

# Geophysical Research Letters

## RESEARCH LETTER

10.1029/2020GL091784

### Key Points:

- Models overestimate the memory of Arctic sea-ice area from summer into the following year, which cannot be explained by internal variability
- Observed summer memory arises from a reemergence in the perennial ice zone dampened by negative correlations in the seasonal ice zone
- The model overestimation emerges in the relation between sea-ice area anomalies in the seasonal ice zone and preceding summer anomalies

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

C. Giesse,  
[celine.giesse@uni-hamburg.de](mailto:celine.giesse@uni-hamburg.de)

### Citation:

Giesse, C., Notz, D., & Baehr, J. (2021). On the origin of discrepancies between observed and simulated memory of Arctic sea ice. *Geophysical Research Letters*, 48, e2020GL091784. <https://doi.org/10.1029/2020GL091784>

Received 24 NOV 2020  
 Accepted 15 MAY 2021

© 2021. The Authors.  
 This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

## On the Origin of Discrepancies Between Observed and Simulated Memory of Arctic Sea Ice

Céline Giesse<sup>1,2</sup> , Dirk Notz<sup>1,3</sup> , and Johanna Baehr<sup>1</sup> 

<sup>1</sup>Institute of Oceanography, Center for Earth System Research and Sustainability (CEN), Universität Hamburg, Hamburg, Germany, <sup>2</sup>International Max Planck Research School on Earth System Modelling, Hamburg, Germany, <sup>3</sup>Max Planck Institute for Meteorology, Hamburg, Germany

**Abstract** To investigate the inherent predictability of sea ice and its representation in climate models, we compare the seasonal-to-interannual memory of Arctic sea ice as given by lagged correlations of sea-ice area anomalies in large model ensembles (Max Planck Institute Grand Ensemble and Coupled Model Intercomparison Project phase 6) and multiple observational products. We find that state-of-the-art climate models significantly overestimate the memory of pan-Arctic sea-ice area from the summer months into the following year. This cannot be explained by internal variability. We further show that the observed summer memory can be disentangled regionally into a reemergence of positive correlations in the perennial ice zone and negative correlations in the seasonal ice zone; the latter giving rise to the discrepancy between observations and model simulations. These findings could explain some of the predictability gap between potential and operational forecast skill of Arctic sea-ice area identified in previous studies.

**Plain Language Summary** Sea ice, as a relatively slowly varying component in the climate system, holds “memory” on seasonal-to-interannual timescales. This means that, based on the current state of the sea ice, meaningful predictions of its state several months into the future can be made, for instance with the use of climate models. Such sea-ice predictions are of growing socioeconomic importance, particularly in the Arctic, where the strong sea-ice loss in the last decades is giving rise to new risks as well as economic opportunities. Here, we provide a comparison of the memory of Arctic sea ice based on model and observational data. We show that current global climate models systematically overestimate the memory of sea ice from one summer into the following year and beyond. We further show that this overestimation arises from how the outer region of the Arctic, which is only seasonally ice-covered, remembers previous summer sea ice. Our findings imply that, first, there is likely a misrepresentation of processes related to the memory of sea ice in climate models and, second, the potential of making skillful sea-ice predictions is less strong than previously assumed based on model simulations.

## 1. Introduction

Over the past decade, major advances in the prediction of sea ice have been made. This includes the development of seasonal sea-ice prediction systems based on coupled global climate models (GCMs) utilized for operational predictions (e.g., Chevallier et al., 2013; Msadek et al., 2014; Peterson et al., 2015; Sigmond et al., 2013; Wang et al., 2013) and “perfect-model” studies (e.g., Blanchard-Wrigglesworth, Bitz, & Holland, 2011; Day et al., 2016; Day, Hawkins, & Tietsche, 2014; Germe et al., 2014; Holland et al., 2011; Koenig & Mikolajewicz, 2009; Tietsche et al., 2014). The latter provide an upper limit of the predictability of sea ice within a given model framework under the assumption of perfect model physics and knowledge of initial conditions (potential predictability). While perfect-model studies show that pan-Arctic sea-ice area (SIA) or extent (SIE) are predictable at 12–36 month lead times, operational predictions of detrended pan-Arctic SIA/SIE are skillful only for lead times of a few months (1–6 months for summer SIE and 1–11 months for winter SIE depending on the prediction system) (Bushuk et al., 2019; Guemas et al., 2016). This gap between potential predictability and operational forecast skill has been noted in previous studies (e.g., Blanchard-Wrigglesworth et al., 2015; Guemas et al., 2016). Bushuk et al. (2019) provide the first consistent assessment of potential and operational forecast skill within one GCM-based prediction system and find a substantial skill gap in nearly all Arctic regions. This predictability gap could indicate a strong

potential for improvements of operational sea-ice predictions, either through improved initialization or improved model physics. However, the skill gap might also hint at a systematic overestimation of sea-ice predictability in state-of-the-art GCMs, as previously suggested by Notz (2017) and Blanchard-Wrigglesworth and Bushuk (2019). This brings up the question of which predictability can be expected based on observations, which we address here.

One way to analyze the inherent predictability of sea ice, arising from the memory/persistence of its initial conditions, in models as well as observations are lagged correlation studies (Blanchard-Wrigglesworth & Bushuk, 2019; Blanchard-Wrigglesworth, Armour, et al., 2011; Bushuk & Giannakis, 2015; Bushuk et al., 2015, 2017; Chevallier & Salas-Mélia, 2012; Day, Tietsche, & Hawkins, 2014; Krikken & Hazeleger, 2015; Ordoñez et al., 2018). As found by Blanchard-Wrigglesworth, Armour, et al. (2011), the memory of pan-Arctic SIA anomalies is characterized by an initial persistence of 2–5 months and two distinct modes of memory reemergence, in which lagged correlations increase again after an initial drop. The first identified mode of memory reemergence occurs between months of the melt and the freezing season (“melt-to-growth season reemergence”) and is related to an imprint of SIA anomalies on sea surface temperature (SST) anomalies in the vicinity of the sea-ice edge, which persist over the summer season. The second mode occurs between the months of one summer and the next (“summer-to-summer reemergence” or in later works also “growth-to-melt season reemergence”) and can be explained by a similar exchange of anomalies between SIA and sea-ice thickness. In addition, the ice-albedo feedback adds to the persistence and reemergence during the summer months. Day, Tietsche, and Hawkins (2014) showed that, despite some inter-model spread in the magnitude of correlations, the memory patterns are robust across different GCMs.

