The Arctic Ocean Observation Operator for 6.9 GHz (ARC3O) - Part 2: Development and evaluation

Clara Burgard[1,2], Dirk Notz[1,3], Leif T. Pedersen[4], and Rasmus T. Tonboe[5]

[1]Max Planck Institute for Meteorology, Hamburg, Germany
[2]International Max Planck Research School for Earth System Modelling, Hamburg, Germany
[3]Institute of Oceanography, Center for Earth System Research and Sustainability, Universität Hamburg, Hamburg, Germany
[4]National Space Institute, Technical University of Denmark, Lyngby, Denmark
[5]Danish Meteorological Institute, Copenhagen, Denmark

Correspondence: Clara Burgard (clara.burgard@mpimet.mpg.de)

Abstract. The observational uncertainty in sea-ice-concentration estimates from remotely-sensed passive-microwave brightness temperatures is a challenge for reliable climate model evaluation and initialization. To address this challenge, we introduce a new tool: the Arctic Ocean Observation Operator (ARC3O). ARC3O allows us to simulate brightness temperatures at 6.9 GHz at vertical polarisation from standard output of an Earth System Model. We evaluate ARC3O by simulating brightness tem-

5 peratures based on three assimilation runs of the MPI Earth System Model (MPI-ESM) assimilated with three different sea-ice concentration products. We then compare these three sets of simulated brightness temperatures to brightness temperatures measured by the Advanced Microwave Scanning Radiometer Earth Observing System (AMSR-E) from space. We find that they differ up to 10 K in the period between October and June, depending on the region and the assimilation run. However, we show that these discrepancies between simulated and observed brightness temperature can be mainly attributed to the underlying

10 observational uncertainty in sea-ice concentration and, to a lesser extent, to the data assimilation process, rather than to biases in ARC3O itself. In summer, the discrepancies between simulated and observed brightness temperatures are larger than in winter and locally reach up to 20 K. This is caused by the very large observational uncertainty in summer sea-ice concentration but also by the melt-pond parametrization in MPI-ESM, which is not necessarily realistic. ARC3O is therefore capable to realistically translate the simulated Arctic Ocean climate state into one observable quantity for a more comprehensive climate

15 model evaluation and initialization.

1 Introduction

The diversity in sea-ice concentration observational estimates affects our understanding of past and future sea-ice evolution as it inhibits reliable climate model evaluation (Notz et al., 2013) and initialization (Bunzel et al., 2016). It also limits our ability to fully exploit relationships between the evolution of sea ice and other climate variables, such as global-mean surface

20 temperature (Niederdrenk and Notz, 2018) and CO_2 emissions (Notz and Stroeve, 2016). To address these issues, we construct an observation operator for the Arctic Ocean at the frequency of 6.9 GHz. This operator provides an alternative approach for climate model evaluation and initialization with satellite observations.

Sea-ice concentration observational estimates are derived from passive microwave brightness temperature measurements from satellites. Some of the uncertainty in these estimates can be induced by the physical noise at the level of the satellite. Most of the uncertainty is however introduced during the interpretation of the measurements because we lack simultaneous observations of the relevant climate variables having an influence on the radiation. The contribution of the individual drivers of the

5 brightness temperature cannot be disentangled unambiguously when the full climate information, such as surface temperature, sea-ice concentration, sea-ice thickness, and snow properties, is not provided. Nevertheless, a variety of algorithms have been developed to retrieve an estimate of sea-ice concentration from these brightness temperatures. As these retrieval algorithms use different combinations of measurements at various frequencies and polarizations, they result in a range of sea-ice concentration products, which differ, sometimes substantially. (Ivanova et al., 2014; Kern et al., 2019). The evaluation of simulated sea-ice

10 concentration in General Circulation Models (GCMs) is therefore influenced by the choice of the sea-ice concentration product against which a simulation is evaluated (Notz et al., 2013).

Observation operators applied to general circulation model (GCM) output have been suggested as a solution to circumvent this observational uncertainty for other climate variables (Flato et al., 2013; Eyring et al., 2019). An observation operator enables us to simulate brightness temperatures based on output from a GCM. This simulated brightness temperature can then

15 be evaluated against the observed brightness temperature. Uncertainty in the evaluation can therefore only be induced by uncertainties in the observation operator and remaining physical noise. We argue that the simulated brightness temperature based on a GCM is a more consistent method less prone to uncertainty than the use of retrieval algorithms because the GCM provides an internally consistent climate state over time and space. Additionally, the climate system as a whole can be evaluated with this approach and not only individual variables.

20 However, the simulation of sea-ice brightness temperatures relies on sea-ice properties not explicitly resolved in most GCMs. In particular, brightness temperatures are driven by the vertical liquid water (or brine) distribution inside the ice and snow, which is driven by temperature and salinity profiles. However, Burgard et al. (2019) showed in a one-dimensional idealized setup that, using a few simple assumptions, the low complexity of GCM output is sufficient to simulate reasonable sea-ice brightness temperatures at 6.9 GHz at vertical polarization. While we focus on the frequency of 6.9 GHz in this study, the

25 framework proposed here, together with the framework presented in Burgard et al. (2019), can be extended to investigate the simulation of brightness temperatures at other frequencies in the future as well, while considering the different influences of snow and atmosphere on the brightness temperature depending on the frequency.

In this study, we first present an Arctic Ocean observation operator that we construct based on the suggestions from Burgard et al. (2019). We then evaluate the brightness temperatures simulated based on assimilation runs against brightness temperatures

30 observed by satellites and investigate potential uncertainty sources in the brightness temperature simulation.

2 The Max Planck Institute Earth System Model

As a baseline for the development of an Arctic Ocean observation operator, we use the Max Planck Institute Earth System Model (MPI-ESM). It is a state-of-the-art Earth System Model that contributed to the Coupled Model Intercomparison Project

in its fifth phase (Taylor et al., 2012) and will contribute to its sixth phase (Eyring et al., 2016). We use its low resolution configuration (MPI-ESM-LR).

The atmosphere component, ECHAM6 (Stevens et al., 2013), has a horizontal resolution of T63 (~1.9° x 1.9°) and a vertical division into 47 levels between surface and 0.01 hPa. The ocean component, MPIOM (Jungclaus et al., 2013), is based on a

5 curvilinear grid with two poles located in South Greenland and Antarctica. The horizontal resolution ranges from 15 km near Greenland to 185 km in the tropical Pacific. Vertically, the ocean is divided into 40 levels between surface and bottom. The sea ice is simulated within MPIOM by a dynamic/thermodynamic sea-ice model based on Hibler (1979). In this simple setup, the sea-ice salinity is kept constant at 5 g/kg, and the ice bottom temperature is kept constant at -1.8 °C. There is no explicit simulation of the ice thickness distribution. Still, the simulation of the mean state and variability of Arctic sea ice is realistic

10 (Notz et al., 2013).

For our observation operator, we use output from the atmosphere component ECHAM6. The sea-ice properties, such as sea-ice concentration, sea-ice thickness and snow thickness, are computed within the ocean component and communicated to the atmosphere component through coupling on a daily frequency (Jungclaus et al., 2013). Based on these properties, ECHAM6 computes the snow cover fraction and the melt pond coverage (Giorgetta et al., 2013), which are needed for a comprehensive

15 assessment of the radiative properties of the surface. Additionally, ECHAM6 provides the atmospheric water and ice content, which are needed for the calculation of the radiation path through the atmosphere (see Sec. 3.2). ECHAM6 therefore provides all variables needed for the simulation of Arctic Ocean brightness temperatures.

3   The Arctic Ocean Observation Operator ARC3O

The purpose of the ARCtic Ocean Observation Operator for 6.9 GHz (ARC3O) is to simulate Arctic Ocean brightness tem-

20 peratures as could be seen at the top of the atmosphere if a theoretical satellite could measure this radiation. This brightness temperature is a result of radiation emitted by the surface, upwelling atmospheric radiation, reflected downwelling atmospheric radiation, atmospheric transmission, and reflected space radiation (Swift and Cavalieri, 1985).

As a consequence, ARC3O is based on two parts. In the first part, an emission model computes the sea-ice surface brightness temperature (see Sec. 3.1). In the second part, an atmospheric radiative transfer model combines the sea-ice surface emission

25 with ocean emission and atmospheric emission, reflection and transmission (see Sec. 3.2). The workflow of ARC3O follows five steps (see Fig. 1), which we explain in the following.

3.1   The contribution of the sea-ice surface to the brightness temperature

The brightness temperature TB_{ice} emitted at an ocean surface completely covered by sea ice at 6.9 GHz, vertical polarization, is defined as:

30 $$\text{TB}_{\text{ice}} = \epsilon_{\text{eff,ice}} \cdot T_{\text{eff,ice}} \tag{1}$$
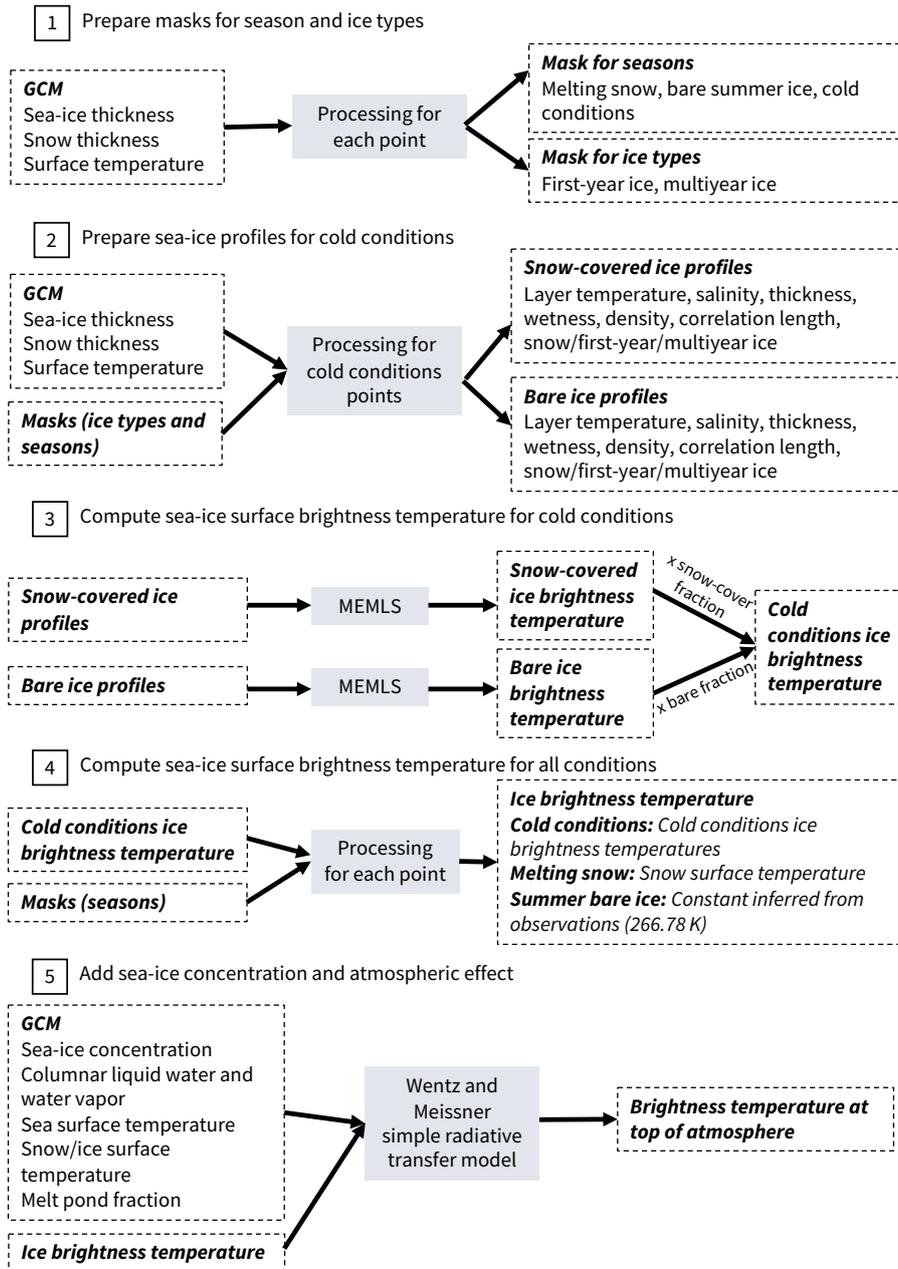
Figure 1. Workflow of the Arctic Ocean Observation Operator ARC3O.

where $\epsilon_{\mathrm{eff,ice}}$ is the emissivity of the emitting part of the ice, i.e. the layers influencing the resulting radiation emitted at the surface and $T_{\mathrm{eff,ice}}$ the integrated temperature over this same emitting part (Hallikainen and Winebrenner, 1992; Shokr and

Sinha, 2015; Tonboe, 2010). The sea-ice emissivity, and therefore the sea-ice surface brightness temperature, is primarily driven by the vertical distribution of the brine volume fraction inside the ice, which principally depends on the temperature and salinity profile (see Burgard et al. (2019) for more details). However, MPI-ESM does not provide temperature and salinity profiles. To circumvent this lack of information, we follow the suggestion of Burgard et al. (2019) to build a simple model for

5   these sea-ice properties, based on the boundary conditions given by the GCM.

