Insights on past and future sea-ice evolution from combining observations and models

Julienne Stroeve[1,2] and Dirk Notz[3]

[1]National Snow and Ice Data Center (NSIDC), University of Colorado (CU), Boulder, CO, USA
[2]University College London, London, UK
[3]Max Planck Institute for Meteorology, Hamburg, Germany

Abstract

We discuss the current understanding of past and future sea-ice evolution as inferred from combining model simulations and observations. In such combined analysis, the models allow us to enhance our understanding behind the observed evolution of sea ice, while the observations allow us to assess how realistically the models represent the processes that govern sea-ice evolution in the real world. Combined, observations and models thus provide robust insights into the functioning of sea ice in the Earth's climate system, and can inform policy decisions related to the future evolution of the ice cover. We find that models and observations agree well on the sensitivity of Arctic sea ice to global warming and on the main drivers for the observed retreat. In contrast, a robust reduction of the uncertainty range of future sea-ice evolution remains difficult, in particular since the observational record is often too short to robustly examine the impact of internal variability on model biases. Process-based model evaluation and model evaluation based on seasonal-prediction systems provide promising ways to overcome these limitations.

1. Introduction

Insights into the functioning of the Earth's climate system can be derived from either observations or from model simulations. However, in both cases an isolated analysis of just one of these two sources of climate data has severe limitations: An isolated analysis of climate-model simulations does not usually allow one to assess whether these simulations capture the main processes that govern the evolution of the Earth's climate system. An isolated analysis of observations does not usually allow one to understand the processes that are responsible for the observed evolution. Both these limitations can be overcome by examining specific aspects of the Earth's climate system in a combined analysis of both observations and model simulations. We here discuss how such combined analysis allows for robust insights related to the past and future evolution of Arctic sea ice, and where the combined approach is still limited by shortcomings of either models or observations.

This study is motivated by the severe impacts that the observed rapid loss of Arctic sea ice is having not only on the Earth's climate system, but also on the Arctic's ecology, communities and economy. For example, in regards to the Earth's climate system, the sea-ice retreat has clearly changed the surface energy budget in the Arctic [*Porter et al.*, 2012]. Through these changes, the ice loss contributes to amplified

warming in the Arctic during autumn [e.g. *Serreze et al.*, 2009], which may impact weather patterns at lower latitudes for example through its winter-time effect on the Northern Annual Mode (NAM) [e.g. *Francis et al.*, 2009; *Jaiser et al.*, 2012; *Liu et al.*, 2012]. Such a link between sea-ice loss and lower-latitude weather is, however, not robustly identified yet, with other studies suggesting that the winter atmospheric response to summer sea ice loss is currently too weak to overcome interannual variability [*Screen et al.*, 2013; *Screen and Simmonds*, 2013].

The loss of sea ice also affects the delicate balance of the Arctic ecosystem. The foundation of the Arctic Ocean food web is made up of the primary production of photosynthesizing organisms such as algae and phytoplankton, on which ultimately all marine life depends. The growth cycle of these photosynthesizing organisms is driven by the annual cycle of nutrients and sun light, both of which change rapidly along with the changing ice cover. A recent study estimated that the shrinking sea ice has already resulted in a 57% loss of total annual primary production in the Arctic Ocean, which in turn is impacting zooplankton that feed on the algae and phytoplankton [*Post et al.*, 2013].

Declining sea ice and warming in the Arctic are also affecting Arctic communities. Indigenous hunters are already seeing changes in weather and sea ice motion patterns that are impacting their day-to-day activities [e.g. *Ford et al.*, 2009] and coastal communities are increasingly exposed to large waves and coastal erosion from increase in open water and thawing of permafrost [e.g. *Jones et al.*, 2009].

There are economic considerations as well. The increase in open water area during summer has attracted substantial attention and generated speculation of increased economic activities in the Arctic, including shipping and resource extraction. Since 2007 the Arctic Ocean along the coast of Russia and Canada has been mostly free of ice in September, opening up these routes to shipping of goods between Europe and the Far East. Transit statistics collected by the Russian Northern Sea Route Administration Office indicate 41 vessels made the transit through the Northern Sea Route in 2011, increasing to 71 in 2013. The rise in ship traffic has the potential to increase pollution and the risk of spills in the region. While the shrinking sea ice cover provides greater access and potentially longer seasons of navigation, the primary driver for ships in the Arctic remains natural resource exploitation.

Given the number of important ramifications that sea-ice loss is already having on regional and, potentially, global climate and weather, Arctic biota, economic development and livelihoods within and outside of the Arctic, the observed loss of sea ice has given rise to two main questions, answering of which has become a major research priority in recent years: First, what drives the observed changes in sea-ice coverage? And second, what will the future evolution of sea-ice look like, both short-term and long-term?

To answer these questions, model simulations are our most helpful research tools. However, as discussed, for robust insights these simulations need to be combined with insights based on the observational record. In the following, we examine how such combination has allowed us to better understand the observed loss of Arctic sea ice, and how it has allowed for better predictions of the possible short-term and long-term future evolution of sea ice. Through such analysis, we identify gaps in our understanding, and identify future observation and modelling needs.

We start in section 2 with a discussion of the available methods for large-scale observations of sea ice, including their uncertainties, and describe how sea ice is represented in climate models. In section 3, we then discuss some of the insights that have been gained regarding the past evolution of sea ice by combining observations and models. This includes, in particular, a discussion of model evaluation based on the observational record. In section 4, we then move on to short-term forecasts of sea ice

that are usually gained by assimilating observations into models. We outline some recent progress in this field and discuss the impact of observational shortcomings. In section 5, we then focus on long-term projections of sea ice, examining in particular ways to reduce the large spread in possible future sea-ice evolution as obtained from model simulations. We close with a summary that also outlines some possible ways forward, including some guidance for the upcoming CMIP6.

2   Sea ice in Observations and Models

2.1 The Observational Data Record of Sea Ice

The ability to understand, predict and adapt to future climate change in the Arctic requires long-term, accurate and consistent climate data records (CDRs). At present, the monitoring of the spatial extent of the sea-ice cover from several successive passive microwave satellite sensors comprises such a CDR. These sensors observe microwave energy emitted by the Earth's surface, which, unlike visible light, passes through clouds and can also be measured during polar night. The continuous sea-ice record began with the launch of the Nimbus-7 Scanning Multichannel Microwave Radiometer (SMMR) in October 1978, and continued with successive Defense Meteorological Satellite Program Special Sensor Microwave Imagers (SSMIs) (1987 to present). These sensors provide a consistently processed sea-ice record at 25 km^2 spatial resolution on a daily basis (every other day for the SMMR record). Higher resolution data (12.5 km^2) are available from the Advanced Microwave Scanning Radiometer (AMSR-E) on NASA's Aqua satellite from 2002 to 2011, and from AMSR-2 onboard the Japanese GCOM-W satellite launched October 2010. Considerable effort has gone into cross calibrating the sensors to provide a long-term, consistently processed CDR useful for analysis of interannual variability and long-term trends, as outlined in section 3.

One of the limits of the passive microwave data record is its rather coarse spatial resolution compared to visible, thermal infrared or active microwave remote sensing. However, for comparison to climate model data, which are at even coarser resolution, the resolution is deemed adequate for model evaluation of metrics such as sea-ice concentration, sea-ice area, and sea-ice extent, where the latter is defined as the total area of all pixels containing at least 15% sea ice concentration.

A more important issue centers on the reliability of the estimates for sea-ice concentration and extent/area. Several studies have assessed the accuracy of commonly used passive microwave sea-ice algorithms through comparisons with measurements of the sea ice edge from other sensors, i.e. visible, thermal infrared or Synthetic Aperture Radar (SAR) observations [e.g. *Meier*, 2005*; Steffen and Schweiger*, 1991]. Among others, these studies have examined the quality of the two most widely used sea-ice algorithms, which are distributed by the National Snow and Ice Data Center (NSIDC) and featured in NSIDC's monthly analysis of the state of the sea ice cover (http://nsidc.org/arcticseaicenews): the NASA Team [*Cavalieri et al.*, 1984] and the Bootstrap [*Comiso*, 1986]. These algorithms show a sea-ice concentration accuracy of about 5% in winter, with the accuracy degrading substantially during summer to about 20% [*Meier and Notz, 2010*], and perhaps even up to 40% during active melt [e.g. *Rosel et al.,* 2012].

Several other algorithms exist that use different combinations of channels and polarizations. As discussed in *Ivanova et al.* [2015], a sea-ice algorithm's sensitivity to emissivity and physical temperature of the sea ice are the two main error sources in the derived sea-ice concentration fields, since they determine, together with sea-ice concentration, the brightness temperatures used in any algorithm. To minimize the

additional impact of atmospheric water vapor, clouds and ocean wind effects, combinations of brightness temperatures at different frequencies are used, with varying success in minimizing the atmospheric influence. Melt ponds also pose a problem in summer since their radiometric signature is similar to that of open water, resulting in an underestimation of the sea-ice concentration during periods of active melt. The sensitivity of the algorithms to melt ponds is the reason why NSIDC reports on the total ice extent rather than the true ice area in their monthly reports on the state of Arctic and Antarctic sea ice.

In the context of model evaluation, such reporting of sea-ice extent however is not ideal. Because of its non-linearity, many factors affect the magnitude of sea ice extent that are not related to the quality of a particular model [*Notz, 2014*]. For example, even for identical sea ice coverage, sea-ice extent will always be larger for larger grid cells, which, for the numerical grids used by CMIP5 models, causes a systematic bias simply from grid-cell geometry of on average about 1 million km² in winter sea-ice extent compared to observations. In addition, with both sea-ice extent and area, regional biases can cancel such that the overall sea-ice coverage is close to the observed one despite large regional differences. This could be overcome for example by the calculation of root-mean square differences between observed and simulated sea-ice concentration, where, however, observational uncertainty again limits the quantitative reliability of the resulting metric.

While satellite-based observations provide a unique and consistent CDR to study sea ice changes, it remains rather short, limiting our ability to accurately characterize sea ice variability and change. To expand the observed time-series past the modern satellite era (starting in October 1978), a blended record of earlier observations from the Hadley Centre Global Sea Ice and Sea Surface Temperature (Had1SST) data record [*Rayner et al.*, 2003], the merged early/late passive microwave times-series (e.g. 1972-2002 ESMR-SMMR-SSM/I) [*Cavalieri et al.*, 2003] and data from the National Snow and Ice Data Center (NSIDC) Sea Ice Index [*Fetterer et al.*, 2002], from 1953 through 2014 has been produced [*Meier et al.,* 2007]. Analysis of the statistical properties of this time series indicates that its time evolution can be explained by a constant internal variability, to which a linear downward trend is added during the satellite period [*Notz and Marotzke, 2012*]. This indicates that the internal variability of the pre-satellite time series is largely consistent with that during the more reliable satellite time series, giving further credibility to the long-term consistency of the merged record.

