Multiweek Prediction Skill Assessment of Arctic Sea Ice Variability in the CFSv2

YANYUN LIU
INNOVIM, LLC, Greenbelt, and NOAA/NWS/NCEP/Climate Prediction Center, College Park, Maryland

WANQIU WANG AND ARUN KUMAR
NOAA/NWS/NCEP/Climate Prediction Center, College Park, Maryland

(Manuscript received 19 March 2018, in final form 21 August 2018)

ABSTRACT

Skillful Arctic Sea ice prediction is becoming increasingly important because of its societal, industrial, and economic impacts over the polar regions and potential influence on lower-latitude weather and climate variability. In this work, we evaluate the multiweek forecast skill of Arctic sea ice using the Climate Forecast System, version 2 (CFSv2). To the authors’ knowledge, this is the first effort to diagnose and assess the skill of multiweek Arctic sea ice prediction from a coupled atmosphere–ocean model. Analysis of a suite of retrospective 45-day forecasts spanning 1999–2015 shows that CFSv2 captures general features of sea ice concentration (SIC) variability. Total SIC variability is dominated by interannual variability, which accounts for more than 60% of the total variance. Submonthly variability accounts for 29% of the total variance in December, 20% in March and June, and 12.5% in September. We assess the ability of CFSv2 to predict the pan-Arctic SIC, as well as regional SIC in nine Arctic regions. Results show that the SIC prediction skill is highly region dependent (e.g., higher prediction skill for Kara/Barents Seas and lower for the Canadian Archipelago). Overall, the maximum anomaly correlation coefficient (ACC) of SIC for both melt and freeze-up seasons is near the marginal zones, and their spatial distribution shows a relationship with the distribution of the variance. If the ACC of 0.5 is taken as the critical value for skillful prediction, the predictability of weekly SIC near the marginal zones is about 5–6 weeks. Prediction skill for Arctic sea ice extent is above 0.6 for the entire six target weeks and has a large contribution from interannual variability.

1. Introduction

Sea ice plays an important role in the global climate system through its impact on the water and energy budgets (McBean et al. 2005; Budikova 2009; Guemas et al. 2016). Sea ice reflects most of the incident solar radiation through high albedo, thus reducing the amount of energy absorbed by the surface. It also modulates the exchange of moisture and heat fluxes between the ocean and the atmosphere. The Arctic sea ice has changed dramatically over the past few decades and is expected to continue to decline throughout the twenty-first century (Stroeve et al. 2012; Serreze et al. 2007). The recent decline of Arctic sea ice has been attributed to both natural and anthropogenic causes (Overland et al. 2008), the latter being associated with increased levels of greenhouse gases. The decline of Arctic sea ice has important societal and economic impacts (e.g., more commercial opportunities, such as opening of new shipping routes and Arctic fishing).

Previous studies have shown that sea ice variability is largely affected by the atmospheric circulation (Fang and Wallace 1994; Deser et al. 2000; Serreze et al. 2003; Ding et al. 2017). For example, the Arctic Oscillation (AO) can affect Arctic sea ice variability (Rigor et al. 2002; Mysak and Venegas 1998; Serreze et al. 2003; Rigor and Wallace 2004). Other low-frequency large-scale climate variations (e.g., Arctic dipole, ENSO, PNA) also have an impact on Arctic sea ice (Wu et al. 2006; Wang and Overland 2009; L’Heureux et al. 2008). On the other hand, Arctic sea ice variability can also affect the Arctic weather, ocean, and climate (Serreze et al. 2007). A number of studies suggest that Arctic sea ice decline can impact high-latitude temperature amplification (Screen and Simmonds 2010; Kumar et al. 2010; Colloff et al. 2018) and midlatitude cooling (Mori et al. 2014; Screen et al. 2015; Kug et al. 2015), while some other studies indicated that no evidence is found regarding the link between Arctic sea ice loss and midlatitude cooling.

Corresponding author: Yanyun Liu, yanyun.liu@noaa.gov

DOI: 10.1175/WAF-D-18-0046.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
extremes (McCusker et al. 2016); thus, a controversy about the Arctic–midlatitude weather linkage still exists (Collow et al. 2018). Arctic sea ice decline can also regulate ocean circulation [e.g., weaken the Atlantic meridional overturning circulation (AMOC) and its poleward heat transport (Sévelle et al. 2017)]. These findings suggest that sea ice is a covarying component in the coupled atmosphere–ocean system. Therefore, understanding the causes for the Arctic sea ice prediction skill and improving the skill of sea ice predictions has become increasingly important to the advance of weather and climate services.

