

The Importance of Initial Conditions in Seasonal Predictions of Antarctic Sea Ice

Elio Campitelli^{1, 2}, Ariaan Purich^{1, 2}, Julie Arblaster^{1, 2}, Eun-Pa Lim³, Matthew C. Wheeler³, and Phillip Reid³

¹School of Earth, Atmosphere and Environment, Monash University, Kulin Nations, Clayton, Victoria, Australia.

²ARC Special Research Initiative for Securing Antarctica's Environmental Future, Clayton, Kulin Nations, Victoria, Australia

³Research, Bureau of Meteorology, Melbourne, Australia.

Correspondence: Elio Campitelli (elio.campitelli@monash.edu)

Abstract. Accurate Antarctic sea-ice forecasts are crucial for climate monitoring and operational planning, yet they remain challenging due to model biases and complex ice-ocean-atmosphere interactions. The two versions of the Australian Bureau of Meteorology's ~~ACCESS~~ seasonal forecast system, ACCESS-S1 and ACCESS-S2, use identical model configuration ~~and but~~ differ only in their initial conditions; primarily in that ACCESS-S2 does not assimilate
5 sea-ice observations, whereas ACCESS-S1 does. This provides a convenient opportunistic experiment to assess the role of initial conditions on Antarctic sea-ice forecasts using more than 20 years of fully coupled simulations with two ~~9-member~~ nine-member ensembles. Our analysis reveals that both systems experience an extended melt season and delayed growth phase compared with observations. This leads to a significant negative sea-ice extent bias, which is ~~corrected only~~ reduced in ACCESS-S1 by the data assimilation system. The impact of the differing initial conditions
10 on forecast errors varies dramatically by season: summer and autumn ~~initial conditions~~-(January-April) sea-ice initial conditions provide predictive skill for up to three months, with February initial conditions being particularly crucial. In contrast, winter forecasts of the two systems are statistically indistinguishable after only two weeks. Regional analysis of forecast skill suggests that this winter ~~predictability~~ predictive skill barrier is most dramatic over East Antarctica, where even ACCESS-S1 shows negative skill. These findings highlight the critical importance
15 of comprehensive year-round sampling in predictability studies and suggest that operational sea-ice data assimilation efforts should prioritise the summer-autumn period when initial conditions have maximum impact on forecast skill.

1 Introduction

Accurately modelling Antarctic sea ice is essential for understanding climate processes and improving climate projections to inform adaptation strategies. Accurate ~~seasonal to~~ sub-seasonal to seasonal forecasts are also crucial
20 for operation contingency planning in and around the Antarctic continent, including scientific missions, fisheries, and

tourism (De Silva et al.; Wagner et al.). Improvements in modelled ~~sea-ice~~ sea ice might also help improve weather forecasts over and away from sea-ice regions (Rinke et al.; Wang et al., b; Semmler et al.).

However, progress in Antarctic sea-ice forecasting ~~system~~ has lagged behind Arctic sea-ice ~~forecasts~~ forecasting due to model biases ~~,~~ and inherent large variability and complexity (~~Zampieri et al.; ?~~) (Zampieri et al.; Gao et al.)
25 . Dynamical seasonal forecasts of summer Antarctic sea ice have been shown to perform worse than relatively simple ~~er~~ statistical methods (Massonnet et al.) and machine learning approaches (e.g. Dong et al., Lin et al.), which also underscores the need for better understanding and physical modelling of sea-ice dynamics ~~,~~ and drivers of its variability.

Good initial conditions are ~~generally required~~ an essential element for a good forecast~~,~~; however, it is not entirely
30 known to what extent accurate sea-ice initial conditions affect the quality of the forecast and at what timescales. Exploring seasonal predictions of Arctic sea ice, Guemas et al. found that sea-ice initial conditions are important in autumn to predict summer Arctic sea ice, but the impact ~~wasn't~~ was not as dramatic when predicting winter Arctic sea ice. Day et al. also found seasonally-varying differences in the effect of initialisation, noting that accurate Arctic sea-ice thickness leads to improved sea-ice forecasts initialised in July but not when initialised in January.

For the Antarctic, Holland et al. studied the initial-value predictability of Antarctic sea ice in a perfect model study using the CCSM3-Community Climate System Model Version 3 model. They found that sea-ice and ocean initial conditions provide predictive information to forecast sea-ice edge location several months in advance and that some predictability is retained for up to two years thanks to ocean heat content anomalies that are advected eastward. Similarly, Morioka et al. studied decadal forecasts of Antarctic sea ice and found that initialising
40 ocean and sea ice improved the correlation between simulated and observed sea-ice concentration evolution in the Amundsen-Bellingshausen Sea. This is in contrast with Marchi et al., who ran perfect model experiments to argue that uncertainty in the predicted atmospheric state and evolution is the main driver of uncertainty in Antarctic sea-ice extent prediction on seasonal timescales, with sea-ice and ocean initial conditions having lesser importance. More recently, ~~Morioka et al. studied decadal forecasts of~~ Xiu et al. contrasted Antarctic sea ice ~~and found that initialising ocean~~
45 ~~and sea ice improved the correlation between simulated and observed~~ hindcasts based on different data assimilation sources. They concluded that hindcasts initialised only with atmospheric data lead to good predictions of sea-ice area in many regions that were comparable, if not better, than hindcasts initialised only with ocean data or with ocean and sea-ice concentration evolution in the Amundsen-Bellingshausen Sea. ~~data. Furthermore, they find that assimilation of sea-ice concentration data offered little benefit on top of assimilation of ocean data.~~

It is hard to compare these studies since they are based on forecasts initialised at different times of the year and different frameworks~~,~~. Holland et al. ran 20 ensemble members initialised on the 1st of January of a particular year, Marchi et al. ran forecasts from the 1st of March and 1st of September, and Morioka et al. ran forecasts only from the 1st of March. Xiu et al. was the most comprehensive study, using hindcasts initialised on the 1st of January, April,

July and October. Marchi et al. also used a coupled ocean–sea-ice model instead of a fully coupled model like Holland et al. and Xiu et al. did. ~~Morioka et al. used observed sea-ice initial conditions and compared and~~ Xiu et al. both compared forecast with observations, ~~while~~ but the former used observed sea-ice initial conditions while the latter used initial conditions derived from data assimilationand. On the other hand, Marchi et al. and Holland et al. were perfect model studies.

In October 2021 the Australian Bureau of Meteorology (BoM) upgraded the Australian Community Climate and Earth System Simulator – Seasonal (ACCESS-S) from version S1 to S2 (Wedd et al.). While the base model remained the same, the change in version was focused on using ocean, sea-ice and land initial conditions generated by the BoM instead of depending on the UK Met Office. Crucially, compared to ACCESS-S1, ACCESS-S2 does not assimilate sea-ice observations, so sea ice is only affected by the ocean and atmospheric data assimilation via the coupled integration.

Since model configuration is identical between ACCESS-S1 and ACCESS-S2, they form ~~a sort of an~~ “opportunistic experiment” where the same forecasting model was run over a long period of time with multiple ensemble forecasts initialised throughout the year, with the only difference being the initial conditions. This provides an opportunity to test the effect of sea-ice initial conditions on the forecast of sea-ice concentrations and the climate.

In this study we compare sea-ice hindcasts produced by ACCESS-S1 and ACCESS-S2. We focus on seasonality of errors and biases and the effect of the data assimilation system. This comparison will inform future work with the prediction system as a research tool to better understand the dynamics and variability of the Antarctic sea ice and its impacts on the climate system as well as to explore the potential of using its sea-ice forecasts for decision-making. The work will also serve as a benchmark for future prediction systems to attempt to improve upon.

2 Data and methods

75 2.1 ~~ACCESS-S2~~Forecasting systems

ACCESS-S2 (Wedd et al.) is the ~~Bureau of Meteorology~~BoM’s seasonal forecast system which became operational in October 2021, replacing the ACCESS-S1 system (Hudson et al.). The model components of both ACCESS-S2 and ACCESS-S1 are identical with the same numbers of levels and resolution. They consist of the Global Atmosphere 6.0 (GA6) (Williams et al.; Waters et al., a), the Unified Model’s Global Land 6.0 (Best et al.; Waters et al., a), NEMO 80 Global Ocean 5.0 (Gurvan et al.; Megann et al.) and Global Sea Ice 6.0 [CICE; ~~?~~Rae et al.]. The atmosphere has a N216 horizontal resolution (~60 km in the mid-latitudes) with 85 vertical levels. The land model uses the same horizontal grid as the atmosphere with four soil levels. The ocean component has a nominal horizontal resolution of $1/4^\circ$ with 75 vertical levels. The sea-ice component, based on CICE version 4.1, has the same resolution as the ocean component and five sea-ice thickness categories as well as an open water category.

85 Both systems take atmospheric initial conditions derived from ERA-interim (Dee et al.) for their hindcasts. The ~~main-only~~ difference between the hindcasts of the two systems are the ocean and sea-ice initial conditions ~~and the~~ ensemble generation procedure.

ACCESS-S1's ocean and sea-ice initial conditions come from the Met Office FOAM system, which uses a multivariate, incremental three-dimensional variational (3D-Var), first-guess-at-appropriate-time (FGAT) data assimilation scheme
90 (Waters et al., b) and assimilates sea surface temperature (SST), sea surface height (SSH), in situ temperature and salinity profiles, and satellite observations of sea-ice concentration using the EUMETSAT OSISAF product described in the next section. If sea-ice concentration innovations are positive, these are added to the first ice category with a fixed thickness of 50 cm. If they are negative, they are removed from the thinnest category first and from thicker categories if needed. Therefore, although this system only assimilates sea-ice concentration data, sea-ice thickness is
95 modified.

ACCESS-S2, on the other hand, is initialised from ocean conditions generated by the BoM weakly coupled ensemble data assimilation scheme described in Wedd et al.. This scheme uses an optimal interpolation method and assimilates temperature and salinity profiles from EN4 (Good et al.). SSTs are nudged to Reynolds OISSTv2.1 (Reynolds et al.) in areas where SSTs are over 0°C and ~~Sea Surface Salinity~~ sea surface salinity is weakly nudged to the World Ocean
100 Atlas 2013 climatology (Zweng et al.).

Of most relevance for this work, sea-ice concentrations are not assimilated in ACCESS-S2. Assimilation cycles are performed daily. The coupled model runs for 24 hours initialised from the previous cycle. Then the restart file fields of the ocean component are used as first guess in the data assimilation cycle and the innovations are used to build the next ocean initial conditions for the following cycle. The atmosphere fields from that daily integration are not
105 used and instead the model atmosphere is initialised using ERA-Interim. The sea-ice initial conditions for the next cycle are the unaltered output of the previous daily integration. Then the cycle starts again and the coupled model runs for another 24 hours. During this integration the sea-ice component is affected by the ocean innovations and the new atmosphere initial conditions via the coupler.

The ACCESS-S1 hindcast set is made up of nine members created by perturbing the atmospheric fields only with a
110 random field perturbation (Hudson et al.) and runs for 217 days for the period 1990–2012 initialised at the first of every month. The ACCESS-S2 hindcast set used in this study runs for the period 1981–2018. Ensemble members are created in the same manner as ACCESS-S1 members~~;~~ however, due to computing cost limitations, only three members per forecast initialisation date were run for 279 days. Bigger ensembles were generated by aggregating ~~several~~ three-member ensembles initialised on successive days (Wedd et al.). Here, we build a nine-member time-lagged
115 ensemble from three consecutive three-member forecasts initialised at the first of every month and the two previous days and run for 279 days. We analyse the ensemble mean hindcasts unless otherwise specified and, although the

hindcast periods differ for each forecasting system, we only use data from the overlapping period (1990–2012) to compute error measures, this excludes the recent period of low sea-ice extent for our analysis.

120 Anomalies for each hindcast set are taken with respect to their own climatology specific to each initialisation date and forecast lead time, for the period 1990–2012. This serves as a first-order correction of model bias and drift. For monthly means, we define “0 lead time months” as the monthly mean ~~forecast of the same month of forecasts~~ immediately after the initialisation.

Besides sea-ice concentration, we also analyse mean sea-ice thickness, which we compute as total sea-ice volume divided by total sea-ice area.

125 2.2 Verification datasets

For verification we use satellite-derived sea-ice concentration, which estimates the proportion of each grid area that is covered with ice. ~~Datasets~~ Different datasets are derived using different algorithms and satellite platforms, and each have their own biases and uncertainties. Estimates of inter-product uncertainty of sea-ice extent (~~SIE, defined here defined~~ as the total region of the Southern Ocean with at least 15% sea-ice cover (Cavalieri et al., a)) are of the 130 order of 0.5 million km^2 (Meier and Stewart). As will be shown below, this spread is minimal compared with the typical errors in the ACCESS-S2 and ACCESS-S1 forecasts, so the overall conclusions of this study are independent of the verification dataset used.

We use NOAA/NSIDC’s Climate Data Record V4 [CDR; Meier et al. (b)] as the primary sea-ice verification dataset. It takes the maximum value of the NASA Team (Cavalieri et al., b) and NASA Bootstrap (Comiso) sea-ice concentration 135 products to reduce their low concentration bias (Meier et al., b, a). Both source algorithms use data from the Scanning Multichannel Microwave Radiometer (SMMR) on the Nimbus-7 satellite and from the Special Sensor Microwave/Imager (SSM/I) sensors on the Defense Meteorological Satellite Program’s (DMSP) -F8, -F11, and -F13 satellites. The data have a spatial resolution of 25 by 25 km and daily from November 1978 onwards.

The European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice Satellite 140 Application Facility [OSI; Facility] based on the SSMIS sensor is another satellite-derived sea-ice concentration product. It is based on mostly the same sensors as the NOAA CDR but computed independently using different algorithms. Figures prepared with this dataset are provided in the appendix and do not differ significantly from the ones prepared using CDR.

2.3 Error measures

145 For evaluation purposes, we use a series of measures. ~~Sea-ice extent~~ We evaluate sea-ice extent anomaly hindcasts by their Root Mean Squared Error (RMSE) and correlation coefficient. The hindcast ensemble mean sea-ice extent is defined as the area of the ocean at least 15% covered by mean sea-ice ~~. This threshold is motivated by the limitations~~

~~in satellite retrieval, which is increasingly unreliable for lower sea-ice concentrations (Cavalieri et al., a).~~ extent of individual ensemble members.

150 Pan-Antarctic (net) sea-ice extent serves as a hemispheric measure of the amount of sea ice, but it does not take into account the spatial distribution. A model could have a relatively accurate extent of the net ice but with different regional distributions. To account for location errors, we computed the ~~Root Mean Squared Error (RMSE)~~ RMSE of grid-point sea-ice concentration anomalies.

We compute RMSE as the square root of the area-averaged squared differences between grid-point forecasted
155 and observed sea-ice concentration anomalies. We compute a pan-Antarctic RMSE by averaging over the whole NOAA/NSIDC CDRV4 Southern Hemisphere domain, and also a zonally-varying RMSE computed over 15 longitude slices 24° wide around Antarctica.

All error measures were computed on the NOAA/NSIDC CDRV4 domain grid, to which model output was bilinearly interpolated. Note that the ACCESS CICE model grid has resolution between two and three times higher than
160 NOAA/NSIDC CDRV4.

Forecast errors are also compared with hypothetical forecasts based on the persistence of anomalies and on climatology. The persistence forecast is generated by extending the observed sea-ice concentration anomalies from the day of the forecast initialisation and comparing it with the actual anomalies observed. The climatological forecast error is computed as the standard deviation of daily anomalies.

165 As a measure of forecast improvement over the hypothetical forecast, we use the skill score (Murphy and Daan), defined as:

$$S = 1 - \frac{RMSE_f}{RMSE_r}$$

$$S = 1 - \frac{RMSE_f}{RMSE_r} \tag{1}$$

170 Where $RMSE_f$ is the RMSE of the forecast, $RMSE_r$ is the RMSE of the reference forecast. Negative skill score indicates that the forecast is worse than the reference forecast while positive values indicate an improvement. A perfect forecast would have zero RMSE and thus a skill score of 1. Furthermore, we also compute sea-ice concentration bias, defined as the mean difference between the forecast and observations.

2.4 Computational procedures

175 We performed all analyses in this paper using the R programming language (R Core Team), using data.table (Dowle and Srinivasan) and metR (?)(Campitelli, a) packages. Significant processing was performed using the CDO

command line operators (Schulzweida). All graphics are made using ggplot2 (Wickham). The paper was rendered using knitr and Quarto (Xie; Allaire et al.).

3 Results and discussion

180 3.1 Bias

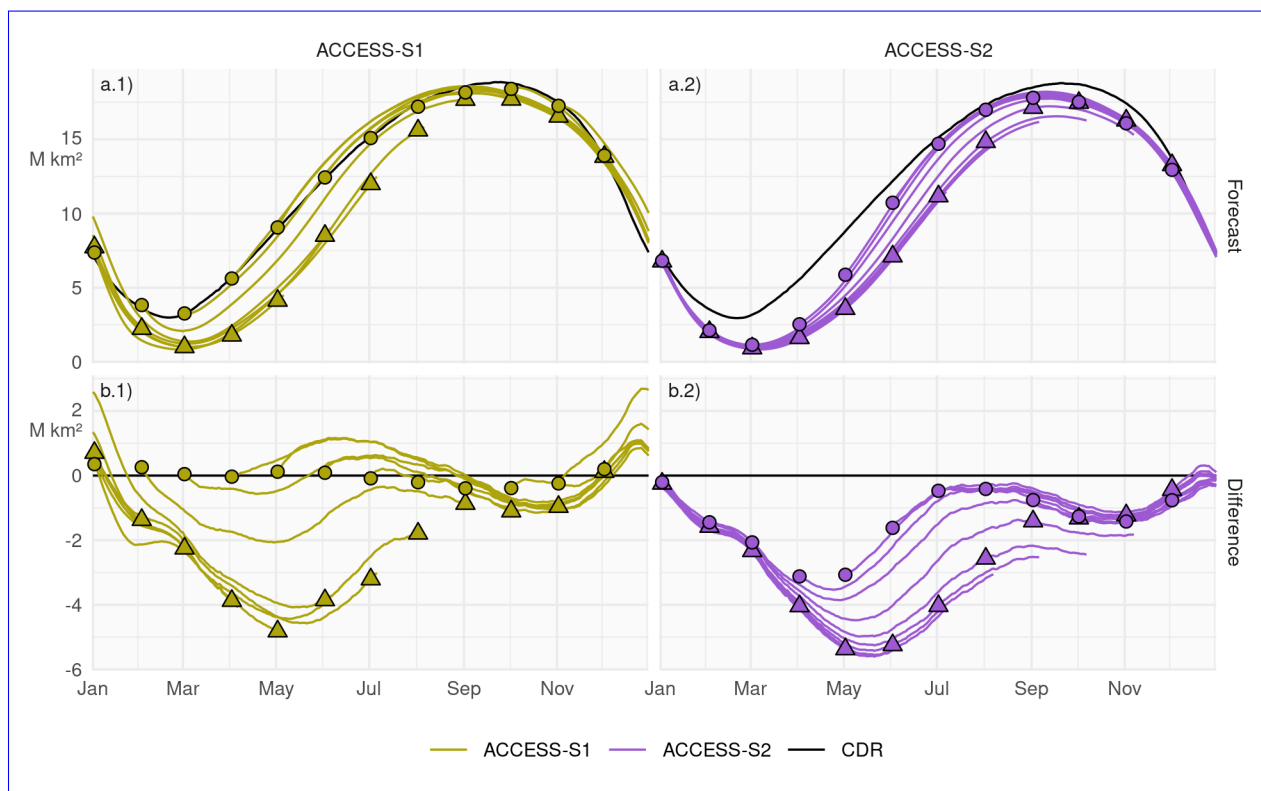


Figure 1. Row a: Pan-Antarctic daily mean sea-ice extent for all hindcasts initialised on the first of each calendar month for ACCESS-S1 (column 1; green) and ACCESS-S2 (column 2; purple). Observed mean sea-ice extent in each corresponding hindcast period is shown in black. Row b: Mean differences between the forecast and the observed values. Circles represent the initial conditions at the start of forecasts (i.e., the first of every month), and triangles represent the mean values [forecasted for the first of every month](#) at the lead time corresponding to the maximum lead time in S1 (between 213 and 216 days, depending on the month).

Figure 1 (Figure A1 for OSI) shows mean sea-ice extent of the ACCESS-S1 and ACCESS-S2 hindcasts (row a) and their differences from mean sea-ice extent of NOAA/NSIDC CDRV4 (row b). Mean extent at the first of every month is indicated with circles for the initial conditions and with triangles for the [leadtime corresponding to the longest lead time possible for each model \(between 274 and 277 days for ACCESS-S2 and between ACCESS-S1 \(between](#)

185 213 and 216 ~~days for ACCESS-S1~~). At this long lead time, information of the initial conditions is essentially lost and the forecast reverts close to each model's preferred equilibrium state.