Unfortunately, the spatial coverage of sea ice, as expressed by sea-ice extent and area, is only one sea ice metric, and does not say anything about how thick the ice is, or how much snow is on top of the ice, both important metrics for the exchange of heat/moisture between the ocean and the atmosphere. Sea ice thickness records primarily come from a combination of early ice draft records derived from submarine sonar and upward looking sonar (ULS) instruments on bottom anchored moorings, airborne electromagnetic (EM) induction instruments, as well as thickness derived from either radar or laser altimeters flown on aircraft (e.g. from NASA's Operation IceBridge) and satellite (e.g. ERS-1/2, ICESat and CryoSat-2). These thickness estimates have been processed into a consistent data format in attempts to produce a Unified Sea Ice Thickness CDR [*Lindsay*, 2010], yet different measurement techniques and retrieval approaches, as well as inconsistent spatial and temporal coverage remain problematic. Even fewer observations of snow depth on sea ice exist, mostly limited to manual measurements, though IceBridge flights include a snow radar, greatly adding to our understanding of interannual variability in snow depth.

Since, climate data records for other sea ice variables currently do not exist sea-ice extent/area and concentration have been the most widely used variables to compare with climate model output [e.g. *Stroeve et al.*, 2007, 2012a, *Massonnet et al.*, 2012].

Nevertheless, there have been recent attempts to synthesize the disparate data records of ice thickness [e.g. *Lindsay and Schweiger*, 2015] and snow depth [*Webster et al.*, 2014] and some of these data have been used to assess climate model's ability to represent the observed sea ice thickness state [e.g. *McLaren et al.*, 2006; *Stroeve et al.*, 2014b]. Other studies have compared thickness estimates from the Pan-Arctic Ice-Ocean Modelling and Assimilation System [PIOMAS, *Zhang and Rothrock*, 2003] to climate model output [*Stroeve et al.*, 2014b; *Notz*, 2015]. PIOMAS consists of an ocean-ice model that assimilates satellite-derived sea ice concentration and is forced by NCAR/NCEP atmospheric reanalysis data to simulate sea ice thickness, concentration and ice volume. While limited comparisons have suggested that sea ice estimates from PIOMAS agree well with observations [e.g. *Schweiger et al.*, 2011; *Laxon et al.*, 2013], biases remain [*Stroeve et al.*, 2014b].

2.2 Sea ice in Global Climate Models

Global climate models (GCMs), and their more complex successors Earth System Model (ESMs), are our key tools for simulating the evolution of the Earth's climate system as well as understanding interlinked processes that comprise our climate system. On their numerical grids, these models solve mathematical formulations that describe the complex processes that together form the drivers of the Earth's climate system. In doing so, they usually contain individual blocks of code describing the ocean, the atmosphere and the land. These individual sub-programs usually use different numerical time steps, representative of the speed of the relevant processes in each system. To represent the exchange of momentum, heat and tracers at the interfaces between the different model components, a special component of the models, called a "coupler", ensures communication between them. The time step with which the coupler exchanges information between the sub-components is called the coupling frequency, ranging for the applications discussed here typically from a few hours up to daily coupling.

Implementing a sea ice component into these models is challenging, because the sea ice component of the models provides a boundary condition to both the atmosphere and the ocean, and also simultaneously interferes with both [see, for example, *Hunke et al., 2010*]. To remain numerically stable and to conserve energy, the sea ice surface as seen by the atmospheric model must react to changes in atmospheric heat fluxes almost immediately, while at the same time, the ice is advected by ocean currents and winds, grows and melts at its bottom, and new ice is forming in open water. These processes are all faster than a typical coupling frequency, which is why the handling of sea ice is a major design decision in the construction of any large-scale climate model, and often contains the key programming code through which information between the ocean and the atmosphere are exchanged.

The ultimate decision as to how such coupling is technically implemented is rarely decided by the sea-ice model developers, but instead is a design decision that lies at the heart of model development. This is because the technical details of the implementation and the scientific aims for which a particular model is going to be used often depend on each other. For example, the physically most consistent coupling strategy that might be necessary for the detailed examination of sea-ice related processes may be computationally too expensive to be used for models whose core usage lies in the simulation of climate evolution over many decades or centuries [*see Belcher et al., 2015*].

A standard approach for handling these difficulties relies on the splitting up of the sea ice code into a main component that resides in the ocean model code, and an auxiliary component that resides in the atmosphere model [*e.g., Hewitt et al., 2011;*

Notz et al., 2013]. The auxiliary routine in the atmospheric model consists of a simplified thermodynamic model whose surface temperature changes in response to changes in the incoming atmospheric fluxes, allowing the sea-ice surface to always be consistent with the atmospheric forcing. The heat fluxes that penetrate into the ice are accumulated by the auxiliary model, and then passed to the full sea-ice code that resides in the ocean model at each coupling time step. There, the accumulated fluxes are applied, changing for example ice thickness, and the new sea-ice state is passed back to the atmosphere. Additionally, at each ocean time step, the ice is advected in response to the wind forcing provided by the coupler and the oceanic currents. A detailed description of this approach is provided in the model description of HadGEM3 by *Hewitt et al.,* [2011], who in particular show a detailed flow chart of the various fluxes that are exchanged between the atmosphere model and the sea-ice/ocean model. They also give a detailed description of issues related to the interpolation of these fluxes between the different grids.

A more consistent approach for handling the simultaneous forcing from above and below is based on calculating the fluxes and the response of the sea ice in the sea-ice model itself, which then becomes a fourth entity in the model system in addition to the ocean, the land and the atmosphere components. This approach allows for a more internally consistent simulation of the sea ice, but on the other hand puts a greater burden on the complexity of the sea-ice code itself. Because of its superior physical consistency, this approach is the recommended approach for using the possibly most advanced sea ice model CICE in a climate-modelling context [*Hunke et al., 2015*].

There are two distinctly different methods for using climate models to establish the future evolution of the climate system. The most commonly used method provides long-term projections of the climate system, but usually does not allow for inferences about the short-term evolution of the climate system. For this method, the models are started from a pre-industrial control simulation, during which the model was exposed to the same external forcing for many hundred simulation years until the model has reached an equilibrium state, such that properties of the deep ocean, with its very slow response time, no longer change. Starting from such a pre-industrial control run, the model is then exposed to a time-varying forcing, where for example solar irradiance, volcanic eruptions and the evolution of the atmospheric carbon dioxide fields are inferred from observations.

A major hindrance in comparing such simulations to the observed evolution of the climate system derives from the internal variability of the climate system. For any given change in forcing, the real world will follow only one of the infinite numbers of possible trajectories, while climate-model simulations will follow some other trajectories that are compatible with the change in external forcing. To take this into account, one often uses ensembles of simulations, where the same model simulates the possible evolution of the climate system several times with only slightly different initial conditions. On short time scales, these simulations will differ widely from one another, and also differ from the evolution of the real world. Therefore, inferences about model quality are hard to obtain simply by comparing model simulations to observations over relatively short periods of time. In periods of rapid changes in the mean climate state, such as experienced currently, often not even the standard period of 30 years is sufficiently long to minimize the impact of internal variability to a useful degree [*Notz, 2015*]. This is discussed in more detail in Section 5.

A fundamentally different approach of using climate models is related to simulations of the climate evolution on much shorter time scales than just discussed. For this approach, one tries to push the modeled state of the climate system as closely to the observed state as possible, using various methods of data assimilation. As long as the climate system remembers some properties of its initial state, one can hope to

actually predict the short-term evolution of the climate system on time scales from a few seasons up to a couple years (see section 4 for details). Linguistically, one differentiates between actual *predictions* obtained by this approach and *projections* obtained from the freely evolving simulations described before.

3 Using Models and Observations to Understand the Present

As discussed in the introduction, combining observations with model simulations allows for the most robust insights regarding the evolution of any climate observable. In this section, we exemplify this for sea ice, describing those findings on the ongoing evolution of Arctic sea ice that are compatible with both the observational record and with the available model simulations. We also outline in as how much model simulations are not compatible with the observational record, which then provides an estimate of the reliability of model simulations of the future evolution of sea ice discussed in the following two sections.

The longest consistent record of sea-ice evolution is based on observations of the global sea-ice cover from multichannel passive microwave satellite sensors. Despite their limitations outlined in the previous section, these observations have been invaluable in documenting the dramatic changes occurring in Arctic sea ice over the last several decades. Since the beginning of the continuous satellite record in 1979, passive microwave observations have revealed a significant decline in the spatial extent of sea ice [e.g. *Stroeve et al.,* 2012b], which is one of the clearest records of climatic changes that we have available.

The observed decline of the ice cover is statistically significant in all calendar months [e.g. *Serreze et al*., 2007], but it is most pronounced at the end of the melt season in September, with an average loss of 87,200 km^2 yr^{-1} during the period 1979-2014 [*Serreze and Stroeve*, 2015]. In particular, the 8 lowest September extents all occurred during the past 8 years. In September 2012, the ice extent shrank to 3.41 million km^2, which is the lowest level yet recorded by satellites, equivalent to a 50% reduction compared to the typical observed ice coverage during the 1980s.

As discussed in section 2.1, similarly robust long-term records of sea-ice thickness do not exist, yet the limited observations draw a clear picture, agreeing on a pronounced thinning of the ice cover during recent decades [e.g. *Lindsay and Schweiger*, 2015; *Laxon et al*., 2013; *Rothrock et al*., 1999]. There is also an overall reduction in snow depth [*Webster et al*., 2014] as the sea ice freezes later in autumn than it used to [*Stroeve et al.,* 2014c]. The observed thinning of both snow and ice is among others caused by the large-scale replacement of thick multiyear ice by thinner first-year [e.g. *Maslanik et al*., 2007]. The thin ice makes the ice more prone to complete melting during summer, contributing to the ongoing decrease of sea-ice extent.

Pushing the observational record further back in time, reconstructions of past sea ice conditions from terrestrial proxies (e.g. ice cores, tree-rings, lake sediments) show that the current decline in the sea ice cover is likely unprecedented for the past 1,450 years [*Kinnard et al*., 2011].

These observed changes are largely compatible with changes in Arctic sea ice as simulated by CMIP5 models, which show the sea-ice declines rapidly during the past few decades, both in regards to its spatial coverage and its thickness. More striking, and from a policy perspective more relevant, the model simulations additionally agree with the observational record on the large-scale sensitivity of Arctic sea ice to global warming. This sensitivity is usually expressed as the loss of sea ice per degree Celsius in annual mean temperature change, which has been found to be surprisingly linear in the observational record. *Mahlstein and Knutti* [2012], for example, show that the

evolution of summer sea ice in the observational record is well described by a loss of roughly 2.4 million km^2 of September sea-ice extent per degree of global warming during the period 1980-2007. The CMIP3 models that they analyzed also exhibited a roughly linear relationship between global mean temperature and sea-ice extent, but showed a lower sensitivity than the observational record. This was also found in an earlier analysis of CMIP3 simulations by *Winton* [2011].