Comparing the memory of pan-Arctic SIA in model simulations and observations, previous studies noted generally higher lagged correlations in the models than in the observations as well as differences in the occurrence of reemergence (Blanchard-Wrigglesworth, Armour, et al., 2011; Day, Tietsche, & Hawkins, 2014; Krikken & Hazeleger, 2015). While the melt-to-growth season reemergence is present in observational data, there is no significant signal of summer-to-summer reemergence. As pointed out by Day, Tietsche, and Hawkins (2014), the attribution of discrepancies to potential causes is complicated by several factors of uncertainty, such as the shortness of the observational record and the detrending of the time series.

With the present study, we aim to systematically analyze differences in the memory of Arctic sea-ice in model simulations and observations: Can they be attributed to internal variability or errors in the model physics? Where do they occur regionally? In contrast to previous studies, we base our lagged correlation analysis of SIA anomalies on a multitude of simulated and observational data: the Max Planck Institute Grand Ensemble (MPI-GE, Maher et al., 2019), a Coupled Model Intercomparison Project phase 6 (CMIP6; Eyring et al., 2016) multi-model ensemble, and several observational data products. By comparing lagged correlations from observational data to the range of model internal variability of large ensemble simulations covering the same period, we systematically identify time lags at which simulated memory is over-/underestimated. By analyzing not only the memory of pan-Arctic SIA but also regional memory of SIA, we gain insights into the spatial origin of discrepancies between models and observations.

## 2. Data and Methods

### 2.1. Sea-Ice Concentration Data Sets

For our analysis, we use monthly sea-ice concentration (SIC) data of the period 1979–2018 from various model and observational data products. We analyze model data from the MPI-GE (Maher et al., 2019), combining historical simulations (1979–2005) and representative concentration pathway 4.5 (RCP4.5) simulations (2006–2018) performed with the Max Planck Institute Earth System Model (MPI-ESM, Giorgetta et al., 2013) from 100 model ensemble members. Additionally, we use a CMIP6 (Eyring et al., 2016) multi-model ensemble consisting of 240 members from 37 different models. For this ensemble, we use all available historical simulations (period 1979–2014) except those performed with MPI-ESM. This allows judgment on whether results obtained with the MPI-GE are model-specific or can be generalized for state-of-the-art GCMs. Note that, due to the consideration of all available simulations, the individual models are weighted differently depending on the amount of provided ensemble members, but qualitatively similar

results are obtained when analyzing only one member per model. A table listing the contributing CMIP6 models with their number of ensemble members is provided in the supporting information (Table S1).

Furthermore, we use three observational products of SIC retrieved from satellite records with different retrieval algorithms, namely Bootstrap (Comiso, 2017) and NASA Team (Cavalieri et al., 1996) data from the National Snow and Ice Data Center (NSIDC) and EUMETSAT Ocean and Sea Ice Satellite Application Facility (OSI SAF) data (Lavergne et al., 2019; OSI SAF, 2017, 2019). The usage of different observational products allows us to take into account the uncertainty in observed SIC (e.g., Kern et al., 2019, 2020).

## 2.2. Quantification of Memory

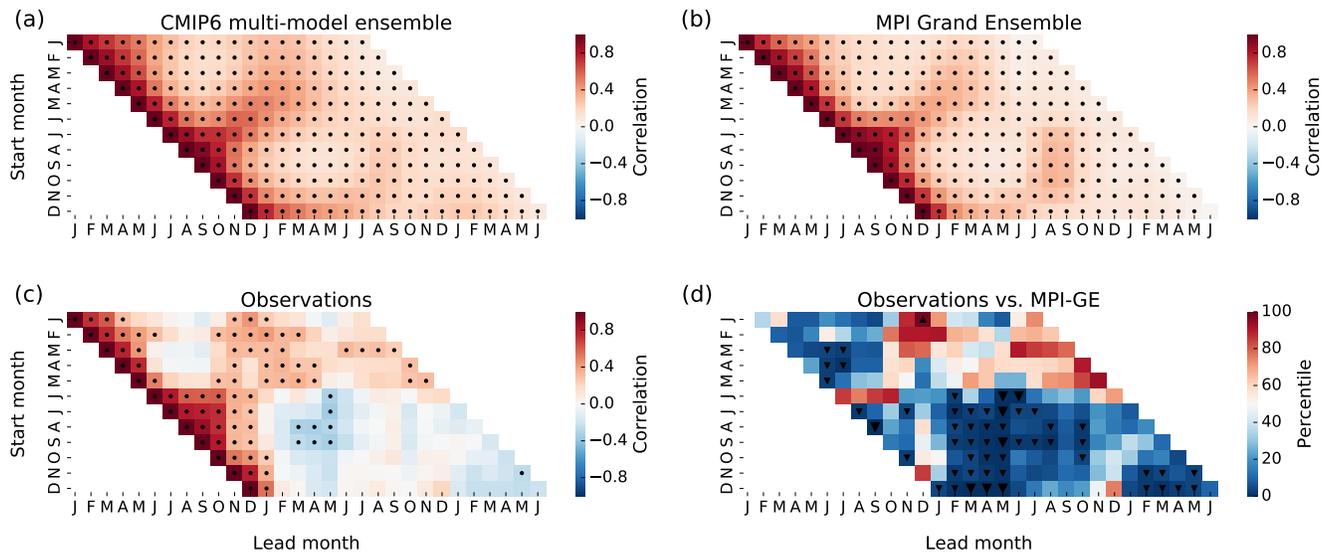
We quantify memory of Arctic sea ice in terms of lagged correlations of SIA anomalies. From the SIC data sets, we compute monthly time series of pan-Arctic and regional SIA, differentiating between a seasonal ice zone (SIZ) and a perennial ice zone (PIZ). We define the PIZ to consist of all grid cells with a September SIC of  $\geq 0.15$  (corresponding to the annual minimum ice extent) in at least 80% of the years based on the NSIDC Bootstrap data; all other grid cells are considered as SIZ. For the CMIP6 multi-model ensemble, we consider only pan-Arctic SIA, determined as described in SIMIP Community (2020). To remove externally driven long-term trends, we detrend the time series of individual months using locally weighted scatterplot smoothing (LOWESS; Cleveland, 1979). This local regression provides a more accurate representation of the sea-ice decline than a linear regression, as the negative trend is increasing with time, particularly in the sea-ice minimum months (e.g., Serreze & Stroeve, 2015). In the supporting information, we provide a visual comparison of the LOWESS and linear detrending (Figure S1) and show some key results based on linearly detrended time series, allowing for a direct comparison to previous studies. From the resulting monthly SIA anomalies, we calculate lagged correlations with time lags of up to 18 months using Pearson's correlation coefficient  $r$ . For details on the computation of SIA and the statistical methods applied for the combination of correlation coefficients of ensemble members, the computation of statistical significance, and the detrending, we refer to the supporting information (Text S1).

