

Understanding the Forecast Skill of Rapid Arctic Sea Ice Loss on Subseasonal Time Scales

MARIE C. MCGRAW,^a EDUARDO BLANCHARD-WRIGGLESWORTH,^a ROBIN P. CLANCY,^a AND CECILIA M. BITZ^a

^a*Department of Atmospheric Science, University of Washington, Seattle, Washington*

(Manuscript received 14 April 2021, in final form 3 November 2021)

ABSTRACT: The predictability of sea ice during extreme sea ice loss events on subseasonal (daily to weekly) time scales is explored in dynamical forecast models. These extreme sea ice loss events (defined as the 5th percentile of the 5-day change in sea ice extent) exhibit substantial regional and seasonal variability; in the central Arctic Ocean basin, most subseasonal rapid ice loss occurs in the summer, but in the marginal seas rapid sea ice loss occurs year-round. Dynamical forecast models are largely able to capture the seasonality of these extreme sea ice loss events. In most regions in the summertime, sea ice forecast skill is lower on extreme sea ice loss days than on nonextreme days, despite evidence that links these extreme events to large-scale atmospheric patterns; in the wintertime, the difference between extreme and nonextreme days is less pronounced. In a damped anomaly forecast benchmark estimate, the forecast error remains high following extreme sea ice loss events and does not return to typical error levels for many weeks; this signal is less robust in the dynamical forecast models but still present. Overall, these results suggest that sea ice forecast skill is generally lower during and after extreme sea ice loss events and also that, while dynamical forecast models are capable of simulating extreme sea ice loss events with similar characteristics to what we observe, forecast skill from dynamical models is limited by biases in mean state and variability and errors in the initialization.

SIGNIFICANCE STATEMENT: We studied weather model forecasts of changes in Arctic sea ice extent on day-to-day time scales in different regions and seasons. We were especially interested in extreme sea ice loss days, or days in which sea ice melts very quickly or is reduced due to diverging forces such as winds, ocean currents, and waves. We find that forecast models generally capture the observed timing of extreme sea ice loss days. We also find that forecasts of sea ice extent are worse on extreme sea ice loss days compared to typical days, and that forecast errors remain elevated following extreme sea ice loss events.

KEYWORDS: Arctic; Ice loss/growth; Forecast verification/skill; Subseasonal variability

1. Introduction

Seasonal sea ice predictability has been a focal point of polar climate research in recent years, motivated by the steep decline in Arctic sea ice extent. Since 2008, the Sea Ice Outlook has evaluated predictions from different statistical and dynamical models, with a special focus on predicting Arctic-wide September sea ice extent (e.g., Stroeve et al. 2014; Blanchard-Wrigglesworth et al. 2015, 2017). Initial assessments indicated that dynamical model predictions did not outperform statistical model predictions, and that overall skill was comparable to a simple linear trend forecast (e.g., Stroeve et al. 2014). These dynamical models also struggled to predict each other's sea ice, highlighting the contributions of model uncertainty to forecast uncertainty (e.g., Blanchard-Wrigglesworth et al. 2015, 2017;

Harnos et al. 2019), although some of this model uncertainty can be reduced by using bias correction (e.g., Blanchard-Wrigglesworth et al. 2017; Dirkson et al. 2019; Batté et al. 2020).

Estimates of potential predictability skill with perfect model studies (i.e., where forecast skill is by construction free of error in the initial conditions or model physics) suggest that predictions of pan-Arctic sea ice extent or area could be skillful at lead times of over 12 months, depending on when the forecasts are initialized (e.g., Blanchard-Wrigglesworth et al. 2011b; Bushuk et al. 2017, 2019). However, comparisons of perfect model experiments and operational forecasts using the same modeling system reveal that there is still a large skill gap between theoretical and actual predictability (e.g., Bushuk et al. 2019). These gaps are even more pronounced for regional predictions of sea ice extent, and vary throughout the year (e.g., Bushuk et al. 2019). Less effort has been given to evaluate forecasts in detail for lead times in the first month. Yet, there is significant value to stakeholders in this time frame, especially when extreme sea ice change brings hazards (e.g., Eicken 2013; Zampieri et al. 2018).

The recent push toward improved Earth system predictions on subseasonal-to-seasonal (S2S) time scales has provided an opportunity to investigate sea ice forecasts (including retrospective forecasts) in the first ~6 weeks of lead times. The S2S

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/jcli-d-21-0301.s1>.

McGraw's current affiliation: Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado.

Corresponding author: Marie C. McGraw, mmcgraw@rams.colostate.edu

DOI: 10.1175/JCLI-D-21-0301.1

© 2022 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) (www.ametsoc.org/PUBSReuseLicenses).

archive (Vitart et al. 2017) permits detailed multimodel comparisons across a range of dynamical forecasts models. So far, Liu et al. (2018), Zampieri et al. (2018), and Wayand et al. (2019) have investigated S2S predictions of pan-Arctic sea ice concentration (Liu et al. 2018) and edge (Zampieri et al. 2018; Wayand et al. 2019) and have identified a range of skill in current forecast modeling systems: some models show skill over climatology at lead times of up to 6 weeks, while other models do not show skill at any lead times. These studies concluded that the wide range of skill indicates the importance of initial conditions and initialization strategies across forecast models; models that assimilated observed sea ice concentrations or thickness upon initialization showed increased forecast skill over models that did not assimilate any observations (e.g., Zampieri et al. 2018; Wayand et al. 2019). Even among models that employ data assimilation, large initial errors may still exist, as shown by the substantial spread in reanalysis estimates of sea ice as well (e.g., Chevallier et al. 2017; Meier and Stewart 2019). Forecast skill at subseasonal time scales also varies seasonally. Skill is generally highest in the summer months and often has a secondary peak in the winter, while it is lowest for forecasts initialized in the spring (e.g., Zampieri et al. 2018; Wayand et al. 2019; Bonan et al. 2019). These first looks at S2S models are a valuable starting point for evaluating the forecast potential of rapid sea ice loss events, which is the focus of our study.

Much of the observed short-term variability in pan-Arctic September sea ice extent (SIE) has been explained by variability in the atmosphere (e.g., Ding et al. 2017; Olonscheck et al. 2019). In particular, stronger anticyclonic circulation anomalies are associated with lower Arctic-wide September SIE (e.g., Ding et al. 2017; Wernli and Papritz 2018). However, when we focus on synoptic-scale events, the role of individual cyclones on sea ice loss is less clear. For example, some studies have linked the record-low September SIE in 2012 to a very large polar cyclone in August 2012 (e.g., Parkinson and Comiso 2013). Other studies have suggested, however, that the impacts of even this large cyclone were local and short-lived (e.g., Zhang et al. 2013; Stern et al. 2020). Statistical studies suggest that the response of sea ice to the August 2012 cyclone may be atypical. Schreiber and Serreze (2020) found that cyclones are often associated with increases in SIE. Within the marginal ice zone, sea ice loss on average occurs to the east of cyclones, while ice is typically gained to the west of cyclones (e.g., Clancy et al. 2022). Recent studies focused on extreme sea ice loss events on synoptic time scales (e.g., Wang et al. 2020) suggest that events are actually driven by a cyclone–anticyclone dipole and enhanced by ocean surface waves (e.g., Blanchard-Wrigglesworth et al. 2021), and a higher number of extreme sea ice loss events is correlated with a lower September sea ice minimum.

Given the diversity of results regarding the impacts of synoptic-scale events such as cyclones on changes in sea ice, the ability of dynamical sea ice forecasting models to faithfully reproduce subseasonal sea ice variability and extreme sea ice events on subseasonal time scales is worth exploring. Thus, we ask the following questions:

- How do extreme sea ice loss events on subseasonal time scales vary regionally, seasonally, and across different forecast models?
- Can dynamical forecast models make skillful predictions of sea ice extent during extreme sea ice loss events? How do these predictions vary regionally, and across different forecast models?
- Do forecast errors “jump” during an extreme sea ice loss event and remain high for the rest of the forecast period?

In section 2, we describe the observations and model reforecasts used in this study. Section 2 also provides our definition of extreme sea ice loss events, how we assess our forecast models, how we assess statistical significance, and the construction of our benchmark forecast. Section 3 discusses our results, including the biases in sea ice extent in the S2S models, seasonal and regional distribution of very rapid ice loss events (VRILE) days in the S2S models, the forecast skill of sea ice on VRILE days as compared to non-VRILE days, and forecast error growth following VRILEs. In section 4, we discuss our results and offer concluding remarks.

2. Data and methods

a. Observations

We use observations of sea ice from passive microwave satellite retrievals using the NASA Bootstrap algorithm for sea ice concentration (SIC; Comiso 2017; <https://nsidc.org/data/NSIDC-0079>). These observations are gridded on NSIDC's 25 km × 25 km polar stereographic grid (Knowles 1993); sea ice extent is calculated as the cumulative area of all grid cells in a region that have a SIC of at least 15%. As different algorithms can provide different estimates of SIE, we have further verified our analysis using the NASA Team sea ice concentration (Cavalieri et al. 1996; <https://nsidc.org/data/NSIDC-0051/versions/1>). We will focus on the results obtained using the NASA Bootstrap dataset, as the models we analyze tend to compare more faithfully to Bootstrap than NASA Team. Nonetheless, our results are similar using either observational dataset, and our conclusions are not dependent on the observational product we have selected.

We evaluate pan-Arctic sea ice extent as well as regional sea ice extent, defined in Fig. 1 and based on the regions originally defined by Parkinson et al. (1999). We omit parts of the Arctic in our evaluation that are disconnected from the main Arctic sea ice pack (e.g., the Sea of Okhotsk and the St. John region). We consolidate adjacent regions in which the sea ice extent exhibits similar seasonal variability; for example, we combine the East Siberian, Beaufort, and Chukchi Seas into one region, and the Kara and Laptev Seas into another region. We analyze the Barents Sea, the East Greenland Sea, Bering Sea, and the central Arctic as separate regions. We do not analyze Baffin Bay or the Canadian Islands separately due to distortion from land masks and regridding (see Fig. S7 in the online supplemental material for an example of how much land masks can vary in these regions), but they are included in the pan-Arctic sea ice extent estimate. We create

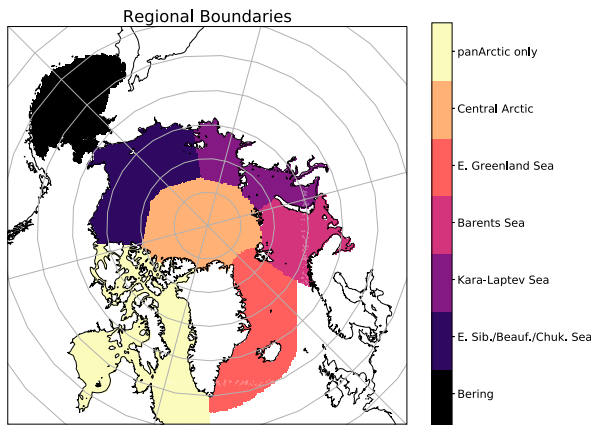


FIG. 1. Aggregated regions used for regional analysis of sea ice extent, based on Parkinson et al. (1999). The pan-Arctic sea ice extent is calculated by combining all regions shown above. The area in yellow indicates the sea ice regions that are included in the pan-Arctic sea ice extent analysis but are not analyzed separately.

an observed climatology based on daily sea ice extent from 1999 to 2014 for each region.

b. Model reforecasts

The model output analyzed here has been obtained from the S2S Prediction project database (Vitart et al. 2017; <https://confluence.ecmwf.int/display/S2S>). We use the reforecasts from four fully coupled numerical models that produce dynamical sea ice forecasts—the U.K. Met Office model (UKMO), the European Center for Medium-Range Weather Forecasting model (ECMWF), the National Centers for Environmental Prediction model (NCEP), and the Météo-France model (METEOFR). Following Wayand et al. (2019), the model reforecasts have been regridded using the nearest-neighbor method from the S2S archive resolution ($1.5^\circ \times 1.5^\circ$ for UKMO, ECMWF, and NCEP and $1^\circ \times 1^\circ$ for METEOFR) to the passive microwave satellite products grid (PM grid) at $25 \text{ km} \times 25 \text{ km}$ horizontal resolution. This step ensures that all models and observations are on a common grid. As we regrid the models to a finer grid, the nearest-neighbor regridding roughly preserves the land mask of the S2S archive grids. Finally, we mask out land using a common mask that classifies a grid cell as land if any of the models or the observations have land in that cell. This step helps to standardize the land masks across the different models, as they can vary quite substantially (see Fig. S7), and differences in land masks can account for some of the intermodel spread in sea ice forecast skill and in model error compared to observations.