For CMIP5 models, however, we find that the multi-model mean agrees almost perfectly with the observed sensitivity (**Figure 1 and Table 1).** The models give on average a sensitivity of 2.3 $\pm$ 0.6 million km^2 loss of September sea-ice per degree of global warming, which is narrowed down to 2.4 $\pm$ 0.4 million km^2 if the two models CSIRO-Mk3-6-0 and FIO-ESM are excluded from the ensemble, since they both clearly have too low a sensitivity. This sensitivity range compares favorably with the sensitivity of 2.4 million km^2 sea-ice loss per degree of global warming that *Mahlstein and Knutti* [2012] derive from the observational record.

We do not yet understand why the models agree so well with the observational record for the sensitivity of sea ice, but the agreement is so striking that the relationship between September sea-ice extent and global mean air temperature is very likely robust. In particular, the good agreement between modelled and observed sea-ice sensitivity very strongly suggests that the models simulate the integrated impact of feedbacks realistically. This is important because the loss of sea ice gives rise to some potentially very strong positive feedbacks, most notably the ice-albedo feedback, which in isolation can self-accelerate the ongoing loss of sea ice until it passes a so-called tipping point, beyond which the loss of the remaining ice would become irreversible. This would then de-couple the evolution of sea ice from the evolution of global mean temperature.

However, a number of studies have shown that such a tipping point is unlikely to exist. By analyzing the short-term evolution of sea ice from one year to the next, *Notz and Marotzke* [2012] showed that the observational record of the retreat of Arctic summer sea ice has so far not been suggestive of a dominating role of positive feedbacks, since the year-to-year changes of the summer Arctic sea ice evolution are significantly negatively correlated: after every strong drop in sea-ice coverage in one summer, the ice always recovered somewhat in the following year. This is also the prevalent behavior of Arctic summer sea ice in CMIP5 simulations [**Figure 2**], which hence agree with the observational record on a net stabilizing role of feedbacks during the ongoing loss of Arctic summer sea ice.

Other model-based studies have also found that negative feedbacks slow down the loss of Arctic summer sea ice, be it the idealized model work of *Eisenman and Wettlaufer* [2009], the conceptual study by *Notz* [2009], the dedicated model simulations by *Tietsche et al.* [2011], or the modelling studies that simulate a reversal of the CO_2 increase as described by *Armour et al.* [2011], *Ridley et al.* [2012] and *Li et al.* [2013]. These studies have in particular found that the stabilizing negative feedbacks are related to (1) the rapid growth of thin ice, (2) snow-related feedbacks, and (3) the increase in outgoing longwave radiation after a sudden loss of summer sea ice [e.g, *Eisenman and Wettlaufer,* 2009; *Notz,* 2009, *Tietsche et al.,* 2012].

These findings allow us to draw some conclusions regarding the driver of the ongoing loss of Arctic sea ice. The negative auto-correlation of year-to-year sea-ice changes and the clear coupling of sea-ice evolution to global mean temperature that are found in both the observations and simulations, suggest that the loss of sea ice is not self-accelerating. Hence, the observed ice loss is only explicable by an external driver. Using only the observational record, *Notz and Marotzke* [2012] showed that the most likely driver of the ongoing loss of Arctic sea ice is the increase in atmospheric greenhouse gas concentration. The same is found in individual CMIP5 model simulations [e.g., *Kay et al.,* 2011, *Notz et al.,* 2013] and across the CMIP5 ensemble,

where a larger proportion of model ensembles show a total loss of Arctic summer sea ice as atmospheric CO_2 concentrations increase further [*Notz*, 2015].

Related to the future evolution of Arctic sea ice, this discussion shows that the models have generally good skill in simulating the large-scale evolution of Arctic sea ice relative to changes in the forcing. This then also allows for robust insights in the future evolution of Arctic sea ice, as discussed in the following sections. We here should, however, note that despite their generally good quality in representing the large-scale evolution of sea ice, the models do not always show good agreement in their representation of details of Arctic sea-ice evolution. This is exemplified in studies that examine, for example, the seasonal cycle in sea ice extent [*Stroeve et al.*, 2012b; *Jahn et al.*, 2012], or the mean and trends in sea ice extent [*Stroeve et al.*, 2012b; *Massonnet et al.*, 2012] or area [*Snape and Forster*, 2014]. While these studies have concluded that the seasonal cycle in ice extent and ice volume are well represented by most climate models, there often are large differences between observed trends and the trends simulated by individual models. It is, however, very challenging to conclude from this mismatch that the models need to be improved, as outlined briefly in section 6 and discussed in more detail by *Notz* [2015]. This is mostly because much of these differences can simply be caused by internal variability of the climate system.

For other metrics, shortcomings of the models are, however, clear. For example, *Pithan et al.*, [2013] found that a number of models have difficulties in correctly reproducing the observed structure of the Arctic atmospheric boundary layer. Also the annual cycle of sea ice albedo, which crucially affects the melt rate driven by short-wave radiation, differs widely across CMIP5 models, and the differences are too large to simply be caused by internal variability [*Koenigk et al.*, 2014].

Regarding sea-ice dynamics, CMIP3 models had severe difficulties in capturing the annual cycle of drift speed [*Rampal et al.*, 2011], and several CMIP5 models fail to simulate the general pattern of atmospheric circulation in the Arctic (e.g. fail to represent the Beaufort Sea high) [e.g. *Stroeve et al.*, 2014b].

To identify the reasons for such shortcomings is usually not straightforward, even if internal variability can be excluded as the primary explanation. On the one hand, the model biases could, obviously, simply be indicative of shortcomings in the formulation of the sea-ice models. On the other hand, shortcomings in the forcing of the sea-ice cover by the atmospheric or the oceanic component of the climate models severely affect the quality of sea-ice simulations in any coupled model as well. The importance of this forcing for model performance is exemplified by the different simulations of sea ice in CMIP5 compared to those in CMIP3. For most models that contributed to both intercomparison studies, the sea-ice component has remained largely unchanged. Hence, the often striking differences in the sea-ice simulations between CMIP3 and CMIP5 are most likely caused by changes in the atmospheric or the oceanic component of the models. Because of these difficulties, process oriented model-evaluation studies, such as the one by *Pithan et al.* [2013], are the most promising route towards a further improvement of large scale sea-ice simulations.

In summary, despite some shortcomings in the details, modern sea-ice simulations capture apparently much of the physics that govern the observed Arctic sea-ice evolution. Models and observations agree on the sensitivity of Arctic sea ice to global warming, they agree on greenhouse gas emissions as the most likely driver for the observed ice loss, and they agree on a main role of negative feedbacks to slow down the ongoing retreat of Arctic sea ice. This agreement on key aspects of Arctic sea-ice evolution suggests that we can also use these models to gain insights into the short-term and long-term future evolution of sea ice on our planet. How observations foster or still limit such insights, and how they can help us to overcome model shortcomings is discussed in the following sections.

4 Using Models and Observations to Predict the Short-term Future

In this section, we review the interplay between observations and models relevant for short-term predictions of sea-ice evolution on time frames from a few months to a few years. Interest in such short-term prediction of Arctic sea ice is spurred in part by the increase in economic activity in the Arctic that has become possible through the ongoing loss of the ice cover. For example, for the economic routing of ship traffic through the Arctic, seasonal predictions of sea-ice evolution are of key value. In addition to such practical applications, short-term predictions also provide insights into the functioning of the climate system on short time scales, which allows model simulations to directly be tested against observations, allowing for ever more realistic representations of the real world in these models.

The increased interest in short-term predictions has led to a rapid evolution of reliable methods that allow one to use numerical models to predict sea ice conditions a few months to a few years in advance [e.g., *Lindsay et al.,* 2008*; Sigmond et al.,* 2013, *Stroeve et al.*, 2014c; *Tietsche et al.*, 2014; *Guemas et al.*, 2014; *Peterson et al.,* 2014]. In order to allow for such forecasts from numerical models, the present conditions of sea ice, ocean and atmosphere need to be assimilated into the model, which then allows the model state to become close to the state of the real world at the onset of the prediction. These forecasts are also generally run as ensemble forecasts in order to capture some of the atmosphere-ocean inherent variability.

A natural question that arises in this context relates to the possible time frame for which meaningful forecasts are possible. This question is best examined in so-called 'perfect model' studies, where a dynamical model is used to generate ensemble pseudo-predictions that are initialized from a slightly perturbed reference simulation. In such setup, the reference simulation is taken as a substitute for the "real world", and one can examine in much detail why and where the pseudo-predictions differ from that "real-world" simulation. Such studies have suggested that monthly pan-Arctic sea ice extent could be significantly predictable for 2 years after initialization. Since these experiments assume perfect knowledge of initial conditions, they provide an upper limit of predictability for a particular dynamic model [see *Guemas et al.*, 2014 for a review].

Such studies also allow one to identify the key parameters that allow for sea-ice predictability, which are hence those parameters that are most relevant from the observational record. The studies agree that their predictive skill is primarily related to memory in the ocean heat content, as expressed by sea surface temperatures (SSTs), and memory in the sea ice itself, as expressed by the distribution of sea-ice volume. For example, sea ice reemergence, where anomalies in sea ice concentrations or area recur at a 5 to 12 month time-lags [*Blanchard-Wrigglesworth et al.*, 2011], has been found to be related to persistence of SST anomalies between spring and fall, and to sea ice thickness anomalies between fall and spring [e.g. *Bushuk et al.*, 2014, 2015]. In another study, *Lindsay et al.*, [2008] examined an extensive set of predictors, including sea ice extent from previous periods, atmospheric circulation indices, ice thickness and ocean temperatures. They found sea-ice concentration to be the best predictor for lead times of 1 to 2 months, and ocean temperatures for longer forecast lead-times, in agreement with *Bushuk et al.*, [2014, 2015]. Atmospheric circulation may play an additional role in sea ice anomaly reemergence from winter to winter [*Bushuk et al.,* 2015].

While these results from "perfect model" studies are promising, the models' skill in predicting the evolution of sea-ice in the real world is usually much lower. Examining the models' predictions of truly observed Arctic sea ice conditions, *Stroeve et al.* [2015] found that they only provided skillful forecasts for a period of 2 to 5 months during summer. Indeed, statistical relationships between various combinations of atmospheric,

sea ice and oceanographic variables often outperform the dynamical models in their skill to forecast the sea-ice state some months in advance. One example for such relationship is given by *Drobot et al*. [2006], who developed a statistical model to predict the pan-Arctic minimum extent at monthly intervals from February through August using combinations of one or two weighted indices of sea ice concentration, skin temperature, albedo and downwelling longwave radiation. All combinations performed better than persistence.