3.1.1   Identifying different periods and ice types

Burgard et al. (2019) showed that the simulation of sea-ice surface brightness temperatures relies on different assumptions, depending on the conditions of the ice. A sea-ice year can be divided into three periods: cold conditions, melting snow, and bare ice near 0 °C. Additionally, sea-ice brightness temperatures depend on the ice type, for example first-year or multiyear

10   ice.

We therefore flag the different type of periods and different ice types based on the sea-ice properties given by the MPI-ESM output (Step 1 in Fig 1). Grid cells containing melting snow are flagged as "melting snow periods", grid cells containing bare ice in July, August and September are flagged as "bare ice near 0°C", and the remaining grid cells are flagged as "cold conditions". To flag the different grid cells as "first-year ice", "multiyear ice" and "open water only", we consider the ice thickness evolution.

15   If the ice thickness is zero, the ice type is set to "open water only", if the ice thickness is larger than zero but there has been at least one "open water only" timestep in the year preceding the timestep evaluated, the ice type is set to "first-year ice". If none of the two before apply, the ice type is set to "multiyear ice". This is a simplification, assuming that there is no ice drift and that the ice present at one point in time and space will be the same, but older, ice later in time.

3.1.2   Cold conditions

20   In periods of cold conditions, Burgard et al. (2019) showed that the sea-ice surface brightness temperature can be simulated with similar low uncertainty using a linear vertical temperature profile and a function of depth for the salinity interpolated to five, seven or ten layers. In this study, we choose to use 10 layers. We construct profiles (Step 2 in Fig 1), divided into eleven layers, namely ten layers of ice and one layer of snow. The ice layers are equidistant, based on the ice thickness given by MPI-ESM, and the snow layer thickness is equal to the snow thickness given by MPI-ESM.

25   We construct temperature profiles based on the ice surface temperature given by MPI-ESM, which represents the temperature at the top of the snow and ice column. For each grid cell, we construct two sets of profiles. One set of profiles interprets the surface temperature as the snow surface temperature. This profile is a combination of two linear profiles, one in the snow, defined by the snow thermal conductivity, and one in the ice, defined by the ice thermal conductivity (see formula in App. A of the Supp. Info.). The other set of profiles interprets the surface temperature as the ice surface temperature and is a linear profile

30   between surface and bottom temperature. The ice bottom temperature is taken as constant at -1.8 °C in both cases.

The salinity is taken as a function of depth z, as formulated by Griewank and Notz (2015). The salinity is defined as follows for first-year ice:

$$S_{fy}(z) = \frac{z}{a + bz} + c \tag{2}$$

with a = 1.0964, b = -1.0552 and c = 4.41272

and as follows for multiyear ice:

$$S_{my}(z) = \frac{z}{a} + (\frac{z}{b})^{1/c} \tag{3}$$

with a = 0.17083, b = 0.92762 and c = 0.024516.

We set the snow salinity to zero. Note, however, that the validity of this assumption is slightly uncertain as the lowest layer of the snow can be saline, especially above first-year ice (Barber et al., 1998; Shokr and Sinha, 2015; Nandan et al., 2017), enabling the presence of liquid water at the base of the snow.

The vertical profile of the ice density ρ_i is computed based on the temperature and salinity profiles, with the following formula applied to each ice layer (Notz, 2005):

$$\rho_i = \Phi_b \cdot \rho_w + (1 - \Phi_b) \cdot \rho_0 \tag{4}$$

with the pure ice density $\rho_0 = 916.18 - 0.1403T$, the brine salinity $S_b = -17.6T - 0.389T^2 - 0.00362T^3$, the density of seawater $\rho_w = 1000.3 + 0.78237S_b + 2.8008 \cdot 10^{-4}S_b^2$, and the brine volume fraction $\Phi_b = S/S_b$. The density of the snow layer is set to 300 kg/m^3, like in MPI-ESM (Giorgetta et al., 2013).

The vertical profile of the correlation length, a measure for the scatterer size (snow particles, brine inclusions, air bubbles), depends on the ice type. If the ice layer is in the upper 20 cm of first-year ice, the correlation length is set to 0.35 mm. If the ice layer is located below the upper 20 cm, the correlation length is set to 0.25 mm (Tonboe, 2010). For multi-year ice, the correlation length is set to 1.5 mm for all ice layers (Burgard et al., 2019). The correlation length of the snow layer is set to 0.15 mm (Tonboe, 2010).

The sea-ice surface brightness temperature is simulated based on the temperature, salinity, density, thickness, and correlation length profiles described above. A slightly modified version of the Microwave Emission Model for Layered Snowpacks (MEMLS, Wiesmann and Mätzler, 1999) extended for sea ice (Tonboe et al., 2006) is used for the brightness temperature simulation. It relates the snow and ice properties to emission, absorption and scattering of the microwave radiation in each layer. Hence, MEMLS simulates the path of the radiation through the ice and snow from bottom to top, resulting in a brightness temperature emitted at the surface.

MPI-ESM provides a snow cover fraction, which means that the ice is not always fully covered by snow. To account for the effect of both snow-covered ice and bare ice on the radiation, we simulate two sea-ice surface brightness temperatures for each grid cell. One set of brightness temperatures is simulated using the temperature profiles computed through snow and ice, where the surface temperature given by MPI-ESM is interpreted as the snow surface temperature. The other set is simulated using linear temperature profiles computed through ice only, assuming that there is no snow cover on the ice and the

surface temperature given by MPI-ESM is interpreted as the ice surface temperature (Step 3 in Fig 1). These surface brightness temperatures are then combined, weighted by the snow cover fraction given by MPI-ESM, resulting in one mean sea-ice surface brightness temperature.

3.1.3 Melting snow

5   In spring, temperatures increase across the Arctic Ocean, leading to the melting of the snow covering the sea ice. Wet snow strongly affects the emitted microwave radiation. This effect mainly depends on the water content of the snow, on the density and on the size and form of the snow particles (Chang and Gloersen, 1975; Ulaby et al., 1986; Shokr and Sinha, 2015). Already low liquid water fractions inside the snow lead to a very high emissivity near 1 at 6.9 GHz (Hallikainen et al., 1986). As we do not have detailed information about the snow properties in MPI-ESM to compute the wet snow brightness temperature more

10 precisely, we assume that the melting snow emissivity is 1 and therefore the brightness temperature of ice covered by melting snow is equal to the temperature of the snow surface (see Eq. 1).

3.1.4 Summer bare ice near 0 °C

In summer, after the snow has fully melted away, the salinity profile inside the ice cannot necessarily be represented by a simple function of depth. As the near-surface ice can be assumed to be isothermal close to 0°C during summer, the brine

15 volume fraction at the surface, the upper few cm of the ice, increases and melt ponds form. Above an ice surface brine volume fraction of 0.2, the brightness temperature is proportional to the ice surface brine volume fraction (Burgard et al., 2019). The ice surface brine volume fractions above 0.2 can be interpreted as a measure for the melt-pond fraction as they mean that the surface is very wet. For these warm conditions, the conditions for the ice which is not covered by melt ponds are similar over the whole Arctic Ocean. Burgard et al. (2019) therefore suggest a very simple approach: use a constant brightness temperature

20 for the ice surface fraction, i.e. the ice fraction not covered by melt ponds.

   To find a surface brightness temperature representing the sea-ice surface in summer, we use the observational dataset Round Robin Data Package (RRDP, Pedersen et al., 2018) developed as part of the European Space Agency (ESA) sea-ice Climate Change Initiative (SICCI). These data cover the period from May to mid-August 2011. The RRDP contains amongst others microwave brightness temperatures between 6 and 89 GHz measured by the Advanced Microwave Scanning Radiometer 2

25 (AMSR2) collocated with the melt-pond fraction product by Istomina et al. (2015b) over areas estimated to be close to 100% sea-ice concentration. Using the combination of the melt-pond fraction and the observed brightness temperatures, we can infer the summer brightness temperature of melt-pond-free sea ice (Fig. 2). We do so by taking the mean brightness temperature for melt-pond fractions between zero and 0.2, as Burgard et al. (2019) showed that above a brine volume fraction of 0.2 in the upper few cm of ice, the brightness temperature is proportional to this ice surface brine volume fraction, and in this case we

30 assume that melt ponds have the same effect. Below 0.2, the effect of the ice on the brightness temperature is stronger than the effect of melt ponds. This method results in a summer ice surface brightness temperature of 262.29±3.56 K.

   The radiative transfer model we use to compute the brightness temperature at the top of the atmosphere (see Sec. 3.2) adds the atmospheric contribution to the mean surface brightness temperature, the combination of ocean and ice emission.
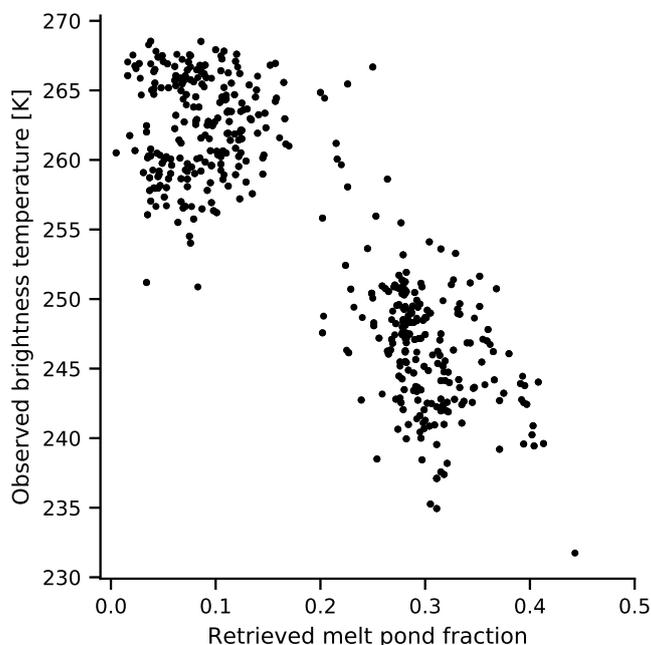
Figure 2. Brightness temperatures measured by AMSR2 against the melt pond fraction product by Istomina et al. (2015b) for different points represented in the Round Robin Data Package from May to mid-August 2011.

Therefore we need the brightness temperature at the ice surface rather than at the top of the atmosphere in the summer case as well. The constant we just inferred is however a measure for both the surface and the atmospheric contribution to the brightness temperature. As an attempt to remove the atmospheric contribution, we infer a mean atmospheric correction by running ARC3O once on MPI-ESM output presented in Sec. 4.2, in regions where the simulated sea-ice concentration is 99% or more in summer and setting all melt pond fractions to zero. We then subtract the constant ice brightness temperature from this top-of-the-atmosphere simulated brightness temperature. This gives us a mean atmospheric effect of 4.49 K, which we add to our inferred brightness temperature of 262.29 K, resulting in a constant brightness temperature of 266.78 K as a constant brightness temperature representing the radiation emitted at the summer bare ice surface. This bare ice surface brightness temperature can then be used in the radiative transfer model to be combined with open water brightness temperatures, representing the melt-pond fraction and open ocean, to obtain the mean surface brightness temperature for summer (see Sec. 3.2).

3.2 The contribution of ocean and atmosphere to the brightness temperature

The sea-ice surface brightness temperature is set for each grid cell depending on the three periods presented above (Step 4 in Fig. 1). As the Arctic Ocean is not covered by 100% of sea ice, the brightness temperature of the surface also depends on the open water surface brightness temperature. Additionally, to simulate the radiation reaching the top of the atmosphere, the

atmospheric contribution has to be added to the surface brightness temperature. We use a geophysical model developed by Wentz and Meissner (2000) to take into account the oceanic and atmospheric contributions (Step 5 in Fig. 1).