The four models we have chosen each have different sea ice models as well as different initialization frequencies and ensemble sizes, detailed in Table 1. All four models' atmospheric components are initialized with 4DVAR assimilation schemes. While all four models have active sea ice components and initialize their sea ice concentrations with 3DVAR

assimilation schemes, none of the model reforecasts assimilate sea ice thickness. As the reforecast period of each model is slightly different, we analyze the 15-yr period spanning 1 January 1999–13 December 2014 that is common to all models. The exception to this is the NCEP model, which has a shorter reforecast period (1999–2011); we simply use the entire NCEP reforecast period. Our observational period is therefore limited to 1999–2014 to match the reforecasts as closely as possible.

Following established procedures for subseasonal forecasts (e.g., Vitart 2004; Zhu et al. 2014; Sun et al. 2018), a lead-time-dependent model climatology is created for each region in each S2S model. Model forecasts are averaged in weekly groups based on lead time (7 days) and day of year, meaning that for each region in each model, every lead week will have its own climatology. That is, for each region in each model,

$$\text{SIE}_{\text{anom}}(V, f) = \text{SIE}_{\text{forecast}}(V, f) - \text{SIE}_{\text{clim}}(\text{doy}_v, f), \quad (1)$$

where V is the forecast valid date (the date which is being forecast), f is the forecast lead time (days since initialization), and doy_v is the day of year of the forecast valid date.

c. Defining extreme sea ice loss events

In this study, we are interested in very rapid ice loss events (VRILEs). Following Wang et al. (2020) (who call these events LDSILs), we estimate large multiday losses in SIE by evaluating the p -day change in SIE, centered on day n . Here, we present our results for $p = 5$, but our analysis was overall similar for $p = 3$ and $p = 7$ (not shown). So, for $p = 5$, we calculate the 5-day change in sea ice extent for day n as follows:

$$d_{\text{SIE}}(n) = \text{SIE}(n + 2) - \text{SIE}(n - 2). \quad (2)$$

Here, VRILE days are defined as those days in which the 5-day change in sea ice extent is at or below its 5th percentile for the whole period of the record. That is, they are the days with the most extreme changes in sea ice loss on subseasonal time scales. Our results are similar when we set the VRILE threshold at the 10th percentile of the 5-day change in SIE.

d. Damped anomaly forecast

We assess the skill of the dynamical forecast models against a benchmark of a damped persisted anomaly of satellite observations of sea ice extent (henceforth called a damped anomaly). Following Wayand et al. (2019), we construct a damped anomaly forecast ($\text{SIE}_{\text{damped}}$) from a climatological trend and a damped anomaly based on the persistence of SIE:

$$\text{SIE}_{\text{damped}}(I, V) = \text{SIE}_{\text{clim}}(\text{doy}_v) + D(I, V), \quad (3)$$

where D is the damped anomaly, I is the forecast initialization date, V is the forecast valid date, doy_I and doy_v are the day of

TABLE 1. Dynamical sea ice forecasting model reforecasts and observational data used in this study. “Reforecast period” refers to the reforecast period used in this study, not the full model reforecast available. More information about all models used in this study is available at <https://confluence.ecmwf.int/display/S2S/Models>.

Label	Organization	Reforecast period	Coupled model	Ocean model	Sea ice model	Ensemble size	Initialization frequency	Forecast length
UKMO	U.K. Met Office	1 Jan 1999–31 Dec 2014	GloSea5	Global Ocean 6.0	CICE5.1	6	4 times per month	60 days
ECMWF	European Centre for Medium-Range Weather Forecasts	1 Jan 1999–31 Dec 2014	IFS CY43R1	NEMO3.4.1	LIM2	10	2 times per week	46 days
METEOFR	Météo France	1 Jan 1999–31 Dec 2014	CNRM-CM6.1	NEMO3.6	GELATO v6	10	Weekly	47 days
NCEP	National Centers for Environmental Prediction	1 Jan 1999–31 Dec 2011	CFSv2	GFDL MOM4	SIS	16	6-hourly	45 days

Label	Organization	Included years	Dataset	Frequency	Resolution
NASA Bootstrap	National Snow and Ice Data Center	1999–2014	NASA Bootstrap	Daily	25 km × 25 km
NASA Team	National Snow and Ice Data Center	1999–2014	NASA Team	Daily	25 km × 25 km

year of I and V , respectively, and SIE_{clim} is the climatological trend. The climatological trend (SIE_{clim}) is created using a 10-yr rolling mean of SIE for each day of the year (doy_v), in each region. The damped anomaly, D , is approximated by a first-order autoregressive process:

$$D(I, V) = [\text{SIE}_{\text{obs}}(I) - \text{SIE}_{\text{clim}}(\text{doy}_I)]\alpha(\text{doy}_I, \text{doy}_V), \quad (4)$$

where SIE_{obs} is the observed sea ice extent for a given date (I), and α is the correlation between sea ice extent anomalies $[\text{SIE}_{\text{obs}}(I) - \text{SIE}_{\text{clim}}]$ on the initialization day of year (doy_I) and the valid day of year (doy_V), based on the observed data. We estimate α using observed SIE (based on NASA Bootstrap) with the linear trend removed from 1989 to 2017. We note that since the common reforecast period used in this study is 1999–2014, the time period used to estimate α extends slightly beyond our common reforecast period. However, observational analyses of daily sea ice persistence indicate that our results are not overly sensitive to the exact time period used to estimate α . Furthermore, the damped persistence benchmark in Eq. (3) is computed over the common reforecast period only (1999–2014). We repeat this process to create a damped anomaly forecast using the NASA Team observational dataset and find that the results are overall similar (not shown).

e. Assessing forecast skill and significance

We evaluate the prediction skill of the S2S forecast models and the damped persistence benchmark forecast using the root-mean-square error (RMSE). For the SIE forecast in each forecast region, we calculate the RMSE as a function of lead time and valid date:

$$\text{RMSE}(L) = \sqrt{\frac{\sum_{I=1}^N [\text{SIE}_{\text{pred}}(I, L) - \text{SIE}_{\text{obs}}(V)]^2}{N}}, \quad (5)$$

where L is the forecast lead time, I is the forecast initialization date, V is the forecast valid date ($I + L$), and N is the number of initialization dates used to compute the mean-square error. Unless otherwise stated, for a given S2S model SIE_{pred} is the ensemble mean.

To assess the significance of our results (i.e., to ensure that our forecast error on VRILE days is significantly different from the error on non-VRILE days), we perform a leave-one-out cross-validation based on each year. That is, for every year in the common reforecast period (1999–2014), we remove each year from the observations and calculate the number of VRILE days over all other years. We choose this approach for several reasons. First, VRILE days are frequently not independent or randomly distributed in time; they often occur consecutively, or as part of an event. Second, we want to ensure that our results are not overly dominated by one or two extreme years, such as the Great Arctic Cyclone of 2012 (e.g., Zhang et al. 2013; Parkinson and Comiso 2013). We then evaluate our S2S model forecasts on VRILE and non-VRILE days for each subset of observations. We assess the significance of the error on VRILE days and the error on non-VRILE days using a two-sided Student's t test. We perform this procedure 16 times, removing each year from the observations from 1999 to 2014. To reject the null hypothesis (i.e., that error on VRILE days is not significantly different from non-VRILE days at 95% confidence), the two RMSEs must be significantly different for at least 13 of 16 folds, as determined by a binomial significance test.

Sea Ice Extent Forecasts, AUG12, East Siberian-Beaufort-Chukchi Sea

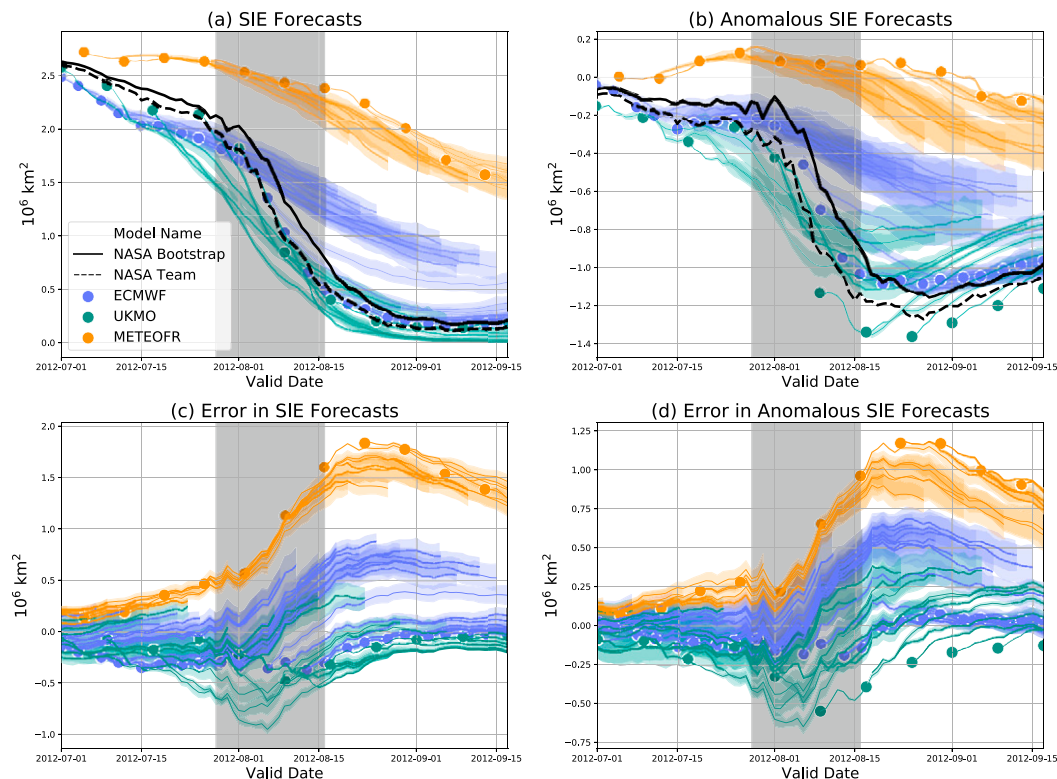


FIG. 2. (a),(b) Example forecasts and (c),(d) forecast errors for (left) sea ice extent (SIE) and (right) anomalous sea ice extent (i.e., a lead-dependent climatological cycle of SIE has been removed from each model) in summer 2012 for the East Siberian–Beaufort–Chukchi Seas. Forecasts from three different S2S models are represented here—ECMWF (blue), UKMO (green), and METEOFR (orange)—with the shading indicating the range of ensemble forecasts of sea ice extent, and circles indicating each forecast initialization. For forecasts of SIE and anomalous SIE, observations are in black (solid: NASA Bootstrap, dashed: NASA Team). Errors are calculated relative to NASA Bootstrap in (c) and (d). The gray shading indicates the dates of the VRILE event.