## 3. Results

### 3.1. Memory of Pan-Arctic Sea-Ice Area

Analyzing the lagged correlations of pan-Arctic SIA, all data sets show an initial decline of memory associated with the persistence of SIA anomalies (Figures 1a–1c and Figure S2 for the individual observational data sets). Related to the seasonal cycle, two persistence regimes can be differentiated: one centered around the sea-ice maximum (winter persistence, January to May start months) and one centered around the sea-ice minimum (summer persistence, June to December start months). The  $e$ -folding decorrelation time ranges between 1 and 6 months depending on the initial month and data set, which is consistent with previous studies (Blanchard-Wrigglesworth, Armour, et al., 2011; Day, Tietsche, & Hawkins, 2014; Krikken & Hazeleger, 2015). Furthermore, all data sets show a melt-to-growth-season reemergence of memory (high correlations between pairs of months around the sea-ice minimum, i.e., August–September, July–October, etc.; Blanchard-Wrigglesworth, Armour, et al., 2011). The correlation between pairs of months around the sea-ice minimum is less clear-cut in the observations than in the models. However, the relation from winter to winter is stronger in the observations than in the MPI-GE, as also noted in previous studies (Blanchard-Wrigglesworth, Armour, et al., 2011; Krikken & Hazeleger, 2015). The CMIP6 ensemble reproduces the observed winter-to-winter memory better than the MPI-GE.

For the summer-to-summer memory, differences between observations and model simulations are more apparent than for other time lags. The model ensembles show a clear summer-to-summer reemergence (high correlations between the summer minimum months, particularly August/September, from one year to the next; Blanchard-Wrigglesworth, Armour, et al., 2011). The signal is more pronounced in the MPI-GE than in the CMIP6 ensemble (September 1-year lag correlation of 0.31 and 0.24, respectively). In the observations, the correlations from the summer months beyond the persistence timescale are substantially lower than in the models (e.g., September 1-year lag correlation of  $-0.06$ ), with even significant negative correlations from summer to spring of the next year. Despite these low correlation coefficients, there is an increase in correlations (from  $-0.35$  at minimum to around zero) in the summer months. This could



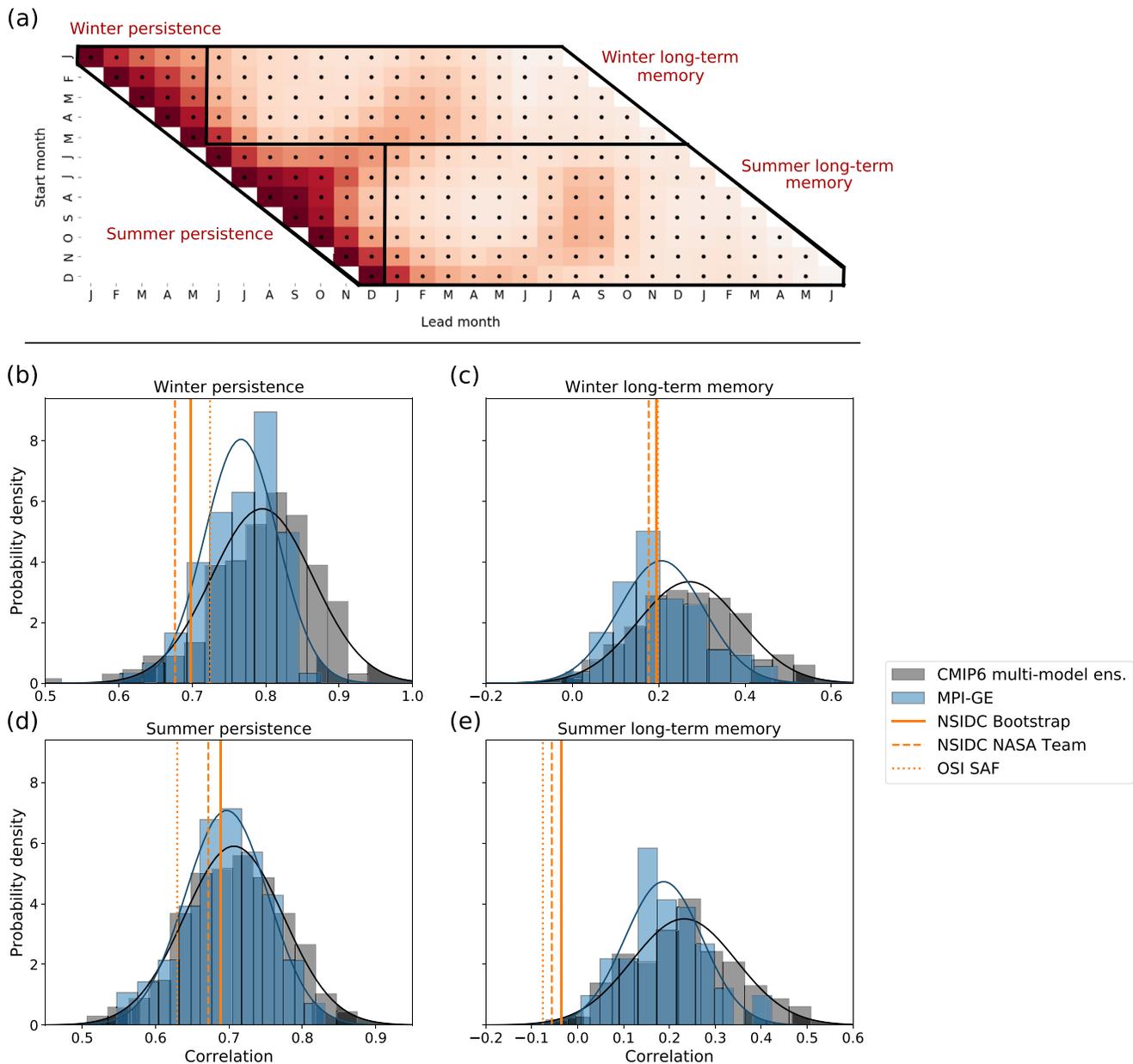
**Figure 1.** (a–c) Lagged correlations of monthly pan-Arctic sea-ice area anomalies in (a) the Coupled Model Intercomparison Project phase 6 (CMIP6) multi-model ensemble, (b) the Max Planck Institute Grand Ensemble (MPI-GE), and (c) the observational products (National Snow and Ice Data Center [NSIDC] Bootstrap, NSIDC NASA Team, and Ocean and Sea Ice Satellite Application Facility [OSI SAF] data) combined to a “3-member ensemble.” Correlation coefficients of individual ensemble members or data products are combined using a Fisher’s z-transformation. Black dots indicate statistical significance on the 99% level. (d) Percentile of MPI-GE members with a lower correlation than observational data for the respective time lag. Downward and upward triangles mark values within the 5th and 95th percentile. Time lags with correlations coefficients outside of the model range (0th and 100th percentile) are marked with a larger triangle.