ACCESS-S2 initial conditions (circles in Fig. 1 column 2) show an overall negative bias, especially ~~in the late summer-early autumn~~ between late summer and early winter, while ACCESS-S1 initial conditions (circles in Fig. 1 column 1) are very close to observations, as expected from the assimilation of sea-ice observations to produce
190 the initial conditions of ACCESS-S1. Both systems' equilibrium states (triangles) show negative biases of sea-ice extent, particularly in the growth phase of late-autumn and winter months. This is due primarily to the melt season being longer in the model than in observations and with faster melt between January and March and the growing seasons being shorter with slower growth during March and April. This is then followed by faster growth between May and July (Figure 2 and Figure A2). Many sea-ice models exhibit this systematic underestimation during the
195 sea-ice minimum and early freezing season (~~Massonnet et al.~~) (Bushuk et al.; Massonnet et al.), which could indicate problems in the representation of thermodynamics in the model (Zampieri et al.). It is also not surprising that both forecasting systems converge to a similar equilibrium state because they share the same model formulation.

The difference between the initial conditions (circles) and the model equilibrium state (triangles) can be mostly attributed to the effect of data assimilation, which in ACCESS-S2 is due solely to the coupling of ~~sea-ice-sea ice~~
200 with the atmosphere and the ocean. From ~~May to October~~ April to September, in ACCESS-S2, circles are closer to observations than the triangles are, indicating that the information from the ocean and atmosphere data assimilation is affecting sea ice and improving the initial conditions. During these months, ACCESS-S1 can overestimate the sea-ice extent at short lead time. For the rest of the year circles are overlaid with triangles in ACCESS-S2, indicating that the ocean and atmosphere data assimilation is not affecting sea ice and that this component of the model is
205 virtually free-running.

To further understand the bias in ACCESS-S2, Figure 3 (Figure A3) shows spatial patterns of the differences of monthly mean sea-ice concentrations between NOAA/NSIDC CDRV4 and ACCESS-S2 hindcasts at the shortest monthly lead time. From October to May, the model underestimates sea-ice concentrations in most regions except for the inner Weddell Sea in April and May, where sea-ice concentrations saturate to 1 both in the observations and
210 forecasts. In winter, the differences are limited to a narrow band around the sea-ice edge with slight positive biases in the African sector of East Antarctica and negative biases around the Indian Ocean sector which partially compensate, resulting in the near-zero extent bias seen in those months (Figure 1).

ACCESS-S1 has a comparatively smaller overall bias (Figure 4 and Figure A4). The largest values are found between April and June, when the faster growth results in large positive bias along the sea-ice edge, and in January, when the
215 faster melt leads to large negative bias in the Weddell and Amundsen-Bellingshausen Seas.

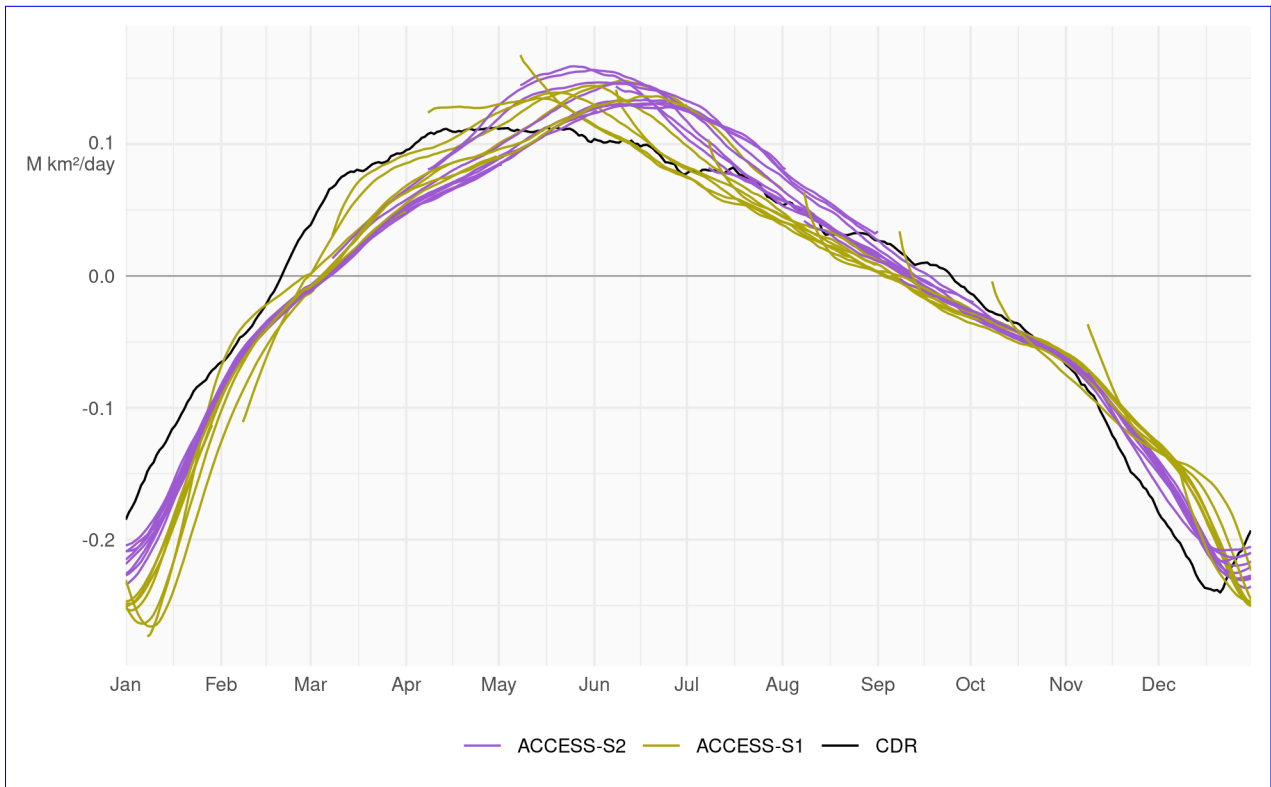


Figure 2. Mean daily sea-ice extent growth ($10^6 \text{ km}^2/\text{day}$) in ACCESS-S1 (green) and ACCESS-S2 (purple) hindcasts and observations (black), computed as the mean daily differences in sea-ice extent between each date and the next for each forecast month. Values are smoothed with a 11-day running mean.

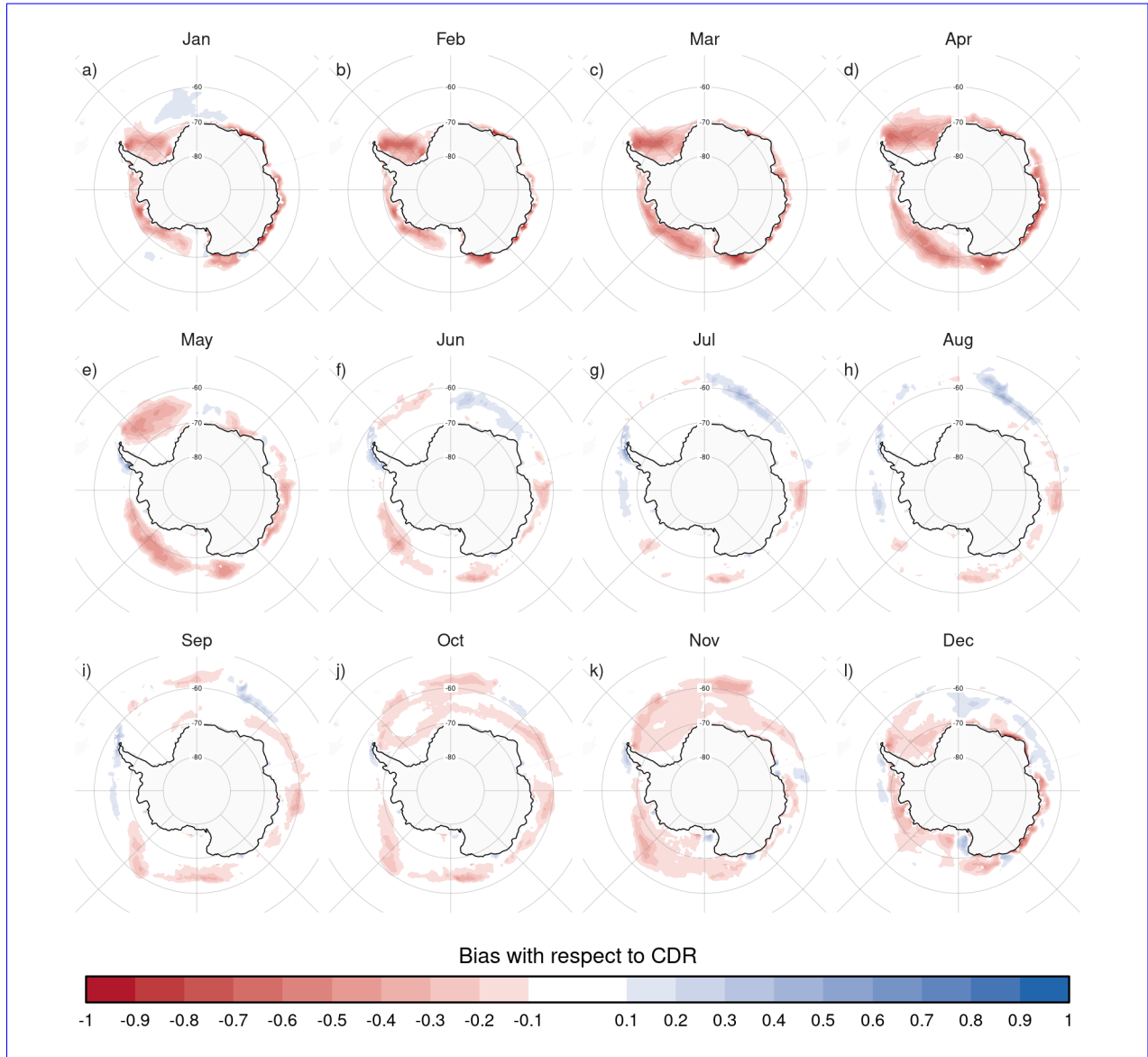


Figure 3. Ensemble mean difference between monthly sea-ice concentration of ACCESS-S2 ensemble mean forecast at 0-month lead time (monthly mean values forecasted from the forecast initialised at the first of the month) and observations (CDR).

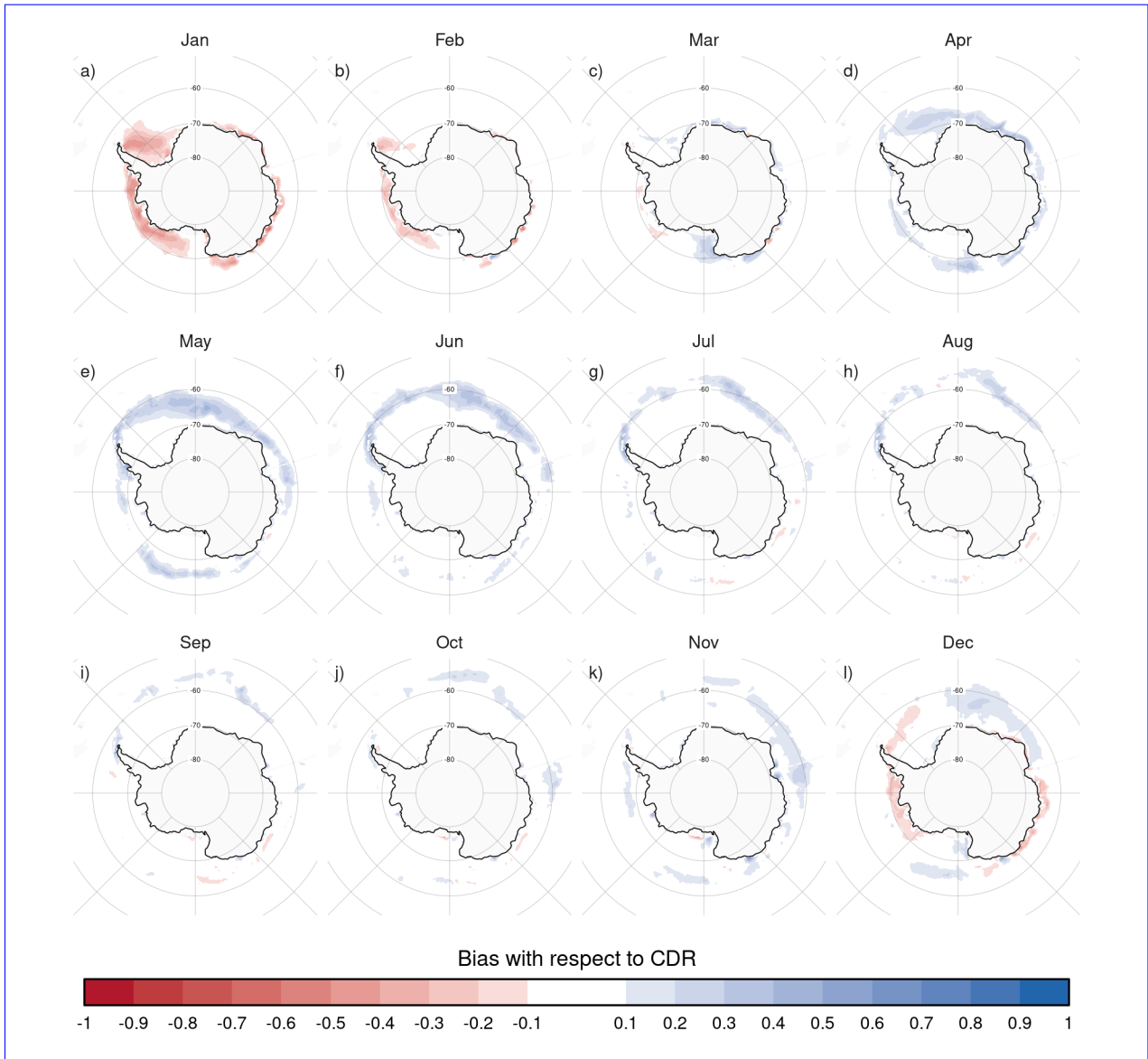


Figure 4. Same as Figure 3 but for ACCESS-S1.

3.2 RMSE Anomaly errors

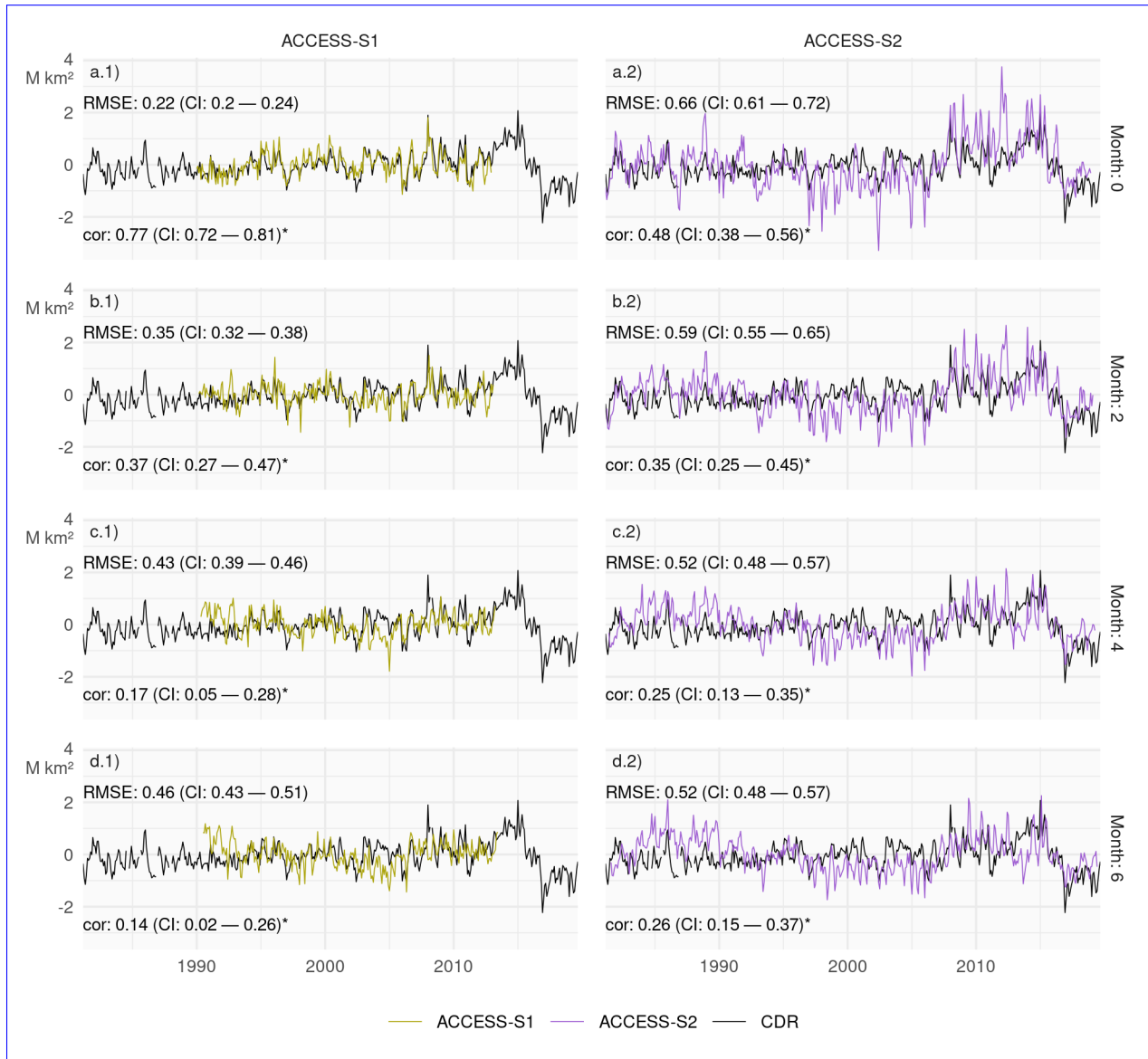


Figure 5. Monthly mean sea-ice extent anomalies of the observations (black) and forecasts from ACCESS-S1 (right column; purple) and ACCESS-S2 (left column; green) at lead times of 0, 2, 4, and 6 months. The RMSE and correlation [with their respective 95% confidence interval](#) during the overlapping period of ACCESS-S1 and ACCESS-S2 hindcasts (1990–2013|1990–2012) are shown on the top left and bottom left of each panel respectively. [Statistically significant correlations at 95% confidence level are marked with an asterisk.](#)

Figure 5 (Figure A5) shows monthly sea-ice extent anomalies forecasted at selected lead times. Compared with ACCESS-S1, ACCESS-S2 anomaly forecasts are relatively poor (large RMSE) even for the first month (lead time 0), ~~whereas which is worse than ACCESS-S1 forecasts stay relatively skilful~~ even at a lead time of ~~three six~~ months. ACCESS-S2 shows much larger interannual variability than observations, with dramatic lows between 1995 and 2007, and highs between 2007 and 2015.

Unexpectedly, for ACCESS-S2, RMSE improves with lead time, even though the correlation degrades with lead time. This ~~is effect is seen in all months except from July to September (Figure A6). This is~~ puzzling behaviour that goes contrary to what is usually seen in prediction models. The explanation seems to be the ~~mentioned increased~~ ~~aforementioned increase in~~ interannual variability. Figure 6 (Figure A7) shows the interannual standard deviation of monthly sea-ice extent of the forecasts as a function of lead time compared with observations. ACCESS-S1 standard deviation lies within the observed standard deviation regardless of lead time, while ACCESS-S2 standard deviation is more than twice that of observations at zero lead time and only approaches the observed value at nine months lead time for most months.

