Arctic sea-ice evolution as modeled by Max Planck Institute for Meteorology's Earth system model[1]

Dirk Notz,[1] F. Alexander Haumann,[1,2] Helmuth Haak,[1] Johann H. Jungclaus,[1] and Jochem Marotzke[1]

Received 27 July 2012; revised 10 December 2012; accepted 8 January 2013; published 18 April 2013.

[1] We describe the evolution of Arctic sea ice as modeled by the Max Planck Institute for Meteorology's Earth System Model (MPI-ESM). The modeled spatial distribution and interannual variability of the sea-ice cover agree well with satellite observations and are improved relative to the model's predecessor ECHAM5/MPIOM. An evaluation of modeled sea-ice coverage based on sea-ice area gives, however, conflicting results compared to an evaluation based on sea-ice extent and is additionally hindered by uncertainties in the observational record. Simulated trends in sea-ice coverage for the satellite period range from more strongly negative than observed to positive. The observed evolution of Arctic sea ice is incompatible with modeled internal variability and probably caused by external forcing. Simulated drift patterns agree well with observations, but simulated drift speed is generally too high. Simulated sea-ice volume agrees well with volume estimates of the PIOMAS reanalysis for the past few years. However, a preceding Arctic wide decrease in sea-ice volume starts much earlier in MPI-ESM than in PIOMAS. Analyzing this behavior in MPI-ESM's ocean model MPIOM, we find that the modeled volume trend depends crucially on the specific choice of atmospheric reanalysis forcing, which casts some doubt on the reliability of estimates of volume trends. In our CMIP5 scenario simulations, we find a substantial delay in sea-ice response to increasing CO_2 concentration; a seasonally ice-free Arctic can result for a CO_2 concentration of around 500 ppm. Simulated winter sea-ice coverage drops rapidly to near ice-free conditions once the mean Arctic winter temperature exceeds $-5°C$.

Citation: Notz, D., F. A. Haumann, H. Haak, J. H. Jungclaus, and J. Marotzke (2013), Arctic sea-ice evolution as modeled by Max Planck Institute for meteorology's Earth system model, *J. Adv. Model. Earth Syst.*, *5*, 173–194, doi:10.1002/jame.20016.

1. Introduction

[2] Being only a thin veneer between the ocean and the atmosphere, the Earth's sea-ice cover reacts rapidly to changes in prevailing climate conditions. The observed retreat of the Arctic sea-ice cover [Fetterer et al., 2002, updated 2012] is therefore a good test bed to examine the capability of modern Earth system models to simulate the sensitivity of certain components of the climate system to changes in the external forcing. To this purpose, we here document the Arctic sea-ice simulations of the Max Planck Institute for Meteorology's Earth system model (MPI-ESM) that have been carried out for the Climate Model Intercomparison Project (CMIP5) [see *Taylor et al.*, 2012]. In doing so, we put special emphasis on a number of issues related to the quality of observations that are of general relevance for any intercomparison between model and observations related to Arctic sea ice.

[3] This contribution is part of a special issue of *Journal of Advances in Modeling Earth Systems* on the climate simulated by MPI-ESM. Of particular relevance for researchers interested in sea ice are, in addition to this rather focused contribution, the contributions that describe the atmospheric model component ECHAM6 [*Stevens et al.*, 2013] the modeled climate variability [*Giorgetta et al.*, 2013], the tuning of MPI-ESM [*Mauritsen et al.*, 2012], the impact of sea-ice melt ponds [*Roeckner et al.*, 2012], and the ocean model MPIOM [*Jungclaus et al.*, 2013].

[4] Despite the existence of these more detailed descriptions of other components of MPI-ESM, we aim for keeping the present paper sufficiently self-contained. Therefore, in section 2 we briefly describe the model components of MPI-ESM, with an obvious focus on the sea-ice model component. We also discuss the impact of resolution on our simulations and describe the process of tuning the MPI-ESM sea-ice model.

[1]Max Planck Institute for Meteorology, Hamburg, Germany.
[2]Now at Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, Switzerland.

1942-2466/13/10.1002/jame.20016

[5] We then move on to the major purpose of this contribution, namely an assessment of the quality of our sea-ice simulations relative to the observed evolution of the Earth's sea-ice cover. Such assessment crucially depends on reliable measurements of the time evolution of the sea-ice cover. In section 3, we therefore briefly discuss the quality of the data sets and reanalysis data we used for the comparison of MPI-ESM model results with observations.

[6] In section 4, we compare MPI-ESM simulations to these data. To do so, we use the CMIP5 historical simulations [*Taylor et al*., 2012] that were carried out with MPI-ESM. Three such historical simulations spanning the period 1850–2005 were carried out with MPI-ESM. All are based on identical model configurations but slightly different initial conditions in 1850. This allows us to examine the impact of internal variability on the modeled time evolution of sea ice. In this section, we focus on differences in sea-ice simulations with MPI-ESM compared to this model's predecessor ECHAM5/MPIOM. The sea-ice simulations carried out with ECHAM5/MPIOM for CMIP3 were analyzed in detail by *Koldunov et al*. [2010], who found in particular too large a simulated sea-ice cover in summer, and a lack of variability on interannual time scales. As we will see in section 4, both these aspects are much improved in MPI-ESM, and the model does reasonably well in simulating the observed evolution of the Arctic sea-ice cover.

[7] Because of this, we can assume that also the modeled internal variability of the sea-ice cover is close to the real internal variability. Based on this assumption, we focus in section 5 on our preindustrial control simulations to derive a range of Arctic sea-ice trends and extreme values that can be caused by internal variability. For these simulations, the external forcing was held constant for 1000 years, and all variability in the modeled sea-ice cover is caused by internal variability. We use the derived estimates of internal variability to examine the chances that the observed sea-ice evolution could simply have happened because of internal variability.

[8] In section 6, we examine the modeled sea-ice cover in our CMIP5-scenario simulations. For the period 2005–2100, three simulations were carried out for each of the representative concentration pathways (RCP) RCP2.6, RCP4.5, and RCP8.5 [*Meinshausen et al*., 2011]. These were extended with a single simulation each until the year 2300. The numbers specifying the various RCP's indicate the additional forcing (in W/m^2) that is caused by increasing greenhouse gas emissions in that particular scenario until the end of this century. We focus in section 6 on the response of the summer and the winter sea-ice cover to increasing CO_2 concentrations and to changes in the Arctic surface temperature in these scenarios. The paper closes with a short concluding section.

2. Sea-Ice Component of MPI-ESM

2.1. Model Description

[9] The climate system is simulated in MPI-ESM by the atmospheric model component ECHAM6 [*Stevens et al*., 2013] and the oceanic model component MPIOM [*Marsland et al*., 2003; *Jungclaus et al*., 2013]. MPI-ESM also includes model components simulating the ocean biogeochemistry (HAMOCC5) [*Maier-Reimer*, 1993] and land-surface processes (JSBACH) [*Reick et al*., 2013].

[10] Sea ice is represented as part of this model system both within ECHAM6 and within MPIOM. As described in more detail below, the purpose of the sea-ice submodel within ECHAM6 is solely to provide at each atmospheric time-step surface temperature and albedo consistent with the evolution of the atmospheric state. The update of ice thickness and sea-ice dynamics are carried out within the "full" sea-ice submodel that forms part of our ocean model MPIOM.

[11] This "full" sea-ice submodel within MPIOM consists of a thermodynamic-dynamic sea-ice model that is barely changed with respect to MPI-ESM's predecessor ECHAM5/MPIOM that was used for CMIP3 [*Meehl et al*., 2007]. The simulation of sea-ice dynamics within MPIOM is based on a viscous-plastic rheology [*Hibler*, 1979]. The thermodynamics of sea ice in MPIOM are represented by a simple zero-layer model [*Semtner*, 1976]. As such, the sea ice has no heat capacity, which might influence the modeled seasonal cycle of sea-ice growth and decay [*Semtner*, 1984]. At the ice surface, the sea-ice temperature is calculated from a flux balance between conductive heat through the ice, outgoing longwave radiation, and incoming atmospheric fluxes as supplied to the ocean model through the model coupler OASIS3 [*Valcke et al*., 2012]. If the surface temperature calculated in such way would exceed 0°C, the excess energy is used to melt sea ice. The melt water is directly transferred into the uppermost ocean grid cell. The impact of snow accumulation is represented according to the changes in overall heat transfer through the ice and snow, and snow ice is formed whenever the ice's freeboard becomes negative. Changes in ice thickness at the bottom of the ice are calculated from a balance of conductive heat flux through the ice and the oceanic heat flux. The latter is based on a simple "ice-bath" assumption [*Schmidt et al*., 2004] in that the uppermost oceanic grid cell is kept at the freezing temperature whenever there is any sea ice present within that oceanic grid cell. Any excess heat content is used to melt the ice at its bottom. Compared to more realistic parameterizations, this very simplified representation of ice-ocean heat exchange overestimates the amount of oceanic heat that can be used to melt sea ice from below [*Notz et al*., 2003; *Schmidt et al*., 2004].

[12] MPIOM does not allow for a subgrid scale distribution of ice thickness. Each grid cell is covered by a certain fraction A of open water and a fraction $1 - A$ of sea ice of thickness h. The standard output of ice thickness from MPI-ESM is V/area, where V is the total volume of sea ice within a certain grid cell and area is that grid cell's area. Hence, the standard output of ice thickness is the mean thickness that sea ice would have if it maintained its volume but covered the entire grid cell. The salinity of any sea ice that forms in MPI-ESM is set to a constant value of 5 g/kg, with all excess salt being retained by the underlying ocean. In terms of its thermodynamics, our

sea-ice model is certainly not state of the art as described by *Hunke et al*. [2010, 2011]. As such, our simulations allow us to estimate how well a sea-ice cover can be simulated with a very simplified representation of the underlying physics. We are currently improving our sea-ice model and hope to have a more up-to-date version available for the next version of MPI-ESM.

[13] The sea-ice state in the atmospheric model component ECHAM6 is primarily used to provide a lower boundary condition for surface fluxes to the atmosphere and to accumulate atmospheric fluxes to be passed on to the full sea-ice model within MPIOM at each coupling time step (every 24 h). As such, the sea-ice state in ECHAM6 primarily aims at a proper representation of the surface characteristics of the ice, which are updated at each atmospheric time step (600 s for the low resolution (LR) model version MPI-ESM-LR, 450 s for the mixed resolution (MR) model version MPI-ESM-MR). This sea-ice model within ECHAM6 consists of a simple zero-layer model of sea ice, whose thickness is not changed until a new thickness is being provided from the oceanic sea-ice model at each coupling time step. The sea-ice model in ECHAM6 calculates the sea-ice surface temperature based on the incoming atmospheric fluxes, surface albedo, the conductive heat flux through snow and ice, and outgoing longwave radiation. During melting, as a major change in the sea-ice model compared to ECHAM5, the surface albedo is calculated as the weighted average of bare sea ice, snow, and melt ponds within a certain grid cell according to the melt-pond scheme described by *Pedersen et al*. [2009]. This scheme also provides for different conductive heat fluxes of bare ice and melt-pond covered ice, and for runoff of melt-pond water, but these two processes were accidentally not considered in our CMIP5 simulations. Including them, however, only leads to rather small changes in the modeled sea-ice state [*Roeckner et al*., 2012].

[14] Compared to the previous model version ECHAM5/MPIOM, the change in the representation of sea-ice albedo introduced by the melt-pond scheme was the major change in our sea-ice model. MPIOM remained almost unchanged, with only some slight adjustments, such as the zenith dependence of the albedo of open water. Hence, major differences in model code between MPI-ESM and ECHAM5/MPIOM only exist in the atmospheric model component, with a shift from ECHAM5 to ECHAM6 [*Stevens et al*., 2013]. Changes in ECHAM6 relative to ECHAM5 concern primarily the representation of shortwave radiation, a revision of cloud optical properties, changes in the representation of convective clouds, and the fact that the standard model version now includes a representation of the middle atmosphere and of land-surface processes.

2.2. Resolution

[15] For CMIP5, MPI-ESM was set up to run at two different resolutions. The low-resolution version (MPI-ESM-LR) uses the atmospheric model ECHAM6 at a T63 (1.875°) horizontal resolution with 47 vertical levels reaching up to 0.1 hPa. For comparison, our CMIP3 standard setup with ECHAM5 was also based on T63, but with only 31 vertical levels reaching up to 10 hPa. For the ocean, MPI-ESM-LR is based on the same model configuration as our standard CMIP3 model version, with a bipolar grid at a horizontal resolution of about 1.5° and 40 vertical levels. The two poles of this setup are placed in the Antarctic and in Greenland, allowing for a resolution of about 20 km in the Fram-Strait region. For the mixed-resolution model version MPI-ESM-MR, the vertical resolution of the atmosphere is increased to 95 levels, leaving the horizontal resolution at T63. For the ocean, the vertical resolution remains at 40 levels, but the horizontal grid is changed to a tripolar setup with a resolution of 0.4°. Hence, in the ocean this model version is eddy permitting [*Jungclaus et al*., 2013]. The three poles of the tripolar oceanic grid are placed in Canada, in Siberia, and in the Antarctic. Compared to the low-resolution model version, the mixed-resolution setup has somewhat lower resolution in the deep-water formation regions around Greenland and around Antarctica, but higher resolution almost everywhere else.

[16] The change in resolution has only rather minor effects on the simulated sea-ice cover. In our 1000-year long preindustrial control simulation, the MR setup results in slightly more wintertime sea ice around Svalbard than the LR setup. In contrast, there is somewhat less ice in the MR setup in the Sea of Okhotsk, south of the Bering Strait and in the Labrador Sea. These differences are also reflected by the surface air temperature. In general, the MR simulations are in winter about 0.5°C warmer over the Arctic Ocean than the LR simulations. However, in the Sea of Okhotsk and in the Labrador Sea, the MR-control simulation shows in March about 3°C higher air temperatures, whereas it is slightly colder around Svalbard. Regarding the simulated ice thickness, the differences between the two resolutions are quite randomly distributed with a total magnitude of usually less than 10 cm. The most significant differences occur in all seasons within the Canadian Archipelago, where MPI-ESM-MR simulates about 50 cm thinner ice than MPI-ESM-LR.

[17] As we will see later, the regional differences in ice-thickness distribution between the LR and the MR simulation do not cause any large differences in the simulated time evolution of total sea-ice coverage and total sea-ice volume. For example, the March sea-ice area in our control simulation is 14.0 million km^2 MPI-ESM-MR compared to 14.4 million km^2 in MPI-ESM-LR. These differences arise primarily because of the lower ice concentration in the Sea of Okhotsk in MPI-ESM-MR. The differences in other seasons are similarly small, as are the differences in modeled sea-ice volume (compare also Table 1). Since in this paper we focus mostly on the large-scale features of our simulations, we will in the following primarily describe our LR simulations, for which we have more ensemble members. This allows for improved statistics of our analysis. To reflect the similarities between MPI-ESM-LR and MPI-ESM-MR also in the text, we will simply refer to MPI-ESM whenever a certain statement refers to both model versions.

