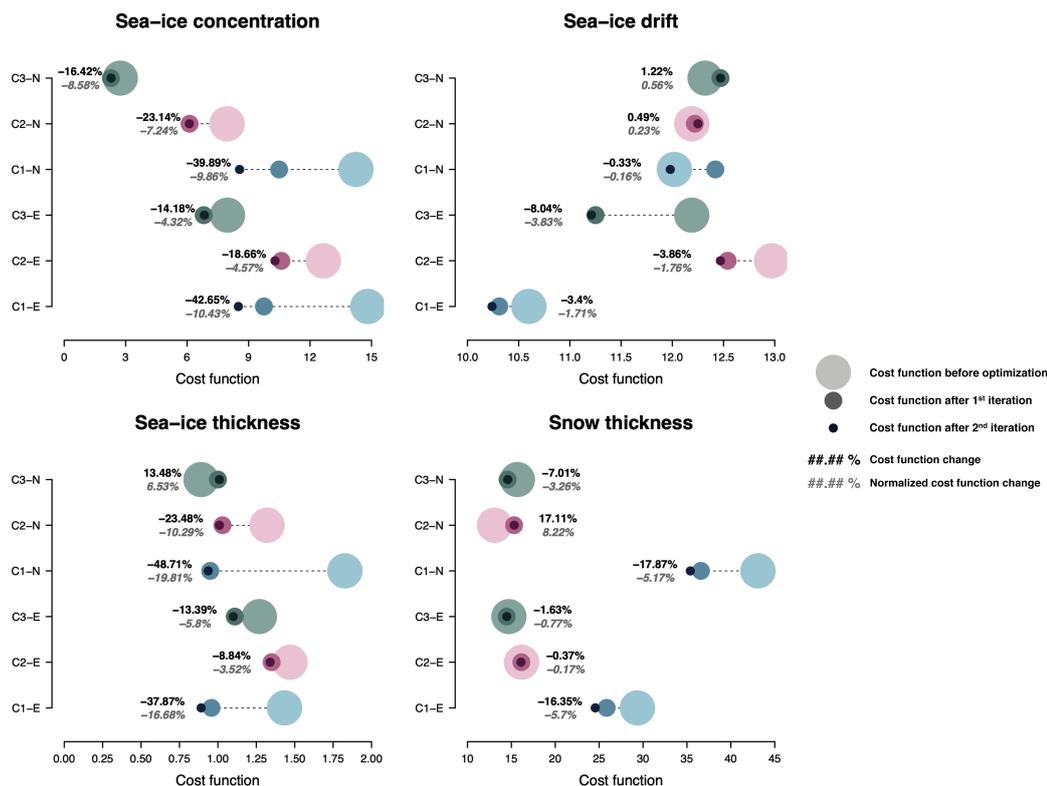
Figure 5.3: Cost function values for the period 2002–2015 at the three stages of the Green's function parameter optimization (x-axis). The cost function measures the average mismatch between the state of six model configurations (y-axis) and four observational products in the Arctic region: sea-ice concentration, drift, thickness, and snow thickness (only the first three observation types are used in the Green's function optimization). The suffixes "-E" and "-N" indicate the employment of the ERA5 and NCEP atmospheric reanalysis used to force the three model setups C1, C2, and C3, respectively. The percentages in black font indicate the cost function change c_F induced by the optimization. The percentages in gray font refer to $\hat{c_F}$, the normalized the cost function change.

position based on two metrics, the Integrated Ice Edge Error (IIEE), and the Absolute Extent Error (AEE; Goessling *et al.* (2016b)), a component of the IIEE (Fig. 5.4). The AEE is defined as the absolute difference in sea-ice extent between model and observations. However, two different configurations of the sea-ice edge can lead to the same sea-ice extent, hence to an AEE = 0. The IIEE is designed to overcome this issue and penalizes situations where sea ice is misplaced in the model simulations compared to the observations.

In terms of IIEE and AEE, the ranking of the six optimized model configurations for the Arctic (Fig. 5.4; top row) confirms what emerges from the analysis of the sea-ice concentration cost function: the C3-N configuration performs best while the C2-E configuration performs worst, exhibiting
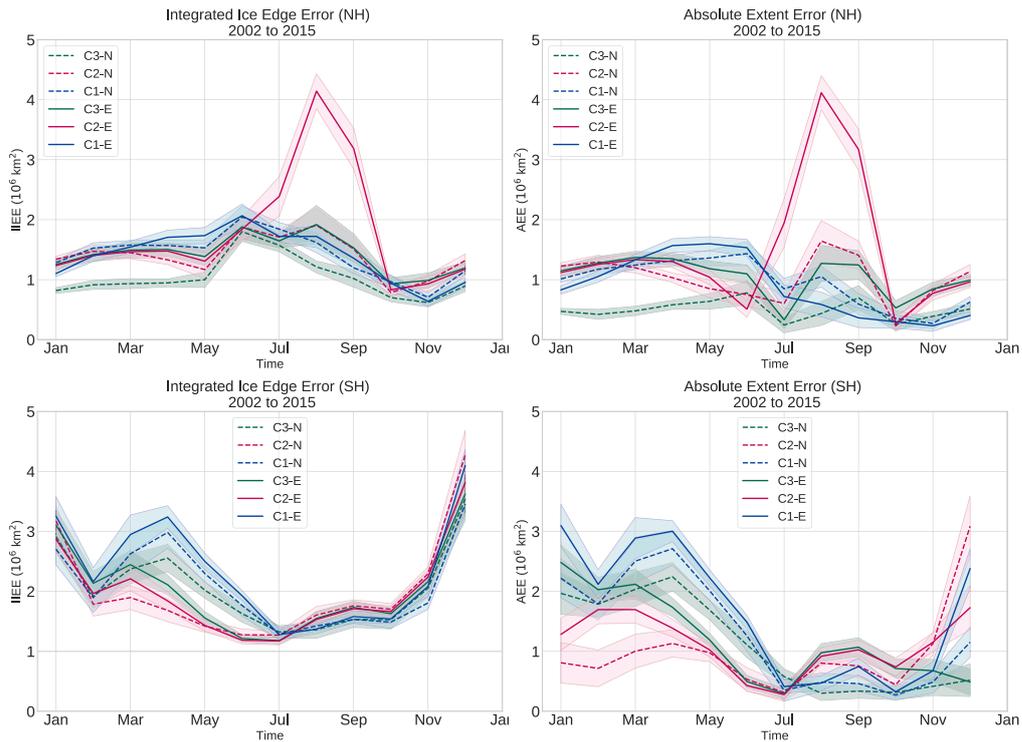
Figure 5.4: Seasonal variation of the northern hemisphere (top) and southern hemisphere (bottom) Integrated Ice Edge Error (IIEE) and Absolute Extent Error (AEE) for six optimized model configurations (C1-E to C3-N) averaged over the period 2002–2015. The IIEE and AEE are computed based on the monthly median ice edge, which is defined as the 15% contour line of the sea-ice concentration. The shading indicates the ∼95% confidence intervals, based on standard errors obtained from the fourteen individual monthly values.

an error peak in summer for both the IIEE and AEE. This error is caused by a strong sea-ice underestimation. Overall, the NCEP forcing leads to a better sea-ice edge representation than ERA5. In all the configurations, both the error magnitude and its variability are largest in late spring and in early summer, while lowest during the winter months. This might suggest a better representation in the model of the physical processes regulating the sea-ice freeze-up compared to those regulating its melting. Furthermore, the 2 m temperature transition across the sea-ice edge in the atmospheric forcing is much sharper during the freezing season than during the melting season, allowing little freedom to the sea-ice model where to place the sea-ice edge and leading to better winter performances.

The ice-edge position analysis has been repeated for the Southern Ocean (Fig. 5.4; bottom row), whose sea-ice observations have not been considered in the parameter optimization. The results evidence some similarities with

the Arctic: the IIEE and AEE are largest during the melting season and lowest in winter when the sea-ice extent reaches its maximum. As for the Arctic, the six configurations exhibit a larger error spread during the summer months. The ranking of the model setups in terms of IIEE and AEE changes substantially in the hemispheres. In Antarctica, the C2 setup, which had the worst performances in the Arctic, exhibits the lowest IIEE and AEE from February to June, followed by the C3 and C1 setups. The situation is inverted from July to January when the differences among the model configurations are however much smaller. Overall, in the Southern Ocean, the Icepack setups C2 and C3 perform comparably or better (depending on the season considered) than the standard FESOM2 formulation C1.

5.3.2 Sea-ice thickness

The analysis of the sea-ice thickness cost function reveals similar performance of different model configurations (Fig. 5.3; bottom-left plot). The cost function values around 1 indicate that, on average, the mismatch between model results and observations is of the same magnitude as the observations uncertainties. After optimization, the model setup C1 exhibits slightly better performance than the C2 and C3 for both atmospheric forcings. Coincidentally, C1 is also the model setup that benefits more from the parameter optimization, with the C1-E and C1-N configurations showing respectively a $\sim -17\%$ and $\sim -20\%$ normalized cost function change. In contrast, the C3-N configuration, which ranks first before optimization, is negatively affected by the optimization and exhibits a $\sim 6\%$ normalized cost function increase.

The model simulations have been compared to three distinct sea-ice thickness observational products (Fig. 5.5): the Envisat and CryoSat-2 products, which target the thicker sea-ice ($>$1m) for different periods, and the merged CryoSat-2/SMOS product, which combines the capability of the SMOS sensor to detect thin sea-ice with the CryoSat-2 measurements in thicker regions. When compared to the observations, the performance of the model configurations changes slightly depending on the choice of the observational product. The Envisat and CryoSat-2 comparison reveal a general underestimation of the average sea-ice thickness by all the model configurations (Fig. 5.5; upper and middle plot). To a certain extent, this underestimation is a consequence of the absence of essentially all thin sea-ice from these observational products, while the thin ice is still present in the model simulations and can be included

in the average thickness computation if the spatial distribution of the sea-ice thickness is different in model simulations and observations. In contrast, the CryoSat-2/SMOS measurements provide a more complete picture of the sea-ice thickness up to the ice edge. It is therefore more compatible with the model results and allows a more robust comparison. Consequently, the agreement between this observational product and the model results is better (Fig. 5.5; bottom plot).

Overall, the sea-ice thickness discrepancies among the optimized model configurations are moderate: on average 25cm, and up to 60cm (Fig. 5.5). The average sea-ice thickness of different configurations tends to converge towards the end of the freezing season, while the spread is slightly larger at its beginning.The results evidence wider discrepancies in terms of model setups than in terms of the atmospheric forcing employed, with C1 having on average a thicker sea-ice cover than C3 and C2. All the model configurations represent fairly well the observed inter-annual variability and the seasonal cycle. For example, both the model simulations and the observations coherently indicate a relatively low sea-ice thickness over the periods 2012–2013 and 2016–2018, and relatively thick sea-ice in 2014–2015. Overall, the model performance in terms of sea-ice thickness is generally better than that of most of the global oceansea ice reanalyses from the Ocean Reanalyses Intercomparison Project (ORA-IP) analyzed by Uotila *et al.* (2018) and Chevallier *et al.* (2017). Note that most of the models analyzed in ORA-IP assimilate sea-ice concentration and/or sea-surface temperature, in addition to other non sea-ice variables.

5.3.3 Sea-ice drift

The sea-ice drift is the model variable for which the parameter optimization procedure is least successful, with a normalized cost function change of on average $\sim -1\%$, and for which the cost function values of different model configurations are most similar (Fig. 5.3; upper-right plot). This behavior can be explained by the fact that the formulation of the dynamic solver has an effect on the simulated sea-ice velocity at least as large (if not more) as the employment of different atmospheric boundary conditions, of sea-ice rheology, and of ice-ocean dynamical interactions (Losch *et al.*, 2010). In this respect, all the model configurations considered here share the same EVP solver for the sea-ice momentum equation, which constrains substantially the model behavior, and which cannot be calibrated through the optimization
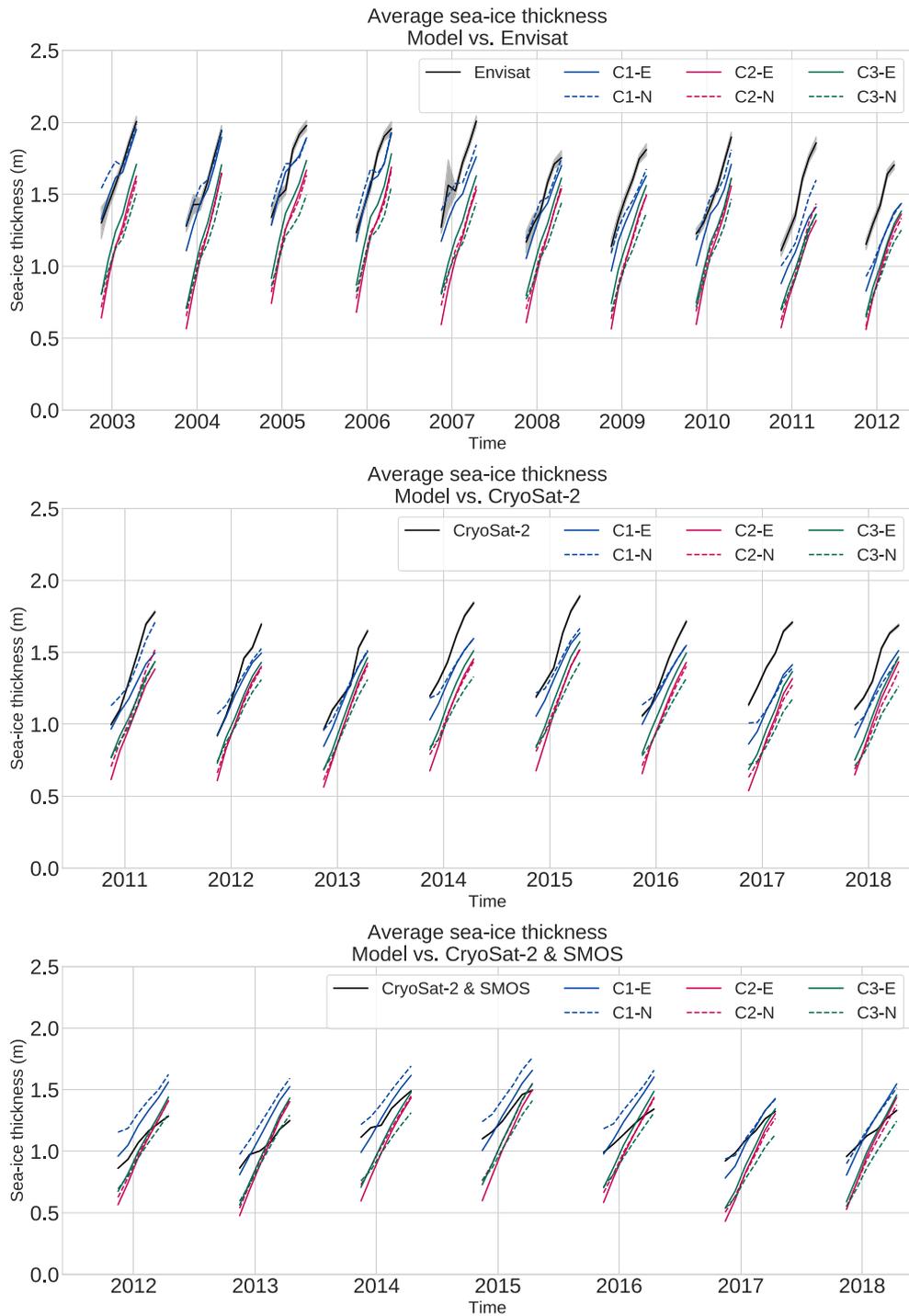
Figure 5.5: November to April average sea-ice thickness for six model configurations (C1-E to C3-N) and for the Envisat (top plot), CryoSat-2 (middle plot), and CryoSat-2/SMOS (bottom plot) satellite observations. The ~95% confidence intervals of the observations are indicated by the gray shading (not visible for CryoSat-2 and CryoSat-2/SMOS), based on 2 standard deviation of the average sea-ice thickness computed through error propagation assuming spatially uncorrelated uncertainties (which is not necessarily the case). The model results have been restricted to the locations within the satellites orbits (< 81.45°N for Envisat and < 87°N for CryoSat-2) where monthly observations are available.

of model parameters. The remaining variability of model performances in terms of sea-ice drift appears to be linked to the choice of the atmospheric forcing. The sea-ice drift optimization is effective only for configurations running under the ERA5 atmospheric forcing, which features a cost function reduction. In contrast, the optimization impact on the configurations running under the NCEP forcing is very small. The poor sea-ice drift performance of C2-E is caused by the summer biases affecting the sea-ice concentration and thickness described in the previous sections.
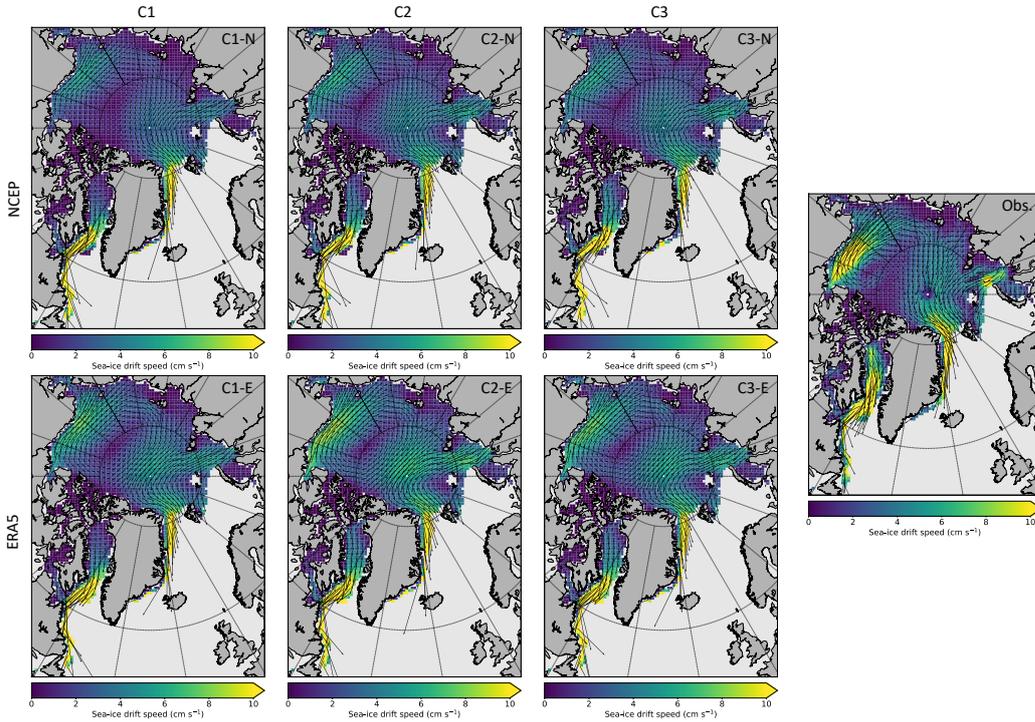
Figure 5.6: April 2015 monthly averaged sea-ice drift speed of six model configurations (C1-N to C3-E) and of the OSI-405 observations.

The simulated sea-ice drift represents well the observed spatial features of the sea-ice circulation in the Arctic, as evidenced by the case study in Fig. 5.6. Here, we limit our analysis to a single month (April 2015) because averaging the sea-ice drift over multiple months and/or years could lead to the cancellation of compensating errors. The anticyclonic circulation in the Beaufort Sea is well represented, as well as the meandering transpolar drift, and the sea-ice export through Fram Strait and the Baffin Bay. The model drift fields are overall smoother and less detailed than the observed drift field. This is caused partially by the finite resolution of the atmospheric forcing and partially by shortcomings of the numerical implementations of the sea-ice

model. A clear aspect that emerges from all the simulations is that the sea-ice
in the model is generally slower than the observations, particularly where the
drift is faster (e.g. cost of Alaska, Baffin Bay, and Kara Sea). This feature
is also evident in Fig. 5.7, which is largely dominated by a positive bias.
However, the ERA5 configurations tend to overestimate the speed of slow
sea-ice ($\mathbf{v}_{ice} < \sim 5$ cm s^{-1}), which results in a too strong sea-ice recirculation
from the transpolar drift into the Beaufort gyre Fig. 5.6. Such a feature is
better captured by the NCEP configurations, whose levels of performance
remain nevertheless worse than ERA5 over most of the Arctic domain.
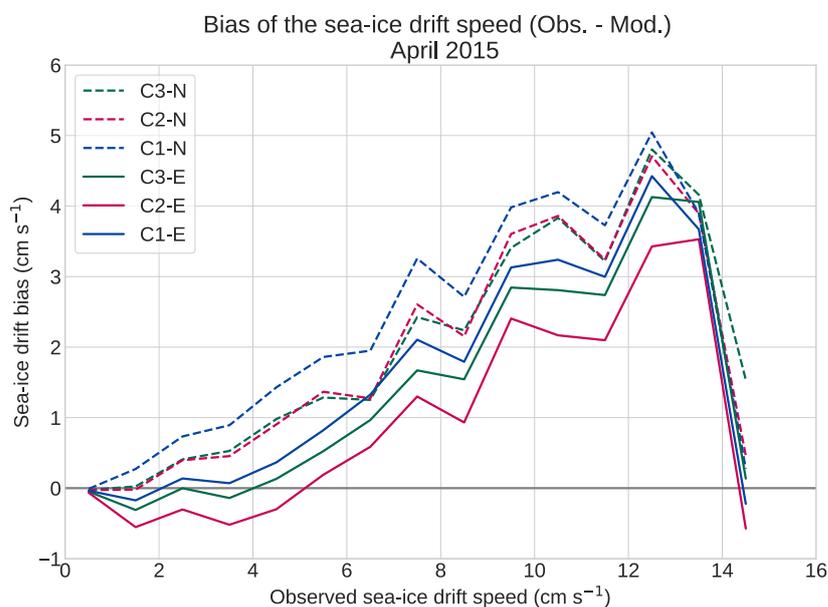
Figure 5.7: April 2015 sea-ice drift speed bias (observation – model; y-axis) for
six model configurations (C1-N to C3-E) as function of the of the observed OSI-
405 sea-ice drift speed (x-axis). The plot is constructed by dividing the observed
sea-ice drift speed in equally spaced intervals of width 1 cm s^{-1}, for which the
corresponding bias values are grouped and averaged. We do not consider observed
sea-ice speeds $\mathbf{v}_{ice} > 15$ cm s^{-1} because of the low number of observational points
and of the consequent low significance of the results.

