Seasonal Prediction and Predictability of Regional Antarctic Sea Ice

MITCHELL BUSHUK,[a,b] MICHAEL WINTON,[a] F. ALEXANDER HAUMANN,[c] THOMAS DELWORTH,[a] FEIYU LU,[a,c]
YONGFEI ZHANG,[a,c] LIWEI JIA,[a,b] LIPING ZHANG,[a,b] WILLIAM COOKE,[a] MATTHEW HARRISON,[a] BILL HURLIN,[a]
NATHANIEL C. JOHNSON,[a] SARAH B. KAPNICK,[a] COLLEEN MCHUGH,[a,d] HIROYUKI MURAKAMI,[a,b]
ANTHONY ROSATI,[a,b] KAI-CHIH TSENG,[a,c] ANDREW T. WITTENBERG,[a] XIAOSONG YANG,[a] AND FANRONG ZENG[a]

[a] *National Oceanic and Atmospheric Administration/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey*
[b] *University Corporation for Atmospheric Research, Boulder, Colorado*
[c] *Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, New Jersey*
[d] *SAIC, Science Applications International Corporation, Reston, Virginia*

(Manuscript received 15 December 2020, in final form 30 March 2021)

ABSTRACT: Compared to the Arctic, seasonal predictions of Antarctic sea ice have received relatively little attention. In this work, we utilize three coupled dynamical prediction systems developed at the Geophysical Fluid Dynamics Laboratory to assess the seasonal prediction skill and predictability of Antarctic sea ice. These systems, based on the FLOR, SPEAR_LO, and SPEAR_MED dynamical models, differ in their coupled model components, initialization techniques, atmospheric resolution, and model biases. Using suites of retrospective initialized seasonal predictions spanning 1992–2018, we investigate the role of these factors in determining Antarctic sea ice prediction skill and examine the mechanisms of regional sea ice predictability. We find that each system is capable of skillfully predicting regional Antarctic sea ice extent (SIE) with skill that exceeds a persistence forecast. Winter SIE is skillfully predicted 11 months in advance in the Weddell, Amundsen/Bellingshausen, Indian, and west Pacific sectors, whereas winter skill is notably lower in the Ross sector. Zonally advected upper-ocean heat content anomalies are found to provide the crucial source of prediction skill for the winter sea ice edge position. The recently developed SPEAR systems are more skillful than FLOR for summer sea ice predictions, owing to improvements in sea ice concentration and sea ice thickness initialization. Summer Weddell SIE is skillfully predicted up to 9 months in advance in SPEAR_MED, due to the persistence and drift of initialized sea ice thickness anomalies from the previous winter. Overall, these results suggest a promising potential for providing operational Antarctic sea ice predictions on seasonal time scales.

KEYWORDS: Antarctica; Sea ice; Climate variability; Seasonal forecasting; Coupled models; Data assimilation

1. Introduction

The Arctic and the Antarctic are Earth's two natural sea ice environments. These regions differ in a number of fundamental aspects, including their continental geometry, ocean stratification and ventilation, atmospheric and oceanic circulations, tropical teleconnections, sea ice thickness, atmospheric chemistry, and interactions with ice sheets (e.g., Maksym 2019; Meredith et al. 2019). While the observed Arctic sea ice extent (SIE) decline was generally projected by climate models, Antarctic SIE has experienced a statistically insignificant increase over the satellite era contrary to the model-projected declines (Roach et al. 2020; Notz et al. 2020). The Arctic SIE decline and associated stakeholder interests have motivated a recent body of research on the seasonal-to-interannual predictability and prediction skill of Arctic sea ice (e.g., Guemas et al. 2016b). A key outcome of this work has been the demonstration that dynamical models can be used to skillfully predict regional Arctic sea ice on seasonal time scales (e.g., Dirkson et al. 2019). Conversely, there have been relatively few assessments of the inherent predictability or seasonal prediction skill of Antarctic sea ice. Antarctic sea ice predictions have potential utility for Southern Ocean fisheries management, shipping, conservation, scientific logistics, tourism, and predicting impacts on the Antarctic ice sheet and ice shelves (e.g., Robel 2017; Massom et al. 2018; Shepherd et al. 2018). The goal of this study is to show that dynamical prediction systems are also a valuable tool for seasonal sea ice predictions in the Antarctic.

Earlier work using coupled Earth system models (ESMs) has demonstrated that Antarctic sea ice is potentially predictable on the seasonal-to-interannual time scale. Holland et al. (2013) used perfect model (PM) experiments—which estimate the upper limit of predictability of a given ESM—to examine the initial value predictability of Antarctic sea ice in the Community Climate System Model version 3. They found that ensembles initialized on 1 January exhibit high potential predictability in all regions for the first three months, and that an eastward propagating predictability signal is retained in some sectors until the sea ice maximum is reached in September. They found that predictability was subsequently lost during the ice retreat season before reemerging the following growth season in certain locations. Holland et al. (2013) attributed this reemergence of wintertime predictability to storage of ocean heat content anomalies from the previous winter beneath the summer mixed

---

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JCLI-D-20-0965.s1.

Corresponding author: Mitch Bushuk, mitchell.bushuk@noaa.gov

DOI: 10.1175/JCLI-D-20-0965.1

layer, which return to the surface as the mixed layer deepens the following autumn.

Marchi et al. (2019) expanded upon these results, providing the first multimodel assessment of Antarctic sea ice predictability. They identified a robust seasonality to predictability, with high potential skill in winter and low potential skill in summer, and verified the winter-to-winter reemergence mechanism identified in Holland et al. (2013). Consistent with this mechanism, Marchi et al. (2019) found that predictability was strongly modulated by the strength of a model's convective mixing, as models with deeper winter mixed layers tended to have higher predictability than those with shallower mixed layers. In addition to these studies, Juricke et al. (2014) used the ECHAM6-FESOM model to show that Antarctic sea ice volume (SIV) was potentially predictable for 11 and 5 months from 1 January and 1 July start dates, respectively. Zunz et al. (2015) used an intermediate-complexity ESM to show that the sea ice edge location was potentially predictable during the first year at most locations and identified a winter reemergence of skill in the second year at some locations. Taken together, these PM studies lay an important foundation for the potential predictability of Antarctic sea ice in dynamical prediction systems. However, whether this predictability can be achieved in nature with initialized dynamical forecasts has remained an open question.

Statistical relationships and proposed physical mechanisms based on observations provide evidence that operational Antarctic sea ice prediction skill could be achievable. A number of previous studies have documented the Antarctic circumpolar wave (ACW), which is a coherent eastward propagating pattern of sea surface temperature (SST), sea ice concentration (SIC), and SIE anomalies with a period of roughly 4 years and a wavelength of 180° (White and Peterson 1996; Gloersen and White 2001; Venegas 2003; Wang et al. 2019). Gloersen and White (2001) emphasized the critical role of eastward propagating SST anomalies in storing the memory of earlier winter/spring sea ice anomalies, allowing these sea ice anomalies to reemerge the following fall/winter during the ice advance season (Holland 2014). Other studies have documented covariability between Antarctic sea ice and modes of climate variability such as El Niño–Southern Oscillation (ENSO) and the southern annular mode (SAM) (Simmonds and Jacka 1995; Yuan and Martinson 2000, 2001; Kwok and Comiso 2002; Stammerjohn et al. 2008; Simpkins et al. 2012; Kwok et al. 2016; Doddridge and Marshall 2017; Schneider and Deser 2018). Both of these modes are linked with the Amundsen Sea low (ASL), which drives regional Antarctic sea ice variability and exhibits high lagged correlation with sea ice in certain regions and seasons (Holland et al. 2017, 2018).

Chen and Yuan (2004) constructed a statistical linear Markov model for seasonal Antarctic sea ice prediction designed to capture some aspects of these observed relationships. Their model, based on SIC and atmospheric input data, showed skillful forecasts of Antarctic SIC up to nine months in advance, with particularly high skill for winter target months in the Amundsen, Bellingshausen, and Weddell Seas. There have been limited attempts to seasonally forecast Antarctic sea ice using dynamical models. Morioka et al. (2019) showed skillful predictions of austral spring [October–December (OND)] SIC for forecasts initialized on 1 September with the SINTEX-F2 dynamical model, but did not assess skill at longer lead times. Zampieri et al. (2019) investigated the subseasonal prediction skill of the Antarctic sea ice edge, finding that dynamical forecasts were skillful up to 30 days in advance and generally had lower skill than Arctic subseasonal forecasts. Guemas et al. (2016a) showed that initialized predictions of pan-Antarctic SIE from 1 May and 1 November were skillful for roughly two months, but with skill values lower than a damped anomaly persistence forecast. The Sea Ice Prediction Network South (SIPN-South) project was established in 2017 as a community effort designed to assess the current ability of prediction systems to forecast summer Antarctic sea ice (Massonnet et al. 2020). The initial three years of SIPN-South forecasts show a large spread in dynamical model–based predictions that exceeds that of the observed climatological spread; however, more forecast years are necessary to rigorously assess the skill of these systems.

In this study, we provide the first comprehensive dynamical model–based assessment of Antarctic seasonal sea ice prediction skill using initialized forecasts from three dynamical prediction systems. We subsequently use these prediction skill findings to investigate the physical mechanisms that underpin Antarctic sea ice predictability and prediction skill. Our findings show that regional Antarctic sea ice predictions are often more skillful than their Arctic counterparts, suggesting a promising potential for skillful operational forecasts of Antarctic sea ice. The outline of this paper is as follows. In section 2, we describe the seasonal prediction systems, initialization techniques, prediction experiments, and methods for forecast skill assessment. In section 3, we evaluate the initial conditions and assess prediction skill for regional SIE and the sea ice edge position. In section 4, we investigate the mechanisms of sea ice predictability in these systems, focusing on the impacts of SIE initialization, sea ice thickness, and advected upper-ocean heat content. In section 5, we consider sources of prediction error, emphasizing the importance of the ocean convective state and sea ice drift. We summarize our results in section 6.

2. Methods

a. *FLOR seasonal prediction system*

In this study, we consider predictions made with three dynamical seasonal prediction systems developed at the Geophysical Fluid Dynamics Laboratory (see Table 1). We begin by describing the Forecast-oriented Low Ocean Resolution (FLOR) system in this subsection and describe two systems based on the Seamless System for Prediction and Earth System Research (SPEAR) in the following subsection. The FLOR prediction system has been shown to skillfully predict regional SIE in the Arctic (Bushuk et al. 2017), which motivates its use for Antarctic sea ice predictions in this study.

FLOR is a global dynamical model that has nominal horizontal resolution of 0.5° in the atmosphere and land components and 1° in the ocean and sea ice components, with 50 vertical ocean levels and 32 vertical atmospheric levels (Vecchi et al. 2014). The sea ice component of FLOR is based upon the sea ice simulator version 1 (SIS1; Delworth et al. 2006), which

TABLE 1. Summary of GFDL seasonal prediction systems and retrospective forecasts considered in this study.

System property	FLOR	SPEAR_LO	SPEAR_MED
Ocean model	MOM5; 1.0°, 50 vertical levels	MOM6; 1.0°, 75 vertical levels	MOM6; 1.0°, 75 vertical levels
Sea ice model	SIS1; 1.0°, 5 category ITD	SIS2; 1.0°, 5 category ITD	SIS2; 1.0°, 5 category ITD
Atmosphere model	AM2.5; 0.5°, 32 vertical levels	AM4; 1.0°, 33 vertical levels	AM4; 0.5°, 33 vertical levels
Land model	LM3; 0.5°	LM4; 1.0°	LM4; 0.5°
Ocean data	Satellite SST, Argo, XBT, moored buoys, CTD, seal data, other WOD profiles; daily	Satellite SST, Argo, XBT, moored buoys; daily	Satellite SST, Argo, XBT, moored buoys; daily
Atmosphere data	3D temperature from NCEP-2; 6-hourly	3D temperature, winds, humidity from CFSR; 6-hourly	3D temperature, winds, humidity from CFSR; 6-hourly
Sea ice data	None	Satellite SIC used to adjust under-ice SST; daily	Satellite SIC used to adjust under-ice SST; daily
Ocean ICs	ECDA	SPEAR ODA	SPEAR ODA
Sea ice ICs	ECDA	SPEAR_LO nudged run	SPEAR_MED nudged run
Atmosphere ICs	AMIP run	SPEAR_LO nudged run	SPEAR_MED nudged run
Land ICs	AMIP run	SPEAR_LO nudged run	SPEAR_MED nudged run
Reforecast period	1981–2018	1992–2018	1992–2018
Ensemble size	12	15	15
Initialization dates	First of each month	First of each month	First of each month
Prediction length	One year	One year	One year

uses an elastic–viscous–plastic rheology to compute internal sea ice stresses (Hunke and Dukowicz 1997), an ice-thickness distribution (ITD) with five ice thickness categories (Thorndike et al. 1975; Bitz et al. 2001), a modified Semtner thermodynamic scheme with two ice layers and one snow layer (Winton 2000), and a surface-temperature-dependent albedo parameterization [see section 3.6.2 of Hunke et al. (2015)]. FLOR's atmosphere, land, and ocean components are based on Atmosphere Model version 2.5 (AM2.5; Anderson et al. 2004; Delworth et al. 2012), Land Model version 3 (LM3; Milly et al. 2014), and an updated version of Ocean Model version 2.1 (OM2.1; Gnanadesikan et al. 2006; Delworth et al. 2012), respectively.

The ocean and sea ice components of the FLOR prediction system are initialized using the GFDL Ensemble Coupled Data Assimilation system (ECDA; Zhang et al. 2007), which is based upon the CM2.1 coupled model (Delworth et al. 2006). ECDA is a weakly coupled assimilation system that uses the ensemble adjustment Kalman filter (EAKF; Anderson 2001) and a full-field initialization approach spanning the time period 1961–2018. The system assimilates 3D atmospheric temperature data from the NCEP–DOE Atmospheric Model Intercomparison Project (AMIP-II) reanalysis (Kanamitsu et al. 2002), which is used to update 3D atmospheric temperature and wind fields via flow-dependent covariances. In the ocean, the system assimilates SSTs from the Met Office Hadley Centre's sea ice and SST dataset (HadISST1; Rayner et al. 2003) (prior to 2011) and NOAA's daily Optimum Interpolation SST dataset (OISST; Reynolds et al. 2007) (post 2011), and ocean temperature and salinity (*T*/*S*) profiles. These *T*/*S* profiles come from the World Ocean Database (WOD; Levitus et al. 2013), the Global Temperature and Salinity Profile Programme (GTSPP; Sun et al. 2010), and the Argo program (Roemmich et al. 2004). The ECDA system does not explicitly assimilate sea ice data, but the sea ice state is constrained via heat fluxes and interfacial stresses from the ocean and atmosphere, associated with the data assimilation in each of these components (Bushuk et al. 2019). The sea ice state variables initialized from ECDA include the ice concentration, thickness, temperature, and snow depth in each ice-thickness category and the sea ice velocity field. FLOR's atmosphere and land initial conditions (ICs) come from a suite of "AMIP-style" atmosphere-land only simulations forced by observed SST and sea ice. This technique was used because the ECDA is based upon CM2.1, which employs a lower-resolution atmosphere than FLOR.

b. SPEAR seasonal prediction system

SPEAR is GFDL's next-generation seasonal prediction system, which recently replaced FLOR and began submitting real-time experimental seasonal predictions to the North American Multi-Model Ensemble (NMME; Kirtman et al. 2014) in February 2021. SPEAR has newly developed ocean, atmosphere, sea ice, and land component models and a fully redesigned initialization system, making it independent of the previous FLOR system (see Table 1). The SPEAR model (Delworth et al. 2020) uses the same components as GFDL's new CMIP6 models, CM4 and ESM4 (Held et al. 2019; Dunne et al. 2020), but with design choices geared toward the computational efficiency required for seasonal-to-decadal climate prediction.

SPEAR's ocean and sea ice components employ a nominal horizontal resolution of 1° with 75 vertical ocean levels and are based upon the Modular Ocean Model version 6 (MOM6) and the Sea Ice Simulator version 2 (SIS2) (Adcroft et al. 2019). Two varieties of SPEAR have been developed, SPEAR_LO and SPEAR_MED, which have different atmospheric and land horizontal resolutions of 1° and 0.5°, respectively. The models both use 33 vertical atmospheric levels and are based on Atmosphere Model version 4 and Land Model version 4 (Zhao et al. 2018a,b). Compared to SIS1, the SIS2 sea ice model features improved shortwave radiation physics based on the delta-Eddington scheme of Briegleb and Light (2007), better vertical resolution with four ice layers and one snow layer, updated thermodynamics with improved conservation properties (Bitz and Lipscomb 1999), and ice dynamics that are solved using a C-grid stencil as opposed to the B-grid used by SIS1 (Bouillon et al. 2009). As in SIS1, SIS2 uses an elastic–viscous–plastic rheology and an ITD with five ice thickness categories. Both the SIS1 and SIS2 models do not include a subgrid ice ridging scheme, a landfast ice parameterization, or a side melt scheme.

The ICs used for the SPEAR_LO and SPEAR_MED predictions come from two separate assimilation experiments spanning 1990–2018. The ocean ICs come from an ocean data assimilation (ODA) system based on the SPEAR_LO model (Lu et al. 2020). The ODA system uses an EAKF to assimilate daily SST from NOAA's OISST product and *T*/*S* profiles from Argo floats, expendable bathythermograph data (XBT), and tropical moorings. The ODA system does not currently assimilate ship-based conductivity–temperature–depth (CTD), drifting buoy, and instrumented marine mammal data, each of which have coverage in the polar regions. The atmospheric state is unconstrained in the ODA system, as this was found to improve the ensemble spread of ocean state variables in the assimilation run and also reduce the magnitude of assimilation increments.