~~The reason for the increased variability in ACCESS-S2 forecasts of sea-ice extent anomalies seem to align moderately well with observations (leading to moderately high correlation) but their magnitude is overestimated (leading to large errors). This could be caused by is not clear. One possible explanation is that ACCESS-S2 sea-ice being much simulates thinner sea ice than ACCESS-S1 (Figure A bottom row) that is more sensitive to atmospheric and oceanic forcing, perhaps due to lower thickness.~~

~~As an example, Figure A shows . To explore this possibility, Figure Figure 7 shows the correlation between sea-ice concentration anomalies (top row) and sea-ice thickness and the difference between the two models (bottom row) for 2 May 2008 initialised one day prior; being that close to initialisation date, these are very approximately the initial conditions. ACCESS-S1 magnitude of sea-ice concentrations anomalies are very close to observations as expected from the system assimilating these data. ACCESS-S2 sea-ice concentration anomalies, which are not assimilated, are not as close, but the large-scale pattern is aligned with observations. The system simulates large positive anomalies in the Weddell and Ross Seas and slight negative anomalies in the Amundsen and Bellingshausen Seas. The fact that ACCESS-S2 can simulate this pattern without assimilating sea-ice data suggests that atmospheric and oceanic forcing were the dominant drivers. However, the magnitude of the sea-ice concentration anomalies is too big. It is plausible that this is due to the thinner ice simulated by in the ACCESS-S2 (bottom row). reanalysis. It shows that in most regions and months, thinner ice is correlated with stronger sea-ice concentration anomalies.~~

~~Extending beyond the one case in Figure A, Figure 8 shows monthly mean sea-ice thickness as a function of lead time for ACCESS-S1 and Figure 8 shows that indeed ACCESS-S2 . Supporting the idea that thinner ice is what causes the increased extent variability in ACCESS-S2, this system simulates thinner sea-ice simulates thinner sea ice compared to ACCESS-S1 overall at almost all lead times and in all months except for summer at short lead times (Dec-Jan, 0-1~~

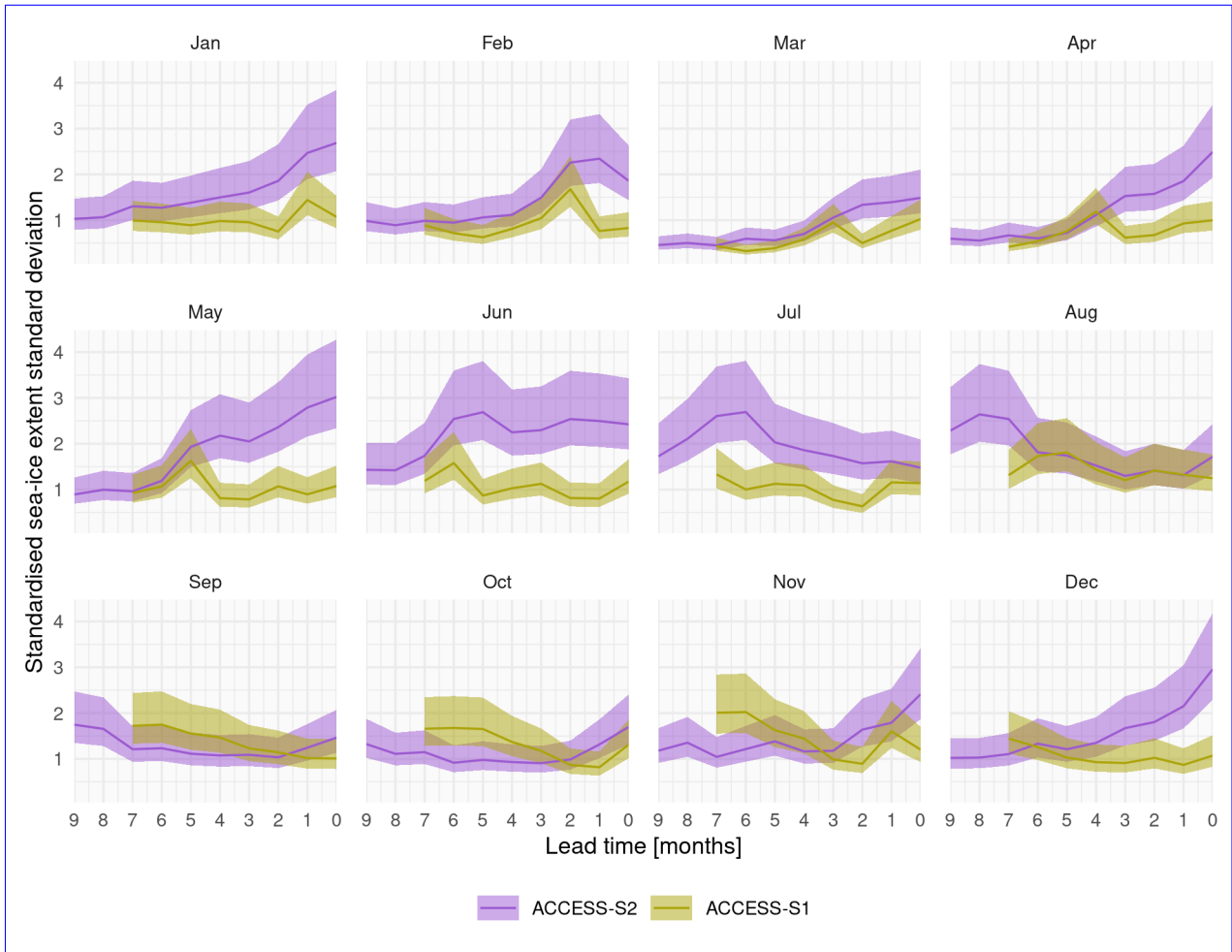


Figure 6. Interannual standard deviation with 95% confidence interval of monthly mean sea-ice extent forecasted for each month ~~divided~~. We standardise the standard deviation by that month's sea-ice extent observation standard deviation. ACCESS-S1 and ACCESS-S2 at different lead times. Each panel indicates the target month. Note the reverse horizontal axis.

250 months; Feb-Mar, 0-2-Dec-Jan, 0-1 months; Feb-Mar, 0-2 months). However, in both systems, forecasted ~~Figure 8~~ also shows that sea-ice is thicker at shorter lead times and then decreases, particularly in the summer months. If thinner ice were a sufficient cause thickness tends to decrease at longer lead times, when ACCESS-S2's variability converges to observed variability. Furthermore, the thickness difference between ACCESS-S1 and ACCESS-S2 is large also in months in which ACCESS-S2 sea-ice extent variability is comparable to ACCESS-S1 and where ACCESS-S2's

255 RMSE does not decrease with lead time, such as August and September.

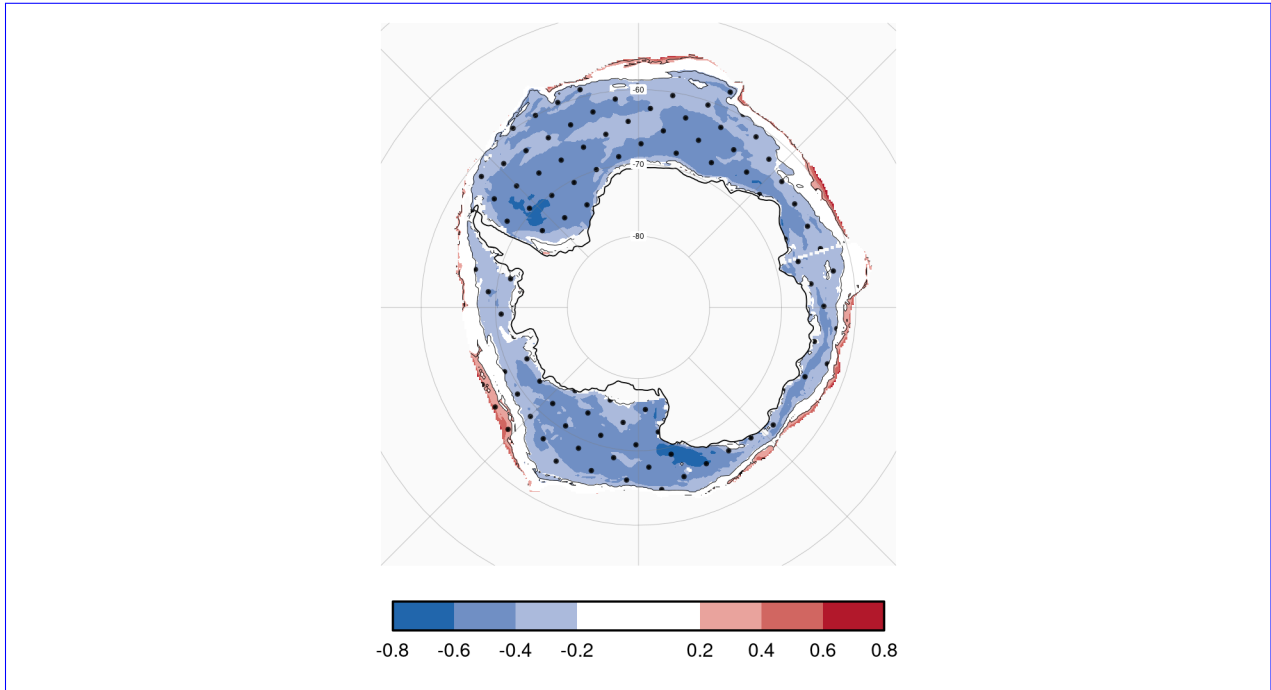
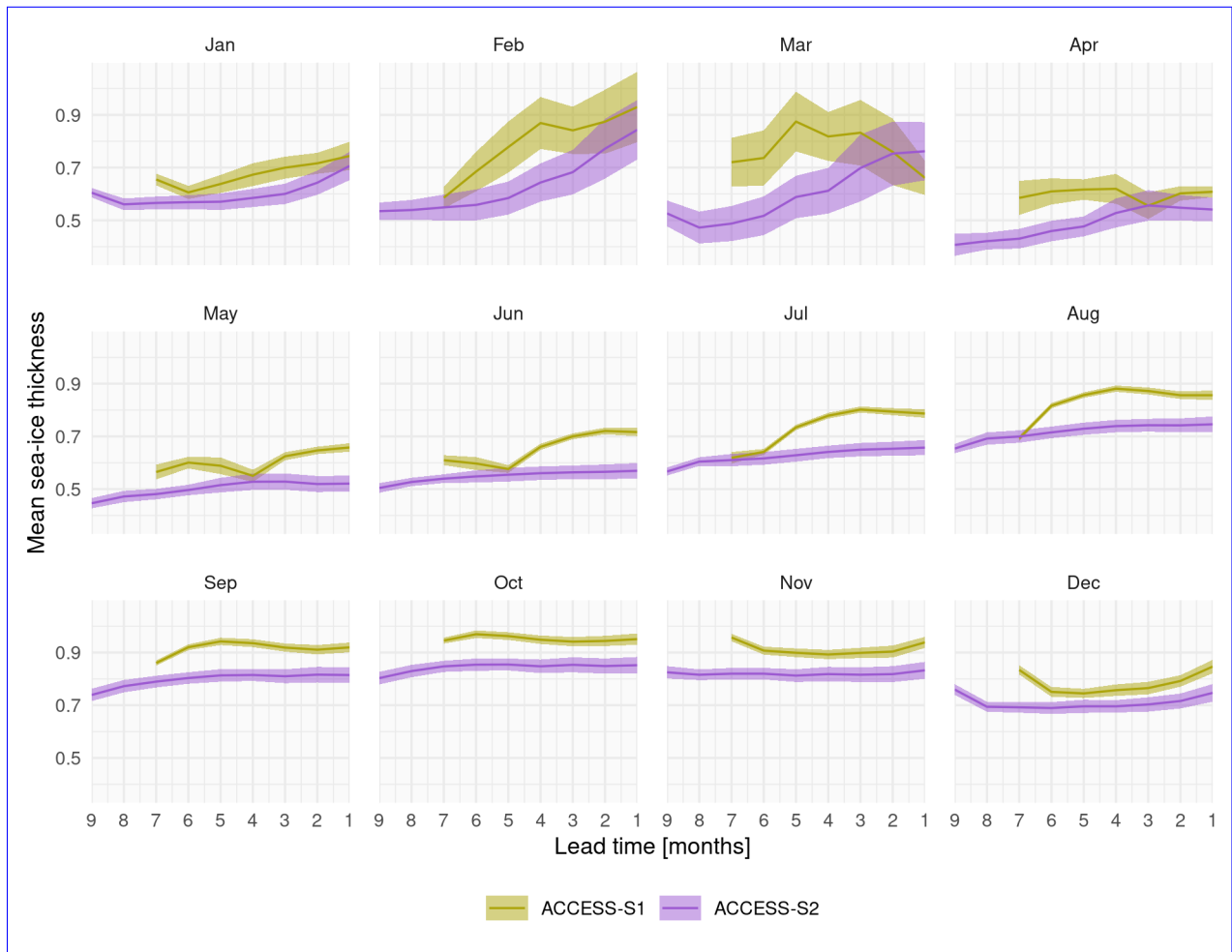


Figure 7. Correlation in the ACCESS-S2 reanalysis between monthly mean sea-ice thickness and the absolute value of monthly mean sea-ice concentration anomalies masked to areas with mean sea-ice concentration greater than 15% more than 50% for each month. Regions where the correlation is significant at p -value < 0.05 are stippled.

Even if ACCESS-S2's sea ice is more sensitive to atmospheric and oceanic forcings than ACCESS-S1 by being overall thinner, this is not a sufficient explanation for the pattern of increased variability ~~then we would expect variability to increase with lead time in both forecasting systems.~~

at shorter lead times and its absence in particular months. Instead, the strengths of the atmospheric and oceanic forcings might also be playing a role. The fact that ACCESS-S1 and ACCESS-S2 share the same model configuration and that the increased variability is more extreme at short lead times (Fig. 6) suggests that the data assimilation procedure ~~is~~ might also be partly responsible. ~~It~~ We speculate that it is possible that ~~sea ice~~ sea ice in the ACCESS-S2 system is left in an unbalanced state ~~after~~ by assimilating atmospheric and oceanic data but not sea-ice data, leading to large responses that are amplified by the ~~thin ice in the initial states which then subside at longer lead times when the model is balanced~~ thinner ice. Teasing this out would require an in-depth analysis of the data assimilation innovations, which is out of scope for this work.

To assess ACCESS-S2 forecasts in more detail, we compute error measures for all hindcasts started on the 1st of every month. Figure 9 (Figure A9) shows the mean RMSE of sea-ice concentration anomalies for ACCESS-S1 and ACCESS-S2 hindcasts compared against persistence and climatological forecasts used as a benchmark. Due to errors



ACCESS-S1 and ACCESS-S2 hindcasts for 2-May-2008 at one day lead time. Top row shows sea-ice concentration anomalies forecasted by each system and the observations. Bottom row shows forecasted sea-ice thickness and the difference between ACCESS-S1 and ACCESS-S2.

Figure 8. Mean and 95% interval of monthly mean sea-ice thickness for ACCESS-S1 and ACCESS-S2 at different lead times. Each panel indicates the target month. Note the reverse horizontal axis.

270 in the initial conditions, it is expected that persistence forecasts would be better than the model forecasts at very short lead times, but that the persistence forecast errors would grow faster and may eventually surpass the model forecast errors. The black line shows that the persistence forecast error indeed grows rapidly and reaches its maximum in about 30 days for most months except for February, when it grows much slower. The ACCESS-S1 forecast errors grow slower than persistence forecast errors and remain lower after less than 10 days on average. The ACCESS-S2
 275 forecast error starts high in all months and is lower than the persistence forecast error after more than 15 days in most months except for forecast initialised in February, when it takes 80 days.

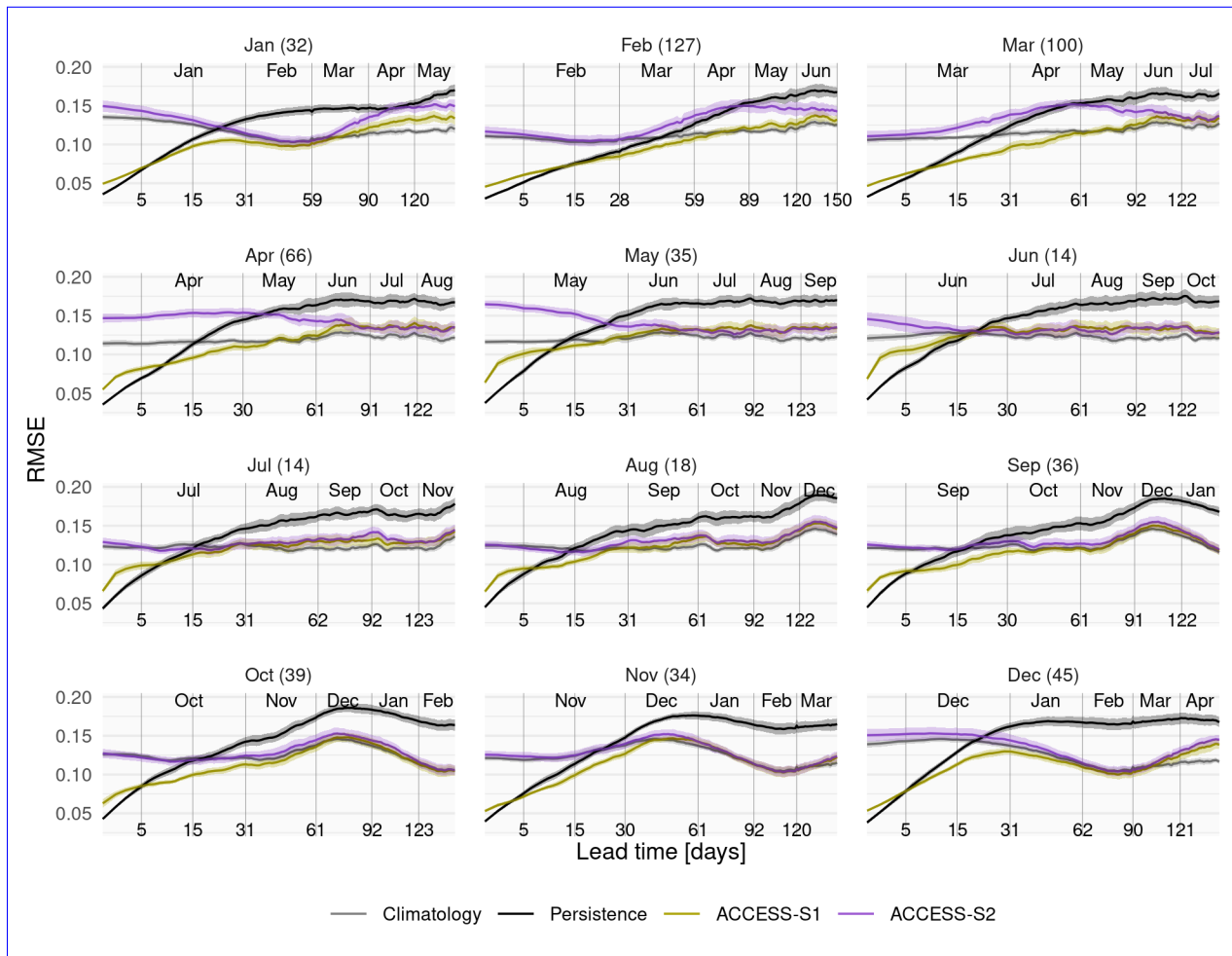


Figure 9. Mean RMSE of sea-ice concentration anomalies as a function of forecast lead time for all forecasts initialised on the first of each month compared with a reference forecast of persistence of anomalies (black) and climatology (gray). [The shading indicates the 95% confidence interval of the mean.](#) Only the first ~~120~~¹⁵⁰ days are shown. In parentheses, the shortest time at which ACCESS-S1 and ACCESS-S2 mean RMSE is not statistically different at the 99% confidence level. [Note that the horizontal axis uses a square root transformation to expand the shorter lead times.](#)

At longer lead times, it is more appropriate to compare errors with the climatological forecast error. The lead time at which ACCESS-S1 forecast error is higher than the climatological forecast error varies between more than 60 [days](#) and less than 20 days depending on forecast initialisation month with the minimum [length](#) in June. ACCESS-S2 forecasts never have lower error than climatology, on the other hand, except marginally in October forecasts.

~~Figure A10 (Figure A11) summarises the lead time window in which each hindcast is better than both the persistence forecast and the climatological forecast as a function of forecast month.~~ [In summary,](#) ACCESS-S1 forecasts have a

285 ~~wider lead time window in longer lead time in which it is skillful in~~ the summer than the other seasons ~~and is not better than both benchmarks at forecasting June sea-ice concentration anomalies.~~ Forecasts initialised ~~in May and June~~ between May and July are particularly poor, and ~~July-June~~ July-June cannot be forecasted better than the benchmarks, ~~suggesting a winter barrier in predictive skill (Figure A10 and Figure A11).~~ This is ~~consistent with similar to~~ the mid-winter loss of predictability observed by Libera et al. for the Weddell Sea, who attributed it to deep warm water entraining into the mixed layer.

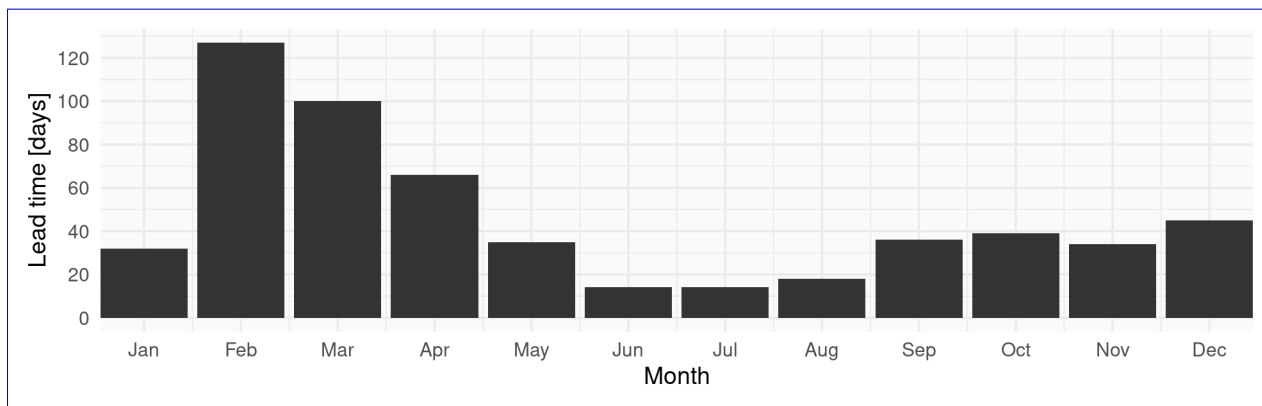


Figure 10. ~~Minimum Shortest~~ lead time at which ~~each forecast's mean RMSE becomes larger than the lower bound of the 95% confidence interval of persistence forecast RMSE (black lines) ACCESS-S1 and maximum lead time at which each forecast's ACCESS-S2 mean RMSE remains lower than is not statistically different at the lower bound of the 9599% confidence interval of climatological forecast RMSE (gray lines) level.~~ Green shading indicates the window where forecasts outperform both persistence (lead times longer than black line) and climatology (lead times shorter than gray line). Text labels show the date corresponding to the maximum lead time at which each forecast outperforms climatology.