Table 1. Overview of March and September Values for Sea-Ice Area, Extent, and Volume[a]

		March						September					
		Area		Extent		Volume		Area		Extent		Volume	
Data Set	Time period	Mean	Trend	Mean	Trend	Mean	Trend	Mean	Trend	Mean	Trend	Mean	Trend
Bootstrap	1979–2007	14.59	−0.40	15.79	−0.42			6.14	−0.70	7.12	−0.71		
Sea-ice index	1979–2007	13.67	−0.34	15.58	−0.47			5.05	−0.65	6.74	−0.72		
PIOMAS	1979–2007					28.14	−2.26					12.91	−2.52
MPI-ESM-LR-1	1979–2007	13.21	−0.21	14.64	−0.26	23.52	−1.28	4.94	−0.40	6.66	−0.45	7.69	−1.21
MPI-ESM-LR-2	1979–2007	13.14	−0.26	14.52	−0.30	23.64	−0.34	5.07	0.03	6.80	0.09	8.00	−0.23
MPI-ESM-LR-3	1979–2007	13.13	−0.37	14.50	−0.48	22.87	−1.28	4.79	−0.37	6.55	−0.37	6.88	−1.25
MPI-ESM-LR-ens	1979–2007	13.16	−0.28	14.55	−0.35	23.34	−0.97	4.93	−0.25	6.67	−0.24	7.52	−0.90
MPI-ESM-MR-ens	1979–2005	13.09	−0.41	14.48	−0.48	23.71	−1.64	4.88	−0.46	6.57	−0.52	7.67	−1.44
ECHAM5/MPIOM-1	1979–2007	13.47	−0.10	14.87	−0.10	37.61	−1.78	6.69	−0.26	8.62	−0.23	22.96	−2.21
ECHAM5/MPIOM-2	1979–2007	13.60	−0.25	15.04	−0.28	37.80	−1.17	6.74	−0.22	8.66	−0.26	23.13	−1.14
ECHAM5/MPIOM-3	1979–2007	13.26	−0.32	14.60	−0.34	35.59	−3.23	6.55	−0.38	8.56	−0.35	21.02	−3.13
ECHAM5/MPIOM-ens	1979–2007	13.44	−0.22	14.84	−0.24	37.00	−2.06	6.66	−0.28	8.61	−0.28	22.37	−2.16
NCEP-forced MPIOM	1979–2007	13.53	−0.34	14.90	−0.35	34.02	−2.78	6.25	−0.44	7.93	−0.40	16.50	−2.52
ERA-forced MPIOM	1979–2007	13.90	−0.44	15.19	−0.47	25.21	−2.06	4.98	−0.49	6.65	−0.57	7.18	−1.59
MPI-ESM-LR	1000 years pi	14.39	−0.00	15.99	−0.00	29.21	−0.01	6.07	−0.00	7.87	−0.00	12.65	−0.01
MPI-ESM-MR	1000 years pi	13.96	−0.00	15.50	−0.00	28.06	0.05	5.81	0.00	7.57	0.00	13.19	−0.05
ECHAM5/MPIOM	500 years pi	13.50	−0.00	14.90	−0.00	34.28	−0.00	6.35	0.00	8.36	0.00	19.17	−0.01

"Bootstrap" refers to the sea-ice concentration derived using the Bootstrap algorithm [*Comiso*, 1986]. The NSIDC "sea-ice index" [*Fetterer et al*., 2002, updated 2012] is based on the NASA Team algorithm [*Cavalieri et al*., 1984]. The third row is based on sea-ice volume reanalysis by PIOMAS [*Zhang and Rothrock*, 2003]. All other rows are model-output based. For MPI-ESM-LR and ECHAM5/MPIOM, the individual values of all three ensemble members (denoted −1, −2, −3) are shown, as is the ensemble mean (denoted -ens). The last three lines show data from model simulations with constant, preindustrial forcing. Units are million km^2 for area and extent mean, million km^2/decade for area and extent trend, 1000 km^3 for volume and 1000 km^3/decade for volume trend.

2.3. Tuning

[18] Many of the characteristic length scales of sea ice are far too small to be resolved by current ESMs, and many of the parameters that are needed to represent the subgrid-scale properties of sea ice are too poorly known and too variable to allow for their direct implementation into a numerical model. For example, the distribution of melt ponds on top of the sea-ice cover during summertime is random, with individual ponds having a size of about 1 m. The interplay between these ponds and the underlying sea ice crucially governs the decay of the summer sea-ice cover, primarily through its impact on the ice albedo. MPI-ESM uses a rather sophisticated scheme to represent these interactions [*Roeckner et al*., 2012]. In this scheme, the albedo of melt ponds, snow, and bare ice are set to observed values for the different surface types. The albedo of sea ice in a certain grid box is then given as the weighted average of the albedo of the individual surface categories (for details, see *Pedersen et al*. [2009] and *Roeckner et al*. [2012]).

[19] Because of the improved subgrid scale representation of surface albedo, we no longer use this parameter to tune our model. Instead, all tuning of our sea-ice model occurs via a parameter that aims at emulating a subgrid scale distribution of ice thickness. In reality, on the scale typical for ESM grid cells, ice floes of very different thicknesses are usually present within the ice cover. This fact is represented in modern sea-ice models by use of a statistical representation of the ice-thickness distribution, with typically 5–10 different ice-thickness classes present within an individual grid cell. In contrast, MPI-ESM currently only allows for a single ice thickness within a certain grid cell. Given that this is a very simplified representation of reality, parameters to describe the evolution of the distribution of that single ice thickness class cannot readily be derived from observations. Hence, tuning is required to achieve results from the sea-ice model that agree reasonably well with observations. A short description of this tuning of our sea-ice model is provided by *Mauritsen et al*. [2012], who discuss in detail the procedure of tuning MPI-ESM. Here, we will give some additional details of the specific tuning of our sea-ice model.

[20] The MPI-ESM sea-ice model contains two main parameters that we use to tune the overall mean state of the sea-ice cover. One of these parameters describes the change in ice-thickness distribution during freezing, and the other parameter describes the change in ice-thickness distribution during melting. The impact of these two parameters can readily be understood as follows: At every time step, the sea-ice cover within a grid cell of MPI-ESM is fully described by the fraction A of that grid cell that is covered with ice, and the thickness of the ice within that ice-covered fraction, h_{oldice}. The first tuning parameter of the sea-ice thickness distribution describes how quickly the fraction $(1 - A)$ of open water shrinks during the formation of new ice [e.g., *Hibler*, 1979]. If new ice of thickness h_{newice} is formed within a certain grid cell, this new ice does not immediately cover the entire open-water fraction of that grid cell. Instead,

it is redistributed to have a thickness $h_0 > h_{\mathrm{newice}}$, which allows one to calculate the change in open-water area as

$$\Delta(1-A) = (1-A) \cdot \frac{h_{\mathrm{newice}}}{h_0}. \quad (1)$$

[21] Given that in MPI-ESM only one ice thickness can be present within an individual grid cell, for the next time step the mean thickness of the old ice becomes the weighted average of the existing old ice and the redistributed new ice. From equation (1) it follows that for a constant ice-formation rate in open water $h_{\mathrm{newice}}/\Delta t$, where Δt is the model time step, the open-water area decreases exponentially in time with a decay rate of $-\Delta t \cdot h_{\mathrm{newice}}/h_0$. Hence, for large h_0, the open-water fraction decreases only relatively slowly, which results in a larger heat exchange between ocean and atmosphere and correspondingly larger ice-formation rates. In MPI-ESM, h_0 is formulated as

$$h_0 = 0.5\,\mathrm{m} + c_{\mathrm{freeze}} \cdot (h_{\mathrm{oldice}} - 0.5\,\mathrm{m}).$$

[22] Here, c_{freeze} is a tuning parameter with values between zero and one. For $c_{\mathrm{freeze}} = 0$, the value for h_0 becomes 0.5 m, as suggested by *Hibler* [1979]. For $c_{\mathrm{freeze}} = 1$, the new ice is redistributed to the same thickness as the existing ice in that grid cell, as suggested by *Parkinson and Washington* [1979]. If $h_{\mathrm{oldice}} < 0.5$ m, always the value $h_0 = 0.5$ m is used within MPI-ESM. For CMIP5 simulations, we used $c_{\mathrm{freeze}} = 2/3$. Decreasing c_{freeze} to the original *Hibler* [1979] value of zero results in overall less sea ice, with the annual mean volume being reduced from around 22,000 km^3 to around 16,000 km^3 (compare with *Mauritsen et al.*, 2012, Figure 5a]). An increase of c_{freeze} to 0.99 and hence a corresponding increase of h_0 to almost h_{oldice} leads to an increase in sea-ice volume to an annual mean of around 25,000 km^3. Despite such rather large change in sea-ice volume, the corresponding sea-ice area changes only little: $c_{\mathrm{freeze}} = 0$ results in an annual mean Arctic sea-ice area of 10.1 million km^2, very close to the annual mean of just below 10.4 million km^2 in our standard control simulation. Increasing c_{freeze} to 0.99 slightly increases the simulated Arctic mean area to just above 10.4 million km^2. Hence, while we can tune the overall sea-ice volume in our sea-ice model using c_{freeze}, such tuning has little effect on the simulated preindustrial control values of sea-ice area and sea-ice extent.

[23] The second parameter that is used to tune the opening or closing of open water within MPI-ESM is only active during melting conditions. This parameter, denoted as c_{melt}, describes how fast the ice-covered area A decreases if an ice volume $\Delta V_{\mathrm{oldice}}$ melts within a certain grid cell. It is applied according to

$$\Delta A = c_{\mathrm{melt}} \times \frac{\Delta V_{\mathrm{oldice}}}{V_{\mathrm{oldice}}}.$$

[24] The new ice thickness is then given as

$$h_{\mathrm{oldice}} = \frac{V_{\mathrm{oldice}} - \Delta V_{\mathrm{oldice}}}{A - \Delta A}.$$

[25] For a larger value of c_{melt}, the ice-covered area decreases relatively quickly while the ice thickness decreases only slowly and vice versa. The value of c_{melt} has very little impact on the modeled sea-ice volume. It does, however, influence the modeled sea-ice area, in particular, during summer. Here, a value of $c_{\mathrm{melt}} = 0.5$, which is twice our standard value of $c_{\mathrm{melt}} = 0.25$, increases the modeled September Arctic sea-ice area during the first 50 years of our control run from 5.9 to 6.3 million km^2. Decreasing c_{melt} to 0.125 results in a mean September sea-ice area of around 5.1 million km^2. Absolute changes are smaller in all other months, being almost zero in winter.

[26] As outlined above, the two single parameters c_{melt} and c_{freeze} that we use to tune our sea-ice simulations act exclusively to differentiate melting and freezing processes between their lateral and their vertical impact. Because MPI-ESM only represents a single ice thickness within any particular grid cell, these parameters can only be found ad hoc as part of our tuning process. In models that employ a subgrid scale ice-thickness distribution, a similar tuning process needs to be carried out to differentiate between lateral and vertical growth and melt. The impact of such tuning in these models will, however, be less pronounced than in our simplified scheme. This can be readily understood as follows: In reality, sea-ice concentration changes in part by lateral melting or freezing, more efficiently, however, by the complete melting or the formation of very thin ice in a particular region. The latter process is explicitly represented in models that employ an ice-thickness distribution, whereas in MPI-ESM both processes must be parameterized in an ad hoc way by means of c_{melt} and c_{freeze}. We therefore expect a smaller bandwidth of possible sea-ice states that can be achieved by tuning these parameters in a model employing a subgrid scale ice-thickness distribution.

[27] Given the expense of running the fully coupled ESM, we only tuned the mean state of the sea-ice component. Our aim was to achieve a preindustrial Arctic sea-ice volume of about 20×10^3 km^3. We are unable to tune for the observed seasonal cycle or trends of the sea-ice cover, of drift patterns, the length of the melt period, and other such variables. Hence, we can use these variables to evaluate MPI-ESM. Before carrying out such model evaluation in section 4, we need to briefly introduce the various sources of observational and reanalysis data that we use to represent reality for such comparison. This we will do in the following section.

3. Observational and Reanalysis Data

[28] One of the major purposes of any Earth system model is to realistically represent the real climate of our Earth. Only if the model is capable of such realistic representation, its results can be related to the real evolution of the Earth's climate. Unfortunately, it is impossible to directly compare the output of any climate model to reality, but we can carry out a comparison to certain measurements that we trust to be close to reality. This fine difference is sometimes neglected, which can lead to misleading results concerning the reliability of models.

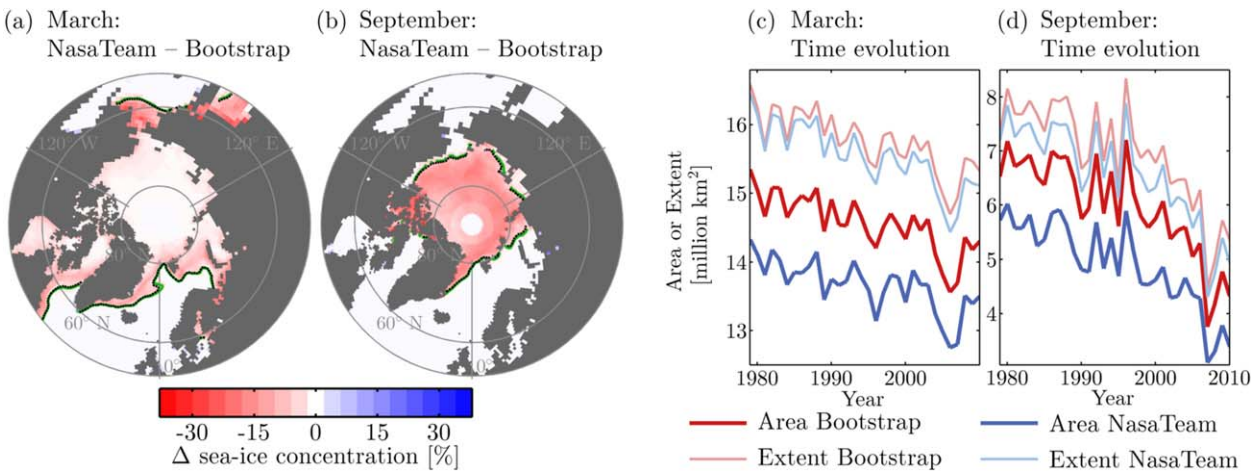

Figure 1. Difference between NASA Team and Bootstrap algorithms to determine sea-ice concentration. (a) Mean difference in March sea-ice concentration for the period 1979–2007; (b) mean difference in September sea-ice concentration for the period 1979–2007. The black and green contour lines denote the sea-ice edges (15% ice concentration) of the NASA Team and Bootstrap algorithms, respectively. Time evolution of (c) March and (d) September area and extent from 1979 until 2010.

[29] We can easily clarify this point by first focusing on measurements of sea-ice concentration from satellites. The standard satellites used for such purpose do not measure sea-ice concentration directly, but the passive microwaves that are emitted from the Earth's surface. A number of algorithms have been developed to determine the sea-ice concentration based on the observed brightness temperature, with the NASA Team algorithm [*Cavalieri et al.*, 1984] and the Bootstrap algorithm [*Comiso*, 1986] being the most widely used. These provide daily estimates of sea-ice concentration with a resolution of approximately 25×25 km^2. The Bootstrap algorithm generally results in a higher concentration estimate around the ice edge in winter and all over the Arctic in summer (Figure 1, cf. *Comiso et al.* [1997]). During the period 1979–2007, the area-weighted mean absolute difference in concentration between the two algorithms averaged over the entire ice cover is around 6% in March and around 12% in September. Locally, the Bootstrap algorithm can result in more than 30 concentration-percent higher sea-ice concentration than the NASA Team algorithm all year round. These differences result in the estimate of sea-ice area to be roughly 1 million km^2 larger based on the Bootstrap algorithm compared to that of the NASA Team algorithm (Figure 1c). The differences in extent are much smaller, since for the extent, all measurement grid cells with an ice concentration larger than 15% are counted as being ice covered. Hence, differences between the two algorithms at sea-ice concentration higher than this threshold vanish for the calculation of sea-ice extent.