The sea ice, atmosphere, and land ICs come from nudged ensemble experiments performed with SPEAR_LO and SPEAR_MED, which incorporate both atmospheric, SST, and SIC constraints. In these nudged runs, the 3D atmospheric temperature, wind, and humidity fields are nudged toward the NOAA/NCEP Climate Forecast System Reanalysis (CFSR; Saha et al. 2010) and the SSTs are nudged toward daily OISST data. The nudged experiments are run as 15-member ensembles initialized from members 1–15 of the SPEAR_LO and SPEAR_MED large ensembles, described in the following subsection. To improve performance near and under sea ice, the raw OISST data are modified prior to nudging and assimilation. Using the daily OISST SIC data, all ice-covered grid points are identified based on a threshold of SIC $\geq$ 30%. At these grid points, the SST value is replaced by the freezing point temperature of seawater (T_f) based on the model-predicted sea surface salinity (SSS) and the relationship $T_f = -0.054 \times \mathrm{SSS}$. The SST nudging uses a piston velocity of 4 m per day (corresponding to a 12.5 day *e*-folding time scale for a 50-m mixed layer) and the atmospheric nudging is performed using a 6-h *e*-folding time scale for temperature and wind, and a 24-h time scale for humidity. These atmospheric data provide critical thermodynamic and dynamic constraints on sea ice thickness (SIT), which are not present in the ODA experiment. The nudged and ODA experiments both utilize a common SST dataset, which allows the ICs from these two runs to be combined in order to initialize the coupled model. The sea ice state variables initialized from the SPEAR_LO and SPEAR_MED nudged runs include the ice concentration, thickness, temperature, and snow depth in each ice-thickness category and the sea ice velocity field.

The SPEAR system also uses an ocean tendency adjustment (OTA) procedure to reduce the model's ocean bias. This procedure applies the climatological increments obtained from a prior ODA run as 3D temperature and salinity tendency terms to the free-running ocean model. This technique reduces model drift and improves both assimilation accuracy and prediction skill in coupled model predictions of El Niño–Southern Oscillation (Lu et al. 2020).

c. *Retrospective seasonal predictions and large ensemble experiments*

We analyze suites of retrospective seasonal prediction experiments performed with each prediction system (see Table 1). The ensemble predictions are initialized on the first of each month and integrated for one year. The FLOR predictions are run with 12 ensemble members and span the period 1981–2018. The SPEAR_LO and SPEAR_MED predictions are both run with 15 ensemble members and span 1992–2018.

We also consider large ensemble (LE) experiments of historical and scenario simulations performed with each model. Ensemble means of these LEs are used to assess the biases of the free-running models. The FLOR LE is a 30-member ensemble that uses historical radiative forcings up to 2005 and representative concentration pathway 8.5 (RCP8.5; Meinshausen et al. 2011) from 2006 to 2100 (Bushuk et al. 2020). The SPEAR_LO and SPEAR_MED large ensembles are 30-member ensembles that use historical forcings up to 2014 and shared socioeconomic pathway 5–8.5 (SSP5-8.5; Riahi et al. 2017) from 2015–2100 (Delworth et al. 2020). The ICs for each LE are taken from different years of preindustrial control simulations performed with each model. The IC years were chosen with 10-yr spacing in FLOR and 20-yr spacing in SPEAR, designed to sample different phases of internal climate variability.

d. *Observational data*

We evaluate the model predictions using monthly averaged passive microwave satellite SIC observations from the National Snow and Ice Data Center processed using the NASA team algorithm (dataset ID: NSIDC-0051; Cavalieri et al. 1996). The observed SIC data are regridded to the model grid in order to compute skill metrics. Sea ice drift is assessed using the low-resolution sea ice drift product of the EUMETSAT Ocean and Sea Ice Satellite Application Facility (OSISAF; Lavergne et al. 2010). We assess the ocean state using 3-monthly temperature anomaly data from the World Ocean Atlas (WOA) Global Ocean Heat and Salt Content dataset (GOHSC; Levitus et al. 2012). We also use an observed mixed-layer depth (MLD) climatology computed from *T*/*S* profiles from Argo (snapshot of Argo Global Data Assembly Centre from 9 April 2020; Argo 2020) and ship-based conductivity–temperature–depth data (downloaded in NetCDF format 9 June 2020 from the NOAA World Ocean Database 2018; Boyer et al. 2018). The MLD is defined as the depth where the surface-referenced density exceeds the 10-m reference value by 0.03 $\mathrm{kg\,m^{-3}}$ (de Boyer Montégut et al. 2004).

e. *Skill metrics and significance testing*

We assess prediction skill using the anomaly correlation coefficient (ACC) and the mean-squared error skill score (MSSS). Throughout the manuscript, the term "target month" will be used to refer to the month that is being predicted. We let

o be an observed time series of some quantity of interest, such as regional SIE in a given target month. We let o_i be the observed value at time i, N be the number of years in the observed time series, K be the number of prediction ensemble members, and τ be the forecast lead time. We let $p_{ij}(\tau)$ be the predicted value of the jth ensemble member that is initialized τ months prior to time i. We take the ensemble-mean prediction $p_i(\tau)$ as our lead τ prediction of o_i, given by

$$p_i(\tau) = \frac{1}{K}\sum_{j=1}^{K} p_{ij}(\tau). \tag{1}$$

We let an overbar denote the temporal mean over the N samples. The ACC is the Pearson correlation coefficient between the predicted and observed time series, given by

$$\mathrm{ACC}(\tau) = \frac{\sum_{i=1}^{N}[p_i(\tau) - \overline{p(\tau)}](o_i - \overline{o})}{\sqrt{\sum_{i=1}^{N}[p_i(\tau) - \overline{p(\tau)}]^2}\sqrt{\sum_{i=1}^{N}(o_i - \overline{o})^2}}. \tag{2}$$

We also compute a "detrended ACC" skill score, where the anomalies are computed relative to a linear trend prediction. This metric removes skill associated with the secular trend, focusing on interannual anomalies. Specifically, for each time i we compute a linear trend prediction o_i^L based on all past observed data (up to time $i - 1$). We similarly compute a linear trend prediction $p_i^L(\tau)$ based on all past predicted data. The detrended ACC is then given by

$$\mathrm{ACC}_{\mathrm{detrend}}(\tau) = \frac{\sum_{i=1}^{N}[p_i(\tau) - p_i^L(\tau)](o_i - o_i^L)}{\sqrt{\sum_{i=1}^{N}[p_i(\tau) - p_i^L(\tau)]^2}\sqrt{\sum_{i=1}^{N}(o_i - o_i^L)^2}}. \tag{3}$$

This approach is chosen to avoid using future data when computing the detrended anomaly in a given year, as is the case when a single linear trend is applied to the full time series. Note that these two detrending approaches produce very similar ACC values for Antarctic SIE. To avoid overfitting to shorter-term variations, we assume a linear trend of zero for the first 10 years.

The MSSS (Murphy 1988) is a skill score based on a comparison of mean-square errors (MSEs) between the model predictions and a reference forecast, given by

$$\mathrm{MSSS}(\tau) = 1 - \frac{\mathrm{MSE}(\tau)}{\mathrm{MSE}_{\mathrm{clim}}}, \tag{4}$$

where

$$\mathrm{MSE}(\tau) = \frac{\sum_{i=1}^{N}[p_i(\tau) - o_i]^2}{N}, \tag{5}$$

and

$$\mathrm{MSE}_{\mathrm{clim}} = \frac{\sum_{i=1}^{N}(\overline{o} - o_i)^2}{N}. \tag{6}$$

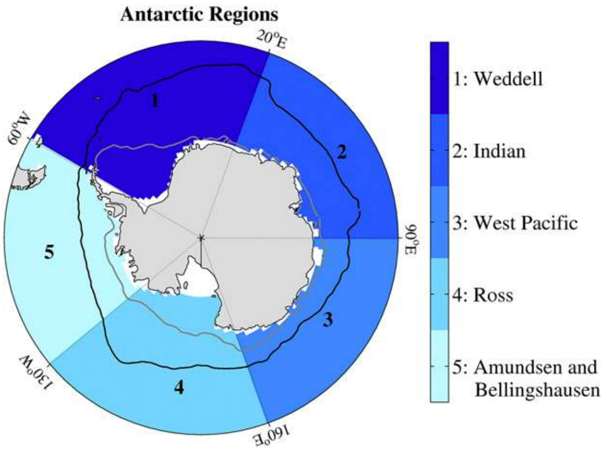

FIG. 1. The Antarctic regions considered in this study. The black and gray contours indicate the observed climatological (1992–2018) sea ice edge in September and March, respectively. The regions have longitude boundaries of 60°W–20°E (Weddell), 20°–90°E (Indian), 90°–160°E (west Pacific), 160°E–130°W (Ross), and 130°–60°W (Amundsen and Bellingshausen), respectively.

The MSSS is directly related to the ACC using the decomposition of Murphy (1988), which shows that

$$\mathrm{MSSS}(\tau) = \mathrm{ACC}^2(\tau) - \left[\mathrm{ACC}(\tau) - \frac{\sigma_p}{\sigma_o}\right]^2 - \frac{[\overline{p(\tau)} - \overline{o}]^2}{\sigma_o^2}, \tag{7}$$

where σ_p and σ_o are standard deviations of the predicted and observed time series, respectively. The second and third terms on the right-hand side make negative definite contributions to the MSSS related to conditional and mean forecast biases, respectively. We also consider a detrended MSSS metric, where MSEs are computed using detrended anomalies as defined above. MSSS values close to 1 indicate a highly skillful forecast, a value of 0 indicates no skill above climatology, and negative values indicate worse performance than a climatological forecast.

We test the ACC and MSSS values for statistical significance using a bootstrapped resampling procedure applied to the prediction ensemble. This approach repeatedly resamples the prediction ensemble (with replacement) in order to produce empirical distribution estimates of the ACC and MSSS statistics (Efron 1982). For each target month, lead time, and region, we compute a 95% confidence interval based on a bootstrapped distribution of 1000 realizations. If the lower limit of this confidence interval exceeds zero, we report the skill to be statistically significant at the 95% confidence level.

All analysis in this manuscript is based on monthly mean fields. The terminology "lead 0" refers to a forecast initialized on the first of the month predicting that month's mean value, and longer-lead forecasts are defined correspondingly. We assess forecast skill for SIE, defined as the areal sum of all grid points with SIC $\geq$ 15%. We also consider forecast skill of the sea ice edge position at each longitude, defined as the maximum northerly extent of sea ice based on a 15% SIC threshold.

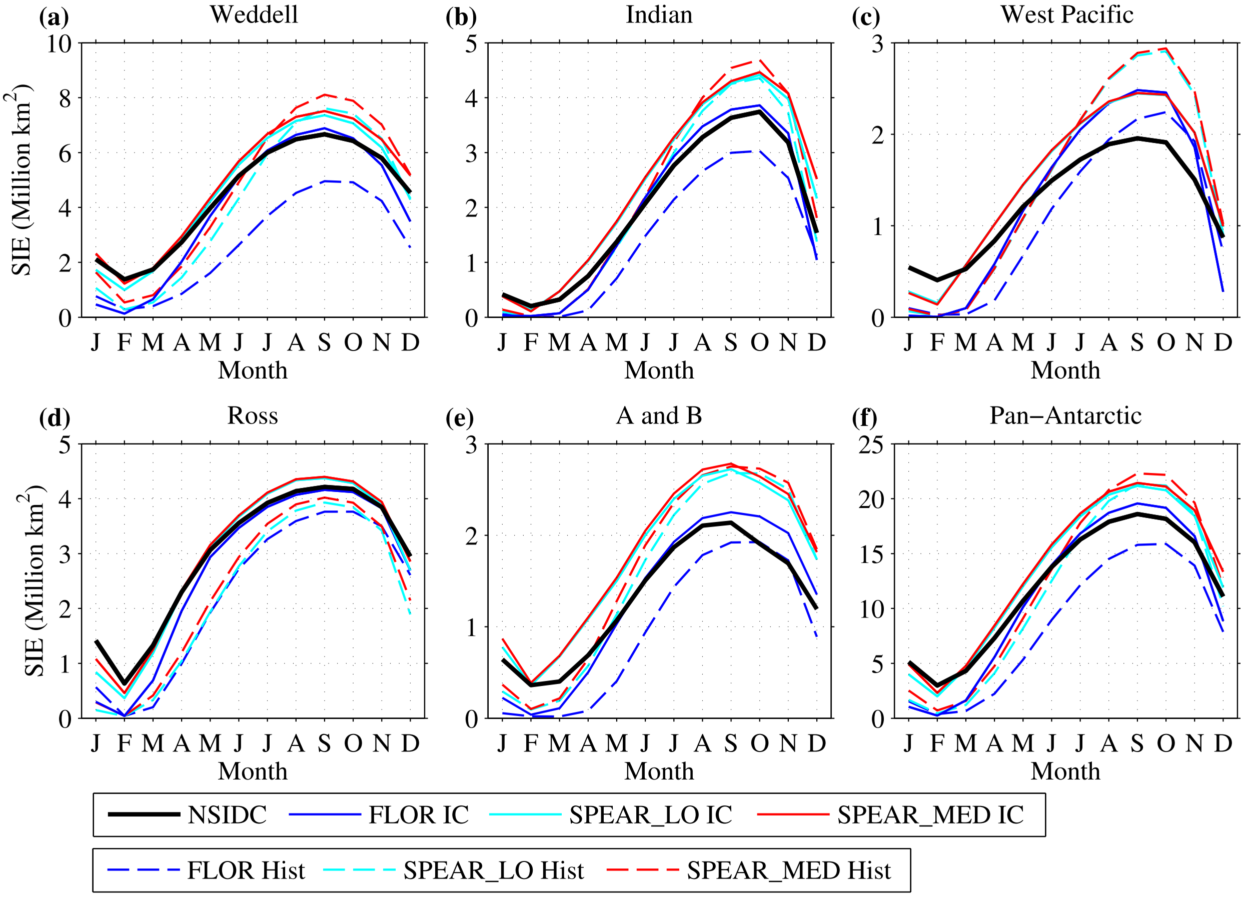

FIG. 2. Antarctic regional SIE climatologies compared to NSIDC observations (black). Solid lines show the SIE climatologies from the assimilation runs used for sea ice ICs for FLOR (blue), SPEAR_LO (cyan), and SPEAR_MED (red). Dashed lines show the SIE climatologies from historical simulations of each model. These climatologies are computed over years 1992–2018, which are common to all experiments.

f. Reference forecasts

In addition to the climatological and linear trend reference forecasts that underpin the skill metrics defined above, we also compare prediction skill to an anomaly persistence reference forecast. The anomaly persistence forecast uses the observed sea ice anomaly at the initial forecast time. These observed anomalies can be either defined relative to the linear trend, termed detrended anomalies, or relative to the climatology, termed nondetrended anomalies. We consider persistence forecasts based on both detrended and nondetrended anomalies. We also consider an advected anomaly persistence forecast for the sea ice edge position, which accounts for zonal advection of the sea ice edge anomaly. Specifically, our advected persistence prediction of the sea ice edge position anomaly $e'(x, t)$ is given by

$$e'(x,t) = e'(x - c\tau, t - \tau), \tag{8}$$

where t is forecast target time, τ is the forecast lead time, x is the longitude, and c is the eastward advection speed. The advection speed is chosen as $c = 360° (7\ \mathrm{yr})^{-1}$ based on the documented period of the ACW (Gloersen and White 2001).

3. Regional Antarctic sea ice prediction skill

a. Climatology and interannual variability of sea ice initial conditions

We first consider the quality of Antarctic sea ice initial conditions (ICs) in the FLOR, SPEAR_LO, and SPEAR_MED systems. The ICs are assessed using monthly mean values from the assimilation runs used to initialize sea ice (see Table 1). We compute regional SIE in the five Antarctic sectors shown in Fig. 1 and over a pan-Antarctic domain. Figure 2 shows regional SIE climatologies over the period 1992–2018, which is common to all experiments. The FLOR ICs have a notable low bias in austral summer SIE in all Antarctic regions, similar to the bias of the free-running FLOR and SPEAR models (see dashed curves for FLOR Hist, SPEAR_LO Hist, and SPEAR_MED Hist). This low summer bias is substantially improved in the SPEAR ICs, which show a reasonably good agreement with observed regional summer SIE. The SPEAR ICs and historical simulations are generally biased high for austral winter SIE, with the exception of the Ross Sea, where the historical simulations have a modest low bias. The FLOR ICs have small winter biases in most regions, except for the

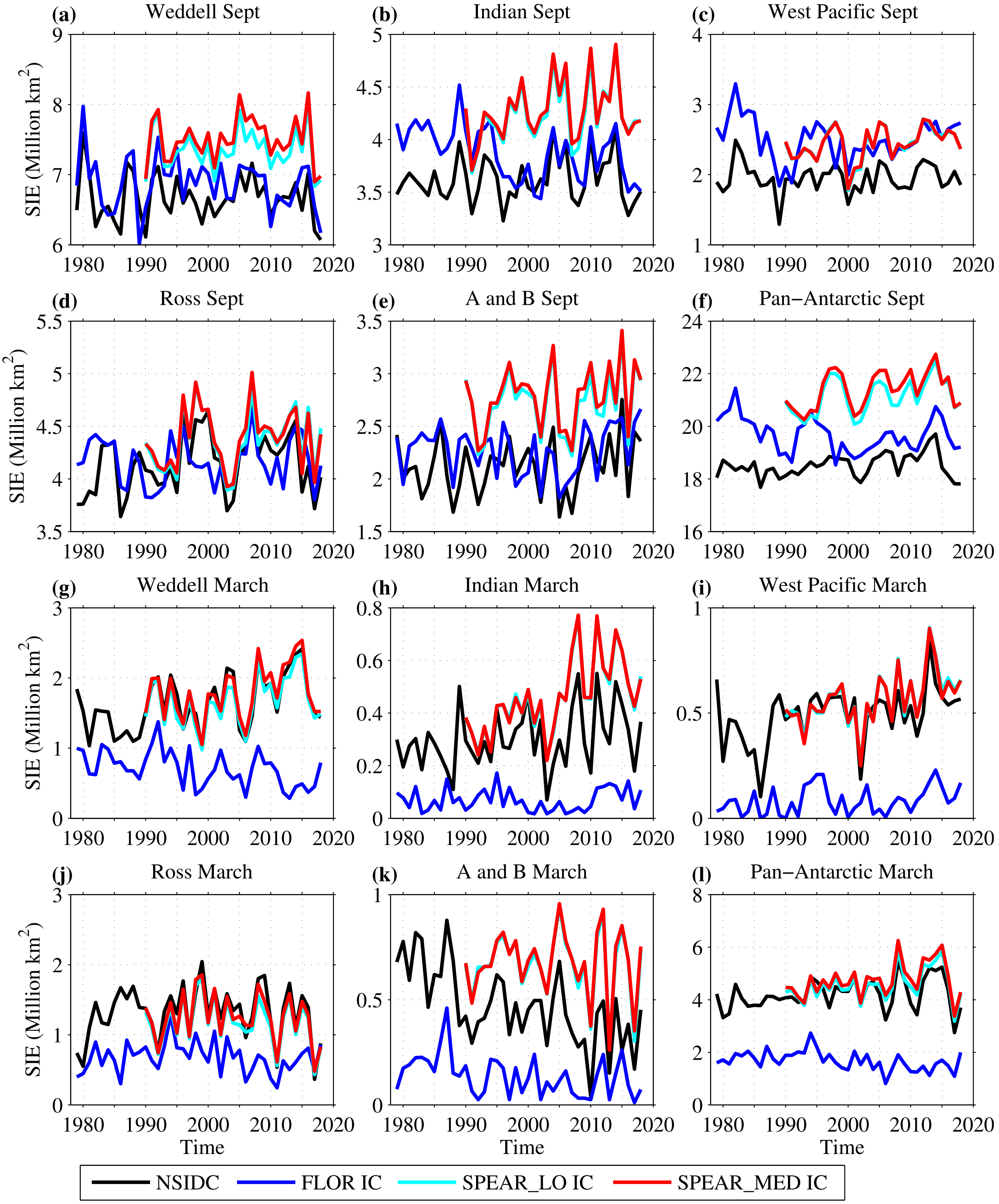

FIG. 3. Regional SIE time series for (a)–(f) September and (g)–(l) March from NSIDC observations (black) and the assimilation runs used for sea ice ICs for FLOR (blue), SPEAR_LO (cyan), and SPEAR_MED (red).