The total brightness temperature of an Arctic Ocean grid cell is computed with this model as a combination of the upwelling atmospheric emission, the upwelling surface emission by ocean, sea ice, and melt ponds, the atmospheric transmittance, the

5   atmospheric emission reflected by the different types of surface, and the reflected background radiation from space. The ocean surface brightness temperature is computed as a function of surface temperature, surface salinity, and wind speed. The melt-pond brightness temperature is computed similarly to the ocean brightness temperature, but setting salinity and wind speed to zero. Finally, as the atmosphere is mostly transparent to radiation in the low microwave range, the radiative transfer through the atmosphere is computed based only on the columnar water vapor and columnar cloud liquid water.

10 ## 4   Evaluation of ARC3O

The approach we use to construct ARC3O was proposed by Burgard et al. (2019), based on an idealized one-dimensional setup that did not involve actual observations. In the following, we evaluate our simulated Arctic Ocean brightness temperatures against brightness temperatures measured by satellites.

We do so by comparing brightness temperatures simulated by ARC3O based on MPI-ESM output from assimilation exper-

15 iments, i.e. experiments where the model is regularly nudged towards observations. Hence, we expect the simulated climate system to be close to reality and the simulated brightness temperature to be close to the observed brightness temperature.

However, the observations used in the data assimilation are reanalysis data for the atmosphere and ocean and retrieved sea-ice concentration products for the sea ice. They are therefore not direct observations but already-processed products prone to differences to reality. Additionally, in the assimilation process, MPI-ESM is nudged towards observations but some character-

20 istic features inherent to the mean model state might remain. The uncertainty of the observed brightness temperature itself is considered to be small and thus neglected here. Hence, differences between observed and simulated brightness temperature can arise from three sources: (1) the difference between real and retrieved climate state due to the difference between retrieval algorithms or reanalysis and the real climate state, (2) the difference between the assimilated climate state and the retrieved or reanalysis product, and (3) biases in ARC3O (Fig. 3). In the following, we try to quantify how the first two uncertainty

25 sources contribute to differences between the simulated and observed brightness temperatures. Any remaining biases can then be attributed to biases of ARC3O itself.

4.1   Observation data

As observed brightness temperatures, we use Calibrated Passive Microwave Daily EASE-Grid 2.0 (CETB) brightness temper-atures processed as part of the NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs)

30 program (Brodzik et al., 2016, Updated 2018). They are an improved, enhanced-resolution, gridded passive microwave Earth System Data Record (ESDR) for monitoring cryospheric and hydrologic time series from the measurement devices Scanning Multi-channel Microwave Radiometer (SMMR), Special Sensor Microwave Imager/Sounder (SSM/I-SSMIS) and Advanced
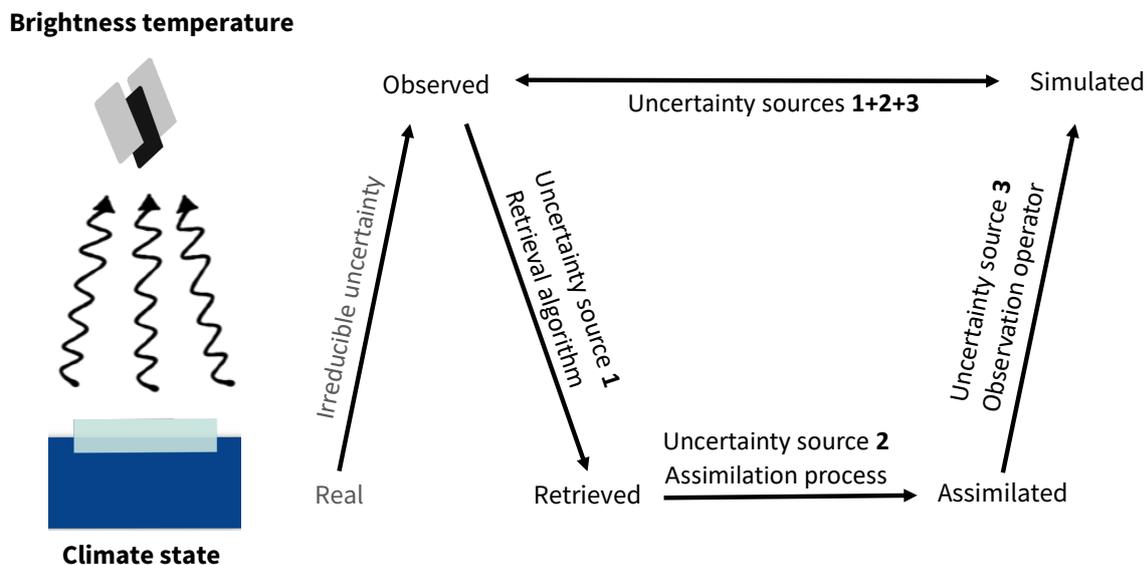
Figure 3. Uncertainty sources possibly introducing differences between simulated and observed brightness temperature.

Microwave Scanning Radiometer - Earth Observing System (AMSR-E). These data cover the period between 1978 and 2017 and are provided on a 25 km x 25 km grid. For the comparison with MPI-ESM data, we focus on the period from 2002 to 2008 and interpolate the observations bilinearly to the model grid (1.9°x1.9°). Again, we concentrate on the frequency of 6.9 GHz, vertical polarization. At this frequency and this time period, the observations stem from AMSR-E.

5 ## 4.2 Model data

We use model data from assimilation runs, as they are nudged towards the observed climate state and are therefore expected to be a reasonable estimate of the real climate state in the model. Differences between simulated and observed brightness temperatures should therefore be small and can be attributed to the three uncertainty sources presented before. To examine the impact of the choice in sea-ice retrieval product, we use three assimilation runs based on three different sea-ice concentration
10 products. The atmosphere and the ocean component are assimilated in the same way in all three cases.

The assimilation experiments cover the period from 2002 to 2008 and were conducted by Bunzel et al. (2016). The assimilation technique used was Newtonian relaxation, also called nudging. Atmospheric, oceanic and sea-ice properties were nudged into the model using full-field data assimilation in all atmospheric and oceanic levels. In the atmosphere, vorticity, divergence, temperature, and surface pressure were nudged into the model with a relaxation time of one day, while salinity and temperature
15 in the ocean were nudged with a relaxation time of ten days. For the assimilation of atmospheric quantities, the ERA-Interim dataset (Dee et al., 2011) was used, while the ocean was nudged toward Ocean Reanalysis System 4 data (Balmaseda et al., 2013).

For sea ice, only sea-ice concentration was assimilated. The three different sea-ice concentration products are the ESA SICCI Version 2 (SICCI2) dataset (Lavergne et al., 2019) as a 50-km-gridded product, and the NASA Team dataset (Cavalieri et al., 1996) and the Bootstrap dataset (Comiso, 2000), both as 25-km-gridded products. We choose these datasets because SICCI2 is a new algorithm combining several existing algorithms with the goal of improving the retrieved sea-ice concentration product,

5 Bootstrap sea-ice concentrations are in the upper range of retrieved sea-ice concentrations, and NASA Team sea-ice concentrations are in the lower range (Ivanova et al., 2014; Kern et al., 2019). The data were interpolated bilinearly to the model grid before assimilation. In grid boxes containing missing values, e.g. the polar observation hole (northward of 87.2°N), no assimilation was applied. The sea ice was then exclusively calculated by the model. To avoid brightness temperature uncertainties due to this free simulation region, we mask out the region northward of 86.72°N, the highest latitude on our grid below the

10 observational hole. The relaxation time was 20 days. Relaxation times differ among the model components to account for the different response times of the components. In order to allow for a realistic relation between ice concentration and thickness, sea-ice thickness was updated in the model proportionally to ice concentration nudging (Tietsche et al., 2013).

4.3 Cold seasons (JFM, AMJ, OND)

4.3.1 Comparison between simulated and observed brightness temperatures

15 The first comparison between simulated and observed brightness temperatures clearly showed a positive bias over the whole Arctic Ocean in the simulated brightness temperatures (Fig. B1, left, in Supp. Info.). The brightness temperature is defined as the product of the emissivity and the physical temperature of the emitting part of the ice (Ulaby et al., 1986). A comparison of the simulated emissivities with emissivities derived from observational data from the RRDP showed that ARC3O systematically overestimates the emissivity. It is however not straightforward to find where the bias is produced in the emission model.

20 We therefore chose to correct the bias by multiplying the inherent sea-ice emissivity by a tuning coefficient at the end of step 3 of the ARC3O workflow (see Fig. 1). The coefficient which yields the best agreement with observations is 0.968 (Fig. B1, right, in Supp. Info.). More information about the tuning process is found in App. B. In the following, we discuss brightness temperatures simulated with this tuning procedure.

The three different sets of simulated brightness temperatures show largely similar behaviours in the cold seasons winter

25 (January/February/March, JFM), spring (April/May/June, AMJ), and autumn (October/November/December, OND) (Fig. 4). Overall, differences between simulated and observed brightness temperatures are very small and are generally lower than 10 K. The pattern of differences appears to be similar across seasons. The simulated brightness temperatures are slightly higher than the observed ones in regions of high sea-ice concentration and thickness, e.g. north of the Canadian Archipelago and the Central Arctic in winter. In contrast, they are lower than the observed ones in regions of low sea-ice concentration and thickness, e.g. in

30 the marginal zones such as the Barents Sea, the Pacific sector, and the Hudson Bay in the SICCI and NASA Team simulations, and higher in the Bootstrap simulation.

The overestimation on the order of 2 to 4 K in the Central Arctic in winter has a similar pattern in all three sets of simulated brightness temperatures. Otherwise, brightness temperatures based on the Bootstrap assimilation run are very close to the
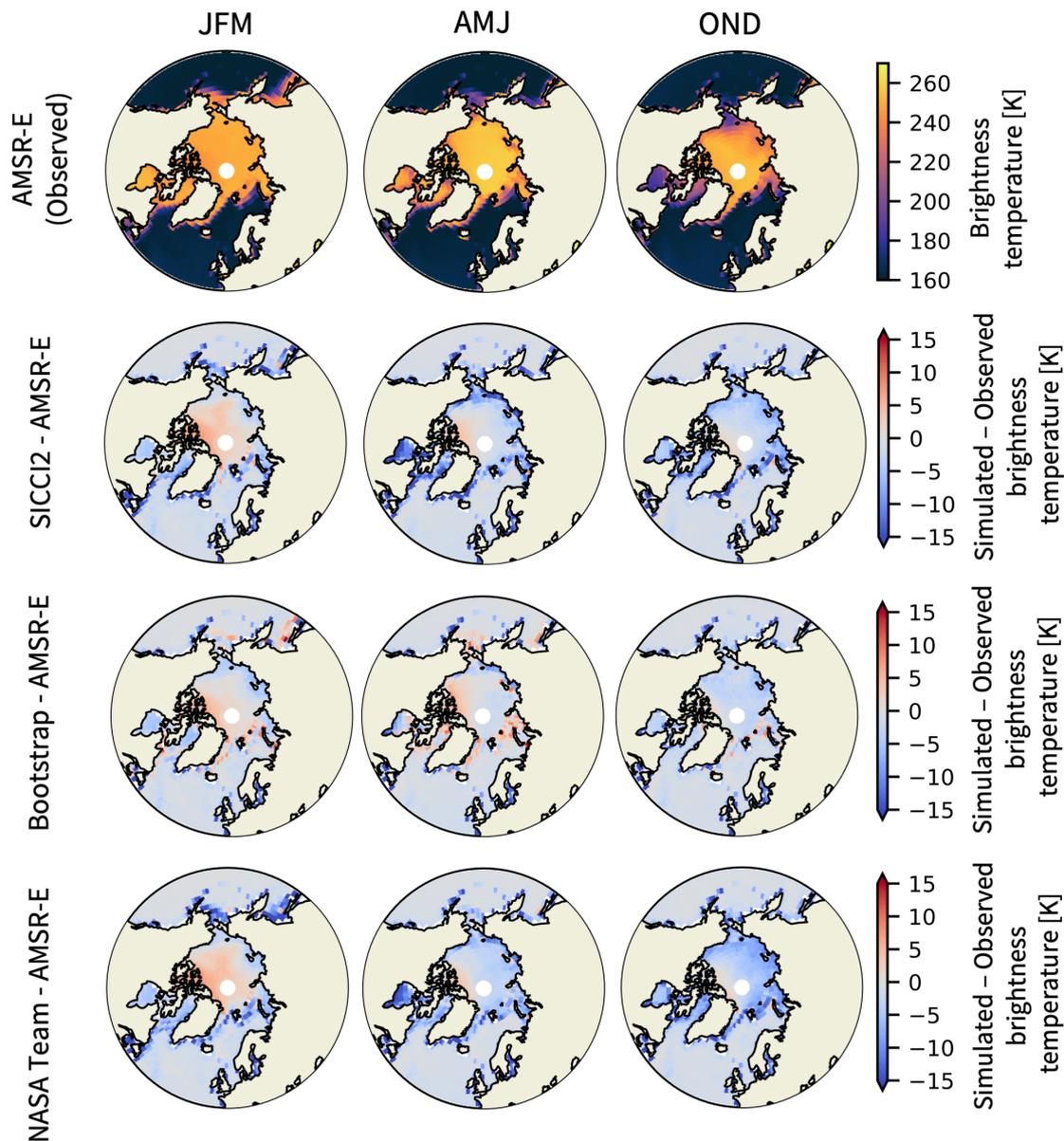
Figure 4. Observed brightness temperatures by AMSR-E (1st row). Differences between brightness temperatures simulated with ARC3O from MPI-ESM output assimilated with SICCI2 (2nd row), Bootstrap (3rd row) and NASA Team (4th row) sea-ice concentration and observed brightness temperatures. The columns stand for the three cold seasons: JFM, AMJ, OND. Summer (JAS) is discussed in Sec. 4.4.