3. Results

a. Sea ice extent forecasts during summer 2012

With the indication of links between extreme sea ice loss on synoptic time scales and Arctic cyclones discussed in the introduction, we first examine sea ice forecasts during the summer of 2012, specifically during the Great Arctic Cyclone. When we compare regional forecasts of SIE in the East Siberian–Beaufort–Chukchi Seas region to observations (Fig. 2), it is clear that forecasts of this event vary substantially from each other and from the observed SIE. While UKMO model forecasts exhibit the correct magnitude of the rapid SIE decline when initialized before or during the observed period of rapid decline in August 2012, when initialized more than a week in advance, the UKMO forecasts had the timing of the rapid loss event too early. The ECMWF and METEOFR model forecasts are much more biased. When we calculate the error in the forecasts (simply the difference between model forecasts of sea ice extent and the observed sea ice extent; Figs. 2c,d), it is clear that some of the forecasts

initialized before or at the beginning of the period of steep decline in sea ice show substantial errors, and these errors remain high for the rest of the forecast (i.e., the model forecasts never recover). Surveying more of these events, such as summer 2007 (Fig. S1), indicates that this intermodel spread is not necessarily consistent across different events, suggesting that sea ice forecast skill during these events is far from clear. The intermodel spread is likely a result of a combination of factors, including mean state biases (discussed subsequently); differences in model initial conditions and initialization products; and differences in coupled model systems, model physics, model native grids, and land masks.

b. Year-round forecasts of sea ice extent in S2S models

First, we evaluate the prediction skill of the S2S forecast models and the damped persistence benchmark forecast throughout the year using the root-mean-square error. Figure 3 shows the RMSE for all SIE forecasts during the common reforecast period for each region of interest. In Fig. 3, it is clear

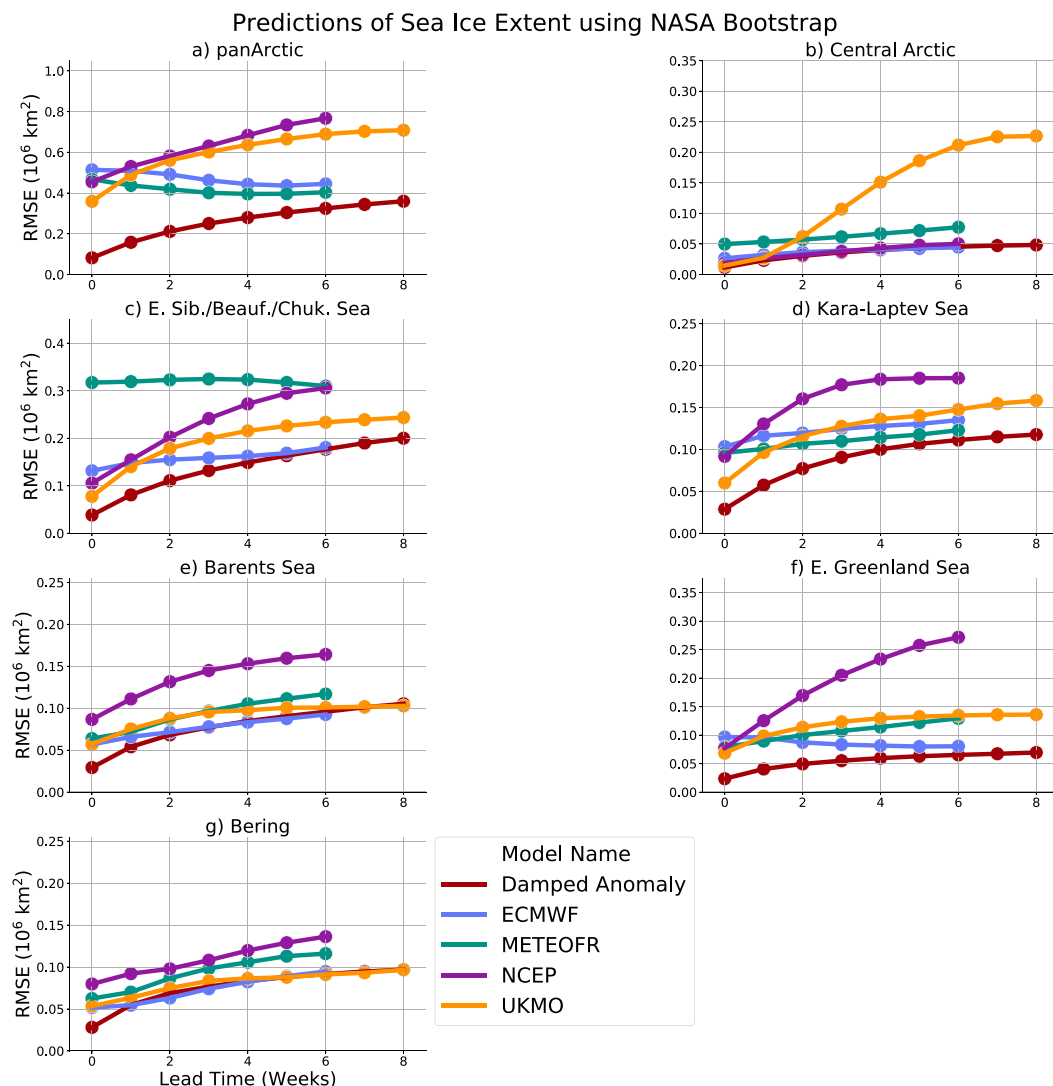


FIG. 3. Root-mean-square error (RMSE) in predictions of sea ice extent for each S2S model as well as the damped anomaly benchmark forecast (red line) for the (a) pan-Arctic, (b) central Arctic, (c) East Siberian–Beaufort–Chukchi Seas, (d) Kara–Laptev Seas, (e) Barents Sea, (f) East Greenland Sea, and (g) Bering Sea regions. RMSE is calculated for all months of the year, on all days.

that error varies widely across different forecast models, even at short lead times. For example, for pan-Arctic SIE predictions (Fig. 3a), the error for lead week 0 (lead days 0–6) ranges from about 0.1 million km^2 for our damped anomaly forecast (red line), to 0.4–0.5 million km^2 for the S2S models. In fact, the errors at early lead times for the ECMWF and METEOFR models are often higher than those at later lead times, highlighting the role of initialization biases. The dynamical models generally struggle to beat the damped anomaly forecast, particularly at shorter lead times. These conclusions are not sensitive to our choice of observations, and results using the NASA Team observed SIE are overall similar (see Fig. S2). We also note that our Fig. 3 seems to potentially contradict previous results, such as Fig. 1 of Zampieri et al. (2018), which shows that the ECMWF

model in particular is the most skillful of the S2S models. We have not fully identified the source of the discrepancies between the raw forecasts in our results and those of Zampieri et al. (2018); however, when model biases are accounted for, as discussed in following sections, these discrepancies are largely resolved.

c. Bias in sea ice extent in S2S models

Because the root-mean-square error for some S2S models is so high even at early lead times (Fig. 3), we compute the bias in SIE predictions for each model as a function of lead time L and forecast valid date V . We then separate the model biases by the forecast valid month and lead time, and average the bias over all years in the common reforecast period. In

Bias in Sea Ice Extent, Lead Days 0-1

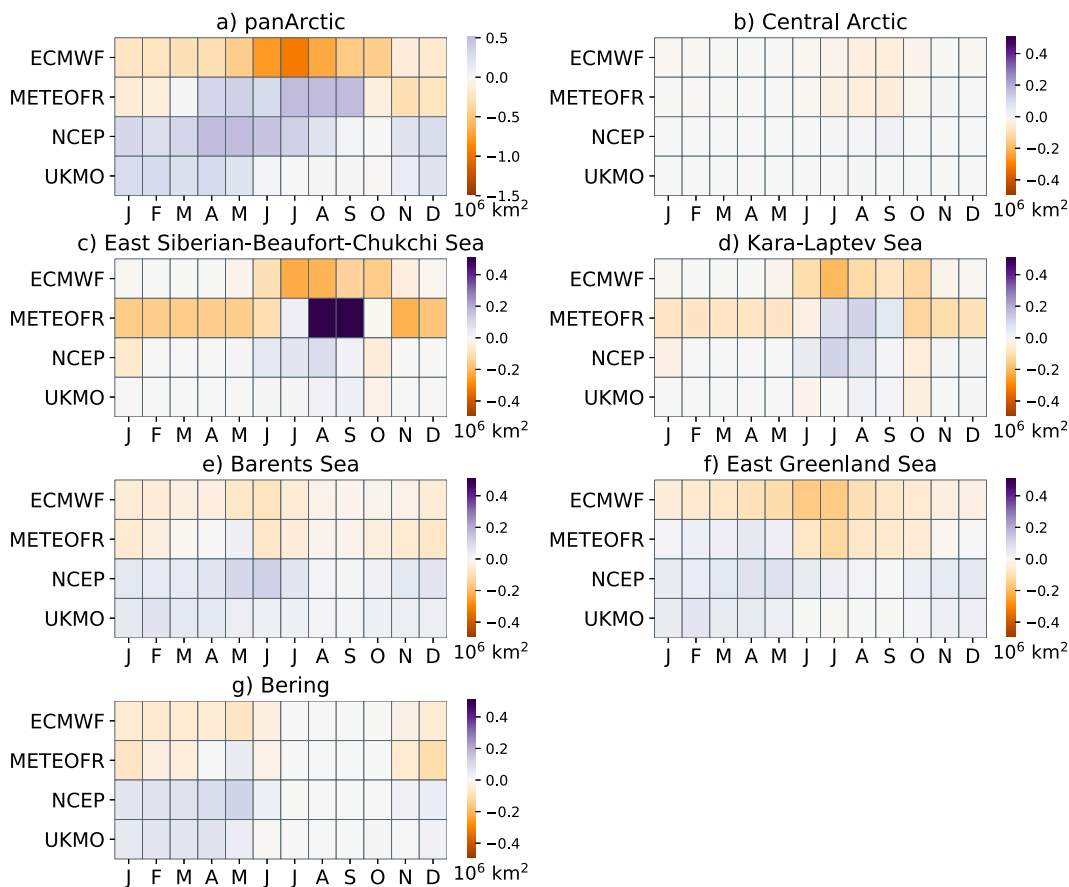


FIG. 4. Bias ($SIE_{\text{model}} - SIE_{\text{obs}}$) in predictions of sea ice extent as a function of forecast valid month. Bias is calculated for lead days 0–1, relative to the NASA Bootstrap observations.

Fig. 4, we compute the bias in SIE predictions for lead times averaged over the first 48 h (lead days 0–1) for each S2S model in each region of interest. For pan-Arctic SIE predictions (Fig. 4a), ECMWF exhibits substantial low biases in SIE throughout the year, even at very short lead times—that is, these models consistently have less sea ice in the pan-Arctic estimates than the observations throughout the year. These biases can be as much as 15% lower than the observed SIE for a given month, and are largest during the summer months. Since these biases are present 24 h after the model run was initialized, this suggests that the error is primarily related to initialization rather than physical processes within the sea ice model. The NCEP, METEOFR, and UKMO estimates of pan-Arctic SIE are closer to the observed SIE (recall that our observed SIE dataset here is NASA Bootstrap).

