Impacts of Sea Ice Thickness Initialization on Seasonal Arctic Sea Ice Predictions

ARLAN DIRKSON
School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia, Canada

WILLIAM J. MERRYFIELD
Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada, University of Victoria, Victoria, British Columbia, Canada

ADAM MONAHAN
School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia, Canada

(Manuscript received 6 June 2016, in final form 28 September 2016)

ABSTRACT
A promising means for increasing skill of seasonal predictions of Arctic sea ice is improving sea ice thickness (SIT) initial conditions; however, sparse SIT observations limit this potential. Using the Canadian Climate Model, version 3 (CanCM3), three statistical models designed to estimate SIT fields for initialization in a real-time forecasting system are applied to initialize sea ice hindcasts over 1981–2012. Hindcast skill is assessed relative to two benchmark SIT initialization methods (SIT-IMs): a climatological initialization currently used operationally and SIT values from the Pan-Arctic Ice Ocean Modeling and Assimilation System (PIOMAS). Based on several measures of skill, sea ice predictions are generally improved relative to a climatological initialization. The accuracy with which the initialization fields represent both the thinning of the ice pack over time and interannual variability impacts predictive skill for pan-Arctic sea ice area (SIA) and regional sea ice concentration (SIC), with the most robust improvements obtained with SIT-IMs that best represent both processes. Similar skill to that achieved by initializing with PIOMAS, including skillful predictions of detrended September SIA from May, is obtained by initializing with two of the statistical models. Regional skill for September SIC is also enhanced using improved SIT-IMs, with an increase in the spatial coverage of statistically significant skill from 10% to 60%–70% of the appreciably varying ice pack. Reduced skill is seen, however, in the Nordic seas using the improved SIT-IMs, resulting from an inherent cold sea surface temperature bias in CanCM3 that is amplified by a thicker initial ice cover.

1. Introduction
Seasonal forecasting of Arctic sea ice has received increased attention in recent years because of a growing demand for forecasts from an array of stakeholders. This demand has grown largely as a result of the increased access to Arctic waterways (Ellis and Brigham 2009), owing to the reduction in sea ice coverage, which is most prominent in the summer months (Serreze et al. 2007). The overall negative trend in pan-Arctic sea ice extent (SIE) is consistent with climate projections that show these reductions continuing into the future under increased greenhouse gas emission scenarios (Stroeve et al. 2012). Methodologies that have been used for seasonal sea ice forecasts include statistical regression-based methods, fully coupled atmosphere–ocean global climate models (AOGCMs), and heuristic approaches (Stroeve et al. 2014b; Guemas et al. 2016). However, few centers currently produce these forecasts operationally.

Arctic sea ice has been shown to be predictable on seasonal to interannual time scales using AOGCMs in both “perfect model” experiments and initialized hindcasts. By construction, perfect model experiments are unaffected by either systematic model errors or by imperfect initial conditions. However, the inherent predictability in the model itself may be biased. By contrast, initialized hindcasts provide an estimate of practical forecast skill, subject to the availability of observations, observational uncertainties, and model biases.
Using the perfect model approach, studies indicate that pan-Arctic sea ice area (SIA) and SIE are inherently predictable for up to 1–2 years, with seasonal reemergence of skill occurring out to 4 years (Blanchard-Wrigglesworth et al. 2011; Day et al. 2014; Tietsche et al. 2014). Through sensitivity studies, it has been shown that the predictability of September SIE and sea ice concentration (SIC)—the fractional ice coverage in a local grid cell—up to 8 months in advance relies on an accurate representation of sea ice thickness (SIT) initial conditions (Day et al. 2014). For forecasts initialized in winter, preconditioning of SIT anomalies contributes to predictive skill for ice area in the summer months, particularly from the memory of ice at least 1.5 m thick (Chevallier and Salas-Mélia 2012).

Initialized hindcasts (Chevallier et al. 2013; Sigmond et al. 2013; Merryfield et al. 2013b; Wang et al. 2013; Msadek et al. 2014; Collow et al. 2015) suggest that September SIE and SIA anomalies are predictable for long lead times (up to a year). A large component of this skill is attributable to the large negative trend in sea ice coverage. Variations around the trend have been found to be much less predictable, limited to maximum lead times between 2 and 6 months for September SIE or SIA. Sigmond et al. (2013) found that the Canadian Seasonal to Interannual Prediction System (CanSIPS), employing a crude SIT initialization (Merryfield et al. 2013a), produced hindcasts of detrended September SIA anomalies that are statistically skillful only when initialized in June or later. With more realistic SIT initialization procedures and using other models, statistically significant hindcast skill of detrended SIE/SIA in September has been obtained from as early as May (Chevallier et al. 2013; Wang et al. 2013) or even March (Msadek et al. 2014; Collow et al. 2015).

Most sea ice predictability studies have focused primarily on integral measures such as SIE or SIA; the predictability of regional sea ice coverage in initial-condition-based hindcasts has received less attention. One such analysis by Collow et al. (2015) showed that skill in predicting regional September SIC is sensitive to the SIT initialization used and also to model physics changes, which reduce model biases in sea surface temperatures (SSTs).