observed ones, with differences to the observations of usually less than 3 K. Only a few individual points in the Atlantic sector show larger biases. Brightness temperatures based on the NASA Team and SICCI2 assimilation run show stronger differences

to observations. The simulated brightness temperatures are up to 10 K lower than the observations in the North Pacific in winter and up to 15 K lower than the observations in the Hudson Bay in spring. In the Central Arctic and the Atlantic Sector, the NASA Team brightness temperatures are 2 to 5 K lower than observations in spring and 5 to 10 K lower than observations in autumn. The pattern of differences between SICCI2 brightness temperatures and observations is similar to the pattern of

5 differences between NASA Team brightness temperatures and observations but the SICCI2 brightness temperatures are about 2 K higher than the NASA Team brightness temperatures.

4.3.2 Investigating uncertainty sources

The total difference Δ_{tot} between simulated and observed brightness temperatures is a consequence of the difference between real and retrieved climate state $\Delta_{\mathrm{retriev}}$, of the difference between retrieved and simulated climate state Δ_{assim}, and of biases in

10 the brightness temperature simulation by ARC3O Δ_{ARC3O} (Fig. 3):

$$\Delta_{\mathrm{tot}} = \Delta_{\mathrm{retriev}} + \Delta_{\mathrm{assim}} + \Delta_{\mathrm{ARC3O}} \tag{5}$$

We set out to investigate Δ_{assim} and $\Delta_{\mathrm{retriev}}$ to gain an estimate of Δ_{ARC3O}.

In a first step, we investigate which drivers the brightness temperature is particularly sensitive to. We concentrate on the variables provided by MPI-ESM, as these are the ones we can quantify in our setup. In the cold seasons, the most important

15 drivers for a the brightness temperature simulation for a surface covered by varying fractions of sea ice and open water are the sea-ice concentration, sea-ice thickness, snow thickness, and surface temperature.

We examine the sensitivity for the month of October, representing the beginning of the freezing period, and for the month of March, representing the end of the freezing period. We use only one of the assimilation runs, the SICCI2 run, as we assume that the physical relationships linking the different variables are the same in all three assimilation runs. For both October and

20 March, we compute, for each grid cell and each variable, the anomaly to the time mean. By taking the standard deviation of these anomalies, we have a representation of the variability of the given variable in the grid cell (Fig. C1 in App. C).

With these variability fields, we conduct sensitivity studies for each variable of interest. For both October and March, we conduct two sets of experiments per variable of interest, one in which we add the variability field and one in which we subtract the variability field. The main message emerging from the results is that the sea-ice concentration variability is the main driver

25 for variations in the brightness temperature (Tab. 1) changing it by up to ≈ 25 K, while the variability in other variables affects the brightness temperature only up to ≈ 3 K. Spatially, the sea-ice concentration is the main driver for variability in regions not completely covered by ice (Fig. 5). In regions covered by near to 100% of ice, the surface temperature has the highest effect on the brightness temperature. Although they both have an indirect influence on the ice surface temperature as well, sea-ice thickness and snow thickness do not play an important role for uncertainties in the total brightness temperature of a grid cell

30 in the simulated variability range as their mean absolute contribution to the brightness temperature variability is on the order of 1 K. To understand the total uncertainty Δ_{tot}, we therefore need to focus on two variables: the sea-ice concentration and the surface temperature.

Table 1. Sensitivity of the simulated brightness temperature to different input variables. We show the 5th and 95th percentile of the difference between modulated and reference brightness temperature and the 5th and 95th percentile of the modulating variable. The latter range is computed as the standard deviation of the anomaly to the time mean in each grid cell. Straight font represents the sensitivity to the increase in the variable, italic font represents the sensitivity to the decrease.

	Range of difference in brightness temperature	Range of variability
March		
Snow thickness	0.02 to 0.98 K	±0.4 to 11 cm
	-1.39 to -0.04 K	
Sea-ice thickness	-0.88 to -0.09 K	±7 to 44 cm
	-0.13±1.02 K	
Sea-ice concentration	0.33 to 17.84 K	±0.05 to 27 %
	-18.65 to -0.56 K	
Surface temperature	0.14 to 3.24 K	±3.16 to 8.00 K
	-3.38±-0.16 K	
October		
Snow thickness	0.06 to 0.43 K	±0.1 to 7 cm
	-0.66 to 0.10 K	
Sea-ice thickness	-0.49 to -0.12 K	±8 to 53 cm
	-0.12±0.54 K	
Sea-ice concentration	1.14 to 16.51 K	±1 to 36 %
	-25.98 to -1.34 K	
Surface temperature	0.80 to 2.00 K	±2.84 to 6.38 K
	-2.03±-0.87 K	

In a next step, we investigate the influence of Δ_{assim} on Δ_{tot}. The goal of a data assimilation is to reach a simulated climate state close to reality. During the data assimilation process, the model is nudged towards three distinct observational datasets: an ocean reanalysis, an atmosphere reanalysis, and a sea-ice concentration product, which are not necessarily consistent with each other. Hence, discrepancies can arise between the variables before and after the assimilation. This is the case for example

5   when a non-zero sea-ice concentration is assimilated at one point but the ocean temperature is too warm to sustain the ice at that point and the ice directly melts away.

As the sea-ice concentration is the main driver for uncertainties in the brightness temperature simulation, we here focus on the effect of the data assimilation on the sea-ice concentration in the three different assimilation runs. This effect is mostly visible in the marginal regions (Fig. 6) and is of similar magnitude for all three sea-ice concentration datasets. At the ice edge,

10 the differences are highest, on the order of 5 %. As a rule of thumb, differences of 1% in sea-ice concentration are equivalent to differences of 1 K in brightness temperatures (see Burgard et al., 2019), so the differences in sea-ice concentration are roughly
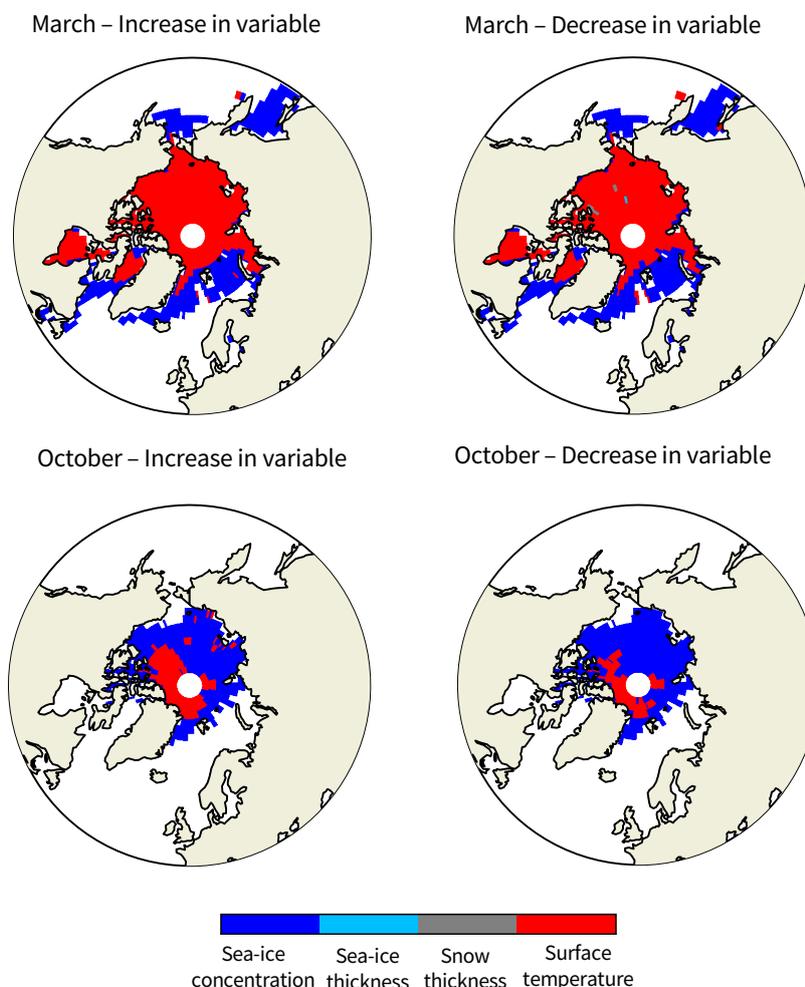
Figure 5. Variable which has the highest absolute mean effect on the brightness temperature in March (top) and October (bottom) when their variability field (one standard deviation) is added to (left) and subtracted from (right) the input variable. Inspired from Fig. 5 in Richter et al. (2018)

equivalent to resulting differences in brightness temperature of around 5 K. Δ_{assim} can therefore account for a large part of the total difference between simulated and observed brightness temperature Δ_{tot} in the ice edge region (see Fig. 4).

Unfortunately, the difference between real and retrieved sea-ice concentration $\Delta_{retriev}$ cannot be as robustly quantified as Δ_{assim}. In-situ observations for a robust evaluation of the retrieved sea-ice concentration products are largely lacking. Although there have been local evaluation approaches, based on regions known to be covered by 100% sea ice (e.g. Tonboe et al., 2016; Lavergne et al., 2019; Ivanova et al., 2015), this lack of in-situ observations inhibits the evaluation of the products on larger scales against reality. The effect of $\Delta_{retriev}$ and Δ_{ARC3O} can therefore not clearly be disentangled.
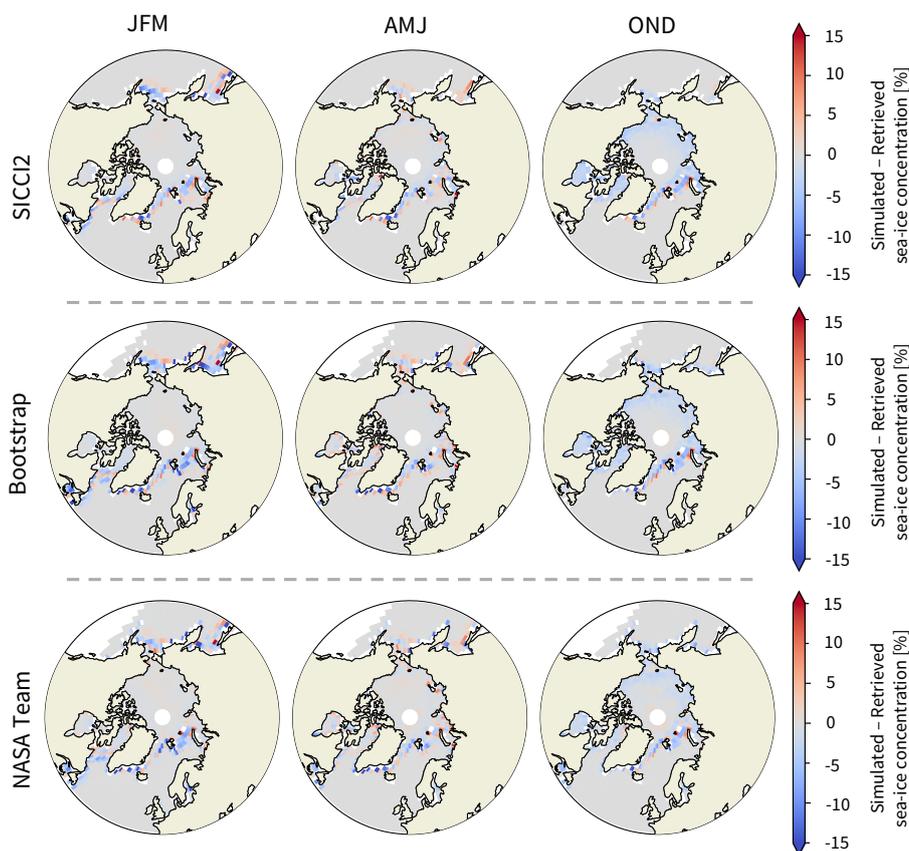
Figure 6. Difference between retrieved and simulated sea-ice concentration for the three assimilation runs.