Previous studies mainly focused on the seasonal forecast of Arctic sea ice concentration (SIC) or sea ice extent (SIE), either using statistical models (Drobot 2007; Lindsay et al. 2008; Yuan et al. 2016; Kapsch et al. 2013, Petty et al. 2017) or using coupled dynamical models (Wang et al. 2013; Zhang et al. 2013, 2016; Chevallier et al. 2013; Merryfield et al. 2013; Blanchard-Wrigglesworth et al. 2016; Tietsche et al. 2014; Bushuk et al. 2017). Using a linear statistical model, Lindsay et al. (2008) found that, except for the trend, much of the predictive information in the ice–ocean system is lost after lead times of 3 months. Using fully coupled dynamical models, Wang et al. (2013), Sigmond et al. (2013), and Chevallier et al. (2013) showed that prediction skill for detrended regional SIE generally exceeds that of an anomaly persistence forecast and the sea ice variability could be predicted up to 2–4 months ahead, except for January/February when forecast skill is moderately high up to an 11-month lead time (Sigmond et al. 2013). Besides the seasonal forecast, there are also several studies that deal with shorter-term forecasts (Posey et al. 2015; Hebert et al. 2015; Smith et al. 2016; Lemieux et al. 2016; Schweiger and Zhang 2015). For example, the U.S. Navy operational Arctic sea ice forecast system, designed to provide shorter-term, 1–7-day forecasts of Arctic sea ice and ocean conditions, was also shown to be skillful compared to the climatology for forecasts up to 102 h (Hebert et al. 2015). Schweiger and Zhang (2015) suggested that Arctic sea ice drift forecasts during the summer of 2004 remain skillful for 8 days using the Marginal Ice Zone Modeling and Assimilation System.

Recently, the National Oceanic and Atmospheric Administration has initiated activities to improve skill of forecasts in the weeks 3–4 time range to extend weather forecast capability beyond the conventional range of 10–15 days. While Arctic sea ice forecasts at seasonal time scales have received considerable attention, very limited work has been done on shorter time scales. Wang et al. (2016) show that summertime Arctic sea ice can be predicted with reasonable skill at lead times of 20–60 days using a data-driven statistical model. With the ongoing advancement in the development of coupled atmosphere–ocean forecast systems during the past years, it is of great interest to investigate the capability of dynamical forecast systems in predicting sea ice at weekly to monthly time scales.

In this study, a fully coupled atmosphere–ocean model is used to evaluate the forecast skill of weekly mean sea ice from week 1 to week 6 for Arctic regions using NCEP’s Climate Forecast System, version 2 (CFSv2). To the authors’ knowledge, this is the first effort to diagnose and assess multiweek Arctic sea ice prediction skill from a coupled atmosphere–ocean model, and the following aspects are analyzed: 1) the overall prediction skill of SIC and SIE in the CFSv2, 2) the contribution to variations and the predictability of weekly sea ice anomalies from long-term trends as well as detrended interannual and submonthly variations, and 3) regional variations in prediction skill, which have a greater practical use than the overall Arctic SIE to stakeholders (Bushuk et al. 2017). Section 2 describes the forecast data from the CFSv2 and observations used for verification. The observed Arctic variability and the contribution of different components (including the trend, interannual, and submonthly variability) to the observed total sea ice variability, and the assessment of the forecast performance of multiweek Arctic sea ice, as well as the regional sea ice prediction skill, are discussed in section 3. Section 3 also presents the Arctic SIE variability and prediction skill in CFSv2 and is followed with a summary and discussion in section 4.

2. The forecast model, data, and processing method

a. Forecast model data and observations

The retrospective forecast dataset analyzed in this study is from the fully coupled CFSv2 model (Saha et al. 2014). This forecast model consists of the NCEP atmospheric Global Forecast System (GFS) at a horizontal resolution of T126 (~100 km) with 64 vertical levels and the Geophysical Fluid Dynamics Laboratory (GFDL) Modular Ocean Model, version 4 (MOM4), with 40 vertical levels. The ocean component includes a thermodynamic and dynamic sea ice model from the GFDL Sea Ice Simulator (Griffies et al. 2004).

The sea ice model in the CFSv2 has two layers for sea ice and one layer for snow. In each ice grid there are five categories of possible sea ice thicknesses: 0–0.1, 0.1–0.3, 0.3–0.7, 0.7–1.1, and >1.1 m. The sea ice dynamics is based on Hunke and Dukowicz (1997) using the
elastic–viscous–plastic technique to calculate ice internal stress, while the sea ice thermodynamics is based on Winton (2000). The sea ice concentration forecast data from the CFSv2 with a horizontal resolution of 0.5° × 0.5° is used.

Retrospective and real-time forecasts initialized from the NCEP Climate Forecast System Reanalysis (CFSR; Saha et al. 2010) for 1999–2015 are used for this study. The CFSR is a weakly coupled assimilation system using a coupled atmosphere–land–ocean–sea ice model. The CFSR assimilates conventional atmospheric observations, satellite radiances, daily Optimum Interpolation Sea Surface Temperature (or daily OISST) of Reynolds et al. (2007), subsurface temperature data from expendable bathythermographs (XBTs), the Tropical Atmosphere–Ocean (TAO) array of moored buoys, and Argo and Argo-like floats, as well as soil moisture from a parallel land assimilation system. For sea ice, only the sea ice concentration from satellite observations is assimilated. Sea ice thickness and its horizontal displacement are determined by the thermodynamic and dynamic balance between the sea ice and the overlying atmosphere and underlying ocean water. The input satellite sea ice concentration