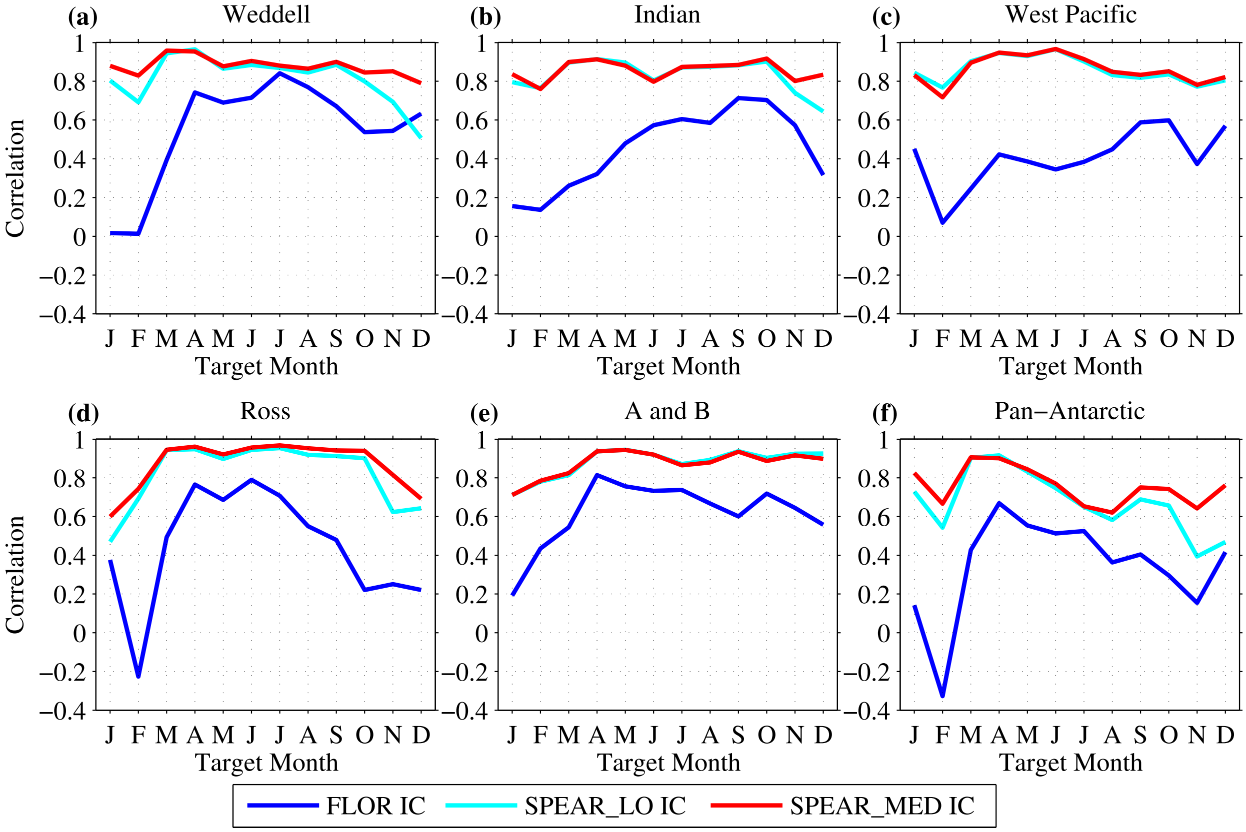

FIG. 4. Detrended correlations with observed regional SIE for the FLOR (blue), SPEAR_LO (cyan), and SPEAR_MED (red) SIE ICs.

west Pacific sector, which shows a positive bias. The FLOR historical simulations are also biased high in the west Pacific and are generally biased low in other regions.

In Fig. 3, we plot time series of regional September and March SIE, respectively, from FLOR, SPEAR_LO, and SPEAR_MED ICs and NSIDC observations. We show March SIE instead of February, the month of the SIE minimum, because FLOR is essentially ice free in February. Both the FLOR and SPEAR ICs capture some aspects of the observed interannual variability of regional Antarctic SIE, with the SPEAR systems consistently outperforming FLOR in all regions and all seasons. This difference between the prediction systems is shown quantitatively in Fig. 4, which plots the detrended correlation values between the ICs and observations. Note that the correlation values and qualitative differences between models are similar if the nondetrended time series are used. The SPEAR IC detrended correlations generally exceed 0.8, whereas the FLOR IC correlations are generally lower than this value. The differences between SPEAR and FLOR are substantial in all seasons, but particularly notable in summer, where the FLOR system has low correlation values. The SPEAR_LO and SPEAR_MED correlations are generally similar, with the exception of the spring and summer months in the Weddell, Ross, and pan-Antarctic domains, which have higher correlation values in the SPEAR_MED system. We also find that, in all three systems, the pan-Antarctic SIE correlations are generally lower than those found in the regional domains. These low pan-Antarctic SIE correlations suggest that some cancellation of regional-scale anomalies occurs when pan-Antarctic SIE is computed, and that the resulting integrated quantity is less well captured than the regional anomalies in these systems.

b. *Regional sea ice extent prediction skill*

Figure 5 shows the detrended ACC for regional SIE predictions in the FLOR, SPEAR_LO, and SPEAR_MED prediction systems. The prediction skill of nondetrended anomalies is similar to the detrended skill (see Fig. S1 in the online supplemental material), indicating that seasonal Antarctic sea ice skill derives primarily from initial-value predictability. This differs from seasonal predictions in the Arctic, where forced sea ice trends represent the dominant source of predictability for nondetrended anomalies (e.g., Sigmond et al. 2013; Wang et al. 2013). In all three systems, we find that the regional SIE skill generically exceeds the skill of an anomaly persistence forecast, as indicated by the triangle markers in Fig. 5. For some target months and some regions, prediction skill is statistically significant at least to 11 months in advance, indicating that these prediction systems are successfully capturing some aspects of interannual climate variability in the Southern Ocean. The correlation skill structures vary by

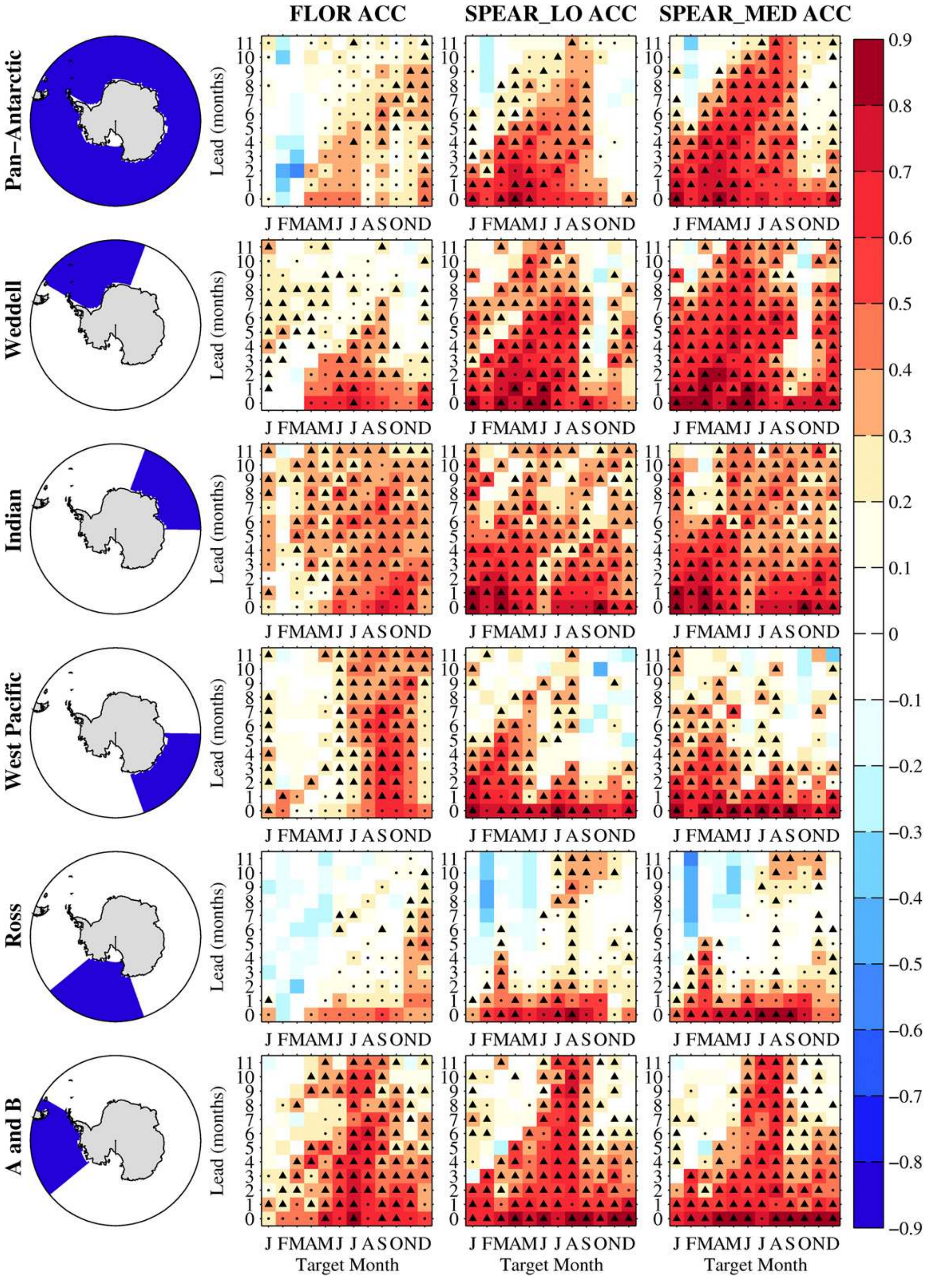

FIG. 5. Seasonal prediction skill (ACC) for detrended regional Antarctic SIE for different target months and forecast lead times. Triangle and dot markers indicate months in which the ACC values are statistically significant at the 95% confidence level based on a bootstrapped resampling procedure. Triangles indicate months where the model's skill beats the persistence forecast, and dots indicate months where the model's skill is statistically significant but does not beat persistence.

region, season, and model. We highlight the key prediction skill features below.

The Weddell Sea is a region of notably high prediction skill, particularly in the SPEAR_MED system, which shows skillful predictions at leads times of 6–11 months for most target months. SPEAR_LO also shows skillful predictions in the Weddell region, but generally has lower skill values than SPEAR_MED. A similar skill reduction from SPEAR_MED to SPEAR_LO is found for pan-Antarctic SIE, suggesting that there could be value to the higher-resolution atmosphere employed by the SPEAR_MED model. This skill difference is further explored in section 5a. The performance of the SPEAR systems is generally similar across the other Antarctic regions. Unlike SPEAR, the FLOR system has low skill for summer sea ice in the Weddell sea. FLOR performs poorly for summer sea ice predictions in all Antarctic regions, due to the low quality of its summer sea ice ICs (Fig. 4) and the negative summer sea ice bias of the free-running model (Fig. 2). The SPEAR systems show notable improvements to FLOR for summer predictions in all Antarctic regions. The sources of this summer prediction skill are further examined in sections 4a and 4b.

FLOR shows skillful predictions of autumn and winter Weddell sea ice, which exceed the persistence forecast at lead times of 3–7 months. Interestingly, the Weddell Sea correlation structures show diagonal skill features in each of the three systems. These diagonal features correspond to a reduction in prediction skill at a fixed initialization month. In SPEAR_LO and SPEAR_MED, the diagonal features correspond to initialization months June or July (i.e., 12 months prior to winter), whereas FLOR shows a diagonal feature corresponding to February- or March-initialized predictions. The fact that SPEAR_LO and SPEAR_MED have winter skill for predictions initialized prior to March implies that the diagonal feature seen in FLOR is not a fundamental aspect of Weddell sea ice predictability, but rather a deficiency of the FLOR system. Interestingly, the pan-Antarctic predictions also display diagonal skill features, corresponding to initialization months November and September in SPEAR_LO and SPEAR_MED, respectively.

The Amundsen and Bellingshausen Seas stand out as a region with high winter prediction skill and a high degree of consistency across the three systems. Target months of June–August are skillfully predicted at least 11 months in advance in each system. We explore the sources of this winter prediction skill in section 4c. On the other hand, the Ross Sea stands out for its notably poor prediction skill across each of the three systems, which display little skill beyond one month lead times. Some aspects of this poor prediction skill are discussed in section 5b.

The FLOR system has winter prediction skill at 11-month lead time in both the Indian and west Pacific sectors. This winter skill is similar to SPEAR in the Indian sector and higher than SPEAR in the west Pacific sector, which does not show continuous skill beyond 4-month lead times. The poor winter skill of SPEAR in the west Pacific is possibly related to the large positive winter SIE bias in the SPEAR ICs and free-running model (Fig. 2c). The SPEAR systems display summer SIE prediction skill in these regions at 3–11-month lead times, which is higher than the skill of the FLOR system.

Pan-Antarctic SIE integrates these diverse regional contributions into a single metric. We find that the SPEAR systems are more skillful than FLOR at the Pan-Antarctic scale. A particular prediction of interest is the 1 December initialized prediction of summer Antarctic SIE, which the SIPN-South project has been collecting from the sea ice community since the 2017/18 melt season (Massonnet et al. 2020). For predictions initialized from 1 December, the SPEAR_MED system has detrended ACC skill for February and March pan-Antarctic SIE of 0.58 and 0.67, respectively, indicating that skillful predictions are realizable at this lead time. In addition to the diagonal correlation structures discussed earlier, the pan-Antarctic SPEAR predictions show a sharp drop in skill for target months October and November. This skill drop off suggests that, in these systems, skillful prediction of the winter sea ice maximum does not imply skillful prediction of spring sea ice anomalies, similar to the finding of Holland et al. (2013). The pan-Antarctic FLOR predictions are limited by their poor Ross Sea skill. We find that prediction skill increases to a level resembling the FLOR Weddell Sea skill if the Ross Sea domain is excluded from the analysis (not shown).

The squared ACC skill can be interpreted as the variance explained by a regression-adjusted forecast, which is free of conditional and systematic biases [see Eq. (7)], whereas the MSSS is sensitive to these biases. Since regression-adjusting using the full hindcast set can artificially inflate MSSS skill, we bias correct the predictions using a leave-one-out linear regression adjustment (Manzanas et al. 2019). This adjustment corrects both the amplitude of predicted anomalies (conditional biases) and the mean predicted value (mean biases). Figure 6 shows the MSSS for detrended regional SIE predictions. We find that the MSSS values mirror the squared ACC plots but generally have lower values (cf. Fig. 6 and Fig. S3), indicating forecast degradations due to conditional biases. This suggests that the perceived skill based on ACC in Fig. 5 cannot always be achieved in practice. In Fig. S4, we consider MSSS skill for predictions that have been mean bias corrected, but have not been corrected for conditional biases. We find that the MSSS skill is substantially degraded in this case, confirming that there are notable conditional biases in these prediction systems.