When we examine different regions of the Arctic, we see a wide range in these biases. The central Arctic, which is typically fully ice-covered throughout the year, has very small biases (Fig. 4b). In the other Arctic Ocean regions, the East Siberian–Beaufort–Chukchi Seas and the Kara–Laptev Seas (Figs. 4c,d), the biases are generally largest in the summertime,

especially in the ECMWF and METEOFR models (these regions, especially the East Siberian–Beaufort–Chukchi Seas, are typically fully ice covered in the winter). Finally, we note that the exact magnitudes of the models' SIE biases will change when we compare them to a different set of observations, such as NASA Team (see Fig. S3). As there is substantial variability in SIE across different observational products, these differences are perhaps not surprising (e.g., Chevallier et al. 2017; Meier and Stewart 2019). Observational uncertainty notwithstanding, many of these general conclusions, in particular, biases varying substantially across models and across regions, and the largest biases occurring in the summer, are not dependent on the specific choice of sea ice dataset, and are present at longer lead times (Fig. S4).

The presence of such large errors just after model initialization highlights the importance of errors in initial conditions, as well as difficulties models have with simulating the sea ice climatological seasonal cycle. These difficulties in simulating sea ice climatology and anomalies may be related to model persistence (e.g., Blanchard-Wrigglesworth et al. 2017; Blanchard-Wrigglesworth and Bushuk 2019), and sea

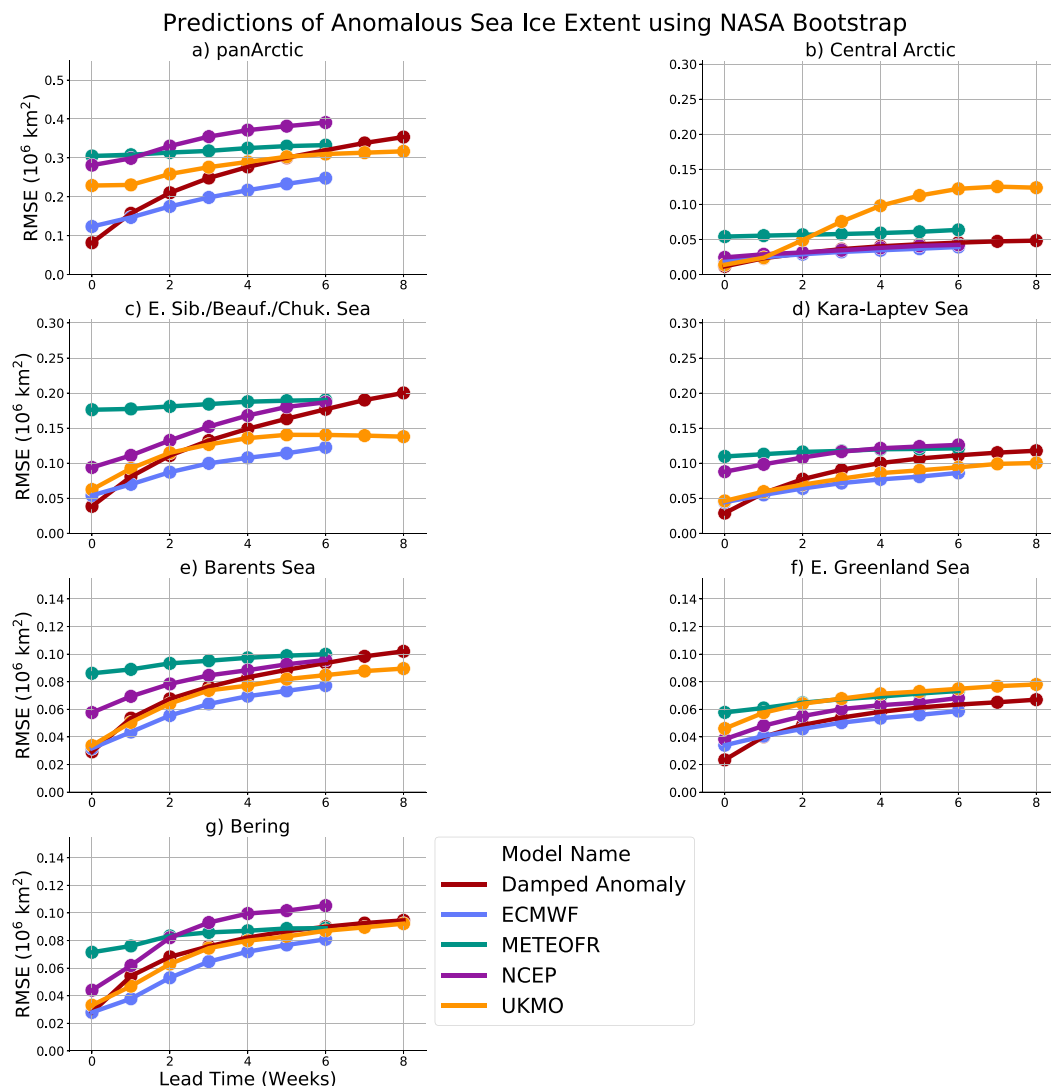


FIG. 5. As in Fig. 3, but for anomalous sea ice extent (climatological cycle has been removed).

ice thickness, which is not assimilated in the S2S reforecasts (e.g., Blockley and Peterson 2018; Balan-Sarojini et al. 2021). When we remove the climatological SIE from the forecasts and instead evaluate how well models predict deviations from climatology, we reduce the impact of these biases. Figure 5 shows the error in forecasts of anomalous SIE for all days throughout the year. Compared to Fig. 3, the intermodel spread has been reduced, and the S2S model error is decreased, especially for the ECMWF, which goes from being moderately skillful at pan-Arctic SIE predictions in Fig. 3 to the most skillful model for pan-Arctic SIE predictions in Fig. 5. In fact, when removing climatology some S2S models' forecasts in some regions exhibit skill compared to the damped anomaly benchmark forecasts as early as lead days 7–13, while other models' forecasts in some regions still exhibit less skill at nearly all lead times than the damped anomaly benchmark forecast. When we compare

the skill of model predictions of VRILE and non-VRILE days, we will focus only on predictions made using anomalous SIE, as these results are less sensitive to the model initial conditions.

We also note that there is some seasonality in prediction skill of S2S forecast models, as has been noted in previous studies (e.g., Blanchard-Wrigglesworth et al. 2011a; Bushuk et al. 2017, 2019). For Arctic Ocean regions as well as pan-Arctic SIE, forecast skill is generally highest in the winter (when these regions are mostly or entirely ice-covered), and lower in the summer (Fig. 6); in regions that are fully ice-covered in the winter like the East Siberian–Beaufort–Chukchi Seas, any error in forecast skill can be attributed to model biases. For regions like the Barents and East Greenland Seas, which have year-round sea ice but are not fully ice covered in the winter, forecast skill does not exhibit a strong seasonality, and remains at similar levels throughout the year. In the

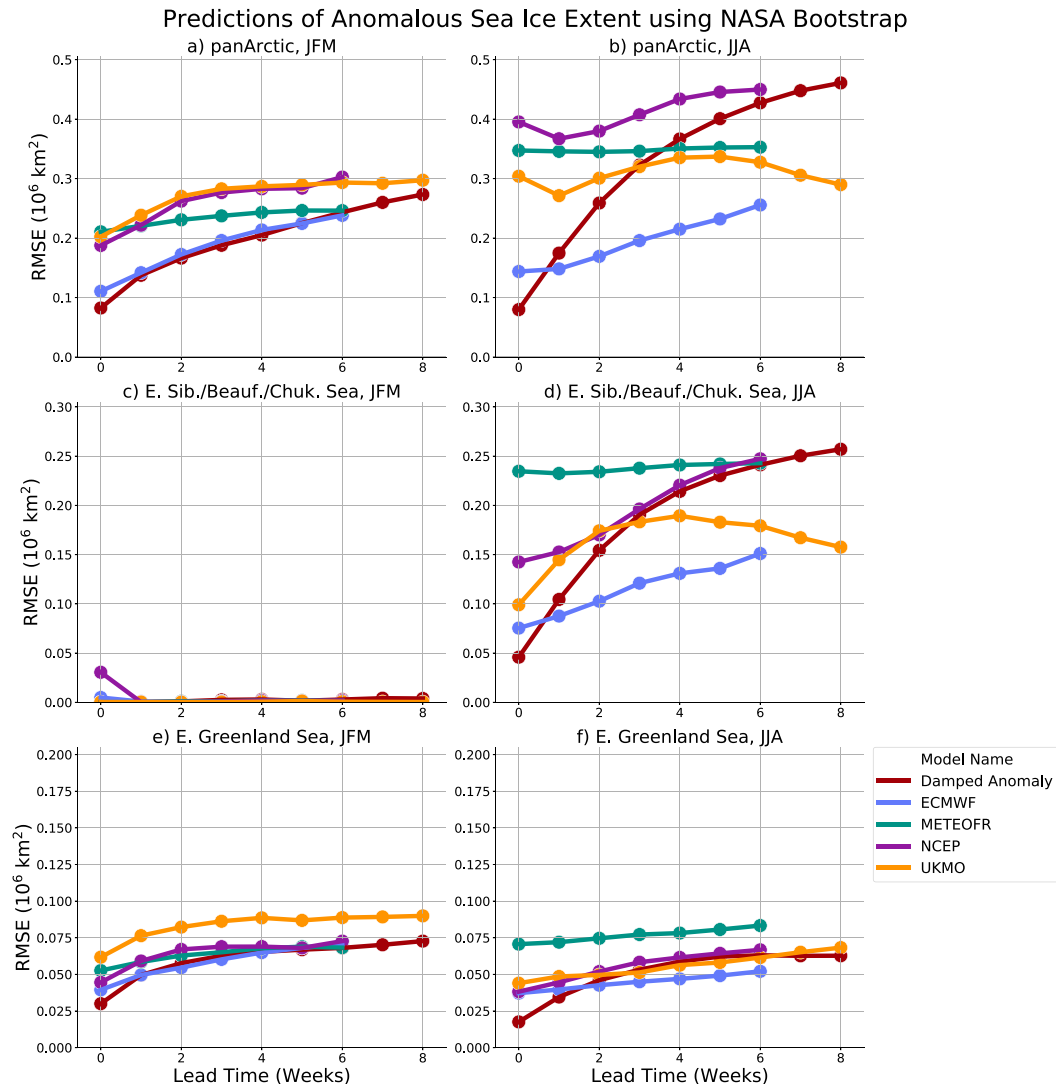


FIG. 6. As in Fig. 3, but for anomalous sea ice extent (climatological cycle has been removed) in (left) January–March (JFM) and (right) June–August (JJA).

Bering Sea (not shown), which is typically ice-free by late summer, forecast skill is higher in summer than in the winter, where sea ice variability is greatest. In this paper, we will focus primarily on sea ice forecasts for specific months or seasons rather than year-round forecasting.

d. Seasonal and regional distribution of VRILE days

Next, we explore the seasonal and regional distribution of VRILE days in observations. Figure 7 breaks down the frequency of occurrence of VRILE days as a function of region and month during the common reforecast period of 1999–2014. We note that our conclusions are not sensitive to this 1999–2014 analysis period; regional and seasonal distributions of VRILE days are very similar when estimated over the full period with daily observations of 1988–2019 (not shown). Across the entire Arctic, the great majority of VRILE days occur during the summer

months—June, July, and August. Summertime is the season of sea ice melt in the Arctic, so it seems intuitive that most of the days with extreme sea ice loss on day-to-day time scales occur during the months with the most sea ice melt overall. Sea ice is also thinner in the summer, meaning that it is easier to melt (e.g., Holland et al. 2006). That there are no VRILE days in September for pan-Arctic estimates of VRILE days may initially seem strange. However, by early September there is simply less ice available to melt and by late September freeze-up has begun; thus, the days that comprise the most extreme 5th percentile of sea ice loss tend to occur earlier in the summer. This is especially true for Arctic-wide estimates.