![Image](attachment:image.png)

**Fig. 1.** Spatial pattern of total weekly anomaly standard deviation of Arctic SIC for four target months of (a) March, (b) June, (c) September, and (d) December during 2000–15, obtained from NSIDC observations. The black lines represent the 15% concentration contour of the 2000–15 average in the NSIDC observations, which is usually used to define the sea ice edge.
observations are from different sea ice analyses. From 1 January 1999 to 29 February 2000, the global ice concentration analysis was the NCEP operational ice analysis. From 1 March 2000 to 29 October 2007 the sea ice analysis is the newer NCEP sea ice analysis system (GRUa) that is applied to archived passive microwave data for DMSP F13, F14, and F15. The newer system is based on the NASA Team 2 algorithm (Markus and Cavalieri 2000). From 30 October 2007 to the present, the operational NCEP passive microwave sea ice concentration analysis was used. More details on the assimilation of sea ice concentration can be found in Saha et al. (2010). Raw forecast data include output from four 45-day forecast runs from 0000, 0600, 1200, and 1800 UTC each day. Retrospective forecasts from 29 February in leap years are not available and are not considered.

For the assessment of the CFSv2 performance, data from the National Snow and Ice Data Center (NSIDC) (Cavalieri et al. 1996; Fetterer et al. 2002) during 2000–15 are used. As suggested by Notz et al. (2013), care must be taken when comparing model results to the “observed” SIC based on just a single algorithm. Thus, we used the average of SIC from the NASA team and bootstrap algorithms as the verification data for forecast assessment to reduce observational uncertainty (Notz et al. 2013). For example, the correlation between the daily CFSR SIE anomalies and the NSIDC SIE
The CFSR by its design assimilates different sea ice concentration datasets. These sea ice datasets are not necessarily the best observations and contain different levels of observational uncertainty and bias. In evaluating the forecast performance, we would like to verify the forecast against independent observations for a realistic assessment of the forecast capability. Evaluating the forecast against the input

**TABLE 1.** Total standard deviation and its components (including trend, detrended interannual, and submonthly variability) for four target months (March, June, September, and December) obtained from observational data. The values are area-weighted averages over the Arctic regions (50°–90°N). The values in parentheses are the percentage contribution of the trend, detrended interannual, and submonthly variance to the total variance.

<table>
<thead>
<tr>
<th>NSIDC SIC standard deviation (percentage variance)</th>
<th>March (%)</th>
<th>June (%)</th>
<th>September (%)</th>
<th>December (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>0.10</td>
<td>0.12</td>
<td>0.15</td>
<td>0.09</td>
</tr>
<tr>
<td>Trend interannual</td>
<td>0.04 (13.6%)</td>
<td>0.04 (10.5%)</td>
<td>0.07 (19.3%)</td>
<td>0.03 (9.3%)</td>
</tr>
<tr>
<td>Detrended interannual</td>
<td>0.08 (64.1%)</td>
<td>0.10 (68.9%)</td>
<td>0.13 (68.2%)</td>
<td>0.07 (61.3%)</td>
</tr>
<tr>
<td>Submonthly</td>
<td>0.05 (22.3%)</td>
<td>0.05 (20.6%)</td>
<td>0.05 (12.5%)</td>
<td>0.05 (29.4%)</td>
</tr>
</tbody>
</table>

**Fig. 3.** As in Fig. 1, but for submonthly SIC anomalies.
data for CFSR might inflate the skill artificially. The mixture of NASA team and bootstrap algorithms provides a better representation of the observation than the individual input sea ice data used in CFSR and allows a more realistic skill measure of the forecast system. The NASA observational dataset has a region close to the pole that cannot be observed owing to the orbit inclination of the satellites. This region, known as the “polar hole,” is masked out from the NSIDC dataset and CFSv2 model output for a consistent evaluation. We also note that the skill assessment may be influenced by differences in the data sources used for forecast initialization and for verification.

b. Processing method

This study focuses on the analysis of weekly mean anomalies from CFSv2 for 0-5-week leads. For each forecast starting day, the following steps are taken to produce the weekly average of the ensemble mean forecast: 1) forecast runs from the latest 3 days were used to form a lagged ensemble of 12 runs; 2) a daily average of the 12-run ensemble mean is computed for 42 target days; 3) a nonoverlapping 7-day weekly average is calculated from the daily ensemble mean for 0-week lead (day 1-7 average) to a 5-week lead (day 36–42 average). For example, the forecast starting on 1 January produces the weekly mean
from 1–7 January (0-week lead) to 5–11 February (5-week lead).