The spread of the prediction ensembles increases with lead time, consistent with lower predictability at longer lead times (not shown). In general, the ensembles are approximately Gaussian distributed, however the ensemble sizes preclude a precise characterization of their probability distribution functions. Applying a chi-square goodness-of-fit test with a 5% significance level to the prediction ensembles (Pearson 1900), we find that the null hypothesis that the ensemble distributions are Gaussian is rejected 7% of the time in FLOR, 8% of the time in SPEAR_LO, and 5% of the time in SPEAR_MED. Therefore, the SPEAR_MED ensembles are generally indistinguishable from Gaussian distributions, whereas the FLOR and SPEAR_LO ensembles occasionally display non-Gaussian features. The non-Gaussianity in FLOR arises for summer target months due to certain regions going ice-free. The non-Gaussian features in SPEAR_LO arise for winter and spring target months at short lead times (0–2 months) due to differences in oceanic convective activity across ensemble

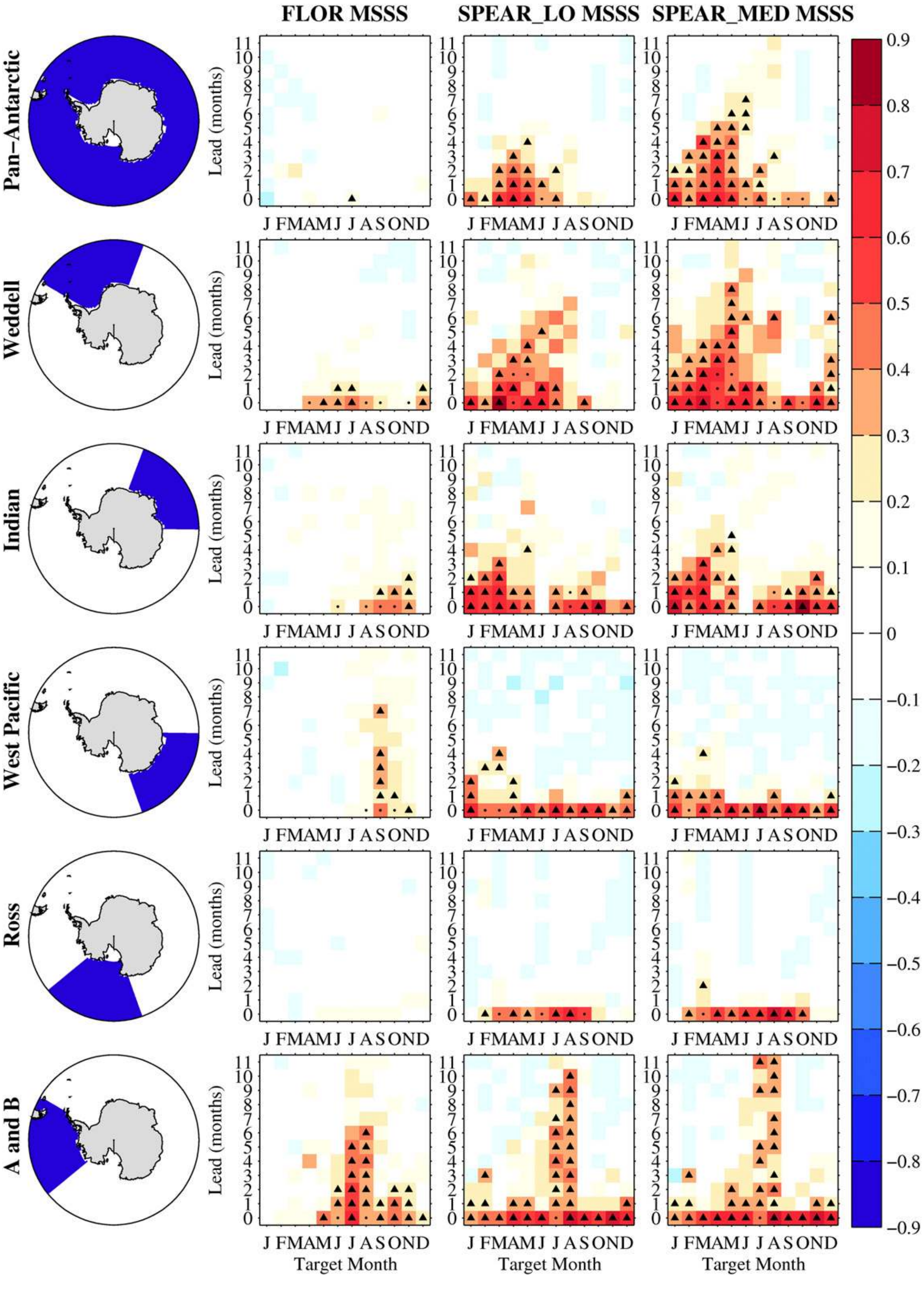

FIG. 6. Seasonal prediction skill (MSSS) for detrended regional Antarctic SIE for different target months and forecast lead times. The predictions have been bias corrected via a leave-one-out linear regression adjustment. Triangle and dot markers indicate months in which the MSSS values are statistically significant at the 95% confidence level based on a bootstrapped resampling procedure. Triangles indicate months where the model's skill beats the persistence forecast, and dots indicate months where the model's skill is statistically significant but does not beat persistence. This metric is comparable to ACC^2 (Fig. S3).

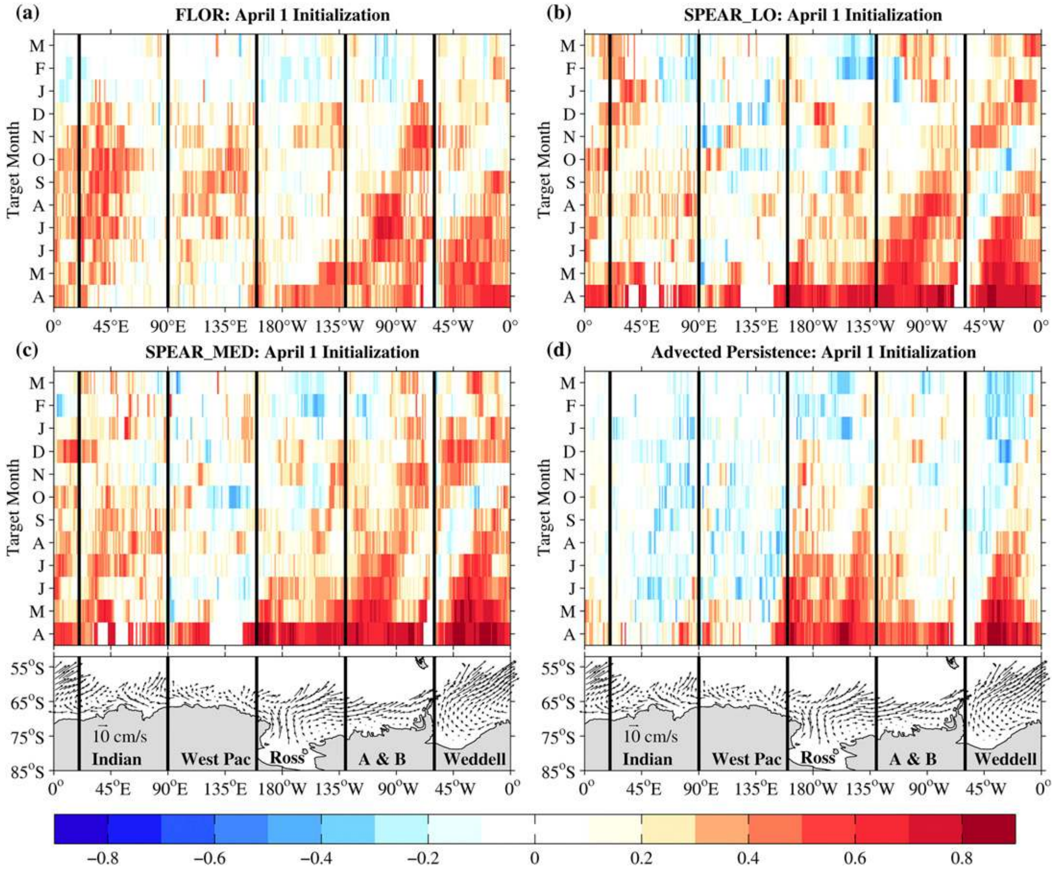

FIG. 7. Prediction skill of the sea ice edge position (defined as maximum northerly extent of sea ice) for forecasts initialized on 1 April. Detrended ACC values are plotted for each longitude and target month. Climatological observed winter [July–September (JAS)] sea ice drift is plotted as vectors.

members in the SPEAR_LO nudged run (see further discussion in section 5a).

c. *Sea ice edge predictions*

Next, we take a more fine-grained approach and consider prediction skill of the sea ice edge position at each longitude. Inspired by the analogous figure of Holland et al. (2013), Fig. 7 shows detrended ACC skill of the ice edge position for predictions initialized on 1 April. Results are qualitatively similar for other choices of initialization month. We find zones of high prediction skill that exhibit an eastward propagation over the autumn and winter months, particularly in the Amundsen/Bellingshausen, Weddell, and Indian sectors. This eastward propagation is consistent with an eastward advection of anomalies via the Antarctic Circumpolar Current, as has been documented in observations (White and Peterson 1996; Gloersen and White 2001).

Interestingly, the initialized prediction skill shown in Fig. 7 bears a close resemblance to the perfect model skill of Holland et al. (2013), suggesting that the prediction systems are capturing some of the predictability mechanisms present in the perfect model context. However, Marchi et al. (2019) showed that while eastward propagation of predictability was a robust feature across models, the spatial zones of high predictability are model dependent. Analysis of other prediction systems is needed to assess the robustness of the skill patterns identified in Fig. 7. The FLOR and SPEAR predictions generally exceed the skill of the advected sea ice persistence forecast [Fig. 7d; Eq. (8)], implying other sources of predictability beyond advected sea ice anomalies.

Figure 8 takes a different vantage point, showing September ice edge prediction skill for different lead times. The FLOR and SPEAR systems each have "gaps" in their prediction skill in the western Ross and western Weddell Seas. Both of these regions are characterized by northward sea ice drift, suggesting that prediction skill may be lower in areas of strong northward ice advection. Conversely, prediction skill is high in the Amundsen/Bellingshausen, eastern Weddell, and western Indian sectors, regions dominated by strong eastward sea ice drift. We discuss the connection between ice drift and prediction skill in section 5b.

4. Mechanisms of regional Antarctic sea ice predictability

In this section, we seek to understand some aspects of the physical mechanisms underlying the prediction skill presented in Figs. 5–8. We focus here on the roles of 1) sea ice extent

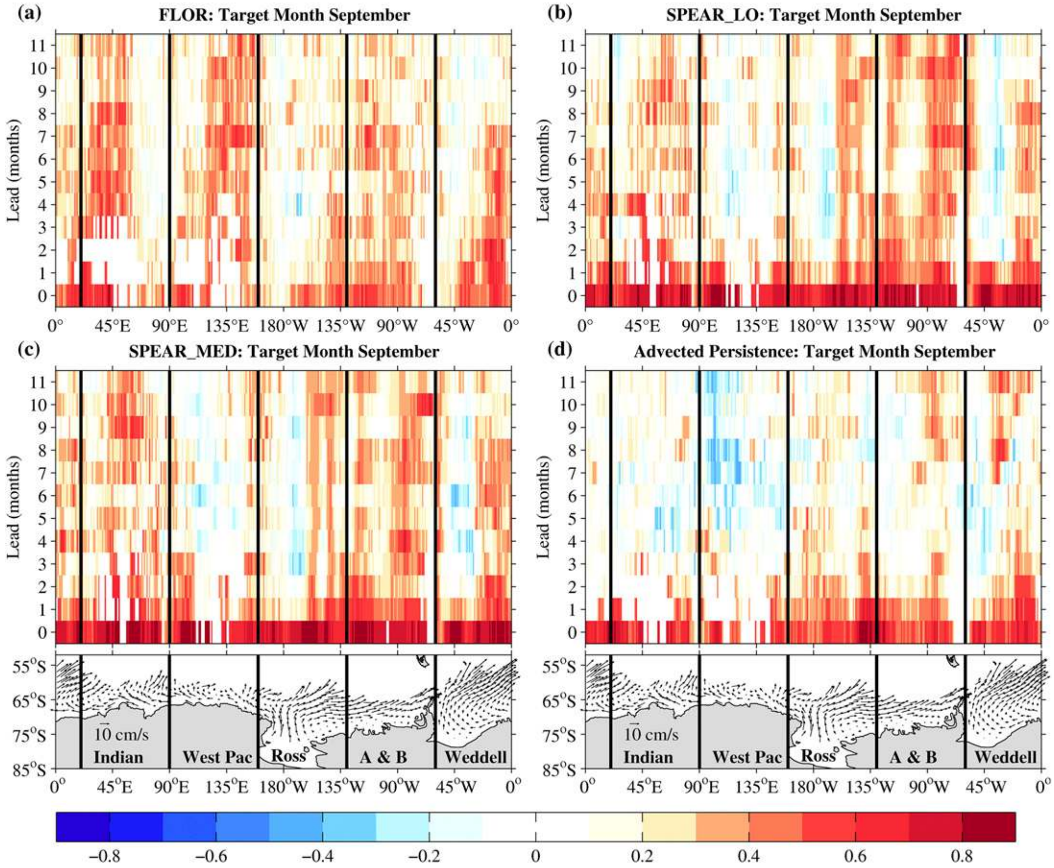

FIG. 8. Prediction skill of the sea ice edge position (defined as maximum northerly extent of sea ice) for target month September. Detrended ACC values are plotted for each longitude and lead time. Climatological observed winter (JAS) sea ice drift is plotted as vectors.

initialization, 2) sea ice thickness, and 3) advected upper-ocean heat content.

a. Sea ice extent initialization

Is there a connection between SIE ICs and prediction skill? To explore this, in Fig. 9 we compare detrended ACC skill at different lead times to detrended correlation values between regional SIE ICs and NSIDC observations. Note that the SPEAR_LO results are qualitatively similar to SPEAR_MED, and are not shown here for visual clarity. The regional SIE ICs set an upper bound to prediction skill (black curves in Fig. 9) and explain some aspects of the regional and intermodel differences in prediction skill. For example, compared to FLOR, the SPEAR_MED predictions clearly benefit from their higher-quality SIE ICs. These SIE ICs directly improve predictions at short lead times (0 and 1 months) via SIE anomaly persistence, whereas other mechanisms become relevant at longer lead times. It is important to note that high-quality SIE ICs do not imply skillful predictions, as evidenced by the SPEAR_MED Ross Sea predictions. Therefore, high-quality SIE initialization represents a necessary but insufficient condition for a skillful Southern Ocean prediction system at lead times longer than 1–2 months.

b. Sea ice thickness

Sea ice thickness (SIT) has been shown to be the crucial source of predictability for summer sea ice predictions in the Arctic (e.g., Chevallier and Salas y Mélia 2012; Bonan et al. 2019) due to the multimonth persistence and relatively large spatial autocorrelation of SIT anomalies (Blanchard-Wrigglesworth and Bitz 2014; Ponsoni et al. 2020). Previous work has suggested that the efficacy of SIT as a predictor in the Antarctic is reduced relative to the Arctic due to its thinner ice pack and smaller fraction of multiyear ice (e.g., Holland et al. 2013; Ordoñez et al. 2018; Marchi et al. 2019). However, the Weddell Sea, which dominates the multiyear ice coverage in the Southern Ocean, is a region where SIT could potentially provide a key source of predictability. In Fig. 10, we investigate this mechanism, plotting detrended correlations between observed March Weddell SIE and the SPEAR_MED SIT ICs that were used to initialize the predictions at different lead times. We find positive correlations between March SIE and earlier SIT ICs, which extend back to the previous winter. These positive correlations are consistent with the physical expectation that anomalously thick ice requires additional energy to melt and therefore leads to positive SIE anomalies during the melt season. These correlations

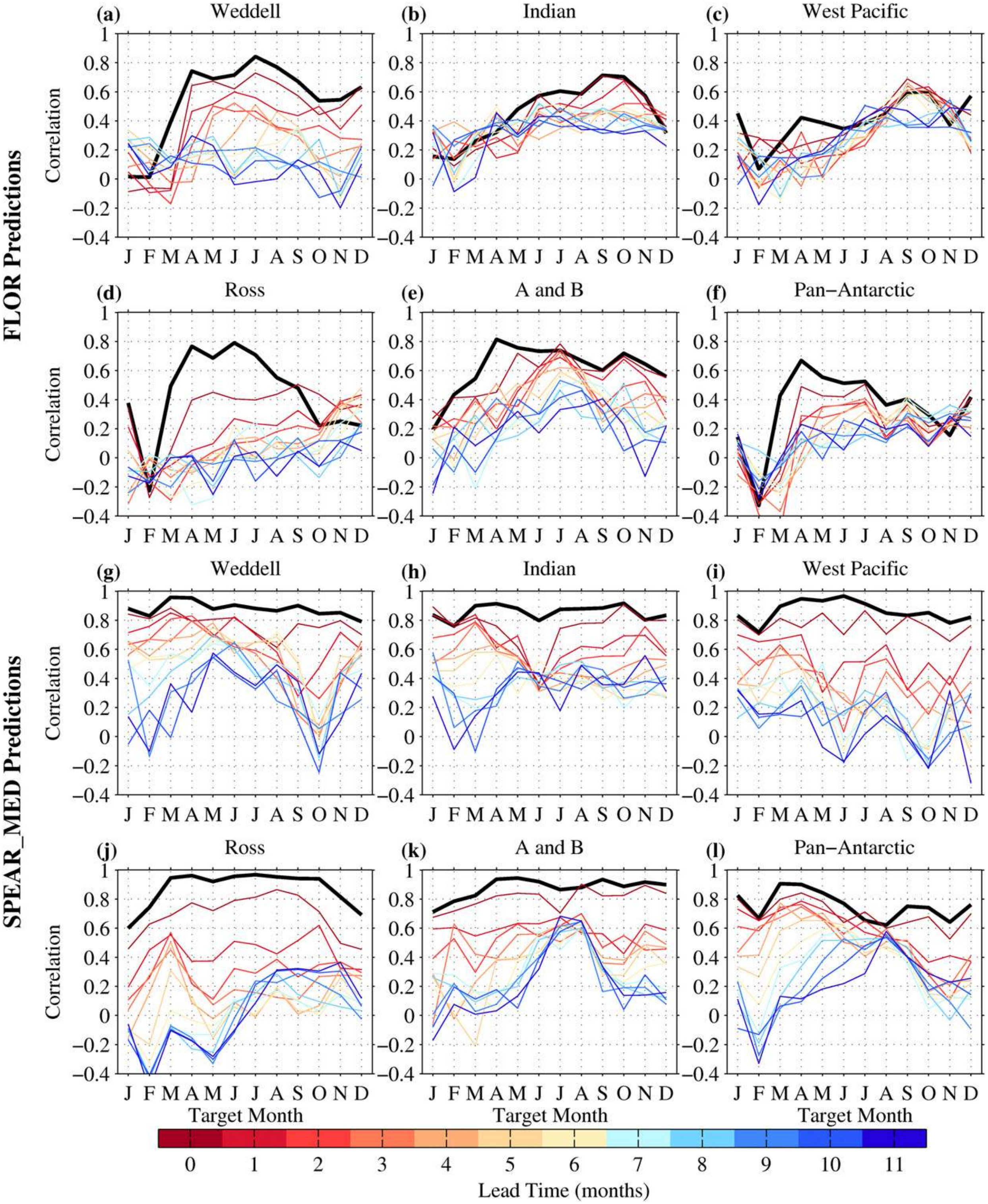

FIG. 9. Relation between regional SIE ICs and prediction skill. Black curves show the monthly detrended correlation for regional SIE between the ICs and NSIDC observations. Colored curves show the detrended ACC prediction skill at various lead times for (a)–(f) FLOR and (g)–(l) SPEAR_MED.