Still, we can give an estimated range for Δ_{ARC3O} by assuming that the real sea-ice concentration lies between the Bootstrap dataset, which is in the higher range of sea-ice area estimates and NASA Team, which is in the lower range (Fig. 7 and Ivanova et al., 2014; Kern et al., 2019). To estimate Δ_{ARC3O} based on this assumption, we first subtract Δ_{assim} from the simulated brightness temperature. Δ_{tot} is now only a sum of $\Delta_{\mathrm{retriev}}$ and Δ_{ARC3O}. Second, for each grid cell and each time step, we

5 evaluate if the observed brightness temperature is located within the range of the brightness temperatures simulated based on the three different sea-ice concentration products. If yes, differences are not necessarily a bias induced by ARC3O. If not, it is likely that ARC3O induces a bias. In this case, the simulated brightness temperature with the lowest absolute distance from the observed brightness temperature represents the smallest plausible estimate of Δ_{ARC3O} (Fig. 8, 2nd row). The largest absolute difference between simulated and observed brightness temperatures in contrast gives an estimate of the largest plausible value

10 of Δ_{ARC3O} (Fig. 8, 3rd row). If the observed brightness temperature is within the range obtained for the different retrieved sea-ice concentration estimates, the estimate of Δ_{ARC3O} is set to zero.
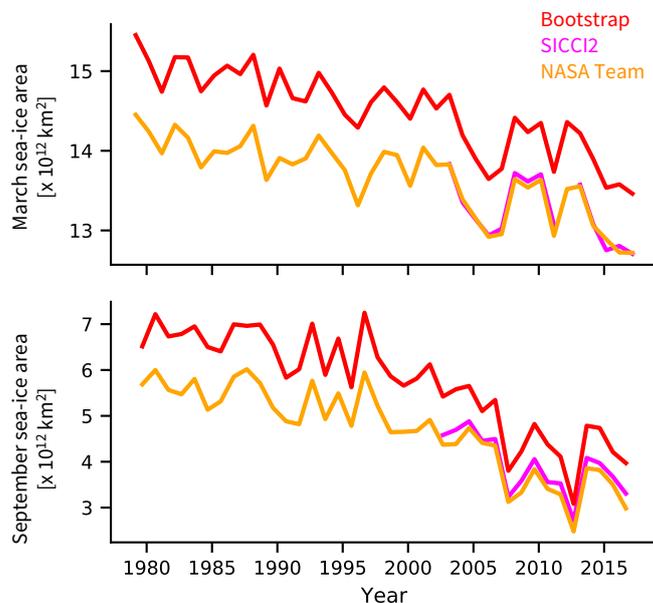
Figure 7. March (top) and September (bottom) sea-ice area for the three observational datasets assimilated in the three assimilation runs used.

The resulting mean estimates of Δ_{ARC3O} are both very small, as the minimal estimates are well below 5 K and the maximal estimates are 5 K or below, except in the Hudson Bay. Additionally, the comparison of the Δ_{ARC3O} estimates to the spread in sea-ice concentration between Bootstrap and NASA Team (Fig. 8, 4th row) shows that biases in ARC3O, i.e. Δ_{ARC3O}, are small compared to the uncertainty in retrievals, i.e. $\Delta_{retriev}$, except in the Central Arctic on the Canadian side.

5   The overestimation on the Canadian side of the Central Arctic is likely due to a MPI-ESM bias in the sea-ice thickness. The sea-ice thickness in our assimilation runs is on the order of 2 m at its thickest north of the Canadian Archipelago (not shown). Observational estimates tend to show thicknesses of rather 4 m or more. This is much more than we varied in our sensitivity study and points to a possible stronger influence of sea-ice thickness on the difference between simulated and observed brightness temperature than estimated from the sensitivity study in this case.

10   The remaining uncertainty contained in Δ_{ARC3O} can have several sources: ARC3O itself, further biases in the simulated climate state, or wrong assumptions. Biases in ARC3O itself can arise from wrong assumptions in the sea-ice emission model MEMLS or in the ocean emission and atmospheric transfer model by Wentz and Meissner (2000), but also from the definition of first-year and multiyear ice. The definition we use does not take into account the dynamics of the ice. As a consequence, if a grid cell is located in a region where sea-ice circulates horizontally and this grid cell therefore contains ice for more than a

15   year, the ice in this grid cell will be defined as multiyear ice. This is the case even if the ice transported through the grid cell is not the same physical ice floe throughout this time period but a different first-year ice floe every day for example.
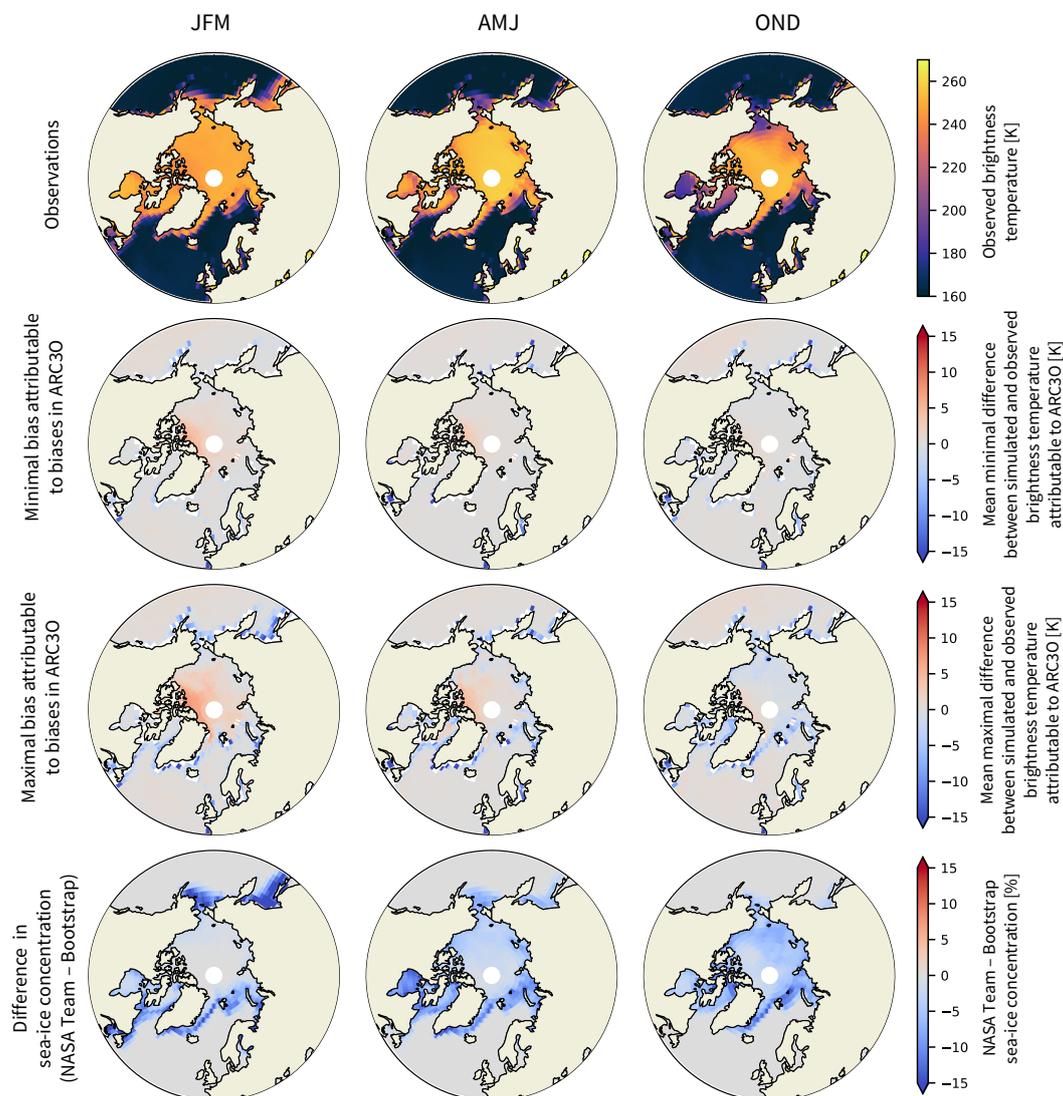
Figure 8. Observed brightness temperatures (1st row), mean minimal (2nd row) and mean maximal (3rd row) estimates of Δ_{ARC3O} and differences in sea-ice concentration (4th row) between the NASA Team and Bootstrap assimilation runs, i.e. maximal estimates of Δ_{retriev}.

Additionally, the brightness temperatures simulated over open ocean by ARC3O tend to be around 3 K too low. More investigation and fine-tuning above the open ocean might therefore slightly lower the uncertainty introduced by ARC3O as well.

Concerning the surface temperature, we expect small biases for the simulated surface temperature, because the ERA-Interim

5   reanalysis compares well to the few available in-situ observations (Lindsay et al., 2014) and satellite-retrieved temperature

measurements could be used if further evaluation was needed. We currently cannot explore the additional impact of model biases in sea-ice thickness and snow thickness as in-situ observations are rare and satellite retrievals of these variables are not necessarily robust yet. Observational estimates for sea-ice and snow thickness are mainly based on retrieval algorithms, similar to sea-ice concentration estimates. Possible biases may therefore remain in these variables compared to reality.

5   Finally, we assumed that the real sea-ice concentration lies in the range between the Bootstrap and the NASA Team estimates. As only limited evaluation against reality is possible, the uncertainty between real and retrieved sea-ice concentration might be different to our assumption.

As a conclusion, we showed in the consistent model setup that the sea-ice concentration is the main driver for large variations in the brightness temperature in regions that are not fully ice-covered. In regions where the sea-ice concentration is very high

10   and does not vary much, such as the Central Arctic in winter, the surface temperature is the main driver of variations in the brightness temperature. Simulated and observed brightness temperatures are generally in good agreement. Most differences are likely driven by the uncertainty brought by the sea-ice concentration products compared to reality. Remaining differences attributable to biases in ARC3O remain below 5 K.

The lack of evaluation possibilities for the observation operator is an indicator for how little is actually known about the

15   real Arctic climate state and processes at work, in particular in regard to the real sea-ice concentration and surface temperature. Extending this observation operator to lower and higher frequencies would be of advantage to fill this gap. Brightness temperatures at different frequencies and polarizations are sensitive to different particular parameters. The combination of different brightness temperatures could enable a comprehensive assessment of the Arctic Ocean surface and atmosphere, and a comprehensive evaluation of the individual observation operators. For example, using an observation operator applied to

20   reanalysis data, Richter et al. (2018) simulated brightness temperatures at the frequency of 1.4 GHz. They found that sea-ice concentration and surface temperature are the main drivers for variations in the brightness temperature in the Central Arctic but that, in regions of thin ice, the ice thickness is the dominating driver at this frequency. Combining frequencies in this case would then enable a climate model evaluation encompassing different perspectives.

4.4   Melting season (JAS)

25   Like in winter, the simulated summer brightness temperatures are primarily a function of sea-ice concentration. Melt ponds are then the main challenge for sea-ice retrieval algorithms, as their passive microwave signature is undistinguishable from open water. This leads to large uncertainties and potential underestimation of the sea-ice concentration in summer (Meier and Notz, 2010; Cavalieri et al., 1990; Comiso and Kwok, 1996; Fetterer and Untersteiner, 1998; Rösel et al., 2012b; Kern et al., 2016). The difference between observed and simulated brightness temperature (Fig. 9, 2nd row) and the difference between

30   observational sea-ice concentration products is therefore much larger in summer than in winter (Ivanova et al., 2015; Kern et al., 2016).