The forecast data of the ensemble-mean weekly average are then rearranged according to lead time and target (or verification) week. For the description of the processing method, we use $F(d, L, yr)$ to represent the forecast $F$ for the weekly mean of the target week beginning on day $d$ (1–365) in year $yr$ (2000–15) at lead time $L$ (0–5 week). Notice that $d$ represents a 7-day target week rather than a day, with $d = 1$ for the average of 1–7 January and $d = 2$ for the average of 2–8 January. For a given target week, say, $d = 1$, $L = 0$ is the 0-week lead forecast from 1 January, $L = 1$ is the 1-week lead forecast from 25 December, $L = 2$ is the 2-week lead forecast from 18 December, and so on. For simplicity, we will use L0–L5 to individuate 0–5-week-lead forecasts. To facilitate the calculation, and for a consistent comparison, the corresponding observational data are rearranged following the CFSv2 forecast structure to form the weekly average of SIC for each target verification week starting from each calendar day between 2000 and 2015 both for NASA team and bootstrap algorithms. The average of SIC using the NASA team and bootstrap algorithms is then calculated and used for verification. The verifying observational data are arranged to follow the forecast

![Fig. 5. ACC between the CFSv2 forecast and observations for total SIC anomalies at L3 for four target months of (a) March, (b) June, (c) September, and (d) December. The black contours indicate the climatology of the observed 15% concentration, and the green contours indicate the climatology of 15% concentration for the CFSv2 forecast at L3.](image-url)
data structure and are denoted as $O(d, \text{yr})$ where $O$ represents the observations.

c. Climatology and anomalies

For both forecasts and observations, a 16-yr (2000–15) average of 7-day mean SIC is calculated for each starting date from 1 January to 31 December ($d = 1–365$); for a forecast, a calculation is performed for each lead time (L0–L5). The climatology for each starting target date $d$ is defined as the 31-day (from $d - 15$ to $d + 15$) running mean of the resulting 16-yr average and is denoted as $F_c(d, L)$ for the forecast and $O_c(d)$ for the observation. As suggested by previous studies (Lindsay et al. 2008; Wang et al. 2013; Chevallier et al. 2013), a large source of sea ice prediction skill is derived from the linear trend. The forecast skill after removing the linear trend is smaller and generally not statistically significant at lead times greater than 2–3 months, except for January/February when the forecast skill is moderately high up to an 11-month lead time (Sigmond et al. 2013). Further, the variability at different time scales may indicate different sources of the predictability in the coupled system components. Therefore, in this work, we divided the total anomaly into the trend, detrended interannual, and submonthly anomalies.

Three types of anomalies are calculated: total anomaly, interannual anomaly, and submonthly anomaly. The total anomaly is calculated as the deviation of the total weekly mean values of SIC from the climatology.
The interannual anomaly is defined as the monthly mean of the total anomaly. The submonthly anomaly is the deviation of the total anomaly from the interannual anomaly (i.e., the monthly mean of the total anomaly). The interannual anomaly is further divided into two components: the trend interannual anomaly and the detrended interannual anomaly. Decomposing the total weekly anomaly into different time scales may help us to understand the sources of the predictability of the weekly mean anomaly. For example, the long-term trend can be attributed to the contribution of anthropogenic forcings, the detrended interannual anomaly may be associated with the low-frequency variability in the ocean and sea ice, while the atmospheric variability may contribute to the high-frequency submonthly anomaly. More detailed descriptions of these anomalies calculations are included in the appendix.

There are alternative ways to define the interannual component, for example, using a 3-month seasonal mean or a low-pass filtering approach. Given the strong month-to-month seasonality of the sea ice anomalies, we think the monthly mean is more appropriate than the seasonal mean to represent the seasonal evolution of interannual variability. One advantage of the decomposition of the total anomaly procedure we adopted here is that the three components (trend interannual, detrended interannual, and submonthly anomalies) are independent and their variances are additive; that is, the sum of the variances of the individual components
equals the variance of the total anomaly. This property would not be satisfied if the interannual component is defined with low-pass filtering.

3. Results

In this section we first examine the observed Arctic variability using the NSIDC SIC. The prediction skill of multiweek pan-Arctic SIC, as well as regional SIC in the CFSv2 forecasts, is then assessed, followed by an analysis of the SIE variability and prediction skill.

a. Observed SIC variability

Figure 1 shows the spatial pattern of the total weekly anomaly standard deviation of NSIDC Arctic SIC anomaly for four target months (March, June, September, and December) during 2000–15. The black curves in Fig. 1 represent the average sea ice edge taken as a 15% contour of the climatological SIC during 2000–15. The climatological sea ice edge latitude based on the 15% contour of the average SIC may be different from the average sea ice edge latitude of individual years. Large variability is generally located near the sea ice edge, as indicated by the 15% SIC black contours. The Arctic SIC shows relatively high variability in different Arctic regions during different months: the Bering Sea and the Sea of Okhotsk during March; the Chukchi, Kara, and Barents Seas during June; the Beaufort, East Siberian, and Laptev Seas during September; and the Kara Sea during December. The variability patterns of the
detrended interannual and submonthly anomalies are similar to that of the total anomalies (Figs. 2 and 3), with the standard deviation of the submonthly anomalies having a smaller amplitude.