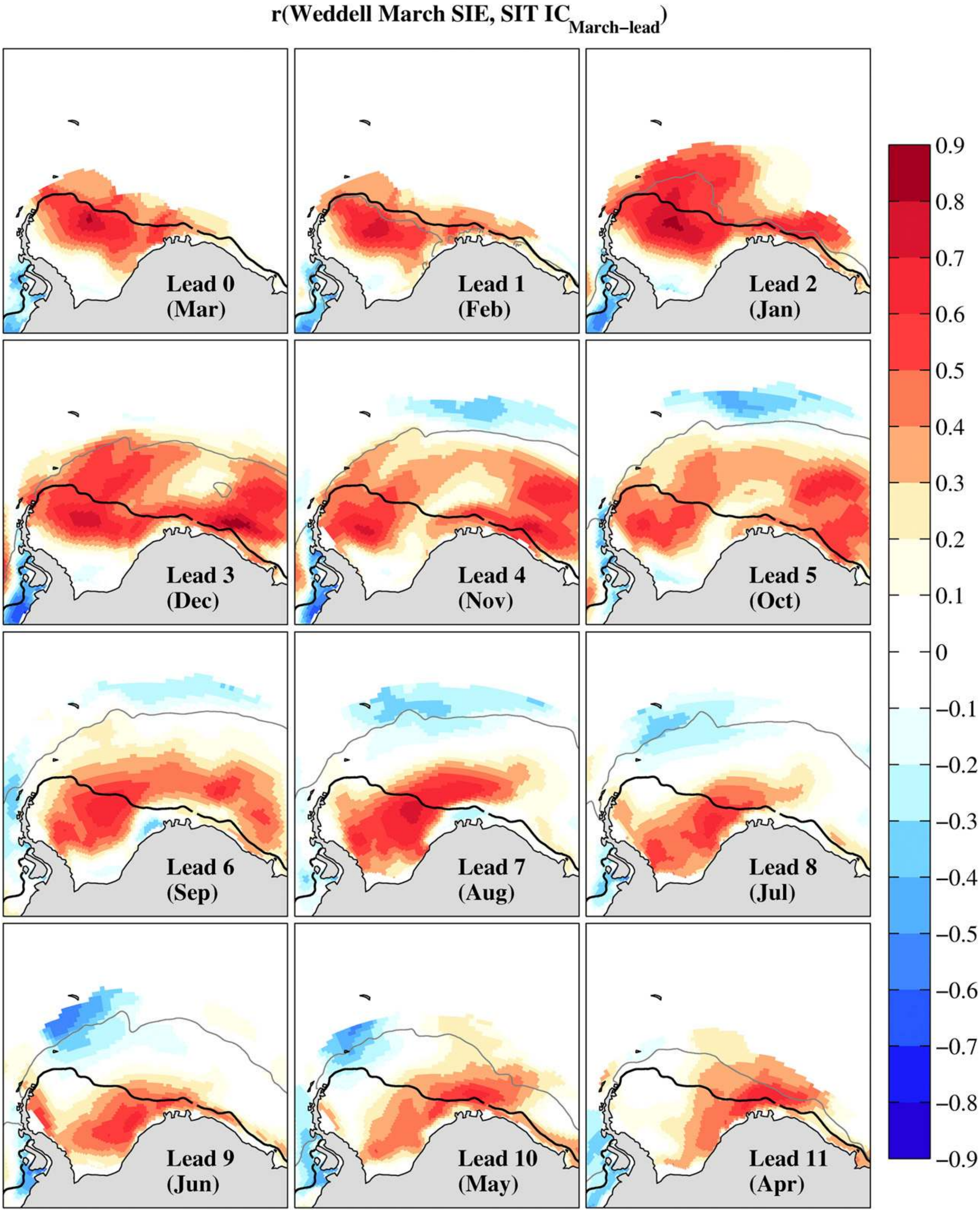

FIG. 10. Detrended correlations between observed March Weddell SIE and SPEAR_MED SIT ICs in earlier months. The black contour shows the observed climatological position of the March sea ice edge. The gray contours show the observed climatological sea ice edge position at different initialization months, which are indicated in parentheses.

suggest that SIT is providing a source of summer SIE prediction skill in the SPEAR_MED system. SPEAR_LO has notably lower correlation values, particularly for lead times greater than four months, consistent with the lower Weddell summer prediction skill in this system (see Fig. S5 in the supplemental material).

The correlations shown in Fig. 10 suggest a clear role for sea ice dynamics in Weddell SIE–SIT coupling. At lead times of 0–5 months, the highest correlations occur near or northward of the March sea ice edge location, suggesting that local persistence of SIT anomalies is providing predictive skill. There may also be a role for spring SIT anomalies north of the summer ice edge in modulating spring sea ice loss and the eventual summer minimum via the ice-albedo feedback. At longer lead times of 6–8 months, the highest correlations occur south of the March sea ice edge. The SIT anomalies at these locations need to advect northward in the Weddell gyre in order to influence summer SIE (see observed drift field in Fig. 8). Indeed, the advective travel time for an ice parcel from the southern Weddell Sea to the summer ice edge position is broadly consistent with these 6–8-month lead times. We also note a dipole correlation pattern present for winter and spring ice thickness (lead times of 4–10 months). This north–south dipole pattern is consistent with Weddell SIT patterns being controlled by wind-driven ice export anomalies. For example, in years with anomalous northward winds, additional ice is exported from the Weddell Sea, producing negative SIT anomalies in the southern Weddell Sea and positive anomalies in the northern Weddell Sea.

How much of the summer Weddell SIE skill can be attributed to SIT initialization? To address this question, we construct three simple linear-regression models to compare to the skill of the dynamical prediction systems. The first uses Weddell SIE as a predictor (see section 4a), the second uses Weddell sea ice volume (SIV) south of 60°S as a predictor, and the third uses advected upper-ocean heat content (OHC) as a predictor (see section 4c). These predictors are computed using the ICs of each system and used to predict the observed Weddell SIE. Figure 11 compares the skill of these statistical predictions to the skill of each dynamical prediction system. We find that SIE ICs are the dominant source of summer prediction skill at short lead times (0–2 months), whereas SIT ICs become the dominant source of skill at longer lead times (3–11 months). Interestingly, the high skill of the SPEAR_MED system in the Weddell Sea can be fully captured by the combination of the SIE and SIV regression models. This statistical reconstruction shows that SIT ICs are the crucial source of long-lead summer Weddell Sea prediction skill in this system. We find a similar breakdown between short and long lead summer prediction skill sources in other Antarctic regions (not shown). The advected OHC predictor, described in the following subsection, is less skillful for summer SIE than the sea ice based predictors. Figure 11 also shows that the lower summer skill in SPEAR_LO can be attributed to the lower-quality SIT ICs in this system. We explore this issue further in section 5a. Finally, it is clear that FLOR's summer SIE skill is severely limited by its poor SIE and SIT initialization.

We also investigate the possible role of snow on sea ice as a source of summer sea ice predictability. We find that a predictor based on the combined mass of sea ice and snow south of 60°S performs very similarly to the SIV-based predictor (not shown). This similar skill is due to the fact that the combined mass is dominated by sea ice (snow mass typically accounts for 0%–8% of the combined mass in summer and 10%–12% in winter) and the fact that snow and sea ice thickness covary strongly in these systems.

c. *Advected upper-ocean heat content*

Perfect model predictability studies have shown that upper OHC provides a key source of predictability for the winter sea ice edge (Holland et al. 2013; Marchi et al. 2019). Here, we ask: Are the FLOR and SPEAR initialized predictions capturing this potential source of predictability? Figure 12 shows regional-mean upper-ocean temperature anomalies in GOHSC observations and the FLOR and SPEAR data assimilation runs that are used for ocean ICs. Note that there is substantial overlap between the observations used for data assimilation and the GOHSC dataset. Also note that the same ocean ICs are used for both SPEAR_LO and SPEAR_MED. We find that both assimilation systems capture upper-ocean temperature variability with some skill, with regional detrended correlation values ranging from 0.31 to 0.79 in FLOR and from 0.45 to 0.85 in SPEAR.

Next, we construct a statistical prediction model that uses these upper-ocean temperature anomalies to predict the position of the sea ice edge. Consistent with earlier work on the Antarctic circumpolar wave (ACW; Gloersen and White 2001), we find that a prediction based on zonally advected temperature anomalies substantially outperforms a prediction based on local temperature anomalies. Specifically, we make a lead τ statistical prediction of the sea ice edge position $e(x, t)$ as

$$e(x,t) = aT'(x - c\tau, t - \tau) + b\,, \tag{9}$$

where t is the forecast target time, τ is the forecast lead time, x is the longitude, T' is the upper-ocean temperature anomaly, c is the upper-ocean eastward advection speed, and a and b are coefficients obtained via linear regression. These predictions are constructed using upper-ocean temperature anomalies from the SPEAR and FLOR ocean ICs. We find that optimal prediction skill is obtained using temperature anomalies in the upper 50 m south of 60°S and using an advection speed of $c = 360° \,(7\ \text{yr})^{-1}$. This advection speed is the same as used earlier in Eq. (8), and the geographic domain corresponds to the dominant region of sea ice variability.

We plot the skill of these advected persistence forecasts in Fig. 13, finding that the skill of the initialized dynamical predictions (Fig. 13a,b) can be reasonably well captured by the statistical predictions (Fig. 13c,d). This high reconstructed skill suggests that initialized upper-ocean temperatures, and the model's ability to advect these anomalies eastward via the simulated ocean circulation, are providing a key source of prediction skill in these systems. In some regions the statistical model notably outperforms the dynamical predictions.

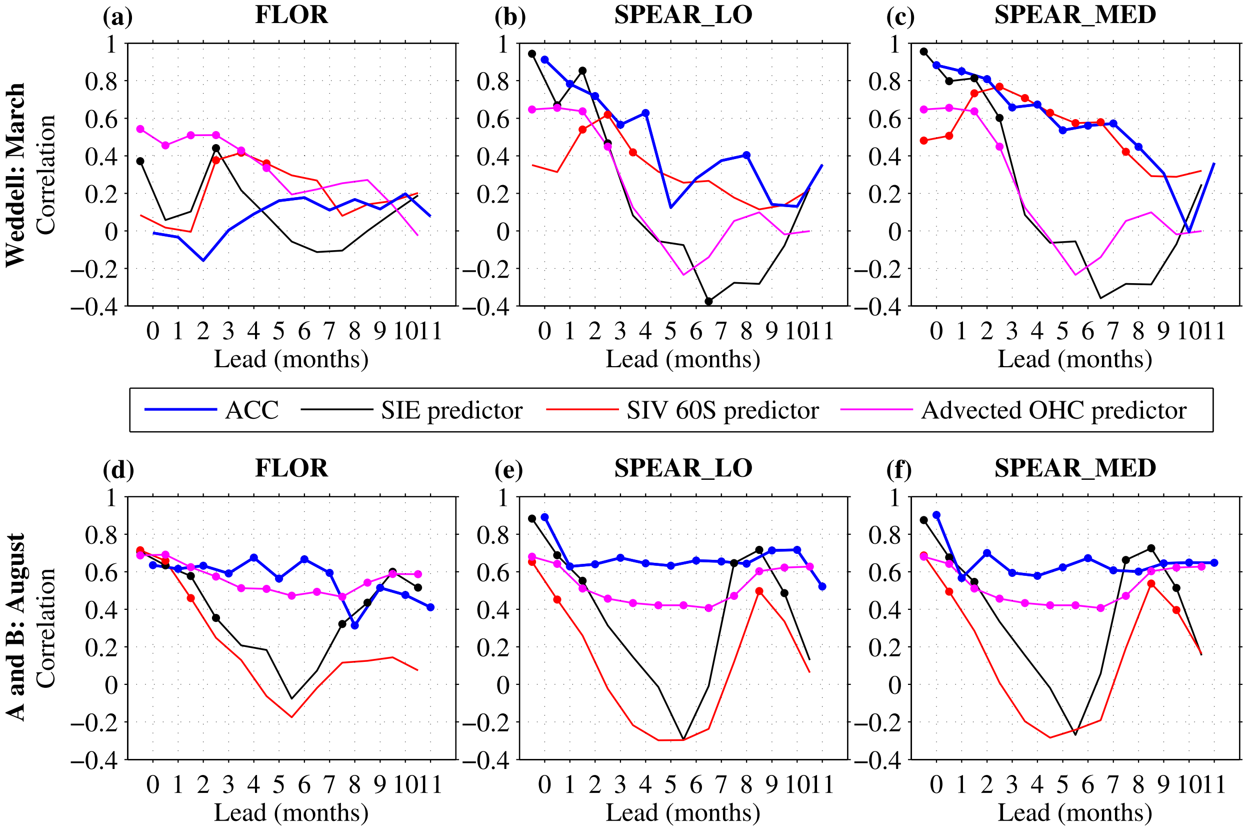

FIG. 11. Sources of prediction skill for (a)–(c) March Weddell SIE and (d)–(f) August Amundsen and Bellingshausen SIE. Blue lines show the detrended ACC skill in FLOR, SPEAR_LO, and SPEAR_MED. Black, red, and magenta lines show the detrended prediction skill of linear regression predictions based on regional SIE ICs, regional SIV ICs, and advected OHC ICs, respectively. The SIV and OHC predictions are based on the region south of 60°S. Correlations that are statistically significant at the 95% level based on a t test are indicated by dots. Note that the statistical predictions are shifted by 0.5-month lead time since these are computed using monthly mean quantities, whereas the dynamical predictions are initialized on the first of each month.

In particular, the Ross Sea stands out as a region where the advected persistence forecast displays prediction skill up to 7-month lead times, whereas the dynamical models have little skill beyond 2-month lead times. The success of the statistical model in the Ross Sea shows that these systems are not capitalizing upon a key source of potential predictability that is present in their ocean ICs, demonstrating clear room for improvement in this region. We further discuss this issue in section 5b.

Figures 11d–f show the skill of the three linear regression models described in section 4b for winter predictions of Amundsen and Bellingshausen SIE, a region of high skill in each of the systems. The advected upper OHC predictor is computed analogously to the advected persistence ice edge forecast, except based on regional-mean temperature anomalies. Similar to the summer SIE predictions, we find that SIE provides the key source of predictability at short lead times, however, unlike summer predictions, SIV does not provide a crucial contribution to the winter prediction skill. Advected OHC provides a key source of predictability at longer lead times, explaining most of the skill in FLOR and some, but not all, of the skill in SPEAR. We also find that SIE persistence provides skill at longer lead times via a winter-to-winter reemergence of SIE anomalies in the Amundsen and Bellingshausen sector.

5. Sources of sea ice prediction errors

In this section, we consider sources of prediction errors, focusing on the roles of 1) the ocean convective state and 2) sea ice drift.

a. Importance of the ocean convective state

We first return to the summer skill differences between SPEAR_MED and SPEAR_LO in the Weddell Sea (Fig. 5, section 4b). The higher prediction skill in SPEAR_MED is attributable to SIT initialization (see Fig. 11), which leads to the follow-up question: Why does the higher resolution atmosphere of SPEAR_MED produce improved SIT ICs? We find that the SPEAR_MED nudged run has a thicker sea ice mean state (Fig. 14a) and longer-lived SIT anomalies relative to SPEAR_LO (Fig. 14b). Therefore, the improved prediction skill in SPEAR_MED appears to derive from this system's thicker, more persistent, sea ice ICs. Figure 14c shows the components of the Weddell sea ice

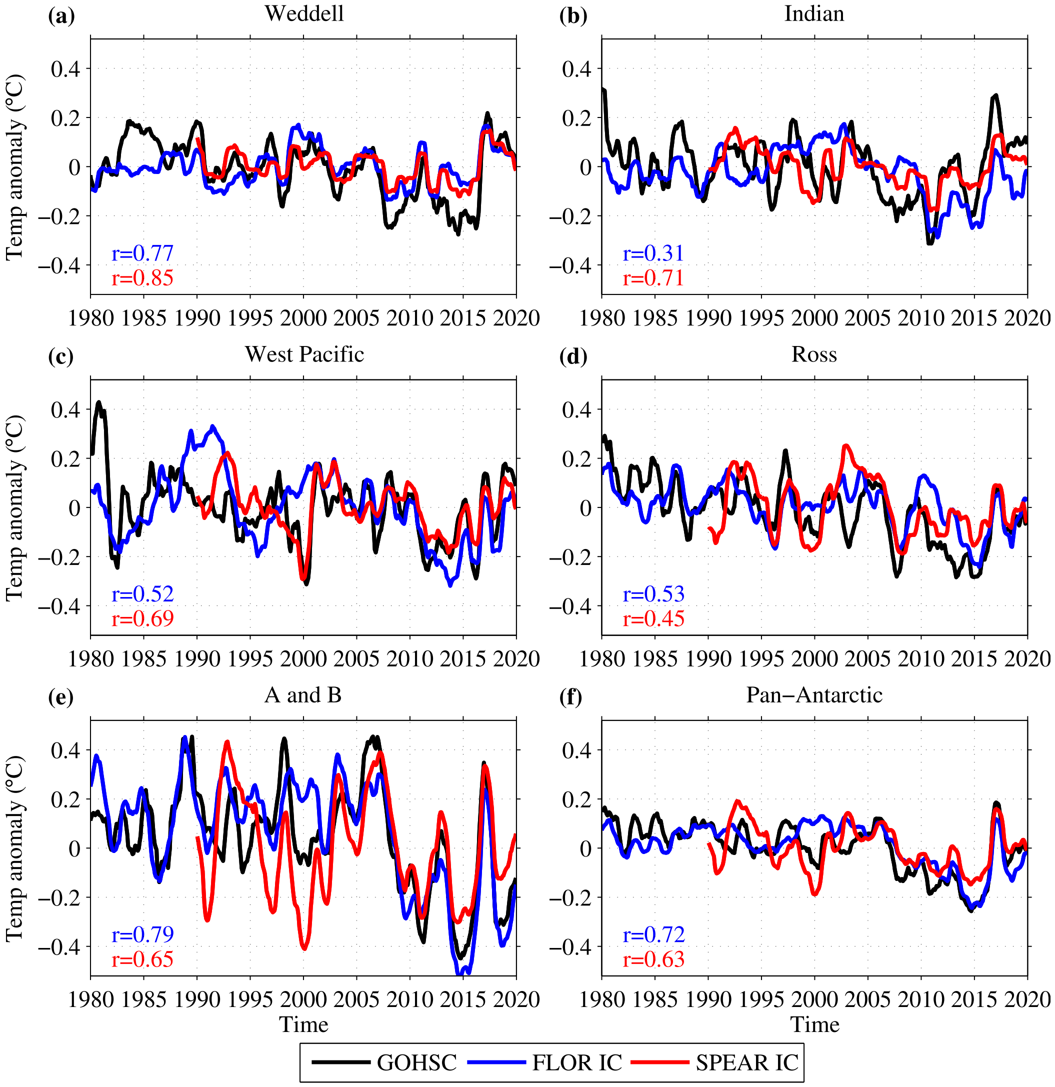

FIG. 12. Regional mean upper ocean temperature anomalies (0–50 m; south of 60°S) in GOHSC observations (black), FLOR ocean ICs (blue), and SPEAR ocean ICs (red). Note that the same ocean ICs are used for both SPEAR_LO and SPEAR_MED. Data are plotted as 1-yr running means of the monthly data. Colored text indicates detrended correlation values between the ocean ICs and GOHSC.

mass budget in the SPEAR_LO (dashed lines) and SPEAR_MED (solid lines) nudged runs. We find that the thinner ice in SPEAR_LO results from decreased winter sea ice growth and increased basal melt in winter. SPEAR_MED has more mass loss via ice export. This difference in ice export is primarily due to the models' mean differences in thickness, as their simulated ice velocities are very similar (Fig. S6).