indicate a reemergence of summer SIA anomalies that is superimposed with negative summer-to-summer correlations caused by a different process. Note that when detrending the time series linearly, the correlation coefficients are higher, rendering the summer-to-summer reemergence in the observations more visible (Figure S3), but we still find statistically significant negative correlations from summer to spring.

As the model correlation values represent an average of many ensemble members and the observations represent only a single time series, differences could be due to internal variability. Still, one would expect the observations to lie within the range of model variability. There are several patterns or individual time lags for which the correlation coefficients from observations are at the edge of model variability (Figure 1d). Most evident is the pattern of time lags from the summer months into the following year, in which the observed correlations are consistently within or below the 5th percentile of MPI-GE correlation coefficients. This is a strong indication for a systematic overestimation of memory related to errors in the model physics. Similar patterns of model over-/underestimation are found when ranking the observed correlations within the CMIP6 ensemble (not shown).

For a more detailed view of the internal model variability, we define four different memory regimes (winter persistence, winter long-term memory, summer persistence, and summer long-term memory regime; Figure 2a) and compare the mean correlation coefficients for each of these memory regimes in the individual data products to their distribution in the MPI-GE and CMIP6 ensemble (Figures 2b–2e). The distributions of correlation coefficients in the MPI-GE and CMIP6 ensemble have a large overlap (75%–90% depending on the memory regime), indicating that the MPI-ESM model behaves similarly to other CMIP6 models in terms of memory. The CMIP6 ensemble has a wider spread than the MPI-GE, which is expected due to the larger ensemble size and the variety of contributing models.

Comparing the model ensembles against observations, we find that for the winter persistence, winter long-term memory, and summer persistence regimes (Figures 2b–2d), despite some spread, all observational products show correlation coefficients that lie within the range of correlations simulated in both the CMIP6 ensemble and the MPI-GE. For the summer long-term memory (Figure 2e), however, the correlations of all three observational data sets are below the model ensemble range (except for two CMIP6 ensemble members having a lower correlation than the NSIDC NASA Team and Bootstrap data). This indicates that



**Figure 2.** (a) Definition of different memory regimes. (b–e) Distribution of mean correlation coefficients for the four different memory regimes shown as histograms for the Coupled Model Intercomparison Project phase 6 (CMIP6) multi-model ensemble (gray) and the Max Planck Institute Grand Ensemble (MPI-GE) (blue), and as lines for the observational data sets (National Snow and Ice Data Center [NSIDC] Bootstrap, NSIDC NASA Team, and Ocean and Sea Ice Satellite Application Facility [OSI SAF] data). The black and blue lines show normal distribution fits to the CMIP6 and MPI-GE data, respectively. Shadings indicate the  $2\sigma$ -range.

the observations are not just an “outlier,” but that climate models systematically overestimate the memory in the summer long-term regime. In the case of linearly detrended time series, the observed correlations are within the range of CMIP6 and MPI-GE internal variability, but also in the lower tail of the distribution (Figure S4).

Note that qualitatively similar results are obtained when analyzing lagged correlations of pan-Arctic SIE (see Figures S5 and S6). As noted by Blanchard-Wrigglesworth, Armour, et al. (2011), SIE anomalies are slightly less persistent than SIA anomalies as they are more sensitive to dynamic wind forcing. Moreover, the SIE does not account for variations in the interior of the ice cover in summer. As a consequence, there is only a weak signal of simulated summer-to-summer reemergence for SIE. Observed summer-to-spring

correlations are negative also for SIE, albeit lower in magnitude than for SIA, suggesting that both variations in the ice pack as well as in the position of the sea-ice edge are involved. Equally as for SIA, all observational data sets show correlations of SIE in the summer long-term memory regime that are below the model ensemble range.