5.3.4   Snow thickness

Although snow thickness winter observations have not been employed in the
Green's function optimization procedure, the analysis of its cost function
gives an interesting insight into the performances of the analyzed model
configurations concerning this variable. Fig. 5.3 (bottom right plot) shows
two distinct behaviors for the Icepack setups C2 and C3, and for the standard

FESOM2 setup C1. The performance of the latter is worse than that of C2 and C3, before and after the parameter optimization procedure, and regardless of the employed atmospheric forcing. At the same time, C1 is the only setup on which the Green's function optimization has a positive impact, suggesting again greater flexibility of this setup compared to the other two. The C1 snow thickness improvements are likely linked to a better-simulated sea-ice concentration, which presence it mandatory for the accumulation of the precipitated snow.
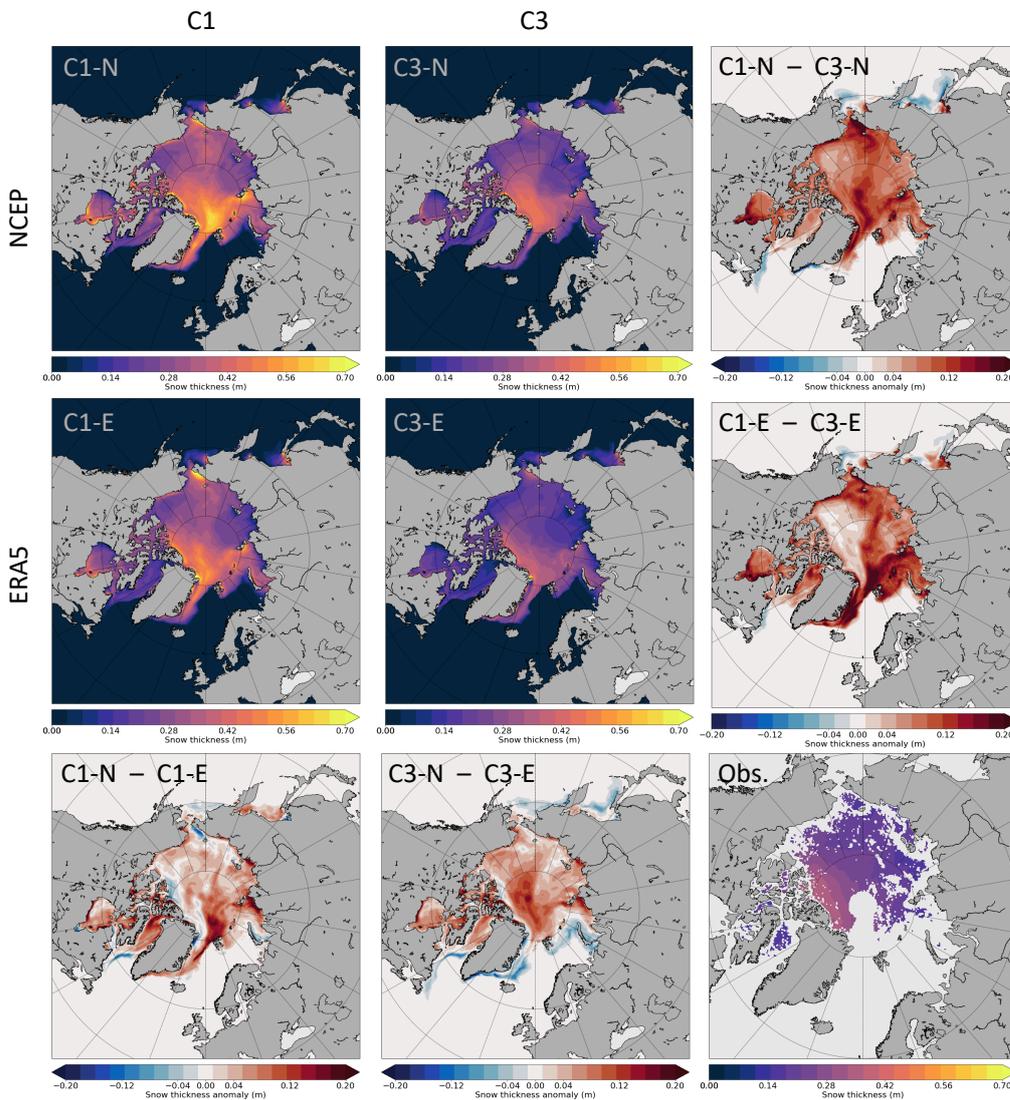
Figure 5.8: April snow thickness and snow thickness anomalies averaged over the period 2002–2015 for four configurations: C1-N, C1-E, C3-N, and C3-E. The C2 setup has not been displayed because its results in terms of snow thickness are very similar to the C3 setup. The April snow thickness observations averaged over the same period are mapped in the bottom-right corner of the panel.

Discrepancies in snow precipitation between different atmospheric reanalysis can be due to the different atmospheric models, data assimilation techniques, and observations used for the production of the reanalysis. Barrett *et al.* (2020) show that this is also the case in the Arctic, where the snow precipitation is higher in the NCEP products compared to ERA5. In this respect, our results are in good agreement with the previous studies: the snow over sea ice in the ERA5 configurations is thinner than that in the NCEP configuration (Fig. 5.8; bottom row). Furthermore, the snow in the C1 setup is overall thicker than that in C2 and C3 for both forcing products (Fig. 5.8; right column). This is likely due to the ridging parameterization adopted in Icepack, which assumes that a fraction of the snow that participates in the ridging (50% in our setups) is lost in the ocean, where it melts eventually. A comparable snow sink is missing in the standard FESIM formulation, hence the thicker snow layer. The observed snow thickness lies in between the NCEP and ERA5 configurations of the C2 and C3 setups. These exhibit comparable cost function values, attributable however to model biases of opposite sign, positive for NCEP and negative for ERA5.

5.4 Discussion

5.4.1 Optimized parameters

Fig. 5.9 compares five optimized parameters for the six model configurations analyzed here. Overall, differences in model formulation appear to have a larger impact on optimized parameter values than differences in atmospheric forcings. Some of the parameters vary more coherently than others. For example, the optimized ice-ocean drag c_{IO} values are systematically larger than the control, for all the setups. In this respect, our results are in good agreement with Sumata *et al.* (2019b), which finds an optimized c_{IO} value of 0.00847 for the NAOSIM model, but they differ from the optimal estimates of Ungermann *et al.* (2017) (0.00664 for the MITgcm model) and Massonnet *et al.* (2014) ([0.00294, 0.00378] for the NEMO-LIM3 model, also associated to a much lower value of P^* compared to our simulations). All the previously mentioned models run with the NCEP atmospheric forcing.

The calibration of P^* leads to minor parameter changes for the setups C1 and C3. In contrast, P^* is reduced in both configurations of the C2 setup. This parameter reduction is likely a consequence of the negative thickness

and concentration biases of this setup, which is mitigated in part by reducing the sea-ice strength. A less stiff sea-ice cover leads to more ridging in winter and, in turn, to an increase of the sea-ice volume and extent. A similar consideration can be made for the relatively high values of C^* for the C2 configurations, which also concur with a reduction of the sea-ice strength. Only the C1-E configuration shows a pronounced reduction of C^*, which implies an increase of the sea ice strength.

The ocean albedo exhibits two different types of behavior: $\alpha_O =\sim 0.085$ for the Icepack setups while $\alpha_O =\sim 0.042$ for the standard FESOM2 setup, a factor-two difference. Note that the treatment of the ocean albedo is equally simplistic in all the model setups considered (no dependency on the incident angle of solar radiation). Therefore, differences in model formulations with respect to this parameter cannot explain the dual behavior observed. Such a feature might be likely linked to different assumptions in the model implementation of the processes regulating the lateral melting of sea-ice, which is impacted by the ocean surface temperature and in turn influenced by α_O. The reader should consider that α_O is the only parameter chosen for the calibration with a substantial impact on the global ocean rather than only on the polar regions. Although both values fall inside the admissible observational range (Jin *et al.*, 2004), a choice in one or the other direction could impact and possibly degrade the model performances concerning the ocean temperatures outside the Arctic. Such a parameter should therefore be manipulated with extreme care, and it could be optimized much more effectively by constraining the optimization procedure with sea-surface temperature observations. Nevertheless, in uncoupled setups varying α_O has a limited effect on the simulated sea surface temperature because this variable is also constrained by the near surface temperature from the atmospheric forcing. Such an assumption does not hold in fully coupled setups, where a correct ocean albedo formulation becomes crucial.

Urrego-Blanco *et al.* (2016) describe the prime role of the snow thermal conductivity k_S in regulating the winter growth of sea-ice in the CICE model. A large k_S allows more heat transfer from the ocean to the atmosphere during winter, enhancing the bottom growth of sea ice and leading to a thicker sea-ice cover. The opposite is true for a low k_S. Apparently, the Green's function parameter optimization effectively exploits this mechanism to reduce the sea-ice thickness biases in the model configurations (Fig. 5.3; bottom-left plot): the Icepack C2-E, C3-E, and C2-N configurations—negatively biased before

the optimization—see an increase of k_S. The C1-E and C1-N configurations, both positively biased in snow and sea-ice thickness before the optimization, experience a reduction of k_S. C3-N, which before the optimization exhibits the best sea-ice thickness correspondence between model results and observations, is the configuration with the lowest k_S change.
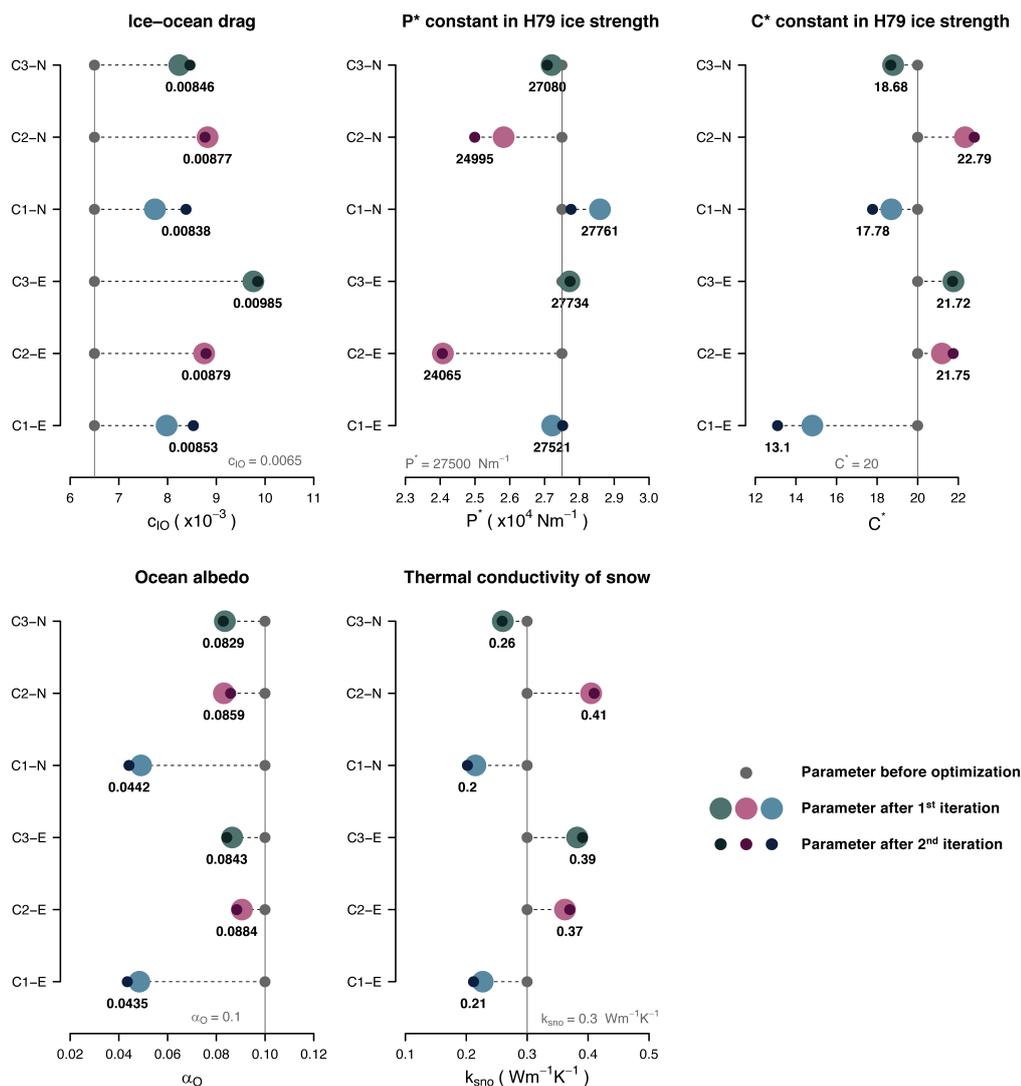

Figure 5.9: Model parameters (x-axis) at three stages of the Green's function parameter optimization. The control values of the parameters are indicated in gray. For each setup, the numerical value of the optimized parameters is reported in black below each point. Only the parameters common to the C1, C2, and C3 model setups are shown. The suffixes "-E" and "-N" indicate respectively the employment of the ERA5 and NCEP atmospheric reanalysis used to force the three model setups.

5.4.2 Shortcomings of the parameter optimization

The first unsatisfactory outcome of the parameter optimizations regards the very weak sea-ice drift performance improvement (Sec. 5.3.3) compared to that of sea-ice concentration and thickness. This behavior is likely caused by our choice of excluding the ice-atmosphere drag coefficient c_{IA} from the optimization. While the results by Massonnet *et al.* (2014) indicate that the optimization of P^* and c_{IO} is sufficient for improving the sea-ice drift, we observe that this does not hold any longer when multiple parameters are optimized simultaneously and constrained with multiple observational types, at least in the FESOM2 model. As a consequence of a slower sea-ice drift in our simulations, an over-optimization of thermodynamic and radiative processes (e.g. enhanced formation of new sea-ice or melting) might have occurred to compensate for the reduced sea-ice transport outside the Arctic. In future studies and tuning exercises, we will revisit this hypothesis by explicitly optimizing c_{IA} in the attempt to better constraining the drift field.

A second aspect that deserves some discussion concerns the overall poor performance of the C2 model setup, and particularly of C2-E. This configuration exhibits a strong negative bias in sea-ice concentration and thickness during summer, which consequently impacts the model performance also in terms of sea-ice drift and snow thickness. This bias likely results from a misrepresentation of the sea-ice radiative processes in the model and, once more, it might be due to an unwise choice concerning the parameters for the optimization. The C2 setup employs the CCSM3 radiation scheme, in which, as described in Sec. 5.2.2, the sea-ice and snow albedo values are split into a visible and an infrared component with a thickness and temperature dependence. These four albedo values have been optimized in the present study (Sec. 5.1). However, the model parameters that regulate the thickness and temperature dependence of the albedo have not been optimized, leading to a poor representation of the melting processes. We observe that both the simpler radiation scheme employed in C1 and the complex delta-Eddington radiation formulation used in C3 respond to the parameter optimization better than the CCSM3 scheme, likely because they can be constrained with fewer model parameters.

5.4.3 Computational costs

The increased complexity of the FESOM2 extended sea-ice model comes with a non-negligible price in terms of computational costs. Fig. 5.10 shows that the sea-ice computations of the Icepack setups C2 and C3 are approximately four times slower than C1, the simpler standard FESOM2 setup. This behavior was expected and caused partially by the more detailed formulation of Icepack thermodynamics, but primarily by the growing number of tracers needed to describe the sea-ice state. These tracers need to be advected separately by the FE-FCT scheme, which translates into a linear increase of the cost for each additional tracer. Furthermore, a set of tests has been implemented to guarantee the conservation of enthalpy, freshwater, and salinity during the advection process, which further increases the computational requirements. An incremental remapping scheme for the advection of sea-ice tracers similar to that implemented in CICE (Lipscomb & Hunke, 2004), which is conservative and becomes very efficient when the number of tracers is large, will be considered in the future for further reducing the computational cost of the FESOM2-Icepack implementation.

Running FESOM2 with Icepack remains nevertheless feasible, and represents a viable option for future modeling studies with a focus on polar regions. The mesh employed for this study is designed with most of the surface nodes in sea-ice active regions, causing the sea-ice computations to account for a substantial part of the model budget, and thus constituting a rather extreme case if compared to CMIP-type applications. The relative cost of the Icepack computations will be lower in meshes with most of the nodes in non-sea-ice regions. Furthermore, in high-resolution simulations (1km to 4km), the contribution of the EVP solver is expected to become predominant over the advection of tracers, due to the increasing number of sub-cycles needed for reaching a converging solution of the momentum equation. An in-depth investigation of the computing performances of the FESOM2-Icepack model for a broader range of scenarios will be the topic of a future study.

5.4.4 Future prospects for the FESOM2 sea-ice representation

As described in Sec. 5.2.2, the options offered by Icepack in terms of sea-ice physics go beyond those explored in this study. In particular, future

Cost of the FESOM2 sea–ice computations
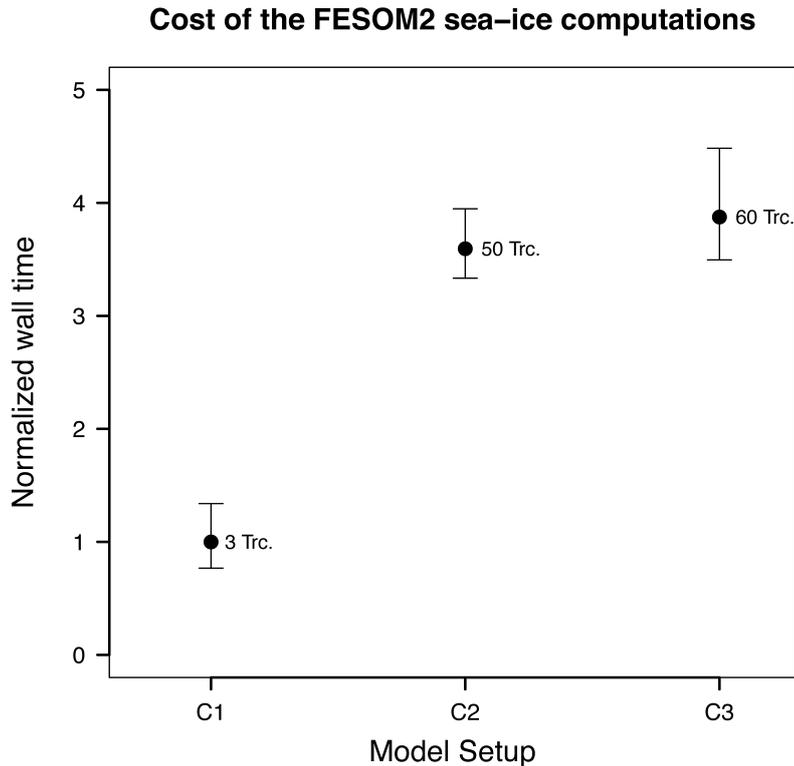
Figure 5.10: Relative computational cost of the sea-ice component of three FESOM2 setups (C1, C2, and C3). The values are normalized by the C1 wall time. All the simulations run on the same machine, with the same computational mesh, and under the ERA5 forcing. The bars indicate the maximum and minimum values registered among the computing CPUs. The number of tracers advected in each setup is also reported.

work will focus on the impact of a highly resolved ITD on the simulated sea-ice thickness and drift (possibly at high spatial resolution), and on the exploration of the floe-size distribution parameterizations. Future FESOM2-Icepack model simulations could also serve as boundary conditions for detailed single-column studies with Icepack in a Lagrangian framework (e.g. Krumpen *et al.* (2020)), allowing to retain a high physical consistency between the driving model and the single-column model.