Previous work has shown that higher-resolution atmospheric forcing of ice–ocean models can increase simulated Antarctic sea ice mass via enhanced ice production in coastal polynyas (Mathiot et al. 2010; Stössel et al. 2011; Barthélemy et al. 2012; Zhang et al. 2015). This increase has been attributed to an improved representation of near-coastline and katabatic winds. We find that coastal polynyas only account for a small fraction (7%) of the ice growth differences between SPEAR_LO and SPEAR_MED, with the dominant contribution coming from open-ocean differences in ice growth (93%; not shown). The fact that SPEAR_LO has both reduced growth and increased basal melt relative to SPEAR_MED is counterintuitive, since these terms typically have a positive covariance that results in a negative (stabilizing) feedback (Martinson 1990; Wilson et al. 2019). In particular, enhanced ice growth leads to enhanced brine rejection, which enhances vertical mixing and entrainment of warm waters at the base of the mixed layer, which enhances basal melt. Conversely, enhanced bottom melt reduces ice thickness, which increases the conductive heat flux and enhances ice growth. The negative covariance between growth and basal melt found here suggests that these processes are being driven by a common oceanic or atmospheric forcing

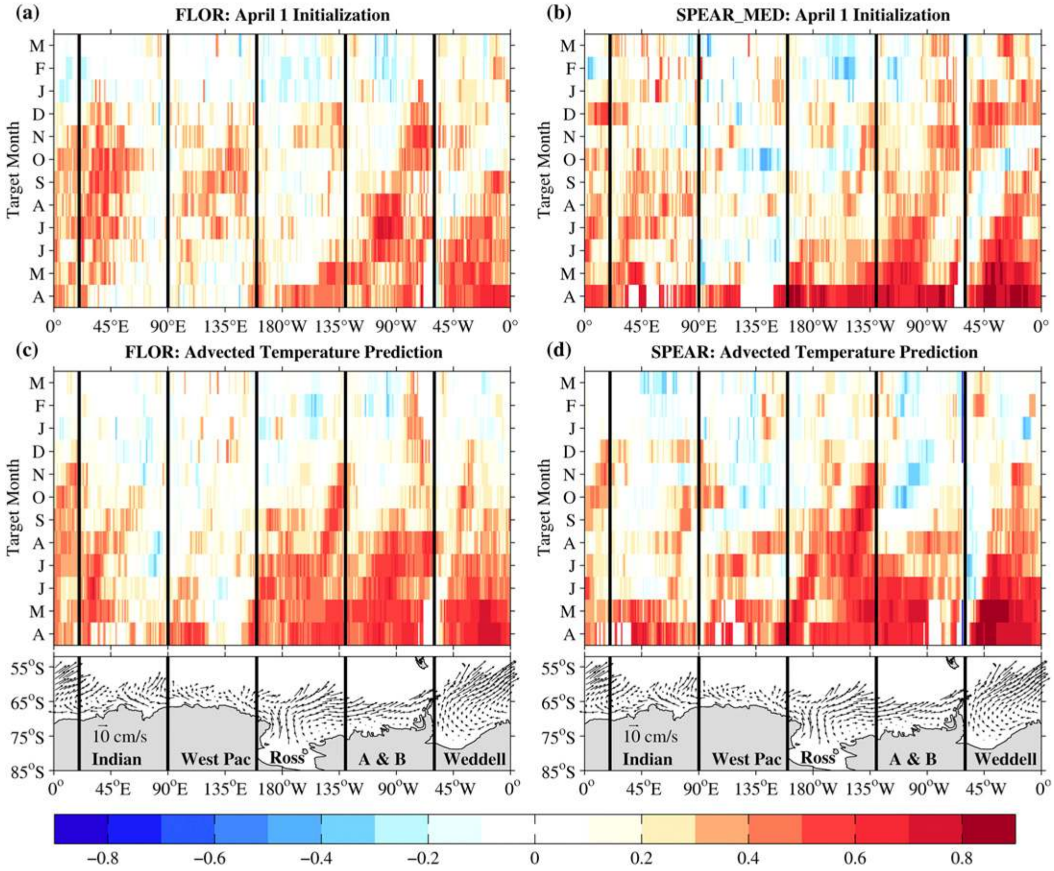

FIG. 13. Prediction skill (detrended ACC) of the sea ice edge position for 1 April initialized forecasts from (a) FLOR and (b) SPEAR_MED and for statistical forecasts based on (c) FLOR ocean ICs and (d) SPEAR ocean ICs. The statistical forecasts use advected persistence of ocean temperature anomalies in the upper 50 m and an eastward advection speed of $c = 360° (7\ \mathrm{yr})^{-1}$ [see Eq. (9)].

which overcomes the negative feedback expected from internal ice–ocean dynamics.

Examination of ocean properties suggests that the upper ocean is providing such a forcing. We find that SPEAR_LO exhibits unrealistic deep wintertime mixing in the Weddell Sea (Fig. 14d), which is substantially deeper than its SPEAR_MED counterpart. These deep mixed layers in SPEAR_LO are accompanied by consistent spatial patterns of thinner sea ice, less areal ice coverage, increased upward heat fluxes at the ocean surface, and increased basal melt (Fig. 15). The mixed layer depths in SPEAR_LO are much deeper than observed, suggesting that the enhanced basal melt and corresponding sea ice anomalies in this run are spurious. We also find warm surface air temperature anomalies that are spatially coincident with the region of deep convection (not shown), suggesting that the lower growth rates in SPEAR_LO are also being driven by this spurious deep mixing. The deep convection mechanism provides a likely explanation for the negative covariance between growth and basal melt in the SPEAR_LO run. This analysis suggests that SPEAR_LO has reduced summer skill due to spurious Weddell Sea deep convection in the nudged run used to produce its sea ice ICs. Future work is required to better understand the connection between this result and atmospheric model resolution.

b. *Sea ice drift*

Observational studies have shown that local Antarctic ice extent correlates with sea ice drift (Holland and Kwok 2012; Haumann et al. 2014). Thus, sea ice drift may influence both the predictability and prediction skill of Antarctic SIE. Figure 8 showed that September ice edge predictions appear to have low prediction skill in regions of strong northward drift. We examine this hypothesis quantitatively in Fig. 16, plotting prediction skill of the September ice edge (skill values averaged over lead times of 0–5 months; colored lines) and the mean northward winter drift speed in OSISAF observations (black). We find that the western Ross and western Weddell Seas have low prediction skill and are characterized by strong northward drift of sea ice. Conversely, regions with weaker northward drift and stronger eastward flow, such as the Amundsen/Bellingshausen, eastern Weddell, and western Indian sectors, tend to have higher prediction skill. We find that ice edge prediction skill is negatively correlated with meridional drift, however the correlation values are fairly modest (−0.38, −0.29, and −0.28 in FLOR,

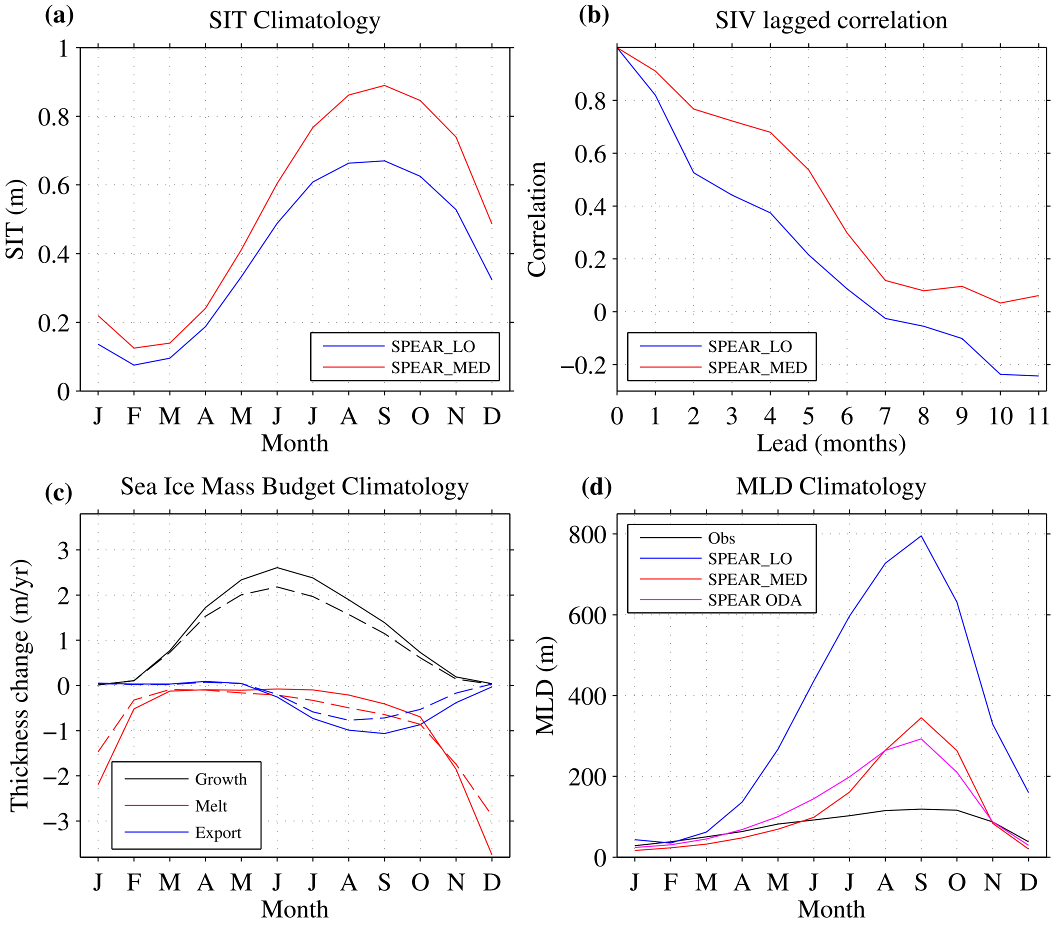

FIG. 14. (a) Weddell SIT climatology in SPEAR_LO (blue) and SPEAR_MED (red) nudged runs. (b) Lagged correlation between Weddell January SIV and earlier SIV in SPEAR nudged runs. (c) Weddell sea ice and snow mass budget climatology in SPEAR_LO (dashed lines) and SPEAR_MED (solid lines). The mass budget consists of sea ice growth, melt, and export. (d) Weddell mixed layer depth climatology in SPEAR nudged runs, SPEAR ODA (magenta), and from Argo and ship-based observations (black). All quantities are computed over the Weddell Sea domain south of 60°S.

SPEAR_LO, and SPEAR_MED, respectively; −0.51, −0.36, and −0.37 when only considering the subset of drift speeds above the 50th percentile). We also find positive correlations with zonal drift speeds of 0.34, 0.49, and 0.60 in FLOR, SPEAR_LO, and SPEAR_MED, respectively. This positive correlation with zonal drift is consistent with the predictability derived from zonal advection of ocean temperature anomalies described in section 4c.

Why could prediction skill be lower in regions of northward sea ice drift? One possibility is that the sea ice edge is inherently less predictable in these regions, owing to a greater role for ice drift in setting the ice edge position. This stronger dependence on drift suggests that unpredictable synoptic events may exert more control on the ice edge in these regions, resulting in reduced predictability. However, the relatively high skill of the statistical advected OHC predictions in the Ross Sea (Fig. 13c,d) suggests that sea ice anomalies in this region are potentially predictable.

Another possible reason is that predictions may be more susceptible to model physics errors in regions of northward drift due to a greater role for ice dynamics. To investigate the latter hypothesis, we examine model biases in predictions of sea ice drift. In Fig. 17, we show predicted winter (July–September) drift patterns from forecasts initialized on 1 April. We find that the general observed ice circulation features, such as eastward drift near the ice edge, westward drift near the coastline, and northward export in the Ross and Weddell Seas, are reasonably well predicted by each of the models. The SPEAR models both have predicted drift speeds that are generally too high and, correspondingly, have too much ice export in the Ross Sea. This enhanced export may contribute to the low prediction skill found in this region. However, we also note that FLOR has a more realistic Ross Sea ice export but also has very poor prediction skill in this region. Future work is required to further explore the connections between ice drift and predictability.

6. Discussion and conclusions

This study has used retrospective seasonal forecasts from three dynamical prediction systems to examine the seasonal

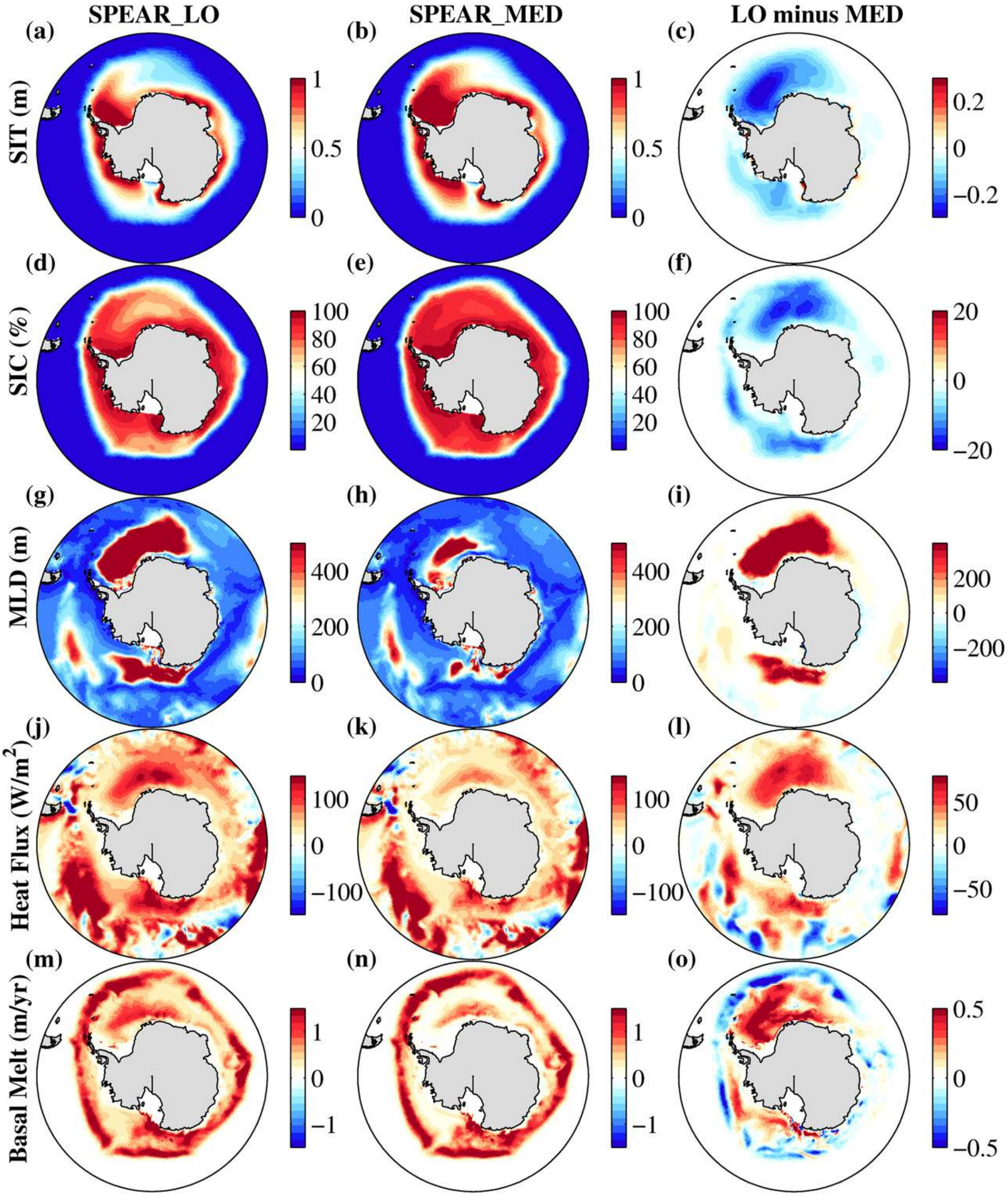

FIG. 15. Winter climatologies (JAS) in nudged runs of SPEAR_LO, SPEAR_MED, and their difference. Plotted are (a)–(c) SIT, (d)–(f) SIC, (g)–(i) mixed layer depth, (j)–(l) net upward heat flux at the ocean surface and ocean–ice interface, and (m)–(o) basal sea ice melt.

prediction skill and predictability of Antarctic sea ice. We have analyzed initialized ensemble forecasts spanning 1992–2018 based on the FLOR, SPEAR_LO, and SPEAR_MED dynamical models developed at the Geophysical Fluid Dynamics Laboratory. Our results demonstrate that regional Antarctic SIE is predictable on seasonal time scales. The prediction skill of the dynamical prediction systems generally exceeds the skill of an anomaly persistence forecast, indicating that the

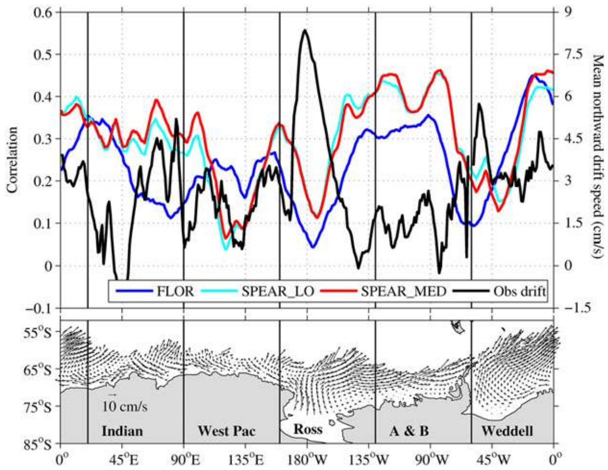

FIG. 16. September sea ice edge prediction skill in FLOR (blue), SPEAR_LO (cyan), and SPEAR_MED (red) and its relation to observed northward winter sea ice drift (black). The northward sea ice drift has been meridionally averaged over months JAS. Prediction skill shows the average detrended ACC over lead times of 0–5 months. Prediction skill values have been smoothed zonally using a 20° running mean. Climatological observed winter sea ice drift (JAS) is plotted as vectors.

dynamical models are providing additional sources of prediction skill beyond SIE persistence.