When we explore the seasonality of VRILE days in different regions of the Arctic, we see two primary regimes of behavior. The regions that incorporate the sea ice of the Arctic Ocean—the central Arctic, the East Siberian–Beaufort–Chukchi Seas,

Percentage of VRILE Days That Fall In Each Month, SIE, week 0

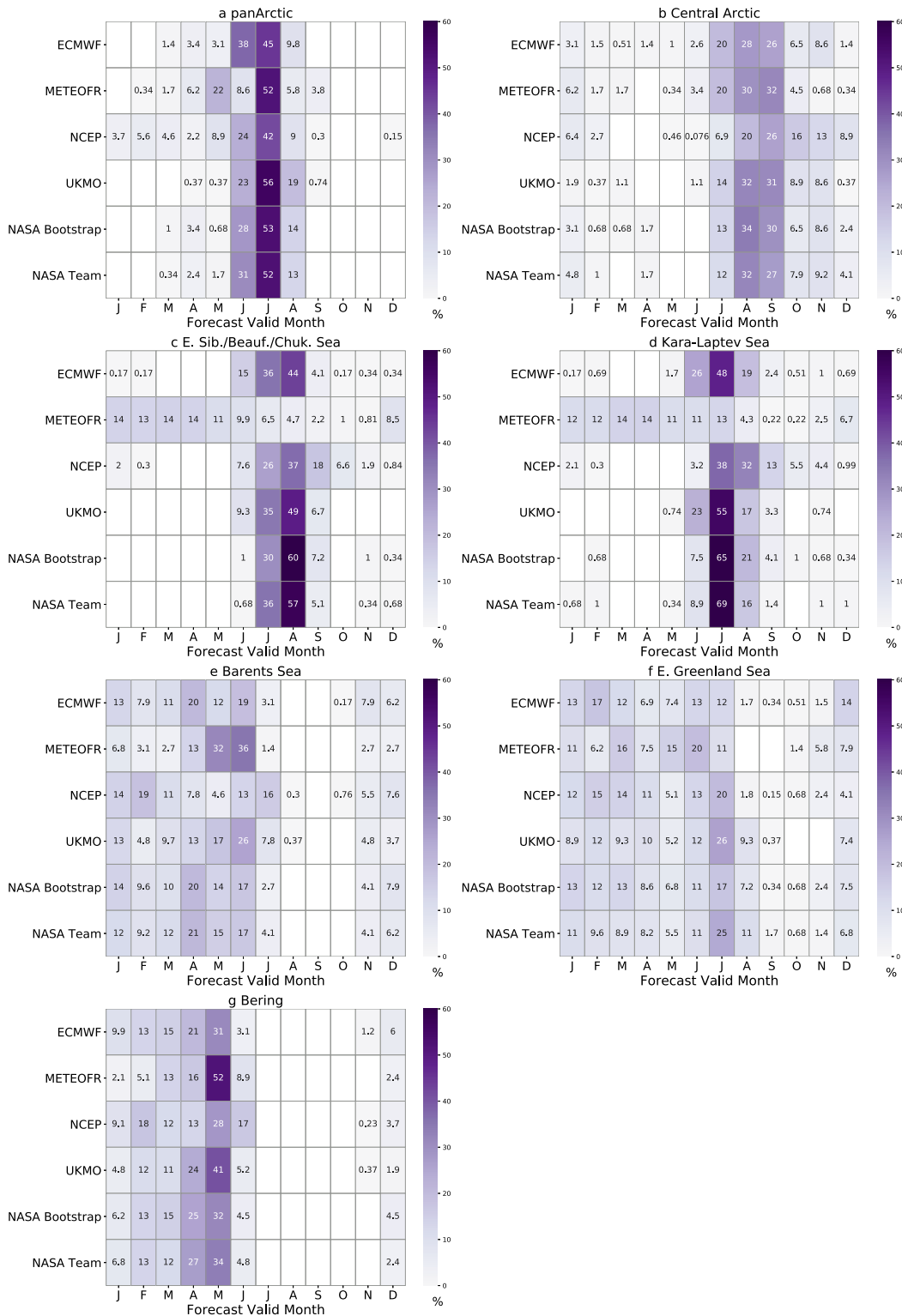


FIG. 7. Percent of VRILE days per month for S2S models and both sets of observations (NASA Bootstrap and NASA Team) in select regions. Model forecasts of sea ice for lead times of up to 1 week (days 0–6) are included. Numbers in the box indicate the percentage of each region's VRILE days that occurred in that month. Empty boxes indicate no VRILE days occurred in that month. Month is based on the forecast valid date. VRILE days are calculated based on the bottom 5th percentile of the 5-day change in sea ice extent.

and the Kara–Laptev Seas—show a similar seasonality to the Arctic-wide estimates of VRILE seasonality. In these regions, most VRILE days occur in summer, primarily in July and August for the Siberian Seas and the Kara–Laptev Seas, and in August and September for the central Arctic. However, in the areas where the Arctic meets the Atlantic Ocean—the Barents and East Greenland Seas—we see that VRILE days occur throughout the year. In the Barents Sea, they are most common in spring and early summer and do not occur in late summer (the Barents Sea is ice-free in late summer); in the East Greenland Sea, where ice occurs year-round, VRILEs occur most frequently in July, but nearly as many VRILEs occur in January, February, and March. Most VRILE days in the Bering Sea, which is ice-free in the summer and early autumn, occur in April and May. Thus, while the characterization of VRILE days as a summertime phenomenon is applicable to Arctic-wide estimates of VRILEs, as well as VRILE behavior in the Arctic Ocean, VRILEs are not universally limited to summertime and can occur throughout the year in some regions.

How well do S2S models capture the observed seasonal and regional variability of VRILE days? Figures 7 and 8 display VRILE days as a function of region and forecast valid month for each S2S model. Forecasts at lead times of up to 1 week are included in Figs. 7 and 8; however, the results look similar when forecasts at longer lead times are included (Fig. S8). Figures 7 and 8 show that for the most part, the S2S models capture the observed seasonal and regional variability of VRILE days. For Arctic-wide VRILE estimates, as well as those regions in the Arctic Ocean basin, we see that most VRILE days occur in summer, whereas VRILE days occur throughout the year in the Barents and East Greenland Seas. Generally, the UKMO model resembles the observations most closely. The NCEP model differs the most from the observations, with VRILEs in the Arctic Ocean regions occurring much later in the year than in the observations, whereas Arctic-wide VRILEs are much more common in the winter in the NCEP model.

e. Sea ice forecasts on VRILE days

Figures 7 and 8 indicate that S2S models are generally capable of capturing the observed seasonality of VRILE events. With that in mind, we explore the sea ice forecast skill on these extreme sea ice loss days. That is, is sea ice more or less predictable when there are extreme sea ice loss events? And does the answer depend on the forecast lead time? We compare the forecast skill of SIE on days with VRILEs to the forecast skill of SIE on days without VRILEs according to observations by computing an error ratio R :

$$R(L) = \frac{\text{RMSE}_{\text{VRILE}}(L)}{\text{RMSE}_{\text{nonVRILE}}(L)}. \quad (6)$$

Thus, when $R > 1$, the error in predictions of SIE on VRILE days ($\text{RMSE}_{\text{VRILE}}$) is greater than the error on typical days ($\text{RMSE}_{\text{nonVRILE}}$), indicating that SIE forecast skill is lower on VRILE days. When $R < 1$, we have the opposite case—the error is lower on VRILE days, suggesting that SIE forecast skill is higher on VRILE days than it is on typical

days. The quantity R is computed at all available forecast lead time L , with VRILEs occurring in the observational record at valid time $V = I + L$.

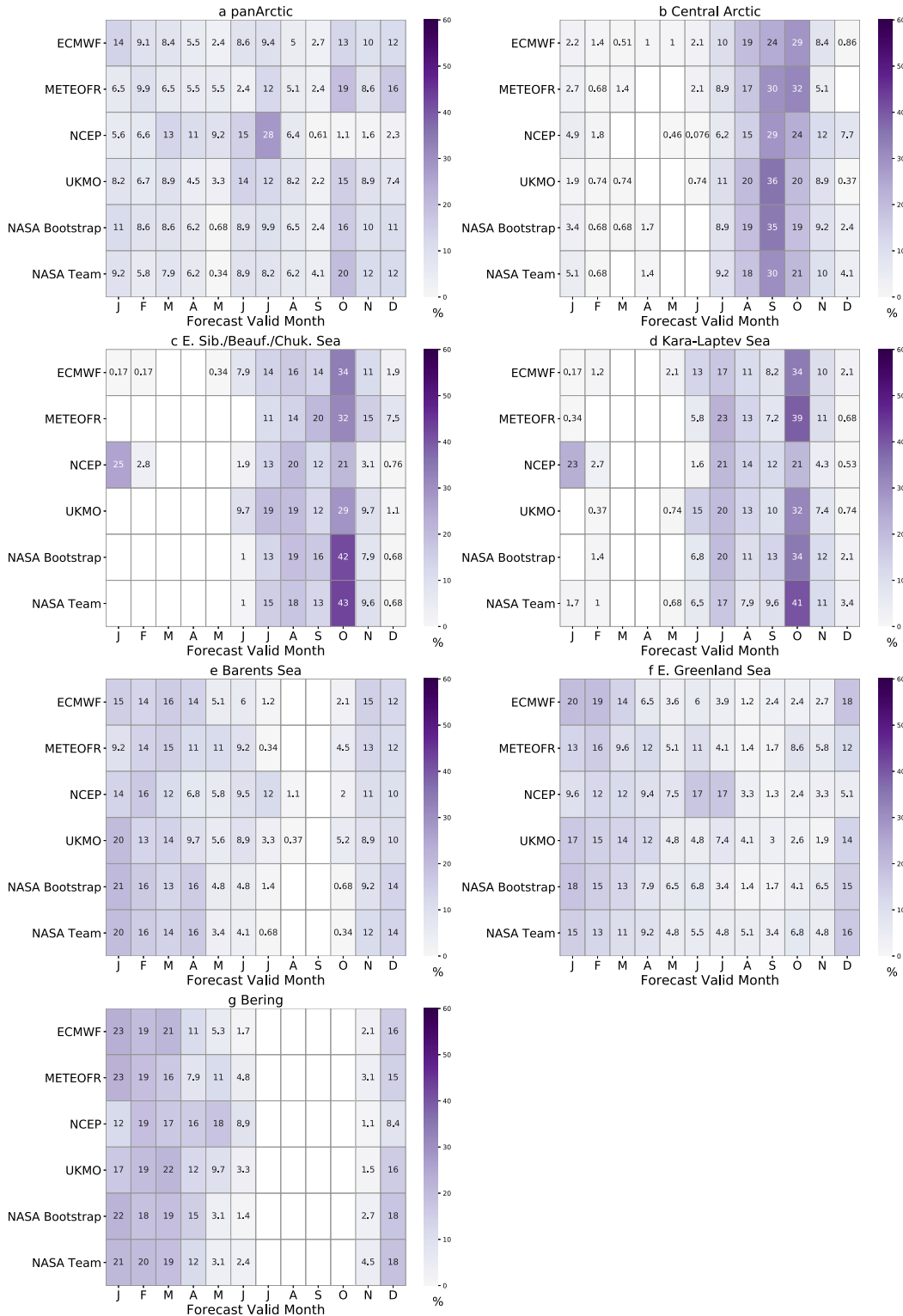
We begin by comparing predictions on VRILE and non-VRILE days in June–August. Figure 9 shows R for anomalous SIE forecasts. We see that generally, for most models in most regions, sea ice is less predictable on VRILE days than on non-VRILE days. The relative error remains high on VRILE days in most regions at all lead times. We note that R often decreases as lead time increases. This decrease in R is not a decrease in the RMSE itself, but rather indicates that forecast skill on non-VRILE days approaches forecast skill on VRILE days at longer lead times, as both sets of days approach their climatological saturation values. No model appears to consistently have greater relative forecast skill on VRILE days (i.e., lower R) than the rest in all regions. Thus, despite evidence of atmospheric circulation anomalies preceding VRILE days (e.g., Wang et al. 2020; Blanchard-Wrigglesworth et al. 2021), sea ice is not actually more predictable during these extreme events. In addition to intermodel spread in forecast error [noted for the 2018 sea ice forecasts by Zampieri et al. (2018) and Wayand et al. (2019)], there is also substantial regional diversity in terms of model performance. For example, in the Arctic Ocean regions (the central Arctic and the East Siberian–Beaufort–Chukchi Seas), the error ratio for all S2S models, including the damped anomaly forecast, is high, suggesting that forecast skill of SIE during VRILE events is low relative to skill on non-VRILE days in these regions.