290 We quantify the skill difference between ACCESS-S1 and ACCESS-S2 in Figure 10 by computing the shortest lead time in which the mean RMSE of each system's forecast are not statistically different. The median difference is around 35 days but the biggest difference is between February and April, where ACCESS-S1 forecasts have lower RMSE than ACCESS-S2's for more between 65 and 125 days. The smallest difference, on the other hand, is in winter, when ACCESS-S1's forecast are statistically indistinguishable from ACCESS-S2's after less than 20 days. Assimilation of sea-ice concentration data seems to have a much greater effect in the late summer than in the winter.

295

To analyse the spatial distribution of the model error, we computed the RMSE of zonal mean sea-ice concentration anomalies on 15 slices of 24° longitude span for each forecasting system. We control for some areas being naturally easier to forecast than others by computing the RMSE skill score with the climatological forecast RMSE as reference.

300 For ACCESS-S1 forecasts (Figure 11 and Figure A12), skill tends to be lower off the coast of Eastern Antarctica even at short lead times; ~~for instance, the skill score for forecasts initialised in May and June are negative between 0° and~~

~~120°E even~~. This is particularly important for forecasts initialised between May and July, where we see that the low skill in early winter (Figure 9 and Figure A10) is only a feature in East Antarctica where skill is even negative at almost zero lead time. ~~This mirrors Libera et al. findings of a “winter predictability barrier”, although they focus on ACCESS-S1 forecasts have relatively high predictive skill in the Weddell Sea and here we show that the effect seems to be stronger more to the east.~~ in winter, consistent with other seasonal prediction systems (Bushuk et al.). July forecasts have positive (albeit not statistically significant) skill with hindcasts initialised as early as February. The winter predictive skill barrier thus seems to come in great part from large errors in East Antarctica.

In West Antarctica there is a hint of easterly-propagating skill in forecasts initialised in February and March. This is consistent with ~~Holland et al. findings that~~, who found that the memory of sea-ice anomalies ~~are~~ is stored in ocean heat content anomalies that are transported eastwards by the Antarctic Circumpolar Current.

ACCESS-S2 forecasts (Figure 12 and Figure A13) also have lower skill over East Antarctica. From July to December, even though the pan-Antarctic average skill is negative at all lead times (Fig. A10), it is positive (but not statistically significant) for up to a month in West Antarctica. Since oceanic and atmospheric ~~forcing is~~ forcings provide the only source of observational information, this suggests that ~~sea-ice~~ sea ice in this region is particularly sensitive to oceanic and atmospheric forcing ~~and suggests a role of the Pacific-South American mode and the Amundsen Sea Low to shape sea-ice concentration anomalies~~. The fact that this is evident in the months in which El Niño–Southern Oscillation teleconnections via de Pacific-South American mode and the Amundsen Sea low are more important for atmospheric circulation (Mo and Paegle; Clem and Fogt; Wang et al., a) also suggests the influence of tropical Pacific variability. February and March are the only two months that can be forecasted with marginally positive skill in large regions.

Finally, Figure 13 (Figure A14) shows the difference in skill between ACCESS-S1 and ACCESS-S2. Large differences in skill indicate areas and months that are most affected by the data assimilation present in ACCESS-S1. Between January and March, ~~which are the months in which~~ when ACCESS-S1 is the most skillful (Figure A10), most of the ~~improvement~~ higher skill compared with ACCESS-S2 is present in the Ross and Weddell Seas. In April and May, the ~~improvement~~ higher skill seems more homogeneous. In June and July, ACCESS-S1 has higher skill only at very short lead times and in limited regions; in particular we see that in June there is no skill difference in the King Haakon Sea –between 0° and 90°E.

3.3 Limitations

~~Decomposition of forecast error spread at 1, 5 and 30 days lead time for ACCESS-S1 and ACCESS-S2 hindcasts across initialization months. The left column shows the mean standard deviation of RMSE errors across ensemble members, while the right column shows the standard deviation of the ensemble mean RMSE error and the spread of the persistence and climatology forecasts errors.~~

In Figure 9 the mean error was shown. Figure ?? (Figure ??) column 1 shows the mean standard deviation of errors among ensemble members at various lead times. At one day lead time (Fig. ?? a.1) This analysis rests on the assumption that the main difference between ACCESS-S1 and ACCESS-S2 has a slightly larger spread than is the lack of data assimilation of sea-ice concentrations. However, although both systems use the same model configuration and atmospheric initial conditions, they also differ in the ensemble generation scheme and their ocean initial conditions.

As mentioned in Section 2.1, ACCESS-S1 due to the way that ensemble members are generated. ACCESS-S1 ensemble members are generated by adding random field perturbations to the atmosphere only, which then are transferred to the other components via the coupled simulation (Hudson et al.) .With this while ACCESS-S2 ensemble members are generated by time-lagged ensemble.

We believe this difference would only alter the ensemble spread and should have minimal impact in the dramatic differences in ensemble mean error shown in Figure 9. In ACCESS-S1's scheme, ensemble members are all but guaranteed to be underdispersed in the ocean and sea-ice components. The time-lag ensemble used for To test the effect of the ensemble generation scheme on ensemble spread, we compute the mean variance of RMSE of individual forecasts as a function of lag for each forecasting system. Figure A15 (Figure A16) shows that indeed ACCESS-S1's initial spread is much lower than ACCESS-S2ensures greater spread's in every month. However, ACCESS-S1's ensemble spread grows quickly, and after less than a week, both systems's spread is comparable. For forecasts initialised in February, March and April, ACCESS-S1 even has larger spread than ACCESS-S2. This difference is gone after about just two days, and both systems have a comparable spread in ensemble member error afterwards (Fig. ?? b1 and c1).

Figure ?? column 2, on the other hand, shows the standard deviation of ensemble mean error of each hindcast and the persistence forecast. At one day lead time, ACCESS-S2 ensemble mean error standard deviation is much larger than Regarding ocean initial conditions, unfortunately we cannot directly compare ACCESS-S1 's, which in turn is comparable to the persistence forecast error standard deviation. At longer lead times, and ACCESS-S2 initial conditions because those data are currently not available. However, we believe there are reasons to consider that the effect would be of second order importance. While the ocean data assimilation process is different, the underlying observations are similar, especially in the spread of ACCESS-S1 high latitudes, where there are few direct observations in total. As an indirect comparison of ocean initial conditions, Wedd et al. in their Figure 9 show the correlation of ocean heat content in the upper 300 m of the ocean between EN4 and persistence forecast standard deviation increases to eventually be comparable to ACCESS-S2 and the standard deviation in climatological forecast errors. ACCESS-S1 reanalysis. In both systems, correlations are relatively high around 60°S and very low in the Antarctic coast; importantly, their correlation patterns are very similar, suggesting that the ocean initial conditions are not too dissimilar between the two systems. A limitation of this comparison is that the low correlation seen in both systems in the ice-covered regions only indicates that both systems have poor ocean initial conditions in that region but does

370 ~~not necessarily mean that their initial conditions are similar. Even though the ocean initial conditions are similarly correlated with the reference dataset in both systems, SST forecasts were shown by Wedd et al. to be highly degraded in the Southern Ocean even at zero months lead time in ACCESS-S2 error standard deviation is fairly independent of lead time and similar to the climatological forecast error standard deviation at all lead times compared with ACCESS-S1. Correlation of SST anomalies in the ice-covered regions is much lower for ACCESS-S2 (Wedd et al. Figure 13) and with a higher positive bias (their Figure 12), which is consistent with the greater negative sea-ice concentration bias in ACCESS-S2 (Figure 1). This raises the possibility that it is the lack of assimilation of sea-ice data that is degrading SST forecasts and not the other way around.~~

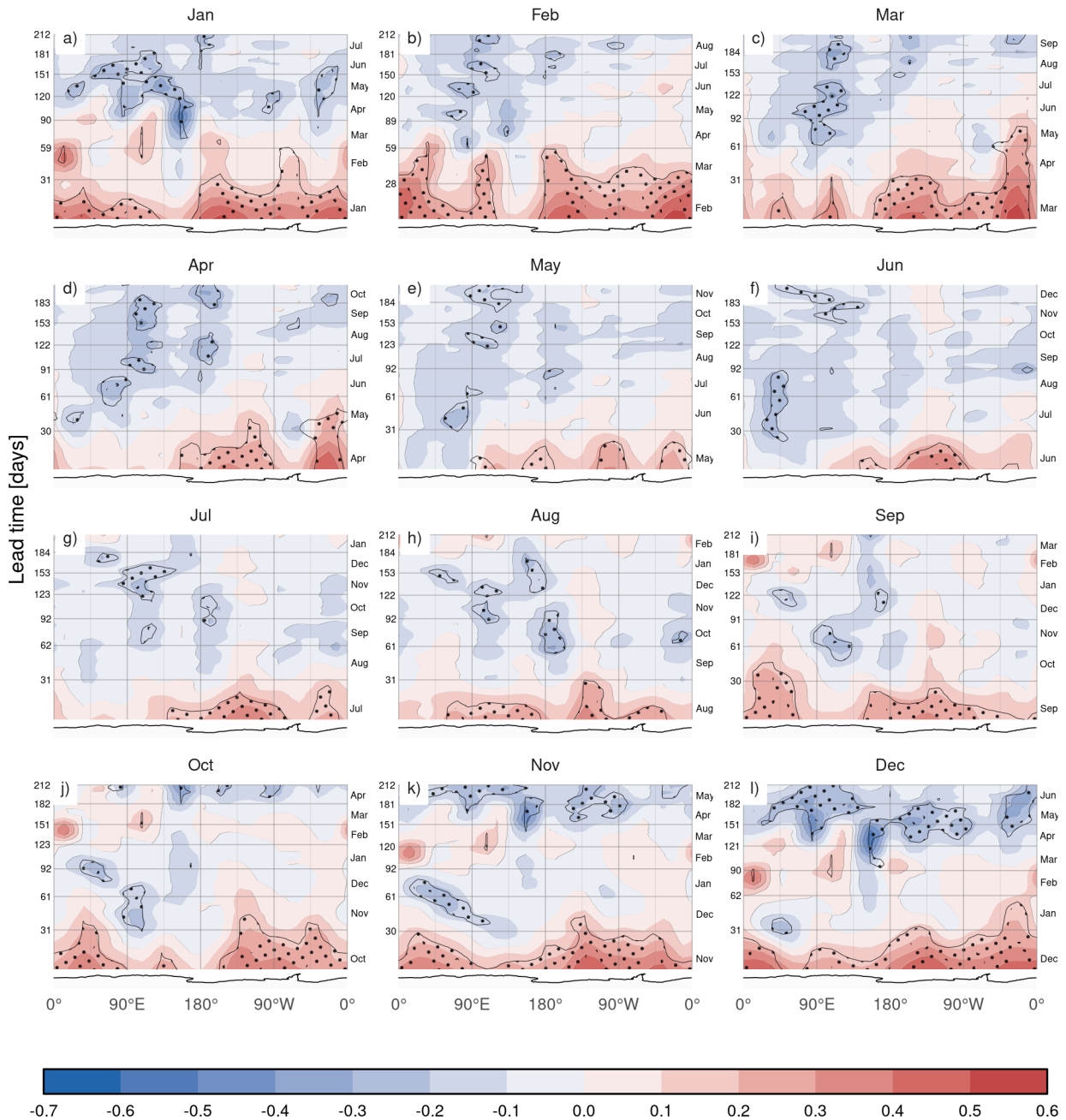


Figure 11. RMSE-MSE skill score of ACCESS-S1 forecasts with climatological forecast as reference computed on 15 meridional slices 24° wide as a function of lead time and longitude. Regions where the difference of means between ACCESS-S1 RMSE and climatological RMSE are significant at p-value < 0.05 are stippled. Antarctica's coastline is shown at the bottom of each panel for reference. Values were smoothed with an 11-day running mean to improve readability. Note that the vertical axis uses a 1.5 root transformation to expand the shorter lead times.

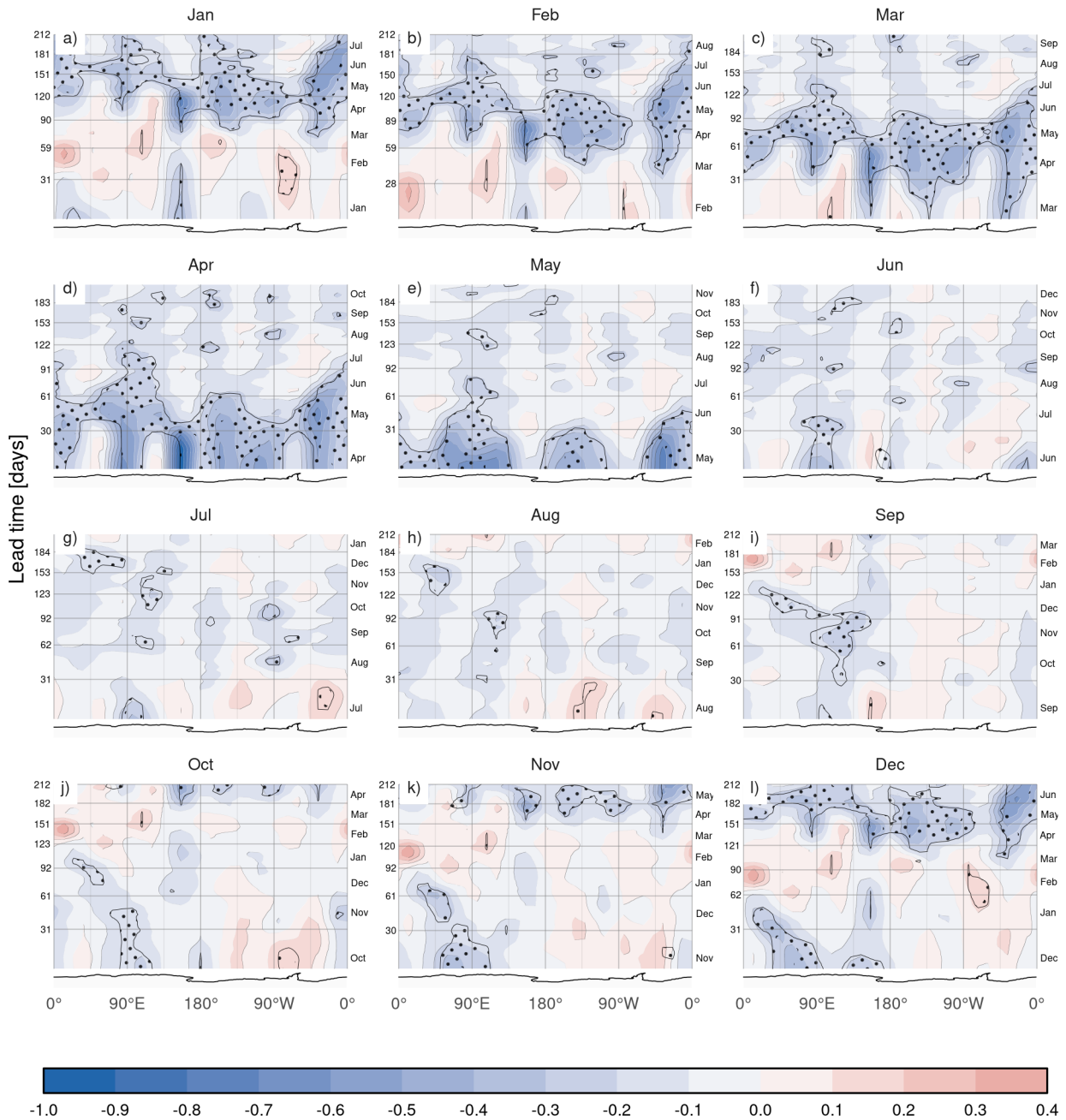


Figure 12. Same as Figure 11 but for ACCESS-S2.

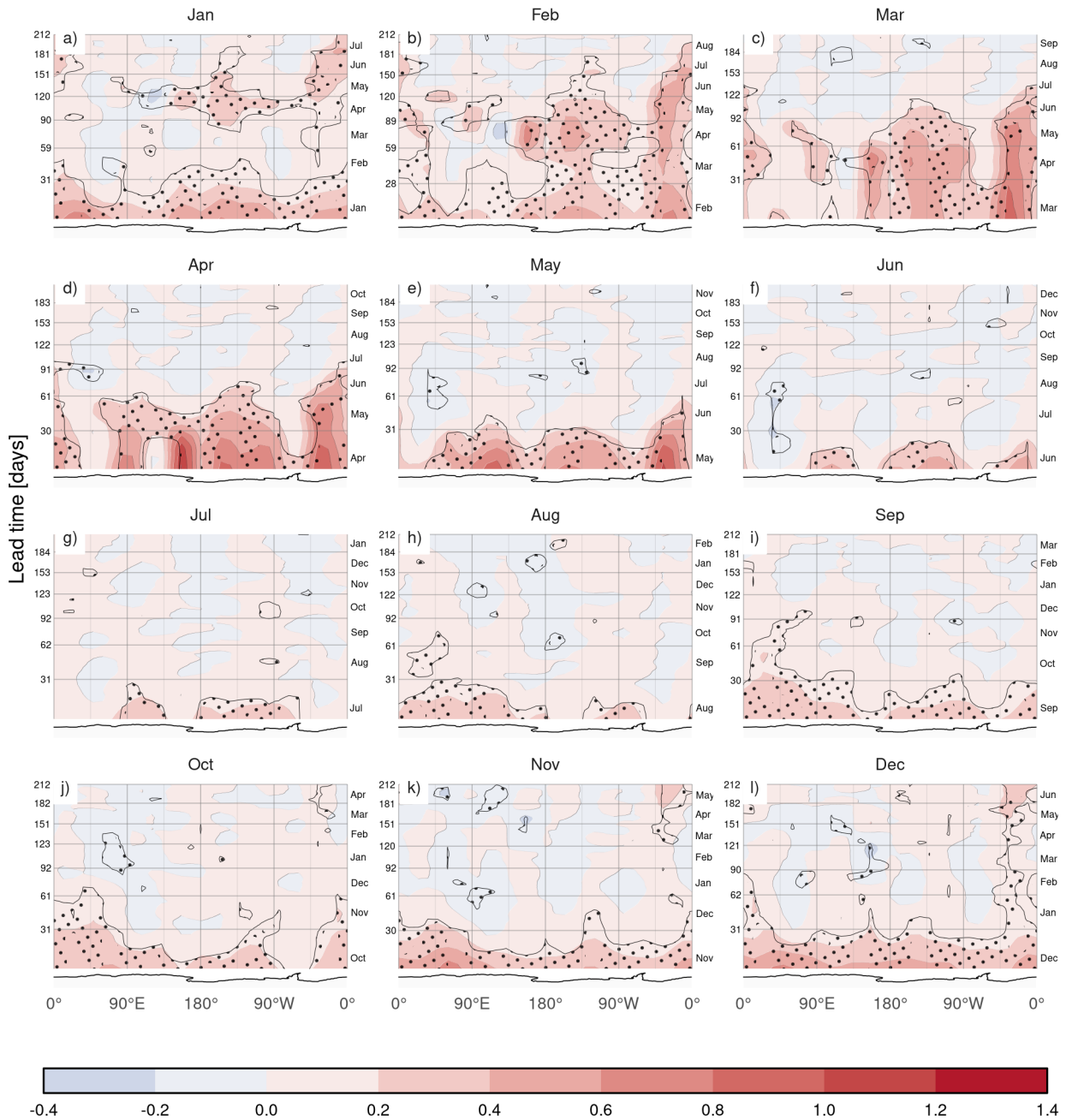


Figure 13. Same as Figure 11 but for the difference between ACCESS-S1 and ACCESS-S2.

4 Conclusions

Sea-ice forecasts from the ACCESS-S2 system show a significant low extent bias, particularly during late summer and early autumn. This bias is attributed to a faster and longer melt season between January and March, and slower growth between March and April. This underestimation during the minimum and early freezing season is a common
380 issue in many ~~seasonal-to-subseasonal (S2S)~~ subseasonal-to-seasonal systems, suggesting potential problems either with the model's thermodynamic representation or with short wave radiation forcing, as shown in other climate models (Zampieri et al.; Roach et al.). Even though ACCESS-S2 shares the same model components as ACCESS-S1, the latter does not suffer from this bias, indicating that assimilating sea-ice concentrations successfully corrects for the negative bias that exists in the free-running model.