Most of the model configurations here analyzed show a minimum in AEE in July (Fig. 5.4; top right), suggesting that the IIEE is mostly caused by sea-ice misplacement rather than by a wrong representation of the sea-ice extent. This behavior could in part reflect the fact that our model cannot simulate the processes leading to land-fast sea-ice formation, both in its standard formulation and with Icepack. The absence of this persistent sea-ice type

impacts the detachment location of the pack ice from the Arctic coastline
and in turn the correctness of the sea-ice edge position for this month. Model
formulations that enable, to a certain extent, the simulation of land-fast sea
ice in shallow seas already exist (Lemieux *et al.*, 2015, 2016) and will be
considered for future versions of the FESOM2 model.

The FESOM sea-ice and ocean model plays a central role in the climate
modeling and forecasting activities at the Alfred Wegener Institute (AWI),
and it is part of different versions of the CMIP6 AWI Climate Model (AWI-
CM; Sidorenko *et al.* (2015); Rackow *et al.* (2016); Sidorenko *et al.* (2019);
Semmler *et al.* (2020)). In this respect, we plan to couple the new FESOM2-
Icepack setup to the latest climate model configuration under development
at AWI, which uses the open-source version of the Integrated Forecast Sys-
tem (OpenIFS) as the atmospheric model. The availability of a more detailed
sea-ice description in a fully coupled setup will enable a better understanding
of the interactions between a warming atmosphere and sea ice. At the same
time, the new coupled configuration will allow to perform sea ice-oriented
climate modeling studies (e.g. Zampieri & Goessling (2019)) under more
physically consistent assumptions. Finally, FESOM2-Icepack will be inte-
grated in the Seamless Sea Ice Prediction System (SSIPS; Mu *et al.* (2020))
and thus equipped with the Parallel Data Assimilation Framework (PDAF;
(Nerger & Hiller, 2013)) for assimilating ocean and sea-ice observations with
an Ensemble Kalman Filter.

5.5 Summary and conclusions

This study presented a new formulation of the sea-ice component of the
unstructured-mesh FESOM2 model. The update, which exploits the state-
of-the-art capabilities of the sea-ice single-column model Icepack, improves
the physical description of numerous sea-ice sub-grid processes while retain-
ing a modular structure that enables the user to adapt the sophistication of
the sea-ice model formulation to the requirements of a specific investigation.
Because of this modularity, the new FESOM2 formulation allows to investi-
gate the impact of the sea-ice model complexity on the performance of the
sea-ice simulations under two different atmospheric forcings. Our findings
indicate that the sophisticated C3 setup performs systematically better than
C2 and C1 concerning the Arctic sea-ice concentration and snow thickness,

supporting the hypothesis that an elaborated model formulation leads to a more appropriate representation of the sea ice. However, the results also indicate that the setup ranking that emerges for the sea-ice concentration in the Arctic does not hold in the Southern Ocean, which has not been included in the optimization; here the C2 setups perform best. The current generation of atmospheric forcings and sea-ice/ocean models is therefore still not fully balanced and fails to guarantee an adequate representation of the sea ice in both hemispheres simultaneously. For other variables, model complexity appears to play only a marginal role in defining the quality of sea-ice simulations. This is the case for sea-ice thickness and drift, for which the differences between the various FESOM2 configurations are small and independent of model sophistication. Finally, we find that the simple C1 setup responds better to the optimization procedure, showing larger improvements compared to C2 and C3, and thus suggesting that a less complex model can be tuned more effectively.

In addition to the model formulation, also the choice of the atmospheric forcing product influences substantially the sea-ice simulations. Concerning the sea-ice concentration, the Icepack setups C1 and C2 perform much better when forced with the NCEP product compared to ERA5, both in the Arctic and in the Antarctic. The C1 setup exhibits similar results for NCEP and ERA5 in the Arctic, while the NCEP forcing outperforms ERA5 in the Antarctic. The opposite is true for the sea-ice drift and the snow thickness variables, which benefit from the employment of the ERA5 product instead of NCEP. In summary, both the atmospheric forcing products here analyzed have strengths and weaknesses that should be considered when employing them to force sea-ice and ocean simulations.

The results of this study are valid for sea-ice/ocean only simulations, where the atmospheric conditions are prescribed from reanalysis products. Some of the findings might not hold in a fully coupled framework, where the atmosphere responds both thermodynamically and dynamically to sea-ice and ocean changes. A similar study could be implemented in a fully coupled configuration by optimizing the climatological sea-ice state of the model using the observational climatology as constraint. We plan to perform such a study for our modeling framework once the FESOM2-Icepack setup will be coupled to the OpenIFS atmospheric model.

We conclude this manuscript by underlining, once more, the importance of the semiautomatic parameter calibration for this study. Without the two

cycles of Green's function optimization, our results would have conveyed a rather different message, erroneously indicating that the Icepack configurations perform systematically better than the standard FESOM2 model for most of the variables considered (Fig. 5.3; large circles). The systematic optimization of the sea-ice parameters is certainly a time-consuming operation that requires a non-negligible amount of computing resources. Nevertheless, we recommend this approach, in some form, in future studies that aim to assess advances in the field of sea-ice modeling to guarantee a fair evaluation of sea-ice models.

6. Summary, conclusions and outlook

6.1 Summary and conclusions

The present section summarizes the main findings of the studies presented in the previous chapters of this dissertation. In doing so, answers to the scientific questions formulated in the introduction are explicitly provided.

Chapter 2 investigates the response of the sea ice and of the climate system to the AIM geoengineering strategy, which aims at intervening on the Arctic icealbedo feedback by mitigating the effects of global warming on polar regions. Based on CMIP5-type climate simulations under a RCP 8.5 scenario forcing and with a modified sea-ice description to account for the AIM effects, I show that keeping the latesummer sea ice cover at the current extent for the next ~ 60 years is in principle possible, from a physical point of view, if at least $\sim 10^7$ AIM devices are deployed in the Arctic (Question **Q1**, page 9). This estimate concerning the number of devices is conservative as the model implementation of AIM is much more flexible in terms of the device positioning and activation than a hypothetical real-world deployment. Compared to scenario simulations without AIM, the increased sea-ice extent and volume generate a significant summer cooling of approximately 1.3 K during the 40-year period 2021–2060 north of the Arctic circle, and of approximately 1.4 K in 2061–2100. However, this cooling signal is confined over sea-ice covered regions and it is not conveyed to lower latitudes, reducing the capacity of this method to be effective on a planetary scale. Besides, the Arctic experiences substantial winter warming in regions with active pumps, as relatively warm water is directly exposed to the cold winter Arctic atmosphere. Overall, the

global annualmean nearsurface air temperature is reduced by just 0.02 K for
the period 2021–2060 in AIM simulations, indicating that this sea ice targeted geoengineering strategy does not have the potential to mitigate climate
change (Question **Q2**, page 9).

Chapter 3 shifts the focus of this thesis from the area of multi-decadal climate projections to the field of operational sea-ice predictions. In this context, I assess, for the first time, the skill of stateoftheart operational sea-ice
forecasts collected in the database of the Subseasonal-to-Seasonal Prediction
Project. The focus of this study is exclusively on the Arctic region. A new
probabilistic verification metric, which quantifies the accuracy of the sea-ice
edge position in a meaningful way, is employed to evaluate the forecast performances concerning this key parameter for potential forecast users. The
results indicate that the inclusion of a prognostic dynamical sea-ice model
is beneficial in terms of forecast skill compared to diagnostic descriptions of
the sea ice based on observational statistics. The forecast systems feature
a surprisingly wide range of skills, with the best system producing skillful
forecasts up to 45 days in advance, while other systems show no predictive
skill at short weather time scales or even at the forecast initialization. Furthermore, the prediction skill of the more skillful forecast system is highest
in late summer, suggesting that valuable subseasonal sea-ice edge forecasts
can be already provided to stakeholders during the most active time of the
year in terms of Arctic navigation (Question **Q3**, page 10). Despite these
promising results, my findings point to several issues that affect the S2S forecast systems, emphasizing large space for improvements in the future. The
most important of these shortcomings is a generalized large forecast error at
initial forecast times, which calls for additional efforts in improving the techniques for the assimilation of sea-ice and ocean observations into the forecast
systems (Question **Q4**, page 10).

Chapter 4 applies the methodologies presented in chapter 3, using the same
forecast database to the investigation of the Antarctic sea-ice prediction capabilities at subseasonal timescales. Compared to the Arctic, Antarctic forecasts show larger errors and, on average, a 30% lower prediction skill, with
only one system remaining more skillful than a climatological benchmark

up to 30 days lead time (Question **Q3**, page 10). The investigation of the
longitudinal variation of the forecast error shows that the prediction skill
tends to be highest in the West Antarctic sector, with the ECMWF system
being skillful up to 44 forecast days, while it is lower for other Southern
Ocean locations. In general, the majority of the forecast systems show an
overestimation of the sea ice edge position and fail to capture the onset and
development of the melting season. The highest forecast skill is found from
March to May, after the onset of the freezing season. As for the Arctic fore-
casts, large errors at initial time affect all the forecast systems. I link these
errors to an inaccurate assimilation and erroneous model representation of
the dispersed sea ice cover (i.e. low sea-ice concentration) in the marginal
ice zone (Question **Q4**, page 10).

Chapter 5 has the double purpose of presenting the new sea-ice formula-
tion of the unstructured FESOM2 sea-ice and ocean model, and of assessing
the impact of the improved sea-ice description on the model performance in
uncoupled sea-ice/ocean simulations. More specifically, I have equipped the
unstructured global sea-ice and ocean model FESOM2 with a set of physical
parameterizations derived from the single-column sea-ice model Icepack. The
simple 0-layer sea-ice and snow thermodynamics have been replaced with a
set of multi-layer parameterizations that take the enthalpy and salinity of
the ice into account. The new system can simulate prognostic thickness and
floe-size distributions (also jointly), accounting for sea-ice ridging and for
processes regulating the break-up and healing of sea-ice floes. A sophisti-
cated delta Eddington multi scattering solar radiation parameterization and
three prognostic melt-pond schemes are also available. The implementation
of Icepack in FESOM2 has been designed to maintain the modular architec-
ture of Icepack, which allows to easily vary the complexity of the sea-ice de-
scription. To compare fairly eventual improvements or drawbacks associated
with the changing model complexity, I optimized a subset of the parame-
ter space of each tested model configuration by applying a Green's function
optimization technique. The results indicate that a complex model formula-
tion leads to a better agreement between modeled and the observed sea-ice
concentration and snow thickness, while differences are smaller for sea-ice
thickness and drift speed (Question **Q5**, page 11). However, the choice of
the atmospheric forcing also impacts the agreement between simulations and

observations, with NCEP-CFSR/CFSv2 being particularly beneficial for the simulated sea-ice concentration and ERA5 for sea-ice drift speed (Question **Q6**, page 11). Furthermore, the results indicate that the parameter calibration can better compensate for differences among atmospheric forcings and for model deficiencies in a simpler model setting (where sea-ice has no heat capacity) compared to more energy consistent formulations with a prognostic ice thickness distribution.

6.2   Outlook

This thesis is a collection of several scientific studies that deal, in various ways, with the investigation of the Arctic and Antarctic sea-ice systems through the application of complex geophysical models. The topics investigated in the main chapters of this dissertation are quite heterogeneous, helping to formulate some more general considerations concerning this exciting field of science.

The first thread of this thesis concerns sea-ice modeling studies that can be connected to two distinct scientific fields: sea-ice projections at multi-decadal timescales and operational sea-ice predictions at subseasonal timescales. As I argue in Sec. 1.4, the approaches and assumptions for the study of sea ice in these two scientific areas differ substantially on multiple levels, mostly because of the largely different timescales addressed. Nevertheless, the sea-ice modeling tools employed in these two fields retain, probably out of convenience, very similar physical formulations. This aspect calls for some further considerations regarding the appropriateness of the various model formulations for answering some of the specific scientific questions discussed in this thesis.

In terms of climate projections, the coupled model that has been chosen for the study of the AIM strategy resulted, in principle, adequate for investigating the impact of this geoengineering approach on the Arctic sea ice and climate. However, the sea-ice model formulation based on a 0-layer thermodynamics did not allow to explicitly resolve the AIM surface water layer in the model, preventing me from drawing any consideration regarding the AIM impact on the vertical temperature and salinity profiles inside the sea ice. In this respect, a prognostic model description of sea-ice enthalpy and salinity (such as that implemented in Icepack) would have been beneficial and could

have enabled a broader and, to some extent, more complete study.

In the context of sea-ice predictions, we demonstrate that a prognostic description of the sea ice through the employment of a geophysical sea-ice model is beneficial, in terms of predictive skills, compared to a prescribed description of the sea ice based on observational statistics and persistence. Nonetheless, the comparison of different forecasting systems revealed no correlation between the forecast skill and the physical complexity of the forecast sea-ice model, with ECMWF—the best performing forecast system both in the Arctic and the Antarctic—featuring only a very simple description of the sea ice. This suggests that model complexity may be less relevant for subseasonal sea-ice predictions than the quality of the initial conditions and the skill of the atmospheric and ocean models coupled to the sea-ice one. Several scientific arguments go against my previous considerations and findings, and call for increased model complexity in the field of sea-ice prediction. Accounting for more physical processes in our sea-ice models could facilitate the development of new observational operators, enabling the assimilation of more and better observations in our models (Burgard *et al.*, 2020a,b). Furthermore, unresolved physical processes could hold back some unexplored sources of predictability. For example, there is evidence that the melt ponds at the beginning of the summer season (currently not explicitly simulated by the S2S forecast system) are a strong predictor for the September sea-ice minimum (Schrder *et al.*, 2014).

In light of the previous considerations, I propose two simple strategies for dealing successfully with the heterogeneous and rapidly evolving field of sea-ice modeling. These strategies can be summarized by the words *modularity* and *community*, and bring this concluding part of the dissertation in touch with the second thread of the thesis: the study and development of sea-ice models.

Modularity is a key requirement for present and future sea-ice modeling infrastructures, as it allows to tailor the model setup to specific scientific applications, optimizing at the same time the use of computational resources and storage space dedicated to model simulations. In this respect, modularity is the compass I followed when designing the upgrade of the thermodynamic sea-ice component of FESOM2 through the integration of the single-column model Icepack into it. The resulting model formulation grants full flexibility in terms of model sophistication, adoption of passive tracers, number of vertical layers in sea-ice and snow, number of ice thickness and floe

size classes, etc. Furthermore, the new sea-ice thermodynamics is embedded into the FESOM2 numerical environment, which, due to the employment of unstructured computational meshes, features an unique flexibility in terms of the placement of spatial resolution. In conclusion, the implementation of Icepack into FESOM2 brings the model a step closer to having a fully seamless formulation.

Working in close collaboration with the rest of the sea-ice modeling *community* is a second strategy that gives the best chances of developing a healthy and well-performing modeling infrastructure. In this respect, having implemented the Icepack subroutines in FESOM2 by following a modular approach guarantees to receive model updates and corrections in an efficient and timely manner. Furthermore, sharing one modeling infrastructure with hundreds of scientists around the world maximizes synergies, which translates into a more direct applicability and assimilation of major scientific advances, and into a larger reach of our own findings and developments.

As described in Section 5.4.4, the development of the FESOM2-Icepack model implementation will allow sea-ice modelling initiatives at at the Alfred Wegener Institute as well as at collaborating institutions to expand. A better representation of the sea-ice physical processes at the subgrid scale makes the FESOM2 model more suitable for cross-field studies, such as those involving biogeochemistry. At the same time, the new model lays the foundations for a physically more consistent coupling between the sea ice, atmospheric, and oceanic modeling components, that will enable more detailed sea-ice studies in the context of climate and paleoclimate projections. Finally, once the FESOM2-Icepack system will be equipped with data assimilation capabilities, the more accurate sea-ice description will enable robust studies that investigate the impact of the sea-ice model formulation on the quality of sea-ice forecasts at different timescales, allowing to properly tailor the model formulation to the requirements of these applications.

A. Appendix to Chapter 2[1]

A.1 Figures and tables
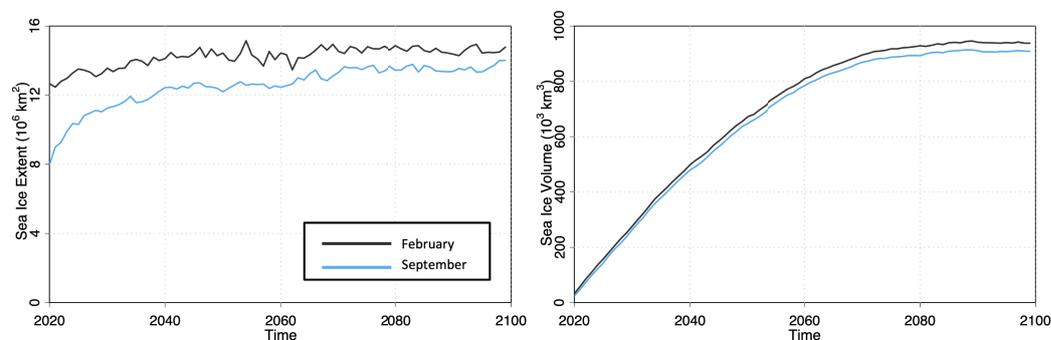
Figure A.1: Evolution of pan-Arctic sea ice extent (left) and volume (right) in February (black curves) and September (blue curves) for theextreme AIM simulation, where a liquid layer is maintained over the whole ice cover during the period 20202100.