In summer, the simulation of brightness temperatures in ARC3O is only based on the combination of a constant bare ice brightness temperature and melt-pond brightness temperature, weighted by the melt-pond fraction. We derived the constant brightness temperature from direct observations and it is representative for summer bare ice on the order of ± 3.56 K (see
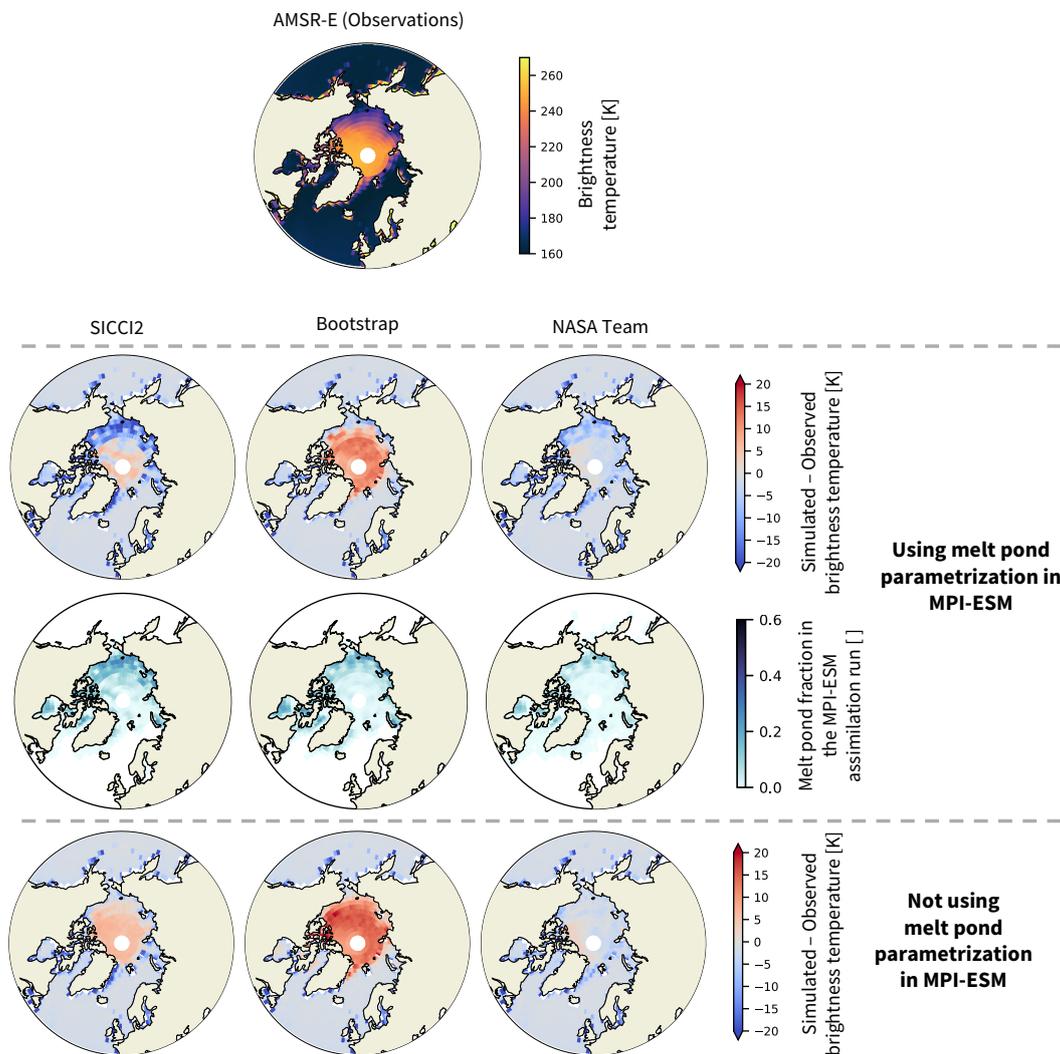
Figure 9. Experiment to compare brightness temperatures in summer (July/August/September) simulated based on assimilation runs assimilated with different sea-ice concentration products (SICCI2, Bootstrap, NASA Team) to brightness temperatures measured by AMSR-E. Observations by AMSR-E are shown in the 1st row. Difference between simulated and observed brightness temperatures if melt ponds are taken into account (2nd row) and if they are set to zero (4th row). In the 3rd row, the melt pond fraction simulated by MPI-ESM is shown.

Sec. 3.1.4). Assumptions about the chosen constant ice surface brightness temperature can influence the analysis but we assume that the uncertainty between simulated and observed brightness temperatures in summer is mainly driven by two parameters: the difference between real and retrieved sea-ice concentration and the difference between real and simulated melt-pond fraction.

In the following, we can therefore evaluate the sea-ice concentration products in summer and their relationship to the melt
5 pond fraction. To do so, we distinguish between two types of sea-ice concentration: the total sea-ice concentration and the

pond-free sea-ice concentration. The pond-free sea-ice concentration is the concentration of sea ice visible by the satellite, assuming that melt ponds are open water. In MPI-ESM, we know both the total sea-ice concentration and the pond-free sea-ice concentration as melt ponds are represented through a melt-pond parametrization, which is a function of the surface energy budget and water drainage to the ocean (Roeckner et al., 2012). In the SICCI2 algorithm, melt ponds are not explicitly accounted for but the dynamic tie-points are based on observed brightness temperatures in areas of high sea-ice concentration, which are covered by melt ponds in summer. The retrieved sea-ice concentration will therefore implicitly account for the melt-ponds, potentially reducing the underestimation of the sea-ice concentration (Kern et al., 2016; Lavergne et al., 2019). In the Bootstrap algorithm, a correction is applied to the sea-ice concentration to account for the effect of melt ponds by synthetically increasing the retrieved sea-ice concentration (Comiso and Kwok, 1996; Bunzel et al., 2016), while in the NASA Team algorithm, no correction is applied (Comiso et al., 1997; Bunzel et al., 2016). By switching on and off the melt-pond parametrization in MPI-ESM, we can evaluate the ability of observational products to produce a reasonable pond-free and total sea-ice concentration.

In this experiment, we run ARC3O on the three MPI-ESM assimilation runs setting the melt-pond fraction to zero everywhere. We then compare this set of simulated brightness temperatures to observed brightness temperatures (Fig. 9, 4th row) and to the set of brightness temperatures simulated taking into account the melt-pond distribution simulated by MPI-ESM (Fig. 9, 2nd row). The results give different insights depending on the sea-ice concentration product used for the assimilation.

For the SICCI2 product, the simulated brightness temperature of the pond-free sea ice is higher than the observed brightness temperature in the Central Arctic. Melt ponds cover the whole ice-covered Arctic Ocean (Rösel et al., 2012a; Istomina et al., 2015a). Adding their effect in the brightness temperature simulation could therefore reduce the difference between simulated and observed brightness temperature. However, while the melt-pond parametrization in MPI-ESM reduces the overall brightness temperature, the reduction is very heterogeneous so that the brightness temperature is now largely underestimated in the Pacific sector but still overestimated in most of the Central Arctic. This means that the dynamic tie-point approach of the SICCI2 algorithm seems to take into account the effect of melt ponds in a reasonable way, therefore yielding a too high pond-free sea-ice concentration. The brightness temperatures simulated using the melt-pond parametrization suggest that SICCI2 does not represent well the total sea-ice concentration. However, the melt-pond distribution in MPI-ESM seems to be too heterogeneous and therefore unrealistic in some regions, as most simulated melt ponds concentrate in the Pacific sector and not many can be found over the Central Arctic (see Fig. 9, 3rd row, and Roeckner et al., 2012). The latitudinal gradient is realistic, showing more melt ponds in lower latitudes, where the incoming solar radiation and air temperatures are higher, than in higher latitudes. However, the melt pond fraction in the Central Arctic was observed to be slightly higher than simulated and the melt pond fraction in some of the marginal regions was observed to be slightly lower than simulated (Rösel et al., 2012a; Istomina et al., 2015a). Corrections to the simulated melt pond fraction in these directions might reduce the simulated brightness temperature in the Central Arctic and increase it in marginal regions, approaching the observed brightness temperature. As a consequence, it seems that the total sea-ice concentration might be well represented in the SICCI2 dataset but we cannot robustly confirm this assumption with our setup due to the apparently somewhat unrealistic melt-pond parametrization provided by MPI-ESM.

For the Bootstrap product, if the ice is assumed to be pond-free, the simulated brightness temperature is more than 10 K higher than the observed brightness temperature over the whole Central Arctic. Due to this large difference, adding melt ponds on top of the ice is not sufficient to counteract this overestimation of the brightness temperature, which remains on the order of 10 K. This means that Bootstrap tends to overestimate both the pond-free and total sea-ice concentration in summer.

5   For the NASA Team product, the simulated brightness temperature of the pond-free ice is very close to the observed brightness temperature. As a consequence, the addition of melt ponds leads mainly to a negative bias compared to observations. This means that the NASA Team dataset represents well the pond-free sea-ice concentration, in agreement with previous results by Ivanova et al. (2015) but that it tends to underestimate the total sea-ice concentration in summer.

The main conclusions show that the main driver for differences between simulated and observed summer brightness tem-
10   peratures are again the differences between retrieved and real sea-ice concentration. However, the melt-pond parametrization used in MPI-ESM is too heterogeneous and unrealistic and therefore contributes to the difference between simulated and observed brightness temperatures as well. For further analysis, the melt-pond parametrization could however be replaced by a climatology using observational melt-pond estimates, such as Rösel et al. (2012a) or Istomina et al. (2015a). This could reduce the uncertainty induced by the melt-pond parametrization.

15   5   Conclusions

In this study, we present the first observational operator for the Arctic Ocean that is applied to GCM output, following suggestions from Burgard et al. (2019). It allows us to simulate brightness temperatures at a frequency of 6.9 GHz, vertical polarization, for the whole Arctic Ocean. The results look promising and open up possibilities for further and deeper analysis of simulated and observed Arctic climate state.

20   Simulated and observed brightness temperatures compare well. In winter, differences between observed and simulated brightness temperatures attributable to biases in ARC3O are well below 5 K. In comparison, the total difference between observed and simulated brightness temperatures ranges from well below 5 K up to 10 K. The large differences can be attributed to possible differences between real and retrieved climate state, especially in sea-ice concentration, and, to a lower extent, to the process of data assimilation into the model. In summer, the difference between simulated and observed bright-
25   ness temperatures locally reach more than 15 K. This difference can be attributed to high differences in the underlying sea-ice concentration products and potential biases arising from the melt-pond parametrization in the climate model.

The low estimate of the uncertainty induced by the observation operator ARC3O itself in the comparison between simulated and observed brightness temperatures showed that it is possible to simulate realistic brightness temperatures based on simple output of a GCM. This is a necessary step to open the way for similar observation operators for different frequencies and
30   polarizations and, as a consequence, for new climate model evaluation and model initialization techniques in a hindcast or unconstrained model run. Additionally, ARC3O can be used to evaluate observation products against satellite measurements by using assimilation runs.

An observation operator translates a consistent climate state into one observable quantity. In climate model evaluation, the full simulated Arctic climate state can therefore be evaluated against one observed quantity instead of several different retrieved quantities, which all carry uncertainties with them, especially in the Arctic region (Jakobson et al., 2012; Lindsay et al., 2014; Ivanova et al., 2015; Boisvert et al., 2018). With one observation operator at one single frequency, not all effects can be disentangled clearly, e.g. in this case the influence of sea-ice concentration and surface temperature in the Central Arctic are comparable. Further development of observational operators for different frequencies is essential to use this approach to its fullest. A multi-frequency framework would allow us to investigate this consistent climate state from different perspectives, as different variables affect different frequencies differently. Also at 6.9 GHz, further investigation and work might improve the brightness temperature simulation.

The possibility of comparing the climate state of a climate model in only one observable quantity is also very beneficial to model initialization through data assimilation. The first-guess procedure used in data assimilation methods, such as variational data assimilation (Talagrand and Courtier, 1987; Andersson et al., 1994) or ensemble Kalman filters (Evensen, 1994; Hunt et al., 2007), would then be based on a consistent climate state and be conducted in observation space, independent of retrieval algorithms (Richter et al., 2018; Scott et al., 2012). This is already done and has led to improvements in weather prediction systems for other regions than the Arctic (e.g. Terasaki and Miyoshi, 2017).

The observational uncertainty of the sea-ice concentration is very large in summer. We showed here that, if we are able to reduce the uncertainty in the melt-pond representation of the model, we can relate differences between observed and simulated brightness temperatures directly to differences between retrieved and real total sea-ice concentration. This is a promising perspective as melt ponds are a strong challenge for the retrieval of summer sea-ice concentrations.