The disappointing results of the numerical models in forecasting the evolution of "real" Arctic sea ice may originate from several limitations of the forecast systems, such as insufficient or poor quality initial conditions for data assimilation. These can lead to an initial state of the model that is too different from the real world to allow for meaningful forecasts. This is exemplified by **Figure 3**, which shows how the forecast of Arctic air temperature depends on the sea-ice concentration data set that was used to initialize the model. Everything else unchanged, a mere change from using the Bootstrap sea-ice concentration data set to using the NASA Team data set to initialize a simulation in May causes regional differences in simulated September air temperature of more than 3 °C [*Bunzel et al.,* submitted].

Other observables have simply not been available to be used for model initialization, such as information on sea-ice thickness distribution. This lack contributes to the relatively low skill of models to forecast sea-ice conditions beyond a few months, since sea-ice thickness has been found to be important for improving predictive skill in short-term forecasts (up to 8 months ahead) [*Day et al*., 2014].

The forecast systems also suffer from insufficient knowledge of the uncertainty of observations. Such uncertainty information is often not present at all, or only represents a lower bound on the true uncertainty, as realistic uncertainty information cannot be provided. One example of this shortcoming for the case of sea ice concentration data sets is their inability to provide a robust estimate of the uncertainty stemming from melt-pond coverage during summer. The necessity of uncertainty information in data-assimilation schemes derives from the fact that any such data assimilation always interferes to a certain degree with the model's own physical state. In order to assess the degree to which the model should be trusted versus the degree to which the observational data should be trusted, uncertainty information is crucial. Indeed, sophisticated data assimilation schemes require observational uncertainty be provided.

The poor predictive skill in real-world applications might additionally be related to insufficient or inadequate model physics, and model drift, while it is ultimately possible that inherent predictability in nature may be lower than that in climate models. New model physics developments that address important sea ice processes such as the life cycle of melt ponds on sea ice show promising results in enhancing predictive skill of sea ice extent. For example, *Schroeder et al*. [2014] demonstrated a strong correlation between spring modeled melt pond fraction using improved melt pond parameterizations and observed September sea ice extent. Interestingly, the statistical model of *Schroeder et al*. [2014] currently outperforms dynamical models in predicting the pan-Arctic September sea ice extent.

Finally, the chaotic nature of the atmospheric response to any given oceanic forcing is of relevance for understanding the limited predictability of Arctic sea-ice conditions. In the perfect-model simulations, the oceanic state is fully known, so the limited predictability found by these studies suggests a dominating influence of the atmospheric response to these oceanic conditions on setting the forecast horizon. However, these relationships are not fully understood yet, and more work needs to be done before we can settle on the forecast horizon for Arctic sea ice. One key challenge for such work will be to quantify the limiting contributions of different structural and parameter

uncertainties, similar to the work by *Hodson et al.* [2013] on long-term climate projections.

Such analysis would then also allow us to examine if the forecast horizon of sea ice is currently decreasing. This might be the case because in general, a thinner ice cover exhibits increased interannual variability, as has been argued both conceptually [*Notz*, 2009] and through an analysis of climate-model simulations [e.g. *Holland et al.*, 2010]. However, such increased variability has yet to manifest in the observations. For example, the standard deviation in September extent is $0.59(10^6)$ km^2 from 1979 to 2004 and only slightly higher at $0.68(10^6)$ km^2 for the past decade of significant thinning and ice extent reductions [see also *Serreze and Stroeve*, 2015]. If interannual variability increases in the future as the ice cover thins further, short-term predictability may decrease even further [e.g. *Tietsche et al.,* 2013]. Unfortunately, few model studies addressing this issue have been performed to date and the evolution of predictive skill remains unclear. What is clear is that there is strong need to better characterize the contribution of uncertainty in short-term forecasts from imperfect initial conditions, imperfect model physics and the influence of the chaotic atmospheric and oceanic forcing. Improvements in the observational data sets used for data assimilation will also allow us to better identify shortcomings in the physical representation of key processes that drive sea-ice evolution on these short time scales. Such improvements will then eventually also lead to improvements in long-term projections, which is the topic of the next section.

5   Using Models and Observations to Project the Long-term Future

On the longer-time scale, estimates of future sea ice conditions are typically based on modeled projections of future ice extent under different greenhouse gas (GHG) warming scenarios. Depending on the future evolution of GHG concentrations, global climate models project a possible seasonally ice-free Arctic before the end of this century [e.g. *Stroeve et al.*, 2012b; *Collins et al.,* 2013; *Overland et al.*, 2014], though results differ widely between models and GHG emission scenarios [**Figure 4**]. The large spread in model projections for any given emission pathway (colored shading, which represents $\pm$ 1 standard deviation of the mean value) is in part caused by internal variability of the climate system and in part by a different skill of the models to reproduce the physical processes that drive sea-ice concentration in the real world. The difference between the individual colored regions, in contrast, is entirely caused by the differences in future emission pathways that have been used to force the models.

These uncertainties, which all currently limit our ability to provide concrete long-term predictions of Arctic sea-ice evolution, are hence fundamentally different from each other: We will never be able to overcome the fact that internal variability sets a clear limit on how precisely we can ever simulate the future ice cover. Uncertainty of future emission pathways, in contrast, might decrease substantially, once policy makers agree on specific future emission targets in a binding treaty. Finally, uncertainty derived from the different representation of physics in our models will ideally decrease the more we understand how sea ice, the ocean and the atmosphere interact with each other.

In light of these uncertainties, it is clearly a desirable aim from a policy perspective to narrow down the uncertainty range of future sea-ice evolution based on observational information we have available today. In this section, we review the two most widely used approaches to do so. One of these approaches relies on the use of emergent constraints, where robust relationships are sought from model simulations that link the evolution of a specific observable in the past to the future evolution of the observable of

interest. The other approach relies on model weighting. For this approach, one gives greater weight to simulations from models whose physics best represents the real world, or, indeed, to simulations whose internal variability by chance matches that of the real world over the past few years. We will start with a detailed discussion of this approach in the following subsection, and then move on to discussing the approach based on emergent constraints in section 5.2.

5.1 Reducing uncertainty range through model weighting

A number of studies have tried to narrow down projection ranges of future sea-ice evolution by sub-selecting models based on their performance in simulating the observed past evolution of climate [e.g. *Snape and Forster*, 2014; *Wang and Overland*, 2012; *Massonnet et al.*, 2012]. Depending on the concrete method of model selection, these studies estimate the timing of a seasonally ice-free Arctic to be possibly as early as the late 2020s [*Wang and Overland*, 2012], 2032 [*Snape and Forster*, 2014], or sometime between 2042 and 2060 [*Massonnet et al.*, 2012].

To choose which models to sub-select or weight, climate model simulations are typically evaluated to see how well a particular model can simulate the observed features of the present-day sea ice cover. However, as discussed in detail by *Notz* [2015], internal variability of the climate system will always limit any model's capability to reproduce the observed evolution of the sea ice cover, independent of the model's quality. Because of this, a sub-selection of models based on their performance over the past few years does not necessarily select the best models, but might simply prefer those models whose random sequence of internal variability matches, by chance, the observed random sequence of internal variability. Given the comparably short memory of the Arctic sea-ice cover that has been found in seasonal prediction studies, it is not clear per se that those selected models will also in the future outperform the models that failed the selection criterion.

One way to overcome this issue is to try and minimize the impact of internal variability on the quality assessment of a particular model. This, however, is only possible if the time-period over which the model can be evaluated is long enough to adequately capture important modes of internal variability. This requires a sufficiently long observational record. During periods of rapidly changing climate conditions, the often-used period of 30 years is not "sufficiently long" to minimize the impact of internal variability [*Notz*, 2015]. Hence, for many metrics that are used for model selection, internal variability remains so large that it hinders a meaningful selection of models based on their quality. This is most notably the case for the trend in both sea-ice area and sea-ice extent.

An example of the impact of internal variability on Arctic sea ice is given by the observed shift in the spring sea ice thickness distribution that occurred in the early 1990s. This shift coincides with a strongly positive phase of the Northern Annual Mode that increased the dynamic export of sea ice from the Arctic [*Rigor and Wallace*, 2004]. A coupled climate model could only ever capture the timing of this "preconditioning" of the icepack by chance, which then directly causes the rates of sea-ice loss to differ between observations and models. As another example, the Atlantic Multidecadal Oscillation (AMO) is found to have a significant impact on Arctic sea ice [*Miles et al.*, 2014], such that multidecadal (60-90 year) fluctuations have been observed in the sea ice cover related to natural fluctuations in the North Atlantic. Any modeled trend will therefore depend on where a model climate is in this multidecadal cycle. If the models are not in phase with the current AMO warm phase, they will not be able to reproduce the recent acceleration of ice loss in the Arctic correctly.

To estimate the impact of internal variability, often ensemble simulations are used, where several simulations are started from slightly differing initial conditions. The spread between these simulations then allows for an estimate of the impact of internal variability. For trends or means over many years, the ensemble size can synthetically be increased further by slightly shifting the start and end date used for the averaging period. *Notz* [2015] found that for such synthetically increased ensemble size, the trend of sea-ice area and sea-ice volume simulated by almost all CMIP5 models is no longer significantly different from the observed evolution. In contrast, if only a single time period is considered, for most CMIP5 models, the ensemble size becomes too small to estimate a realistic value of internal variability.

We here exemplify this point further by examining the observed trends in sea-ice extent, which change rapidly because of internal variability. Previously *Stroeve et al.* [2012b] evaluated trends in September sea ice extent from 1979 to 2011 from 56 ensemble members from 20 models in the CMIP5 archive and found that 46 ensemble members had trends outside of 2σ of the observations, or 82%. Extending the analysis to 107 ensemble members from 35 climate models, we here evaluate the sensitivity of the trend analysis for four changing end-dates, namely 2000, 2006, 2011 and 2014. These years are chosen in part to reflect the more recent acceleration in ice extent during the past decade [*Serreze and Stroeve*, 2015]. We find that the significance of modelled trends being either different from zero or being different from observations depends strongly on the observational period [**Table 2**]. For example, at the 99% confidence level, only 10% of the ensembles have trends different from those observed when evaluating trends through 2000; including 2014 increases that number to 47%. There is a larger increase in the number of ensemble trends that are different from 0, from 25% for trends through 2000 to 65% through 2014 (at 99% confidence). This analysis suggests that not only will the time-period of evaluation influence which models may be rejected by trend analysis, but it also suggests that if a model gets a trend correct in a chosen hindcast or "calibration" time-period, this does not imply the same model will get the evolution of the ice cover correct in some future time-period. This further underpins the finding that trends and 30-year means are insufficient metrics for evaluating model performance.

As another way to illustrate this graphically, we compare trends for 1953 to1990 to those from 1991 to 2013, using the combined and homogenized Had1SST and NSIDC sea ice record to estimate the observed trends. Results are shown in **Figure 5.** Symbols represent the trends calculated for individual model ensemble members. A large number of individual simulations match the observed trend over the 1953-1990 time-period (within $\pm 2\sigma$). However, selecting just these models in order to obtain a better estimate of sea-ice evolution during the second period would have been counter-productive: the average trend across these models is further away from the observational record over the second period than the average calculated across all simulations.