By contrast, there is less difference in the forecast skill on VRILE and non-VRILE days during the winter months of January–March for pan-Arctic SIE, as well as SIE in the Bering and East Greenland Sea regions (Fig. 10). The forecast skill ratios are overall lower in the wintertime, remaining below 2 (i.e., the difference in forecast skill on VRILE and non-VRILE days is smaller than it is in the summertime). In many regions, the differences in forecast skill on VRILE and non-VRILE days are fairly small, particularly after the first 1–2 lead weeks. Still, broadly speaking, sea ice forecast skill is lower on VRILE days than it is on non-VRILE days, especially at shorter lead times.

f. Error growth following VRILEs

Finally, we analyze the errors in sea ice prediction in the days following VRILE events to investigate whether VRILEs exert a longer-term influence on the sea ice extent forecast skill. For example, consider the summer 2012 sea ice forecasts depicted in Fig. 2. On 28 July 2012, the 5-day change in observed SIE in the East Siberian–Beaufort–Chukchi Seas dropped below the 5th-percentile threshold and remained below that threshold for almost three weeks (indicated by gray shading). The S2S model forecasts have divergent responses: while the observed SIE rapidly decreases shortly after the VRILE begins, the ECMWF forecasts initialized before or right at the beginning of the VRILE do not decrease nearly as quickly, and error in the ECMWF SIE forecasts remains high during and after the VRILE. The

Percentage of VRILE Days That Fall In Each Month, Anomalous SIE, week 0

FIG. 8. As in Fig. 7, but with VRILE days calculated based on the bottom 5th percentile of the 5-day change in *anomalous* sea ice extent.

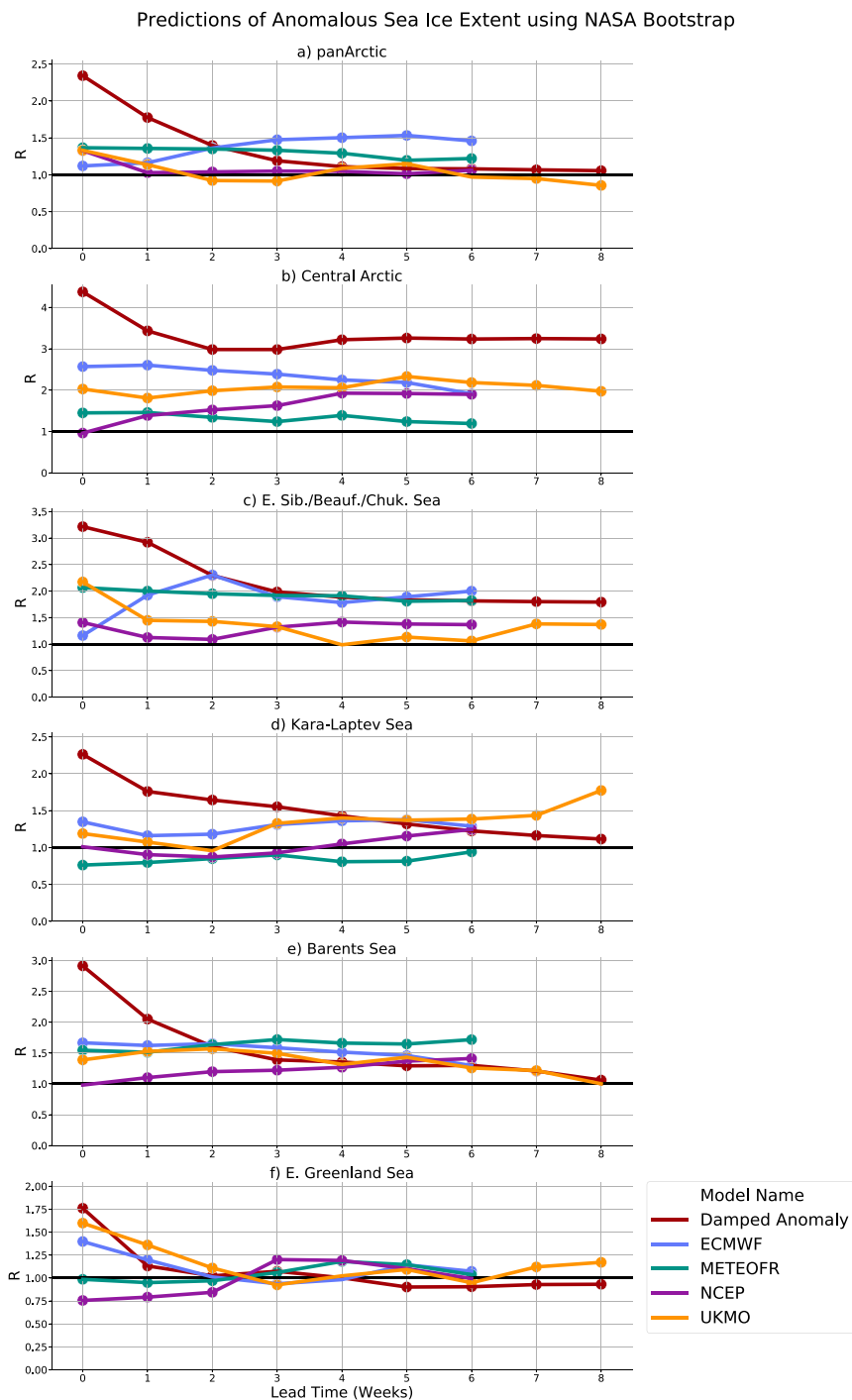


FIG. 9. Ratio of RMSE of sea ice extent forecasts in June–August for VRILE days vs non-VRILE days ($RMSE_{VRILE}/RMSE_{noVRILE}$) for anomalous sea ice extent. Values that are greater than 1 indicate that sea ice is less predictable on VRILE days, while values less than 1 indicate that sea ice on VRILE days is more predictable on VRILE days as compared to non-VRILE days. Filled circles indicate that RMSE on VRILE days is significantly different from RMSE on non-VRILE days at 95% confidence.

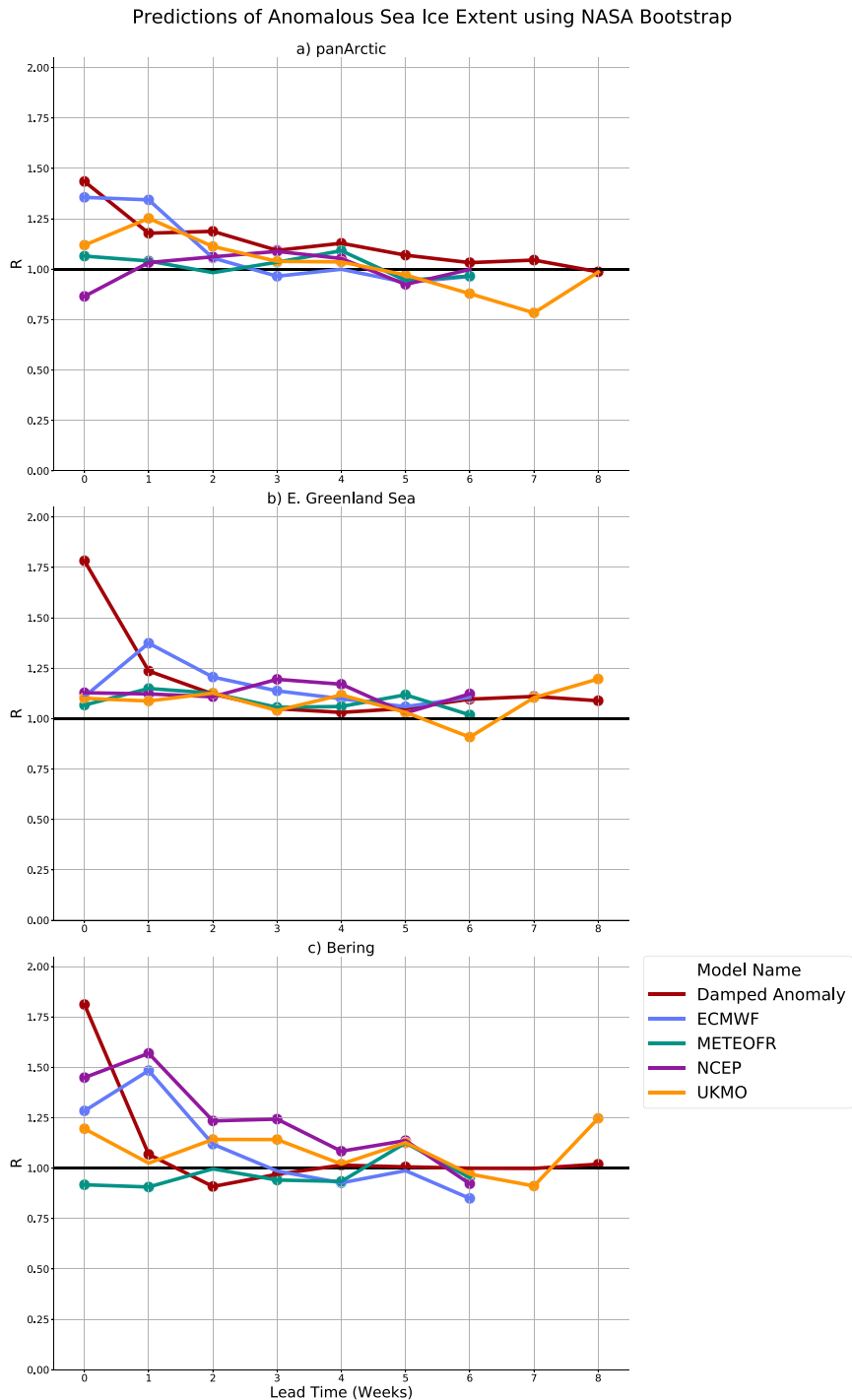


FIG. 10. As in Fig. 9, but for sea ice forecasts in January–March.

UKMO forecasts initialized before the VRILE event demonstrate a remarkably sharp decline at the VRILE start, similar to the observed decline. On the whole, the UKMO forecasts have moderate error during the VRILE that declines once the VRILE event is over. The METEOFR forecasts behave

similarly to the ECMWF forecasts, with larger errors at initialization.