[30] Given that both algorithms were developed to best represent the "real" state of the Arctic sea-ice cover, the rather high differences between the resulting sea-ice concentration can be considered as the upper bound on any reliable comparison between models and observations. Great care must therefore be taken when comparing model results to the "observed" sea-ice concentration based on just a single algorithm. For the purpose of our paper, we will take the following pragmatic approach. Whenever we want to compare the *spatial distribution* of model results to observations, we use the National Oceanic and Atmospheric Administration (NOAA)/National Snow and Ice Data Center (NSIDC) Climate Data Record of Passive Microwave Sea Ice Concentration (CDR) [*Meier et al.*, 2011). This record was compiled from a merging of the Bootstrap and the NASA Team algorithms to provide a consistent time series of sea-ice concentration from 1987 to 2007 based on microwave measurements by the Special sensor microwave/imager (SSMI) satellite. Given that both original algorithms tend to underestimate sea-ice concentration, the merging was carried out by always including the larger value of the two algorithms into the final data set. This causes the CDR to be in all its characteristics very close to those of the Bootstrap algorithm. Prior to 1987, the CDR is extended back to 1979 by also including data obtained from the SMMR satellite. Overall, the CDR hence spans the period 1979–2007, which we will therefore use as our reference period for any comparison between models and data throughout this paper.

[31] For analyzing the *time evolution* of sea-ice area and extent, often the NSIDC sea-ice index [*Fetterer et al.*, 2002, updated 2012) is used, which provides a time series of these two quantities from 1979 until today. This index is primarily based on the NASA Team algorithm, and hence can probably be considered a very conservative lower bound on the real sea-ice extent and area (Figure 1c). Given the widespread use of this index for time-series analysis of sea-ice evolution, we also compare time series of MPI-ESM output to this record, sometimes extended back to 1953 following the procedure described by *Meier et al.* [2012]. We will, however, always also discuss a comparison between

model output and sea-ice area and extent as obtained from the Bootstrap algorithm, which gives almost identical results as calculating area and extent based on the CDR. Because of this similarity to the CDR record, we take observational estimates based on the Bootstrap algorithm to be more reliable than those based on the NASA Team algorithm. We made sure that no artifacts arise from the different grids of MPI-ESM simulations and the observational record by interpolating the observational grid to MPI-ESM grid. Using nearest-neighbor interpolation, we find almost no change in the resulting sea-ice area and sea-ice extent. Note that a bilinear or bicubic interpolation of the observation grid to the model grid results in about 10% smaller estimated sea-ice area and extent, while a distance-weighted average remapping results in a roughly 10% overestimate of both parameters.

[32] For a comparison of our simulated sea-ice drift against observations, we use polar pathfinder daily 25 km EASE-grid sea ice motion vectors from 1979 to 2005 [*Fowler*, 2003, updated 2011], interpolated by a nearest-neighbor technique to the model grid. These observations are primarily based on tracking the movement of sea-ice patterns from one satellite-based estimate of sea-ice concentration to the next. In addition, buoy data have been used to complement the satellite-derived estimates. Biases in this data set are particularly large (and negative) in areas of low ice concentration and in areas of high drift velocity [*Kwok et al.*, 1998; *Schwegmann et al.*, 2011]. Examining the drift velocities in the Weddell sea, *Schwegmann et al.* [2011] find an overestimate of satellite-derived estimates of between 30% and 40%. Drift patterns and directions are, however, generally well represented. In the central Arctic, *Kwok et al.* [1998] find that the Polar Pathfinder data set realistically represents in situ measurements of the large-scale drift patters.

[33] While observation-based estimates of sea-ice concentration and sea-ice drift exist from 1979 onwards, no such observation-based time series exists for sea-ice thickness estimates. We therefore focus our comparison on the Pan-Arctic Ice-Ocean Modeling and Assimilation System (PIOMAS) [*Zhang and Rothrock*, 2003] reanalysis of ice thickness from 1979 until today. This system is based on a coupled sea-ice ocean model that is driven by NCEP/NCAR atmospheric reanalysis data (see next paragraph). Additionally, sea-ice concentration based on the NASA Team algorithm and sea-surface temperature as provided by the NCEP/NCAR reanalysis are assimilated into the model. Estimates from sea-ice thickness from PIOMAS agree well with satellite-derived estimates of sea-ice thickness for the past decade [*Schweiger et al.*, 2011]. For earlier decades, sporadic comparisons with field measurements also show reasonable agreement with PIOMAS estimates, though spatial data coverage for such comparison is rather low.

[34] Most of the analysis presented here is based on the fully coupled setup MPI-ESM. Such setup is fundamentally incapable of simulating the correct timing of observed fluctuations in the sea-ice cover caused by internal variability of the Earth's climate system. To estimate how much of an impact this has on the simulated sea-ice cover, we also carried out simulations with our standalone ocean model MPIOM. For these simulations, the ocean was forced by atmospheric fields provided by either the NCEP/NCAR reanalysis [*Kalnay et al.*, 1996] or by the ERA-Interim reanalysis [*Simmons et al.*, 2007]. For the latter, the sea-ice concentration field that is used for the reanalysis is provided by ECMWF. Comparing these original ERA-Interim fields of sea-ice concentration with the directly satellite-derived data discussed above, we find that the ERA-Interim ice-concentration fields are in between those provided by the NASA Team algorithm and that of the Bootstrap algorithm.

[35] In the next section, we will discuss how the representation of reality that is provided by these various data products during the past few decades compares with the MPI-ESM CMIP5 simulations.

4. Evaluating the Model: The MPI-ESM Historical Simulations

[36] In order for the results of any climate model to be of relevance for the real world, we must at least show that such climate model is capable of representing the known conditions of today's climate. Of particular relevance in this respect is a realistic modeling of the variability of a climate variable's state around some long-term trend, and a realistic modeling of the seasonal cycle. Such assessment allows one to estimate the model's skill to realistically simulate the response of a climate variable to changes in the forcing. Such comparison between the observed and the modeled state of the Arctic sea-ice cover will be the topic of this section. For this comparison, we focus on the "historical" simulations [*Taylor et al.*, 2012] of MPI-ESM-LR, which cover the period for which sea-ice observations are available. Since the historical simulations only run up to the year 2005, here we extend them to the year 2007 by including the first 2 years of our scenario simulations. Between 2005 and 2007, all scenario simulations are very similar, and we decided to use the RCP4.5 scenario simulations to extend the historical simulations. For simplicity, we will in the following apply the term "historical simulation" to this extended simulation. We will also use this term to refer to the 20th century simulations of our CMIP3 contribution based on ECHAM5/MPIOM, extending that simulation to the year 2007 by adding the first years of the CMIP3-A1B scenario simulations.

[37] We start with the de facto standard test on any large-scale sea-ice simulation, namely a comparison of observed and modeled evolution of sea-ice extent. Focusing first on the differences between our three ensemble members, we find that they lie close together throughout the 20th century (Figure 2). The root-mean-square difference in simulated sea-ice extent between two ensemble members varies from around 0.45×10^6 km^2 in June, November, and December to around 0.65×10^6 km^2 in February, March, April, and September, with intermediate values in the other months.

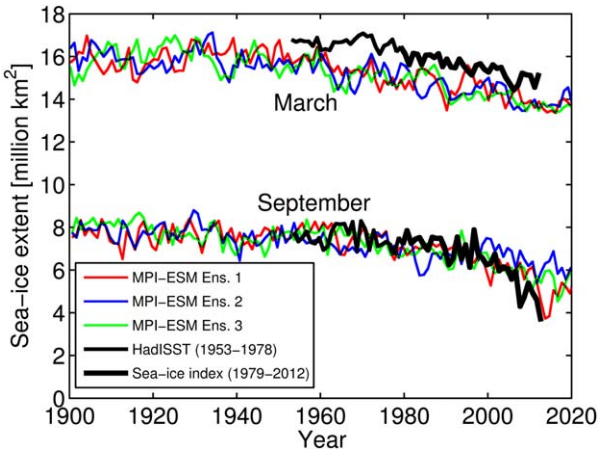

Figure 2. Time evolution from 1900 to 2020 of Arctic sea-ice extent in March and September. The colored lines show individual ensemble members of MPI-ESM simulations, the bold black line shows observations of sea-ice extent from 1953 until 2012 following *Meier et al*. [2012].

Repeating such analysis for our RCP45 scenario simulations throughout the 21st century, the difference between individual ensemble members increases in particular in summer, reaching values of more than 1.0×10^6 km^2 from August to October. This is indicative of a loss of predictability and an increase in interannual variability in summer sea-ice extent throughout the 21st century [cf., *Notz*, 2009; *Goosse et al*., 2009; *Holland et al*., 2011]. Such increase in interannual variability is also directly seen within the individual ensemble members. There, year-to-year differences in sea-ice extent are normally distributed with a standard deviation of between 0.4 and 0.5 $\times$ 10^6 km^2 all year round throughout the 20th century. In the 21st century, the standard deviation decreases to around 0.3 from January to May, but increases to values between 0.7 and 0.9 $\times$ 10^6 km^2 from August to October.

[38] Compared to observations, we find that during the period 1953–2012 MPI-ESM simulates too little ice in winter, with a root–mean-square error in simulated March sea-ice extent of 1.2 $\times$ 10^6 km^2, 1.3 $\times$ 10^6 km^2, and 1.4 $\times$ 10^6 km^2 for our three ensemble members. The summer-time difference between simulations and observations is close to the simulated difference between individual ensemble members and amounts in September to 0.7 $\times$ 10^6 km^2, 0.9 $\times$ 10^6 km^2, and 0.7 $\times$ 10^6 km^2 for our three ensemble members. Note, in particular, that the simulated sea-ice evolution of ensemble member 1 is similar to the observed evolution up to the year 2012 observational estimate of summer sea-ice extent. While such comparison of observed and simulated sea-ice extent has become the de-facto standard in assessing the quality of large-scale sea-ice simulations, we will see later that there are significant issues with focusing on this particular variable.

[39] We therefore now turn to the more meaningful assessment of the spatial distribution of sea-ice concentration. In wintertime, the comparison of MPI-ESM model simulations (Figure 3a) with any of the two satellite records NSIDC-CDR (Figures 3b and 3c) and NSIDC-NASA Team (not shown) gives consistent results: The large-scale features of the observed spatial distribution of Arctic sea ice are well simulated, with a mean absolute deviation of around 8 % in sea-ice concentration compared to both algorithms. Regionally, MPI-ESM simulates too low a sea-ice concentration in the Sea of Okhotsk, East of Greenland in the Odden ice tongue, in the Barents Sea, and in the Labrador Sea (Figure 3c). South of the Bering Strait, the model shows too high a sea-ice concentration. The same results hold qualitatively and quantitatively also for the ECHAM5/MPIOM simulations (Figure 3d). The reanalysis-forced simulations (Figures 3e and 3f) result in a slightly smaller mean absolute deviation of around 5% in sea-ice concentration compared to the satellite data. Both reanalysis-driven simulations result in too much ice around Greenland and too little ice in the Bering Strait and in the Sea of Okhotsk. For the latter, agreement between observations and model is however much improved for the ERA-Interim forced simulation compared to all other simulations (Figure 3f).

[40] In summertime, the MPI-ESM simulations show generally smaller biases in concentration than the ECHAM5/MPIOM simulations (Figure 4). The value of the mean bias relative to concentration depends, however, on the reference observational data set: Relative to the ice-concentration distribution provided by the NASA Team algorithm, the mean absolute concentration bias of MPI-ESM is less than 10%, while the bias is about 14% relative to the Bootstrap algorithm. Concentration biases in the historical simulations of ECHAM5/MPIOM exceed on average 16% compared to both algorithms. Regionally, the improvement in MPI-ESM relative to ECHAM5/MPIOM is primarily caused by the improved representation of sea ice close to Siberian and North-American shore lines. Here, ECHAM5/MPIOM simulated too much sea ice, with an absolute concentration bias exceeding 50% compared to observations. The largest biases in the MPI-ESM simulations compared to observations are around 20% in limited regions of the Kara Sea, North of Greenland, and in the Canadian Archipelago, where simulations show too little ice, and in parts of the Laptev Sea, where simulations show too much ice.

[41] Our reanalysis simulations show a rather peculiar behavior in terms of their regional biases (Figures 4e and 4f): While the spatial patterns and biases in the NCEP-forced simulation are very similar to those of the ECHAM5/MPIOM simulations, the spatial patterns and biases of the ERA-Interim-forced simulation are very similar to those of MPI-ESM. This allows us to draw conclusions regarding the overall quality of these two reanalysis data sets for the Arctic during the past 30 years: Since the overall ice cover in summertime is an integrative response to the combined forcing of wind, temperature, humidity, precipitation, and radiation, the comparably poor spatial distribution of the sea-ice cover resulting from NCEP-forcing makes it very likely

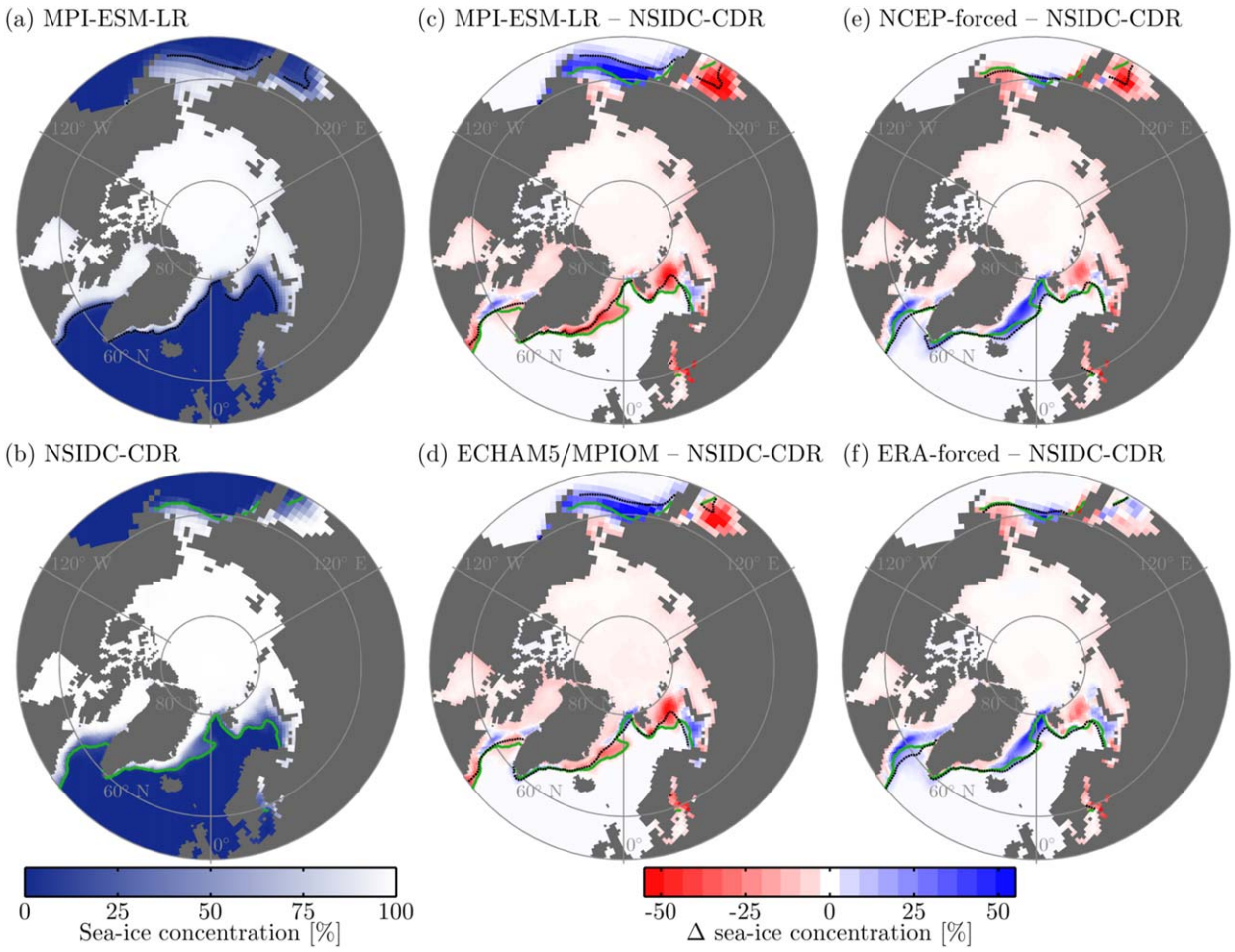

Figure 3. Mean sea-ice concentration in March for the period 1979 to 2007. (a) MPI-ESM-LR; (b) NSIDC-CDR satellite observations; (c) difference between MPI-ESM-LR and NSIDC-CDR; (d) difference between ECHAM5/MPIOM and NSIDC-CDR; (e) difference between NCEP-forced MPIOM and NSIDC-CDR; and (f) difference between ERA-INTERIM-forced MPIOM and NSIDC-CDR. The green contour line denotes the NSIDC-CDR ice edge and the black one the respective modeled ice edge.

that the overall climate state of the Arctic is better described by the ERA-Interim reanalysis.