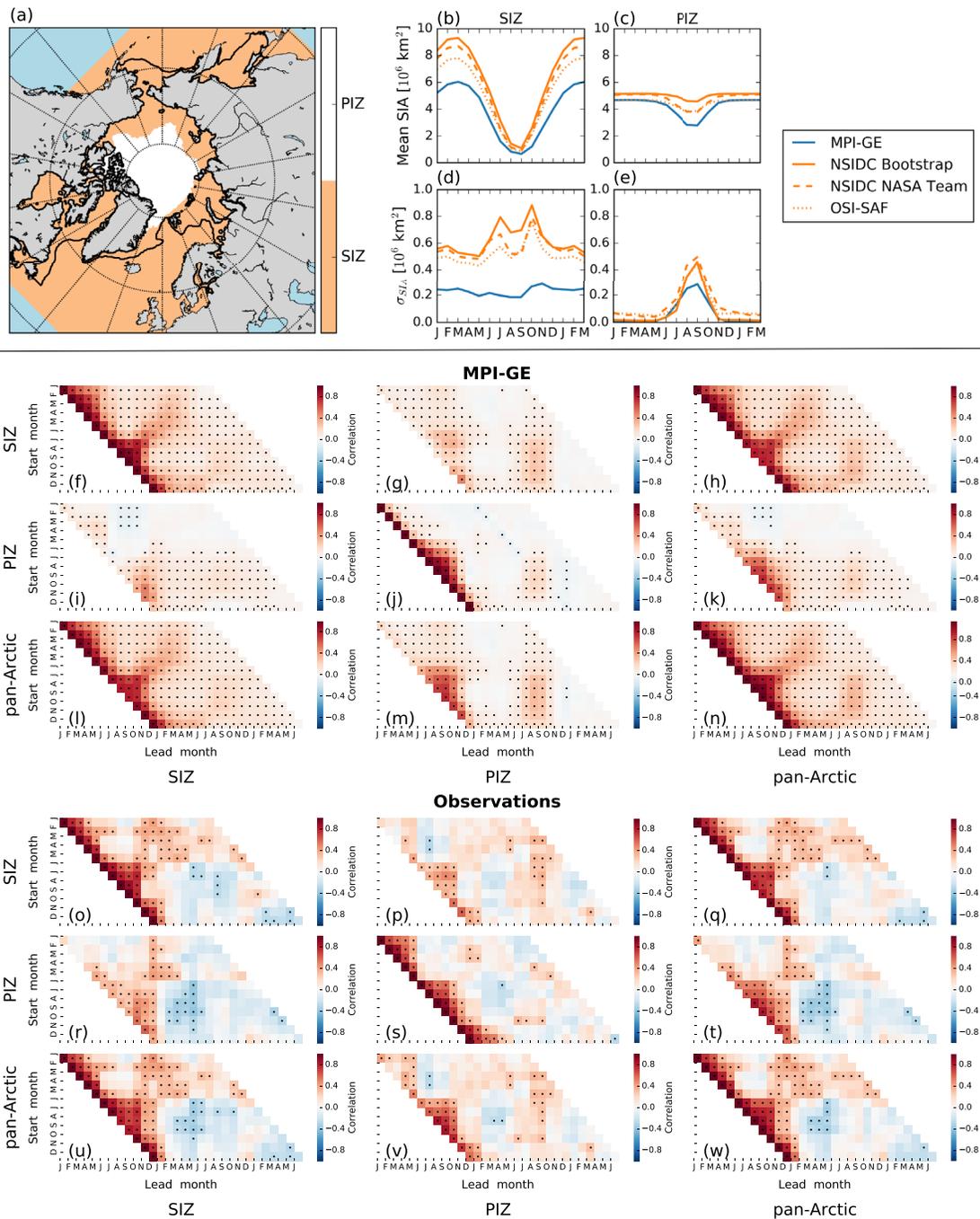
### 3.2. Regional Memory of Sea-Ice Area

To investigate whether some of the memory properties and differences between the data sets have a certain spatial origin, we analyze the memory of SIA on a regional level. As shown by Ordoñez et al. (2018), the memory of regional SIA can vary substantially between different Arctic basins: It is impacted on the one hand by the geographic location and associated ocean dynamics, and on the other hand by the seasonal cycle of the regional SIA and its variability. For simplicity, we here choose a variability-based regional separation, differentiating only between the SIZ, which is characterized by thin, seasonal ice in the vicinity of the ice edge, and the PIZ, which contains thick, multi-year ice in the center of the Arctic Ocean (see map in Figure 3a). While the SIA of the SIZ has a pronounced seasonal cycle and year-round variability (Figures 3b and 3d), the SIA of the PIZ is practically constant throughout most of the year with a dip and substantial interannual variability in the months around the sea-ice minimum (Figures 3c and 3e). From analyzing the lagged correlations between different combinations of SIZ, PIZ, and pan-Arctic SIA anomalies (Figures 3f–3w), we can gain information on the spatial occurrence and origin of memory.

The different memory characteristics, identified on the pan-Arctic scale, show different regional occurrences. The persistence of SIA anomalies is strongly connected to the seasonal cycle. As both ice zones exhibit seasonal variations of ice area in summer, SIA anomalies of both SIZ (Figures 3f and 3o) and PIZ (Figures 3j and 3s) persist during summer and contribute to the summer persistence on the pan-Arctic scale (Figures 3h, 3k, 3q, and 3t). In winter, the ice area in the SIZ shows strong seasonal variations, while the ice area in the PIZ is practically constant. Thus, only the SIZ (Figures 3f and 3o) shows a pronounced signal of winter persistence, reflected also on the pan-Arctic scale (Figures 3h and 3q). The observations also show an intra-regional persistence of winter SIA anomalies in the PIZ (Figure 3s) not present in the model. However, these correlations result from only small fluctuations of the otherwise full ice cover and do not transfer any memory to the pan-Arctic scale (Figure 3t). Similar to the winter persistence, the melt-to-growth season reemergence is only apparent in the SIZ (Figure 3, left column) but not in the PIZ (Figure 3, middle column), as it is related to the imprint of SIA anomalies to the SST in the vicinity of the sea-ice edge. Overall, the regional memory in the persistence and winter long-term memory regimes is consistent between MPI-GE and observations.

The most striking result on the pan-Arctic scale is the overestimation of the summer long-term memory, which is characterized by a summer-to-summer reemergence in the model simulations and negative correlations in the observations. The inter-regional correlations show that summer SIA anomalies from both ice zones (especially from the SIZ) reemerge in the PIZ (Figure 3, middle column) but barely in the SIZ (Figure 3, left column). Albeit weaker than in the MPI-GE data, the reemergence signal is also present in the observations, implying that it is a real-world phenomenon and not just a model artifact. As the summer-to-summer reemergence is explained by an imprint of the SIA anomalies to the ice thickness that persists throughout the winter (e.g., Blanchard-Wrigglesworth, Armour, et al., 2011), its occurrence in the PIZ but not in the SIZ is plausible. In the SIZ, instead of a reemergence, the MPI-GE shows low, positive correlations in the summer long-term memory regime, whereas the observations show negative correlations in spring and summer of the next year. The observed negative correlations arise primarily from summer SIA anomalies in the PIZ (summer-to-spring, Figure 3r) and to a smaller extent from summer SIA anomalies in the SIZ (mainly summer-to-summer, Figure 3o).