---

[1]This appendix contains the supplementary information of the published paper '*Sea Ice Targeted Geoengineering Can Delay Arctic Sea Ice Decline but not Global Warming*' by Zampieri & Goessling (2019)
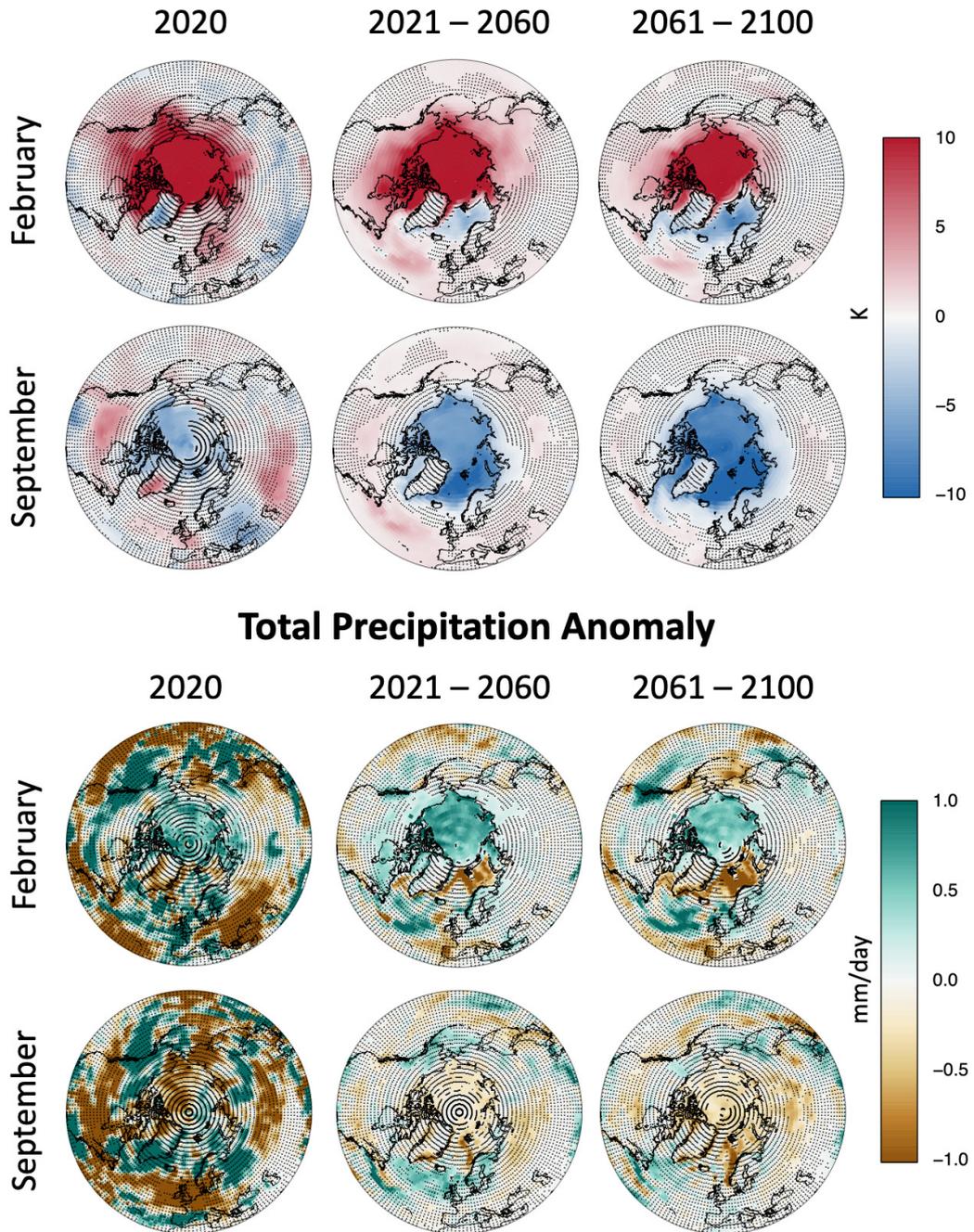
Figure A.2: Top: Near-surface (2 m) temperature anomalies (extreme case simulation minus control ensemble-mean) for the periods 2020, 2021–2060 and 2061–2100. Bottom: total precipitation anomalies (extreme case simulation minus control ensemble-mean) for the periods 2020, 2021–2060 and 2061–2100. Stippling indicates local statistical non-significance of the anomaly at the 95% confidence level according to a two-tailed t-test.
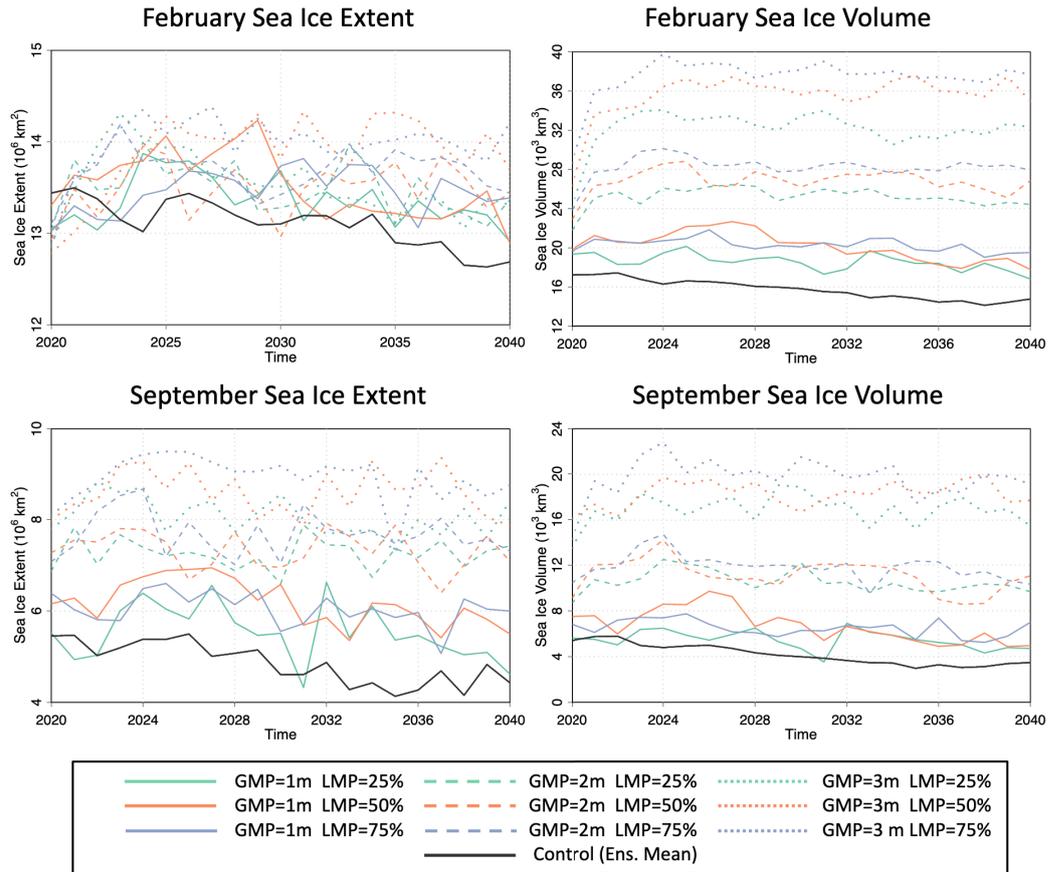
Figure A.3: Evolution of pan-Arctic sea ice extent (left side) and volume (right side) in February (top plots) and September (bottom plots) for the 9 sensitivity simulations (2020–2040) and for the control ensemble-mean (black lines). The combinations of the Global and Local Modulation Parameters (GMP and LMP) in the legend defines each sensitivity simulation.
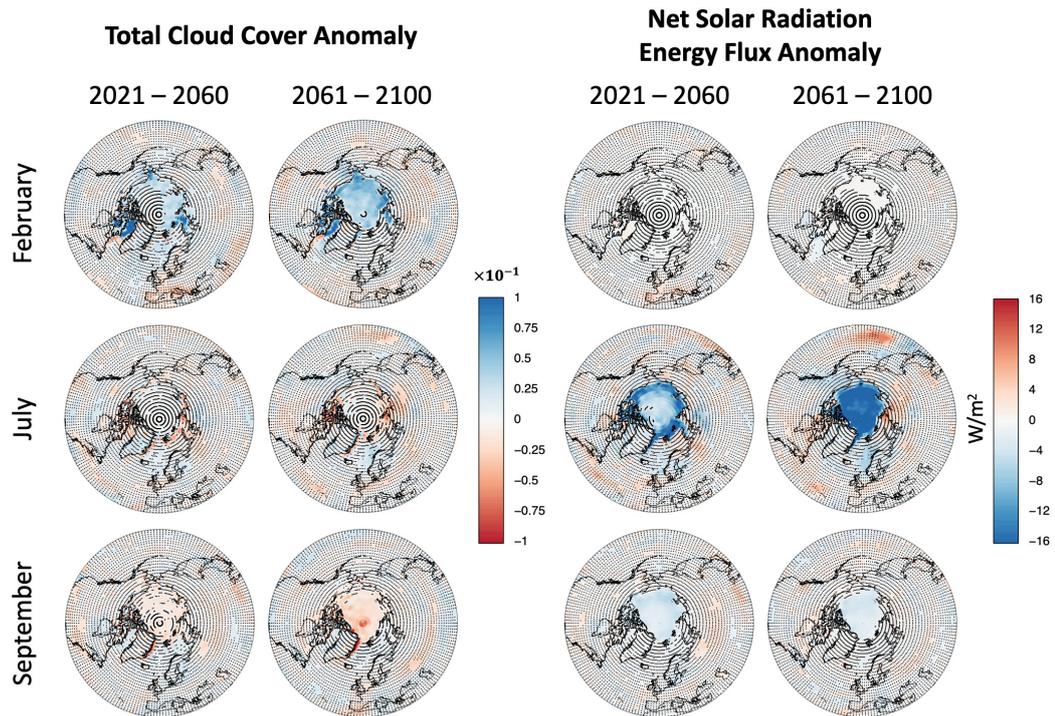
Figure A.4: Left: Total cloud cover anomalies (AIM ensemble-mean minus control ensemble-mean) for the periods 2021–2060 and 2061–2100. Right: Net solar radiation energy flux anomalies (AIM ensemble-mean minus control ensemble-mean) for the periods 2021–2060 and 2061–2100. Stippling indicates local statistical non-significance of the anomaly at the 95% confidence level according to a two-tailed t-test.
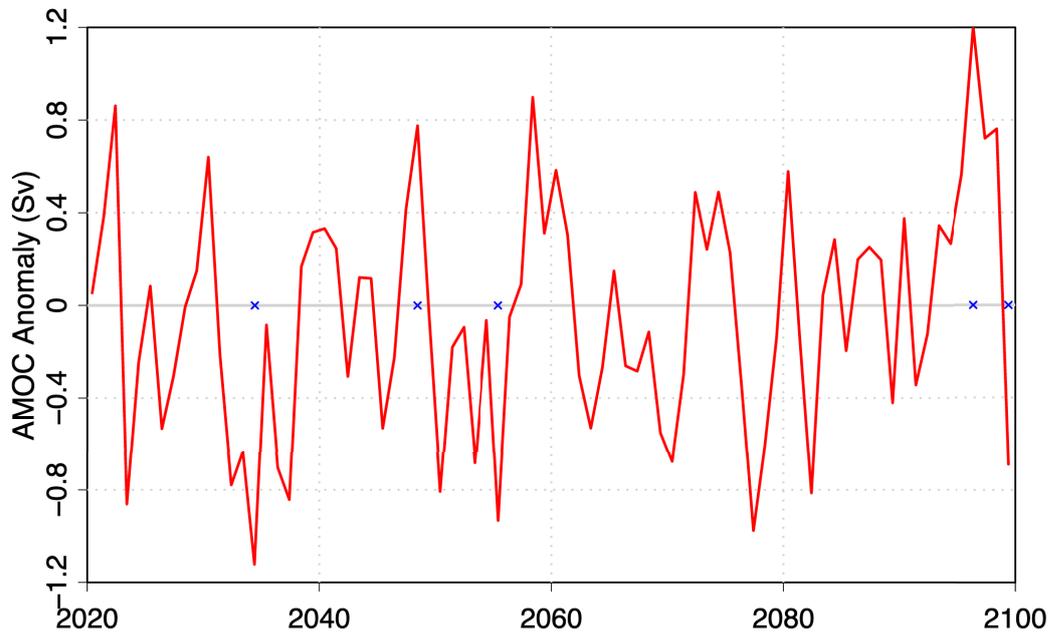
Figure A.5: Annual mean maximum North Atlantic meridional overturning circulation (AMOC) anomaly (AIM ensemble-mean minus control ensemble-mean) computed between 43N and 47N. The blue crosses indicate years with statistically significant anomalies at the 95% confidence level according to a two-tailed t-test.
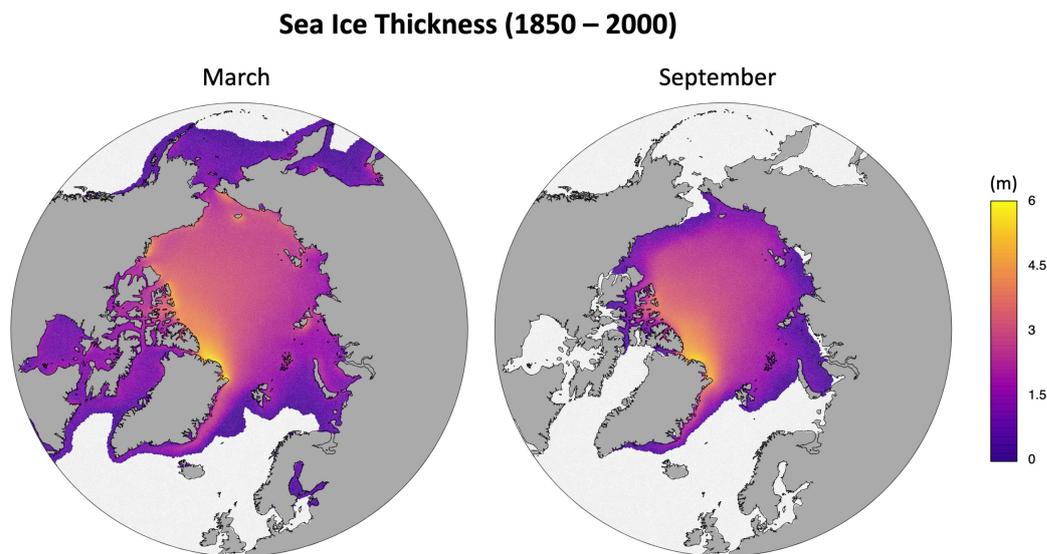
Figure A.6: March and September mean sea ice thickness over the historical period 1850–2000.

March Sea Ice Volume Trends

	Trend (km^3/year)	Intercept ($\times 10^3$ km^3)	r^2	p-value
Control Ens. Mem. 1	-159 ± 4	22.4 ± 0.2	0.95	< 2.2 × 10^{-16}
Control Ens. Mem. 2	-152 ± 5	22.8 ± 0.2	0.93	< 2.2 × 10^{-16}
Control Ens. Mem. 3	-171 ± 5	23.5 ± 0.2	0.95	< 2.2 × 10^{-16}
Control Ens. Mem. 4	-168 ± 4	22.8 ± 0.2	0.96	< 2.2 × 10^{-16}
Control Ens. Mean	-163 ± 2	22.9 ± 0.1	0.99	< 2.2 × 10^{-16}
AIM Ens. Mem. 1	-177 ± 5	34.5 ± 0.3	0.96	< 2.2 × 10^{-16}
AIM Ens. Mem. 2	-181 ± 5	34.6 ± 0.3	0.96	< 2.2 × 10^{-16}
AIM Ens. Mem. 3	-193 ± 6	34.8 ± 0.4	0.95	< 2.2 × 10^{-16}
AIM Ens. Mem. 4	-176 ± 5	34.4 ± 0.4	0.94	< 2.2 × 10^{-16}
AIM Ens. Mean	-182 ± 3	34.6 ± 0.2	0.98	< 2.2 × 10^{-16}

September Sea Ice Volume Trends

	Trend (km^3/year)	Intercept ($\times 10^3$ km^3)	r^2	p-value
Control Ens. Mem. 1	-114 ± 5	7.5 ± 0.2	0.90	< 2.2 × 10^{-16}
Control Ens. Mem. 2	-118 ± 4	8.1 ± 0.2	0.92	< 2.2 × 10^{-16}
Control Ens. Mem. 3	-134 ± 5	8.8 ± 0.2	0.90	< 2.2 × 10^{-16}
Control Ens. Mem. 4	-118 ± 4	7.9 ± 0.2	0.94	< 2.2 × 10^{-16}
Control Ens. Mean	-121 ± 2	8.1 ± 0.1	0.98	< 2.2 × 10^{-16}
AIM Ens. Mem. 1	-139 ± 4	15.8 ± 0.3	0.94	< 2.2 × 10^{-16}
AIM Ens. Mem. 2	-139 ± 6	15.7 ± 0.3	0.90	< 2.2 × 10^{-16}
AIM Ens. Mem. 3	-154 ± 5	16.3 ± 0.3	0.93	< 2.2 × 10^{-16}
AIM Ens. Mem. 4	-144 ± 4	16.2 ± 0.3	0.94	< 2.2 × 10^{-16}
AIM Ens. Mean	-144 ± 2	16.0 ± 0.2	0.98	< 2.2 × 10^{-16}

September Sea Ice Extent Trends

	Trend ($\times 10^4$ km^2/year)	Intercept ($\times 10^6$ km^2)	r^2	p-value
Control Ens. Mem. 1	-8.2 ± 0.4	6.8 ± 0.2	0.83	< 2.2 × 10^{-16}
Control Ens. Mem. 2	-7.9 ± 0.4	7.1 ± 0.2	0.81	< 2.2 × 10^{-16}
Control Ens. Mem. 3	-8.4 ± 0.4	7.3 ± 0.2	0.87	< 2.2 × 10^{-16}
Control Ens. Mem. 4	-8.6 ± 0.4	7.1 ± 0.2	0.84	< 2.2 × 10^{-16}
Control Ens. Mean	-8.3 ± 0.3	7.1 ± 0.1	0.92	< 2.2 × 10^{-16}
AIM Ens. Mem. 1	-6.0 ± 0.3	9.6 ± 0.2	0.83	< 2.2 × 10^{-16}
AIM Ens. Mem. 2	-6.0 ± 0.3	9.5 ± 0.2	0.84	< 2.2 × 10^{-16}
AIM Ens. Mem. 3	-6.8 ± 0.3	9.9 ± 0.2	0.86	< 2.2 × 10^{-16}
AIM Ens. Mem. 4	-6.2 ± 0.3	9.8 ± 0.2	0.85	< 2.2 × 10^{-16}
AIM Ens. Mean	-6.2 ± 0.2	9.7 ± 0.2	0.92	< 2.2 × 10^{-16}

Table A.1: Sea ice volume (March and September) and sea ice extent (September) linear fits for the Control and AIM ensembles. Two different 70 years periods have been chosen for the two ensembles: 1980–2050 for the control ensemble and 2030–2100 for the AIM ensemble. The March sea ice extent has not been considered because not sensitive to the AIM implementation. The intercept values are all computed by considering 1980 as origin of the x axis.

B. Appendix to Chapter 3[1]

B.1 Introduction

This Supporting Information further details some aspects that have been discussed only briefly in the method section of the paper. Text B.2 provides a theoretical description of the SPS and the $IIEE$, the verification metrics employed in our analysis. This includes the error decompositions of the $IIEE$, which are used to characterize model shortcomings in more detail. Although the metrics themselves have been already described in previous literature, we believe that reporting here a brief summary is beneficial for the readership of the paper. Moreover, an additional new aspect of our paper is to consider the SPS as the sum of the $IIEE$ for the ensemble-median edge and a residual term, the properties of which are discussed here. Text B.3 gives a deeper insight into the choice of the resolution at which the analysis is performed. Furthermore, we investigate the sensitivity of the results to different observational products. Finally, Text B.4 describes the main features of two different observation-based benchmark forecasts and outlines their respective advantages and drawbacks.

The main features of the S2S forecast systems considered in the analysis are summarized in Tab. B.1. Fig. B.1 provides some details on the SPS of the climatological benchmark forecasts ($CSPS$) and of the persistence benchmark forecasts ($PSPS$), including uncertainties and error decompositions. The error components of the $IIEE$ for the S2S forecast systems are shown in Figs. B.2 and B.3, while the maps in Fig. B.4 illustrate the comparison between the older and the new ECMWF forecast system. The figure also depicts multiple observed ice edges to exemplify observational uncertainties, as well as the common (maximum-overlap) land mask that is applied to all

---

[1]This appendix contains the supplementary information of the published paper '*Bright prospects for Arctic sea ice prediction on subseasonal time scales*' by Zampieri *et al.* (2018)

model data and observational data for our analysis.

B.2 Text S1

The following paragraphs provide some details on the verification metrics employed in our analysis, thereby complementing the "Methods" section of the paper.

To account for the probabilistic nature of the S2S ensemble forecasts we apply the recently introduced SPS, which has been devised specifically for the verification of probabilistic contour forecasts (Goessling & Jung (2018)). The SPS is defined as the spatial integral of the local (Half) Brier Scores. Applied to the sea ice edge (the 15% sea ice concentration contour), we have

$$SPS = \int_S \left(P_o \left[sic \geq 15\% \right] (\vec{x}) - P_f \left[sic \geq 15\% \right] (\vec{x}) \right)^2 \, dS \, , \qquad \text{(B.1)}$$

with $P_o \left[sic \geq 15\% \right]$ and $P_f \left[sic \geq 15\% \right]$ defined as the ensemble-based local probability of a sea ice concentration above 15%, respectively for observation and forecast. In this study, the integration domain S is the entire Arctic. Note that the SPS is also meaningful when applied to deterministic sea ice edge forecasts with a binary probability distribution. In this case, the system is described only by two probability values, $1 = ice$ and $0 = no\ ice$, and the SPS is equivalent to the $IIEE$. This allows the evaluation of both deterministic and probabilistic forecasts in the same framework, with the SPS as a common verification metric.

The $IIEE$ can be decomposed into Overestimation (O) and Underestimation (U) components or, alternatively, into Absolute Extent Error (AEE) and Misplacement Error (ME) components, which provides insight into the origin of forecast errors (Goessling *et al.* (2016a)):

$$IIEE = O + U \qquad \text{(B.2)}$$

and

$$IIEE = AEE + ME \, , \qquad \text{(B.3)}$$

where O is the spatial integral of all areas where the forecast sea ice concentration is above 15% but the observed sea ice concentration is below 15%,

and U is the spatial integral of all areas where the forecast sea ice concentration is below 15% but the observed sea ice concentration is above 15%. Further, it is $AEE = |O - U|$, which implies that $ME = 2 \cdot \min(O, U)$. The two types of decomposition are not independent of each other: A forecast with either too much ice everywhere ($O/IIEE \approx 100\%$) or too little ice everywhere ($O/IIEE \approx 0\%$) must also be characterized by a dominance of the Absolute Extent Error ($AEE/IIEE \approx 100\%$). Nevertheless, the two decompositions are not completely redundant and reveal different aspects of forecast errors. For example, a forecast system that is unbiased with respect to the total sea ice extent ($\overline{O} \approx \overline{U}$, where the overline denotes a time average) can be dominated either by absolute extent errors or by misplacement errors, depending on the spatial pattern of regional biases as well as the spatial degrees of freedom of ice-edge variations (Goessling *et al.* (2016a)).