Another issue with the approach of model weighting pertains as to what sea ice metrics, or other aspects of the climate system are most useful for predicting the future evolution of the ice cover. In general, such metric must project onto the future projection of the target quantity, such as the timing of when the Arctic will be ice-ice free in September (defined as less than 1 million km^2). If such projection is not robust, then the weighting on that metric in an average prediction is not particularly useful (nor will it change the answer appreciably). One candidate for a useful metric could be the current distribution of ice thickness, which is used in a number of statistical models to provide seasonal forecasts of sea-ice extent. However, during the period 1979-2013, detrended September sea ice extent shows comparably little correlation with detrended March ice thickness as obtained from the PIOMAS reanalysis [**Figure 6**]. The

comparably low correlation (r=0.58) indicates that other factors, such as summer atmospheric circulation, often overrule the importance of the initial ice thickness on seasonal time scales. For the future, however, the relationship between the sea-ice thickness in spring the sea-ice extent at the end of summer might become more robust: *Holland and Stroeve* [2011] showed that the variance of September sea ice extent anomalies explained by winter-spring ice thickness increases as the ice-cover thins and transitions towards a seasonal ice cover, with summer circulation anomalies playing a lesser role.

The relatively low current correlation between sea-ice thickness anomalies and anomalies in sea-ice extent some months later exemplifies that in addition to the modeled sea-ice state itself, the forcing of the sea-ice cover by the ocean and the atmosphere is a dominating driver of sea-ice evolution. For example, *Stroeve et al.* [2014b] showed that not all climate models capture the general sea-level pressure pattern over the Arctic, with important implications on the spatial pattern of ice thickness. The long-term annual mean Arctic sea level pressure pattern is defined by high pressure over the Beaufort Sea and a trough of low pressure extending from the Icelandic Low northeastward into the Kara Sea. This circulation pattern helps to drive the clockwise Beaufort Gyre and the Transpolar Drift stream, which in turn helps to sequester thick ice in the Canada Basin and transport ice out of the Arctic Ocean through Fram Strait, respectively. The failure of some models to reproduce these and other similarly important aspects of the observed forcing of the sea-ice coverage can contribute to unrealistic projections of sea-ice evolution, but studies are largely lacking that take such more indirect metrics into account when selecting the most realistic sea-ice projections.

The discussion so far has outlined some of the conceptual difficulties that one faces when trying to reduce the uncertainty range of sea-ice projections by weighting models according to their performance relative to observations. Even worse, the uncertainty range that one reaches by model selection might be a less realistic representation of the most likely future evolution than that given by the full ensemble. This can even be the case for metrics that show a clear correlation between the variable of interest and a performance metric. For example, while the 107 CMIP5 simulations that we analyze here show a clear correlation of 0.8 between mean September sea ice extent in 1953-1990 and sea-ice extent in 2050-2070, the correlation decreases to 0.67 if we exclude those 50 simulations from 22 models whose trends from 1979 to 2014 do not match the observed trend. This would argue against removing models based on analysis of trends in extent. Furthermore, removing models has little impact on the timing of when the Arctic first becomes ice-free. A random sample of ensembles (100 samples) compared to the distribution from models selected by matching the observed trend shows that random selection of models yields ice-free dates within the 25th to 27th percentile of those sub-selected, and often matches the median extent from the "best fit" models [**Figure 7**].

5.2 Reducing uncertainty using all models through emergent constraints

As outlined in the previous section, it is very difficult to use model weighting in order to robustly narrow down the uncertainty range of, say, the timing of the loss of Arctic summer sea ice [compare also *Notz*, 2015 for a related discussion]. There is, however, a different approach to combine observations with model simulations that is not based on model weighting, but instead is most robustly used if the largest possible number of simulations is kept for the analysis. This method is based on emergent constraints, which are climate observables that constrain the most likely future evolution of another climate observable [c.f., *Santer et al.*, 2008; *Knutti et al.*, 2010;

Collins et al., 2012; *Notz*, 2015]. In practice, one establishes such constraints by examining either the observational record or an ideally wide range of model simulations for a statistical relationship between the past evolution of the constraining observable and the future evolution of the observable of interest. If such statistical relationship is established, one can then use the observed value of the constraining variable to estimate the future evolution of the observable of interest.

For the case of sea ice, a number of studies have used this approach to narrow down the uncertainty range of future sea-ice evolution. *Mahlstein and Knutti* [2012] used the relationship between global mean temperature and sea-ice evolution described in section 3 to estimate that Arctic summer sea ice will largely be gone for a global warming of 2°C. *Boe et al.* [2009] found a correlation in CMIP3 model simulations between the trend in sea-ice extent from 1979-2007 and the percentage of remaining Arctic sea ice in 2021-2040, using this relationship to estimate a loss of Arctic summer sea ice for a medium-range emission scenario within this century. A similar analysis was also given by *Collins et al.* [2012] in the broader context of constraining future projections. Examining CMIP5 model simulations, *Liu et al.* [2012] showed that not only the trend, but also the sea-ice extent itself during the period 1979-2007 is directly related to the sea-ice extent during the period 2018-2022. They use this relationship to estimate 2054-2058 as the most likely time period for the loss of Arctic summer sea ice in a high-emission scenario.

Despite possibly having some conceptual advantages over the approach of selecting or weighting models to narrow down projection ranges of future sea-ice evolution, the approach of using emergent constraints has clear limitations, too. This is already obvious from the different estimates of future sea-ice evolution that are seen in studies that differ in their underlying metrics. Hence, even if clear statistical relationships are established, the robustness of these relationships is not necessarily an indication for their relevance for the real world, as exemplified for Arctic mean temperature in a recent study by *Bracegirdle and Stephenson* [2013]. For example, if the statistical relationship that one establishes is based on climate-model simulations, these models might all share similar shortcomings that could impact the statistical relationship one establishes. In addition, in particular for relationships based on observations, one has to assure that the relationships are truly based on emergent constraints, rather than simply on a random fluctuation that might overshadow a true relationship. For relationships from multi-model ensembles, one additionally has to establish that all models contribute similarly to the statistical relationships one establishes. Outliers, for example, could otherwise obtain too much weight in establishing a specific statistical relationship. Finally, as discussed for Arctic sea ice by *Notz* [2015], it is not always clear that a relationship that held in the past will also hold for the future. For example, in the case of sea ice, the net thinning of thick ice is usually much larger than the net thinning of thin ice for any given warming, which is related to the fact that thin ice more easily re-gains its summer ice loss during the following winter [compare, *Bitz and Roe*, 2004]. Hence, it is not clear per se that a statistical relationship that holds for the thick sea ice that used to cover the Arctic Ocean in the past will also hold for the much thinner sea ice that will cover the Arctic Ocean in the future. Indeed, it is likely that this will not be the case, as exemplified by the slowdown of sea-ice loss that is shown across CMIP5 model simulations once September sea-ice extent drops below a few million km^2 [*Massonnet et al.*, 2012]. These factors all limit the reliability of emergent constraints on reducing future uncertainty ranges, which can be minimized to focus on those relationships that hold both for the observational record and a wide range of model simulations. Hence, again, most robust insights are gained from a combined analysis of both observations and simulation.

6 Discussion and Conclusions

The decline in Arctic sea-ice extent over the past several decades is one of the most visible indicators of large-scale environmental change. For the future, climate models project an ongoing decline of the sea-ice cover, as the climate warms further in response to rising concentrations of atmospheric greenhouse gas concentrations. This ongoing substantial ice loss has profound ecological, climatic, economic and societal consequences, including the extinction of marine species, substantial shifts in the prevailing livelihood of indigenous population, large-scale climate feedbacks and increased economic activity. Such changes obviously have important ramifications for public policy and decision making at all levels of government.

A promising approach to most reliably inform policy makers about the future evolution of sea ice, as well as to understand the main drivers of its ongoing loss, is based on the combination of insights from the observational record with large-scale model simulations. As reviewed in this contribution, such combination allows us to both estimate the reliability of the models and to possibly narrow down the uncertainty range of future sea-ice evolution.

We have shown that in many of their large-scale aspects, modern climate models capture much of the physics of the observed evolution of the Arctic sea-ice cover. Their simulated sensitivity of the sea-ice cover to global warming matches the observational record very well, and the models agree with the observations that as a whole, sea-ice related feedbacks currently slow down the loss of sea ice that is driven by the increase in atmospheric greenhouse gas emissions.

Despite their skill in capturing some of the integrated features of Arctic sea-ice evolution, many models do not capture more specific aspects of the observed Arctic sea-ice state. For example, on average the modelled decline of Arctic sea ice is somewhat slower than has been observed [e.g., *Stroeve et al.*, 2012; *Massonnet et al.*, 2012], the models do not capture the observed evolution of sea-ice albedo [*Koenigk et al.*, 2014], and often fail to represent the historical ice thickness pattern [*Stroeve et al.*, 2014b]. However, the reason for these mismatches between observations and models are often hard to identify. For example, internal variability alone can explain why sea-ice retreat is slower in many models than it is in observations [*Notz*, 2015]. The erroneous patterns of ice thickness, in contrast, are most likely caused by unrealistic forcing of sea-ice movement in the models that stems from unrealistic modeled patterns in air pressure [*Stroeve et al.*, 2014b]. Finally, obviously shortcomings in the sea-ice model component will contribute to all these issues, but the relative weighting of the various contributions to model biases remains difficult.

This ambiguity also hinders us to robustly narrow down the uncertainty range of future sea-ice evolution. This is in particularly the case for approaches that are based on model sub-selection. As discussed here, it remains often unclear if models are selected because they better represent the relevant physics, or simply because by chance they match the timing of the observed internal variability with little relevance for the future evolution of the ice cover. In the light of substantial internal variability, the full multi-model mean therefore can indeed provide a more reliable projection of future conditions than can any of its subset, because the full ensemble samples the largest possible range of internal variability.

It is, however, not fully clear if answering the question as to when the Arctic will have lost its summer sea ice is of true relevance for stakeholders, or whether it primarily is useful as an "Alarm Bell" headline, a wake-up call that the loss of sea ice will have large impacts on society and the planet. The fact that sea ice will decrease further for further increasing greenhouse gas emissions is established largely beyond doubt, and the exact decade in which the Arctic Ocean may become seasonally ice-free makes little

difference to emission policies. Rather it may have some implications in future international treaties on mineral exploration and trade and on national security policies.