Finally, ARC3O is a simple observation operator as it is based on variables simulated by all GCMs, can be applied to output from any kind of GCM simulation, and does not require extensive computational power. It is therefore a powerful tool which has the potential to uncover model biases and improve model initialization by providing a new perspective on the Arctic climate system.

Code and data availability. Primary data and scripts used in this study are archived by the Max Planck Institute for Meteorology and can be obtained by contacting publications@mpimet.mpg.de.

References

Andersson, E., Pailleux, J., Thépaut, J.-N., Eyre, J., McNally, A., Kelly, G., and Courtier, P.: Use of cloud-cleared radiances in three/four-dimensional variational data assimilation, Q. J. Roy. Meteor. Soc., 120, 627–653, https://doi.org/10.1002/qj.49712051707, 1994.

Balmaseda, M., Mogensen, K., and Weaver, A.: Evaluation of the ECMWF ocean reanalysis system ORAS4, Q. J. Roy. Meteor. Soc., 139, 1132–1161, https://doi.org/10.1002/qj.2063, 2013.

Barber, D., Fung, A., Grenfell, T., Nghiem, S., Onstott, R., Lytle, V., Perovich, D., and Gow, A.: The role of snow on microwave emission and scattering over first-year sea ice, IEEE T. Geosci. Remote, 36, 1750–1763, https://doi.org/10.1109/36.718643, 1998.

Boisvert, L., Webster, M., Petty, A., Markus, T., Bromwich, D., and Cullather, R.: Intercomparison of Precipitation Estimates over the Arctic Ocean and Its Peripheral Seas from Reanalyses, J. Climate, 31, 8441–8462, https://doi.org/10.1175/JCLI-D-18-0125.1, 2018.

Brodzik, M., Long, D., Hardman, M., Paget, A., and Armstrong, R.: MEaSUREs Calibrated Enhanced-Resolution Passive Microwave Daily EASE-Grid 2.0 Brightness Temperature ESDR, Version 1, https://doi.org/10.5067/MEASURES/CRYOSPHERE/NSIDC-0630.001, [Accessed in October 2018], 2016, Updated 2018.

Bunzel, F., Notz, D., Baehr, J., Müller, W., and Fröhlich, K.: Seasonal climate forecasts significantly affected by observational uncertainty of Arctic sea ice concentration, Geophys. Res. Lett., 43, 852–859, https://doi.org/10.1002/2015GL066928, 2016.

Burgard, C., Notz, D., Pedersen, L., and Tonboe, R.: The Arctic Ocean Observation Operator for 6.9 GHz (ARC3O) - Part 1: How to obtain sea-ice brightness temperatures at 6.9 GHz from climate model output, Submitted to The Cryosphere, 2019.

Cavalieri, D., Burns, B., and Onstott, R.: Investigation of the effects of summer melt on the calculation of sea ice concentration using active and passive microwave data, J. Geophys. Res-Oceans, 95, 5359–5369, https://doi.org/10.1029/JC095iC04p05359, 1990.

Cavalieri, D. J., Parkinson, C., Gloersen, P., and Zwally, H.: Sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS passive microwave data, version 1, https://doi.org/10.5067/8GQ8LZQVL0VL, [Accessed in August 2014, updated yearly], 1996.

Chang, T. and Gloersen, P.: Microwave Emission from dry and wet snow, in: Operational Applications of Satellite Snowcover Observations, edited by Rango, A., chap. 27, pp. 399–407, NASA, 1975.

Comiso, J.: Bootstrap sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS, version 2, https://doi.org/10.5067/J6JQLS9EJ5HU, [Accessed in August 2014, updated yearly], 2000.

Comiso, J. and Kwok, R.: Surface and radiative characteristics of the summer Arctic sea ice cover from multisensor satellite observations, J. Geophys. Res-Oceans, 101, 28 397–28 416, https://doi.org/10.1029/96JC02816, 1996.

Comiso, J., Cavalieri, D., Parkinson, C., and Gloersen, P.: Passive Microwave Algorithms for Sea Ice Concentration: A Comparison of Two Techniques, Remote Sens. Environ., 60, 357–384, https://doi.org/10.1016/S0034-4257(96)00220-9, 1997.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hersbach, H., Holm, E., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A., Monge-Sanz, B., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thébaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res-Oceans, 99, 10 143–10 162, https://doi.org/10.1029/94JC00572, 1994.

Eyring, V., Bony, S., Meehl, G., Senior, C., Stevens, B., Stouffer, R., and Taylor, K.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Eyring, V., Cox, P., Flato, G., Gleckler, P., Abramowitz, G., Caldwell, P., Collins, W., Gier, B., Hall, A., Hoffman, F., Hurtt, G., Jahn, A., Jones, C., Klein, S., Krasting, J., Kwiatkowski, L., Lorenz, R., Maloney, E., Meehl, G., Pendergrass, A., Pincus, R., Ruane, A., Russell, J., Sanderson, B., Santer, B., Sherwood, S., Simpson, I., Stouffer, R., and Williamson, M.: Taking climate model evaluation to the next level, Nat. Clim. Change, https://doi.org/10.1038/s41558-018-0355-y, 2019.

Fetterer, F. and Untersteiner, N.: Observations of melt ponds on Arctic sea ice, J. Geophys. Res-Oceans, 103, 24 821–24 835, https://doi.org/10.1029/98JC02034, 1998.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M.: Evaluation of Climate Models, book section 9, pp. 741–866, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/CBO9781107415324.020, 2013.

Giorgetta, M., Roeckner, E., Mauritsen, T., Bader, J., Crueger, T., Esch, M., Rast, S., Kornblueh, L., Schmidt, H., Kinne, S., Hohenegger, C., Möbis, B., Krismer, T., Wieners, K., and Stevens, B.: The atmospheric general circulation model ECHAM6: Model description, Tech. Rep. Reports on Earth System Science, 135/2013, Max Planck Institute for Meteorology, 2013.

Griewank, P. and Notz, D.: A 1-D modelling study of Arctic sea-ice salinity, Cryosphere, 9, 305–329, https://doi.org/10.5194/tc-9-305-2015, 2015.

Hallikainen, M. and Winebrenner, D.: The Physical Basis for Sea Ice Remote Sensing, in: Microwave Remote Sensing of Sea Ice, edited by Carsey, F., chap. 4, pp. 29–46, American Geophysical Union, 1992.

Hallikainen, M., Ulaby, F., and Abdelrazik, M.: Dielectric properties of snow in the 3 to 37 GHz range, IEEE T Antenn. Propag., 34, 1329–1340, https://doi.org/10.1109/TAP.1986.1143757, 1986.

Hibler, W.: A Dynamic Thermodynamic Sea Ice Model, J. Phys. Oceanogr., 9, 815–846, 1979.

Hunt, B., Kostelich, E., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, Physica D, 230, 112–126, https://doi.org/10.1016/j.physd.2006.11.008, 2007.

Istomina, L., Heygster, G., Huntemann, M., Marks, H., Melsheimer, C., Zege, E., Malinka, A., Prikhach, A., and Katsev, I.: Melt pond fraction and spectral sea ice albedo retrieval from MERIS data - Part 2: Case studies and trends of sea ice albedo and melt ponds in the Arctic for years 2002-2011, Cryosphere, 9, 1567–1578, https://doi.org/10.5194/tc-9-1551-2015, 2015a.

Istomina, L., Heygster, G., Huntemann, M., Schwarz, P., Birnbaum, G., Scharien, P., Polashenski, C., Perovich, D., Zege, E., Malinka, A., Prikhach, A., and Katsev, I.: Melt pond fraction and spectral sea ice albedo retrieval from MERIS data - Part 1: Validation against in situ, aerial, and ship cruise data, Cryosphere, 9, 1551–1566, https://doi.org/10.5194/tc-9-1551-2015, 2015b.

Ivanova, N., Johannessen, O. M., Pedersen, L. T., and Tonboe, R. T.: Retrieval of Arctic Sea Ice Parameters by Satellite Passive Microwave Sensors: A Comparison of Eleven Sea Ice Concentration Algorithms, IEEE T. Geosci. Remote, 52, 7233–7246, https://doi.org/10.1109/TGRS.2014.2310136, 2014.

Ivanova, N., Pedersen, L., Kern, S., Heygster, G., Lavergne, T., Sørensen, A., Saldo, R., Dybkjaer, G., Brucker, L., and Shokr, M.: Intercomparison and evaluation of sea ice algorithms: towards further identification of challenges and optimal approach using passive microwave observations, Cryosphere, 9, 1797–1817, https://doi.org/10.5194/tc-9-1797-2015, 2015.

Jakobson, E., Vihma, T., Keernik, H., and Jaagus, J.: Validation of atmospheric reanalyses over the central Arctic Ocean, Geophys. Res. Lett.,
39, https://doi.org/10.1029/2012GL051591, 2012.

Jungclaus, J., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., Mikolajewicz, U., Notz, D., and von Storch, J.: Characteristics of
the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model, J. Adv.
Model Earth Sy., 5, 422–446, https://doi.org/10.1002/jame.20023, 2013.

Kern, S., Rösel, A., Pedersen, L., Ivanova, N., Saldo, R., and Tonboe, R.: The impact of melt ponds on summertime microwave brightness
temperatures and sea-ice concentrations, Cryosphere, 10, 2217–2239, https://doi.org/10.5194/tc-10-2217-2016, 2016.

Kern, S., Lavergne, T., Notz, D., Pedersen, L., Tonboe, R., Saldo, R., and A.M., S.: Satellite Passive Microwave Sea-Ice Concentration Data
Set Intercomparison: Closed Ice and Ship-Based Observations, The Cryosphere, 13, 3261–3307, https://doi.org/10.5194/tc-13-3261-2019,
2019.

Lavergne, T., Macdonald Sørensen, A., Kern, S., Tonboe, R., Notz, D., Aaboe, S., Bell, L., Dybkjær, Eastwood, S., Gabarro, C., Heygster,
G., Killie, M., Brandt Kreiner, M., Lavelle, J., Saldo, R., Sandven, S., and Pedersen, L.: Version 2 of the EUMETSAT OSI SAF and ESA
CCI sea-ice concentration climate data records, Cryosphere, 13, 49–78, https://doi.org/10.5194/tc-13-49-2019, 2019.

Lindsay, R., Wensnahan, M., Schweiger, A., and Zhang, J.: Evaluation of Seven Different Atmospheric Reanalysis Products in the Arctic, J.
Climate, 27, 2588–2606, https://doi.org/10.1175/JCLI-D-13-00014.1, 2014.

Meier, W. and Notz, D.: A note on the accuracy and reliability of satellite-derived passive microwave estimates of sea-ice extent, CliC Arctic
sea ice working group consensus document, World Climate Research Program, 2010.

Nandan, V., Geldsetzer, T., Yackel, J., Mahmud, M., Scharien, R., Howell, S., King, J., Ricker, R., and Else, B.: Effect of Snow Salinity on
CryoSat-2 Arctic First-Year Sea Ice Freeboard Measurements, Geophys. Res. Lett., 44, 419–426, https://doi.org/10.1002/2017GL074506,
2017.

Niederdrenk, A. and Notz, D.: Arctic sea ice in a 1.5°C warmer world, Geophys. Res. Lett., 45, https://doi.org/10.1002/2017GL076159,
2018.

Notz, D.: Thermodynamic and Fluid-Dynamical Processes in Sea Ice, Ph.D. thesis, University of Cambridge, 2005.

Notz, D. and Stroeve, J.: Observed Arctic sea-ice loss directly follows anthropogenic CO_2 emission, Science,
https://doi.org/10.1126/science.aag2345, 2016.

Notz, D., Haumann, A., Haak, H., and Marotzke, J.: Arctic sea-ice evolution as modeled by Max Planck Institute for Meteorology's Earth
system model, J. Adv. Model Earth Sy., 5, 173–194, https://doi.org/10.1002/jame.20016, 2013.

Pedersen, L., Saldo, R., Ivanova, N., Kern, S., Heygster, G., Tonboe, R., Huntemann, M., Ozsoy, B., Ardhuin, F., and Kaleschke, L.: Ref-
erence dataset for sea ice concentration, https://doi.org/10.6084/m9.figshare.6626549.v6, https://figshare.com/articles/Reference_dataset_
for_sea_ice_concentration/6626549/3, 2018.

Richter, F., Drusch, M., Kaleschke, L., Maaß, N., Tian-Kunze, X., and Mecklenburg, S.: Arctic sea ice signatures: L-band brightness tem-
perature sensitivity comparison using two radiation transfer models, Cryosphere, 12, 921–933, https://doi.org/10.5194/tc-12-921-2018,
2018.