To take advantage of the additional information these decompositions provide, we consider the SPS as the sum of the $IIEE$ for the ensemble-median ice edge and a residual:

$$SPS = IIEE + r \ . \tag{B.4}$$

The ensemble-median ice edge is the 50%-contour of the forecast sea ice probability ($P_f\,[sic \geq 15\%]$). The residual r is zero for a deterministic forecast and can be either positive or negative for a probabilistic forecast. For example, if the observed ice edge coincides exactly with the forecast median ice edge, then r is positive, meaning that a larger error is assigned to the full probabilistic forecast compared to the perfect ensemble-median ice edge. Typically, however, non-binary forecast probabilities tend to reduce the error, at least as long as they are reliable (Goessling & Jung (2018)). In this case, r tends to be negative, which is consistent with the premise that probabilistic forecast information should add value to a "deterministic" forecast such as the one defined by the ensemble median.

B.3   Text S2

To test the sensitiveness of our results to the employment of different observational products, we repeated parts of our analysis with sea ice concentration data derived from passive microwave measurements using the ASI retrieval algorithm (Spreen *et al.*, 2008). Given the higher resolution of this obser-

vational product ($\sim$6 km instead of $\sim$25 km of OSI-SAF), the resulting 15%
contour line encompasses in this case also small areas of open water within
the main ice edge (Fig. B.4, orange contour line). When the analysis is per-
formed on the observational grids, the result differs depending on the chosen
sea ice concentration product. The ASI-based forecast errors are on average
25% larger than the OSI-SAF-based errors, mostly because of the small areas
of open water within the main ice edge (not shown). However, for the same
reason, the reference benchmark error is also higher when based on the ASI-
derived sea ice edge, leading to equivalent considerations in terms of the S2S
forecast skills. On the contrary, if the analysis is performed on the coarser
model grid, the choice of the observational product plays a minor role. In
fact, the ASI and OSI-SAF ice edges are almost equivalent if interpolated to
the 1.5°$\times$1.5°S2S longitude-latitude grid. Regridding the observations to the
model grid has therefore been preferred to guarantee higher stability to the
analysis.

The S2S forecasts data are provided on a common 1.5°$\times$1.5° longitude-
latitude grid, even though the models originally run at higher resolutions
varying between 0.25° and 1°. To test whether our results are affected by the
interpolation operated on the forecasts, we have repeated our analysis for 5
ECMWF forecast provided at the model native resolution (0.25°). The test
reveals that the sensitivity to the remapping procedure is small. The SPS
is on average 10% higher when computed on the high-resolution native grid
compared to the low-resolution S2S mesh. Once again, this discrepancy does
not affect our estimations of the forecast predictive skill, since the reference
benchmark error would also increase when computed at higher resolution and
therefore compensate for the slight SPS increase.

In conclusion, the previous tests reveal that the best and most conve-
nient configuration to perform the analysis is that where the OSI-SAF ob-
servations are interpolated by a first-order conservative remapping to the
coarse-resolution forecast grid.

B.4 Text S3

Here we provide further details about the main features of the benchmark
forecasts. As mentioned in the main body of the paper, two strategies have
been followed to define the observation-based benchmarks: a climatological

forecast ($CSPS$) and a 1-month persistence forecast ($PSPS$)(Blanchard-Wrigglesworth *et al.*, 2010). Given the different nature of the two approaches, the mechanisms leading to eventual predictive skills are also different and highly seasonal dependent.

The CSPS is built by definition on the climatological record, therefore it is generally more skillful (i.e. lower error) when the sea ice edge undergoes a low year-to-year variability and less skillful when the variability is higher. This feature is well captured by the $CSPS$, as shown in Fig. B.1 (top-left plot). The $CSPS$ reaches its minimum value ($\sim 0.35 \times 10^6 \, \mathrm{km}^2$) in June and November when the sea ice edge is shorter and constrained by the topographic position of the coastlines. Therefore, the seasonality in the $CSPS$ is primarily influenced by a corresponding seasonality in the length of the Arctic ice edge. Two maximums are reached in correspondence with the maximum and minimum Arctic sea ice extent in March and September. In these months the sea ice edge is longer and less coherent because influenced by the high atmospheric and ocean year-to-year variability and by the decadal sea ice declining trend. In particular, the $CSPS$ is least skillful in September ($\sim 0.75 \times 10^6 \, \mathrm{km}^2$), when the declining sea ice extent trend is more pronounced. Consistently, the error decompositions reveal that the $CSPS$ is slightly dominated by overestimation, meaning that the climatology predicts, in general, more ice than observed.

The skill of the persistence-based benchmark forecast relay on an ocean-memory mechanism. The surface heat content of the Arctic ocean is highly correlated with the amount of time for which the ocean surface is ice-free and therefore exposed to the atmosphere. The ocean heat content has a strong impact on the onset of the melting and freezing seasons, becoming therefore a source of potential predictability. In light of the previous considerations, one would expect the 1-month persistence benchmark to be most skillful for at most 1 month after the maximum and minimum sea ice extent are reached, and least skillful during transient periods of rapid advance and retreat of the ice edge. This feature is well captured by the PSPS in Fig. B.1 (bottom-left plot, turquoise curve). A sharp minimum ($\sim 0.6 \times 10^6 \, \mathrm{km}^2$) can be observed at the beginning of October, less than one month after the September minimum, while a second less pronounced minimum is evident in late March, after the smooth transition between freezing and melting season. The O–U decomposition shows a strong bimodal trend which is coherent with the definition of persistence benchmark: overestimation is dominant during the melt

season and underestimation during the freezing season.

To conclude, both the CSPS and PSPS appear to be meaningful choices to evaluate the predictive skill of the S2S forecast systems, with seasonally dependent advantages and drawbacks. However, the CSPS is systematically lower than the PSPS, with the only exception of a relatively short 20-days window between September and October. We, therefore, consider the CSPS to be generally a more restrictive benchmark with lower seasonal variability. The CSPS represents in our opinion a better and fairer reference to assess the skill of subseasonal sea ice forecasts in representing the ice edge position. As we show in the paper, exceptions are certainly possible for specific periods in some years. The persistence benchmark is much more skillful, and therefore restrictive, than the climatological benchmark for forecasts targeting a period shortly after the September minimum in years with an exceptionally low sea ice extent.

B.5 Figures and tables

Forecast System	Initialisation Freq.	Ensemble Size	Range
ECMWF	×2 weekly	10+1	46 days
UKMO	×4 monthly	6+1	60 days
KMA	×4 monthly	2+1	60 days
NCEP	daily	3+1	44 days
MF	×2 monthly	14+1	61 days
CMA	daily	3+1	60 days
ECMWF Pres.	×2 weekly	single sea ice state	46 days

Forecast System	Sea Ice Model	SIC Assimilated
ECMWF	LIM 2	yes
UKMO	CICE 4.1	yes
KMA	CICE 4.1	yes
NCEP	GFDL SIS	yes
MF	GELATO 5	no
CMA	GFDL SIS	no
ECMWF Pres.	none	no

Table B.1: Summary of some key characteristics of the reforecasts from the S2S database used in this study. The ensemble size is given as $n+1$ to emphasize the availability of n perturbed ensemble members and 1 unperturbed control run. Sea ice concentration is abbreviated by SIC. The forecast systems are indicated by the following acronyms: ECMWF = "Europen Centre for Medium Weather Forecasts", UKMO = "United Kingdom Meteorological Office", KMA = "Korean Meteorological Administration", NCEP = "National Centers for Environmental Protection", MF = "Météo France", CMA = "Chinese Meteorological Administration"
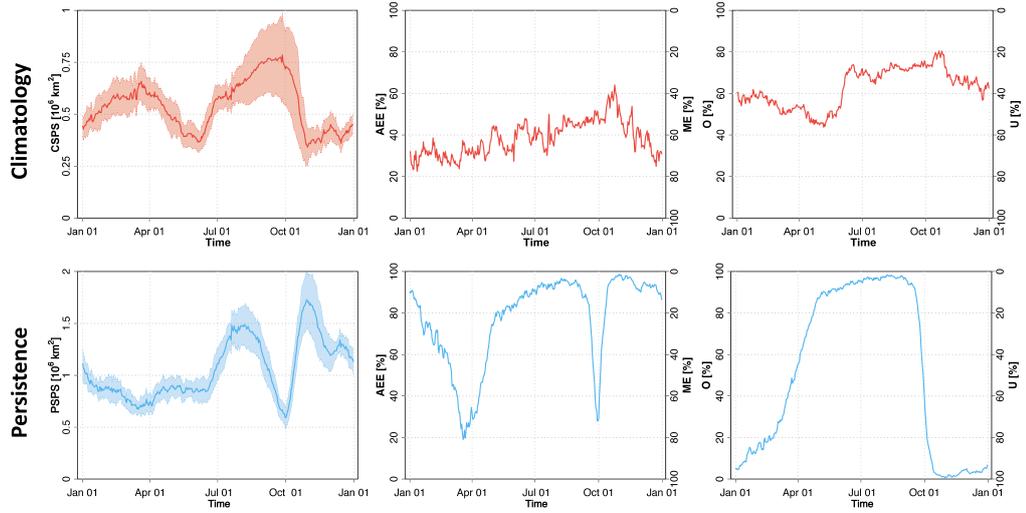
Figure B.1: First row left: Spatial Probability Score of the climatological benchmark forecasts (CSPS) based on the respective previous ten years as a function of the time of the year, averaged over the period 1999–2010. The shading and dashed lines delineate ∼95% confidence intervals, based on two times the standard error obtained from the twelve individual annual means. Second row left: Spatial Probability Score of the 1-month persistence benchmark forecasts (PSPS) as before. Middle column: Relative contributions to the Integrated Ice Edge Error of the climatological ensemble-median edge (first row) and persistence edge (second row) from Absolute Extent Error (AEE) versus Misplacement Error (ME); Right column: As middle column, but contributions from Overestimation (O) versus Underestimation (U).
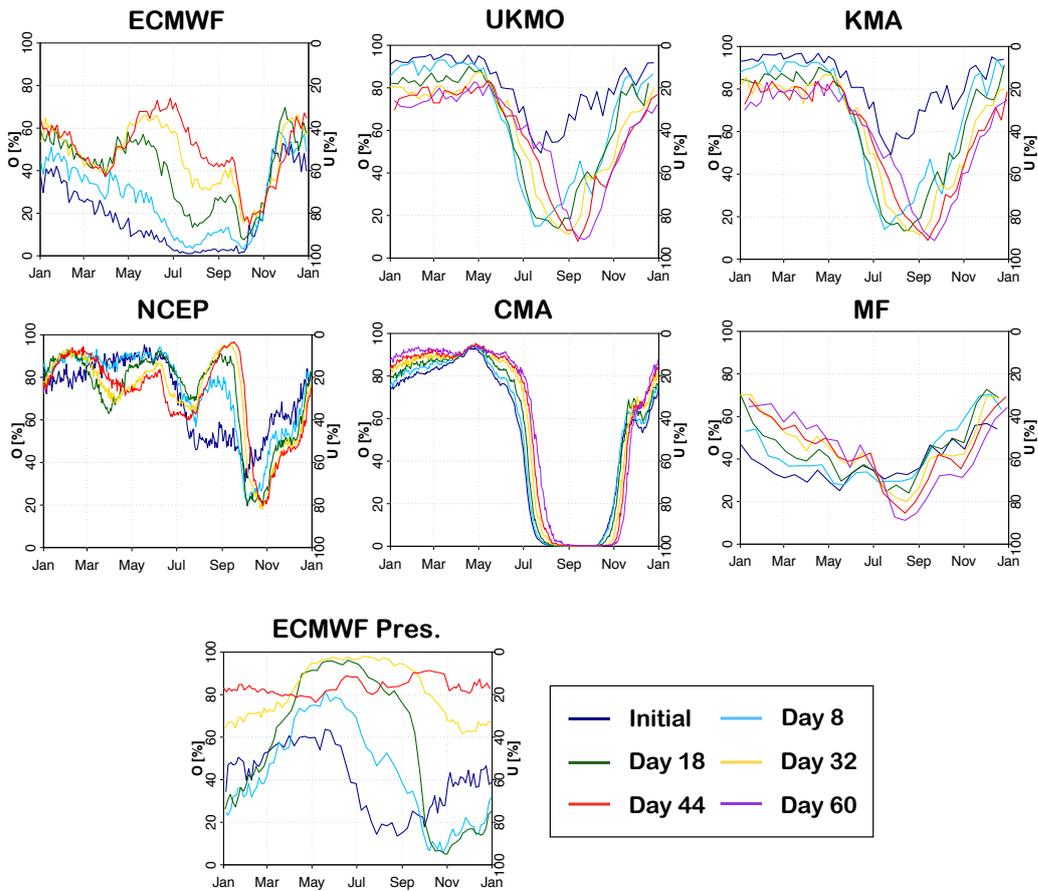
Figure B.2: Relative contributions to the Integrated Ice Edge Error of the ensemble-median ice edge from Overestimation (O) versus Underestimation (U) of individual S2S systems as a function of the time of the year (target date) and for six different lead times (see legend). Results have been averaged over the common reforecast period 1999–2010. Note that Day 60 is missing for NCEP and ECMWF (both versions) due to their shorter lead time ranges, and that *Initial Time* corresponds to Day 1 for all systems except NCEP and MF where it corresponds to Day 2 for technical reasons.
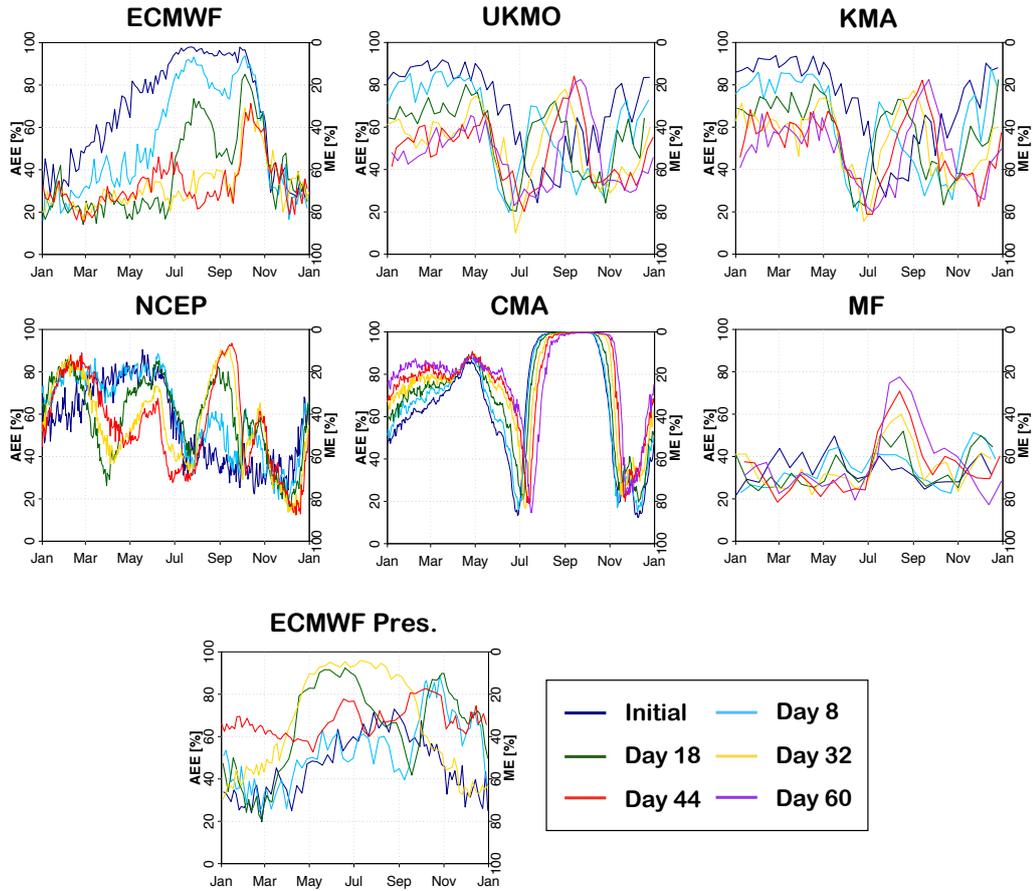
Figure B.3: Relative contributions to the Integrated Ice Edge Error of the ensemble-median ice edge from Absolute Extent Error (AEE) versus Misplacement Error (ME) of individual S2S systems as a function of the time of the year (target date) and for six different lead times (see legend). Results have been averaged over the common reforecast period 1999–2010. Note that Day 60 is missing for NCEP and ECMWF (both versions) due to their shorter lead time ranges, and that *Initial Time* corresponds to Day 1 for all systems except NCEP and MF where it corresponds to Day 2 for technical reasons.
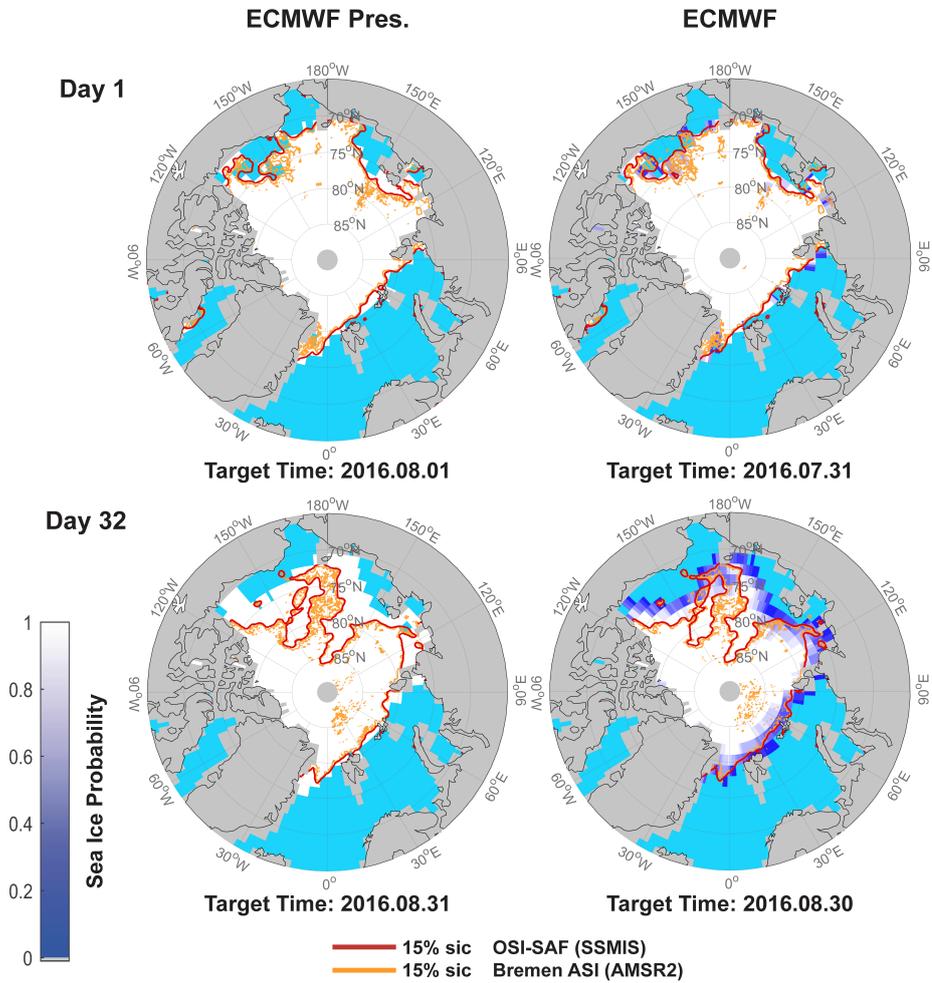
Figure B.4: Forecasts of the probability of occurrence of sea ice (probability that sea ice concentration exceeds 15%) for August 2016 with simplified (ECMWF Pres, left column) and advanced treatment (ECMWF, right column) of sea ice in the ECMWF forecast system (see text for detail on model formulation). Top row: Day-1 forecasts (≈ initial state). Bottom row: Day-32 forecasts. The absolute date is shifted by one day due to different initialization times. The forecast sea ice Probability (probability that the sea ice concentration exceeds 15%), shown by the color scale, is overlaid with 15% sea ice concentration contours based on three different observational products (see legend).