385 Ensemble spread grows quickly even when perturbations are only implemented in the atmosphere component (in ACCESS-S1), indicating that sea ice is indeed responding quickly to atmospheric perturbations. However, our analysis suggests that the atmosphere and ocean data assimilation implemented in ACCESS-S2 is only effectively influencing sea-ice initial conditions from June to October, while the rest of the year, the sea-ice component runs virtually free, reverting to its biased equilibrium state. Zhou and Alves had previously evaluated sea-ice forecasts in ACCESS-S2
390 and also highlighted the poor performance of this forecasting system attributed to the lack of good initial conditions.

~~Analysis of the error spread shows that ACCESS-S2 initial conditions from December to May not only have large errors, but that the initial error spread is very large compared with ACCESS-S1. This spread is not due to the perturbation scheme, since the mean error variance for individual forecasts is low and comparable with ACCESS-S1. Instead, it is due to large variance of the mean error of individual forecasts, which is comparable to the climatology spread. This is further evidence that individual initial conditions are not being affected by the data assimilation scheme.~~

Although ACCESS-S1 only assimilates sea-ice concentration, it is clear that sea-ice thickness is also affected through the assimilation process. ACCESS-S1 simulates significantly thicker ice than ACCESS-S2 and in both systems ~~sea-ice~~ sea ice is thicker at shorter lead times than at longer lead times. Both the explicit data assimilation in ACCESS-S2
400 and the effects of atmospheric and oceanic data assimilation in ACCESS-S1 might be nudging simulated sea ice to be thicker than the model equilibrium state. It still remains unclear why ACCESS-S2 simulates much larger interannual variability than ACCESS-S1 at shorter lead times. We suggest that the thinner sea ice in ACCESS-S2 contributes to the large sea-ice extent variance, but other mechanisms, such as unbalanced initial conditions might also be important.

405 Given that ACCESS-S2 sea-ice extent is not directly initialised by sea-ice observations, and assuming that the effects of the slightly different ocean initial conditions and ensemble generation are of second-order importance, comparing

its forecasts with those of ACCESS-S1 allows us to estimate the time-scale over which initial conditions are important. We find that initial conditions affect Antarctic sea-ice forecasts in the order of a few months, but that effect is seasonally dependent. January to April initial conditions improve forecasts for up to three months. February initial conditions in particular are shown to be crucial for determining sea-ice evolution at least up to May. Arctic sea-ice forecasts also show greater sensitivity to initial conditions in boreal summer, compared with boreal winter (Day et al.; Bunzel et al.), suggesting a similar mechanism might be playing a role.

These results contrast with Xiu et al., which used a very similar methodology to our own. Based on their claims that assimilation of sea-ice concentration data doesn't improve sea-ice forecasts compared with only assimilating ocean data and that atmospheric conditions are enough for a skillful sea-ice forecast, we would have expected skillful forecasts from ACCESS-S2's atmosphere initial conditions. It is possible that model biases are important to explain this discrepancy. Their forecasting system simulated more sea ice than observations; the reverse of ACCESS-S1 and ACCESS-S2, which tended to drift into lower sea-ice concentration that needed to be corrected with data assimilation. However, since they used anomaly field assimilation, which assimilates anomalies on top of the model's own climatology, their initial conditions were close to the model's own equilibrium state, reducing the strong initial shock of the type we detected in ACCESS-S1 and which we speculate might be degrading variability in ACCESS-S2.

Forecasts initialised in the winter have very little skill and ACCESS-S1 and ACCESS-S2 forecast errors are statistically indistinguishable after just two weeks. ~~This is consistent with Libera et al.'s finding of a "winter predictability barrier"~~ Contrasting to Libera et al. but consistent with Xiu et al., we find that this predictive skill barrier is not apparent in the Weddell Sea ; although they describe the barrier as a sharp loss of predictability in July, and here we find a gradual reduction in skill compared with climatology around June. This difference might be due to our use of pan-Antarctic RMSE, since our regional analysis indicates that the degraded skill is most dramatic in the King Haakon Sea and instead comes from low skill in the Indian Ocean sector. However, we cannot make a direct comparison between Libera et al. and our study due to significant methodological differences. Their results were based on persistence of sea-ice area, whereas here we study errors in sea-ice concentration anomalies of a dynamical model. Further exploration of these differences could lead to more insight in how persistence-based measures of predictability translate to actual predictive skill.

These findings have important implications for both operational forecasting, model development and predictability studies. For operational centers, our results suggest that efforts to improve sea-ice data assimilation should prioritize the summer and autumn months when initial conditions have the greatest impact on forecast skill. Additionally, the substantial bias in ACCESS-S2 both models highlights the need for improved model physics, particularly in the representation of sea-ice thermodynamics and radiation processes. Crucially, our results suggest dramatic seasonal variations in sea-ice predictability. ~~Future~~ Our conclusions are limited by the fact that ACCESS-S1 and ACCESS-S2 also differ in the ocean data assimilation system and the ensemble generation method. Although we believe these

differences to be of second-order importance compared with the wildly divergent sea-ice initial conditions, they motivate the need for more targeted experiments. Predictability studies should therefore ~~use initial conditions through sample initial conditions throughout~~ the whole year ~~rather than instead of~~ focusing on a limited number of initialisation dates.

445 5 References

- Allaire, J., Teague, C., Xie, Y., and Dervieux, C.: Quarto, <https://doi.org/10.5281/ZENODO.5960048>.
- Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, 4, 677–699, 450 <https://doi.org/10.5194/gmd-4-677-2011>.
- Bunzel, F., Notz, D., Baehr, J., Müller, W. A., and Fröhlich, K.: Seasonal climate forecasts significantly affected by observational uncertainty of Arctic sea ice concentration, 43, 852–859, <https://doi.org/10.1002/2015GL066928>, [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2015GL066928](https://onlinelibrary.wiley.com/doi/pdf/10.1002/2015GL066928).
- Bushuk, M., Winton, M., Haumann, F. A., Delworth, T., Lu, F., Zhang, Y., Jia, L., Zhang, L., Cooke, W., Harrison, M., Hurlin, 455 B., Johnson, N. C., Kapnick, S. B., McHugh, C., Murakami, H., Rosati, A., Tseng, K.-C., Wittenberg, A. T., Yang, X., and Zeng, F.: Seasonal Prediction and Predictability of Regional Antarctic Sea Ice, 34, 6207–6233, <https://doi.org/10.1175/JCLI-D-20-0965.1>.
- Campitelli, E.: metR: Tools for Easier Analysis of Meteorological Fields, <https://CRAN.R-project.org/package=metR>, a.
- Campitelli, E.: Data for "Evaluating the Importance of Initial Conditions for Antarctic Sea Ice Seasonal Predictability with a 460 Fully Coupled Forecast Model", <https://doi.org/10.5281/ZENODO.17479537>, b.
- Cavalieri, D. J., Crawford, J. P., Drinkwater, M. R., Eppler, D. T., Farmer, L. D., Jentz, R. R., and Wackerman, C. C.: Aircraft active and passive microwave validation of sea ice concentration from the Defense Meteorological Satellite Program special sensor microwave imager, 96, 21989–22008, <https://doi.org/10.1029/91JC02335>, [_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/91JC02335](https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/91JC02335), a.
- 465 Cavalieri, D. J., Gloersen, P., and Campbell, W. J.: Determination of sea ice parameters with the NIMBUS 7 SMMR, 89, 5355–5369, <https://doi.org/10.1029/JD089iD04p05355>, [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/JD089iD04p05355](https://onlinelibrary.wiley.com/doi/pdf/10.1029/JD089iD04p05355), b.
- Clem, K. R. and Fogt, R. L.: Varying roles of ENSO and SAM on the Antarctic Peninsula climate in austral spring, 118, 11481–11492, <https://doi.org/10.1002/jgrd.50860>, [_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/jgrd.50860](https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/jgrd.50860).
- 470 Comiso, J.: Bootstrap Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS. (NSIDC-0079, Version 4), <https://doi.org/https://doi.org/10.5067/X5LG68MH013O>.
- Day, J. J., Hawkins, E., and Tietsche, S.: Will Arctic sea ice thickness initialization improve seasonal forecast skill?, 41, 7566–7575, <https://doi.org/10.1002/2014GL061694>, [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2014GL061694](https://onlinelibrary.wiley.com/doi/pdf/10.1002/2014GL061694).
- De Silva, L. W. A., Inoue, J., Yamaguchi, H., and Terui, T.: Medium range sea ice prediction in support of Japanese 475 research vessel MIRAI's expedition cruise in 2018, 43, 223–239, <https://doi.org/10.1080/1088937X.2019.1707317>, [_eprint: https://doi.org/10.1080/1088937X.2019.1707317](https://doi.org/10.1080/1088937X.2019.1707317).

- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, 137, 553–597, <https://doi.org/10.1002/qj.828>.
- Dong, X., Yang, Q., Nie, Y., Zampieri, L., Wang, J., Liu, J., and Chen, D.: Antarctic sea ice prediction with A convolutional long short-term memory network, 190, 102386, <https://doi.org/10.1016/j.ocemod.2024.102386>.
- 485 Dowle, M. and Srinivasan, A.: data.table: Extension of 'data.frame', <https://CRAN.R-project.org/package=data.table>.
- Facility, E. O. a. S. I. S. A.: Global sea ice concentration climate data record 1978-2020 (v3.0, 2022), OSI-450-a.
- Gao, Y., Xiu, Y., Nie, Y., Luo, H., Yang, Q., Zampieri, L., Lv, X., and Uotila, P.: An Assessment of Subseasonal Prediction Skill of the Antarctic Sea Ice Edge, 129, e2024JC021499, <https://doi.org/10.1029/2024JC021499>, [_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024JC021499](https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024JC021499).
- 490 Good, S. A., Martin, M. J., and Rayner, N. A.: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates, 118, 6704–6716, <https://doi.org/10.1002/2013JC009067>, [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2013JC009067](https://onlinelibrary.wiley.com/doi/pdf/10.1002/2013JC009067).
- Guemas, V., Chevallier, M., Déqué, M., Bellprat, O., and Doblas-Reyes, F.: Impact of sea ice initialization on sea ice and atmosphere prediction skill on seasonal timescales, 43, 3889–3896, <https://doi.org/10.1002/2015GL066626>, [_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2015GL066626](https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2015GL066626).
- 495 Gurvan, M., Bourdallé-Badie, R., Bouttier, P.-A., Bricaud, C., Bruciaferri, D., Calvert, D., Chanut, J., Clementi, E., Coward, A., Delrosso, D., Ethé, C., Flavoni, S., Graham, T., Harle, J., Iovino, D., Lea, D., Lévy, C., Lovato, T., Martin, N., Masson, S., Mocavero, S., Paul, J., Rousset, C., Storkey, D., Storto, A., and Vancoppenolle, M.: NEMO ocean engine, <https://doi.org/10.5281/zenodo.1475234>.
- 500 Holland, M. M., Blanchard-Wrigglesworth, E., Kay, J., and Vavrus, S.: Initial-value predictability of Antarctic sea ice in the Community Climate System Model 3, 40, 2121–2124, <https://doi.org/10.1002/grl.50410>, [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/grl.50410](https://onlinelibrary.wiley.com/doi/pdf/10.1002/grl.50410).
- Hudson, D., Alves, O., Hendon, H. H., Lim, E.-P., Liu, G., Luo, J.-J., MacLachlan, C., Marshall, A. G., Shi, L., Wang, G., Wedd, R., Young, G., Zhao, M., and Zhou, X.: ACCESS-S1 The new Bureau of Meteorology multi-week to seasonal prediction system, 67, 132–159, <https://doi.org/10.1071/es17009>.
- 505 Libera, S., Hobbs, W., Klockner, A., Meyer, A., and Matear, R.: Ocean-Sea Ice Processes and Their Role in Multi-Month Predictability of Antarctic Sea Ice, 49, e2021GL097047, <https://doi.org/10.1029/2021GL097047>, [_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021GL097047](https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021GL097047).
- Lin, Y., Yang, Q., Li, X., Dong, X., Luo, H., Nie, Y., Wang, J., Wang, Y., and Min, C.: Ice-kNN-South: A Lightweight Machine Learning Model for Antarctic Sea Ice Prediction, 2, e2024JH000433, <https://doi.org/10.1029/2024JH000433>, [_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024JH000433](https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024JH000433).
- 510 Marchi, S., Fichet, T., and Goosse, H.: Respective influences of perturbed atmospheric and ocean-sea ice initial conditions on the skill of seasonal Antarctic sea ice predictions: A study with NEMO3.6-LIM3, 148, 101591, <https://doi.org/10.1016/j.ocemod.2020.101591>.

- 515 Massonnet, F., Barreira, S., Barthélemy, A., Bilbao, R., Blanchard-Wrigglesworth, E., Blockley, E., Bromwich, D. H., Bushuk, M., Dong, X., Goessling, H. F., Hobbs, W., Iovino, D., Lee, W.-S., Li, C., Meier, W. N., Merryfield, W. J., Moreno-Chamorro, E., Morioka, Y., Li, X., Niraula, B., Petty, A., Sanna, A., Scilingo, M., Shu, Q., Sigmond, M., Sun, N., Tietsche, S., Wu, X., Yang, Q., and Yuan, X.: SIPN South: six years of coordinated seasonal Antarctic sea ice predictions, 10, <https://doi.org/10.3389/fmars.2023.1148899>.
- 520 Megann, A., Storkey, D., Aksenov, Y., Alderson, S., Calvert, D., Graham, T., Hyder, P., Siddorn, J., and Sinha, B.: GO5.0: the joint NERC–Met Office NEMO global ocean model for use in coupled and forced applications, 7, 1069–1092, <https://doi.org/10.5194/gmd-7-1069-2014>.
 Meier, W. N. and Stewart, J. S.: Assessing uncertainties in sea ice extent climate indicators, 14, 035 005, <https://doi.org/10.1088/1748-9326/aaf52c>.
- 525 Meier, W. N., Fetterer, F., Windnagel, A. K., and Stewart, J. S.: NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, <https://doi.org/https://doi.org/10.7265/efmz-2t65>, a.
 Meier, W. N., Peng, G., Scott, D. J., and Savoie, M. H.: Verification of a new NOAA/NSIDC passive microwave sea-ice concentration climate record, <https://doi.org/10.3402/polar.v33.21004>, b.
- Mo, K. C. and Paegle, J. N.: The Pacific–South American modes and their downstream effects, 21, 1211–1229, <https://doi.org/https://doi.org/10.1002/joc.685>, _eprint: <https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/joc.685>.
- 530 Morioka, Y., Iovino, D., Cipollone, A., Masina, S., and Behera, S. K.: Decadal Sea Ice Prediction in the West Antarctic Seas with Ocean and Sea Ice Initializations, 3, 189, <https://doi.org/10.1038/s43247-022-00529-z>.
- Murphy, A. H. and Daan, H.: Forecast Evaluation, in: Probability, Statistics, And Decision Making In The Atmospheric Sciences, CRC Press, num Pages: 59.
- 535 R Core Team: R: A language and environment for statistical computing, <https://www.R-project.org/>, tex.organization: R Foundation for Statistical Computing.
- Rae, J. G. L., Hewitt, H. T., Keen, A. B., Ridley, J. K., West, A. E., Harris, C. M., Hunke, E. C., and Walters, D. N.: Development of the Global Sea Ice 6.0 CICE configuration for the Met Office Global Coupled model, 8, 2221–2230, <https://doi.org/10.5194/gmd-8-2221-2015>.
- 540 Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., and Schlax, M. G.: Daily High-Resolution-Blended Analyses for Sea Surface Temperature, 20, 5473–5496, <https://doi.org/10.1175/2007JCLI1824.1>.
- Rinke, A., Maslowski, W., Dethloff, K., and Clement, J.: Influence of sea ice on the atmosphere: A study with an Arctic atmospheric regional climate model, 111, <https://doi.org/10.1029/2005JD006957>, _eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2005JD006957>.
- 545 Roach, L. A., Dörr, J., Holmes, C. R., Massonnet, F., Blockley, E. W., Notz, D., Rackow, T., Raphael, M. N., O’Farrell, S. P., Bailey, D. A., and Bitz, C. M.: Antarctic Sea Ice Area in CMIP6, 47, e2019GL086 729, <https://doi.org/10.1029/2019GL086729>, _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019GL086729>.
- Schulzweida, U.: CDO User Guide, <https://doi.org/10.5281/ZENODO.10020800>, version Number: 2.3.0.
- Semmler, T., Kasper, M. A., Jung, T., and Serrar, S.: Remote impact of the Antarctic atmosphere on the southern mid?latitudes, 550 25, 71–77, <https://doi.org/10.1127/metz/2015/0685>, place: Stuttgart, Germany.

- Wagner, P. M., Hughes, N., Bourbonnais, P., Stroeve, J., Rabenstein, L., Bhatt, U., Little, J., Wiggins, H., and Fleming, A.: Sea-ice information and forecast needs for industry maritime stakeholders, 43, 160–187, <https://doi.org/10.1080/1088937X.2020.1766592>, _eprint: <https://doi.org/10.1080/1088937X.2020.1766592>.
- 555 Wang, J., Luo, H., Yu, L., Li, X., Holland, P. R., and Yang, Q.: The Impacts of Combined SAM and ENSO on Seasonal Antarctic Sea Ice Changes, 36, 3553–3569, <https://doi.org/10.1175/JCLI-D-22-0679.1>, a.
- Wang, Z., Fraser, A. D., Reid, P., Coleman, R., and O’Farrell, S.: The Influence of Time-Varying Sea Ice Concentration on Antarctic and Southern Ocean Numerical Weather Prediction, 39, 293–310, <https://doi.org/10.1175/WAF-D-22-0220.1>, b.
- Waters, J., Bell, M. J., Martin, M. J., and Lea, D. J.: Reducing ocean model imbalances in the equatorial region caused by data assimilation, 143, 195–208, <https://doi.org/10.1002/qj.2912>, _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2912>, a.
- 560 Waters, J., Lea, D. J., Martin, M. J., Mirouze, I., Weaver, A., and While, J.: Implementing a variational data assimilation system in an operational 1/4 degree global ocean model, 141, 333–349, <https://doi.org/10.1002/qj.2388>, _eprint: <https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2388>, b.
- Wedd, R., Alves, O., Burgh-Day, C. d., Down, C., Griffiths, M., Hendon, H. H., Hudson, D., Li, S., Lim, E.-P., Marshall, A. G., Shi, L., Smith, P., Smith, G., Spillman, C. M., Wang, G., Wheeler, M. C., Yan, H., Yin, Y., Young, G., Zhao, M., Xiao, Y., and Zhou, X.: ACCESS-S2: the upgraded Bureau of Meteorology multi-week to seasonal prediction system, 72, 218–242, <https://doi.org/10.1071/ES22026>.
- 565 Wickham, H.: ggplot2: Elegant Graphics for Data Analysis, Use R!, Springer-Verlag, <https://doi.org/10.1007/978-0-387-98141-3>.
- Williams, K. D., Harris, C. M., Bodas-Salcedo, A., Camp, J., Comer, R. E., Copesey, D., Fereday, D., Graham, T., Hill, R., Hinton, T., Hyder, P., Ineson, S., Masato, G., Milton, S. F., Roberts, M. J., Rowell, D. P., Sanchez, C., Shelly, A., Sinha, B., Walters, D. N., West, A., Woollings, T., and Xavier, P. K.: The Met Office Global Coupled model 2.0 (GC2) configuration, 8, 1509–1524, <https://doi.org/10.5194/gmd-8-1509-2015>.
- Xie, Y.: Dynamic documents with R and knitr, Chapman and Hall/CRC, 2 edn., <https://yihui.org/knitr/>.
- 575 Xiu, Y., Wang, Y., Luo, H., Garcia-Oliva, L., and Yang, Q.: Impact of Ocean, Sea Ice or Atmosphere Initialization on Seasonal Prediction of Regional Antarctic Sea Ice, 17, e2024MS004382, <https://doi.org/10.1029/2024MS004382>, _eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024MS004382>.
- Zampieri, L., Goessling, H. F., and Jung, T.: Predictability of Antarctic Sea Ice Edge on Subseasonal Time Scales, 46, 9719–9727, <https://doi.org/10.1029/2019GL084096>, _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019GL084096>.
- 580 Zhou, X. and Alves, O.: Evaluating sea ice in ACCESS-S2, <http://www.bom.gov.au/research/publications/researchreports/BRR-066.pdf>.
- Zweng, M. M., Reagan, J. R., Antonov, J. I., Locarnini, R. A., Mishonov, A. V., Boyer, T. P., Garcia, H. E., Baranova, O. K., Johnson, D. R., Seidov, D., and Biddle, M. M.: World ocean atlas 2013. Volume 2, Salinity, <https://doi.org/10.7289/V5251G4D>.