Because of the difficulties just outlined, much of our recent insights into the functioning of sea ice in response to changes in atmospheric or oceanic forcing has not only been obtained from studies on the long-term evolution of the ice, but additionally from the increasing body of literature examining the evolution of sea ice on seasonal and inter-annual time scales. On these time scales, forecasts of ice conditions are particularly helpful to support local communities, marine transport operations, tourism, and resource extraction operations. While dynamic models have demonstrated predictive skill may be possible for 2 years after initialization using perfect model experiments, this has yet to translate into useful forecasts beyond a couple of months. This is in part based on the importance that sea ice dynamics plays in prediction. Currently numerical weather forecasts can give good forecasts out to about 7 days, after which there is little information (or skill). Thus, while perfect model forecasts offer some hope that the models can provide useful short-term forecasts, they often lack the data initialization and forcing data to improve forecast skill. At present, statistical models [e.g. *Schroeder et al.*, 2014] have provided the greatest skill in forecasting September sea ice extent. Notwithstanding any progress in the field of sea-ice predictability, it seems rather certain that the temporal limit of predictability will not be longer than a few years at maximum.

The seasonal-prediction studies have also established a new focus on the regional evolution of the sea-ice pack. For longer-term projections, most studies to date have in contrast focused on predicting total pan-Arctic ice extent. This information however, is not informative to conducting operations in the Arctic, such as a shipping company trying to decide how long the Northern Sea Route will be navigable each year. More useful information includes spatial maps on the probability of sea ice presence, thickness of the ice, and timing of freeze-up or melt onset. In response, the Sea Ice Prediction Network (SIPN, http://arcus.org/sipn) has been encouraging the research community to include spatial forecasts and maps in their seasonal predictions of the September sea ice, using statistical or dynamical model forecast systems.

For the future, much promise to develop better and more realistic simulations of Arctic sea ice evolution lies in model-evaluation studies that aim at improved simulations either through better understanding of key processes, or by examining the short term evolution of sea ice on seasonal and annual time scales that allow for a robust model evaluation in response to short-term changes in the forcing. In general, we recommend increased efforts to differentiate between the various factors that contribute to model biases, most importantly internal variability, uncertainties in the forcing, and shortcomings of the sea-ice model itself [compare *Notz*, 2015]. In particular the upcoming CMIP6 protocol should ensure that sufficient model output is made available for the different MIPs as well as for the planned DECK experiments that allow for such improved understanding of model biases.

To conclude, in this contribution we have examined the relationship between observations and models to obtain robust insights regarding the future evolution of sea ice. Such combination can allow one to constrain the most likely future evolution of Arctic sea ice, but the robustness of related methods is often not clear.

Despite the fact that the models clearly can be improved further to provide more reliable insights into the future evolution of sea ice, the combination of observations with model simulations already now provides for a number of robust insights regarding the past and future evolution of sea ice. From such combination, it has by now become clear that the ongoing loss of Arctic sea ice is directly coupled to increasing greenhouse gas emissions, that it the ice loss will continue as long as global temperature keeps rising, and that the loss of summer sea ice is not substantially self-accelerating. This, in

turn, also means that a binding treaty that keeps global CO_2 emissions on par with, say, the RCP2.6 emission scenario would imply a high probability that the complete loss of Arctic summer sea ice can still be avoided.

Acknowledgements

Table 1. Dependence of September sea ice extent on global and Arctic (60-90°N) 2m air temperatures from 26 CMIP5 climate models under the RCP8.5 emission scenario. The multi-model mean indicated by * is based on excluding CSIRO-Mk3-6-0 and FIO-ESM from the ensemble.

Model	Global 2m Air Temperature vs. September SIE		Arctic 2m Air Temperature vs. September SIE	
	Slope	Std. Dev.	Slope	Std. Dev.
ACCESS1-0	-2.43	0.37	-1.22	0.45
ACCESS1-3	-2.65	0.23	-1.21	0.19
bcc-csm1-1	-2.69	0.45	-1.20	0.37
BNU-ESM	-1.74	0.33	-0.73	0.34
CanESM2	-1.84	0.28	-0.84	0.27
CCSM4	-2.12	0.14	-1.06	0.10
CESM1-CAM5	-3.03	0.45	-1.34	0.36
CNRM-CM5	-3.16	0.38	-1.35	0.24
CSIRO-Mk3-6-0	-1.04	0.05	-0.73	0.06
EC-EARTH	-1.79	0.36	-0.71	0.32
FGOALS-g2	-2.81	0.35	-1.15	0.29
FIO-ESM	-0.20	0.35	-0.64	0.20
GISS-E2-R	-2.25	0.19	-1.28	0.11
HadGEM2-AO	-1.86	0.45	-0.80	0.49
HadGEM2-CC	-2.47	0.21	-1.00	0.27
HadGEM2-ES	-2.14	0.29	-0.85	0.37
inmcm4	-2.25	0.35	-1.05	0.26
IPSL-CM5A-LR	-2.04	0.26	-1.04	0.24
IPSL-CM5A-MR	-2.51	0.46	-1.35	0.33
MIROC-ESM	-2.51	0.75	-1.11	0.64
MIROC-ESM-CHEM	-2.70	0.80	-1.14	0.65
MIROC5	-2.87	0.28	-1.16	0.15
MPI-ESM-LR	-2.59	0.30	-1.24	0.24
MPI-ESM-MR	-2.91	0.29	-1.45	0.21
MRI-CGCM3	-2.77	0.41	-1.27	0.29
NorESM1-M	-2.53	0.26	-0.94	0.20
Multi-model mean	**-2.30** **(-2.44)**[*]	**0.63** **(0.40)**[*]	**-1.07**	**0.22**

Table 2. Number of models for which September sea ice extent trends for 1979 to 2000, 2006, 2011 and 2014 are statistically significant from the observed trend or from 0 at 90%, 95% and 99% confidence levels.

End of Period	Observed vs. Modeled Trend			Different from 0		
	Sig 90%	Sig 95%	Sig 99%	Sig 90%	Sig 95%	Sig 99%
2000	28	54	11	54	44	27
2006	33	28	20	81	70	53
2011	63	57	46	95	93	71
2014	62	59	50	99	93	79

7 References

Armour, K., I. Eisenman, E. Blanchard-Wrigglesworth, K. McCusker, and C. M. Bitz, (2011): The reversibility of sea ice loss in a state-of-the-art climate model. *Geophys. Res. Lett.*, 38, L16 705, doi:10.1029/2011GL048 739.

Belcher, S.E., H. T. Hewitt, A. Beljaars, E. Brun, B. Fox-Kemper, J.-F. Lemieux, G. Smith, S. Valcke (2015): Ocean-waves-sea ice-atmosphere interactions. In: Seamless Prediction of the Earth System: from Minutes to Months, WMO-No. 1156 edited by Gilbert Brunet, Sarah Jones, Paolo Ruti, World Meteorological Organization.

Bitz, C. M., and G. H. Roe (20040, A Mechanism for the High Rate of Sea Ice Thinning in the Arctic Ocean, *J. Climate* 17(18), 3623–3632.

Blanchard-Wrigglesworth, E., K.C. Armour, C.M. Bitz and E. DeWeaver (2011), Persistence and inherent predictability of Arctic sea ice in a GCM ensemble and observations, *J. Climate*, 24, 231-250.

Boé, Julien, Alex Hall, and Xin Qu. (2009), September Sea-Ice Cover in the Arctic Ocean Projected to Vanish by 2100, *Nat. Geosci.,* 2 (5): 341–43. doi:10.1038/ngeo467.

Bracegirdle, T. J., and D. B. Stephenson (2012), On the Robustness of Emergent Constraints Used in Multimodel Climate Change Projections of Arctic Warming. *J. Clim.* 26 (2): 669–78. doi:10.1175/JCLI-D-12-00537.1.

Bushuk, M., D. Giannakis, and A. J. Majda (2014), Reemergence mechanisms for North Pacific sea ice revealed through nonlinear Laplacian spectral analysis, *J. Climate,* 27, 6265-2687.

Bushuk, M., D. Giannakis, and A. J. Majda (2015), Arctic sea-ice reemergence: The role of large-scale oceanic and atmospheric variability, *J. Climate*, doi:10.1175/JCLI-D-14-00354.1, in press.

Cavalieri, D. J., C. L. Parkinson, and K. Y. Vinnikov (2003), 30-year satellite record reveals contrasting Arctic and Antarctic decadal sea ice variability, *Geophys. Res. Lett.* 30, doi: 10.1029/2003GL018031.

Cavalieri, D. J., P. Gloersen, and W. J. Campbell (1984), Determination of Sea Ice Parameters with the NIMBUS-7 SMMR, *J. Geophys. Res.,* 89(D4):5355-5369.

Collins, M., R. E. Chandler, P. M. Cox, J. M. Huthnance, J. Rougier, and D. B. Stephenson (2012), Quantifying future climate change, *Nature Clim. Change*, 2 (6), 403–409, doi:10.1038/nclimate1414.

Collins M, Knutti R, et al. (2013), Long-term climate change: projections, commitments and irreversibility. In Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Edited by Stocker T, Qin D, Plattner GK, Tignor M, Allen S, Boschung J, Nauels A, Xia Y, Bex V, Midgley P, Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press 2013:1029–1136.

Comiso, J.C., (1986), Characteristics of Arctic winter sea ice from satellite multispectral microwave observations, *J. Geophys. Res.,* 91, doi:10.1029/JC091iC01p00975.

Day, J.J., E. Hawkins and S. Tietsche (2014), Will Arctic sea ice thickness initialization improve seasonal forecast skill?, *Geophys. Res. Lett.*, 41(21), 7566-7575, doi:10.1002/2014GL061694.

Eisenman, I. and J. S. Wettlaufer, 2009: Nonlinear threshold behavior during the loss of Arctic sea ice. *Proc. Natl. Acad. Sci.*, 106, 28–32.

Fetterer, F., K. Knowles, W. Meier, and M. Savoie (2002), updated 2009. *Sea Ice Index,* Boulder, CO: National Snow and Ice Data Center. Digital media.

Ford, J.D., W.A. Gough, G.J. Laidler, J. MacDonald, C. Irngaut and K. Qrunnut (2009), Sea ice, climate change and community vulnerability in nortern Fox Basin, Canada, *Clim. Res.,* 38, doi:10.3354/cr00777.

Francis, J., W. Chan, D.J. Leathers and J.R. Miller, (2009), Winter Northern Hemisphere weather patterns remember summer Arctic sea-ice extent, *Geophys. Res. Lett.*, 36, doi:10.1029/2009Gl037274.

Guemas, V., E. Blanchard-Wrigglesworth, M. Chevallier, J.J. Day, M. Déqué, F.J. Doblas-Reyes, N.S. Fučkar, A. Germe, E. Hawkins, S. Keeley, T. Koenigk, D. Salas y Mélia, and S. Tietsche (2014), A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales: Arctic Sea-Ice Predictability and Prediction, Quarterly Journal of the Royal Meteorological Society, n/a-n/a, doi:10.1002/qj.2401.

Hewitt, H. T., D. Copsey, I. D. Culverwell, C. M. Harris, R. S. R. Hill, A. B. Keen, A. J. McLaren and E. C. Hunke, Design and implementation of the infrastructure of HadGEM3: the next-generation Met Office climate modelling system, Geosci. Model Dev., 4, 223-253, doi:10.5194/gmd-4-223-2011,2011.