Hence, the superposition of reemergence and negative correlations, as seen on the pan-Arctic scale, can be disentangled regionally and the discrepancies between model simulations and observations arise from a different relation between SIA anomalies in the SIZ and preceding summer anomalies. This finding is further reinforced by comparing the inter-regional lagged correlations in the observations to their internal model variability in the MPI-GE (Figures S7 and S8). While the memory of pan-Arctic summer SIA anomalies in the PIZ agrees well between the data sets, in the SIZ the correlation coefficients of all observational data



**Figure 3.** (a) Map showing the geographical area of the perennial ice zone (PIZ; grid cells with September sea-ice concentration larger than 0.15 in 80% of the years in the National Snow and Ice Data Center [NSIDC] Bootstrap time series) and seasonal ice zone (SIZ; remaining grid cells). The bold black line indicates the average annual maximum sea-ice extent. (b and c) Seasonal cycle of mean sea-ice area in the SIZ and PIZ in Max Planck Institute Grand Ensemble (MPI-GE) and observational data sets (NSIDC Bootstrap, NSIDC NASA Team, and Ocean and Sea Ice Satellite Application Facility [OSI SAF] data). (d and e) Seasonal cycle of the standard deviation of sea-ice area in the SIZ (upper row), PIZ (middle row), and entire Arctic (lower row) with succeeding sea-ice area anomalies in the SIZ (left column), PIZ (middle column), and entire Arctic (right column) in the MPI Grand Ensemble. (m–u) Same as (d–l) but for the observational ensemble.

sets are below the MPI-GE range of correlation values, indicating a significant overestimation of memory, similar to the pan-Arctic scale (Figure 2e).

#### 4. Discussion

We presented a comprehensive overview and comparison of Arctic sea-ice memory/persistence in a large set of model and observational data based on lagged correlations of SIA anomalies. Our results are consistent with previous studies (e.g., Blanchard-Wrigglesworth, Armour, et al., 2011; Day, Tietsche, & Hawkins, 2014; Krikken & Hazeleger, 2015) in identifying the same persistence and reemergence characteristics of pan-Arctic SIA and noting an overestimation of the memory from summer into the following year (summer long-term memory) in model simulations compared to observations. While previous studies point out the lack of summer-to-summer reemergence in observations, we additionally note an even larger discrepancy between models and observations in the persistence of summer anomalies into the following spring, where observational data consistently show negative correlations which are not found in model simulations. Comparing our results to Blanchard-Wrigglesworth and Bushuk (2019), CMIP6 models show better agreement with observations than CMIP5 models, particularly in the winter-to-winter memory (see their Figure 1e). These differences could be related to model improvements or changes in the forcing, but may also be influenced by differences in the methodology (i.e., different time periods, detrending methods, and memory regime definitions).

Beyond that, this study shows the robustness of models overestimating the summer long-term memory in many aspects. By analyzing the distribution of lagged correlations in large model ensembles for the same period as the observational record, we show that the overestimation cannot be explained by internal variability. This reduces the likelihood of the discrepancy being caused by a “sampling error” due to the shortness of the observational time series as suggested by Day, Tietsche, and Hawkins (2014). The overestimation is present not only within a single-model ensemble but also in the CMIP6 multi-model ensemble, showing its robustness across state-of-the-art GCMs. Moreover, the overestimation of summer long-term memory is independent of the considered observational data set. Using three observational data products (NSIDC Bootstrap, NASA Team, OSI SAF) that use different retrieval algorithms to determine SIC from satellite measurements, we reduce the uncertainty associated with observations. However, it should be noted that the discrepancy between model simulations and observations is related to SIC anomalies in the summer months, in which observations have their largest uncertainty due to the presence of melt ponds (e.g., Kern et al., 2020). Another factor of uncertainty is the applied method of detrending, which could either not fully capture the long-term trend or remove parts of the low-frequency internal variability. Applying a linear detrending instead of the LOWESS detrending, as done for instance by Blanchard-Wrigglesworth, Armour, et al. (2011) and Day, Tietsche, and Hawkins (2014), yields higher correlations and observed summer long-term memory correlations that are no longer outside the range of model internal variability, but still in the lower tail of the distribution (Figures S3 and S4). This could be due to the remaining non-linear part of the trend, which may be stronger in observations than model simulations.

While on the pan-Arctic scale the summer-to-summer reemergence is only detectable in model simulations, we could show that, in the PIZ, summer SIA anomalies reemerge also in observational data. The discrepancy between models and observations, however, is found in the relation between SIA anomalies in the SIZ and preceding summer SIA anomalies, where observational data show significant negative correlations not present in the model simulations. The negative correlations arise primarily from SIA anomalies in the PIZ, which suggests a non-local mechanism. However, a part of the negative correlations arises in the SIZ, indicating that also the position of the ice edge is involved. This is reinforced by the finding that the model overestimation of summer long-term memory is significant not only for pan-Arctic SIA but also for SIE (Figures S5 and S6). While the variability-based separation between PIZ and SIZ nicely disentangles the summer-to-summer reemergence from the negative correlations, it does not reflect the geographical complexity of the Arctic Ocean and regional sea-ice dynamics. As shown by Ordoñez et al. (2018), the strength of persistence and reemergence features strongly depends on the geographical location. Moreover, it should be noted that the fixed separation between SIZ and PIZ can only be an approximation as it does not reflect the changing mean sea-ice state in the period of interest. Still, our regional analysis provides guidance for future work identifying the causes of the discrepancy between models and observations.

The findings of this study have important implications for the predictability of sea ice. As previously suggested by Notz (2017) and Blanchard-Wrigglesworth and Bushuk (2019), the overestimation of memory of pan-Arctic SIA in the summer long-term memory regime could explain a part of the predictability gap between perfect-model experiments and operational forecasts (e.g., Bushuk et al., 2019). This would imply that perfect-model studies overestimate the potential predictability of pan-Arctic SIA arising from knowledge of summer sea-ice conditions and that the potential for improvement of sea-ice predictions is less strong than these studies suggest. Nevertheless, this can only be a partial explanation of the year-round predictability gap and does not diminish the potential for improved operational sea-ice predictions, for instance, through a better initialization. Moreover, there are additional sources of sea-ice predictability that are not considered here, such as ice thickness/volume and oceanic variables. Regarding future research, it should be of high priority to identify the causes of the overestimation of summer long-term memory in state-of-the-art GCMs.