6 Appendix

585 The following are the same figures from the main paper but using the OSI dataset instead of CDR.

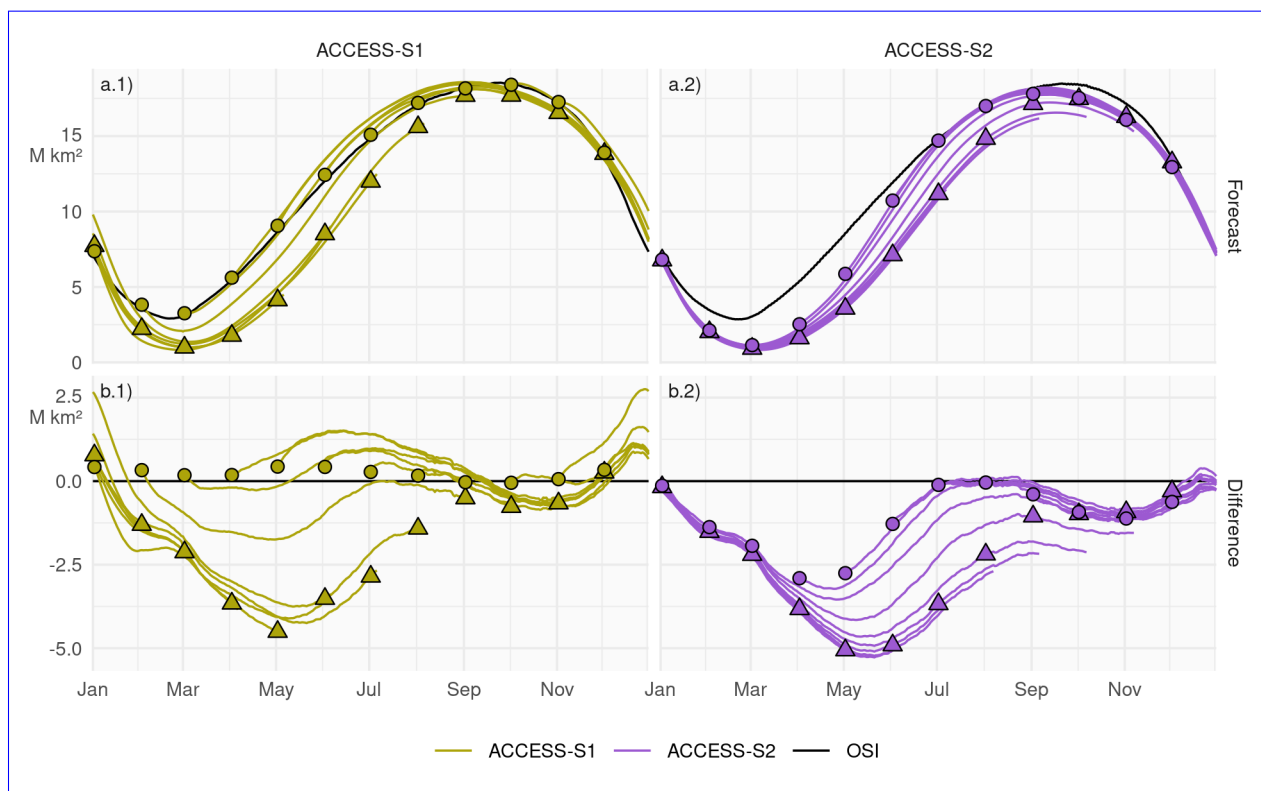


Figure A1. Row a: Pan-Antarctic daily mean sea-ice extent for all hindcasts initialised on the first of each calendar month for ACCESS-S1 (column 1; green) and ACCESS-S2 (column 2; purple). Observed mean sea-ice extent in each corresponding hindcast period is shown in black. Row b: Mean differences between the forecast and the observed values. Circles represent the initial conditions at the start of forecasts (i.e., the first of every month), and triangles represent the mean values [forecasted for the first of every month](#) at the lead time corresponding to the maximum lead time in S1 (between 213 and 216 days, depending on the month).

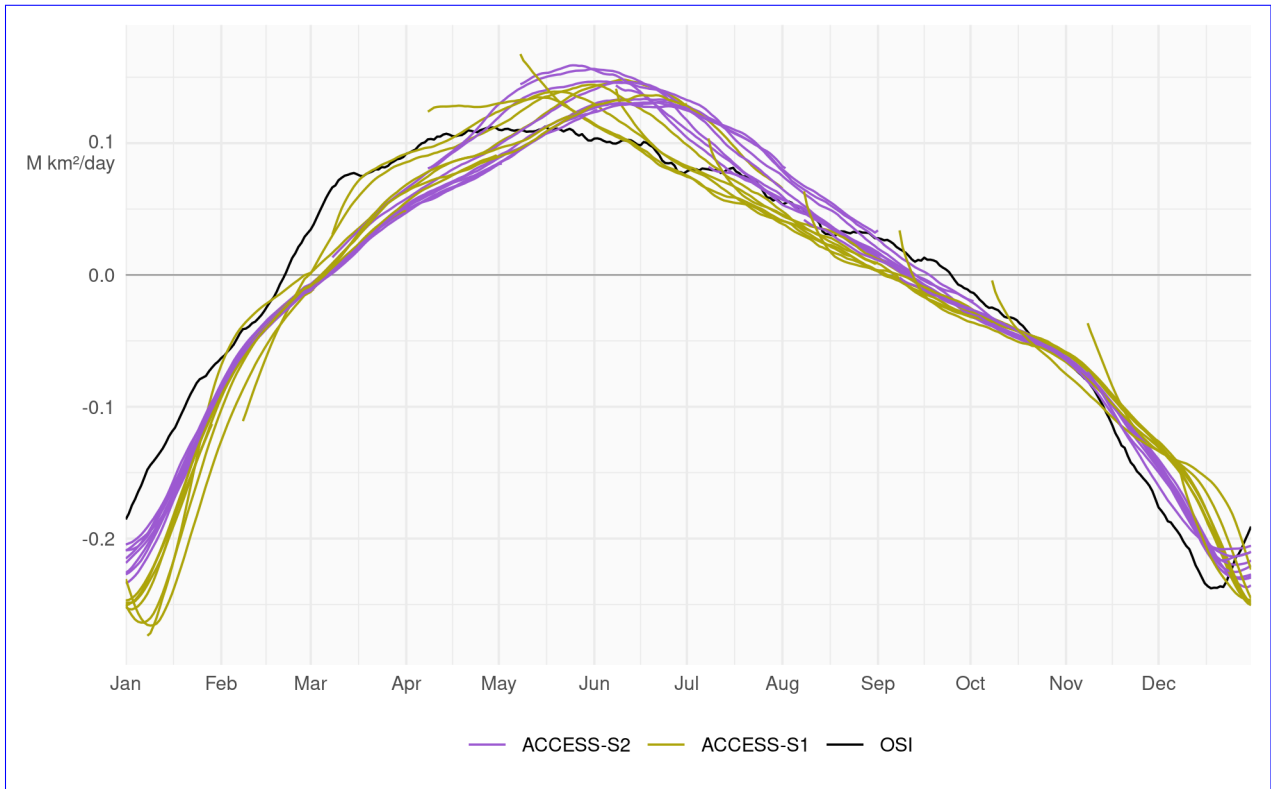


Figure A2. Mean daily sea-ice extent growth ($10^6 km^2/day$) in ACCESS-S1 (green) and ACCESS-S2 (purple) hindcasts and observations (black), computed as the mean daily differences in sea-ice extent between each date and the next for each forecast month. Values are smoothed with a 11-day running mean.

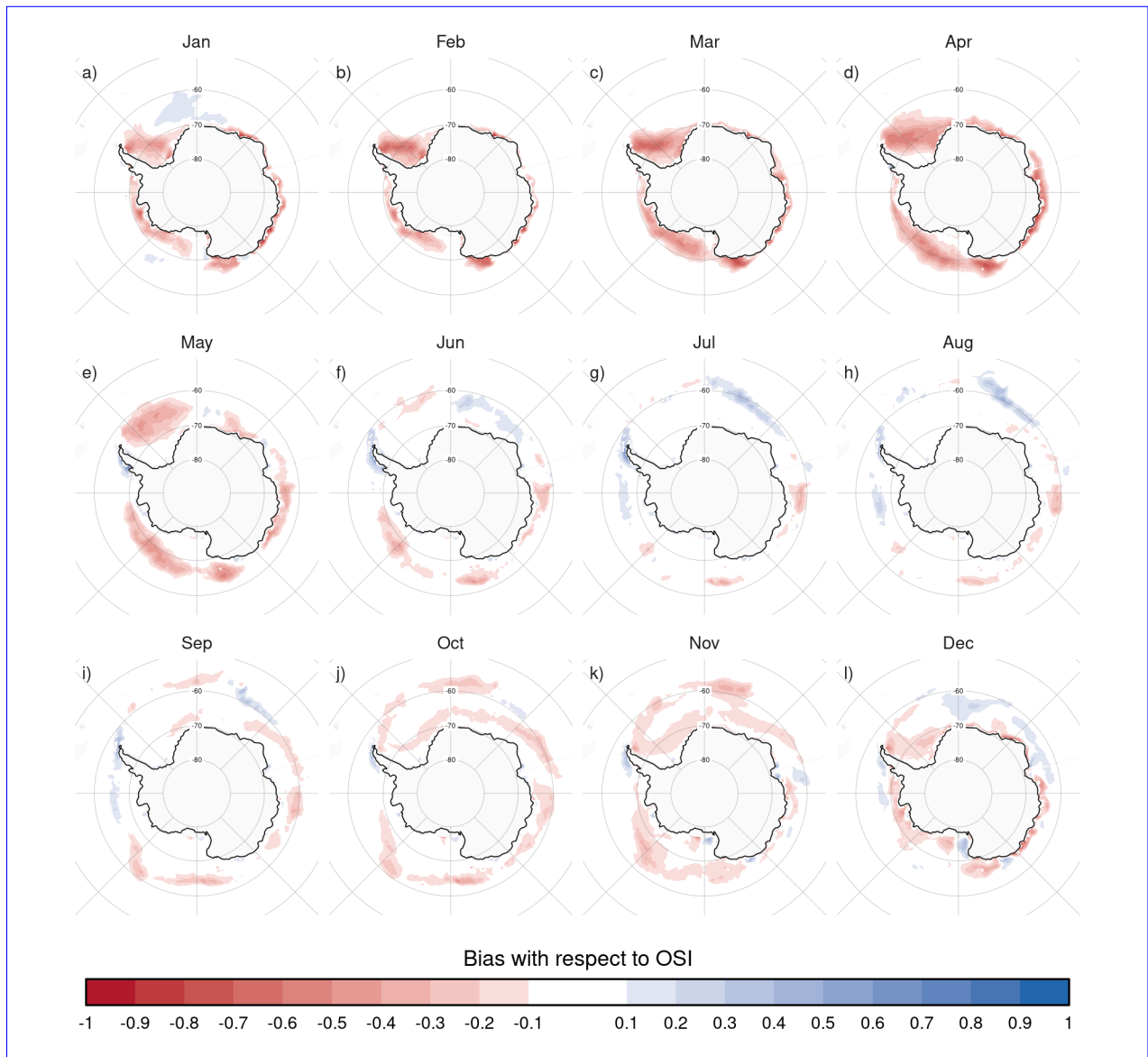


Figure A3. Ensemble mean difference between monthly sea-ice concentration of ACCESS-S2 ensemble mean forecast at 0-month lead time (monthly mean values forecasted from the forecast initialised at the first of the month) and observations (OSI).

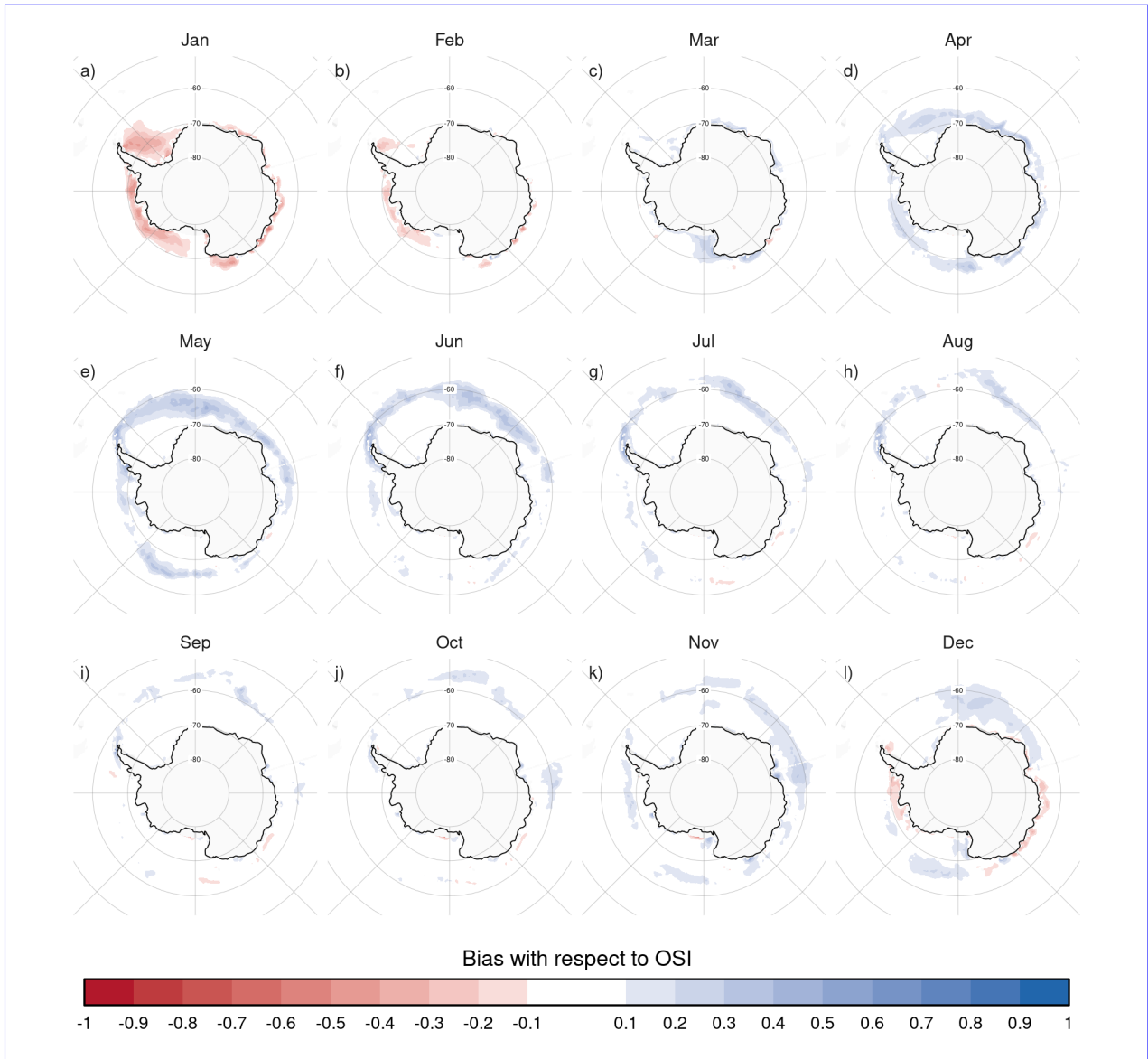


Figure A4. Same as Figure 3 but for ACCESS-S1.

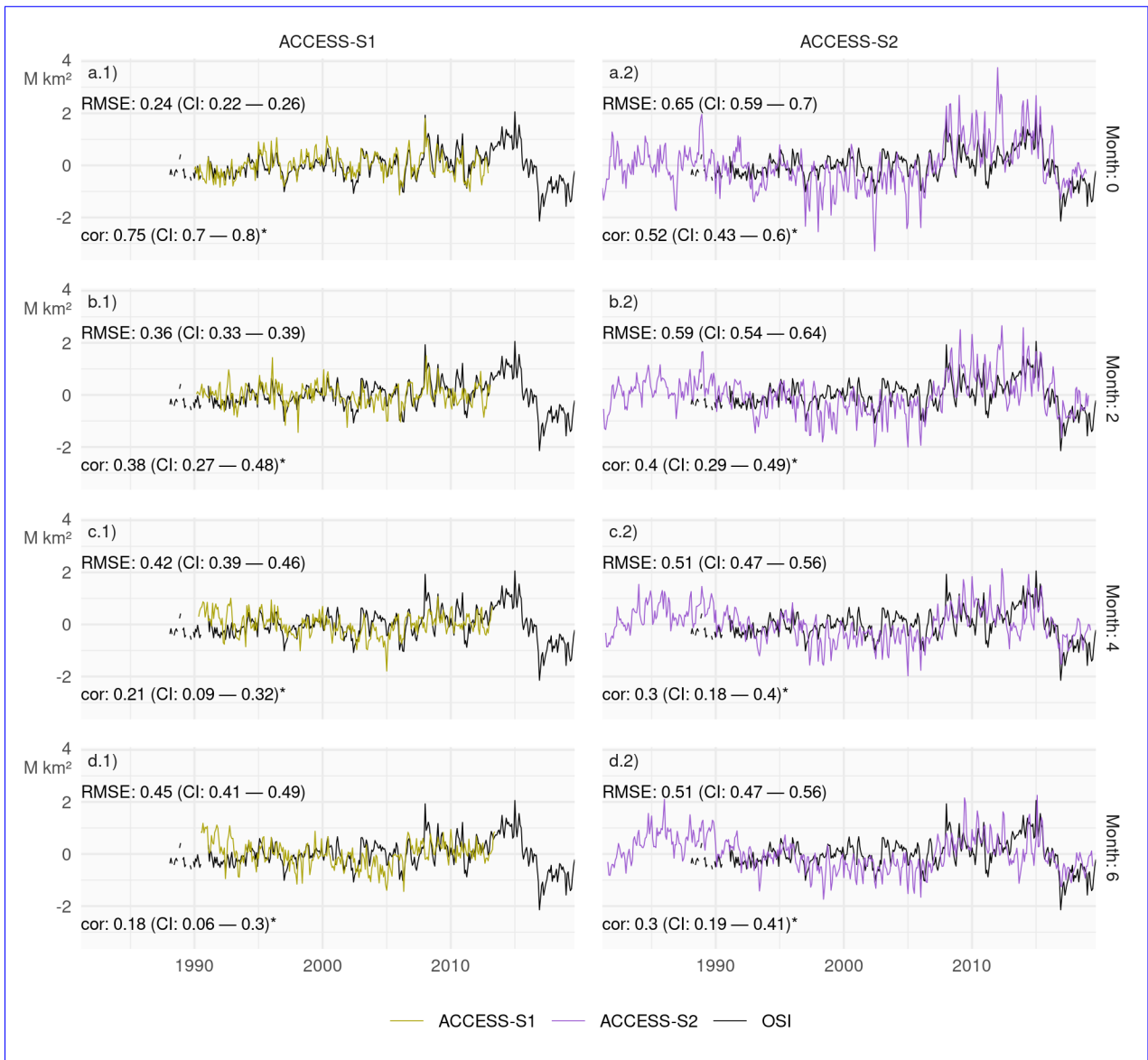


Figure A5. Monthly mean sea-ice extent anomalies of the observations (black) and forecasts from ACCESS-S1 (right column; purple) and ACCESS-S2 (left column; green) at lead times of 0, 2, 4, and 6 months. The RMSE and correlation [with their respective 95% confidence interval](#) during the overlapping period of ACCESS-S1 and ACCESS-S2 hindcasts (1990–2013|1990–2012) are shown on the top left and bottom left of each panel respectively. [Statistically significant correlations at 95% confidence level are marked with an asterisk.](#)

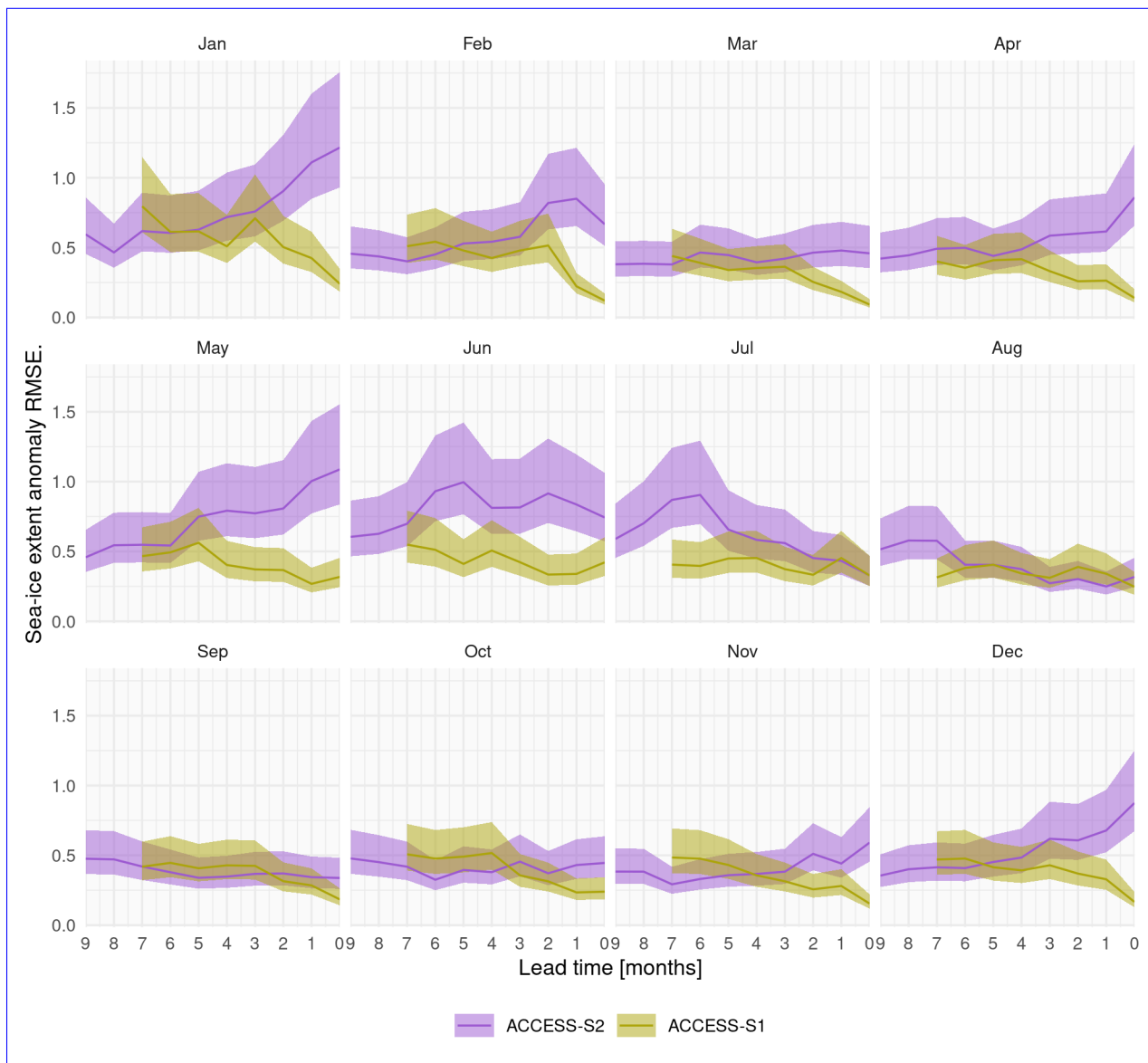


Figure A6. RMSE of monthly mean sea-ice extent anomalies as a function of lead time (months) for ACCESS-S1 (green) and ACCESS-S2 (purple). RMSE is computed over the overlapping period of ACCESS-S1 and ACCESS-S2 hindcasts (1990–2013). Each panel indicates the target month. Note the reverse horizontal axis.