In the previous section, we showed that the forecast error is typically enhanced on VRILE days. Here, we ask if it remains enhanced after the VRILE, or does the error return to its

Predictions of Anomalous Sea Ice Extent

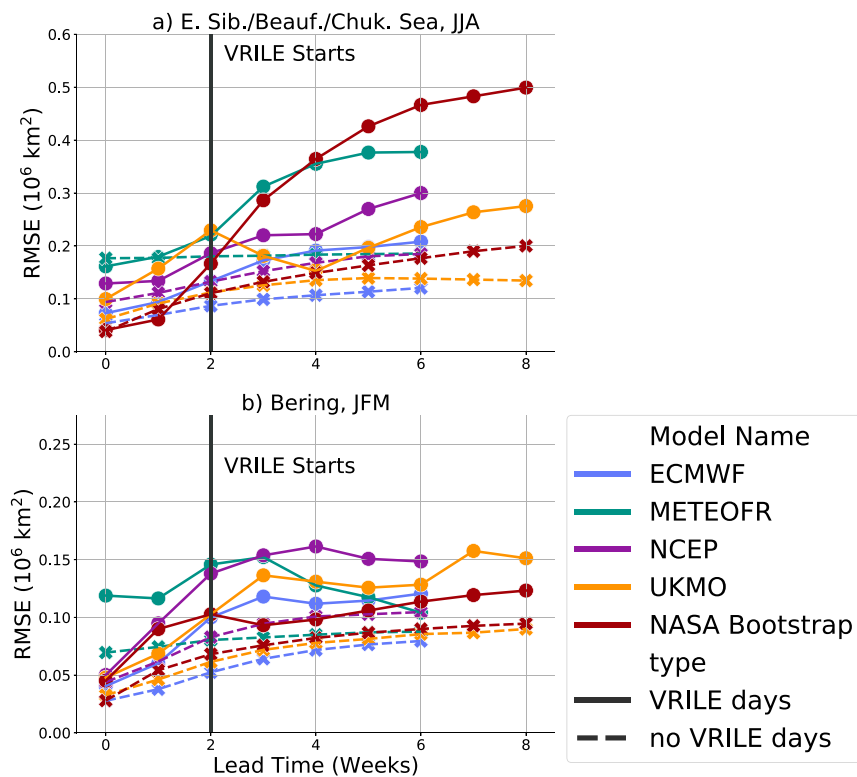


FIG. 11. Predictions of anomalous sea ice extent forecasts initialized 2 weeks before the onset of a VRILE event in S2S models for the (a) East Siberian–Beaufort–Chukchi Seas in June–August and (b) Bering Sea in January–March. The black line indicates the onset of the VRILE events. Filled circles indicate that RMSE on VRILE days is significantly different from RMSE on non-VRILE days at 95% confidence.

typical non-VRILE value? To explore this question more systematically, we composite RMSE on VRILE events aligned by a common lead time of 14 days. Given our results in the previous section, we expect the forecast error to rise sharply on day 14, when the VRILE occurs. Here we seek to determine whether the forecast error remains high following the VRILE. We found it important to ensure that there is a non-VRILE baseline period prior to the VRILE on which we composite to ensure that we can cleanly identify the error dependence on lead time. Therefore, we stipulate that all VRILE events must be separated from the next VRILE day by at least 14 days. Since forecast initialization frequency varies by model, we expand our VRILE onset days slightly to encompass VRILEs that start anywhere from 11 to 14 days after the forecast initialization.

Figure 11 shows the forecast error propagation, with forecasts initialized 14 days prior to the onset of the VRILE event. Here, we focus on two specific regions as examples—the East Siberian–Beaufort–Chukchi Seas in June–August, and the Bering Sea in January–March. For summertime sea ice forecasts in the East Siberian–Beaufort–Chukchi Seas (Fig. 11a), we see that at lead weeks 0–1 (0–13 days before the VRILE begins), there is

little difference between the error on VRILE and non-VRILE days. (We note that while, in many cases, the differences in RMSEs on VRILE and non-VRILE days during lead weeks 0–1 are *statistically* significant in Fig. 11, the actual magnitude of these differences is quite small, and unlikely to be physically meaningful.) Beginning at lead week 2 (day 14), when our VRILE begins, the RMSE starts to increase. In the weeks following the onset of the VRILE, the error in most S2S models and the damped anomaly forecast increases much faster than the RMSE on non-VRILE days, and remains elevated relative to non-VRILE days during the rest of the forecast period. The exception to this pattern is the UKMO model, whose RMSE at forecast lead times of 3 and 4 weeks decreases before increasing again starting in lead week 5. We suspect that this behavior is related to the small number of VRILE events and the low initialization frequency of the UKMO model, and we note that in other regions and at other time intervals (i.e., UKMO forecasts initialized 1 week or 3 weeks before the VRILE onset) this behavior is not present (Fig. S5).

In the wintertime (Fig. 11b), the pattern is fairly similar—14 days before the onset of the VRILE, there is little

difference in the RMSE for VRILE and non-VRILE forecasts (the exception is METEOR, which generally has the largest initialization biases). At lead week 2, when the VRILE begins, the error increases much more for VRILE forecasts than for non-VRILE forecasts. The error stays higher after the VRILE events in the following weeks, although this effect is somewhat less pronounced in the wintertime than it is in the summer.

This pattern of behavior of RMSE remaining enhanced relative to typical RMSE values following VRILE events is quite robust in the damped anomaly forecast. In the S2S models, there is more variability across regions and models, and less consistency with this behavior (although generally the forecast errors do remain high after the VRILE), highlighting again the role of regional variability and different model physics and configurations.

4. Discussion and conclusions

In this study, we have explored extreme sea ice loss on subseasonal time scales in a subset of dynamical forecast models in the S2S project. We have determined that the dynamical forecast models are generally capable of simulating extreme sea ice loss days with a regional and seasonal variability that is similar to observed seasonality and regional variability. In particular, we note that while these extreme sea ice loss days are often considered to be a summertime phenomenon, in some regions, such as the Barents and East Greenland Seas, these days actually occur throughout the year.

At subseasonal to seasonal time scales, atmospheric initial conditions are especially important (e.g., Mohammadi-Aragh et al. 2018), and extreme sea ice loss events on subseasonal time scales have been linked to atmospheric circulation anomalies in observations (e.g., Wang et al. 2020). Previous observational studies have associated large-scale anticyclonic circulation anomalies with lower Arctic-wide September SIE (e.g., Ding et al. 2017; Wernli and Papritz 2018), and summers with more VRILE days have also been linked to lower September SIE (e.g., Wang et al. 2020). Furthermore, regional and seasonal variability in cyclones is mirrored somewhat by regional and seasonal variability in VRILEs; polar cyclones are most common in the subpolar seas in winter, and in the Arctic Ocean basin in the summer (e.g., Clancy et al. 2022), similar to the VRILE distributions shown here in Figs. 7 and 8. Thus, if VRILE events and large-scale atmospheric circulation anomalies are closely coupled, it is reasonable to hypothesize that these events might be more predictable in forecast models than a typical day.

However, we find that SIE forecasts on VRILE days typically have larger forecast error than SIE forecasts on non-VRILE days (Figs. 9–11), suggesting that the atmospheric impacts on VRILE events are far from straightforward, especially in forecasting models. While some observational studies have suggested links between large-scale circulation and seasonal SIE (e.g., Ding et al. 2017; Wernli and Papritz 2018), and between VRILE events and seasonal SIE (e.g., Wang et al. 2020), the true role of individual storms on sea ice loss is

less clear. For example, different studies draw different conclusions about the role of the August 2012 polar cyclone on the September 2012 sea ice minimum (e.g., Parkinson and Comiso 2013; Zhang et al. 2013; Stern et al. 2020); other studies have linked cyclones to increases in SIE (e.g., Schreiber and Serreze 2020) or have emphasized the heterogeneous response of sea ice to cyclones (e.g., Clancy et al. 2022). Wang et al. (2020) suggest that summertime VRILE events are often preceded by a coupled high and low pressure anomaly rather than a cyclone alone. Furthermore, even if atmospheric events are properly forecast, the sea ice response to atmospheric forcing may or may not be properly represented in forecast models (e.g., Blanchard-Wrigglesworth et al. 2021). Finally, when we calculate the error propagation in the damped anomaly forecasts following VRILE events (Fig. 11), we see that the error remains high for the rest of the forecast period following the VRILE event, suggesting that these events could wield some longer-term influence. Thus, a fuller understanding of the atmospheric dynamics that drive VRILE events is necessary to both understand these events better themselves, and to improve model forecasts of VRILE events.

Dynamical forecast models still frequently exhibit higher forecast errors in predicting SIE when compared to a damped anomaly forecast. There are several possible reasons for this. One is that the dynamical forecast models exhibit biases relative to the observed SIE, even at very short lead times. This bias in initialization has been previously noted (e.g., Zampieri et al. 2018; Wayand et al. 2019), and can be at least somewhat reduced by removing the climatological cycle of SIE and predicting anomalous SIE (Fig. 5). Removing the climatological cycle also reduces but does not eliminate the intermodel spread. We also note that there is variability across different observed and reanalysis estimates of sea ice concentration (e.g., Chevallier et al. 2017; Meier and Stewart 2019); however, biases at lead days 0–1 remained large when compared to both of our observational datasets. Reducing these initialization biases is still critical to improving dynamical forecast model performance for subseasonal time scales; a recent study using a single climate model suggests that initial conditions biases can dominate forecast biases for 3 weeks or longer (Cruz-García et al. 2021). Improved data assimilation is one way to attack this problem, as noted by Zampieri et al. (2018), Wayand et al. (2019). We also note, however, that the models are biased not just in their mean states but also in their variability, suggesting that improved data assimilation alone may not be enough (Fig. S6).

Another possible area for improvement is the assimilation of sea ice thickness (SIT). The four sets of S2S model reforecasts used in this study do not assimilate SIT, nor do they provide forecasts for SIT. Assimilation of SIT has been shown to improve seasonal forecasts of SIC and SIE (e.g., Day et al. 2014), and SIT anomalies have strong persistence, which means they can impact SIC anomalies on long time scales (i.e., Blanchard-Wrigglesworth and Bitz 2014; Blanchard-Wrigglesworth et al. 2017; Blanchard-Wrigglesworth and Bushuk 2019). Recent studies have highlighted the potential for assimilation of SIT to improve sea ice forecasts. Blockley and Peterson (2018) and Balan-Sarajini et al. (2021) found

that assimilation SIT in the UKMO and ECMWF models, respectively, improved seasonal forecasts of sea ice by reducing biases in sea ice extent, area, and volume, as well as biases in near-surface temperature and pressure over sea ice. Blockley and Peterson (2018) show a clear relationship between biases in wintertime SIT and errors in summertime SIE and SIC, and the results of both Blockley and Peterson (2018) and Balan-Sarojini et al. (2021) discuss the role of SIT assimilation in sea ice melting and freezing. Finally, regridding and land masking choices can also have substantial impacts on model biases in sea ice extent, and decisions regarding regridding and land masking (Fig. S7).

There is reason to believe that dynamical forecast models can be useful for predicting extreme sea ice loss events going forward. While dynamical forecast models are not more skilled at predicting sea ice on VRILE days than on typical days, they are often not vastly less skilled, and their relative skill is often superior to our damped anomaly forecast. That is, the damped anomaly forecast is often worse at predicting SIE on VRILE days relative to its skill on non-VRILE days, especially at short lead times. So, while SIE on VRILE days is not overall more predictable than it is on non-VRILE days, the dynamical forecast models may still be able to make better predictions on these days than our damped anomaly forecast.