Long-lead ACC prediction skill at least 11 months in advance was found for autumn and winter SIE in the Weddell, Indian, west Pacific, and Amundsen and Bellingshausen sectors. Skill horizons based on MSSS skill were found to be generally shorter due to conditional forecast biases. The high autumn and winter skill is partially attributable to upper-ocean heat content (OHC) anomalies that are initialized via ocean data assimilation and subsequently advected eastward by the simulated upper-ocean circulation. Predictions of the sea ice edge position show clear zones of high prediction skill that propagate eastward, closely resembling the eastward propagating predictability features identified in earlier perfect model studies (Holland et al. 2013; Marchi et al. 2019). We found that a simple statistical prediction based on advected upper-ocean temperature anomalies reproduced most of the ice edge prediction skill captured by the dynamical prediction systems. The Ross Sea was found to have comparatively low skill in all three systems with little skill beyond 2-month lead times. The advected temperature statistical forecast clearly outperformed the dynamical models in this region with skill up to 7-month lead times, suggesting that these systems are failing to capitalize on a key source of potential prediction skill. We also found that sea ice drift properties may contribute to the lower prediction skill in the Ross Sea. Regions of strong northward sea ice drift, such as the western Ross and western Weddell Seas, were found to have lower prediction skill whereas regions of strong eastward drift, such as the Amundsen/Bellingshausen, eastern Weddell, and western Indian sectors, were found to have higher prediction skill. The skill degradation in northward drift regions could be due to either a greater sensitivity to model physics errors in these zones or reduced inherent predictability due to a greater role for unpredictable synoptic variability in driving sea ice variations, however, this second explanation appears unlikely given the success of the advected temperature prediction in the Ross Sea.

The SPEAR systems exhibit marked improvements over FLOR for summer SIE predictions. These summer skill improvements are primarily associated with improved SIE and SIT initial conditions. SPEAR_MED, which has higher summer skill than its lower-resolution counterpart SPEAR_LO, was found to skillfully predict Weddell summer SIE at 7–9-month lead times and pan-Antarctic summer SIE at 5–7-month lead times. We constructed three simple linear regression-based prediction models, based on regional SIE, SIV, and advected upper OHC, respectively, to examine the sources of SIE predictability in these systems. The statistical models showed that the dominant source of summer SIE prediction skill at short lead times of 0–2 months is SIE anomaly persistence whereas SIT anomaly persistence provides the dominant source of prediction skill at longer lead times of 3–11 months. Earlier perfect model studies have emphasized that SIT plays a limited role in multiannual predictability (Holland et al. 2013; Ordoñez et al. 2018; Marchi et al. 2019); however, our results show that SIT is a critical source of Weddell Sea predictability on seasonal time scales. The regression-based models also showed that SIE is the key source of predictability at short lead times for winter SIE predictions, whereas advected OHC provides a crucial source of predictability at longer lead times. We investigated the differences in summer Weddell SIE skill between SPEAR_LO and SPEAR_MED, finding that SPEAR_MED's higher prediction skill is associated with thicker and more persistent Weddell SIT. The thinner sea ice used to initialize the SPEAR_LO predictions was found to result from unrealistically deep wintertime mixing in the Weddell Sea and corresponding enhanced basal melt, enhanced upward surface heat flux, and reduced winter sea ice growth.

The regional Antarctic SIE skill scores reported in this study are generally higher than previously documented regional SIE skill in the Arctic (e.g., Sigmond et al. 2016; Bushuk et al. 2017; Dirkson et al. 2019; Kimmritz et al. 2019). The long-lead winter skill shown in the Weddell, Amundsen/Bellingshausen, Indian, and west Pacific sectors is matched in the Arctic only in the Barents and Labrador Seas. A unifying feature across these regions of long-lead skill is that the initialization of upper OHC anomalies provides the key source of winter prediction skill (Bushuk et al. 2019). The zonal propagation of predictable OHC anomalies is an essential aspect of Antarctic sea ice predictability, whereas Arctic winter skill has primarily been attributed to locally persistent OHC anomalies. Similar to earlier work on Arctic sea ice, we have found that SIT provides the key source of prediction skill for summer SIE predictions (Bonan et al. 2019). A key difference, however, is that summer Weddell SIE can be predicted 7–9 months in advance whereas regional Arctic summer SIE can only be skillfully predicted 1–4 months in advance due to an Arctic sea ice spring predictability barrier (Day et al. 2014; Bushuk et al. 2020). These longer Antarctic skill horizons suggest that an analogous predictability barrier is not present in the Antarctic and that long-lead

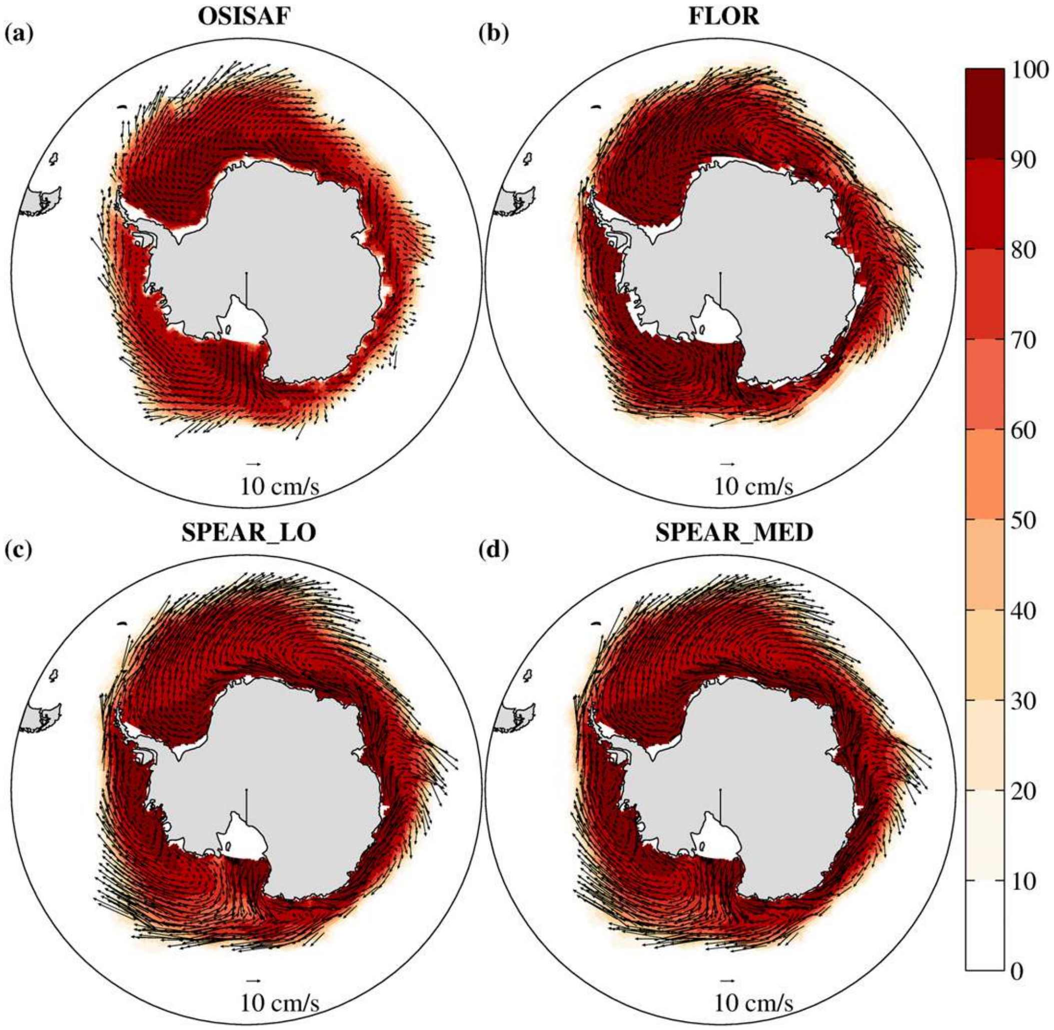

FIG. 17. Predicted climatological JAS sea ice drift patterns from 1 April initialized forecasts in (b) FLOR, (c) SPEAR_LO, and (d) SPEAR_MED compared with (a) observed JAS drift from OSISAF. Observed and predicted JAS SIC (%) is plotted in colors.

summer SIE predictions are possible, with important consequences for fisheries management, shipping, conservation, scientific expeditions, and tourism.

There are a number of future research directions that emerge from this study. First, comparisons to other dynamical prediction systems are needed to further assess Antarctic sea ice prediction skill and to evaluate the mechanisms of predictability identified in this work. Comparisons to systems based on higher-resolution ice–ocean models would be particularly valuable, as these models may have improved representation of Southern Ocean convective processes. Second, this study has highlighted long-lead summer predictability associated with SIT and short-lead predictability associated with SIE persistence. The SIT and SIC initial conditions used in the GFDL prediction systems are constrained by SSTs, atmospheric temperatures, and winds, but do not directly assimilate sea ice observations. Continuous satellite SIC observations are available dating back to 1979; however, satellite SIT products have large uncertainties and only exist for brief periods of time in the Antarctic. The recent *CryoSat-2* and *IceSat-2* satellite missions represent an opportunity to create these SIT data products, which could critically improve seasonal Antarctic sea ice predictions (Kacimi and Kwok 2020). Future work is needed to assess the potential benefits of sea ice data assimilation on Antarctic sea ice prediction skill. Third, the identified autumn and winter predictability associated with advected upper OHC was based on temperature anomalies in the upper 50 m. There is potentially additional predictability to be leveraged from subsurface heat content in deeper layers, which is now being sampled in greater detail via autonomous Argo floats, marine mammals, and ship-based CTD data. Finally, additional research is required on the inherent predictability of Antarctic sea ice. Potential topics to explore include the roles of sea ice drift, the ocean convective state, and forecast initialization month in determining sea ice predictability. Overall, the findings of this study suggest a promising potential for providing skillful operational Antarctic sea ice predictions on seasonal time scales.

Acknowledgments. We thank three anonymous reviewers for insightful comments that improved the manuscript. We also thank Olga Sergienko and Graeme MacGilchrist for comments on a preliminary version of this manuscript. This research from the Geophysical Fluid Dynamics Laboratory is supported by NOAA's Science Collaboration Program and administered by UCAR's Cooperative Programs for the Advancement of Earth System Science (CPAESS) under Awards NA16NWS4620043 and NA18NWS4620043B. F.A.H. was supported by the SNSF Grants P2EZP2_175162 and P400P2_186681, and NSF's Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM; PLR-1425989) Project.

Data availability statement. The NASA team sea-ice concentration observations used in this study are available from the National Snow and Ice Data Center website (http://nsidc.org/data/NSIDC-0051/versions/1). The OSISAF low-resolution sea ice drift product is available via the OSISAF web portal (http://www.osi-saf.org/?q=content/sea-ice-products). The World Ocean Atlas Global Ocean Heat and Salt Content dataset are available from the NOAA National Centers for Environmental Information website (https://www.nodc.noaa.gov/OC5/3M_HEAT_CONTENT). The FLOR predictions analyzed in this work are available through the North American Multi-Model Ensemble (NMME) Phase-II data (http://www.cpc.ncep.noaa.gov/products/NMME/data.html). The SPEAR prediction data are available via a public ftp server (ftp://nomads.gfdl.noaa.gov/users/Mitchell.Bushuk/Antarctic_Sea_Ice_Predictions/SPEAR).

REFERENCES

Adcroft, A., and Coauthors, 2019: The GFDL global ocean and sea ice model OM4. 0: Model description and simulation features. *J. Adv. Model. Earth Syst.*, **11**, 3167–3211, https://doi.org/10.1029/2019MS001726.

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903, https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

——, and Coauthors, 2004: The new GFDL global atmosphere and land model AM2–LM2: Evaluation with prescribed SST simulations. *J. Climate*, **17**, 4641–4673, https://doi.org/10.1175/JCLI-3223.1.

Argo, 2020: Argo float data and metadata from Global Data Assembly Centre (Argo GDAC). SEANOE, accessed 9 April 2020, https://doi.org/10.17882/42182.

Barthélemy, A., H. Goosse, P. Mathiot, and T. Fichefet, 2012: Inclusion of a katabatic wind correction in a coarse-resolution global coupled climate model. *Ocean Modell.*, **48**, 45–54, https://doi.org/10.1016/j.ocemod.2012.03.002.

Bitz, C. M., and W. H. Lipscomb, 1999: An energy-conserving thermodynamic model of sea ice. *J. Geophys. Res. Oceans*, **104**, 15669–15677, https://doi.org/10.1029/1999JC900100.

——, M. Holland, A. Weaver, and M. Eby, 2001: Simulating the ice-thickness distribution in a coupled climate model. *J. Geophys. Res. Oceans*, **106**, 2441–2463, https://doi.org/10.1029/1999JC000113.

Blanchard-Wrigglesworth, E., and C. M. Bitz, 2014: Characteristics of Arctic sea-ice thickness variability in GCMs. *J. Climate*, **27**, 8244–8258, https://doi.org/10.1175/JCLI-D-14-00345.1.

Bonan, D. B., M. Bushuk, and M. Winton, 2019: A spring barrier for regional predictions of summer Arctic sea ice. *Geophys. Res. Lett.*, **46**, 5937–5947, https://doi.org/10.1029/2019GL082947.

Bouillon, S., M. A. M. Maqueda, V. Legat, and T. Fichefet, 2009: An elastic–viscous–plastic sea ice model formulated on Arakawa B and C grids. *Ocean Modell.*, **27**, 174–184, https://doi.org/10.1016/j.ocemod.2009.01.004.

Boyer, T. P., and Coauthors, 2018: World Ocean Database 2018. NOAA Atlas NESDIS 87, https://www.ncei.noaa.gov/products/world-ocean-database.

Briegleb, B. P., and B. Light, 2007: A delta-Eddington multiple scattering parameterization for solar radiation in the sea ice component of the Community Climate System Model. NCAR Tech. Note NCAR/TN-472+STR, 108 pp.

Bushuk, M., R. Msadek, M. Winton, G. Vecchi, R. Gudgel, A. Rosati, and X. Yang, 2017: Skillful regional prediction of Arctic sea ice on seasonal timescales. *Geophys. Res. Lett.*, **44**, 4953–4964, https://doi.org/10.1002/2017GL073155.

——, X. Yang, M. Winton, R. Msadek, M. Harrison, A. Rosati, and R. Gudgel, 2019: The value of sustained ocean observations for sea-ice predictions in the Barents Sea. *J. Climate*, **32**, 7017–7035, https://doi.org/10.1175/JCLI-D-19-0179.1.

——, M. Winton, D. B. Bonan, E. Blanchard-Wrigglesworth, and T. Delworth, 2020: A mechanism for the Arctic sea ice spring predictability barrier. *Geophys. Res. Lett.*, **47**, e2020GL088335. https://doi.org/10.1029/2020GL088335.

Cavalieri, D. J., C. L. Parkinson, P. Gloersen, and H. J. Zwally, 1996: Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data, version 1. NASA National Snow and Ice Data Center Distributed Active Archive Center, accessed 14 April 2020, https://doi.org/10.5067/8GQ8LZQVL0VL.

Chen, D., and X. Yuan, 2004: A Markov model for seasonal forecast of Antarctic sea ice. *J. Climate*, **17**, 3156–3168, https://doi.org/10.1175/1520-0442(2004)017<3156:AMMFSF>2.0.CO;2.

Chevallier, M., and D. Salas y Mélia, 2012: The role of sea ice thickness distribution in the Arctic sea ice potential predictability: A diagnostic approach with a coupled GCM. *J. Climate*, **25**, 3025–3038, https://doi.org/10.1175/JCLI-D-11-00209.1.

Day, J., S. Tietsche, and E. Hawkins, 2014: Pan-Arctic and regional sea ice predictability: Initialization month dependence. *J. Climate*, **27**, 4371–4390, https://doi.org/10.1175/JCLI-D-13-00614.1.

de Boyer Montégut, C., G. Madec, A. S. Fischer, A. Lazar, and D. Iudicone, 2004: Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *J. Geophys. Res. Oceans*, **109**, C12003, https://doi.org/10.1029/2004JC002378.

Delworth, T. L., and Coauthors, 2006: GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *J. Climate*, **19**, 643–674, https://doi.org/10.1175/JCLI3629.1.

——, and Coauthors, 2012: Simulated climate and climate change in the GFDL CM2.5 high-resolution coupled climate model. *J. Climate*, **25**, 2755–2781, https://doi.org/10.1175/JCLI-D-11-00316.1.

——, and Coauthors, 2020: SPEAR: The next generation GFDL modeling system for seasonal to multidecadal prediction and projection. *J. Adv. Model. Earth Syst.*, **12**, e2019MS001895, https://doi.org/10.1029/2019MS001895.