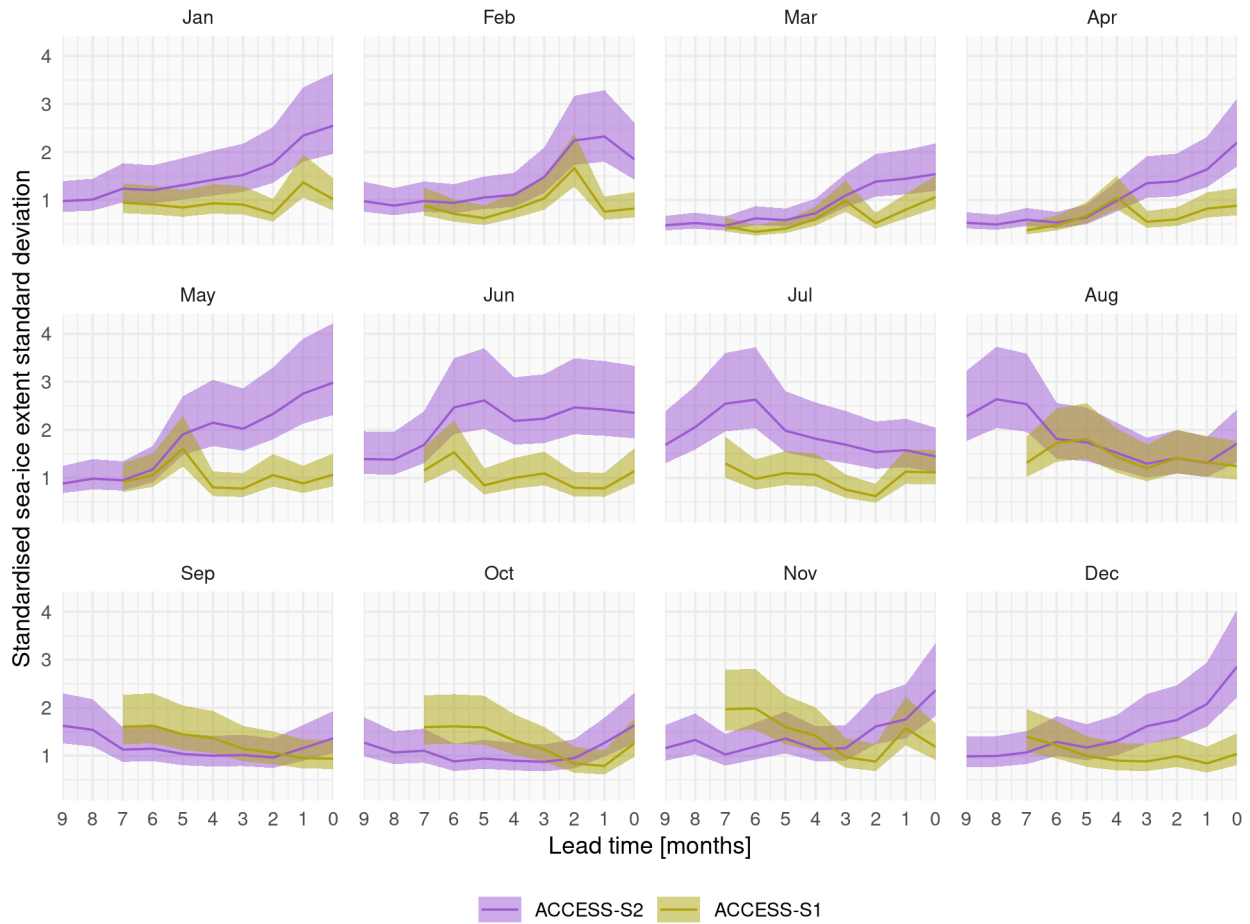
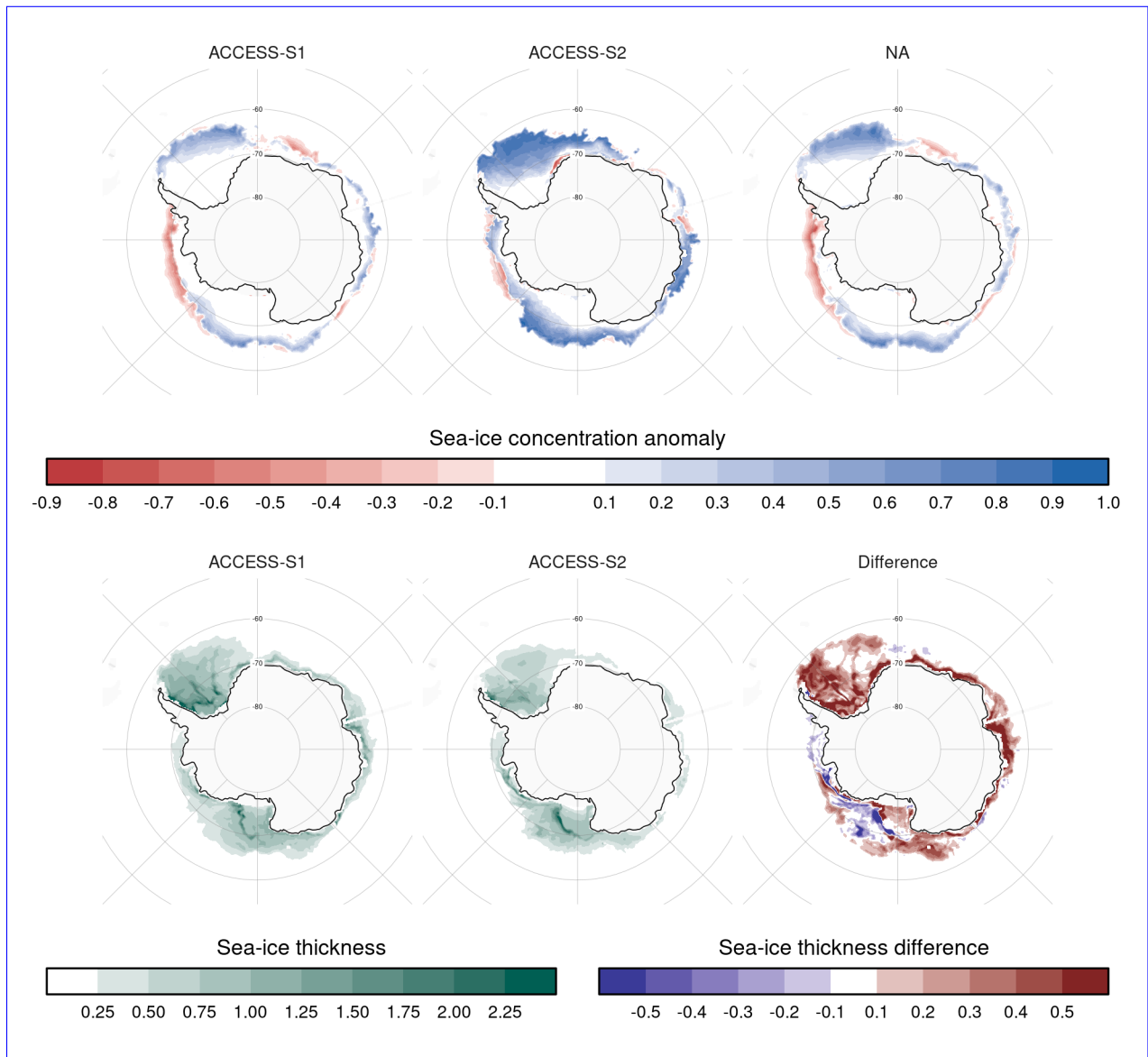


Figure A7. Interannual standard deviation with 95% confidence interval of monthly mean sea-ice extent forecasted for each month ~~divided~~. We standardise the standard deviation by that month's sea-ice extent observation standard deviation. ~~ACCESS-S1 and ACCESS-S2 at different lead times~~. Each panel indicates the target month. Note the reverse horizontal axis.



Minimum lead time at which each forecast's mean RMSE becomes larger than ACCESS-S2 simulates large sea-ice concentration anomalies in the lower bound of Weddell and Ross Seas and slight negative anomalies in the 95% confidence interval of persistence forecast RMSE (black lines) Amundsen and maximum lead time at which each forecast's mean RMSE remains lower than Bellingshausen Seas that align fairly well with observations, but the lower bound magnitude of the 95% confidence interval of climatological forecast RMSE (gray lines) anomalies is too big. Green shading indicates the window where forecasts outperform both persistence (lead times longer than black line) and climatology (lead times shorter than gray line) ACCESS-S1's sea-ice concentration anomalies are close to observations. Text labels show At the date corresponding to the maximum lead same time at which each forecast outperforms climatology, ACCESS-S2 simulates thinner ice than ACCESS-S1.

Minimum lead time at which each forecast's mean RMSE becomes larger than ACCESS-S2 simulates large sea-ice concentration anomalies in the lower bound of Weddell and Ross Seas and slight negative anomalies in the 95% confidence interval of persistence forecast RMSE (black lines) Amundsen and maximum lead time at which each forecast's mean RMSE remains lower than Bellingshausen Seas that align fairly well with observations, but the

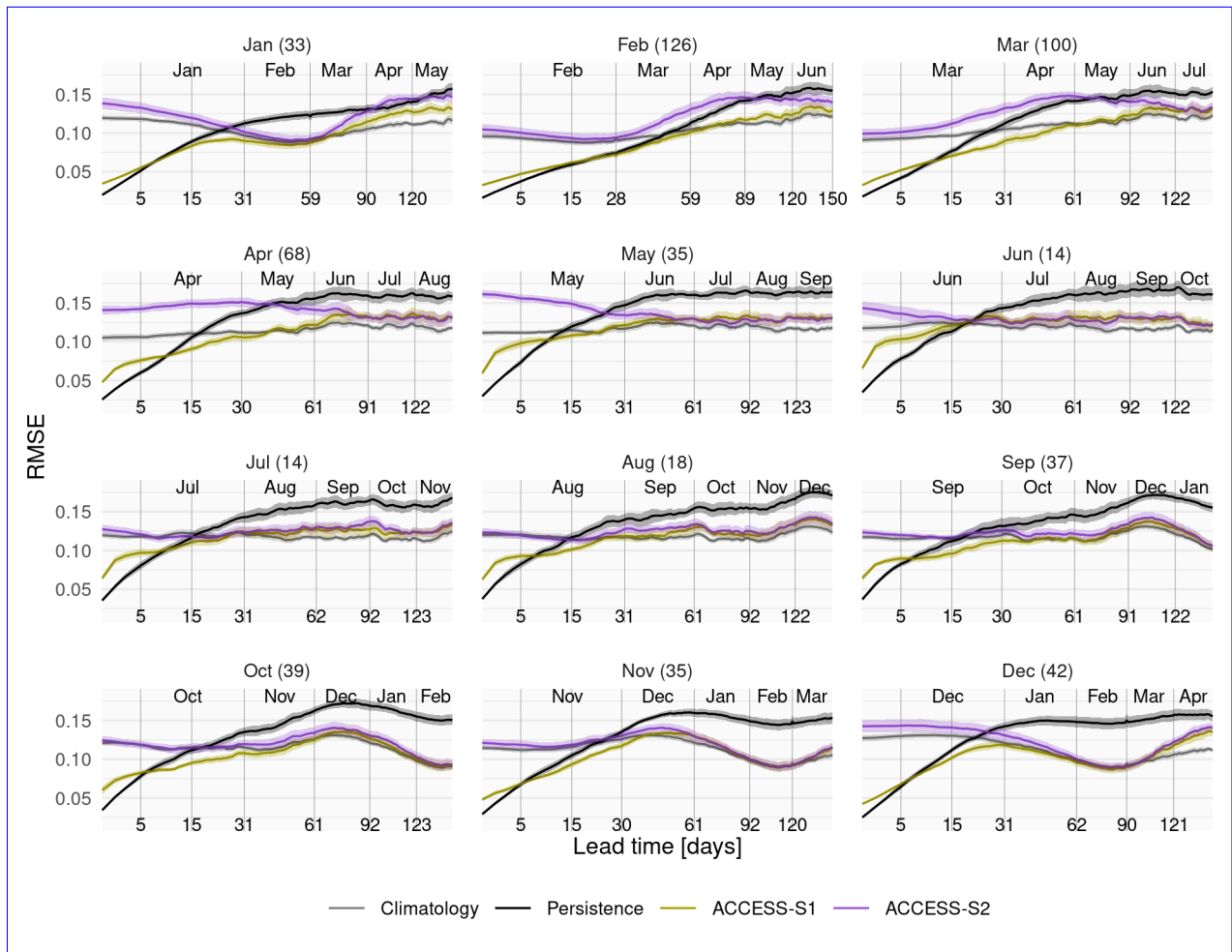


Figure A9. Mean RMSE of sea-ice concentration anomalies as a function of forecast lead time for all forecasts initialised on the first of each month compared with a reference forecast of persistence of anomalies (black) and climatology (gray). The shading indicates the 95% confidence interval of the mean. Only the first 150 days are shown. In parentheses, the shortest time at which ACCESS-S1 and ACCESS-S2 mean RMSE is not statistically different at the 99% confidence level. Note that the horizontal axis uses a square root transformation to expand the shorter lead times.

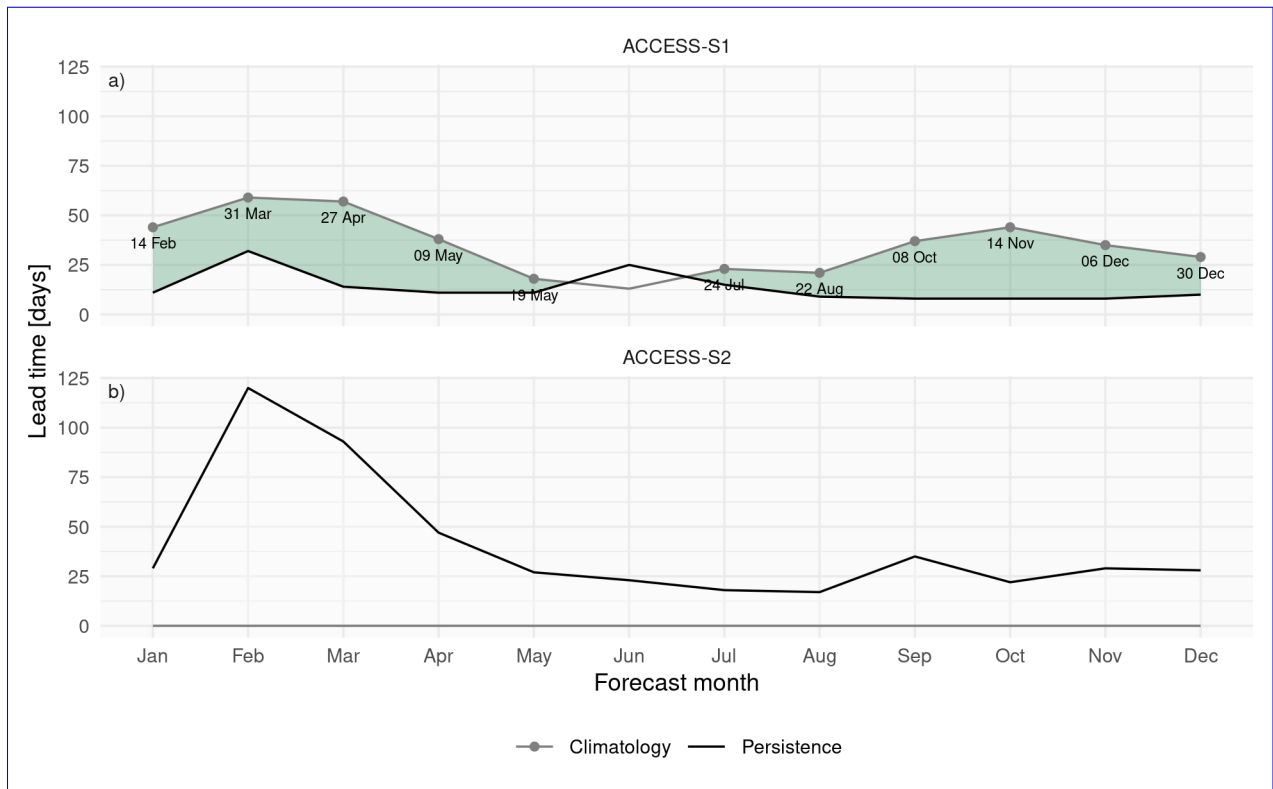


Figure A10. Minimum lead time at which each forecast's mean RMSE becomes lower than the lower bound of the 95% confidence interval of persistence forecast RMSE (black lines) and maximum lead time at which each forecast's mean RMSE remains lower than the lower bound of the 95% confidence interval of climatological forecast RMSE (gray lines). Green shading indicates the window where forecasts outperform both persistence (lead times longer than black line) and climatology (lead times shorter than gray line). Text labels show the date corresponding to the maximum lead time at which each forecast outperforms climatology. Using CDR as the validation dataset.