[42] Moving now from such first assessment of the regional distribution of sea-ice concentration to a quantitative comparison of observed and modeled estimates of sea-ice area and sea-ice extent, an issue arises that gives rise to some concern regarding any study comparing modeled and observed sea-ice coverage. Usually, such comparisons are based on sea-ice extent because of the smaller observational uncertainty in this quantity (Figure 1). Climatically, sea-ice extent is, however, much less relevant than the sea-ice area that determines, for example, the ice-ocean heat exchange or albedo changes. Comparisons of sea-ice extent are therefore only a substitute for comparisons of sea-ice area (or sea-ice concentration). This differentiation between sea-ice area and sea-ice extent would not matter for intercomparisons between models and observations if both metrics would result in the same result regarding a model's quality. We find, however, that this is not the case: there is sometimes good agreement between one of our simulations and observations in sea-ice area with far less good agreement in sea-ice extent and vice versa.

[43] This becomes obvious when we examine the seasonal cycle of sea-ice area and extent. Focusing first on extent, we find good agreement between the MPI-ESM simulated seasonal cycle and the observed seasonal cycle from both satellite products (Figure 5a). In particular in summertime, there is almost perfect agreement with observations, while there is slightly too little ice in MPI-ESM in wintertime. In contrast, the comparably poor performance in summertime regarding sea-ice extent of both the ECHAM5/MPIOM simulations and the NCEP-forced simulations stand out. In particular, the poor performance of our ocean model MPIOM if driven by NCEP reanalysis is notable, since MPIOM driven by ERA-Interim reanalysis results in a good agreement of the modeled ice extent with observations.

[44] Examining now the climatically relevant seasonal cycle of sea-ice area, we find a very different picture (Figure 5b). In summertime, there is still almost perfect agreement of the MPI-ESM simulations with the sea-

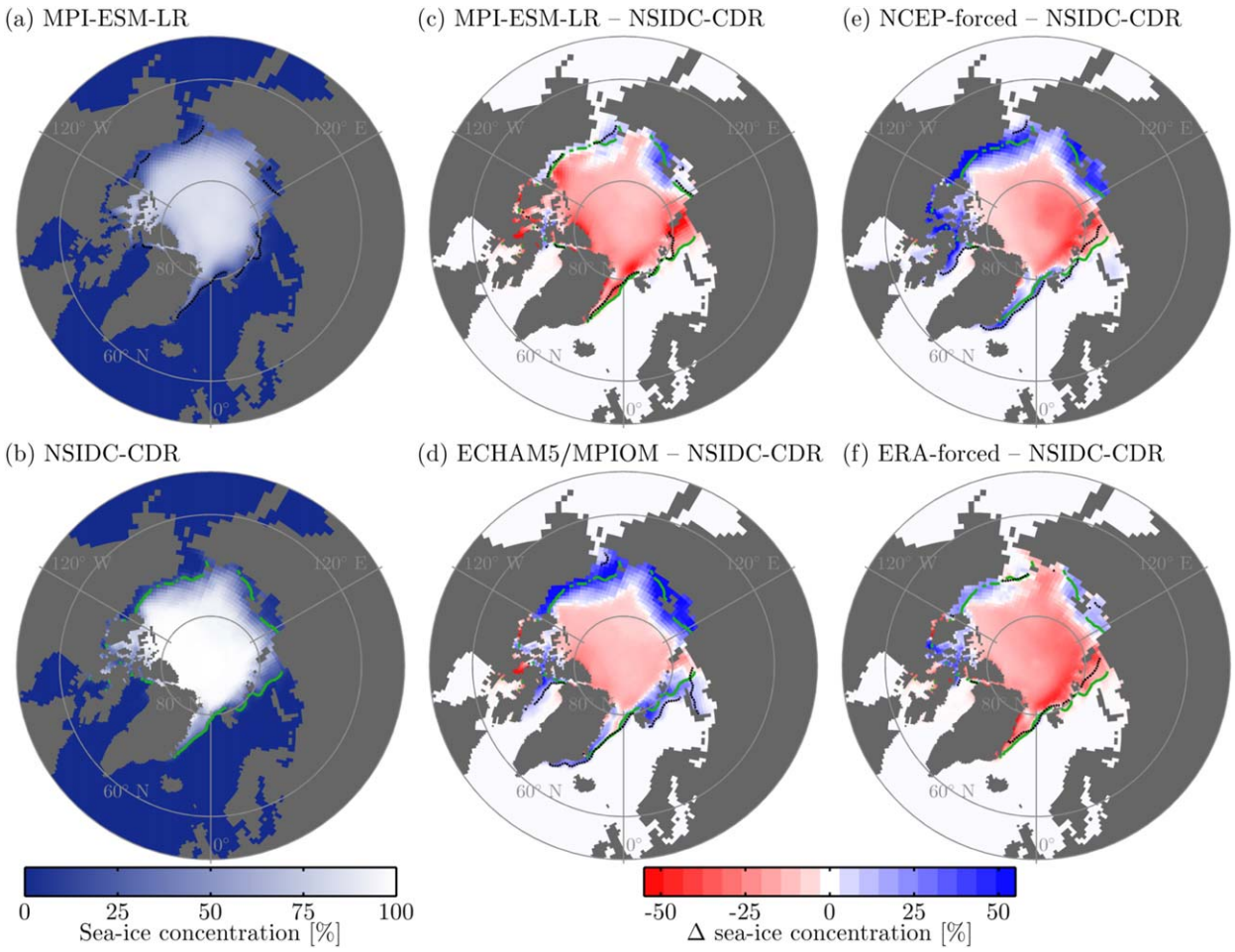

Figure 4. Mean sea-ice concentration in September for the period 1979–2007. (a) MPI-ESM-LR; (b) NSIDC-CDR satellite observations; (c) difference between MPI-ESM-LR and NSIDC-CDR; (d) difference between ECHAM5/MPIOM and NSIDC-CDR; (e) difference between NCEP-forced MPIOM and NSIDC-CDR; and (f) difference between ERA-INTERIM-forced MPIOM and NSIDC-CDR. The green contour line denotes the NSIDC-CDR ice edge and the black one the respective modeled ice edge.

ice area as estimated from the NASA Team based sea-ice index, with slightly too little ice in wintertime compared to this data set. If, however, we focus on the sea-ice area as derived from the Bootstrap algorithm, MPI-ESM simulates too little ice throughout the year, while now there is at least in summer almost perfect agreement of the MPIOM simulations forced by NCEP reanalysis with this observational record. The bootstrap estimate of sea-ice area is also relatively close to the modeled sea-ice area from ECHAM5/MPIOM.

[45] Putting these findings into numbers, and focusing on the probably more reliable estimate of sea-ice concentration based on the Bootstrap algorithm, the issues regarding sea-ice area versus sea-ice extent become obvious (Table 1): While the simulated sea ice of ECHAM5/MPIOM has a roughly 20% too large sea-ice extent in summer, it only has a roughly 7% too large sea-ice area. In contrast, MPI-ESM has worse agreement in sea-ice area, with about 20% too little sea ice, but much better agreement in sea-ice extent, with a deviation of only 6% relative to the Bootstrap estimate. Focusing on the geophysically irrelevant sea-ice extent, one would conclude from these numbers that the sea-ice coverage simulated by MPI-ESM agrees better with observations than that simulated by ECHAM5/MPIOM, while the opposite holds regarding the geophysically relevant sea-ice area. We find qualitatively the same regarding the differences between our NCEP-forced simulations (with in summer almost perfect agreement in sea-ice area with the Bootstrap algorithm) and our ERA-Interim forced simulations (with good agreement in sea-ice extent). As discussed by Notz (Sea-ice extent provides a limited metric of model performance, submitted manuscript, 2013), this finding of different results regarding model quality based on sea-ice area as opposed to sea-ice extent also holds for any other CMIP5 model. The quantitative reliability of studies comparing models and observations primarily based on sea-ice extent could become questionable.

[46] It might therefore be desirable to focus more on sea-ice area in model-data intercomparisons. However, such focus is currently made difficult by the poor

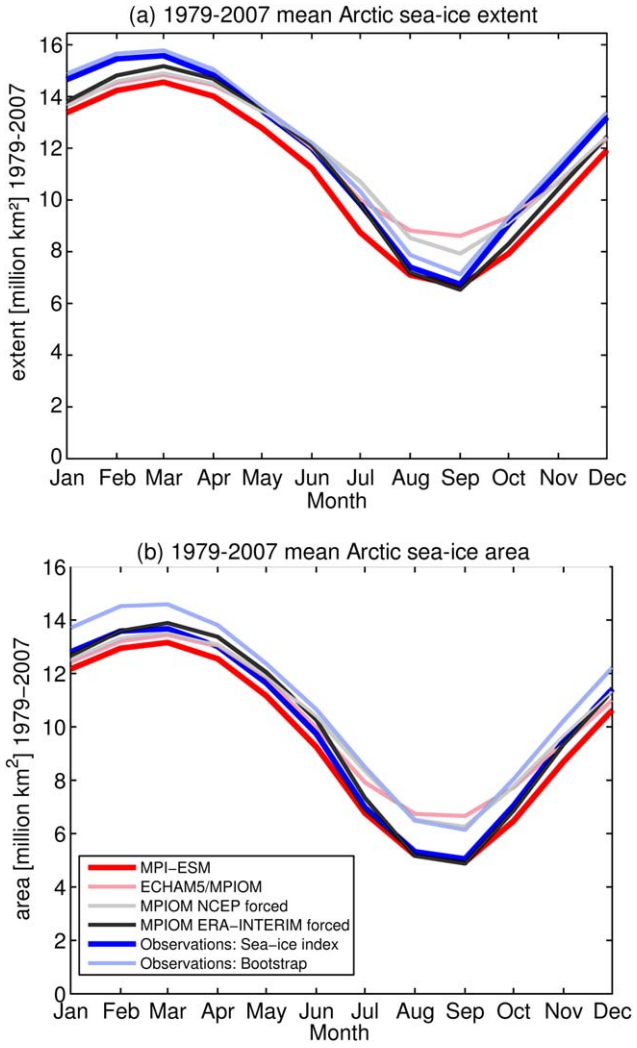

Figure 5. Comparison of the mean seasonal cycle in (a) sea-ice extent and (b) sea-ice area for the period 1979–2007.

agreement of the different observational data sets regarding this quantity: For example, the MPI-ESM simulated ensemble-mean September area and extent for the period 1979–2007 are almost identical to the value obtained from the NSIDC sea-ice index (i.e., NASA Team algorithm), but they are (as just seen) below the observational value following the Bootstrap algorithm (Table 1). Hence, when considering the total summertime sea-ice area, a comparison with observations is crucially affected by the uncertainty arising from the different algorithms for obtaining sea-ice concentration.

[47] Furthermore, any comparison of the mean state of the seasonal cycle between observations and models for a particular period might be misleading because the timing of internal variability cannot be represented by unforced models. As a measure for such internal variability, we use the standard deviation around a 29-year long linear trend that we subtract from the data. This length was chosen to match the length of the time period 1979–2007 that we use as our reference period throughout this paper. For the observational record, we find that the variability in sea-ice area around the long-term linear trend is smallest in wintertime, with standard deviation of around 0.25 million km^2 (Figure 6a). The variability then increases throughout summer, reaching a maximum of 0.52 million km^2 in October. Given that (a) the timing of this variability is caused by the chaotic nature of the climate system and (b) that the variability is changing in a changing climate (see below), we cannot expect any climate model to reproduce the exact same variability pattern at exactly the same time. Hence, it is not surprising that MPI-ESM does not exactly match the observed variability for the time period 1979–2007: During that period, compared to the NSIDC-observations the model has more variability in winter and less variability in summertime.

[48] However, shifting the analyzed model-period by 10 years to 1989–2017 results in very good agreement of

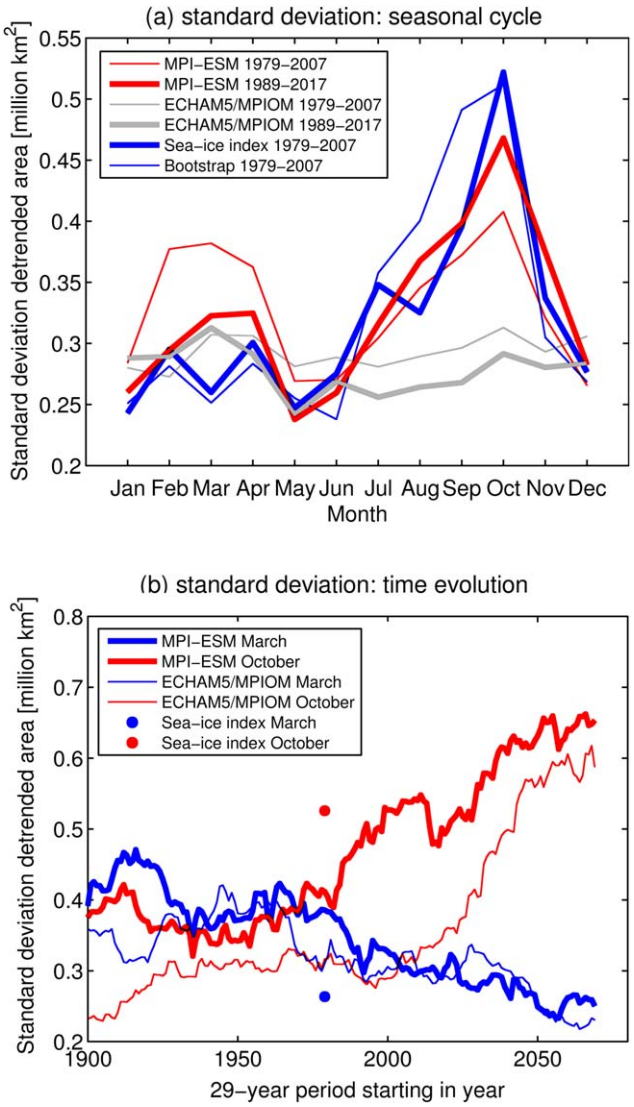

Figure 6. (a) Comparison of the standard deviation of sea-ice area around a 29 year long linear trend; (b) Time evolution of the standard deviation around a 29 year long trend.

the seasonal cycle compared to observations. Hence, a mere snapshot of the variability during a specific period might give a misleading picture, because of the substantial time evolution of internal variability in a changing climate (Figure 6b). While the summertime variability is increasing steadily after 1990 in our simulations, the wintertime variability decreases simultaneously. The increasing variability in September is most likely caused by the thinning of the sea-ice pack [e.g., *Notz*, 2009; *Goosse et al*., 2009], which allows for larger year-to-year variability of the ice pack. The decreasing variability in wintertime, however, is currently not well understood. A geometric argument related to the land-mass distribution as suggested by *Eisenman* [2010] can be excluded since the decreasing wintertime sea-ice cover should then result in an *increased* variability of the ice-covered area.