C.   Appendix to Chapter 4[1]

C.1   Figures and tables
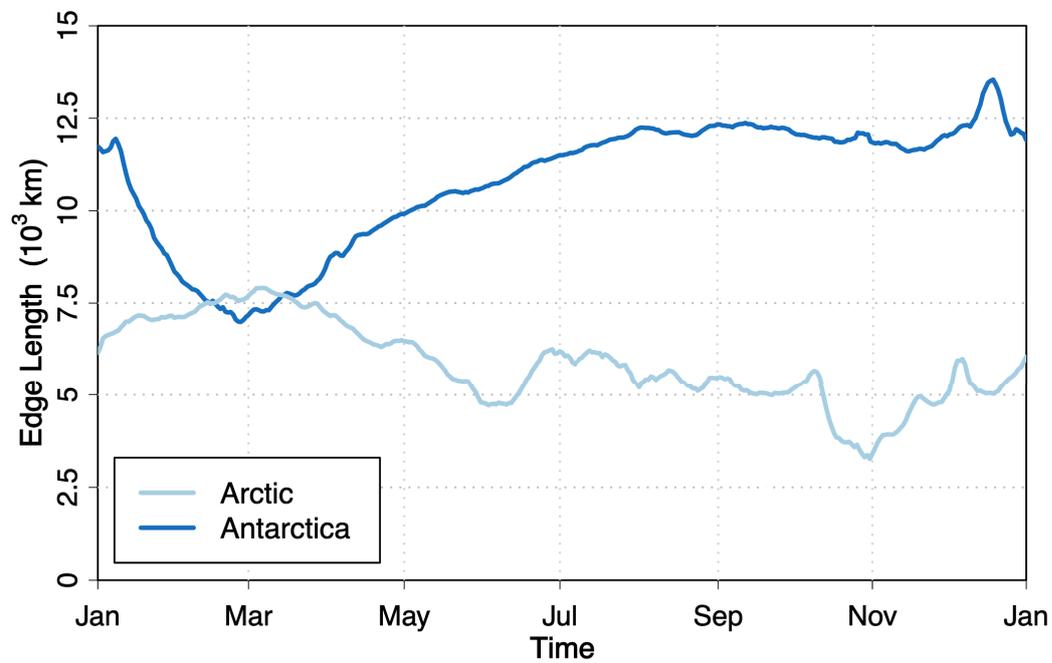
Figure C.1: Seasonal variation of the Arctic and Antarctic climatological sea ice edge length based on satellite observations. The sea ice edge climatology is based on the 12-years period 1999–2010.

---

[1]This appendix contains the supplementary information of the published paper 'Pre-
dictability of Antarctic sea ice edge on subseasonal time scales' by Zampieri *et al.* (2019)
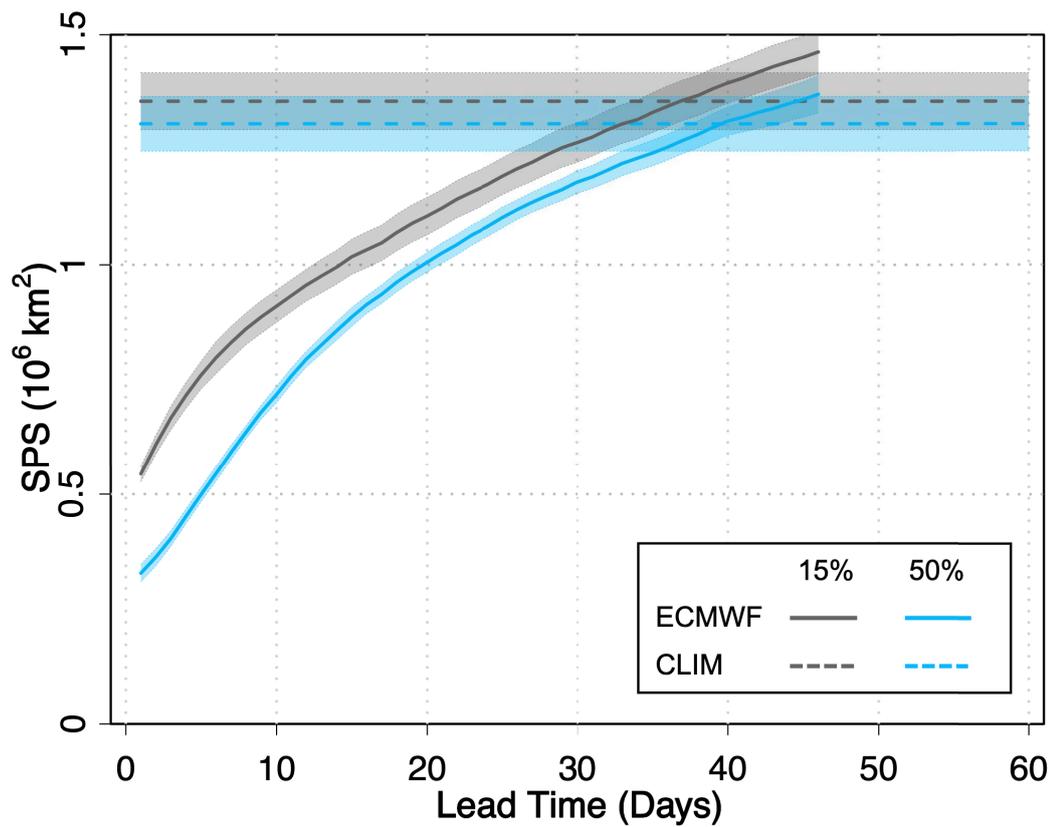
Figure C.2: Annual-mean forecast skill in predicting the sea ice edge location in terms of the SPS of the ECMWF forecastsystems (solid line) and of the climatological benchmark (dashed line) for two different sea ice concentration contours (15% and 50%). ECMWF = European Centre for Medium-Range Weather Forecasts.

D. Additional papers and datasets related to this thesis

In the framework of my Ph.D., I contributed to other scientific studies outside the scope of this thesis and to the compilation of scientific datasets. I give a brief overview of these activities in the following sections.

D.1 Co-author papers

"Leads and ridges in Arctic sea ice from RGPS data and a new tracking algorithm" has been published in the journal '*The Cryosphere*' by Hutter *et al.* (2019) (including L. Zampieri). This manuscript describes two methods for detecting and tracking linear kinematic features (LKFs) in sea ice deformation data and it builds an LKF dataset for the entire observing period of the RADARSAT Geophysical Processor System (RGPS). I rewrote the original version of the tracking algorithm in Python, which has been used as a basis for further developments by the lead author of the study.

"Toward a Data Assimilation System for Seamless Sea Ice Prediction Based on the AWI Climate Model" has been published in the '*Journal of Advances in Modeling Earth Systems*' by Mu *et al.* (2020) (including L. Zampieri). This manuscript describes and evaluates the data assimilation component of a seamless sea-ice prediction system for the formulation of sea-ice forecast at different timescales. The system is based on the fully coupled climate model developed at the Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research (AWICM, v1.1). In this context, I gave technical support in relation to the atmospheric model component ECHAM6.

I contributed to writing both papers.

D.2 Datasets

"Linear Kinematic Features (leads & pressure ridges) detected and tracked in RADARSAT Geophysical Processor System (RGPS) sea-ice deformation data from 1997 to 2008" has been published on the 'PANGAEA' repository by Hutter *et al.* (2019) (including L. Zampieri) as supplement to Hutter *et al.* (2019).

"Sea ice targeted geoengineering simulation with the AWI Climate Model" has been published on the 'PANGAEA' repository by Zampieri & Goessling (2019) as supplement to Zampieri & Goessling (2019).

Acknowledgments

List of Figures

List of Tables

Abbreviations

Table D.1: Abbreviations

AEE	Absolute Extent Error
aEVP	adaptive EVP
AIDJEX	Arctic Ice Dynamics Joint Experiment
AIM	Arctic Ice Management
AMOC	Atlantic Meridional Overturning Circulation
AMSR	Advanced Microwave Scanning Radiometer
AWI	Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research
AWI-CM	Alfred Wegener Institute Climate Model
BL99	Bitz & Lipscomb (1999) thermodynamics
CCSM3	Community Climate System Model version 3
CESM	Community Earth System Model
CLIM	Probabilistic climatological forecast
CMA	China Meteorological Administration
CMIP	Coupled Model Intercomparison Project
CMIP3	3rd phase of CMIP
CMIP5	5th phase of CMIP
CMIP6	6th phase of CMIP
D17	Desch *et al.* (2017)
DEM	Discrete Element Model
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA5	ECMWF Reanalysis, 5th Generation
ESM	Earth System Models
EVP	Elastic Viscous Plastic
FCT	Flux Corrected Transport

Continued on next page

Table D.1 – *Continued from previous page*

FE	Finite Element
FESIM	Finite Element Sea-Ice Model
FESOM2	Finite-volumE Sea ice-Ocean Model version 2
GMP	Global Modulation Parameter
H79	Hibler (1979) sea-ice strength formulation
IIEE	Integrated Ice Edge Error
ITD	Ice Thickness Distribution
KMA	Korea Meteorological Administration
LIM3	Louvain-la-Neuve sea Ice Model version 3
LMP	Local Modulation Parameter
ME	Misplacement Error
mEVP	modified EVP
MF	Météo-France
MITgcm	Massachusetts Institute of Technology General Circulation Model
MOSAiC	Multidisciplinary drifting Observatory for the Study of Arctic Climate
NAOSIM	North Atlantic Arctic Ocean Sea Ice Model
NCEP	National Centers for Environmental Prediction
NCEP-CFS	NCEP Climate Forecast System
NEMO	Nucleus for European Modelling of the Ocean
NSIDC	National Snow and Ice Data Center
O	Overestimation
ORA-IP	Ocean Reanalyses Intercomparison Project
OSI-SAF	Ocean and Sea Ice Satellite Application Facility
PDAF	Parallel Data Assimilation Framework
PERS	Deterministic persistence forecast
PIOMAS	PanArctic Ice Ocean Modeling and Assimilation System
PMW	Passive Microwave
PWL	Persistent Water Layer
R75	Rothrock (1975) sea-ice strength formulation
RCP	Representative Concentration Pathway
TOA	Top Of the Atmosphere
S2S	Subseasonal to Seasonal
SHEBA	Surface Heat Budget of the Arctic

Continued on next page

Table D.1 – *Continued from previous page*

SI	Sea Ice
SIO	Sea Ice Outlook
SIMIP	Sea Ice Model Intercomparison Project
SIP	Sea Ice Probability
SIPN	Sea Ice Prediction Network
SMOS	Soil Moisture Ocean Salinity
SPS	Spatial Probability Score
SSIPS	Seamless Sea Ice Prediction System
SSM/I	Special Sensor Microwave/Imager
SSMIS	Special Sensor Microwave Imager/Sounder
SST	Sea Surface Temperature
UKMO	UK Met Office
U	Underestimation
VP	Viscous Plastic

Bibliography

ALLEY, R. B., EMANUEL, K. A., & ZHANG, F. 2019. Advances in weather prediction. *Science*, **363**(6425), 342–344.

BALAN-SAROJINI, B., TIETSCHE, S., MAYER, M., ALONSO-BALMASEDA, M., & ZUO, H. 2019. Towards Improved Sea Ice Initialization and Forecasting with the IFS. *Technical Memorandum*.

BARRETT, A. P., STROEVE, J. C., & SERREZE, M. C. 2020. Arctic Ocean Precipitation From Atmospheric Reanalyses and Comparisons With North Pole Drifting Station Records. *Journal of Geophysical Research: Oceans*, **125**(1), e2019JC015415.

BATRAK, Y., & MÜLLER, M. 2019. On the warm bias in atmospheric reanalyses induced by the missing snow over Arctic sea-ice. *Nature Communications*, **10**(1), 4170.

BAUER, P., THORPE, A., & BRUNET, G. 2015. The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55.

BELLAMY, R., LEZAUN, J., & PALMER, J. 2017. Public perceptions of geoengineering research governance: An experimental deliberative approach. *Global Environmental Change*, **45**, 194 – 202.

BITZ, C. M., & LIPSCOMB, W. H. 1999. An energy-conserving thermodynamic model of sea ice. *Journal of Geophysical Research: Oceans*, **104**(C7), 15669–15677.

BITZ, C. M., HOLLAND, M. M., WEAVER, A. J., & EBY, M. 2001. Simulating the ice-thickness distribution in a coupled climate model. *Journal of Geophysical Research: Oceans*, **106**(C2), 2441–2463.

BLACKSTOCK, J. J., & LONG, J. C. S. 2010. The Politics of Geoengineering. *Science*, **327**(5965), 527–527.

BLANCHARD-WRIGGLESWORTH, E., BARTHÉLEMY, A., CHEVALLIER, M., CULLATHER, R., FUČKAR, N., MASSONNET, F., POSEY, P., WANG, W., ZHANG, J., ARDILOUZE, C., *et al.* 2017. Multi-model seasonal forecast of Arctic sea-ice: forecast uncertainty at pan-Arctic and regional scales. *Climate Dynamics*, **49**(4), 1399–1410.

BLANCHARD-WRIGGLESWORTH, E., ARMOUR, K. C., BITZ, C. M., & DEWEAVER, E. 2010. Persistence and Inherent Predictability of Arctic Sea Ice in a GCM Ensemble and Observations. *J. Climate*, **24**(1), 231–250.

BLOCKLEY, E. W., & PETERSON, K. A. 2018. Improving Met Office seasonal predictions of Arctic sea ice using assimilation of CryoSat-2 thickness. *The Cryosphere*, **12**(11), 3419–3438.

BLOCKLEY, E., VANCOPPENOLLE, M., HUNKE, E., BITZ, C., FELTHAM, D., LEMIEUX, J.-F., LOSCH, M., MAISONNAVE, E., NOTZ, D., RAMPAL, P., TIETSCHE, S., TREMBLAY, B., TURNER, A., MASSONNET, F., LASON, E., ROBERTS, A., AKSENOV, Y., FICHEFET, T., GARRIC, G., IOVINO, D., MADEC, G., ROUSSET, C., SALAS Y MELIA, D., & SCHROEDER, D. 2020. The Future of Sea Ice Modeling: Where Do We Go from Here? *Bulletin of the American Meteorological Society*, **101**(8), E1304–E1311.

BURGARD, C., NOTZ, D., PEDERSEN, L. T., & TONBOE, R. T. 2020a. The Arctic Ocean Observation Operator for 6.9 GHz (ARC3O) – Part 2: Development and evaluation. *The Cryosphere*, **14**(7), 2387–2407.

BURGARD, C., NOTZ, D., PEDERSEN, L. T., & TONBOE, R. T. 2020b. The Arctic Ocean Observation Operator for 6.9 GHz (ARC3O) – Part 1: How to obtain sea ice brightness temperatures at 6.9 GHz from climate model output. *The Cryosphere*, **14**(7), 2369–2386.

CARRIERES, T., BUEHNER, M., LEMIEUX, J., & PEDERSEN, L. (eds). 2017. *Sea Ice Analysis and Forecasting: Towards an Increased Reliance on Automated Prediction Systems*. Cambridge University Press.

CHEN, D., & YUAN, X. 2004. A Markov Model for Seasonal Forecast of Antarctic Sea Ice. *Journal of Climate*, **17**(16), 3156–3168.

CHEUNG, W. W. L., LAM, V. W. Y., SARMIENTO, J. L., KEARNEY, K., WATSON, R., ZELLER, D., & PAULY, D. 2010. Large-scale redistribution

of maximum fisheries catch potential in the global ocean under climate change. *Global Change Biology*, **16**(1), 24–35.

CHEVALLIER, M., SMITH, G., DUPONT, F., LEMIEUX, J.-F., FORGET, G., YOSUKE, F., HERNANDEZ, F., MSADEK, R., PETERSON, K., STORTO, A., TOYODA, T., VALDIVIESO, M., VERNIERES, G., ZUO, H., BALMASEDA, M., CHANG, Y.-S., FERRY, N., GARRIC, G., HAINES, K., & WANG, X. 2016. Intercomparison of the Arctic sea ice cover in global ocean-sea ice reanalyses from the ORA-IP project. *Climate Dynamics*.

CHEVALLIER, M., SMITH, G. C., DUPONT, F., LEMIEUX, J.-F., FORGET, G., FUJII, Y., HERNANDEZ, F., MSADEK, R., PETERSON, K. A., STORTO, A., TOYODA, T., VALDIVIESO, M., VERNIERES, G., ZUO, H., BALMASEDA, M., CHANG, Y.-S., FERRY, N., GARRIC, G., HAINES, K., KEELEY, S., KOVACH, R. M., KURAGANO, T., MASINA, S., TANG, Y., TSUJINO, H., & WANG, X. 2017. Intercomparison of the Arctic sea ice cover in global oceansea ice reanalyses from the ORA-IP project. *Climate Dynamics*, **49**(3), 1107–1136.

COLLINS, M., KNUTTI, R., ARBLASTER, J., DUFRESNE, J.-L., FICHEFET, T., FRIEDLINGSTEIN, P., GAO, X., GUTOWSKI, W., JOHNS, T., KRINNER, G., SHONGWE, M., TEBALDI, C., WEAVER, A., & WEHNER, M. 2013. *Long-term Climate Change: Projections, Commitments and Irreversibility*. Cambridge University Press. Book Section 12, pages 1029–1136.

COMISO, J. C., & NISHIO, F. 2008. Trends in the sea ice cover using enhanced and compatible AMSR-E, SSM/I, and SMMR data. *Journal of Geophysical Research: Oceans*, **113**(C2).

COMISO, J. C. 2012. Large Decadal Decline of the Arctic Multiyear Ice Cover. *Journal of Climate*, **25**(4), 1176–1193.

COOLEY, S. W., RYAN, J. C., SMITH, L. C., HORVAT, C., PEARSON, B., DALE, B., & LYNCH, A. H. 2020. Coldest Canadian Arctic communities face greatest reductions in shorefast sea ice. *Nature Climate Change*, **10**(6), 533–538.

Cornwall, W. 2015. Inside the Paris climate deal. *Science*, **350**(6267), 1451–1451.

Cvijanovic, I., Caldeira, K., & MacMartin, D. G. 2015. Impacts of ocean albedo alteration on Arctic sea ice restoration and Northern Hemisphere climate. *Environ. Res. Lett.*, **10**(4), 044020.

Danilov, S., Wang, Q., Timmermann, R., Iakovlev, N., Sidorenko, D., Kimmritz, M., Jung, T., & Schröter, J. 2015a. Finite-Element Sea Ice Model (FESIM), version 2. *Geoscientific Model Development*, **8**(6), 1747–1761.

Danilov, S., Wang, Q., Timmermann, R., Iakovlev, N., Sidorenko, D., Kimmritz, M., Jung, T., & Schröter, J. 2015b. Finite-Element Sea Ice Model (FESIM), version 2. *Geoscientific Model Development*, **8**(6), 1747–1761.

Danilov, S., Sidorenko, D., Wang, Q., & Jung, T. 2017. The FinitevolumE Sea ice–Ocean Model (FESOM2). *Geoscientific Model Development*, **10**(2), 765–789.

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

Desch, S. J., Smith, N., Groppi, C., Vargas, P., Jackson, R., Kalyaan, A., Nguyen, P., Probst, L., Rubin, M. E., Singleton, H., Spacek, A., Truitt, A., Zaw, P. P., & Hartnett, H. E. 2017. Arctic ice management. *Earth's Future*, **5**(1), 107–127.

Dieckmann, G. S., & Hellmer, H. H. 2010. *The Importance of Sea Ice: An Overview*. John Wiley Sons, Ltd. Chap. 1, pages 1–22.

Director, H. M., Raftery, A. E., & Bitz, C. M. 2017. Improved Sea Ice Forecasting through Spatiotemporal Bias Correction. *Journal of Climate*, **30**(23), 9493–9510.

Döscher, R., Vihma, T., & Maksimovich, E. 2014. Recent advances in understanding the Arctic climate system state and change from a sea ice perspective: a review. *Atmospheric Chemistry and Physics*, **14**(24), 13571–13600.

Eijgelaar, E., Thaper, C., & Peeters, P. 2010. Antarctic cruise tourism: the paradoxes of ambassadorship, last chance tourism and greenhouse gas emissions. *Journal of Sustainable Tourism*, **18**(3), 337–354.

Emmerson, C., & Lahn, G. 2012. *Arctic Opening: Opportunity and Risk in the High North*. Chathman House.

EUMETSAT Ocean and Sea Ice Satellite Application Facility. 2017. *Global sea ice concentration climate data record 1979-2015 (v2.0)*.

EUMETSAT Ocean and Sea Ice Satellite Application Facility. 2019. *Global sea ice concentration interim climate data record 2016 onwards (v2.0)*.

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

Field, L., Ivanova, D., Bhattacharyya, S., Mlaker, V., Sholtz, A., Decca, R., Manzara, A., Johnson, D., Christodoulou, E., Walter, P., & Katuri, K. 2018. Increasing Arctic Sea Ice Albedo Using Localized Reversible Geoengineering. *Earth's Future*, **6**(6), 882–901.

Flocco, D., Feltham, D. L., & Turner, A. K. 2010. Incorporation of a physically based melt pond scheme into the sea ice component of a climate model. *Journal of Geophysical Research: Oceans*, **115**, C08012.

Flocco, D., Schroeder, D., Feltham, D. L., & Hunke, E. C. 2012. Impact of melt ponds on Arctic sea ice simulations from 1990 to 2007. *Journal of Geophysical Research: Oceans*, **117**(C9).

Fowler, C., Maslanik, J., Emery, W., & Tschudi, M. 2013. *Polar Pathfinder Daily 25 km EASE-Grid Sea Ice Motion Vectors, Version 2 (daily and mean gridded field)*.