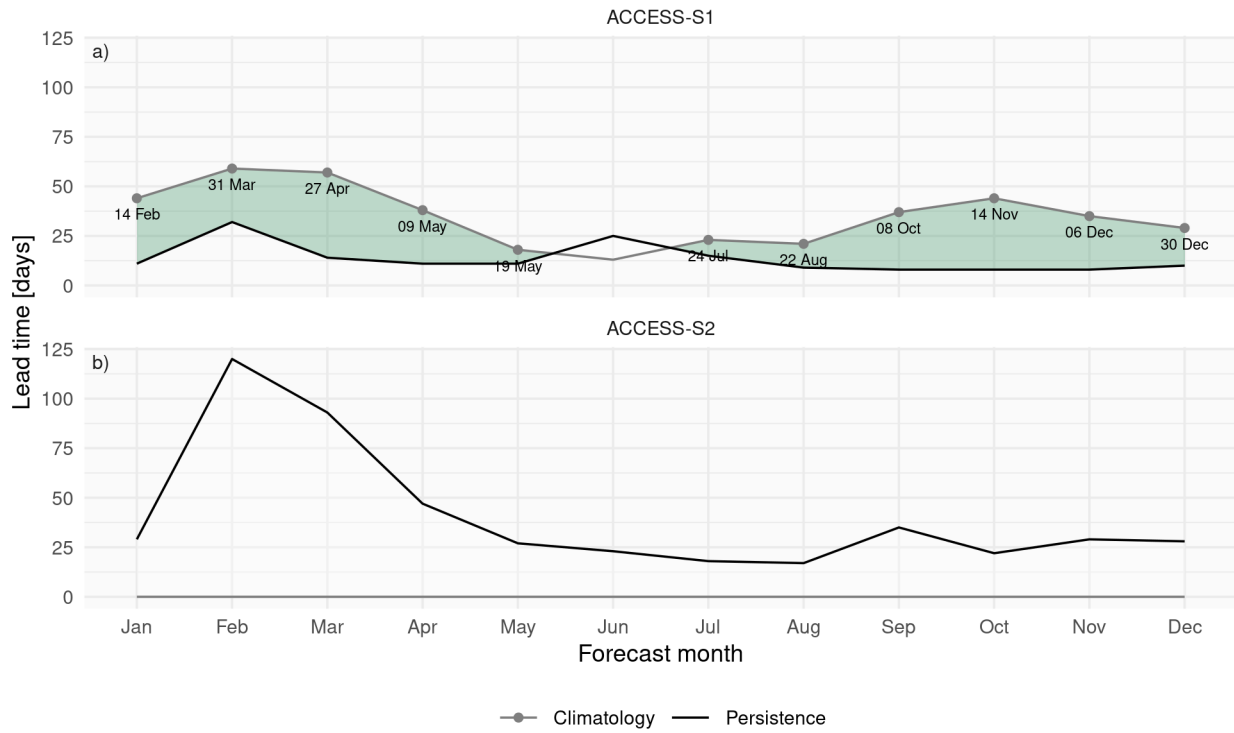


Figure A11. [Same as Figure A10 but using OSI as the validation dataset.](#)

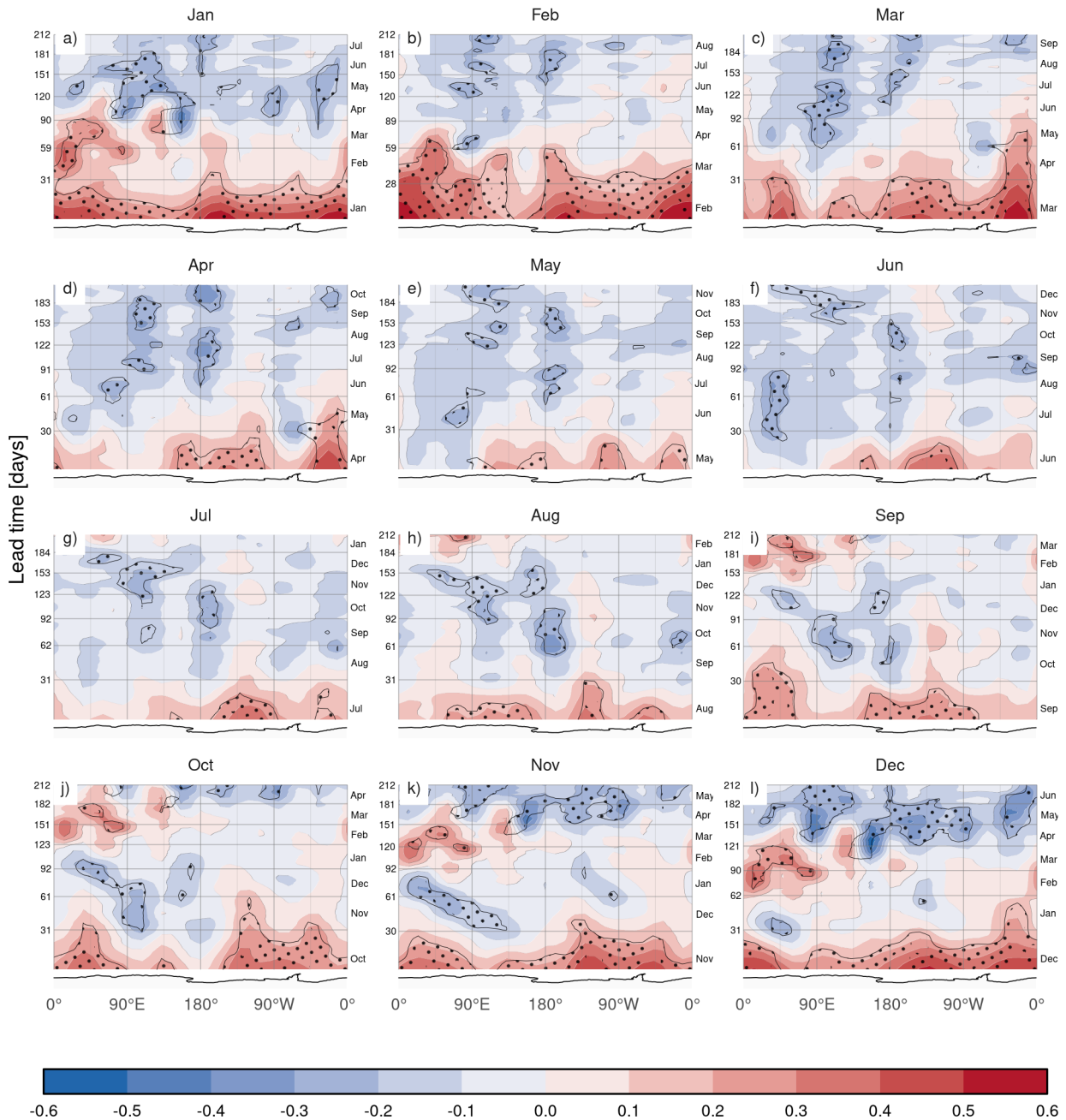


Figure A12. RMSE-MSE skill score of ACCESS-S1 forecasts with climatological forecast as reference computed on 15 meridional slices 24° wide as a function of lead time and longitude. Regions where the difference of means between ACCESS-S1 RMSE and climatological RMSE are significant at p-value < 0.05 are stippled. Antarctica's coastline is shown at the bottom of each panel for reference. Values were smoothed with an 11-day running mean to improve readability. Note that the vertical axis uses a 1.5 root transformation to expand the shorter lead times.

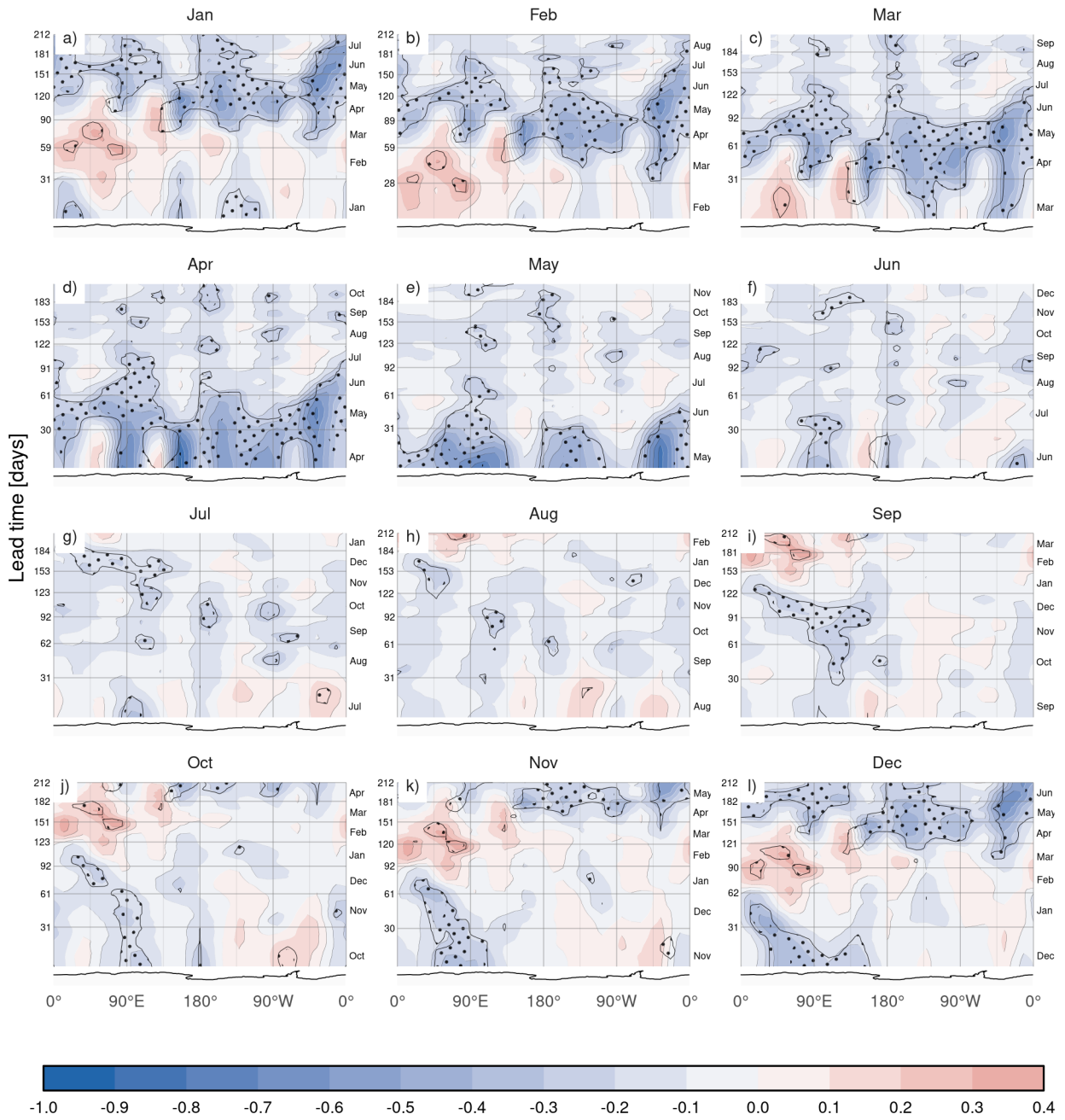


Figure A13. Same as Figure 11 but for ACCESS-S2.

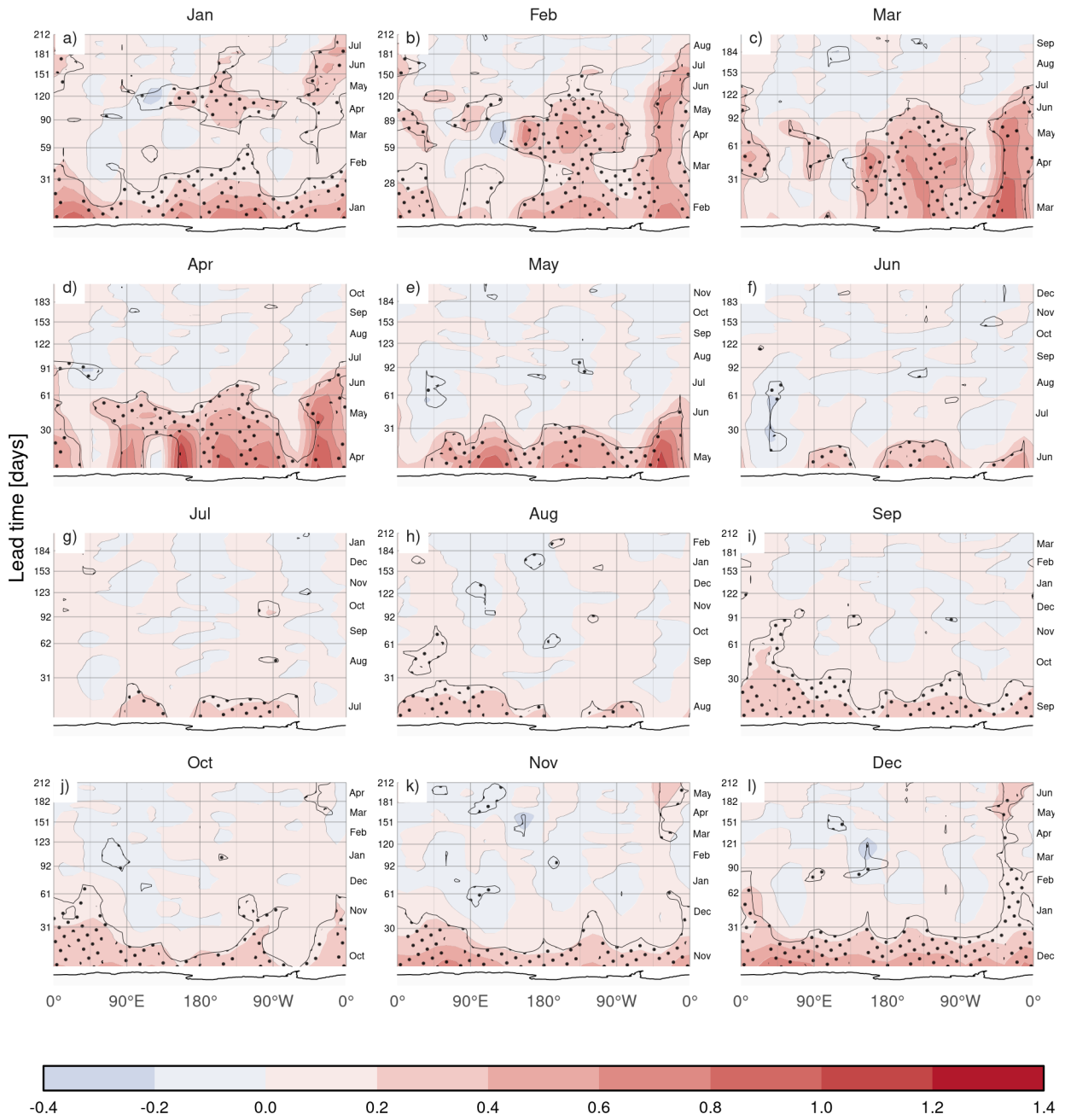


Figure A14. Same as Figure 11 but for the difference between ACCESS-S1 and ACCESS-S2.

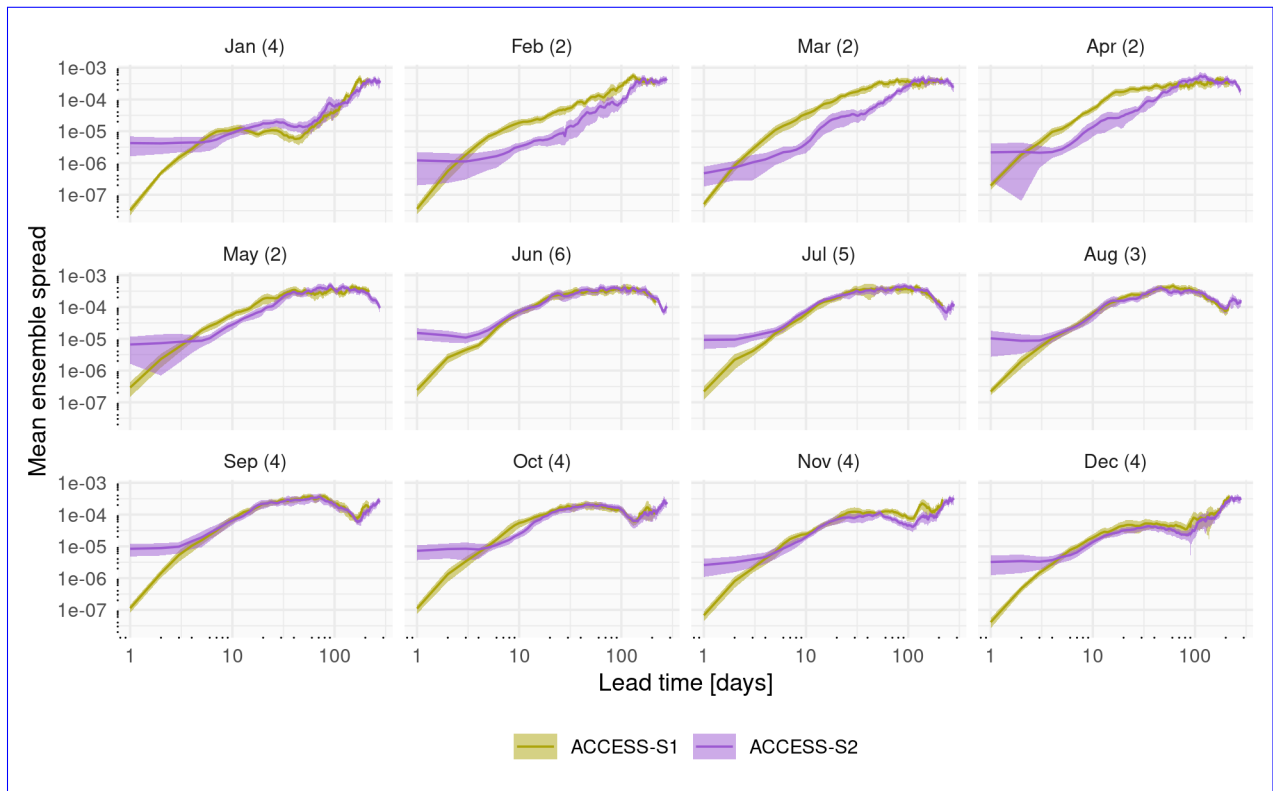


Figure A15. Decomposition—Mean variance of RMSE between ensemble members of each forecast error spread at 1. In parentheses, 5 and 30 days the shortest minimum lead time for at which ACCESS-S1 and ACCESS-S2 hindcasts across initialization months. The left column shows the mean standard deviation of RMSE errors across ensemble members, while spread becomes larger than the right column shows the standard deviation lower bound of the ensemble mean RMSE error and the spread 95% confidence interval of ACCESS-S2 spread. Note the persistence and climatology forecasts errors double log scale.

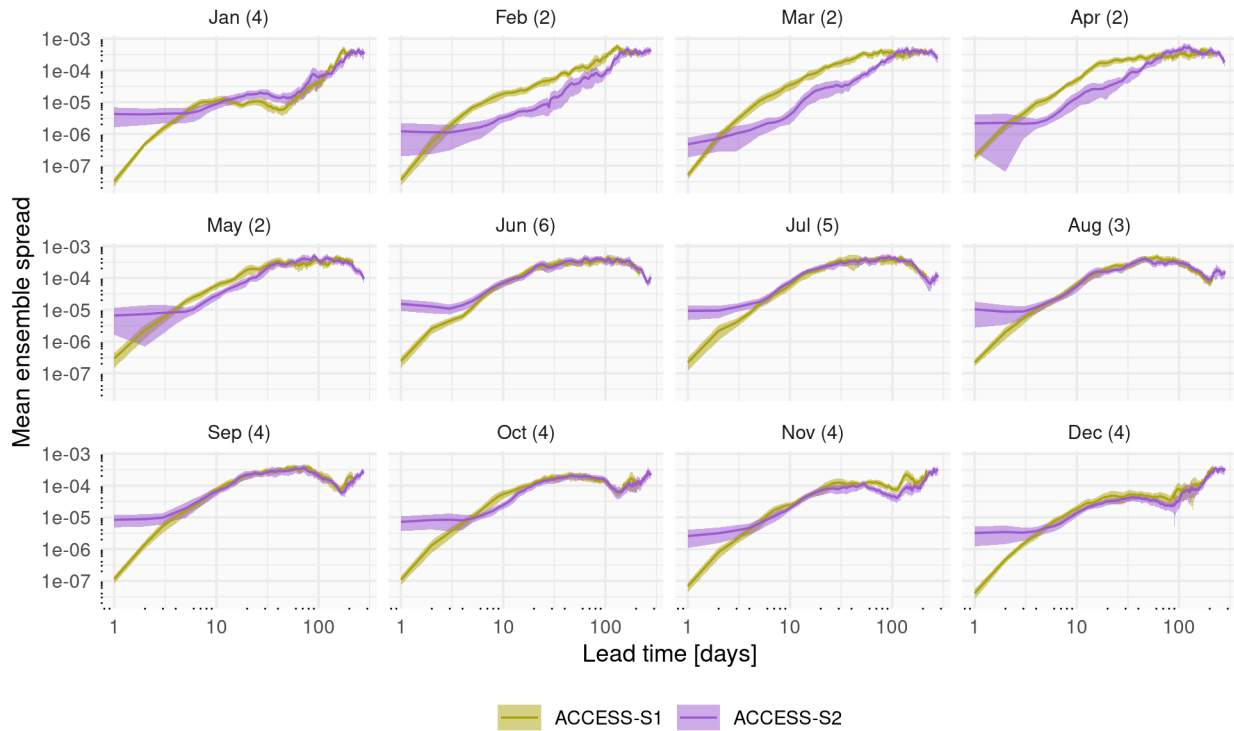


Figure A16. Mean variance of RMSE between ensemble members of each forecast. In parentheses, the shortest minimum lead time at which ACCESS-S1 mean spread becomes larger than the lower bound of the 95% confidence interval of ACCESS-S2 spread. Note the double log scale.

. B Code/Data availability

The underlying code for this study is available on GitHub: https://github.com/eliocamp/access-s2_ice-eval. Raw data of +S1 and +S2 hindcast are not available due to size. Derived datasets required to reproduce the results, including extent timeseries and error measures, are available in this Zenodo repository: <https://zenodo.org/records/17479538> (Campitelli, b)

590 . C Author contributions

EC performed the data analysis and wrote the manuscript draft. AP, JA, EL, MW and PR, performed interpretation of the results, and reviewed and edited the draft. All authors read and approved the final manuscript.

. D Competing interests

The authors declare no competing interests.

595 . E Acknowledgements

We thank the internal reviewers Bethan White and Xiaobing Zhou for their comments and feedback. This work benefited from earlier unpublished work by Laura Davies, Phil Reid, Andrew G. Marshall. This research was undertaken with the assistance of resources from the National Computational Infrastructure (NCI Australia), an NCRIS enabled capability supported by the Australian Government. This work was supported by ARC SRIEAS Grant SR200100005 Securing Antarctica's Environmental
600 Future.