[49] Despite the long-term evolution of internal variability, the poor performance of the ECHAM5/MPIOM simulations in simulating the observed seasonal cycle of internal variability is obvious: the variability is almost constant throughout the year for today's climate condition (Figure 6a), and only in 2040 reaches the observed summer-time values (Figure 6b). This lack of realistic variability in ECHAM5/MPIOM for the 20th century was found by *Koldunov et al*. [2010] to be one of the major issues with our earlier model version, and the improvement in MPI-ESM is hence a clear indication for an overall improved simulation of the Arctic sea-ice cover.

[50] Another metric that is often used to compare climate simulations with observations is the long-term trend. For example, *Stroeve et al*. [2007] pointed out that CMIP3 models significantly underestimated the observed trend in Arctic sea-ice decline. However, it is a rather difficult task to estimate if such discrepancy between observed and simulated trends are caused by model deficiencies or simply by internal variability of the real climate that lead to an untypically large observed trend. This latter hypothesis is underpinned by our CMIP5 simulations. Analyzing 29-year trends throughout the 20th and 21st century, we find that the model is, in principle, capable of modeling as large a trend as has been observed in reality (Figure 7): The observed trend is reproduced by one of the MPI-ESM ensemble members for the 29-year long September trend in sea-ice area starting any time between 1984 and 1992. However, such strong negative trend is unusually large. Equally long modeled trends starting after 1992 are usually weaker, which lends some support to the notion that the extreme magnitude of the observed long-term trend is caused by a combination of internal variability and an external driver. If indeed internal variability contributed to the extreme loss of sea ice, the rate of sea-ice decline is likely to slow down in the future. The large magnitude of internal variability is also reflected by the fact that one ensemble member of the MPI-ESM simulation even results in a slightly *positive* trend for the 29-year long period starting in 1979. Hence, in MPI-ESM even positive trends in sea-ice area are possible on climatic time scales even if the background climate is warming [cf. *Kay et al*., 2011]. The large modeled variability of long-term trends ranging from slightly positive to as extremely negative as has been observed is a clear indication that a direct comparison of modeled trends with observations are meaningless unless the models' internal variability is carefully taken into account.

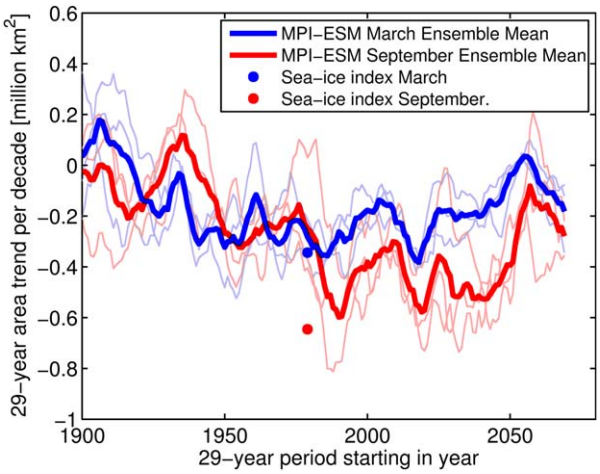

Figure 7. Time evolution of 29 year trends in sea-ice area based on the mean of the MPI-ESM historical simulations, extended by RCP 4.5 scenario simulations. The thin lines denote individual ensemble members.

[51] Analyzing the seasonal cycle of sea-ice melt, we find generally good agreement with observations. In the NSIDC observations based on the Bootstrap algorithm, the minimum area of the Arctic sea-ice cover was on average reached on 6 September (not shown), 1 day earlier than the mean for our historical simulations over the same period. The maximum area of the Arctic sea-ice cover in the observations was on average reached on 1 March, 6 days earlier than in our simulations. Hence, the total melt period as defined by the loss of sea-ice area in the observations was on average 188 days long, whereas MPIOM-ESM simulates for the same period on average a melt duration of 181 days. Given the standard deviation of 11 days of the simulated duration of the melt period, the differences to the observations are not significant. We do not find significant trends in our simulation, and with simulated melt periods ranging during the period 1979–2007 rather randomly from 151 to 204 days, also the trend of +5 days/decade that is displayed by the NSIDC record is not significant if the duration of the melt period as modeled by MPI-ESM reflects the real internal variability.

[52] Regarding sea-ice volume, a comparison with observations is basically impossible, since we do not have reliable, Arctic-wide measurements of sea-ice thickness. As a substitute for direct observations, we here use the PIOMAS reanalysis of Arctic sea-ice volume [*Zhang and Rothrock*, 2003] to compare against MPI-ESM. In PIOMAS, over the past 30 years, on average 15,400 km^3 of ice melt every year and about 15,000 km^3 new ice form every winter. This imbalance

results in a mean decrease in sea-ice volume by 226 km^3/year in March and by 252 km^3/year in September during the period 1979–2007 (Table 1). For the same period, in MPI-ESM-LR, 15,770 km^3 melt every year, whereas an average amount of 15,660 km^3 new ice form every winter. Hence, in MPI-ESM-LR the sea-ice volume decreases more slowly than in the PIOMAS reanalysis, with an average decrease of 97 km^3/year in March and by 90 km^3/year in September. The drop in sea-ice volume from a September value of around 15,000 km^3 in 1979 to about 5000 km^3 in 2011 that is displayed by the PIOMAS reanalysis is simulated by MPI-ESM over a much longer period: While we do have a similar volume of around 5000 km^3 in our simulations at around 2011, the drop in sea-ice volume from values of around 15,000 km^3 starts between 1910 and 1930 in all three ensemble members of MPI-ESM-LR (see discussion around Figure 13c below). If we include the entire historical simulation in our analysis, the mean seasonal cycle in sea-ice volume in MPI-ESM-LR is almost identical to the seasonal cycle as displayed by the PIOMAS reanalysis.

[53] Some insight into the underlying cause of the very different trends in sea-ice volume can be gained by considering our NCEP and ERA-Interim-forced MPIOM simulations (Table 1). Both these simulations have an overall trend in sea-ice volume that is much closer to the PIOMAS reanalysis than the fully coupled MPI-ESM simulations. Despite a generally higher sea-ice volume, the NCEP-forced simulations match the PIOMAS September trend in sea-ice volume exactly. This agreement is striking because the PIOMAS reanalysis is based on NCEP atmospheric forcing: In contrast, we find smaller trends in sea-ice volume for ERA-Interim atmospheric forcing in our simulations, which is indicative of the major role that the atmospheric forcing has on the heat-flux convergence and divergence that ultimately governs the trend in sea-ice volume. It is worrying that the trend in sea-ice volume is so strongly dependent on the choice of the reanalysis data set, which is indicative of the large uncertainty that might be associated with the reanalyzed trends in sea-ice volume as displayed by the PIOMAS-reanalysis. We should also note that one ensemble member of the mixed-resolution model MPI-ESM-MR shows larger trends both in September and in March compared to those obtained from the PIOMAS reanalysis. Hence, again, the internal variability of individual trends is too large to allow for a meaningful assessment of the quality of individual model simulations.

[54] In addition to a comparison of the overall volume, the PIOMAS data set also allows for a comparison of the spatial distribution of sea-ice thickness (Figure 8). Compared to ECHAM5/MPIOM simulations, the MPI-ESM simulations give a more realistic spatial distribution of ice thickness. In wintertime, the central Arctic is now covered by ice with an average thickness of about 2.2 m, with the thickest ice being found North of Greenland and the Canadian Archipelago. Such spatial distribution of ice thickness agrees much better with the PIOMAS reanalysis than the rather evenly distributed 4 m thick ice that ECHAM5/MPIOM produced for the CMIP3 simulations. In summertime, the mean ice thickness in MPI-ESM during the period 1979–2007 is about 1 m, which is too thin compared to the PIOMAS mean ice thickness of about 1.6 m. ECHAM5/MPIOM produced unrealistically thick ice with a mean thickness of about 2.4 m in summer. The mean sea-ice thickness of MPI-ESM since the beginning of the modeled decline in sea-ice thickness around 1920 until today compares well with the mean PIOMAS thickness. As mentioned earlier, we do currently not understand why sea-ice thickness starts to decrease much earlier in our simulations than in the PIOMAS reanalysis.

[55] As our final comparison of the modeled sea-ice state against data, we will now briefly turn to the drift pattern of Arctic sea ice simulated by MPI-ESM (Figure 9). The general flow direction in March agrees almost everywhere with the Polar Pathfinder observations. This is encouraging since the representation of flow direction has been found to be very reliable in the Polar Pathfinder data set [*Kwok et al*., 1998; *Schwegmann et al*., 2011]. In comparison to satellite-derived drift speed, the modeled sea ice moves generally too fast. Differences are particularly large in the Fram Strait and in the Davis Strait, where the model estimates flow speeds that are more than twice as fast as the Polar Pathfinder data. Especially in these regions of rather high flow speeds, however, the Polar Pathfinder data might underestimate the real flow speeds. However, the underestimate of 30%–40% that *Schwegmann et al*. [2011] found for Antarctic sea ice cannot fully explain the biases in modeled drift speed.

[56] Overall, there is improvement in many aspects of the simulated Arctic sea-ice cover in MPI-ESM compared to ECHAM5/MPIOM. Because the ocean and the sea-ice model are almost identical between these two model versions, we can easily identify the two factors that are possibly responsible for this improvement: changes in the atmospheric model and a tuning of the sea-ice model to an average lower sea-ice volume. Of these two, the tuning is probably less important for the overall improvement, because MPI-ESM simulates despite the tuning a sea-ice cover similar to ECHAM5/MPIOM if forced by NCEP reanalysis atmospheric data. Therefore, the main reason for the improved simulation of the Arctic sea-ice cover must be the improved representation of atmospheric processes owing to changes in ECHAM6 compared to ECHAM5. These changes are described in detail by *Stevens et al*. [2013], and primarily concern the representation of solar radiation, the representation of deep convection and the extension of the vertical grid to 0.1 hPa. These changes have led to clear improvements of the simulated atmospheric state compared to ECHAM5/MPIOM. Compared to the ERA-Interim reanalysis (which seems superior to the NCEP reanalysis in particular in the Arctic), we find for example that MPI-ESM has basically gotten rid of the significant winter cold bias over European and Asian high latitudes that existed in ECHAM5/MPIOM (Figures 10a–10c). There remains,

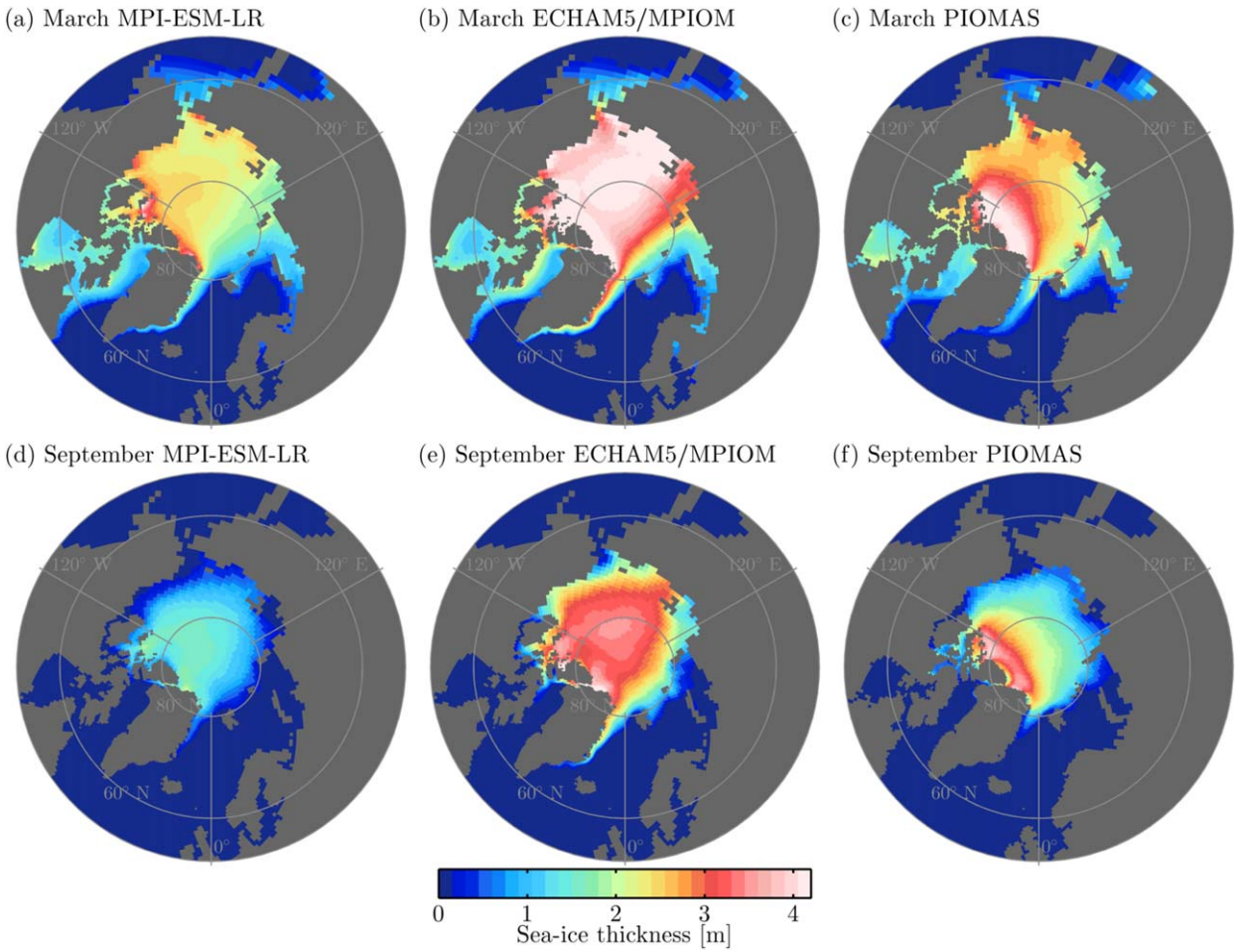

Figure 8. Comparison of mean sea-ice thickness for the period 1979–2007.