Fraser, A. D., Massom, R. A., Ohshima, K. I., Willmes, S., Kappes, P. J., Cartwright, J., & Porter-Smith, R. 2020. High-resolution mapping of circum-Antarctic landfast sea ice distribution, 2000–2018. *Earth System Science Data Discussions*, **2020**, 1–18.

GASCARD, J.-C., ZHANG, J., & RAFIZADEH, M. 2019. Rapid decline of Arctic sea ice volume: Causes and consequences. *The Cryosphere Discussions*, **2019**, 1–29.

GHERARDI, M., & LAGOMARSINO, M. C. 2015. Characterizing the size and shape of sea ice floes. *Scientific Reports*, **5**(1), 10226.

GIVENS, J. E. 2018. Geoengineering in Context. *Nature Sustainability*, **1**, 459 460.

GOESSLING, H. F., & JUNG, T. 2018. A probabilistic verification score for contours: Methodology and application to Arctic ice edge forecasts. *Quarterly Journal of the Royal Meteorological Society*.

GOESSLING, H. F., TIETSCHE, S., DAY, J. J., HAWKINS, E., & JUNG, T. 2016a. Predictability of the Arctic sea ice edge. *Geophysical Research Letters*, **43**, 1642–1650.

GOESSLING, H. F., TIETSCHE, S., DAY, J. J., HAWKINS, E., & JUNG, T. 2016b. Predictability of the Arctic sea ice edge. *Geophysical Research Letters*, **43**(4), 1642–1650.

GOESSLING, H. F., JUNG, T., KLEBE, S., BAESEMAN, J., BAUER, P., CHEN, P., CHEVALLIER, M., DOLE, R., GORDON, N., RUTI, P., BRADLEY, A., BROMWICH, D. H., CASATI, B., CHECHIN, D., DAY, J. J., MASSONNET, F., MILLS, B., RENFREW, I., SMITH, G., & TATUSKO, R. 2016c. Paving the Way for the Year of Polar Prediction. *Bulletin of the American Meteorological Society*, **97**(4), ES85–ES88.

GUEMAS, V., BLANCHARD-WRIGGLESWORTH, E., CHEVALLIER, M., DAY, J. J., DQU, M., DOBLAS-REYES, F. J., FUKAR, N. S., GERME, A., HAWKINS, E., KEELEY, S., KOENIGK, T., SALAS Y MLIA, D., & TIETSCHE, S. 2016. A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales. *Quarterly Journal of the Royal Meteorological Society*, **142**(695), 546–561.

HAMILTON, C. 2013. No, we should not just at least do the research. *Nature News*, **496**(7444), 139.

HENDRICKS, S., PAUL, S., & RINNE, E. 2018a. *ESA Sea Ice Climate Change Initiative (Sea_Ice_cci): Northern hemisphere sea ice thickness from the CryoSat-2 satellite on a monthly grid (L3C), v2.0.*

138

HENDRICKS, S., PAUL, S., & RINNE, E. 2018b. *ESA Sea Ice Climate Change Initiative (Sea_Ice_cci): Northern hemisphere sea ice thickness from the Envisat satellite on a monthly grid (L3C), v2.0.*

HERMAN, A. 2016. Discrete-Element bonded-particle Sea Ice model DESIgn, version 1.3a – model description and implementation. *Geoscientific Model Development*, **9**(3), 1219–1241.

HERSBACH, H., BELL, B., BERRISFORD, P., HIRAHARA, S., HORNYI, A., MUOZ-SABATER, J., NICOLAS, J., PEUBEY, C., RADU, R., SCHEPERS, D., SIMMONS, A., SOCI, C., ABDALLA, S., ABELLAN, X., BALSAMO, G., BECHTOLD, P., BIAVATI, G., BIDLOT, J., BONAVITA, M., DE CHIARA, G., DAHLGREN, P., DEE, D., DIAMANTAKIS, M., DRAGANI, R., FLEMMING, J., FORBES, R., FUENTES, M., GEER, A., HAIMBERGER, L., HEALY, S., HOGAN, R. J., HLM, E., JANISKOV, M., KEELEY, S., LALOYAUX, P., LOPEZ, P., LUPU, C., RADNOTI, G., DE ROSNAY, P., ROZUM, I., VAMBORG, F., VILLAUME, S., & THPAUT, J.-N. 2020. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, (146), 1999–2049.

HIBLER, W. D., I. 1979. A Dynamic Thermodynamic Sea Ice Model. *Journal of Physical Oceanography*, **9**(4), 815–846.

HOLLAND, M. M., BAILEY, D. A., BRIEGLEB, B. P., LIGHT, B., & HUNKE, E. C. 2012. Improved Sea Ice Shortwave Radiation Physics in CCSM4: The Impact of Melt Ponds and Aerosols on Arctic Sea Ice. *Journal of Climate*, **25**(5), 1413–1430.

HOLLAND, M. M., BLANCHARD-WRIGGLESWORTH, E., KAY, J., & VAVRUS, S. 2013. Initial-value predictability of Antarctic sea ice in the Community Climate System Model 3. *Geophysical Research Letters*, **40**(10), 2121–2124.

HUDSON, S. R. 2011. Estimating the global radiative impact of the sea icealbedo feedback in the Arctic. *Journal of Geophysical Research: Atmospheres*, **116**(D16).

HUNKE, E. C., & DUKOWICZ, J. K. 1997. An ElasticViscousPlastic Model for Sea Ice Dynamics. *Journal of Physical Oceanography*, **27**(9), 1849–1867.

HUNKE, E. C., HEBERT, D. A., & LECOMTE, O. 2013. Level-ice melt ponds in the Los Alamos sea ice model, CICE. *Ocean Modelling*, **71**, 24–42.

HUNKE, E. C., ALLARD, R., BAILEY, D. A., BLAIN, P., CRAIG, A., DUPONT, F., DUVIVIER, A., GRUMBINE, R., HEBERT, D., HOLLAND, M., JEFFERY, N., LEMIEUX, J., OSINSKI, R., RASMUSSEN, T., RIBERGAARD, M., ROACH, L., ROBERTS, A., TURNER, M., & WINTON, M. 2020a. CICE-Consortium/CICE: CICE Version 6.1.2 (Version 6.1.2). *Zenodo*.

HUNKE, E. C., ALLARD, R., BAILEY, D. A., BLAIN, P., CRAIG, A., DUPONT, F., DUVIVIER, A., GRUMBINE, R., HEBERT, D., HOLLAND, M., JEFFERY, N., LEMIEUX, J., OSINSKI, R., RASMUSSEN, T., RIBERGAARD, M., ROACH, L., ROBERTS, A., TURNER, M., & WINTON, M. 2020b. CICE-Consortium/Icepack: Icepack 1.2.1 (Version 1.2.1). *Zenodo*.

HUNKE, E., ALLARD, R., BLAIN, P., BLOCKLEY, E., FELTHAM, D., FICHEFET, T., GARRIC, G., GRUMBINE, R., LEMIEUX, J.-F., RASMUSSEN, T., RIBERGAARD, M., ROBERTS, A., SCHWEIGER, A., TIETSCHE, S., TREMBLAY, B., VANCOPPENOLLE, M., & ZHANG, J. 2020c. Should Sea-Ice Modeling Tools Designed for Climate Research Be Used for Short-Term Forecasting? *Current Climate Change Reports*.

HUNKE, E. C., LIPSCOMB, W. H., & TURNER, A. K. 2010. Sea-ice models for climate study: retrospective and new directions. *Journal of Glaciology*, **56**(200), 11621172.

HUTTER, N., ZAMPIERI, L., & LOSCH, M. 2019. Leads and ridges in Arctic sea ice from RGPS data and a new tracking algorithm. *The Cryosphere*, **13**(2), 627–645.

HUTTER, N., ZAMPIERI, L., & LOSCH, M. 2019. *Linear Kinematic Features (leads & pressure ridges) detected and tracked in RADARSAT Geophysical Processor System (RGPS) sea-ice deformation data from 1997 to 2008*.

JAHN, A. 2018. Reduced probability of ice-free summers for 1.5°C compared to 2 °C warming. *Nature Climate Change*, **8**(5), 409–413.

JIN, Z., CHARLOCK, T. P., SMITH JR., W. L., & RUTLEDGE, K. 2004. A parameterization of ocean surface albedo. *Geophysical Research Letters*, **31**(22).

JORDAN, R. E., ANDREAS, E. L., & MAKSHTAS, A. P. 1999. Heat budget of snow-covered sea ice at North Pole 4. *Journal of Geophysical Research: Oceans*, **104**(C4), 7785–7806.

JUNG, T., & MATSUEDA, M. 2016. Verification of global numerical weather forecasting systems in polar regions using TIGGE data. *Quarterly Journal of the Royal Meteorological Society*, **142**(695), 574–582.

JUNG, T., GORDON, N. D., BAUER, P., BROMWICH, D. H., CHEVALLIER, M., DAY, J. J., DAWSON, J., DOBLAS-REYES, F., FAIRALL, C., GOESSLING, H. F., *et al.* 2016. Advancing polar prediction capabilities on daily to seasonal time scales. *Bulletin of the American Meteorological Society*, **97**(9), 1631–1647.

KAY, J. E., HOLLAND, M. M., & JAHN, A. 2011. Inter-annual to multi-decadal Arctic sea ice extent trends in a warming world. *Geophysical Research Letters*, **38**(15).

KEITH, D. W. 2001. Geoengineering. *Nature*, **409**(420).

KIMMRITZ, M., DANILOV, S., & LOSCH, M. 2015. On the convergence of the modified elasticviscousplastic method for solving the sea ice momentum equation. *Journal of Computational Physics*, **296**, 90–100.

KIMMRITZ, M., DANILOV, S., & LOSCH, M. 2016. The adaptive EVP method for solving the sea ice momentum equation. *Ocean Modelling*, **101**, 59–67.

KIMURA, N., NISHIMURA, A., TANAKA, Y., & YAMAGUCHI, H. 2013. Influence of winter sea-ice motion on summer ice cover in the Arctic. *Polar Research*, **32**.

KRUMPEN, T., BIRRIEN, F., KAUKER, F., RACKOW, T., VON ALBEDYLL, L., ANGELOPOULOS, M., BELTER, H. J., BESSONOV, V., DAMM, E., DETHLOFF, K., HAAPALA, J., HAAS, C., HARRIS, C., HENDRICKS, S., HOELEMANN, J., HOPPMANN, M., KALESCHKE, L., KARCHER, M., KOLABUTIN, N., LEI, R., LENZ, J., MORGENSTERN, A., NICOLAUS,

M., Nixdorf, U., Petrovsky, T., Rabe, B., Rabenstein, L., Rex, M., Ricker, R., Rohde, J., Shimanchuk, E., Singha, S., Smolyanitsky, V., Sokolov, V., Stanton, T., Timofeeva, A., Tsamados, M., & Watkins, D. 2020. The MOSAiC ice floe: sediment-laden survivor from the Siberian shelf. *The Cryosphere*, **14**(7), 2173–2187.

Kuzmin, D. 2009. Explicit and implicit FEM-FCT algorithms with flux linearization. *Journal of Computational Physics*, **8**(7), 2517–2534.

Kwok, R. 2018. Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018). *Environmental Research Letters*, **13**(10), 105005.

Kwok, R. 2002. Sea ice concentration estimates from satellite passive microwave radiometry and openings from SAR ice motion. *Geophysical Research Letters*, **29**(9), 25–1–25–4.

Larsen, J. N., & Fondahl, G. (eds). 2015. *Arctic Human Development Report : Regional Processes and Global Linkages*. TemaNord, no. 2014:567.

Lavergne, T., Eastwood, S., Teffah, Z., Schyberg, H., & Breivik, L.-A. 2010. Sea ice motion from low-resolution satellite sensors: An alternative method and its validation in the Arctic. *Journal of Geophysical Research: Oceans*, **115**, C10032.

Lavergne, T., Sørensen, A. M., Kern, S., Tonboe, R., Notz, D., Aaboe, S., Bell, L., Dybkjær, G., Eastwood, S., Gabarro, C., Heygster, G., Killie, M. A., Brandt Kreiner, M., Lavelle, J., Saldo, R., Sandven, S., & Pedersen, L. T. 2019. Version 2 of the EUMETSAT OSI SAF and ESA CCI sea-ice concentration climate data records. *The Cryosphere*, **13**(1), 49–78.

Lemieux, J.-F., Tremblay, L. B., Dupont, F., Plante, M., Smith, G. C., & Dumont, D. 2015. A basal stress parameterization for modeling landfast ice. *Journal of Geophysical Research: Oceans*, **120**(4), 3157–3173.

Lemieux, J.-F., Dupont, F., Blain, P., Roy, F., Smith, G. C., & Flato, G. M. 2016. Improving the simulation of landfast ice by combining tensile strength and a parameterization for grounded ridges. *Journal of Geophysical Research: Oceans*, **121**(10), 7354–7368.

LEPPRANTA, M. 2009. Sea Ice Dynamics. *Pages 159 – 169 of:* STEELE, J. H. (ed), *Encyclopedia of Ocean Sciences (Second Edition)*, second edn. Oxford: Academic Press.

LINDSAY, R., & SCHWEIGER, A. 2015. Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite observations. *The Cryosphere*, **9**(1), 269–283.

LIPSCOMB, W. H., HUNKE, E. C., MASLOWSKI, W., & JAKACKI, J. 2007. Ridging, strength, and stability in high-resolution sea ice models. *Journal of Geophysical Research: Oceans*, **112**, C03S91.

LIPSCOMB, W. H., & HUNKE, E. C. 2004. Modeling Sea Ice Transport Using Incremental Remapping. *Monthly Weather Review*, **132**(6), 1341–1354.

LOSCH, M., MENEMENLIS, D., CAMPIN, J.-M., HEIMBACH, P., & HILL, C. 2010. On the formulation of sea-ice models. Part 1: Effects of different solver implementations and parameterizations. *Ocean Modelling*, **33**(1), 129 – 144.

LHNER, R., MORGAN, K., PERAIRE, J., & VAHDATI, M. 1987. Finite element flux-corrected transport (FEMFCT) for the euler and NavierStokes equations. *International Journal for Numerical Methods in Fluids*, **7**(10), 1093–1109.

MAITRA, U., & ZARE, R. N. 2016. Fall and rise of a D_2O ice cube in liquid H_2O. *Reson*, **21**, 453–456.

MANABE, S., & STOUFFER, R. J. 1980. Sensitivity of a global climate model to an increase of CO2 concentration in the atmosphere. *Journal of Geophysical Research: Oceans*, **85**(C10), 5529–5554.

MASLANIK, J., STROEVE, J., FOWLER, C., & EMERY, W. 2011. Distribution and trends in Arctic sea ice age through spring 2011. *Geophysical Research Letters*, **38**(13), L13502.

MASSONNET, F., FICHEFET, T., GOOSSE, H., VANCOPPENOLLE, M., MATHIOT, P., & KÖNIG BEATTY, C. 2011. On the influence of model physics on simulations of Arctic and Antarctic sea ice. *The Cryosphere*, **5**(3), 687–699.

MASSONNET, F., GOOSSE, H., FICHEFET, T., & COUNILLON, F. 2014. Calibration of sea ice dynamic parameters in an ocean-sea ice model using an ensemble Kalman filter. *Journal of Geophysical Research: Oceans*, **119**(7), 4168–4184.

MASSONNET, F., REID, P., LIESER, J. L., BITZ, C. M., FYFE, J., HOBBS, W., & KUSAHARA, K. 2018. Assessment of February 2018 sea-ice forecasts for the Southern Ocean. *Technical Note, Université catholique de Louvain.*

MASSONNET, F., REID, P., LIESER, J. L., BITZ, C. M., FYFE, J., & HOBBS, W. 2019. Assessmentof summer 2018–2019 sea-ice forecasts for the Southern Ocean. *Technical Note, Université catholique de Louvain.*

MAYKUT, G. A., & UNTERSTEINER, N. 1971. Some results from a time-dependent thermodynamic model of sea ice. *Journal of Geophysical Research (1896-1977)*, **76**(6), 1550–1575.

McCUSKER, K. E., ARMOUR, K. C., BITZ, C. M., & BATTISTI, D. S. 2014. Rapid and extensive warming following cessation of solar radiation management. *Environmental Research Letters*, **9**(2), 024005.

MELSOM, A., PALERME, C., & MÜLLER, M. 2019. Validation metrics for ice edge position forecasts. *Ocean Science Discussions*, **2019**, 1–27.

MENEMENLIS, D., & WUNSCH, C. 1997. Linearization of an Oceanic General Circulation Model for Data Assimilation and Climate Studies. *Journal of Atmospheric and Oceanic Technology*, **14**(6), 1420–1443.

MENEMENLIS, D., FUKUMORI, I., & LEE, T. 2005. Using Green's Functions to Calibrate an Ocean General Circulation Model. *Monthly Weather Review*, **133**(5), 1224–1240.

MENEMENLIS, D., CAMPIN, J.-M., HEIMBACH, P., HILL, C., LEE, T. L., NGUYEN, A., SCHODLOK, M., & ZHANG, H. 2008. ECCO2: High Resolution Global Ocean and Sea Ice Data Synthesis. *Mercator Ocean Quarterly Newsletter*, 13–21.

MENGIS, N., MARTIN, T., KELLER, D. P., & OSCHLIES, A. 2016. Assessing climate impacts and risks of ocean albedo modification in the Arctic. *Journal of Geophysical Research: Oceans*, **121**(5), 3044–3057.

MILLER, P. A., LAXON, S. W., FELTHAM, D. L., & CRESSWELL, D. J. 2006. Optimization of a Sea Ice Model Using Basinwide Observations of Arctic Sea Ice Thickness, Extent, and Velocity. *Journal of Climate*, **19**(7), 1089–1108.

MILLER, P. A., LAXON, S. W., & FELTHAM, D. L. 2007. Consistent and contrasting decadal Arctic sea ice thickness predictions from a highly optimized sea ice model. *Journal of Geophysical Research: Oceans*, **112**(C7).

MOHAMMADI-ARAGH, M., GOESSLING, H. F., LOSCH, M., HUTTER, N., & JUNG, T. 2018. Predictability of Arctic sea ice on weather time scales. *Scientific Reports*, **8:6514**, 1–6.

MU, L., YANG, Q., LOSCH, M., LOSA, S. N., RICKER, R., NERGER, L., & LIANG, X. 2017. Improving sea ice thickness estimates by assimilating CryoSat-2 and SMOS sea ice thickness data simultaneously. *Quarterly Journal of the Royal Meteorological Society*, **144**, 529–538.

MU, L., NERGER, L., TANG, Q., LOZA, S. N., SIDORENKO, D., WANG, Q., SEMMLER, T., ZAMPIERI, L., LOSCH, M., & GOESSLING, H. F. 2020. Toward a Data Assimilation System for Seamless Sea Ice Prediction Based on the AWI Climate Model. *Journal of Advances in Modeling Earth Systems*, **12**(4), e2019MS001937. e2019MS001937 10.1029/2019MS001937.

MUELLER, B. L., GILLETT, N. P., MONAHAN, A. H., & ZWIERS, F. W. 2018. Attribution of Arctic Sea Ice Decline from 1953 to 2012 to Influences from Natural, Greenhouse Gas, and Anthropogenic Aerosol Forcing. *Journal of Climate*, **31**(19), 7771–7787.

NERGER, L., & HILLER, W. 2013. Software for ensemble-based data assimilation systemsImplementation strategies and scalability. *Computers Geosciences*, **55**, 110 – 118. Ensemble Kalman filter for data assimilation.

NGUYEN, A. T., MENEMENLIS, D., & KWOK, R. 2011. Arctic ice-ocean simulation with optimized model parameters: Approach and assessment. *Journal of Geophysical Research: Oceans*, **116**(C4).

NIEDERDRENK, A. L., & NOTZ, D. 2018. Arctic Sea Ice in a 1.5°C Warmer World. *Geophysical Research Letters*, **45**(4), 1963–1971.

NOTZ, D. 2014. Sea-ice extent and its trend provide limited metrics of model performance. *The Cryosphere*, **8**(1), 229–243.

NOTZ, D. 2005. Thermodynamic and Fluid-Dynamical Processes in Sea Ice. *Ph.D. dissertation at the University of Cambridge*.

NOTZ, D. 2012. Challenges in simulating sea ice in Earth System Models. *WIREs Climate Change*, **3**(6), 509–526.

NOTZ, D., & STROEVE, J. 2016. Observed Arctic sea-ice loss directly follows anthropogenic CO_2 emission. *Science*, **354**(6313), 747–750.

NOTZ, D., & STROEVE, J. 2018. The Trajectory Towards a Seasonally Ice-Free Arctic Ocean. *Current Climate Change Reports*, **4**(4), 407–416.

O'NEILL, B. C., TEBALDI, C., VAN VUUREN, D. P., EYRING, V., FRIEDLINGSTEIN, P., HURTT, G., KNUTTI, R., KRIEGLER, E., LAMARQUE, J.-F., LOWE, J., MEEHL, G. A., MOSS, R., RIAHI, K., & SANDERSON, B. M. 2016. The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, **9**(9), 3461–3482.