Hodson, D. L. R., S. P. E. Keeley, A. E. West, J. Ridley, E. Hawkins and H. T. Hewitt, Identifying uncertainties in Arctic climate change projections, Clim. Dyn., doi: 10.1007/s00382-012-1512-z, 2012.

Holland, M.M., D.A. Bailey and S. Vavrus (2010), Inherent sea ice predictability in a rapidly changing Arctic enviroment of the Community Climate System Model, version 3, *Clim. Dyn.*, doi:10.1007/s00382-010-0792-4.

Holland, M.M., and J.C. Stroeve (2011), Changing seasonal sea ice predictor relationships in a changing Arctic climate, *Geophys. Res. Lett.,* doi:10.1029/2011GL049303.

Hunke E.C., Lipscomb W.H. and A.K. Turner (2010), Sea-ice models for climate study: retrospective and new directions. *J. Glaciol.*, 56:1162–1172. doi:10.3189/002214311796406095.

Hunke, E.C., W.H. Lipscomb, A. K. Turner, N. Jeffery, S. Elliott (2015), CICE: the Los Alamos Sea Ice Model Documentation and Software User's Manual Version 5.1, Los Alamos National Laboratory, Los Alamos, NM, US.

Ivanova, N., and others, (2015), Satellite passive microwave measurements of sea ice concentration: an optimal algorithm and challenges, *The Cryosphere Discuss.*, 9, doi:10.5194/tcd-9-1269-2015.

Jahn, A., K. Sterling, M.M. Holland, J.E. Kay, J.A. Maslanik, C.M. Bitz, D.A. Bailey, J. Stroeve, E.C. Hunke, W.H. Lipscomb, D. Pollak (2012), Late 20th century simulation of Arctic sea ice and ocean properties in the CCSM4, *J. Climate*, doi:10.1175/JCLI-D-11-0-0201.1.

Jaiser, R., K. Dethloff, D. Handor, A. Rinke and J. Cohen, (2012), Impact of sea ice cover changes on the Northern Hemisphere atmospheric winter circulation, *Tellus*, 64, 11595, doi:10.3402/tellusa.v64i0.11595.

Jones, B.M., C.D. Arp, M.T. Jorgenson, K.M. Hinkel, J.A. Schmutz and P.L. Flint (2009), Increase in the rate and uniformity of coastline erosion in Arctic Alaska, *Geophys. Res. Lett.*, 36, doi:10.1029/2008GL036205.

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

Kinnard, C., C.M. Zdanowicz, D.A. Fisher, E. Isaksson, A. de Vernal and L.G. Thompson, (2011), Reconstructed changes in Arctic sea ice over the past 1,450 years, *Nature*, 479, 509-512, doi:10.1038/nature10581.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl, (2010), Challenges in Combining Projections from Multiple Climate Models, *J. Climate*, 23(10), 2739–2758.

Koenigk, T., A. Devasthale, A., and K.-G. Karlsson (2013).: Summer Arctic sea ice albedo in CMIP5 models, *Atmos. Chem. Phys*., 14, 1987-1998, doi:10.5194/acp-14-1987-2014, 2014.

Laxon, S.W., K.A. Giles, A.L. Ridout, D.J. Wingham, R. Willatt, R. Cullen, R. Kwok, A. Schweiger, J. Zhang, C. Haas, S. Hendricks, R. Krishfield, N. Kurtz, S. Farrell, M. Davidson, (2013), CryoSat-2 estimates of Arctic sea ice thickness and volume, *Geophys. Res. Lett., 40(4): 732-737.*

Li, Chao, Dirk Notz, Steffen Tietsche, and Jochem Marotzke, (2013), The Transient versus the Equilibrium Response of Sea Ice to Global Warming, *J. Climate,* 26 (15): 5624–36. doi:10.1175/JCLI-D-12-00492.1.

Lindsay, R. W., J. Zhang, A. J. Schweiger, and M. A. Steele (2008), Seasonal predictions of ice extent in the Arctic Ocean, *J. Geophys. Res*., 113, C02023, doi:10.1029/2007JC004259.

Lindsay, R. 2013. Unified Sea Ice Thickness Climate Data Record Collection Spanning 1947-2012. Boulder, Colorado USA: National Snow and Ice Data Center. *http://dx.doi.org/10.7265/N5D50JXV*.

Lindsay, R. and A. Schweiger, (2015), Arctic sea ice thickness loss determined using subsurface, aircraft and satellite observations, *The Cryosphere*, 9, doi:10.5194/tc-9-269-2015.

Liu, J.P., J.A. Curry, H. Wang, M. Song, and R.M. Horton, (2012), Impact of declining Arctic sea ice on winter snowfall. *Proc. Natl. Acad. Sci.*, **109**, 4074-4079, doi:10.1073/pnas.1114910109.

Mahlstein, I. and R. Knutti, (2012), September Arctic sea ice predicted to disappear near 2°C global warming above present, *J. Geophys. Res*., 117, D06104, doi:10.1029/2011JD016709.

Maslanik, J.A., C. Fowler, J. Stroeve, S. Drobot, J. Zwally, D. Yi and W. Emery, (2007). A younger, thinner Arctic ice cover: Increased potential for rapid extensive sea ice loss, *Geophys. Res. Lett.*, 34, L24501, dio:10.1029/2007/GL032043.

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

Meier, W. and D. Notz, (2010), A note on the accuracy and reliability of satellite-derived passive microwave estimates of sea-ice extent, CliC Arctic Sea Ice Working Group Consensus Document, World Climate Research Program.

Meier, W.N., J. Stroeve, and F. Fetterer, (2007), Whither Arctic sea ice? A clear signal of decline regionally, seasonally and extending beyond the satellite record, *Ann. Glaciol.*, 46, 428-434.

Meier, W.N., (2005), Comparison of passive microwave ice concentration algorithm retrievals with AVHRR imagery in the Arctic peripheral seas, *IEEE Trans. Geosci. and Rem. Sens.*, 43(6), 1324-1337.

Miles, W.M., D.V. Divine, T. Furevik, E. Jansen, M. Moros and A.E.J. Ogilvie, 2014. A signal of persistent Atlantic multidecadal variability in Arctic sea ice, *Geophys. Res. Lett*., 41, doi:10.1002/2013GL058084.

Notz, D. (2009): The future of ice sheets and sea ice: Between reversible retreat and unstoppable loss*. Proc. Natl. Acad. Sci.*, 106, 20 590–20 595.

Notz, D. (2014), Sea-ice extent and its trend provide limited metrics of model performance, *The Cryosphere,* 8 (1), 229–243, doi:10.5194/tc-8-229-2014

Notz, D. (2015), How well must climate models agree with observations?, *Phil. Trans. Roy. Soc. A*, in press.

Notz, D., and J. Marotzke (2012), Observations reveal external driver for arctic sea-ice retreat, *Geophys. Res. Lett.*, 39 (8), L051,094, doi:10.1029/2012GL051094

Notz, D., A. Haumann, H. Haak, J. Jungclaus, and J. Marotzke (2013), Arctic sea ice as simulated by MPI-ESM, *J. Adv. Model. Earth Syst.*, 5, 173–194, doi:10.1002/jame.20016.

Overland, J. E., J. A. Francis, E. Hanna, and M. Wang, (2012), The recent shift in early summer Arctic atmospheric circulation. *Geophys. Res. Lett.,* 39, L19804, doi:10.1029/2012GL053268.

Overland, J.E., M. Wang, J.E. Walsh, and J. Stroeve, (2014), Mitigation of Arctic Climate Change, *AGU Earth's Future*, doi:2013EF000162R.

Peterson, K.A., Arribas, A., Hewitt, H.T., Keen, A.B., Lea, D.J. and McLaren, A.J., (2014), Assessing the forecast skill of Arctic sea ice extent in the GloSea4 seasonal prediction system, Climate Dynamics, doi:10.1007/s00382-014-2190-9.

Pithan, F., B. Medeiros, and T. Mauritsen (2013), Mixed-Phase Clouds Cause Climate Model Biases in Arctic Wintertime Temperature Inversions, *Clim. Dyn.* 43 (1-2): 289–303. doi:10.1007/s00382-013-1964-9.

Pithan, F. and T. Mauritsen, (2014), Arctic amplification dominated by temperature feedbacks in contemporary climate models, *Nature Geosc*. Doi:10.1038/ngeo2071.

Porter, D.F., J.J. Cassano, and M.. Serreze (2012), Local and large-scale atmospheric responses to reduced Arctic sea ice and ocean warming in the WRF model, *J. Geophys. Res*., 117, D11115.

Post, E., U.S. Bhatt, C.M. Bitz, J.F. Brodie, T.L. Fulton, M. Hebblewhite, J. Kerby, S.J. Kutz, I. Stirling, and D.A. Walker (2013), Ecological consequences of sea-ice decline.,*Science*. 341(6145), 519-524, doi: 10.1126/science.1235225.

Rampal, P. J. Weiss, C. Dubois, and J.-M. Campin, (2011), IPCC climate models do not capture Arctic sea ice drift acceleration: Consequences in terms of projected sea ice thinning and decline, *J. Geophys. Res.-Oceans*, 116(C8), doi:10.1029/2011JC007110.

Rayner, N.A., D.E. Parkler, E.B. Horton, C.K. Folland, L.V. Alexander, D.P. Rowell, E.C. Kent, and A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperatures since the late nineteenth century, *J. Geophy. Res*., 108(D14), 4407, doi:10.1029/2002JD002670.

Ridley, J. K., J. A. Lowe, and H. T. Hewitt (2012): How reversible is sea ice loss? *The Cryosphere*, 6, 193–198.

Rigor, I.G., and J.M. Wallace, 2004. Variations in the age of Arctic sea-ice and summer sea-ice extent, *Geophys. Res. Lett*., 31, L09401, doi:10.1029/2004GL019492.

Rothrock, D.A., Y. Yu, and G.A. Maykut, (1999), Thinning of the Arctic sea-ice cover, *Geophys. Res. Lett*., 26(23): 3469-3472.

Rosel, A., L. Kaleschke, and G. Birnbaum, (2012), Melt ponds on Arctic sea ice determined from MODIS satellite data using an artificial neural network, *The Cryosphere*, 6, doi:10.5194/tc-6-431-2012.

Santer, B.D. et al., 2008. Consistency of modeled and observed temperature trends in the tropical troposphere, *Int. J. Climatol*., 28, 1703-1722, doi:10.1002/joc.1756.

Schröder, D., D.L. Feltham, D. Flocco, and M. Tsamados (2014), September Arctic sea-ice minimum predicted by spring melt-pond fraction. *Nature Climate Change*.

Schweiger A, R. Lindsay, J. Zhang, M. Steele, H. Stern and R. Kwok, (2011), Uncertainty in modeled Arctic sea ice volume. *J. Geophys. Res.* 116:C00D06, doi:10.1029/2011JC007084.