For a quantitative comparison, Table 1 lists the contributions of different components, including the trend interannual, detrended interannual, and submonthly variability, to the observed total SIC standard deviation averaged over Arctic regions ($50^\circ$-$90^\circ$N) for four target months (March, June, September, and December). The percentage variance contributions are given in parentheses. As shown in Table 1, the Arctic SIC shows the largest (smallest) variability during September (December).

The total SIC variability is dominated by detrended interannual variability, which accounts for more than 60% of the total variance. The contribution of submonthly anomalies accounts for about 20% of the total variance in March and June, 12.5% in September, and 29.4% in December. The submonthly variance is 2–3 times as strong as the trend variance for March, June, and December, but not for September, when the submonthly variance is weaker than the trend variance.

b. SIC climatology

The prediction performance in representing Arctic SIC climatology in CFSv2 is assessed in this section.

![Fig. 9. RMSEs between the CFSv2 forecast and observations for total SIC anomalies at L3 for four target months of (a) March, (b) June, (c) September, and (d) December. The black contours indicate the observed 15% concentration, and the orchid contours indicate the 15% concentration for the CFSv2 forecast at L3.](image-url)
We first compare the 16-yr SIC climatology between the forecast and observations. Figure 4 shows the mean SIC biases of the CFSv2 for the 3-week-lead forecast (L3) for four target months (March, June, September, and December). The CFSv2 captures the general features of the seasonal SIC evolution, including the sea ice retreat in summer and fall and its advancement in winter and spring. However, the forecasts also contain systematic biases. The forecast bias is largest in the marginal ice zone region, which is defined as the transition zone between the open ocean and sea ice. There is a positive bias around the Labrador, Barents, and Greenland Seas and a negative bias in the Bering Sea and the Sea of Okhotsk in March. In June, a positive bias is seen in the Kara/Barents Seas, southern Chukchi Sea, Baffin Bay, and Davis Straits. Systematic errors are small in September with a positive bias in the Chukchi Sea and East Siberian Sea. In December, there is a negative bias in Hudson Bay and the Bering Sea.

c. SIC prediction skill

Forecast skill of the sea ice variability is evaluated based on the anomaly correlation coefficient (ACC) and root-mean-square error (RMSE) between forecasts and observations. Spatial maps of the skill of the week 4 (i.e., L3) forecast are first analyzed. Skill as a function of lead time is then diagnosed for individual regions. We focus on the skill for the total anomaly, detrended interannual anomaly, and submonthly anomaly.
Figures 5 and 6 show the spatial ACC of the total and detrended interannual SIC anomalies between the CFSv2 week-4 forecast and observations. CFSv2 shows good skill at L3. The Arctic Ocean shows the highest prediction skill during September, particularly in the regions of the Chukchi Sea, Beaufort Sea, and East Siberian Sea, where the ACC is above 0.6. The ACC for the detrended interannual anomalies is generally higher than that for the total anomalies (Figs. 5 and 6) and much higher than the prediction skill of the submonthly variability (Fig. 7). This suggests the detrended interannual variability is more predictable than the submonthly variability.

For comparison, the ACC prediction skill of the damped persistence forecast is given in Fig. 8. The damped persistence forecast for each target week is taken as the observed weekly mean anomaly of the initial week (L − 1) immediately preceding the start of a forecast scaled by a lagged correlation between the initial and target weeks that has a seasonal and spatial dependence. For example, the damped persistence forecast for target week L3 is the observed anomaly of L − 1 multiplied by the lagged anomaly correlation coefficient between L − 1 and L3. The detailed equation and discussions are included in the appendix. NSIDC data are used as observations for the damped persistence forecast. Compared with CFSv2, the skill of the damped persistence forecast is generally lower than the CFSv2 forecast skill (Figs. 8 and 5).
The forecast performance of CFSv2 sea ice is also assessed using the RMSE. The RMSE increases with lead time for the total and each component. Figures 9–11 show the spatial RMSEs of the total, detrended interannual, and submonthly SIC anomalies between the CFSv2 forecast and observations at L3 for four target months. During March, there are larger RMSEs in the Bering Sea, the Sea of Okhotsk/Japan, and the marginal ice zone. In June and December, the RMSEs are larger in the Kara Sea and Hudson Bay. In September, the RMSE is larger particularly in the Chukchi, Beaufort, and East Siberian Seas. The RMSEs of both the detrended interannual and submonthly anomalies (Figs. 10 and 11) show similar spatial patterns to that of the total anomalies (Fig. 9) but with smaller amplitude. The RMSE of the submonthly anomalies is smaller than that of the detrended interannual anomalies in March and June and is comparable to that of the detrended interannual anomalies in September and December (Figs. 10 and 11).