Acknowledgments. We are grateful to the Office of Naval Research for funding this research via Grant N00014-18-1-2175, and to the three anonymous reviewers who provided valuable feedback on this manuscript. E.B.W. is also supported by NSF (OPP Grant 1751363).

Data availability statement. Observed sea ice concentration data are available from the NASA National Snow and Ice Data Center Distributed Active Archive Center. The sea ice concentration data analyzed using the NASA Bootstrap algorithm are available at <https://doi.org/10.5067/7Q8HCCWS4I0R> as cited in Comiso (2017). The sea ice concentration data analyzed using the NASA Team algorithm are available at <https://doi.org/10.5067/8GQ8LZQVL0VL> as cited in Cavalieri et al. (1996). This work is based on S2S data. S2S is a joint initiative of the World Weather Research Programme (WWRP) and the World Climate Research Programme (WCRP). The original S2S database is hosted at ECMWF as an extension of the TIGGE database. S2S model reforecasts, as described in Vitart et al. (2017), are available at <https://apps.ecmwf.int/datasets/data/s2s-reforecasts-daily-averaged-ecmf/>, copyright 2021 European Centre for Medium-Range Weather Forecasts (ECMWF).

REFERENCES

- Balan-Sarojini, B., S. Tietsche, M. Mayer, M. Balmaseda, H. Zuo, P. de Rosnay, T. Stockdale, and F. Vitart, 2021: Year-round impact of winter sea ice thickness observations on seasonal forecasts. *Cryosphere*, **15**, 325–344, <https://doi.org/10.5194/tc-15-325-2021>.
- Batté, L., I. Välisuo, M. Chevallier, J. Acosta Navarro, P. Ortega, and D. Smith, 2020: Summer predictions of Arctic sea ice edge in multi-model seasonal re-forecasts. *Climate Dyn.*, **54**, 5013–5029, <https://doi.org/10.1007/s00382-020-05273-8>.
- Blanchard-Wrigglesworth, E., and C. 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>.
- , and M. Bushuk, 2019: Robustness of Arctic sea-ice predictability in GCMs. *Climate Dyn.*, **52**, 5555–5566, <https://doi.org/10.1007/s00382-018-4461-3>.
- , K. Armour, C. Bitz, and E. DeWeaver, 2011a: Persistent and inherent predictability of Arctic sea ice in a GCM ensemble and observations. *J. Climate*, **24**, 231–250, <https://doi.org/10.1175/2010JCLI3775.1>.
- , C. Bitz, and M. Holland, 2011b: Influence of initial conditions and climate forcing on predicting Arctic sea ice. *Geophys. Res. Lett.*, **38**, L18503, <https://doi.org/10.1029/2011GL048807>.
- , R. Cullather, W. Wang, J. Zhang, and C. Bitz, 2015: Model forecast skill and sensitivity to initial conditions in the seasonal sea ice outlook. *Geophys. Res. Lett.*, **42**, 8042–8048, <https://doi.org/10.1002/2015GL065860>.
- , and Coauthors, 2017: Multi-model seasonal forecast of Arctic sea-ice: Forecast uncertainty at pan-Arctic and regional scales. *Climate Dyn.*, **49**, 1399–1410, <https://doi.org/10.1007/s00382-016-3388-9>.
- , A. Donohoe, L. Roach, A. DuVivier, and C. Bitz, 2021: High-frequency sea ice variability in observations and models. *Geophys. Res. Lett.*, **48**, e2020GL092356, <https://doi.org/10.1029/2020GL092356>.
- Blockley, E., and K. Peterson, 2018: Improving Met Office seasonal predictions of Arctic sea ice using assimilation of CryoSat-2 thickness. *Cryosphere*, **12**, 3419–3438, <https://doi.org/10.5194/tc-12-3419-2018>.
- Bonan, D., 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>.
- 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, A. Rosati, and R. Gudgel, 2019: Regional Arctic sea-ice prediction: Potential versus operational seasonal forecast skill. *Climate Dyn.*, **52**, 2721–2743, <https://doi.org/10.1007/s00382-018-4288-y>.
- Cavalieri, D., C. Parkinson, P. Gloersen, and H. 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 21 February 2020, <https://doi.org/10.5067/8GQ8LZQVL0VL>.
- Chevallier, M., and Coauthors, 2017: Intercomparison of the Arctic sea ice cover in global ocean-sea ice reanalyses from the ORA-IP project. *Climate Dyn.*, **49**, 1107–1136, <https://doi.org/10.1007/s00382-016-2985-y>.
- Clancy, R., C. Bitz, E. Blanchard-Wrigglesworth, M. McGraw, and S. Cavallo, 2022: A cyclone-centered perspective on the drivers of asymmetric patterns in the atmosphere and sea ice during Arctic cyclones. *J. Climate*, **35**, 73–89, <https://doi.org/10.1175/JCLI-D-21-0093.1>.
- Comiso, J., 2017: Bootstrap sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS, version 3. NASA National Snow and Ice Data Center Distributed Active Archive Center, accessed 21 February 2020, <https://doi.org/10.5067/7Q8HCWS4I0R>.

- Cruz-García, R., P. Ortega, V. Guemas, J. C. A. Navarro, F. Massonnet, and F. J. Doblas-Reyes, 2021: An anatomy of Arctic sea ice forecast biases in the seasonal prediction system with EC-Earth. *Climate Dyn.*, **56**, 1799–1813, <https://doi.org/10.1007/s00382-020-05560-4>.
- 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>.
- Ding, Q., and Coauthors, 2017: Influence of high-latitude atmospheric circulation changes on summertime Arctic sea ice. *Nat. Climate Change*, **7**, 289–295, <https://doi.org/10.1038/nclimate3241>.
- 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**, GL083831, <https://doi.org/10.1029/2019GL083831>.
- Eicken, H., 2013: Arctic sea ice needs better forecasts. *Nature*, **497**, 431–433, <https://doi.org/10.1038/497431a>.
- Harnos, K., M. L'Heureux, Q. Ding, and Q. Zhang, 2019: Skill of seasonal Arctic sea ice extent predictions using the North American Multimodel Ensemble. *J. Climate*, **32**, 623–638, <https://doi.org/10.1175/JCLI-D-17-0766.1>.
- Holland, M., C. Bitz, and B. Tremblay, 2006: Future abrupt reductions in the summer Arctic sea ice. *Cryosphere*, **33**, GL028024, <https://doi.org/10.1029/2006GL028024>.
- Knowles, K. W., 1993: Points, pixels, grids, and cells: A mapping and gridding primer. National Snow and Ice Data Center, https://www.cpc.ncep.noaa.gov/products/predictions/short_range/tools/ppgc.html.
- Liu, Y., W. Wang, and A. Kumar, 2018: Multiweek prediction skill assessment of Arctic sea ice variability in the CFSv2. *Wea. Forecasting*, **33**, 1453–1476, <https://doi.org/10.1175/WAF-D-18-0046.1>.
- Meier, W., and J. Stewart, 2019: Assessing uncertainties in sea ice extent climate indicators. *Environ. Res. Lett.*, **14**, 035005, <https://doi.org/10.1088/1748-9326/aaf52c>.
- Mohammadi-Aragh, M., H. F. Goessling, M. Losch, N. Hutter, and T. Jung, 2018: Predictability of Arctic sea ice on weather time scales. *Sci. Rep.*, **8**, 6514, <https://doi.org/10.1038/s41598-018-24660-0>.
- Olonscheck, D., T. Mauritsen, and D. Notz, 2019: Arctic sea-ice variability is primarily driven by atmospheric temperature fluctuations. *Nat. Geosci.*, **12**, 430–434, <https://doi.org/10.1038/s41561-019-0363-1>.
- Parkinson, C., and J. Comiso, 2013: On the 2012 record low Arctic sea ice cover: Combined impact of preconditioning and an August storm. *Geophys. Res. Lett.*, **40**, 1356–1361, <https://doi.org/10.1002/grl.50349>.
- , D. Cavalieri, P. Gloersen, H. Zwally, and J. Comiso, 1999: Arctic sea ice extents, areas, and trends, 1978–1996. *J. Geophys. Res.*, **104**, 20 837–20 856, <https://doi.org/10.1029/1999JC900082>.
- Schreiber, E., and M. Serreze, 2020: Impacts of synoptic-scale cyclones on Arctic sea-ice concentration: A systematic analysis. *Ann. Glaciol.*, **61**, 139–153, <https://doi.org/10.1017/aog.2020.23>.
- Stern, D., J. Doyle, N. Barton, P. Finocchio, W. Komaromi, and E. Metzger, 2020: The impact of an intense cyclone on short-term sea ice loss in a fully-coupled atmosphere–ocean–ice model. *Geophys. Res. Lett.*, **47**, L085580, <https://doi.org/10.1029/2019GL085580>.
- Stroeve, J., L. Hamilton, C. Bitz, and E. Blanchard-Wrigglesworth, 2014: Predicting September sea ice: Ensemble skill of the SEARCH Sea Ice Outlook 2008–2013. *Geophys. Res. Lett.*, **41**, 2411–2418, <https://doi.org/10.1002/2014GL059388>.
- Sun, S., B. Green, R. Bleck, and S. Benjamin, 2018: Subseasonal forecasting with an icosahedral, vertically quasi-Lagrangian coupled model. Part II: Probabilistic and deterministic forecast skill. *Mon. Wea. Rev.*, **146**, 1619–1639, <https://doi.org/10.1175/MWR-D-18-0007.1>.
- Vitart, F., 2004: Monthly forecasting at ECMWF. *Mon. Wea. Rev.*, **132**, 2761–2779, <https://doi.org/10.1175/MWR2826.1>.
- , and Coauthors, 2017: The Subseasonal to Seasonal (S2S) prediction project database. *Bull. Amer. Meteor. Soc.*, **98**, 163–173, <https://doi.org/10.1175/BAMS-D-16-0017.1>.
- Wang, Z., J. Walsh, S. Szymborski, and M. Peng, 2020: Rapid Arctic sea ice loss on the synoptic time scale and related atmospheric circulation anomalies. *J. Climate*, **33**, 1597–1617, <https://doi.org/10.1175/JCLI-D-19-0528.1>.
- Wayand, N., C. Bitz, and E. Blanchard-Wrigglesworth, 2019: A year-round subseasonal-to-seasonal sea ice prediction portal. *Geophys. Res. Lett.*, **46**, 3298–3307, <https://doi.org/10.1029/2018GL081565>.
- Wernli, H., and L. Papritz, 2018: Role of polar anticyclones and mid-latitude cyclones for Arctic summertime sea-ice melting. *Nat. Geosci.*, **11**, 108–113, <https://doi.org/10.1038/s41561-017-0041-0>.
- Zampieri, L., H. Goessling, and T. Jung, 2018: Bright prospects for Arctic sea ice prediction on subseasonal time scales. *Geophys. Res. Lett.*, **45**, 9731–9738, <https://doi.org/10.1029/2018GL079394>.
- Zhang, J., R. Lindsay, A. Schweiger, and M. Steele, 2013: The impact of an intense summer cyclone on 2012 Arctic sea ice retreat. *Geophys. Res. Lett.*, **40**, 720–726, <https://doi.org/10.1002/grl.50190>.
- Zhu, H., M. Wheeler, A. Sobel, and D. Hudson, 2014: Seamless precipitation prediction skill in the tropics and extratropics from a global model. *Mon. Wea. Rev.*, **142**, 1556–1569, <https://doi.org/10.1175/MWR-D-13-00222.1>.