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

Rösel, A., Kaleschke, L., and Birnbaum, G.: Melt ponds on Arctic sea ice determined from MODIS satellite data using an artificial neural
network, Cryosphere, 6, 431–446, https://doi.org/10.5194/tc-6-431-2012, 2012a.

Rösel, A., Kaleschke, L., and Kern, S.: Influence of melt ponds on microwave sensors' sea ice concentration retrieval algorithms, 2012 IEEE International Geoscience and Remote Sensing Symposium, pp. 3261–3264, https://doi.org/10.1109/IGARSS.2012.6350608, 2012b.

Scott, K. A., Buehner, M., Caya, A., and Carrieres, T.: Direct Assimilation of AMSR-E Brightness Temperatures for Estimating Sea Ice Concentration, Mon. Weather Rev., 140, 997–1013, https://doi.org/10.1175/MWR-D-11-00014.1, 2012.

5 Shokr, M. and Sinha, N.: Remote Sensing Principles Relevant to Sea Ice, in: Sea Ice: Physics and Remote Sensing, Geophysical Monograph 209, First Edition, edited by Union, A. G., chap. 7, pp. 271–335, John Wiley & Sons, Inc., 2015.

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblueh, L., Lohmann, U., Pincus, R., Reichler, T., and Roeckner, E.: Atmospheric component of the MPI-M Earth System Model: ECHAM6, J. Adv. Model Earth Sy., 5, 146–172, https://doi.org/10.1002/jame.20015, 2013.

10 Swift, C. and Cavalieri, D.: Passive microwave remote sensing for sea ice research, Eos, Transactions American Geophysical Union, 66, 1210–1212, https://doi.org/10.1029/EO066i049p01210, 1985.

Talagrand, O. and Courtier, P.: Variational Assimilation of Meteorological Observations With the Adjoint Vorticity Equation. I: Theory, Q. J. Roy. Meteor. Soc., 113, 1311–1328, https://doi.org/10.1002/qj.49711347812, 1987.

Taylor, K., Stouffer, R., and Meehl, G.: An Overview of CMIP5 and the Experiment Design, Bull. Amer. Meteor. Soc., 93, 485–498, 2012.

15 Terasaki, K. and Miyoshi, T.: Assimilating AMSU-A Radiances with the NICAM-LETKF, J. Meteorol. Soc. Jpn. Ser. II, 95, 433–446, https://doi.org/10.2151/jmsj.2017-028, 2017.

Tietsche, S., Notz, D., Jungclaus, J., and Marotzke, J.: Assimilation of sea-ice concentration in a global climate model - physical and statistical aspects, Ocean Sci., 9, 19–36, https://doi.org/10.5194/os-9-19-2013, 2013.

Tonboe, R.: The simulated sea ice thermal microwave emission at window and sounding frequencies, Tellus, 62A, 333–344, https://doi.org/10.1111/j.1600-0870.2010.00434.x, 2010.

20 Tonboe, R., Andersen, S., Toudal, L., and Heygster, G.: Sea ice emission modelling, in: Thermal Microwave Radiation - Applications for Remote Sensing, edited by Mätzler, C., Rosenkranz, P., Battaglia, A., and Wigneron, J., pp. 382–400, IET Electromagnetic Waves Series 52, 2006.

Tonboe, R., Eastwood, S., Lavergne, T., Sørensen, A., Rathmann, N., Dybkjær, G., Pedersen, L., Høyer, J., and Kern, S.: The EUMETSAT sea ice concentration climate data record, Cryosphere, 10, 2275–2290, https://doi.org/10.5194/tc-10-2275-2016, 2016.

25 Ulaby, F., Moore, R., and Fung, A.: Passive microwave sensing of the ocean, in: Microwave Remote Sensing, Active and Passive Volume III, From Theory to Applications, chap. 18, pp. 1412–1521, Artech House, Inc., 1986.

Wentz, F. and Meissner, T.: Algorithm theoretical basis document (atbd), version 2, Tech. Rep. AMSR Ocean Algorithm, RSS Tech. Proposal 121599A-1, Remote Sensing Systems, Santa Rosa, CA, 2000.

30 Wiesmann, A. and Mätzler, C.: Microwave emission model of layered snowpacks, Remote Sens. Environ., 70, 307–316, 1999.

Appendix A: Temperature profile in snow and ice

The temperature at the interface between ice and snow is computed as follows:

$$T_{\text{ice,surf}} = \frac{T_{\text{snow,surf}} \cdot \frac{k_s}{h_s} + T_{\text{bottom}} \cdot \frac{k_i}{h_i}}{\frac{k_s}{h_s} + \frac{k_i}{h_i}} \tag{A1}$$

with k_s the thermal conductivity of snow (= 0.31 W/Km), k_i the thermal conductivity of ice (= 2.17 W/Km), h_s the snow thickness, h_i the ice thickness, $T_{snow,surf}$ the temperature at the surface of the snow, T_{bottom} the temperature at the bottom of the ice, set to -1.8 °C.

Appendix B: Tuning of the temperature profiles

5   The brightness temperatures initially produced by ARC3O were clearly too bright (Fig. B1, left). A comparison of the simulated emissivities with emissivities derived from observational data from the RRDP showed that ARC3O systematically overestimates the emissivity. The brightness temperature is defined as the product of the emissivity and the physical temperature of the emitting part of the ice (Ulaby et al., 1986). As it is not straightforward to find where the bias is produced in the emission model, we chose to multiply the inherent sea-ice emissivity with a tuning coefficient to counteract the systematic bias. To do

10   so, we selected all points with a sea-ice concentration of 99.7% or more to avoid influence from open water, in the year 2004. We then multiplied the sea-ice surface brightness temperature by a range of coefficients between 0.96 and 0.975. We found the best agreement between simulated and observed brightness temperatures for a coefficient of 0.963 in the months January, February, and March and for a coefficient of 0.973 in the months October, November, December. As a consensus, we therefore chose a coefficient of 0.968 to apply to the sea-ice brightness temperatures, which yields a more reasonable distribution of

15   brightness temperatures (Fig. B1, right).
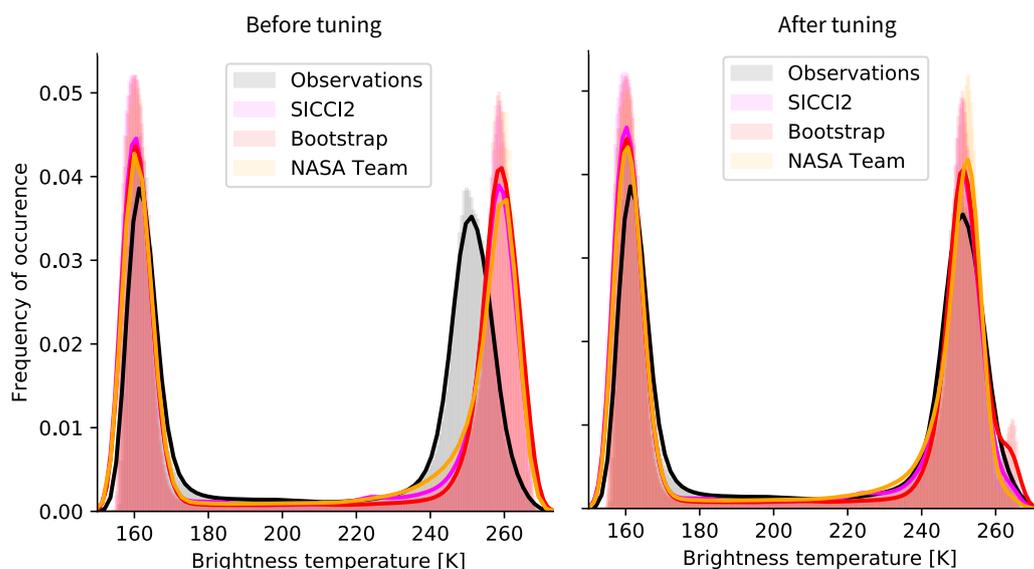
Figure B1. Density distribution of the brightness temperatures in the three simulated cases and in observations in the untuned (left) and tuned (right) version for the years 2005 to 2008.

Appendix C: Variability in climate parameters
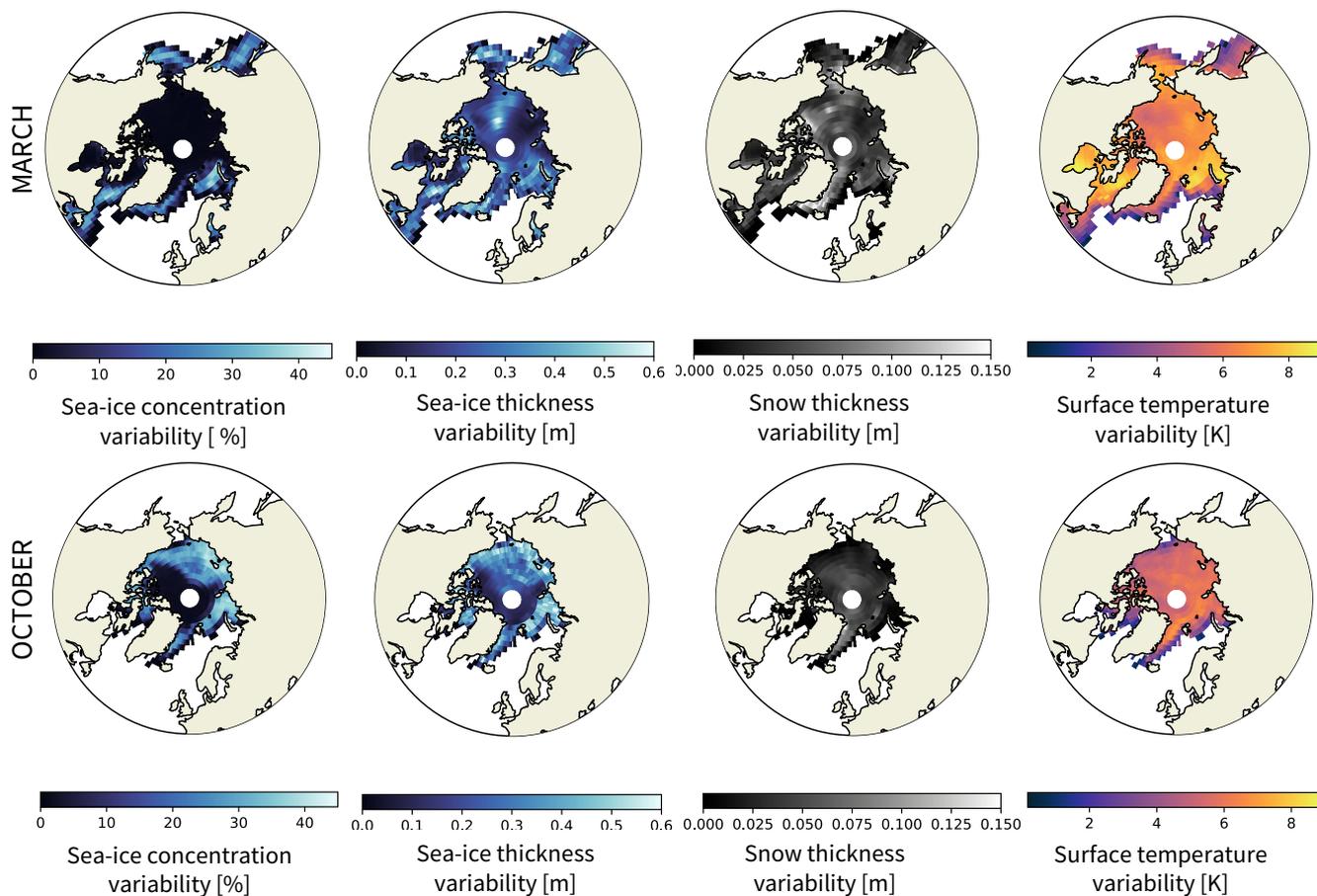
Figure C1. Standard deviation of the anomaly with regard to the time mean for each grid cell. This is the parameter used to modulate the input variables in the sensitivity studies of Sec. 4.3.2

Author contributions. C.B., D.N. and L.T.P. developed the original idea of this manuscript. C.B. carried out all analyses and wrote the manuscript. All authors contributed to discussions.

Competing interests. No competing interests are present.

Acknowledgements. We thank Stefan Kern for constructive comments and discussions. This work was funded by the project "ESA CCI Sea Ice Phase 2".