The utility of a forecast greatly depends on the geographic characteristics of the forecast skill. Regional SIC forecast skill is especially important to the stakeholders for their decision-making. To further assess the multiweek prediction skill of regional SIC, we calculate the ACC and RMSE in the pan-Arctic Ocean as well as the nine Arctic regions (shown in Fig. 12). The spatial average of temporal ACC of the total anomalies as a function of lead time for the pan-Arctic region and the nine Arctic regions is shown in Fig. 13. It is noted that no skill is shown for months where there is no sea ice (e.g., in the Gulf of St Lawrence during June and September). The ACC values are calculated at every grid point and then averaged across the nine Arctic regions. The ACC decreases with lead time and is region and season dependent. The ACC is relatively high for the Kara and Barents Seas for all seasons. During March, the ACC is higher in the Bering Sea, Sea of Okhotsk/Japan, and Gulf of St. Lawrence. During June, the ACC is higher in the Greenland Sea and Hudson Bay. During September, the ACC is higher in the Greenland Sea, Arctic Ocean, and Canadian Archipelago. During December, higher prediction skill is found in the Bering Sea, Hudson Bay and Sea of Okhotsk/Japan. To further assess the skill for different time scales, prediction skills for the total, detrended interannual, and submonthly anomalies are shown in Fig. 14 for individual regions. Each region is plotted for the month of the largest variance in the observations. Generally, the skill of the detrended interannual anomalies is higher than that of the total because of the lower skill of the submonthly anomalies.

The spatial average of RMSE for pan-Arctic region and the nine Arctic regions is shown in Fig. 15. The RMSE increases with lead time and the rate of increase varies with region and season. During March, the RMSE is larger in the Sea of Okhotsk/Japan and Bering Sea. The RMSE is larger in the Kara/Barents Seas and Hudson Bay during June. During September,
the RMSE is larger in the Arctic Ocean and Canadian Archipelago. During December, the RMSE is higher in Hudson Bay and the Kara/Barents Seas.

Spatial and seasonal variations of the forecast ACC skill show that the sea ice at a specific location is most predictable when this location is near the marginal zone having large sea ice variance (Figs. 1 and 5). This means that for each location there exist months for which the prediction skill is higher than the adjacent months. The skill in the month of maximum ACC (MMA) represents the best performance of the forecast system. Given that the physical processes that contribute to the predictability are different between the melt and freeze-up seasons, there may be two MMAs at each location, one during the melt season and the other during the freeze-up season.

The distribution of the maximum ACC skill among the months of the melt season (April–September) at L3 is plotted in Fig. 16a and corresponding months when the maximum ACC skill is realized are shown in Fig. 16b with values 4–9 corresponding to months of April–September. Overall, the maximum ACC during the melt season progressively moves from south to north.
north (Fig. 16b) with the mean sea ice concentration in the MMA increasing with latitude (Fig. 16c). The maximum ACC varies mostly from 0.5 to 0.8 (Fig. 16a), and its spatial distribution shows a relationship with the distribution of the variance (Fig. 16d); that is, areas of larger variance correspond to larger values of ACC. For example, the areas of relatively larger variance surrounding the central Arctic match the relatively larger values of the maximum ACC. This is consistent with the finding of Van den Dool and Toth (1991) and Kumar (2009) that high signal-to-noise regimes lead to high expected values of skill. Similar features of statistics of maximum ACC for freeze-up season (October–March) can be seen in Fig. 17, except that there is a north-to-south seasonal progression. On average over the entire Arctic region (and for each lead time), the melt season prediction is more skillful for the first 3 weeks than the freeze-up season (Fig. 18). Taking 0.5 as the critical maximum ACC for skillful prediction, the predictability of the weekly mean sea ice concentration is about 5–6 weeks (Fig. 18).
**d. SIE prediction skill**

SIE is defined as the sum of ice-covered areas with concentrations of at least 15%. The SIE provides an overall measure of the pan-Arctic sea ice extent and has been used as the key parameter in the seasonal outlook of the Sea Ice Prediction Network (Stroeve et al. 2014). The SIE variability and its prediction at lead times of 0–5 weeks are examined in this section. Figures 19a–c show the observed total, detrended interannual, and submonthly SIE anomalies and the forecast from CFSv2 at L1, L3, and L5. The SIE forecasts match well with the observed SIE anomalies, including the extreme negative anomalies during September 2007 and 2012. These extreme anomalies are primarily associated with interannual variations and are well captured in CFSv2 at lead times of multiple weeks (Fig. 19b). Figure 19c shows the observed submonthly variability and its prediction for 2012–13 at L1, L3, and L5. It is clear that CFSv2 is capable of reproducing the observed submonthly variability to some extent (Fig. 19c). Figures 19d–f show the scatterplots of the observed total, detrended interannual, and submonthly SIE anomalies versus the CSFv2 forecast at L3. The scatterplots also show that the SIE forecasts at 3-week lead are in good agreement with the observed SIE anomalies.