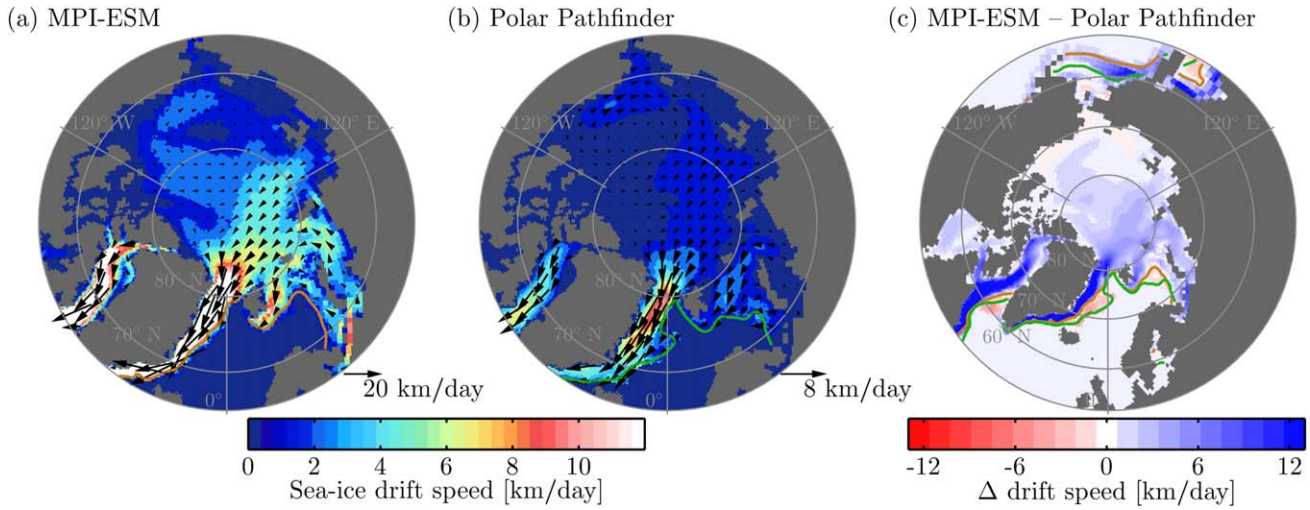

Figure 9. Comparison between mean modeled and observed March Arctic sea-ice drift vectors (only every eighth vector is plotted) and drift speeds in the period 1979 to 2005: (a) simulated MPI-ESM-LR; (b) satellite-derived Polar Pathfinder; and (c) simulated minus satellite-derived drift speed. The contour lines denote the modeled (orange) and observed (green) ice edges (15% ice concentration).

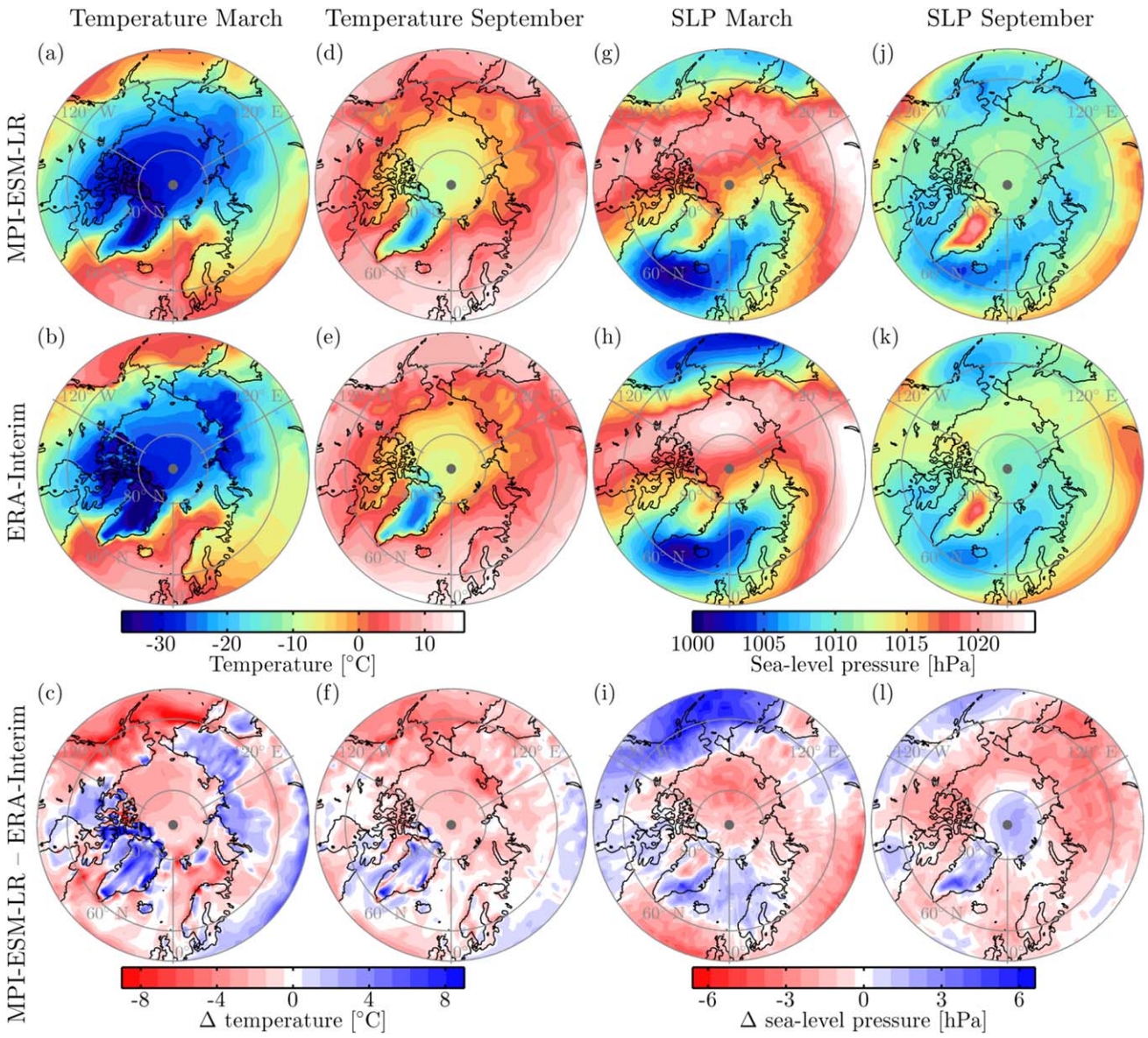

Figure 10. (left) Mean surface temperature and (right) sea-level pressure in MPI-ESM-LR compared to ERA-Interim for the period 1979–2007.

however, a cold bias over Alaska and some warm bias over Northern Canada and Northern Greenland. This improvement in simulated high-Northern temperature goes along with significant improvements in simulated winter surface pressure patterns (Figures 10g–10i). As described by *Koldunov et al.* [2010], ECHAM5/MPIOM had significant positive biases in Northern hemisphere surface pressure patterns. These strong biases are almost absent in MPI-ESM-LR, with only some positive bias larger than 2 hPa remaining in winter surface pressure over Southern Alaska and a slight negative bias over the central Arctic/Beaufort gyre region. The same general improvement in simulated temperature and pressure fields also holds in summer (Figures 10d–10f, 10j–10l). While there is still some slight summertime cold bias over much of the Arctic, this cold bias is significantly reduced compared to ECHAM5/MPIOM. There are still some minor issues with simulated pressure patterns in summertime, with generally too low surface pressure in MPI-ESM reaching from Canada and Alaska all over Siberia, but the large positive bias in surface pressure in the central Arctic Ocean that was present in ECHAM5/MPIOM is now almost gone.

[57] It is beyond the scope of this study to analyse in detail which changes in ECHAM6 relative to ECHAM5 are responsible for the improved representation of the atmospheric circulation in high Northern latitudes. An initial analysis suggests that much of the improved pressure patterns is caused by a stronger focus on a realistic representation of Arctic pressure patterns during the tuning of the orographic gravity wave drag in ECHAM6. The improved variability in our simulations

are likely related to the improved representation of processes in the upper stratosphere, which allows for the better representation of the variability of near-surface processes that are triggered by stratosphere-troposphere interaction (see also Omrani et al., Stratosphere key for wintertime atmospheric response to warm Atlantic decadal conditions, submitted to *Climate Dynamics,* 2013*)*. We should note that ECHAM6 still has issues with spatial patterns of modeled precipitation. This leads, in particular, to too large a runoff of Siberian rivers into the Arctic Ocean [*Hagemann et al.*, 2012].

[58] To summarize, from this analysis of our historical CMIP5 simulations we can conclude that the MPI-ESM is well capable of simulating observed large-scale features of the Arctic sea-ice cover. The modeled seasonal cycle of sea-ice extent and sea-ice area is very close to observations, as are the modeled variability around the long-term trend. One ensemble member of our simulations reproduces the observed long-term trend in sea-ice extent and area, while other ensemble members display a much slower decline of the ice pack. In the light of these findings, it seems fair to assess that MPI-ESM can be used as a tool to examine the past and possible future evolution of the real sea-ice cover. Hence, having positively evaluated our historical simulations against reality, we can in the following draw conclusions about the real sea-ice cover from our control and scenario simulations.

5. Assessing Internal Variability: The MPI-ESM Control Simulation

[59] As part of our CMIP5-simulations, we carried out a 1000-year long preindustrial control simulation. For this simulation, the annual cycle of the forcing was held constant at preindustrial conditions. For all sea-ice related parameters, the long-term trends over the entire control simulation are less than 0.001%, and we can safely assume the control simulation to be stationary. The annual mean volume of sea ice in our control run is 21.7×10^3 km^3, close to our aim of around 20×10^3 km^3 when tuning the sea-ice component of our coupled model. Compared to the control simulation of ECHAM5/MPIOM, we now have about 20% less sea-ice volume (see Table 1). However, with about 16.5×10^3 km^3, the overall change in sea-ice volume between March and September is only slightly larger for the MPI-ESM control run compared to the ECHAM5/MPIOM control run, where the amplitude of seasonal melting and growth amounts to 15.1×10^3 km^3. This indicates that the mean annual cycle of near-surface heat-flux convergence in the Arctic is only slightly affected by the changes made in the atmospheric component of MPI-ESM. For sea-ice extent, however, the seasonal cycle is different between ECHAM5/MPIOM and MPI-ESM control simulations, whereas the annual mean sea-ice extent is almost the same. Regional differences between the ECHAM5/MPIOM pre-industrial control simulation and the MPI-ESM preindustrial control simulation are very similar to the differences in the historical simulations discussed in section 4.

[60] Since the annual cycle of the forcing was held constant for the preindustrial control simulations, we can use this simulation to analyze the internal variability of the Arctic sea-ice cover. We first focus on the significance of short-term and long-term trends (Figure 11a). Doing so, we find that changes in sea-ice area from 1 year to the next are significant at the 5% level of an underlying normal distribution if they exceed 0.94 million km^2 in March and 0.82 million km^2 in September. In comparison, we find that 10-year long trends are significant if their absolute magnitude exceeds 0.14 million km^2 per year in March or 0.11 million km^2 per year in September.

[61] In the following, we will use the modeled internal variability to examine if internal variability can explain the observed retreat of Arctic sea ice. Such analysis hinges crucially on the assumption that the modeled internal variability is representative for the internal variability of the real sea-ice cover. Unfortunately, we have insufficient observational data to reliably test this assumption and some uncertainty will necessarily remain as to the validity of any transfer from the model to the real world. Nevertheless, the good agreement of the seasonal and long-term variability of our 20th century simulations with observed variability lends support to our assumption that the modeled internal variability for preindustrial climate conditions is representative for pre-industrial climate conditions also in the real world.

[62] Based on this, we first use the estimate of significant trends in the model simulations for pre-industrial climate conditions to examine if trends in the observed sea-ice coverage are significant. We find that all positive and negative short-term and long-term trends in observed wintertime sea-ice area are not significant at the 5 % significance level of an underlying normal distribution. In September, we find no significant positive trends on time scales longer than 1 year. However, all negative trends longer than 20 years are unlikely to be caused by preindustrial internal variability at the 5% significance level since 2000 (Figure 11b). On shorter time scales, all negative trends longer than a decade are significant that include the extreme minimum of 2007.

[63] Focusing on extreme values rather than extreme trends, we can use the MPI-ESM pre-industrial control simulation to estimate for each month which sea-ice area could simply occur by chance for preindustrial climate conditions. We find that at the 5% significance level of an underlying normal distribution the March sea-ice cover of a single observation could vary naturally between 13.4 and 15.4 million km^2. Given the length of 33 years of the entire NSIDC record of sea-ice area, chances are less than 2% that the low winter sea-ice area that was observed in 2006, 2007, and 2011 simply happened by chance. For September sea-ice extent, we find a range of internal variability between 5.2 and 6.9 million km^2. Here, 16 out of the 23 years since 1990 show a sea-ice area that is below internal variability at the 5% significance level (Figure 12).

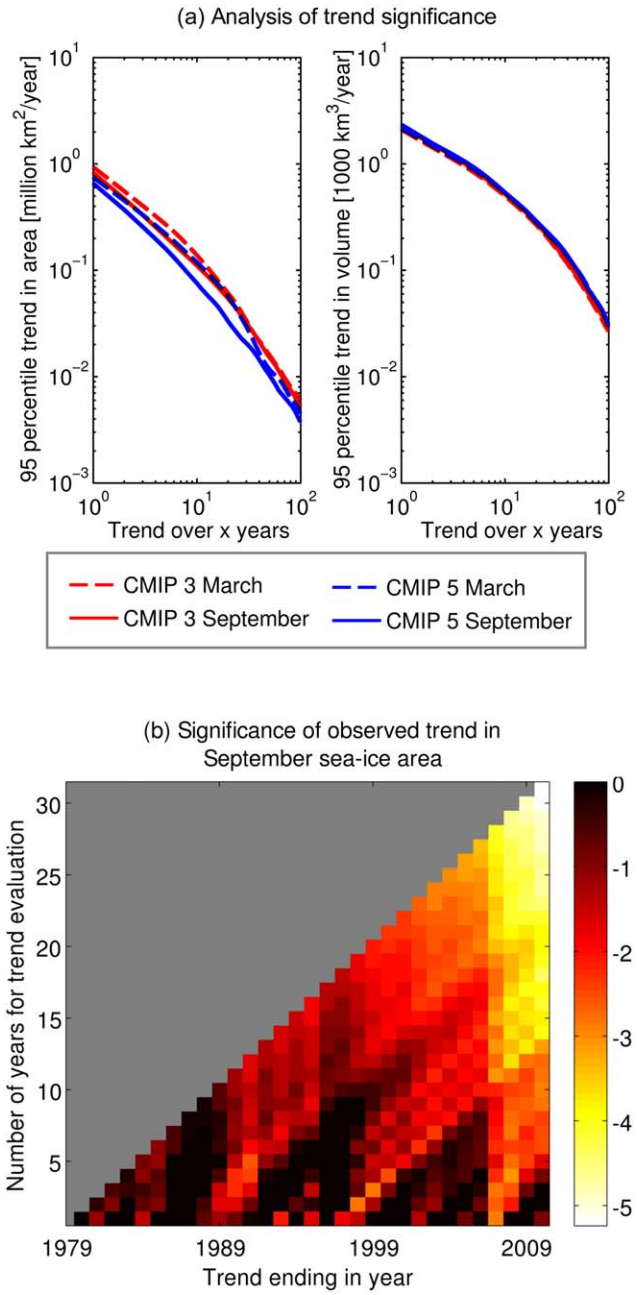

Figure 11. (a) Analysis of trend significance. A trend over the number of years indicated on the x axis is significant if its absolute value exceeds the reduction in sea-ice area (left) or the reduction in sea-ice volume (right) indicated on the y-axis. (b) Analysis of observed trends in September sea-ice area. The color bar indicates the number of standard deviations that a trend is larger than modeled internal variability of the MPI-ESM control simulations. A trend that is less than ±2 standard deviations different to modeled internal variability is not significant (black areas).