Dirkson, A., B. Denis, and W. Merryfield, 2019: A multimodel approach for improving seasonal probabilistic forecasts of regional Arctic sea ice. *Geophys. Res. Lett.*, **46**, 10 844–10 853, https://doi.org/10.1029/2019GL083831.

Doddridge, E. W., and J. Marshall, 2017: Modulation of the seasonal cycle of Antarctic sea ice extent related to the Southern Annular Mode. *Geophys. Res. Lett.*, **44**, 9761–9768, https://doi.org/10.1002/2017GL074319.

Dunne, J. P., and Coauthors, 2020: The GFDL Earth System Model version 4.1 (GFDL-ESM 4.1): Overall coupled model description and simulation characteristics. *J. Adv. Model. Earth Syst.*, e2019MS002015, https://doi.org/10.1029/2019MS002015.

Efron, B., 1982: *The Jackknife, the Bootstrap, and Other Resampling Plans*. SIAM, 85 pp.

Gloersen, P., and W. B. White, 2001: Reestablishing the circumpolar wave in sea ice around Antarctica from one winter to the next. *J. Geophys. Res. Oceans*, **106**, 4391–4395, https://doi.org/10.1029/2000JC000230.

Gnanadesikan, A., and Coauthors, 2006: GFDL's CM2 global coupled climate models. Part II: The baseline ocean simulation. *J. Climate*, **19**, 675–697, https://doi.org/10.1175/JCLI3630.1.

Guemas, V., M. Chevallier, M. Déqué, O. Bellprat, and F. Doblas-Reyes, 2016a: Impact of sea ice initialisation on sea ice and atmosphere prediction skill on seasonal timescales. *Geophys. Res. Lett.*, **43**, 3889–3896, https://doi.org/10.1002/2015GL066626.

——, and Coauthors, 2016b: A review on Arctic sea ice predictability and prediction on seasonal-to-decadal timescales. *Quart. J. Roy. Meteor. Soc.*, **142**, 546–561, https://doi.org/10.1002/qj.2401.

Haumann, F. A., D. Notz, and H. Schmidt, 2014: Anthropogenic influence on recent circulation-driven Antarctic sea ice changes. *Geophys. Res. Lett.*, **41**, 8429–8437, https://doi.org/10.1002/2014GL061659.

Held, I., and Coauthors, 2019: Structure and performance of GFDL's CM4. 0 climate model. *J. Adv. Model. Earth Syst.*, **11**, 3691–3727, https://doi.org/10.1029/2019MS001829.

Holland, M. M., E. Blanchard-Wrigglesworth, J. Kay, and S. Vavrus, 2013: Initial-value predictability of Antarctic sea ice in the Community Climate System Model 3. *Geophys. Res. Lett.*, **40**, 2121–2124, https://doi.org/10.1002/grl.50410.

——, L. Landrum, M. Raphael, and S. Stammerjohn, 2017: Springtime winds drive Ross Sea ice variability and change in the following autumn. *Nat. Comms.*, **8**, 731, https://doi.org/10.1038/s41467-017-00820-0.

——, ——, M. N. Raphael, and R. Kwok, 2018: The regional, seasonal, and lagged influence of the Amundsen Sea low on Antarctic sea ice. *Geophys. Res. Lett.*, **45**, 11–227https://doi.org/10.1029/2018GL080140.

Holland, P. R., 2014: The seasonality of Antarctic sea ice trends. *Geophys. Res. Lett.*, **41**, 4230–4237, https://doi.org/10.1002/2014GL060172.

——, and R. Kwok, 2012: Wind-driven trends in Antarctic sea-ice drift. *Nat. Geosci.*, **5**, 872–875, https://doi.org/10.1038/ngeo1627.

Hunke, E. C., and J. K. Dukowicz, 1997: An elastic-viscous-plastic model for sea ice dynamics. *J. Phys. Oceanogr.*, **27**, 1849–1867, https://doi.org/10.1175/1520-0485(1997)027<1849:AEVPMF>2.0.CO;2.

——, W. H. Lipscomb, A. K. Turner, N. Jeffery, and S. Elliot, 2015: CICE: the Los Alamos Sea Ice Model Documentation and Software User's Manual version 5.1. LA-CC-06-012, T-3 Fluid Dynamics Group, Los Alamos National Laboratory, 675 pp.

Juricke, S., H. F. Goessling, and T. Jung, 2014: Potential sea ice predictability and the role of stochastic sea ice strength perturbations. *Geophys. Res. Lett.*, **41**, 8396–8403, https://doi.org/10.1002/2014GL062081.

Kacimi, S., and R. Kwok, 2020: The Antarctic sea ice cover from ICESat-2 and CryoSat-2: Freeboard, snow depth and ice thickness. *Cryosphere*, **14**, 4453–4474, https://doi.org/10.5194/tc-14-4453-2020.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. Hnilo, M. Fiorino, and G. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631–1644, https://doi.org/10.1175/BAMS-83-11-1631.

Kimmritz, M., F. Counillon, L. H. Smedsrud, I. Bethke, N. Keenlyside, F. Ogawa, and Y. Wang, 2019: Impact of ocean and sea ice initialisation on seasonal prediction skill in the Arctic. *J. Adv. Model. Earth Syst.*, **11**, 4147–4166, https://doi.org/10.1029/2019MS001825.

Kirtman, B. P., and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc.*, **95**, 585–601, https://doi.org/10.1175/BAMS-D-12-00050.1.

Kwok, R., and J. Comiso, 2002: Southern Ocean climate and sea ice anomalies associated with the Southern Oscillation. *J. Climate*, **15**, 487–501, https://doi.org/10.1175/1520-0442(2002)015<0487:SOCASI>2.0.CO;2.

——, ——, T. Lee, and P. Holland, 2016: Linked trends in the South Pacific sea ice edge and Southern Oscillation index. *Geophys. Res. Lett.*, **43**, 10–295, https://doi.org/10.1002/2016GL070655.

Lavergne, T., S. Eastwood, Z. Teffah, H. Schyberg, and L.-A. Breivik, 2010: Sea ice motion from low-resolution satellite sensors: An alternative method and its validation in the Arctic. *J. Geophys. Res. Oceans*, **115**, C10032, https://doi.org/10.1029/2009JC005958.

Levitus, S., and Coauthors, 2012: World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. *Geophys. Res. Lett.*, **39**, 1955–2010, https://doi.org/10.1029/2012GL051106.

——, and Coauthors, 2013: The World Ocean Database. *Data Sci. J.*, **12**, WDS229–WDS234, https://doi.org/10.2481/dsj.WDS-041.

Lu, F., and Coauthors, 2020: GFDL's SPEAR seasonal prediction system: Ocean data assimilation (ODA), ocean tendency adjustment (OTA) and coupled initialization. *J. Adv. Model. Earth Syst.*, **12**, e2020MS002149, https://doi.org/10.1029/2020MS002149.

Maksym, T., 2019: Arctic and Antarctic sea ice change: Contrasts, commonalities, and causes. *Annu. Rev. Mar. Sci.*, **11**, 187–213, https://doi.org/10.1146/annurev-marine-010816-060610.

Manzanas, R., J. M. Gutiérrez, J. Bhend, S. Hemri, F. J. Doblas-Reyes, V. Torralba, E. Penabad, and A. Brookshaw, 2019: Bias adjustment and ensemble recalibration methods for seasonal forecasting: A comprehensive intercomparison using the C3S dataset. *Climate Dyn.*, **53**, 1287–1305, https://doi.org/10.1007/s00382-019-04640-4.

Marchi, S., T. Fichefet, H. Goosse, V. Zunz, S. Tietsche, J. J. Day, and E. Hawkins, 2019: Reemergence of Antarctic sea ice predictability and its link to deep ocean mixing in global climate models. *Climate Dyn.*, **52**, 2775–2797, https://doi.org/10.1007/s00382-018-4292-2.

Martinson, D. G., 1990: Evolution of the Southern Ocean winter mixed layer and sea ice: Open ocean deepwater formation and ventilation. *J. Geophys. Res. Oceans*, **95**, 11 641–11 654, https://doi.org/10.1029/JC095iC07p11641.

Massom, R. A., T. A. Scambos, L. G. Bennetts, P. Reid, V. A. Squire, and S. E. Stammerjohn, 2018: Antarctic ice shelf disintegration triggered by sea ice loss and ocean swell. *Nature*, **558**, 383–389, https://doi.org/10.1038/s41586-018-0212-1.

Massonnet, F., P. Reid, J. Lieser, C. Bitz, J. Fyfe, and W. Hobbs, 2020: Assessment of summer 2019–2020 sea-ice forecasts for the Southern Ocean. Tech. rep., Université Catholique de Louvain. Available at https://fmassonn.github.io/sipn-south.github.io/.

Mathiot, P., B. Barnier, H. Gallée, J. M. Molines, J. Le Sommer, M. Juza, and T. Penduff, 2010: Introducing katabatic winds in global ERA40 fields to simulate their impacts on the Southern Ocean and sea-ice. *Ocean Modell.*, **35**, 146–160, https://doi.org/10.1016/j.ocemod.2010.07.001.

Meinshausen, M., and Coauthors, 2011: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, **109**, 213–241, https://doi.org/10.1007/s10584-011-0156-z.

Meredith, M., and Coauthors, 2019: Polar regions. IPCC Special Report on the Ocean and Cryosphere in a Changing Climate, H.-O. Pörtner et al., Eds., in press.

Milly, P. C., and Coauthors, 2014: An enhanced model of land water and energy for global hydrologic and Earth-system studies. *J. Hydrometeor.*, **15**, 1739–1761, https://doi.org/10.1175/JHM-D-13-0162.1.

Morioka, Y., T. Doi, D. Iovino, S. Masina, and S. K. Behera, 2019: Role of sea-ice initialization in climate predictability over the Weddell Sea. *Sci. Rep.*, **9**, 1–11, https://doi.org/10.1038/s41598-019-39421-w.

Murphy, A. H., 1988: Skill scores based on the mean square error and their relationships to the correlation coefficient. *Mon. Wea. Rev.*, **116**, 2417–2424, https://doi.org/10.1175/1520-0493(1988)116<2417:SSBOTM>2.0.CO;2.

Notz, D., and Coauthors, 2020: Arctic sea ice in CMIP6. *Geophys. Res. Lett.*, **47**, e2019GL086749, https://doi.org/10.1029/2019GL086749.

Ordoñez, A. C., C. M. Bitz, and E. Blanchard-Wrigglesworth, 2018: Processes controlling Arctic and Antarctic sea ice predictability in the Community Earth System Model. *J. Climate*, **31**, 9771–9786, https://doi.org/10.1175/JCLI-D-18-0348.1.

Pearson, K., 1900: X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *London Edinburgh Dublin Philos. Mag. J. Sci.*, **50**, 157–175, https://doi.org/10.1080/14786440009463897.

Ponsoni, L., F. Massonnet, D. Docquier, G. Van Achter, and T. Fichefet, 2020: Statistical predictability of the Arctic sea ice volume anomaly: Identifying predictors and optimal sampling locations. *Cryosphere*, **14**, 2409–2428, https://doi.org/10.5194/tc-14-2409-2020.

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

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution-blended analyses for sea surface temperature. *J. Climate*, **20**, 5473–5496, https://doi.org/10.1175/2007JCLI1824.1.

Riahi, K., and Coauthors, 2017: The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environ. Change*, **42**, 153–168, https://doi.org/10.1016/j.gloenvcha.2016.05.009.

Roach, L. A., and Coauthors, 2020: Antarctic sea ice area in CMIP6. *Geophys. Res. Lett.*, **47**, e2019GL086729, https://doi.org/10.1029/2019GL086729.

Robel, A. A., 2017: Thinning sea ice weakens buttressing force of iceberg mélange and promotes calving. *Nat. Commun.*, **8**, 14596, https://doi.org/10.1038/ncomms14596.

Roemmich, D., S. Riser, R. Davis, and Y. Desaubies, 2004: Autonomous profiling floats: Workhorse for broad-scale ocean observations. *Mar. Technol. Soc. J.*, **38**, 21–29, https://doi.org/10.4031/002533204787522802.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1058, https://doi.org/10.1175/2010BAMS3001.1.

Schneider, D. P., and C. Deser, 2018: Tropically driven and externally forced patterns of Antarctic sea ice change: Reconciling observed and modeled trends. *Climate Dyn.*, **50**, 4599–4618, https://doi.org/10.1007/s00382-017-3893-5.

Shepherd, A., H. A. Fricker, and S. L. Farrell, 2018: Trends and connections across the Antarctic cryosphere. *Nature*, **558**, 223–232, https://doi.org/10.1038/s41586-018-0171-6.

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

——, M. Reader, G. Flato, W. Merryfield, and A. Tivy, 2016: Skillful seasonal forecasts of Arctic sea ice retreat and advance dates in a dynamical forecast system. *Geophys. Res. Lett.*, **43**, 43, https://doi.org/10.1002/2016GL071396.

Simmonds, I., and T. Jacka, 1995: Relationships between the interannual variability of Antarctic sea ice and the Southern Oscillation. *J. Climate*, **8**, 637–647, https://doi.org/10.1175/1520-0442(1995)008<0637:RBTIVO>2.0.CO;2.

Simpkins, G. R., L. M. Ciasto, D. W. Thompson, and M. H. England, 2012: Seasonal relationships between large-scale climate variability and Antarctic sea ice concentration. *J. Climate*, **25**, 5451–5469, https://doi.org/10.1175/JCLI-D-11-00367.1.

Stammerjohn, S., D. Martinson, R. Smith, X. Yuan, and D. Rind, 2008: Trends in Antarctic annual sea ice retreat and advance and their relation to El Niño–Southern Oscillation and Southern Annular Mode variability. *J. Geophys. Res. Oceans*, **113**, C03S90, https://doi.org/10.1029/2007JC004269.

Stössel, A., Z. Zhang, and T. Vihma, 2011: The effect of alternative real-time wind forcing on Southern Ocean sea ice simulations. *J. Geophys. Res. Oceans*, **116**, C11021, https://doi.org/10.1029/2011JC007328.

Sun, C., and Coauthors, 2010: The data management system for the global temperature and salinity profile programme. ESA Publication WPP-306, 8 pp., https://doi.org/10.5270/OceanObs09.cwp.86.

Thorndike, A. S., D. Rothrock, G. Maykut, and R. Colony, 1975: The thickness distribution of sea ice. *J. Geophys. Res.*, **80**, 4501–4513, https://doi.org/10.1029/JC080i033p04501.

Vecchi, G. A., and Coauthors, 2014: On the seasonal forecasting of regional tropical cyclone activity. *J. Climate*, **27**, 7994–8016, https://doi.org/10.1175/JCLI-D-14-00158.1.

Venegas, S. A., 2003: The Antarctic Circumpolar Wave: A combination of two signals? *J. Climate*, **16**, 2509–2525, https://doi.org/10.1175/1520-0442(2003)016<2509:TACWAC>2.0.CO;2.

Wang, W., M. Chen, and A. Kumar, 2013: Seasonal prediction of Arctic sea ice extent from a coupled dynamical forecast system. *Mon. Wea. Rev.*, **141**, 1375–1394, https://doi.org/10.1175/MWR-D-12-00057.1.

Wang, X., D. Giannakis, and J. Slawinska, 2019: The Antarctic circumpolar wave and its seasonality: Intrinsic travelling modes and El Niño–Southern Oscillation teleconnections. *Int. J. Climatol.*, **39**, 1026–1040, https://doi.org/10.1002/joc.5860.

White, W. B., and R. G. Peterson, 1996: An Antarctic circumpolar wave in surface pressure, wind, temperature and sea-ice extent. *Nature*, **380**, 699–702, https://doi.org/10.1038/380699a0.

Wilson, E. A., S. C. Riser, E. C. Campbell, and A. P. Wong, 2019: Winter upper-ocean stability and ice–ocean feedbacks in the sea ice–covered Southern Ocean. *J. Phys. Oceanogr.*, **49**, 1099–1117, https://doi.org/10.1175/JPO-D-18-0184.1.

Winton, M., 2000: A reformulated three-layer sea ice model. *J. Atmos. Oceanic Technol.*, **17**, 525–531, https://doi.org/10.1175/1520-0426(2000)017<0525:ARTLSI>2.0.CO;2.

Yuan, X., and D. G. Martinson, 2000: Antarctic sea ice extent variability and its global connectivity. *J. Climate*, **13**, 1697–1717, https://doi.org/10.1175/1520-0442(2000)013<1697:ASIEVA>2.0.CO;2.

——, and ——, 2001: The Antarctic dipole and its predictability. *Geophys. Res. Lett.*, **28**, 3609–3612, https://doi.org/10.1029/2001GL012969.

Zampieri, L., H. F. Goessling, and T. Jung, 2019: Predictability of Antarctic sea ice edge on subseasonal time scales. *Geophys. Res. Lett.*, **46**, 9719–9727, https://doi.org/10.1029/2019GL084096.

Zhang, S., M. Harrison, A. Rosati, and A. Wittenberg, 2007: System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. *Mon. Wea. Rev.*, **135**, 3541–3564, https://doi.org/10.1175/MWR3466.1.

Zhang, Z., T. Vihma, A. Stössel, and P. Uotila, 2015: The role of wind forcing from operational analyses for the model representation of Antarctic coastal sea ice. *Ocean Modell.*, **94**, 95–111, https://doi.org/10.1016/j.ocemod.2015.07.019.

Zhao, M., and Coauthors, 2018a: The GFDL global atmosphere and land model AM4.0/LM4.0: 1. Simulation characteristics with prescribed SSTs. *J. Adv. Model. Earth Syst.*, **10**, 691–734, https://doi.org/10.1002/2017MS001208.

——, and Coauthors, 2018b: The GFDL global atmosphere and land model AM4.0/LM4.0: 2. Model description, sensitivity studies, and tuning strategies. *J. Adv. Model. Earth Syst.*, **10**, 735–769, https://doi.org/10.1002/2017MS001209.

Zunz, V., H. Goosse, and S. Dubinkina, 2015: Impact of the initialisation on the predictability of the Southern Ocean sea ice at interannual to multi-decadal timescales. *Climate Dyn.*, **44**, 2267–2286, https://doi.org/10.1007/s00382-014-2344-9.