**Fig. 15.** Spatial average of RMSEs between the CFSv2 SIC and observations from L0 to L5 for four target months. The dashed lines are the observed standard deviations for the four target months.
especially for the total and detrended interannual anomalies. A quantitative assessment of the SIE forecast skill is shown in Fig. 20, which includes the ACCs and RMSEs of the total, detrended interannual, submonthly anomalies for CFSv2 forecasts from L0 to L5. The ACC results of the total and detrended interannual SIE anomalies are comparable, while the ACC of the submonthly anomalies is lower especially after L3 (Fig. 20a), indicating that the skill of the multiweek SIE forecast is strongly affected by interannual variability. Variations in RMSE show that forecast errors in total SIE come from both interannual and submonthly errors with detrended interannual errors being the dominant factor (Fig. 20b).

e. Possible sources for the sea ice prediction skill

Possible sources of Arctic sea ice prediction skill come from the prediction of the atmospheric variability, the impacts of external forcing, and the evolution of the upper-ocean and sea ice conditions. For example, the Arctic Oscillation, Arctic dipole, and PNA can affect the Arctic sea ice prediction (Rigor et al. 2002; Wang and Overland 2009). Henderson et al. (2014) suggested that atmospheric intraseasonal variability...
[e.g., the Madden–Julian oscillation (MJO)] can modulate Arctic sea ice in both summer and winter. The external forcing includes the anthropogenic variability, responsible for the sea ice decline in recent decades. For subseasonal prediction, the upper-ocean/sea ice initial conditions are especially important. Bushuk et al. (2017) suggested that the initial conditions of SST and sea ice can provide a crucial source of predictability for both winter and summer regional SIE. Previous modeling studies suggest the anomalies of sea ice thickness have persistency at seasonal-to-interannual time scales and are as important as SST in controlling the persistence characteristics of the sea ice cover (Bitz et al. 1996; Chevallier and Salas-Méria 2012; Blanchard-Wrigglesworth and Bitz 2014). Our analysis indicates the detrended interannual skill in CFSv2 is much higher than the submonthly skill, showing that most skill is coming from physical mechanisms acting on seasonal-to-interannual time scales. These suggest factors that have impacts on seasonal predictions, such as sea ice thickness and oceanic conditions, are also major sources of subseasonal predictability, while atmospheric variations, such as tropical intraseasonal oscillations, have relatively smaller contributions to subseasonal predictability. Further studies are required to determine what physical processes are responsible for individual contributions of the prediction.

Fig. 17. As in Fig. 16, but for the freeze-up season (from October to March).
skill and how the different sources of predictability interact.

Only the sea ice concentration from satellite observations is assimilated in the current initialization system. The lack of real-time sea ice thickness observations needs to be resolved in data assimilation systems to further enhance the skill of the subseasonal and seasonal sea ice predictions. In addition to the need for higher quality observations, improvements of subseasonal predictions also depend on reducing model biases and better understanding the detailed physical mechanisms and their representation in the forecast model.

4. Summary and discussion

This study examined Arctic sea ice weekly mean variability and multiweek prediction skill in the CFSv2. While many studies on dynamical seasonal sea ice predictions have been published, this study, to the authors’ knowledge, appears to be the first effort to analyze sea ice prediction skill at multiweek lead times using a dynamical forecast system. Analysis of weekly mean NSIDC SIC indicates that there is relatively high variability in some Arctic regions: the Bering Sea and Sea of Okhotsk during March; the Chukchi, Kara, and Barents Seas during June; the Beaufort, East Siberian, and Laptev Seas during September; and the Kara Sea during December. Seasonally, the largest Arctic SIC variability occurs in September. Comparison among different components (trend interannual, detrended interannual, and submonthly variability) shows that detrended interannual variability is dominant in the weekly mean sea ice variability, compared to the trend and submonthly variabilities. On average, the detrended interannual variability accounts for more than 60% of the total variance. The contributions of submonthly variability account for about 20% of the total variance in March and June, 12.5% in September, and 29.4% in December.

Multiweek sea ice prediction skill for the pan-Arctic region, as well as individual Arctic regions, is analyzed. CFSv2 captures general features of the Arctic sea ice variability, with the largest bias in marginal ice zones. The results demonstrate the multiweek prediction skill is dominated by the prediction of interannual variability although predictions of both interannual and submonthly anomalies contribute to the prediction skill. The RMSEs of the total SIC anomalies mainly come from interannual anomalies with the submonthly variability playing a smaller role. Regional Arctic sea ice analysis shows that the regional prediction skill of SIC in CFSv2 is highly region and season dependent. Prediction skill is generally higher in the Kara/Barents Sea and lower for the Canadian Archipelago. Overall, the maximum ACCs for both the melt and freeze-up seasons are near the marginal zones and their spatial distribution shows a relationship with the distribution of the variance; that is, areas of larger variance correspond to larger values of ACC. Taking 0.5 as a useful level of ACC for skillful prediction, the predictability of weekly mean sea ice concentration near marginal zones is about 5–6 weeks. The prediction skill for Northern Hemisphere SIE is above 0.6 for the entire six target weeks and receives a large contribution from detrended interannual variability.