[64] Carrying out a similar analysis based on the historical simulations with MPI-ESM, we find for September sea-ice coverage similar statistical values at similar times as in the observations. In March we find higher significance of the modeled retreat than observed, because the historical simulations by MPI-ESM display generally a somewhat smaller ice area than observed.

6. Sea-Ice Response in the MPI-ESM Scenario Simulations

[65] The overall good agreement of the variability and temporal evolution of the Arctic sea-ice cover in our historical simulation not only allows us to use MPI-ESM to estimate internal variability of Arctic sea ice. It also lends credibility to the projection of the future evolution of Arctic sea ice that we simulated following the Representative Concentration Pathways RCP2.6, RCP4.5, and RCP8.5 that are all part of the CMIP5 protocol [*Taylor et al.*, 2012]. With MPI-ESM-LR, we carried out three simulations of each of these three scenarios up to 2100, with one simulation each being extended to 2300. With MPI-ESM-MR, we carried out one simulation for each scenario up to 2100. Since we do not find large differences in any of the parameters analyzed between the LR and the MR model version, we will in the following only discuss details of the LR version, because we have a larger ensemble and longer simulations available.

[66] We first discuss our RCP2.6 simulations, which are based on increasing CO_2 concentration up to 443 ppm in 2053, and declining CO_2 concentration afterward to reach 421 ppm by 2100 and 361 ppm by 2300. In this scenario, the summer sea-ice area continues to decline to a level of around 3 million km^2 by 2020, staying close to this value throughout the remainder of this century (Figure 13b, green line). In the 22nd and 23rd centuries, summer sea-ice area recovers slowly to reach by 2300 values close to the extreme sea-ice minima as observed in 2007 and 2011. Given the substantial restructuring of the world's economy that would be needed to achieve a concentration

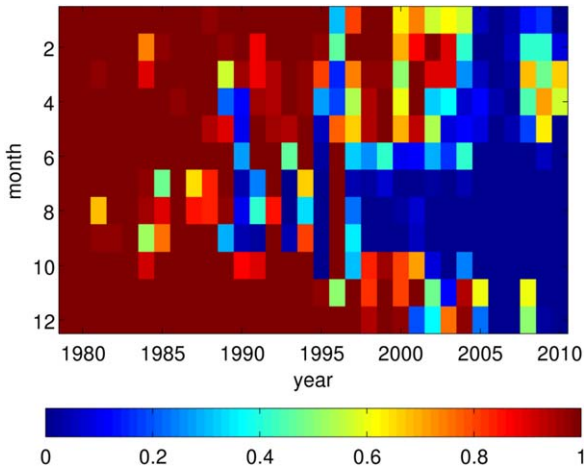

Figure 12. Analysis of observed values of sea-ice area. The color coding indicates the chances that a certain observed area could have happened by chance given the internal variability modeled by the MPI-ESM preindustrial control simulation.

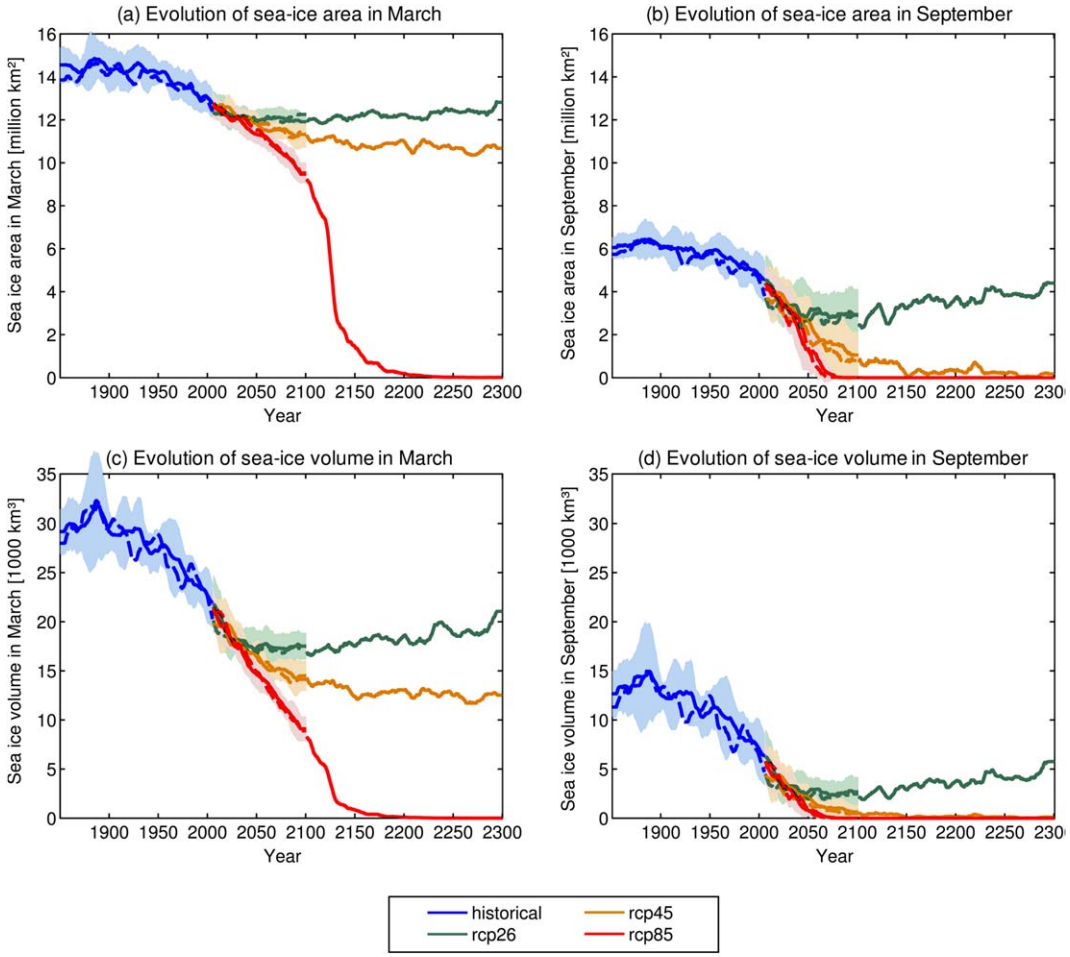

Figure 13. Modeled time evolution of (a, b) sea-ice area and (c, d) volume in MPI-ESM. The lines indicate the 10 year running mean ensemble mean simulated by MPI-ESM-LR, with the shading indicating ±2 standard deviations of the three ensemble members. The dashed lines show the evolution as modeled by MPI-ESM-MR. From 2100 onward, only one ensemble member was simulated.

pathway as assumed in RCP2.6, we can assume that this scenario provides a lower bound on future global warming. Hence, in the absence of any events or actions that will cool the climate additionally, it is unlikely that the Arctic summer sea-ice area will increase much over today's sea-ice area throughout the coming centuries. Given, on the other hand, that the scenario RCP2.6 is at least physically plausible, our results also suggest that a total loss of Arctic summer sea ice is, in principle, still avoidable. The absolute loss of winter sea ice is about 40% smaller than the absolute loss of summer sea ice throughout the 21st century (Figure 13a, green line). This behavior, which is also found in the two other scenarios, is explicable by geometric constraints [*Eisenman*, 2010]. In absolute numbers, we find a decrease in winter sea-ice area to values of around 12 million km^2 throughout the 21st century, which then increases slowly toward 13 million km^2 by the end of the 23rd century.

[67] The RCP4.5 scenario, which is comparable to the CMIP3 scenario B1, is based on a scenario with CO_2 concentration increasing further to reach 539 ppm by 2100, staying almost constant at this level thereafter. In this scenario, the Arctic continues to rapidly lose ice in summer, reaching an overall area of around 1 million km^2 by the end of the 21st century. Once CO_2 concentration remains constant from then onward, the sea-ice loss slows down, with the Arctic becoming fully ice free around 2150 in the single ensemble member whose simulation continues beyond 2100 (Figure 13b, orange line). Wintertime sea-ice cover follows this general pattern of fast retreat throughout the 21st century and slower retreat for the 22nd and 23rd century (Figure 13a).

[68] The RCP8.5 scenario, which is comparable to CMIP3 scenario A2, is based on a much stronger increase in CO_2 concentration than the other two scenarios. Here, CO_2 concentration rises to values of 935 ppm by 2100 and 1962 ppm by 2250, staying constant thereafter. In this scenario, the Arctic becomes in our simulations ice free in summertime during the second half of the 21st century (Figure 13b, red line) and stays ice free in summer throughout the remainder of the simulations. In wintertime, sea-ice area decreases roughly linearly to values of around 9.5 million km^2 by the end

of the 21st century. In the first half of the 22nd century, the retreat of winter sea-ice area accelerates substantially, with only around 1 million km^2 of winter sea ice remaining by 2150. The Arctic then becomes fully ice-free all year round throughout the 23rd century (Figure 13a).

[69] The evolution of Arctic sea-ice volume follows the general trend of the sea-ice area. However, the relative loss of sea-ice volume is generally much larger than that of sea-ice area throughout the 20th century, and generally slower throughout the 21st century (Figures 13c and 13d). Throughout the 20th century, in our simulations the Arctic sea-ice volume in summer has decreased by about 60%. During the same time, summer sea-ice area has only decreased by about 30%. Hence, the relative changes in summer volume have been roughly twice as fast as the changes in summer area throughout the 20th century. In the 21st century, the relative change of summer sea-ice volume becomes roughly equal to the relative change of summer sea-ice area. This behavior is consistent with the fact that in general, thick sea ice thins much faster than thin sea ice [cf., *Bitz and Roe*, 2004], which causes a rapid loss of sea-ice volume in the presence of rather thick sea ice. This is also reflected in MPI-ESM, where we find throughout the 21st century a roughly 20% faster absolute reduction of the volume of thick winter sea ice compared to the loss of summer sea-ice volume. Analyzing the relationship between areal and volume decrease in the PIOMAS reanalysis, we again find that the loss of summer sea-ice volume has been about twice as fast as the observed relative loss in sea-ice area. Hence, it seems reasonable to assume that real sea ice follows a behavior similar to that in MPI-ESM, with the relative change in sea-ice volume decreasing considerably over the coming decades. Therefore, the future evolution of sea-ice volume cannot be estimated reliably by a linear or even exponential extrapolation of the past evolution of Arctic sea-ice volume. Contributions to the public debate as to when the Arctic might become ice free in summertime based on such extrapolations therefore do not reflect our current understanding of the past and future evolution of the Arctic sea-ice cover.

[70] We will now finally use our scenario simulations to examine the direct relationship between the external forcing and the sea-ice state. Focusing first on the direct relationship to external CO_2 forcing, we find as expected a general tendency of less summer sea ice for higher CO_2 concentration (Figure 14b). However, there is no one-to-one mapping of a certain CO_2 concentration to a certain sea-ice area or volume. This is nicely illustrated by the RCP4.5 scenario. Here, CO_2 concentration is kept constant at around 540 ppm from 2100 onward, and nevertheless sea-ice area and volume continue to decrease until all summer sea ice is gone. Hence, in the observations we currently only see the transient response of the sea-ice cover to the CO_2 forcing, which because of the rapid increase in CO_2 concentration corresponds to the maximum sea-ice cover that is possible for any given CO_2 concentration. The equilibrium state of the Arctic sea-ice cover for any given CO_2 concentration is much lower, and in equilibrium, the Arctic can become ice free in summer at a CO_2 concentration of around 500 ppm [*Li et al.*, 2013]. In contrast to the response to CO_2, the response of sea ice to Arctic temperature is much more constrained (Figures 14c and 14d). This is a reflection of the fact that, in general, Arctic temperature and sea-ice coverage react on similar time scales to changes in CO_2 concentration. The time scale of both is set by the slow oceanic response [*Li et al.*, 2013].

[71] For the winter sea-ice cover, we find a rapid decrease from a sea-ice area of around 6 million km^2 to around 3 million km^2 in just a few years. This occurs at a CO_2 concentration of around 1200 ppm (Figure 14a) and a mean Arctic winter temperature of around -5°C (Figure 14c). A similar behavior was also observed in the previous model version ECHAM5/MPIOM, and there has been some speculation whether this behavior constitutes a so-called tipping point [*Winton*, 2006]. A more recent analysis has found that the rapid loss of Arctic sea ice in ECHAM5/MPIOM version goes along with the sudden onset of significant convective precipitation in the Arctic [*Li et al.*, 2013]. We find the same for MPI-ESM, where convective precipitation north of 80° in March is less than 0.1 mm/month throughout the 20th and 21st century. From around 2120 onward, monthly mean values of convective precipitation in the same region regularly exceed 1 mm/month, and values exceeding 10 mm/month are common throughout the 23rd century. As in ECHAM5/MPIOM, this is suggestive of an atmospheric convection feedback. Such feedback, as suggested by *Abbot and Tziperman* [2008], would cause an increased number of convective clouds and increased humidity, which both trap outgoing longwave radiation and cause a rapid loss of Arctic winter sea ice.