ORDOÑEZ, A. C., BITZ, C. M., & BLANCHARD-WRIGGLESWORTH, E. 2018. Processes Controlling Arctic and Antarctic Sea Ice Predictability in the Community Earth System Model. *Journal of Climate*, **31**(23), 9771–9786.

OSI-SAF. 2016. *Global Sea Ice Concentration climate data record.* `http://www.osi-saf.org/?q=content/global-sea-ice-concentration-data-record-ssmrssmi`.

OVERLAND, J. E., & WANG, M. 2013. When will the summer Arctic be nearly sea ice free? *Geophysical Research Letters*, **40**(10), 2097–2101.

OWENS, W. B., & LEMKE, P. 1990. Sensitivity studies with a sea icemixed layerpycnocline model in the Weddell Sea. *Journal of Geophysical Research: Oceans*, **95**(C6), 9527–9538.

PALERME, C., MLLER, M., & MELSOM, A. 2019. An intercomparison of verification scores for evaluating the sea ice edge position in seasonal forecasts. *Geophysical Research Letters*, **46**, 4757–4763.

PARKINSON, C. L. 2019. A 40-y record reveals gradual Antarctic sea ice increases followed by decreases at rates far exceeding the rates seen in the Arctic. *Proceedings of the National Academy of Sciences*, **116**(29), 14414–14423.

PARKINSON, C. L., & WASHINGTON, W. M. 1979. A large-scale numerical model of sea ice. *Journal of Geophysical Research: Oceans*, **84**(C1), 311–337.

PEROVICH, D. K. 2003. Complex yet translucent: the optical properties of sea ice. *Physica B: Condensed Matter*, **338**(1), 107 – 114. Proceedings of the Sixth International Conference on Electrical Transport and Optical Properties of Inhomogeneous Media.

PITHAN, F., & MAURITSEN, T. 2014. Arctic amplification dominated by temperature feedbacks in contemporary climate models. *Nature Geoscience*, **7**(3), 181–184.

RACKOW, T., GOESSLING, H. F., JUNG, T., SIDORENKO, D., SEMMLER, T., BARBI, D., & HANDORF, D. 2016. Towards multi-resolution global climate modeling with ECHAM6-FESOM. Part II: climate variability. *Climate Dynamics*, **44**, 757780.

RACKOW, T., SEIN, D. V., SEMMLER, T., DANILOV, S., KOLDUNOV, N. V., SIDORENKO, D., WANG, Q., & JUNG, T. 2019. Sensitivity of deep ocean biases to horizontal resolution in prototype CMIP6 simulations with AWI-CM1.0. *Geoscientific Model Development*, **12**(7), 2635–2656.

RAHMSTORF, S. 1999. Decadal Variability of the Thermohaline Ocean Circulation. *In:* NAVARRA, A. (ed), *Beyond El Nio*. Springer.

RICKER, R., HENDRICKS, S., KALESCHKE, L., TIAN-KUNZE, X., KING, J., & HAAS, C. 2017. A weekly Arctic sea-ice thickness data record from merged CryoSat-2 and SMOS satellite data. *The Cryosphere*, **11**(4), 1607–1623.

RINGEISEN, D., LOSCH, M., TREMBLAY, L. B., & HUTTER, N. 2019. Simulating intersection angles between conjugate faults in sea ice with different viscous–plastic rheologies. *The Cryosphere*, **13**(4), 1167–1186.

147

Roach, L. A., Tett, S. F. B., Mineter, M. J., Yamazaki, K., & Rae, C. D. 2018a. Automated parameter tuning applied to sea ice in a global climate model. *Climate Dynamics*, **50**(1), 51–65.

Roach, L. A., Horvat, C., Dean, S. M., & Bitz, C. M. 2018b. An Emergent Sea Ice Floe Size Distribution in a Global Coupled Ocean-Sea Ice Model. *Journal of Geophysical Research: Oceans*, **123**(6), 4322–4337.

Rogelj, J., den Elzen, M., Hhne, N., Fransen, T., Fekete, H., Winkler, H., Schaeffer, R., Sha, F., Riahi, K., & Meinshausen, M. 2016. Paris Agreement climate proposals need a boost to keep warming well below 2°C. *Nature*, **534**(7609), 631–639.

Rostosky, P., Spreen, G., Farrell, S. L., Frost, T., Heygster, G., & Melsheimer, C. 2018. Snow Depth Retrieval on Arctic Sea Ice From Passive Microwave RadiometersImprovements and Extensions to Multiyear Ice Using Lower Frequencies. *Journal of Geophysical Research: Oceans*, **123**(10), 7120–7138.

Rostosky, P., Melsheimer, C., & Spreen, G. 2019a. *AMSR-2 winter snow depth on Arctic sea ice, Version 1.0 (NetCDF) (2012 to 2018)*.

Rostosky, P., Melsheimer, C., & Spreen, G. 2019b. *AMSR-E winter snow depth on Arctic sea ice, Version 1.0 (NetCDF) (2002 to 2011)*.

Rothrock, D. A. 1975. The energetics of the plastic deformation of pack ice by ridging. *Journal of Geophysical Research (1896-1977)*, **80**(33), 4514–4519.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.-y., Juang, H.-M. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., & Goldberg,

M. 2010. The NCEP Climate Forecast System Reanalysis. *Bulletin of the American Meteorological Society*, **91**(8), 1015–1058.

SAHA, S., MOORTHI, S., WU, X., WANG, J., NADIGA, S., TRIPP, P., BEHRINGER, D., HOU, Y.-T., CHUANG, H.-Y., IREDELL, M., EK, M., MENG, J., YANG, R., MENDEZ, M. P., VAN DEN DOOL, H., ZHANG, Q., WANG, W., CHEN, M., & BECKER, E. 2014. The NCEP Climate Forecast System Version 2. *Journal of Climate*, **27**(6), 2185–2208.

SCHRDER, D., FELTHAM, D. L., FLOCCO, D., & TSAMADOS, M. 2014. September Arctic sea-ice minimum predicted by spring melt-pond fraction. *Nature Climate Change*, **4**(5), 353–357.

SCHWEIGER, A., LINDSAY, R., ZHANG, J., STEELE, M., STERN, H., & KWOK, R. 2011. Uncertainty in modeled Arctic sea ice volume. *Journal of Geophysical Research: Oceans*, **116**, C00D06.

SEIN, D. V., DANILOV, S., BIASTOCH, A., DURGADOO, J. V., SIDORENKO, D., HARIG, S., & WANG, Q. 2016. Designing variable ocean model resolution based on the observed ocean variability. *Journal of Advances in Modeling Earth Systems*, **8**(2), 904–916.

SEIN, D. V., KOLDUNOV, N. V., DANILOV, S., SIDORENKO, D., WEKERLE, C., CABOS, W., RACKOW, T., SCHOLZ, P., SEMMLER, T., WANG, Q., & JUNG, T. 2018. The Relative Influence of Atmospheric and Oceanic Model Resolution on the Circulation of the North Atlantic Ocean in a Coupled Climate Model. *Journal of Advances in Modeling Earth Systems*, **10**(8), 2026–2041.

SEITZ, R. 2011. Bright water: hydrosols, water conservation and climate change. *Climatic Change*, **105**(3-4), 365–381.

SEMMLER, T., DANILOV, S., GIERZ, P., GOESSLING, H. F., HEGEWALD, J., HINRICHS, C., KOLDUNOV, N., KHOSRAVI, N., MU, L., RACKOW, T., SEIN, D. V., SIDORENKO, D., WANG, Q., & JUNG, T. 2020. Simulations for CMIP6 With the AWI Climate Model AWI-CM-1-1. *Journal of Advances in Modeling Earth Systems*, **12**(9), e2019MS002009. e2019MS002009 2019MS002009.

SEMTNER, ALBERT J., J. 1976. A Model for the Thermodynamic Growth of Sea Ice in Numerical Investigations of Climate. *Journal of Physical Oceanography*, **6**(3), 379–389.

SERREZE, M. C., BARRETT, A. P., SLATER, A. G., STEELE, M., ZHANG, J., & TRENBERTH, K. E. 2007. The large-scale energy budget of the Arctic. *Journal of Geophysical Research: Atmospheres*, **112**(D11).

SIDORENKO, D., RACKOW, T., JUNG, T., SEMMLER, T., BARBI, D., DANILOV, S., DETHLOFF, K., DORN, W., FIEG, K., GOESSLING, H. F., HANDORF, D., HARIG, S., HILLER, W., JURICKE, S., LOSCH, M., SCHRÖTER, J., SEIN, D. V., & WANG, Q. 2015. Towards multi-resolution global climate modeling with ECHAM6-FESOM. Part I: model formulation and mean climate. *Climate Dynamics*, **44**(3), 757–780.

SIDORENKO, D., GOESSLING, H., KOLDUNOV, N., SCHOLZ, P., DANILOV, S., BARBI, D., CABOS, W., GURSES, O., HARIG, S., HINRICHS, C., JURICKE, S., LOHMANN, G., LOSCH, M., MU, L., RACKOW, T., RAKOWSKY, N., SEIN, D., SEMMLER, T., SHI, X., STEPANEK, C., STREFFING, J., WANG, Q., WEKERLE, C., YANG, H., & JUNG, T. 2019. Evaluation of FESOM2.0 Coupled to ECHAM6.3: Preindustrial and HighResMIP Simulations. *Journal of Advances in Modeling Earth Systems*, **11**(11), 3794–3815.

SIMIP COMMUNITY. 2020. Arctic Sea Ice in CMIP6. *Geophysical Research Letters*, **47**(10), e2019GL086749. e2019GL086749 10.1029/2019GL086749.

SMETACEK, V., & NICOL, S. 2015. Polar ocean ecosystems in a changing world. *Nature*, **437**, 362–368.

SMITH, G. C., ROY, F., RESZKA, M., SURCEL COLAN, D., HE, Z., DEACU, D., BELANGER, J., SKACHKO, S., LIU, Y., DUPONT, F., LEMIEUX, J., BEAUDOIN, C., TRANCHANT, B., DRVILLON, M., GARRIC, G., TESTUT, C., LELLOUCHE, J., PELLERIN, P., RITCHIE, H., LU, Y., DAVIDSON, F., BUEHNER, M., CAYA, A., & LAJOIE, M. 2015. Sea ice forecast verification in the Canadian Global Ice Ocean Prediction System. *Quarterly Journal of the Royal Meteorological Society*, **142**(695), 659–671.

SPINDLER, M. 1994. Notes on the biology of sea ice in the Arctic and Antarctic. *Polar Biology*, **14**(5), 319–324.

SPREEN, G., KALESCHKE, L., & HEYGSTER, G. 2008. Sea ice remote sensing using AMSR-E 89-GHz channels. *Journal of Geophysical Research*, **113**, C02S03.

STAMMER, D., & WUNSCH, C. 1996. The determination of the large-scale circulation of the Pacific Ocean from satellite altimetry using model Green's functions. *Journal of Geophysical Research: Oceans*, **101**(C8), 18409–18432.

STEELE, M., MORLEY, R., & ERMOLD, W. 2001. PHC: A Global Ocean Hydrography with a High-Quality Arctic Ocean. *Journal of Climate*, **14**(9), 2079–2087.

STEPHENSON, S. R., LAURENCE, C. S., & AGNEW, A. A. 2011. Divergent long-term trajectories of human access to the Arctic. *Nature Climate Change*, **1**, 156–160.

STERN, H. L., SCHWEIGER, A. J., STARK, M., ZHANG, J., STEELE, M., & HWANG, B. 2018. Seasonal evolution of the sea-ice floe size distribution in the Beaufort and Chukchi seas. *Elem Sci Anth*, **6**(1), 48. Number: 1 Publisher: University of California Press.

STEVENS, B., GIORGETTA, M., ESCH, M., MAURITSEN, T., CRUEGER, T., RAST, S., SALZMANN, M., SCHMIDT, H., BADER, J., BLOCK, K., BROKOPF, R., FAST, I., KINNE, S., KORNBLUEH, L., LOHMANN, U., PINCUS, R., REICHLER, T., & ROECKNER, E. 2013. Atmospheric component of the MPI-M Earth System Model: ECHAM6. *Journal of Advances in Modeling Earth Systems*, **5**(2), 146–172.

STROEVE, J., HOLLAND, M. M., MEIER, W., SCAMBOS, T., & SERREZE, M. 2007. Arctic sea ice decline: Faster than forecast. *Geophysical Research Letters*, **34**(9).

STROEVE, J., HAMILTON, L. C., BITZ, C. M., & BLANCHARD-WRIGGLESWORTH, E. 2014. Predicting September sea ice: Ensemble skill of the SEARCH Sea Ice Outlook 2008–2013. *Geophysical Research Letters*, **41**, 2411–2418.

STROEVE, J. C., KATTSOV, V., BARRETT, A., SERREZE, M., PAVLOVA, T., HOLLAND, M., & MEIER, W. N. 2012. Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations. *Geophysical Research Letters*, **39**(16).

SUMATA, H., KAUKER, F., GERDES, R., KÖBERLE, C., & KARCHER, M. 2013. A comparison between gradient descent and stochastic approaches for parameter optimization of a sea ice model. *Ocean Science*, **9**(4), 609–630.

SUMATA, H., LAVERGNE, T., GIRARD-ARDHUIN, F., KIMURA, N., TSCHUDI, M. A., KAUKER, F., KARCHER, M., & GERDES, R. 2014. An intercomparison of Arctic ice drift products to deduce uncertainty estimates. *Journal of Geophysical Research: Oceans*, **119**(8), 4887–4921.

SUMATA, H., KAUKER, F., KARCHER, M., & GERDES, R. 2019a. Covariance of Optimal Parameters of an Arctic Sea IceOcean Model. *Monthly Weather Review*, **147**(7), 2579–2602.

SUMATA, H., KAUKER, F., KARCHER, M., & GERDES, R. 2019b. Simultaneous Parameter Optimization of an Arctic Sea IceOcean Model by a Genetic Algorithm. *Monthly Weather Review*, **147**(6), 1899–1926.

TALBERG, A., THOMAS, S., CHRISTOFF, P., & KAROLY, D. 2018. How geoengineering scenarios frame assumptions and create expectations. *Sustainability Science*, **13**(4), 1093–1104.

TIETSCHE, S., NOTZ, D., JUNGCLAUS, J. H., & MAROTZKE, J. 2011. Recovery mechanisms of Arctic summer sea ice. *Geophysical Research Letters*, **38**(2).

TIETSCHE, S., DAY, J. J., GUEMAS, V., HURLIN, W. J., KEELEY, S. P. E., MATEI, D., MSADEK, R., COLLINS, M., & HAWKINS, E. 2014. Seasonal to interannual Arctic sea ice predictability in current global climate models. *Geophys. Res. Lett.*, **41**, 1035–1043.

TIMMERMANN, R., DANILOV, S., SCHROETER, J., BOENING, C., SIDORENKO, D., & ROLLENHAGEN, K. 2009. Ocean circulation and sea ice distribution in a finite element global sea ice/ocean model. *Ocean Modelling*, **27**(3), 114 – 129.

TRODAHL, H. J., WILKINSON, S. O. F., McGUINNESS, M. J., & HASKELL, T. G. 2001. Thermal conductivity of sea ice; dependence on temperature and depth. *Geophysical Research Letters*, **28**(7), 1279–1282.

TSCHUDI, M., FOWLER, C., MASLANIK, J., & STROEVE, J. 2010. Tracking the Movement and Changing Surface Characteristics of Arctic Sea Ice. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **3**(4), 536–540.

TURNER, A. K., HUNKE, E. C., & BITZ, C. M. 2013a. Two modes of sea-ice gravity drainage: A parameterization for large-scale modeling. *Journal of Geophysical Research: Oceans*, **118**(5), 2279–2294.

TURNER, J., BRACEGIRDLE, T. J., PHILLIPS, T., MARSHALL, G. J., & HOSKING, J. S. 2013b. An Initial Assessment of Antarctic Sea Ice Extent in the CMIP5 Models. *Journal of Climate*, **26**(5), 1473–1484.

UNGERMANN, M., TREMBLAY, L. B., MARTIN, T., & LOSCH, M. 2017. Impact of the ice strength formulation on the performance of a sea ice thickness distribution model in the Arctic. *Journal of Geophysical Research: Oceans*, **122**(3), 2090–2107.

UNITED NATIONS. 2015. *Paris Agreement*. `https://unfccc.int/sites/default/files/english_paris_agreement.pdf`. Accessed: 2018-09-26.

UOTILA, P., GOOSSE, H., HAINES, K., CHEVALLIER, M., BARTHÉLEMY, A., BRICAUD, C., CARTON, J., FUČKAR, N., GARRIC, G., IOVINO, D., KAUKER, F., KORHONEN, M., LIEN, V. S., MARNELA, M., MASSONNET, F., MIGNAC, D., PETERSON, K. A., SADIKNI, R., SHI, L., TIETSCHE, S., TOYODA, T., XIE, J., & ZHANG, Z. 2018. An assessment of ten ocean reanalyses in the polar regions. *Climate Dynamics*, **52**(3), 1613–1650.

URREGO-BLANCO, J. R., URBAN, N. M., HUNKE, E. C., TURNER, A. K., & JEFFERY, N. 2016. Uncertainty quantification and global sensitivity analysis of the Los Alamos sea ice model. *Journal of Geophysical Research: Oceans*, **121**(4), 2709–2732.

UTTAL, T., CURRY, J. A., McPHEE, M. G., PEROVICH, D. K., MORITZ, R. E., MASLANIK, J. A., GUEST, P. S., STERN, H. L., MOORE, J. A.,

Turenne, R., Heiberg, A., Serreze, M. C., Wylie, D. P., Persson, O. G., Paulson, C. A., Halle, C., Morison, J. H., Wheeler, P. A., Makshtas, A., Welch, H., Shupe, M. D., Intrieri, J. M., Stamnes, K., Lindsey, R. W., Pinkel, R., Pegau, W. S., Stanton, T. P., & Grenfeld, T. C. 2002. Surface Heat Budget of the Arctic Ocean. *Bulletin of the American Meteorological Society*, **83**(2), 255–276.

Vancoppenolle, M., Fichefet, T., Goosse, H., Bouillon, S., Madec, G., & Maqueda, M. A. M. 2009. Simulating the mass balance and salinity of Arctic and Antarctic sea ice. 1. Model description and validation. *Ocean Modelling*, **27**(1), 33 – 53.

Vitart, F., Robertson, A. W., & Anderson, D. L. T. 2012. Subseasonal to Seasonal Prediction Project: Bridging the gap between weather and climate. *WMO Bulletin*, **61 (2)**, 23–28.

Vitart, F., *et al.* 2016. The Sub-seasonal to Seasonal Prediction (S2S) Project Database. *Bulletin of the American Meteorological Society*.

Wallace, J., & Hobbs, P. 2006. *Atmospheric Science: An Introductory Survey*. International Geophysics Series. Elsevier Academic Press.

Wang, M., & Overland, J. E. 2009. A sea ice free summer Arctic within 30 years? *Geophysical Research Letters*, **36**(7).

Wang, Q., Koldunov, N. V., Danilov, S., Sidorenko, D., Wekerle, C., Scholz, P., Bashmachnikov, I. L., & Jung, T. 2020. Eddy Kinetic Energy in the Arctic Ocean From a Global Simulation With a 1-km Arctic. *Geophysical Research Letters*, **47**(14), e2020GL088550.

Wang, X., Key, J., Kwok, R., & Zhang, J. 2016. Comparison of Arctic Sea Ice Thickness from Satellites, Aircraft, and PIOMAS Data. *Remote Sensing*, **8**(9).

Wayand, N. E., Bitz, C. M., & Blanchard-Wrigglesworth, E. 2019. A Year-Round Subseasonal-to-Seasonal Sea Ice Prediction Portal. *Geophysical Research Letters*, **46**(6), 3298–3307.

World Meteorological Organization. 2017. *Sea-Ice Information Services in the World, Edition 2017*. World Meteorological Organization. WMO-No. 574.

ZAMPIERI, L., & GOESSLING, H. 2019. *Sea ice targeted geoengineering simulation with the AWI Climate Model.*

ZAMPIERI, L., & GOESSLING, H. F. 2019. Sea Ice Targeted Geoengineering Can Delay Arctic Sea Ice Decline but not Global Warming. *Earth's Future*, **7**(12), 1296–1306.

ZAMPIERI, L., GOESSLING, H. F., & JUNG, T. 2018. Bright Prospects for Arctic Sea Ice Prediction on Subseasonal Time Scales. *Geophysical Research Letters*, **45**(18), 9731–9738.

ZAMPIERI, L., GOESSLING, H. F., & JUNG, T. 2019. Predictability of Antarctic Sea Ice Edge on Subseasonal Time Scales. *Geophysical Research Letters*, **46**(16), 9719–9727.