Wang et al. (2016) show that summertime Arctic sea ice can be predicted with reasonable skill at lead times of 20–60 days in a statistical model using 1979–2012 satellite data. The summertime SIC prediction skill of Wang et al.’s (2016) statistical methodology exceeds that of the damped persistence forecast for lead times longer than 20 days. Our results also indicated that the CFSv2 SIC prediction skill is generally higher than the skill of the damped persistence forecast. Wang et al. (2016) showed relatively high forecast skill levels for two selected marginal seas (Kara and Barents), which is consistent with the results of our analysis of the CFSv2 forecast. Our analysis indicates that the prediction skill in CFSv2 comes from all time scales including a long-term trend, detrended interannual variability, and submonthly variability, with the detrended interannual variability being the dominant source of prediction skill.
The assessment in this study is based on the output from the operational NCEP CFSv2. Collow et al. (2015) indicated that the operational CFSv2 configuration is not optimal and the sea ice prediction at seasonal time scales can be improved with an improved initial sea ice thickness analysis. The importance of initial sea ice thickness in multidecadal time prediction is not clear and needs to be further investigated. In addition, while we have analyzed the contributions of sea ice variations and their prediction from trend interannual, detrended interannual, and submonthly variabilities, it will be useful to determine how these variations are contained in the initial conditions and how the memory of the components (e.g., sea ice, ocean, and atmosphere) of the forecast system affects the subsequent sea ice evolution. Model physics can affect the seasonal forecasts of air temperature and sea ice properties (Yang et al. 2017). Consequently, it will be interesting to find out how prediction skill depends on the uncertainty in each component of the forecast system. Understanding the impact of the initialization of each component will also provide guidance as to where the critical effort in improving the initialization of the forecast system should be made.

Acknowledgments. The authors greatly appreciate the valuable comments by Drs. Thomas Collow and Peitao...
APPENDIX

Sea Ice Anomaly Calculations

The procedure for the forecast anomalies is described below. The observed anomalies are calculated following the same steps by replacing $F$ with $O$.

The total anomaly $F_{\text{tot}}(d, L, \text{yr})$ is calculated as the deviation of total weekly mean values of SIC from the climatology, which is

$$F_{\text{tot}}(d, L, \text{yr}) = F(d, L, \text{yr}) - F_c(d, L),$$

where $d = 1-365$, $L = 0-5$ weeks, and $\text{yr} = 2000-2015$.

The total anomaly can be decomposed into contributions from different time-scale bands. In this study, we decompose the total anomaly into its interannual and submonthly components. The interannual anomaly $F'_{\text{ina}}(m, L, \text{yr})$, which contains time scales longer than a month, is defined as the monthly mean of the total anomaly defined above:

$$F'_{\text{ina}}(m, L, \text{yr}) = \frac{1}{n} \sum_{d=1}^{n} F'_{\text{tot}}(d, L, \text{yr}),$$

where $n = \text{total number of days for each target month } m$; for example, $n = 31$ for the target month March.

The interannual anomaly is further divided into two components: the trend and the detrended interannual anomaly. The trend interannual anomaly $F'_{\text{trend}}(m, L, \text{yr})$ is calculated as the linear trend of $F'_{\text{ina}}(m, L, \text{yr})$, and the detrended interannual anomaly $F'_{\text{ina dtr}}(m, L, \text{yr})$ is calculated as the difference between the total interannual and trend interannual anomalies $F'_{\text{ina}}(m, L, \text{yr}) - F'_{\text{trend}}(m, L, \text{yr})$.

The submonthly anomaly $F'_{\text{sub}}(d, L, \text{yr})$ is the deviation of the total anomaly from the interannual anomaly and is computed as

$$F'_{\text{sub}}(d, L, \text{yr}) = F'_{\text{tot}}(d, L, \text{yr}) - F'_{\text{ina}}(m, L, \text{yr}).$$

In summary, the total anomaly is the sum of trend interannual, detrended interannual, and submonthly anomalies:

$$F'_{\text{tot}}(d, L, \text{yr}) = F'_{\text{trend}}(m, L, \text{yr}) + F'_{\text{ina dtr}}(m, L, \text{yr}) + F'_{\text{sub}}(d, L, \text{yr}).$$

The meaning of “submonthly” comes from the fact that this component is the anomaly on top of the monthly mean anomaly and thus represents the variations of the time scales within the month, while the monthly mean, that is, the interannual component, represents variations at time scales including detrended interannual and the long-term trend.

Damped persistence

The damped anomaly persistence is a useful reference for evaluating the forecast skill, where the amplitudes of the anomalies are assumed to reduce in time exponentially with a time scale of the local autocorrelation time (e.g., Van den Dool 2006). In the limit of long lead time, the amplitude of the anomaly is damped toward zero, and therefore the damped anomaly persistence approaches the climatological mean with increasing lead times.
The damped persistence forecast for each target week is taken as the observed weekly mean anomaly for the week immediately preceding the forecast scaled by lagged correlation $R$ between the initial and target weeks:

$$F_{dp}(d, L, yr) = O_{ano}(d, L, yr) \times R,$$

where $d = 1–365$, $L = 0–5$ weeks, and $yr = 2000–2015$. Here, $R$ is the lagged correlation between the initial and target week for the observed weekly mean. It should be mentioned that the ACC skill of the damped persistence is the same as the lagged correlation. However, the RMSE of the damped persistence forecast is smaller than that of the persistence forecast.

REFERENCES