7. Conclusions

[72] In this contribution, we have provided an overview of the past and future evolution of Arctic sea ice in MPI-ESM. Doing so, we have come to the following conclusions:

[73] The regional distribution of the modeled Arctic sea ice, and the trend and variability in the sea-ice coverage, are improved in MPI-ESM compared to the model's predecessor ECHAM5/MPIOM. We find that these improvements stem from changes in the atmospheric model component between ECHAM5 and ECHAM6, with the latter simulating more realistic surface pressure patterns and a more realistic temperature distribution in the Arctic. Tuning of MPI-ESM to overall less sea-ice volume has contributed to the improvements, but we find the overall contributions of tuning to the improved simulations to be rather small. Regional differences between MPI-ESM and observations occur in wintertime primarily along the ice edge, with too little modeled ice in the Sea of Okhotsk, the Barents Sea, and east of Greenland, in particular, and too much ice south of the Bering Strait. In summertime, the model has slightly too little sea-ice concentration almost

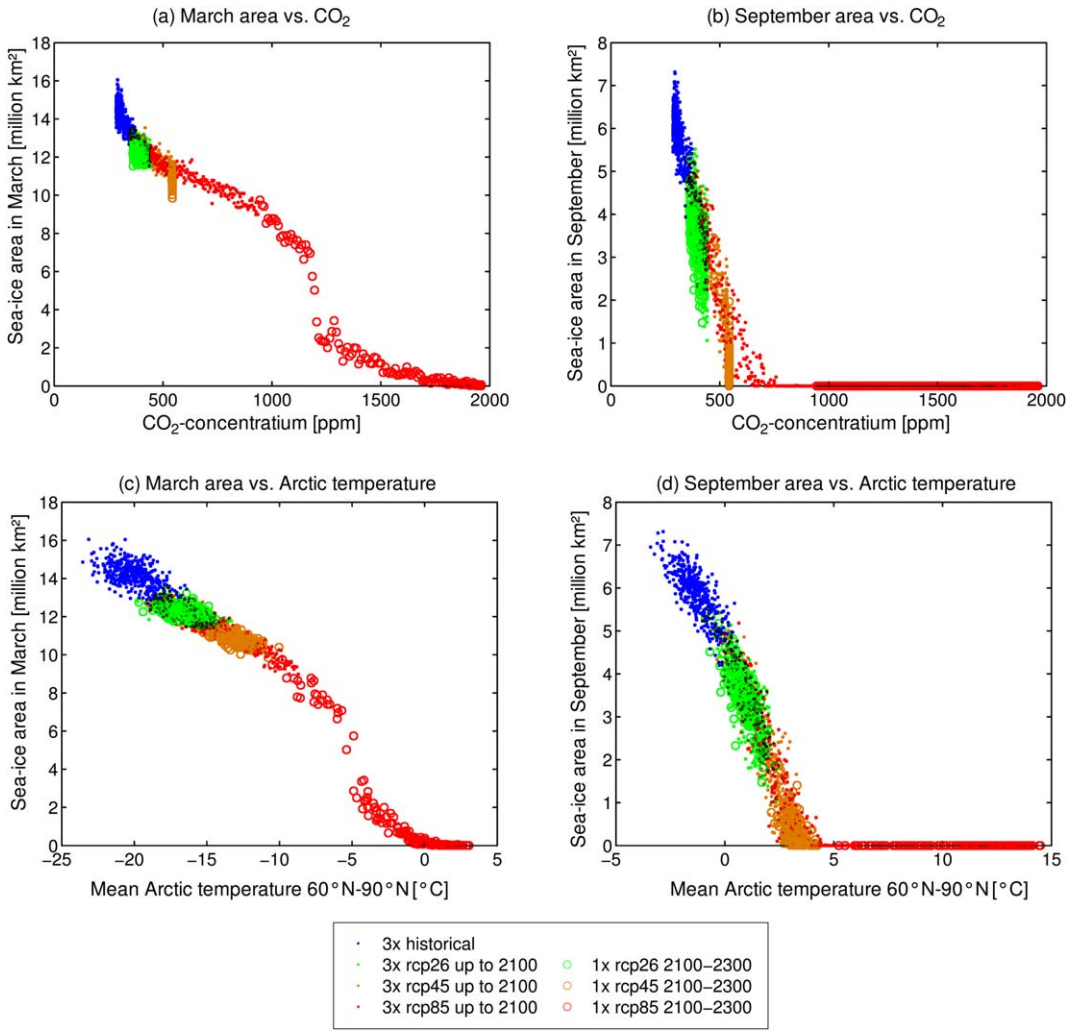

Figure 14. Sea-ice area in (a, c) September and (b, d) March as a function of (a, b) global CO_2 concentration and (c, d) mean surface temperature north of 60°N. Data are plotted for the number of ensemble members indicated in the legend.

everywhere; too much ice is simulated only in the East Siberian Sea.

[74] The commonly used practice of focusing on sea-ice extent in model-data intercomparison can result in misleading results: For example, the sea-ice extent as modeled by MPI-ESM is much closer to the observational estimate of the Bootstrap algorithm than is the one modeled by ECHAM5/MPIOM, whereas in terms of the geophysically more relevant sea-ice area, the simulations of ECHAM5/MPIOM agree better than MPI-ESM with the observational estimate of the Bootstrap algorithm.

[75] Despite being desirable, a primary focus on sea-ice area in model-data intercomparison studies is currently hindered by the uncertainty related to the estimated sea-ice concentration from satellite remote sensing. The NASA Team and the Bootstrap algorithms that were compared for this study show regional differences that in summer can be larger than the differences to MPI-ESM model simulations. A more reliable estimate of sea-ice concentration, including a reliable error estimate, is therefore highly desirable.

[76] Large internal variability makes it impossible to directly judge the quality of a single sea-ice simulation from a direct comparison of trends between observations and simulation. For the same external forcing but slightly different initial conditions, MPI-ESM simulates sea-ice trends during the satellite period that range from more strongly negative than has been observed to positive. This also shows that 30 year-long periods of increasing sea-ice coverage can occur despite a warming of the background climate.

[77] The modeled sea-ice volume for the past few years agrees well between MPI-ESM and the PIOMAS reanalysis. However, in MPI-ESM an Arctic-wide decrease in sea-ice volume started already in the 1920s, far earlier than the decrease shown by the PIOMAS reanalysis. Analyzing this behavior in more detail, we find from reanalysis-forced simulations with MPI-ESM's ocean model MPIOM that the magnitude of the

sea-ice volume decline crucially depends on the choice of the atmospheric reanalysis used to force the model. Using the NCEP reanalysis, on which also PIOMAS is based, we find the same decay rate in summer sea-ice volume as the PIOMAS reanalysis. However, forcing MPIOM with ERA-Interim reanalysis, we find a 40 % smaller decay rate. This dependence of the modeled evolution of sea-ice volume on the particular choice of the atmospheric reanalysis indicates that a reliable estimate of sea-ice volume trends is currently not possible.

[78] The MPI-ESM modeled sea-ice drift patterns agree well with Polar Pathfinder observations. MPI-ESM has, however, on average too high a drift speed, in particular in regions of very high drift speed.

[79] Because MPI-ESM does a reasonable job in simulating the observed variability of Arctic sea ice, we can use the pre-industrial model simulations to estimate the likelihood that the observed trends and extreme minima in sea-ice extent could be caused by internal variability. We find that internal variability can be excluded for both parameters for the observed evolution of summer sea ice. The observed evolution of winter sea ice could still be caused by internal variability.

[80] Examining the dependence of the simulated sea-ice cover on external forcing, we find a significant delay in sea-ice response to increasing CO_2 concentration. The Arctic can become ice free at a CO_2 concentration of around 500 ppm. We also find a sudden drop in winter sea-ice area once the mean Arctic winter temperature becomes warmer than $-5°C$. This drop in winter sea-ice area goes along with a sudden increase in convective precipitation.

[81] The comparison of the modeled evolution with observations and reanalysis provided here has convinced us that the sea-ice cover as modeled by MPI-ESM is reasonably close to the observed one. As such, we see this paper as a starting point for more detailed scientific studies based on MPI-ESM that will help us to better understand the past and future evolution of Arctic sea ice.

[82] **Acknowledgments.** Thanks go to Nils Fischer and to two anonymous referees for very helpful comments on the manuscript. We are grateful to Holger Pohlmann for providing the MPIOM ERA-Interim forced simulations. The simulations described here were carried out at the German Climate Computing Centre (DKRZ). Funding for this study was provided through a Max-Planck-Research Fellowship.

References

Abbot, D. S., and E. Tziperman (2008), Sea ice, high-latitude convection, and equable climates, *Geophys. Res. Lett.*, *35*, L03702, doi:200810.1029/2007GL032286.

Bitz, C., and G. Roe (2004), A mechanism for the high rate of sea ice thinning in the Arctic Ocean, *J. Clim.*, *17*(18), 3623–3632.

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, doi:198410.1029/JD089iD04p05355.

Comiso, J. C. (1986), Characteristics of arctic winter sea ice from satellite multispectral microwave observations, *J. Geophys. Res.*, *91*(C1), 975–994, doi:198610.1029/JC091iC01p00975.

Comiso, J. C., D. J. Cavalieri, C. L. Parkinson, and P. Gloersen (1997), Passive microwave algorithms for sea ice concentration: A comparison of two techniques, *Remote Sens. Environ.*, *60*(3), 357–384, doi:10.1016/S0034-4257(96)00220-9.

Eisenman, I. (2010), Geographic muting of changes in the arctic sea ice cover, *Geophys. Res. Lett.*, *371*, L16501, doi:10.1029/2010GL043741.

Fetterer, F., K. Knowles, W. Meier, and M. Savoie (2002, updated 2012), Sea Ice Index*, Digital Media, National Snow and Ice Data Center*, Boulder, Colo.

Fowler, C. (2003, updated 2011), Polar Pathfinder Daily 25 km EASE-Grid Sea Ice Motion Vectors*, Digital Media*, National Snow and Ice Data Center, Boulder, Colo.

Giorgetta, M., et al. (2013), Climate change from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project 5, *J. Adv. Model. Earth Syst.*, in press.

Goosse, H., O. Arzel, C. M. Bitz, A. de Montety, and M. Vancoppenolle (2009), Increased variability of the Arctic summer ice extent in a warmer climate, *Geophys. Res. Lett.*, *36*, L23702, doi:10.1029/2009GL040546.

Hagemann, S., A. Loew, and A. Andersson (2012), Combined evaluation of MPI-ESM land surface water and energy fluxes, *J. Adv. Model. Earth Syst.*, doi:10.1029/2012MS000173, in press.

Hibler, W. D. (1979), A dynamic thermodynamic sea ice model, *J. Phys. Oceanogr.*, *9*(4), 815–846.

Holland, M. M., D. A. Bailey, and S. Vavrus (2011), Inherent sea ice predictability in the rapidly changing Arctic environment of the Community Climate System Model, version 3, *Clim. Dyn.*, *36*(7–8), 1239–1253, doi:10.1007/s00382-010-0792-4.

Hunke, E. C., W. H. Lipscomb, and A. K. Turner (2010), Sea-ice models for climate study: retrospective and new directions, *J. Glaciol.*, *56*(200), 1162–1172.

Hunke, E. C., D. Notz, A. K. Turner, and M. Vancoppenolle (2011), The multiphase physics of sea ice: A review for model developers, *Cryosphere*, *5*(4), 989–1009.

Jungclaus, J., et al. (2013), Characteristic of the ocean simulations in MPIOM, the ocean component of the MPI-Earth System Model, *J. Adv. Model. Earth Syst.*, in press.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-Year reanalysis project., *Bull. Am. Meteorol. Soc.*, *77*, 437–472.

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.

Koldunov, N. V., D. Stammer, and J. Marotzke (2010), Present-day Arctic sea ice variability in the coupled ECHAM5/MPI-OM model, *J. Clim.*, *23*(10), 2520–2543.

Kwok, R., A. Schweiger, D. A. Rothrock, S. Pang, and C. Kottmeier (1998), Sea ice motion from satellite passive microwave imagery assessed with ERS SAR and buoy motions, *J. Geophys. Res.*, *103*(C4), 8191–8214, doi:199810.1029/97JC03334.

Li, C., D. Notz, S. Tietsche, and J. Marotzke (2013), The transient versus the equilibrium response of sea ice to global warming, *J. Clim.*, in press.

Maier-Reimer, E. (1993), Geochemical cycles in an ocean general circulation model: Preindustrial tracer distributions, *Glob. Biogeochem. Cycles*, *7*, 645–677, doi:10.1029/93GB01355.

Marsland, S., H. Haak, J. Jungclaus, M. Latif, and F. Rske (2003), The Max-Planck-Institute global ocean/sea ice model with orthogonal curvilinear coordinates, *Ocean Modell.*, *5*(2), 91–127, doi:10.1016/S1463-5003(02)00015-X.

Mauritsen, T., et al. (2012), Tuning the climate of a global model, *J. Adv. Model. Earth Syst.*, *4*, M00A01, doi:10.1029/2012MS000154.

Meehl, G., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. Mitchell, R. Stouffer, and K. Taylor (2007), The WCRP CMIP3 multi-model dataset: A new era in climate change research, *Bull. Am. Meteorol. Soc.*, *88*, 1383–1394.

Meier, W., F. Fetterer, M. Savoie, S. Mallory, R. Duerr, and J. Stroeve (2011), NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration*, Digital Media*, National Snow and Ice Data Center, Boulder, Colo.

Meier, W. N., J. Stroeve, A. Barrett, and F. Fetterer (2012), A simple approach to providing a more consistent Arctic sea ice extent time series from the 1950s to present, *Cryosphere*, *6*(6), 1359–1368, doi:10.5194/tc-6-1359-2012.

Meinshausen, M., et al. (2011), The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, *Clim. Change*, *109*(1), 213–241, doi:10.1007/s10584-011-0156-z.

Notz, D. (2009), The future of ice sheets and sea ice: Between reversible retreat and unstoppable loss, *Proc. Natl. Acad. Sci. U. S. A.*, *106*(49), 20,590–20,595, doi:10.1073/pnas.0902356106.

Notz, D., M. G. McPhee, M. G. Worster, G. Maykut, K. H. Schlünzen, and H. Eicken (2003), Impact of underwater-ice evolution on Arctic summer sea ice, *J. Geophys. Res.*, *108*(C3), doi:10.1029/2001JC001173.

Parkinson, C. L., and W. M. Washington (1979), A large-scale numerical model of sea ice, *J. Geophys. Res.*, *84*(C1), 311–337.

Pedersen, C. A., E. Roeckner, M. Lüthje, and J. Winther (2009), A new sea ice albedo scheme including melt ponds for ECHAM5 general circulation model, *J. Geophys. Res.*, *114*, D08101, doi:10.1029/2008JD010440.

Reick, C., T. Raddatz, V. Brovkin, and V. Gayler (2013), The representation of natural and anthropogenic land cover change in MPI-ESM, *J. Adv. Model. Earth Syst.*, in press.

Roeckner, E., T. Mauritsen, M. Esch, and R. Brokopf (2012), Impact of melt ponds on Arctic sea ice in past and future climates as simulated by MPI-ESM, *J. Adv. Model. Earth Syst.*, *4*, M00A02, doi:10.1029/2012MS000157.

Schmidt, G., C. Bitz, U. Mikolajewicz, and L. Tremblay (2004), Ice-ocean boundary conditions for coupled models, *Ocean Model.*, *7*(1–2), 59–74.

Schwegmann, S., C. Haas, C. Fowler, and R. Gerdes (2011), A comparison of satellite-derived sea-ice motion with drifting-buoy data in the Weddell Sea, Antarctica, *Ann. Glaciol.*, *52*(57), 103–110.

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.

Semtner, A. J. (1976), A model for the thermodynamic growth of sea ice in numerical investigations of climate, *J. Phys. Oceanogr.*, *6*(3), 379–389.

Semtner, A. J. (1984), On modelling the seasonal thermodynamic cycle of sea ice in studies of climatic change, *Clim. Change*, *6*(1), 27–37.

Simmons, A., S. Uppala, D. Dee, and S. Kobayashi (2007), ERA-Interim: New ECMWF reanalysis products from 1989 onwards, *ECMWF Newsl.*, *110*, 25–35.

Stevens B., et al. (2013), The atmospheric component of the MPI-M earth system model: ECHAM6, *J. Adv. Model. Earth Syst.*, doi:10.1002/jame.20015, in press.

Stroeve, J., M. M. Holland, W. Meier, T. Scambos, and M. Serreze (2007), Arctic sea ice decline: Faster than forecast, *Geophys. Res. Lett.*, *34*, L09501, doi:10.1029/2007GL029703.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *Bull. Am. Meteorol. Soc.*, *93*(4), doi:10.1175/BAMS-D-11-00094.1.

Valcke, S., A. Caubel, R. Vogelsang, and D. D. (Eds.) (2012), The OASIS3 coupler: a European climate modelling community software, *Geosci. Model Dev. Discuss.*, *5*, 2139–2178, doi:10.5194/gmdd-5-2139-2012.

Winton, M. (2006), Does the Arctic sea ice have a tipping point?, *Geophys. Res. Lett.*, *33*, L23504, doi:10.1029/2006GL028017.

Zhang, J., and D. Rothrock (2003), Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates, *Mon. Weather Rev.*, *131*(5), 845–861.

Corresponding author: D. Notz, Max Planck Institute for Meteorology, Bundesstr. 53, D-20146 Hamburg, Germany. (dirk.notz@zmaw.de)