## 5. Conclusions

In summary, we draw the following conclusions from our analysis and data set-intercomparison of lagged correlations of Arctic SIA anomalies:

- The memory of pan-Arctic SIA from the summer months into the following year and beyond (“summer long-term memory”) is significantly overestimated in model simulations compared to observations. Observed lagged correlations in this memory regime are below the range of internal model variability (MPI-GE) and inter-model variability (CMIP6 multi-model ensemble), showing that the result is robust across state-of-the-art climate models.
- The observed summer long-term memory can be disentangled regionally into a summer-to-summer reemergence in the PIZ and negative correlations in the SIZ. The observed negative relation between summer SIA anomalies in both ice zones, particularly in the PIZ, and succeeding spring and summer SIA anomalies in the SIZ is not present in model simulations, giving rise to the model overestimation.
- The results reinforce that a part of the predictability gap between potential and operational forecast skill of Arctic SIA could be caused by over-persistence of summer SIA in models.

## Data Availability Statement

All data used for this study are publicly available. The MPI Grand Ensemble and CMIP6 data can be accessed from the ESGF (MPI-GE: <https://esgf-data.dkrz.de/projects/mpi-ge/>, CMIP6: <https://esgf-node.llnl.gov/projects/cmip6/>). The observational sea-ice concentration data sets are available online (NSIDC Bootstrap: <https://nsidc.org/data/nsidc-0079/>, NSIDC NASA Team: <https://nsidc.org/data/nsidc-0051>, OSI SAF: [www.osi-saf.org](http://www.osi-saf.org)).

## Acknowledgments

This study is a contribution to the Cluster of Excellence ‘CLICCS - Climate, Climatic Change, and Society’, contribution to the Center for Earth System Research and Sustainability (CEN) of Universität Hamburg. The data analysis for this study was performed with computing resources from the German Climate Computing Center (Deutsches Klimarechenzentrum, DKRZ). The authors thank Jakob Dörr for providing processed CMIP6 data of sea-ice area. The authors acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. The authors thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. Open access funding enabled and organized by Projekt DEAL.

## References

- Blanchard-Wrigglesworth, E., Armour, K. C., Bitz, C. M., & DeWeaver, E. (2011). Persistence and inherent predictability of Arctic sea ice in a GCM ensemble and observations. *Journal of Climate*, 24(1), 231–250. <https://doi.org/10.1175/2010JCLI3775.1>
- Blanchard-Wrigglesworth, E., Bitz, C. M., & Holland, M. M. (2011). Influence of initial conditions and climate forcing on predicting Arctic sea ice. *Geophysical Research Letters*, 38(18). <https://doi.org/10.1029/2011GL048807>
- Blanchard-Wrigglesworth, E., & Bushuk, M. (2019). Robustness of Arctic sea-ice predictability in GCMs. *Climate Dynamics*, 52(9–10), 5555–5566. <https://doi.org/10.1007/s00382-018-4461-3>
- Blanchard-Wrigglesworth, E., Cullather, R. I., Wang, W., Zhang, J., & Bitz, C. M. (2015). Model forecast skill and sensitivity to initial conditions in the seasonal Sea Ice Outlook. *Geophysical Research Letters*, 42(19), 8042–8048. <https://doi.org/10.1002/2015GL065860>
- Bushuk, M., & Giannakis, D. (2015). Sea-ice reemergence in a model hierarchy. *Geophysical Research Letters*, 42(13), 5337–5345. <https://doi.org/10.1002/2015GL063972>
- Bushuk, M., Giannakis, D., & Majda, A. J. (2015). Arctic sea ice reemergence: The role of large-scale oceanic and atmospheric variability. *Journal of Climate*, 28(14), 5477–5509. <https://doi.org/10.1175/JCLI-D-14-00354.1>
- Bushuk, M., Msadek, R., Winton, M., Vecchi, G. A., Gudgel, R., Rosati, A., & Yang, X. (2017). Skillful regional prediction of Arctic sea ice on seasonal timescales. *Geophysical Research Letters*, 44(10), 4953–4964. <https://doi.org/10.1002/2017GL073155>
- Bushuk, M., Msadek, R., Winton, M., Vecchi, G. A., Yang, X., Rosati, A., & Gudgel, R. (2019). Regional Arctic sea-ice prediction: Potential versus operational seasonal forecast skill. *Climate Dynamics*, 52(5–6), 2721–2743. <https://doi.org/10.1007/s00382-018-4288-y>
- Cavaliere, D. J., Parkinson, C. L., Gloersen, P., & Zwally, H. J. (1996). *Sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I-SS-MIS passive microwave data, version 1*. NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/8GQ8LZQVLOVL>
- Chevallier, M., & Salas-Méla, D. (2012). The role of sea ice thickness distribution in the arctic sea ice potential predictability: A diagnostic approach with a coupled GCM. *Journal of Climate*, 25(8), 3025–3038. <https://doi.org/10.1175/JCLI-D-11-00209.1>