Screen, J.A., I. Simmonds, C. Deser and R. Tomas, (2013), The atmospheric response to three decades of observed Arctic sea ice loss *J. Clim.,* **26** 1230–48.

Screen, J.A, and I. Simmonds, (2013), Exploring links between Arctic amplification and mid-latitude weather, *Geophys. Res. Lett.*, **40** 959–64.

Serreze, M.C. and J. Stroeve, (2015), Arctic Sea Ice Trends, Variability and Implications for Seasonal Ice Forecasting, *Phil. Trans. A.,* 373, 20140159, doi:10.1098/rsta.2014.0159.

Serreze, M.C., M.M. Holland, and J. Stroeve, (2007), Perspectives on the Arctic's Shrinking Sea Ice Cover*, Science*, 16, 1533-1536.

Serreze, M. C., A. P. Barrett, J. C. Stroeve, D. N. Kindig, and M. M. Holland, (2009). The emergence of surface-based Arctic amplification. *The Cryosphere* 3: 11–19.

Sigmond, M., J. C. Fyfe, G. M. Flato, V. V. Kharin, and W. J. Merryfield (2013), Seasonal forecast skill of Arctic sea ice area in a dynamical forecast system, *Geophys. Res. Lett.*, 40, 529–534, doi:10.1002/grl.50129.

Snape, T.J. and P.M. Forster, (2012), Decline of Arctic sea ice: Evaluation and weighting of CMIP5 projections, *J. Geophys. Res.-Atmos.*, doi:1002/1013/JD020593.

Steffen, K. and A. Schwieger. 1991. NASA Team algorithm for sea ice concentration retrieval from Defense Meteorological Satellite Program Special Sensor Microwave/Imager: Comparison with Landsat satellite imagery. *Journal of Geophysical Research* 96(C12):21,971-21,988.

Stroeve, J.C., E, Blanchard-Wrigglesworth, V. Guemas, S. Howell, F. Massonnet and S. Tietsche, (2015), Developing user-oriented seasonal ice forecasts in a changing Arctic, *EOS Trans. AGU,* doi:10.1029/2015EO031431.

Stroeve, J.C., T. Markus, L. Boisvert, J. Miller and A. Barrett, (2014c), Changes in Arctic melt season and implications for sea ice loss, *Geophys. Res. Lett.*, doi:10.1002/2013GL058951.

Stroeve, J.C., Barrett, A, Serreze, M. & Schweiger, A. (2014b), Using records from submarine, aircraft and satellite to evaluate climate model simulations of Arctic sea ice thickness. *The Cryosphere*, 8, doi:10.5194/tc-8-1-2014.

Stroeve, J., L.C. Hamilton, C.M. Bitz, and E. Blanchard-Wrigglesworth, (2014a), Predicting September sea ice: Ensemble skill of the SEARCH Sea Ice Outlook, 2008-2013, *Geophys. Res. Lett.*, doi:10.1002/2014GL059388.

Stroeve, J.C., M.C. Serreze, J.E. Kay, M.M. Holland, W.N. Meier and A.P. Barrett (2012b), The Arctic's rapidly shrinking sea ice cover: A research synthesis, *Clim. Change*, doi: 10.1007/s10584-011-0101-1.

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

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

Tietsche, S., D. Notz, J. H. Jungclaus, and J. Marotzke, (2011), Recovery mechanisms of Arctic summer sea ice. *Geophys. Res. Lett.*, 38, L02 707, doi:10.1029/2010GL045 698.

Tietsche, S., D. Notz, J.H. Jungclaus, and J. Marotzke, J. (2013). Predictability of large interannual Arctic sea-ice anomalies. *Climate dynamics*, *41*(9-10), 2511-2526.

Tietsche, S., J. Day, V. Guemas, W. Hurlin, S. Keeley, D. Matei, R. Msadek, M. Collins, and E. Hawkins (2014), Seasonal to interannual Arctic sea ice predictability in current global climate models, *Geophys. Res. Lett.,* 41(3), 1035-1043.

Wang, M. and J.E. Overland, (2012), A sea ice free summer Arctic within 30 years – an update from CMIP5 models, *Geophys. Res. Lett.*, doi:10.1029/2012GL052868.

Webster, M.A., I.G. Rigor, S.V.Nghiem, N.T. Kurtz, S.L. Farrell, D.K. Perovich, and M. Sturm (2014), Interdecadal changes in snow depth on Arctic sea ice, *J. Geophys. Res.-Oceans,* doi:10.1002/2014JC009985.

Winton, Michael (2011), Do climate models underestimate the sensitivity of Northern hemisphere sea ice cover?" *J. Climate*, 24(15), 3924–34. doi:10.1175/2011JCLI4146.1.

Zhang, J.L. and D.A. Rothrock, (2003), Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates, *Mon. Weather Rev.*, 131, 845-861.

8 List of Figures

Figure 1. Relationship between the year-to-year change in summer sea-ice area in a specific year and in the previous year. The figure shows for all CMIP5 simulations and for the observational record the year-to-year change in sea-ice area in a given year (y-axis) if the year-to-year change in summer sea-ice area in the previous year was more than 1 million km^2 (x axis). If positive feedbacks were dominant, the year-to-year change should usually be of the same sign as in the previous year. Since this is not the case, negative feedbacks dominate the short-term evolution of the Arctic sea-ice cover.

Figure 2. Sensitivity of Arctic September sea ice extent to changes in global mean (top) and Arctic mean 2m air temperatures. Results are shown for 26 CMIP5 climate models under the RCP8.5 emission scenario. See also *Mahlstein and Knutti* [2012].

Figure 3: Impact of differences in the used sea-ice satellite product on seasonal temperature predictions. The figure shows the difference in 2 m temperature in September that emerges in a seasonal prediction study from the difference between assimilating either Bootstrap or NASA-Team sea-ice concentration into a seasonal prediction system. The simulations are started from assimilation runs in May and develop freely throughout summer. See *Bunzel et al.*, [submitted] for details.

Figure 4. Observed (black line) and CMIP5 modeled (gray and colored lines/shading) of September sea ice extent from 1900 to 2100. Historical simulations were started in 1850 and run through 2005. To extend the time-series past 2005 we use future projections for different forcing scenarios named Representative Concentration Pathways (RCPs). In each case the RCP number indicates the amount of additional radiative forcing (in units W/m²) from long-lived greenhouse gases that are caused by human activities. The colored lines represent the multi-model mean extent under different RCPs, whereas the shading represents one standard deviation from the mean (± 1σ). The numbers in parenthesis represent the number of climate models used in the figure.

Figure 5. Comparison of observed September sea ice extent trends for 1953 to1990 (x-axis) to those from 1991 to 2013 (y-axis), using the combined and homogenized Had1SST and NSIDC sea ice record to estimate the observed trends (gray shading represents the observed 2 standard deviations (± 2σ)). Symbols represent the trends calculated for individual historical CMIP5 model ensemble member.

Figure 6. Comparison of detrended March thickness anomaly to detrended September ice extent from 1979 to 2013 using the Pan-Arctic Ice-Ocean Modelling and Assimilation System [PIOMAS, *Zhang and Rothrock*, 2003].

Figure 7. Distributions of years of the first occurrence of an ice-free September under the CMIP5 RCP8.5 emission scenario. Grey shading shows the distributions for models selected by matching the observed trend, with the outer light gray shading representing the 10th to 90th percentile and the inner (darker gray) shading the 25th to 75th percentile. The black line represents the 50th percentile (median). The black vertical lines represent the 10th to 90th percentile and median (asterisk) for 100 samples of 10 random models, whereas the distribution from all models is shown in red. To get each

sample, 10 models are selected at random from the available model ensemble means, giving 10 ice-free dates for that sample. This is then repeated 100 times and compared to the ice-free dates from using all models and those selected based on matching the observed trend from 1979 to 2014.

Figure 1. Relationship between the year-to-year change in summer sea-ice area in a specific year and in the previous year. The figure shows for all CMIP5 simulations and for the observational record the year-to-year change in sea-ice area in a given year (y-axis) if the year-to-year change in summer sea-ice area in the previous year was more than 1 million km^2 (x axis). If positive feedbacks were dominant, the year-to-year change should usually be of the same sign as in the previous year. Since this is not the case, negative feedbacks dominate the short-term evolution of the Arctic sea-ice cover.

Figure 2. Sensitivity of Arctic September sea ice extent to changes in global mean (top) and Arctic mean 2meter air temperatures. Results are shown for 26 CMIP5 climate models under the RCP8.5 emission scenario.

2m temperature diff. [K] BT-NT, SEP-mean, MAY-HCs

Figure 3: Impact of differences in the used sea-ice satellite product on seasonal temperature predictions. The figure shows the difference in 2 m temperature in September that emerges in a seasonal prediction study from the difference between assimilating either Bootstrap or NASA-Team sea-ice concentration into a seasonal prediction system. The simulations are started from assimilation runs in May and develop freely throughout summer. See *Bunzel et al.*, [submitted] for details.

Figure 4. Observed (black line) and CMIP5 modeled (gray and colored lines/shading) of September sea ice extent from 1900 to 2100. Historical simulations were started in 1850 and run through 2005. To extend the time-series past 2005 we use future projections for different forcing scenarios named Representative Concentration Pathways (RCPs). In each case the RCP number indicates the amount of additional radiative forcing (in units W/m²) from long-lived greenhouse gases that are caused by human activities. The colored lines represent the multi-model mean extent under different RCPs, whereas the shading represents one standard deviation from the mean (± 1σ). The numbers in parenthesis represent the number of climate models used in the figure.

Figure 5. Comparison of observed September sea ice extent trends for 1953 to1990 (x-axis) to those from 1991 to 2013 (y-axis), using the combined and homogenized Had1SST and NSIDC sea ice record to estimate the observed trends (gray shading represents the observed 2 standard deviations (± 2σ)). Symbols represent the trends calculated for individual historical CMIP5 model ensemble member.

Figure 6. Comparison of detrended March thickness anomaly to detrended September ice extent from 1979 to 2013 using the Pan-Arctic Ice-Ocean Modelling and Assimilation System [PIOMAS, *Zhang and Rothrock*, 2003]

Figure 7. Distributions of years of the first occurrence of an ice-free September under the CMIP5 RCP8.5 emission scenario. Grey shading shows the distributions for models selected by matching the observed trend, with the outer light gray shading representing the 10th to 90th percentile and the inner (darker gray) shading the 25th to 75th percentile. The black line represents the 50th percentile (median). The black vertical lines represent the 10th to 90th percentile and median (asterisk) for 100 samples of 10 random models, whereas the distribution from all models is shown in red. To get each sample, 10 models are selected at random from the available model ensemble means, giving 10 ice-free dates for that sample. This is then repeated 100 times and compared to the ice-free dates from using all models and those selected based on matching the observed trend from 1979 to 2014.