- Chevallier, M., Salas y Mélia, D., Voldoire, A., Déqué, M., & Garric, G. (2013). Seasonal forecasts of the Pan-Arctic sea ice extent using a GCM-based seasonal prediction system. *Journal of Climate*, *26*(16), 6092–6104. <https://doi.org/10.1175/JCLI-D-12-00612.1>
- Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, *74*(368), 829–836. <https://doi.org/10.1080/01621459.1979.10481038>
- Comiso, J. (2017). *Bootstrap sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS, version 3*. NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/7Q8HCCWS4I0R>
- Day, J. J., Hawkins, E., & Tietsche, S. (2014). Will Arctic sea ice thickness initialization improve seasonal forecast skill? *Geophysical Research Letters*, *41*(21), 7566–7575. <https://doi.org/10.1002/2014GL061694>
- Day, J. J., Tietsche, S., Collins, M., Goessling, H. F., Guemas, V., Guillory, A., et al. (2016). The Arctic Predictability and Prediction on Seasonal-to-Interannual Timescales (APPOSITE) data set version 1. *Geoscientific Model Development*, *9*(6), 2255–2270. <https://doi.org/10.5194/gmd-9-2255-2016>
- Day, J. J., Tietsche, S., & Hawkins, E. (2014). Pan-Arctic and regional sea ice predictability: initialization month dependence. *Journal of Climate*, *27*(12), 4371–4390. <https://doi.org/10.1175/JCLI-D-13-00614.1>
- EUMETSAT Ocean and Sea Ice Satellite Application Facility. (2017). *Global sea ice concentration climate data record 1979–2015 (v2.0, 2017)*, OSI-450. [https://doi.org/10.15770/EUM\\_SAF\\_OSI\\_0008](https://doi.org/10.15770/EUM_SAF_OSI_0008)
- EUMETSAT Ocean and Sea Ice Satellite Application Facility. (2019). *Global sea ice concentration interim climate data record 2016 onwards (v2.0, 2019)*, OSI-430b.
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. <https://doi.org/10.5194/gmd-9-1937-2016>
- Germe, A., Chevallier, M., Salas y Mélia, D., Sanchez-Gomez, E., & Cassou, C. (2014). Interannual predictability of Arctic sea ice in a global climate model: Regional contrasts and temporal evolution. *Climate Dynamics*, *43*(9–10), 2519–2538. <https://doi.org/10.1007/s00382-014-2071-2>
- Gorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., et al. (2013). Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *Journal of Advances in Modeling Earth Systems*, *5*(3), 572–597. <https://doi.org/10.1002/jame.20038>
- Guemas, V., Blanchard-Wrigglesworth, E., Chevallier, M., Day, J. J., Déqué, M., Doblas-Reyes, F. J., et al. (2016). A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales. *Quarterly Journal of the Royal Meteorological Society*, *142*(695), 546–561. <https://doi.org/10.1002/qj.2401>
- Holland, M. M., Bailey, D. A., & Vavrus, S. (2011). Inherent sea ice predictability in the rapidly changing Arctic environment of the Community Climate System Model, version 3. *Climate Dynamics*, *36*(7), 1239–1253. <https://doi.org/10.1007/s00382-010-0792-4>
- Kern, S., Lavergne, T., Notz, D., Pedersen, L. T., & Tonboe, R. (2020). Satellite passive microwave sea-ice concentration data set inter-comparison for Arctic summer conditions. *The Cryosphere*, *14*(7), 2469–2493. <https://doi.org/10.5194/tc-14-2469-2020>
- Kern, S., Lavergne, T., Notz, D., Pedersen, L. T., Tonboe, R. T., Saldo, R., & Sørensen, A. M. (2019). Satellite passive microwave sea-ice concentration data set intercomparison: Closed ice and ship-based observations. *The Cryosphere*, *13*(12), 3261–3307. <https://doi.org/10.5194/tc-13-3261-2019>
- Koenigk, T., & Mikolajewicz, U. (2009). Seasonal to interannual climate predictability in mid and high northern latitudes in a global coupled model. *Climate Dynamics*, *32*(6), 783–798. <https://doi.org/10.1007/s00382-008-0419-1>
- Krikken, F., & Hazeleger, W. (2015). Arctic energy budget in relation to sea ice variability on monthly-to-annual time scales. *Journal of Climate*, *28*(16), 6335–6350. <https://doi.org/10.1175/JCLI-D-15-0002.1>
- Lavergne, T., Sørensen, A. M., Kern, S., Tonboe, R., Notz, D., Aaboe, S., et al. (2019). Version 2 of the EUMETSAT OSI SAF and ESA CCI sea-ice concentration climate data records. *The Cryosphere*, *13*(1), 49–78. <https://doi.org/10.5194/tc-13-49-2019>
- Maher, N., Milinski, S., Suarez-Gutierrez, L., Botzet, M., Dobrynin, M., Kornblueh, L., et al. (2019). The Max Planck Institute Grand Ensemble: Enabling the exploration of climate system variability. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2050–2069. <https://doi.org/10.1029/2019MS001639>
- Msadek, R., Vecchi, G. A., Winton, M., & Gudgel, R. G. (2014). Importance of initial conditions in seasonal predictions of Arctic sea ice extent. *Geophysical Research Letters*, *41*(14), 5208–5215. <https://doi.org/10.1002/2014GL060799>
- Notz, D. (2017). Arctic sea ice seasonal-to-decadal variability and long-term change. *PAGES Magazine*, *25*(1), 14–19. <https://doi.org/10.22498/pages.25.1.14>
- Ordoñez, A. C., Bitz, C. M., & Blanchard-Wrigglesworth, E. (2018). Processes controlling Arctic and Antarctic sea ice predictability in the community earth system model. *Journal of Climate*, *31*(23), 9771–9786. <https://doi.org/10.1175/JCLI-D-18-0348.1>
- Peterson, K. A., Arribas, A., Hewitt, H. T., Keen, A. B., Lea, D. J., & McLaren, A. J. (2015). Assessing the forecast skill of Arctic sea ice extent in the GloSea4 seasonal prediction system. *Climate Dynamics*, *44*(1–2), 147–162. <https://doi.org/10.1007/s00382-014-2190-9>
- Serreze, M. C., & Stroeve, J. (2015). Arctic sea ice trends, variability and implications for seasonal ice forecasting. *Philosophical Transactions of the Royal Society A*, *373*(2045), 20140159. <https://doi.org/10.1098/rsta.2014.0159>
- Sigmond, M., Fyfe, J. C., Flato, G. M., Kharin, V. V., & Merryfield, W. J. (2013). Seasonal forecast skill of Arctic sea ice area in a dynamical forecast system. *Geophysical Research Letters*, *40*(3), 529–534. <https://doi.org/10.1002/grl.50129>
- SIMIP Community. (2020). Arctic sea ice in CMIP6. *Geophysical Research Letters*, *47*(10). <https://doi.org/10.1029/2019GL086749>
- Tietsche, S., Day, J. J., Guemas, V., Hurlin, W. J., Keeley, S. P. E., Matei, D., et al. (2014). Seasonal to interannual Arctic sea ice predictability in current global climate models. *Geophysical Research Letters*, *41*(3), 1035–1043. <https://doi.org/10.1002/2013GL058755>
- Wang, W., Chen, M., & Kumar, A. (2013). Seasonal prediction of Arctic sea ice extent from a coupled dynamical forecast system. *Monthly Weather Review*, *141*(4), 1375–1394. <https://doi.org/10.1175/MWR-D-12-00057.1>