Both perfect model and initial-condition-based hindcasts suggest that forecast skill depends strongly on the SIT initialization used (Day et al. 2014; Tietsche et al. 2014; Collow et al. 2015). However, the ability to initialize SIT both in hindcasts and in real time is hampered by the limited observational record of SIT and further complicated by inconsistent SIT observing systems (Lindsay and Schweiger 2015). These difficulties pose a challenge to initializing SIT accurately in real time in a manner that is consistent with the 20–30 years of hindcasts that enable real-time bias correction and calibration.

The goal of the present study is to evaluate sea ice hindcast skill using a range of SIT initialization methods (SIT-IMs) that include statistical models designed to be applicable in both hindcast and real-time forecast settings. We examine sea ice hindcast skill over a 32-yr period spanning 1981–2012. Hindcasts are generated using the Third Generation Canadian Centre for Climate Modelling and Analysis (CCCma) Canadian Climate Model (CanCM3) (Merryfield et al. 2013a). The details of these hindcast simulations are described in section 2. A summary of the five SIT-IMs considered is given in section 3. Hindcast skill is evaluated in section 4, wherein predictive skill for both integrated Arctic SIA and spatially varying SIC is examined. The dependence of SIC skill on differences in hindcast SIT and SST is also considered. Finally, a discussion and conclusions are presented in section 5.

2. Sea ice hindcasts

CanSIPS produces operational seasonal forecasts based on CanCM3 and the Fourth Generation Canadian Coupled Global Climate Model (CanCM4) (Merryfield et al. 2013a). CanCM3 uses the Third Generation Canadian Atmospheric General Circulation Model (CanAM3), whereas CanCM4 uses the Fourth Generation Canadian Atmospheric General Circulation Model (CanAM4). Both CanAM3 and CanAM4 have horizontal grid spacings of approximately 2.8° but differ in their vertical resolutions (31 levels for CanAM3 and 35 levels for CanAM4). CanCM3 and CanCM4 share the same land, ocean, and sea ice components. The ocean model used is the CCCma Fourth Generation Ocean Model (CanOM4). CanOM4 has an approximately 100-km horizontal grid resolution with 40 vertical levels with spacing of 10 m near the surface and increasing with depth. Sea ice is modeled as a cavitating fluid with a one-category ice thickness following Flato and Hibler (1992).

Hindcasts using CanCM3 are considered in this study because of its lower computational expense compared with CanCM4. As multimodel ensembles are generally more skillful than single-model forecasts (e.g., Kharin et al. 2009), our results likely provide a lower-end estimate of Arctic sea ice skill relative to that which could be achieved by the two-model combination employed by CanSIPS.

Merryfield et al. (2013a) show that freely running historical simulations of CanCM3 yield Arctic sea ice biases relative to the observational Hadley Centre Sea Ice and Sea Surface Temperature dataset, version 1.1 (HadISST1.1) (Rayner et al. 2003). Specifically, CanCM3 overestimates SIE in all calendar months and incorrectly simulates the minimum seasonal extent to occur in...
August rather than September. Additionally, CanCM3 shows biases in its climatological SIC distribution. Positive SIC biases of 15%–50% occur in September in the Greenland, Barents, Kara, Laptev, and East Siberian Seas while the western Arctic near the Canadian Archipelago shows negative biases of up to 15%. In March, large positive biases in SIC are seen in the Labrador, Greenland, and Barents Seas, and negative biases of 25%–50% are seen in the Bering Sea.

a. Hindcast configuration

The hindcasts considered in this study extend six target months, initialized on the first day of each of March, May, June, and September. The March, May, and June initializations are chosen to include the spring and summer melt months within the forecast range. September initializations are chosen to include the autumn freeze and winter growth seasons. Each hindcast set consists of 10 ensemble members that have slightly different initial conditions intended to represent observational uncertainties. As described in Merryfield et al. (2013a), initial conditions are obtained from a set of assimilation runs (one for each ensemble member), in which SSTs, SIC, and atmospheric variables are constrained near observation-based values with relaxation time scales of 3 days for SST and SIC and 24 h for atmospheric fields. SIC is nudged toward HadISST1.1, whereas SIT is nudged toward values predicted by each SIT-IM, also with a relaxation time scale of 3 days. Daily sea ice fields are obtained by interpolating between monthly mean fields.

b. Defining interannual variability

Throughout this study, the distinction between sea ice prediction skill associated with the trend and prediction skill associated with interannual variability will be made. In previous studies, the interannual predictive skill for SIA or SIE has been assessed by evaluating the skill of their residual values relative to a best-fit linear trend (Chevallier et al. 2013; Sigmond et al. 2013; Merryfield et al. 2013b; Wang et al. 2013; Collow et al. 2015). To date, no study has assessed the prediction skill for interannual variability of spatially distributed variables like SIC. Here, we offer a cautionary note when defining interannual variability that should be considered for both spatially integrated and spatially varying sea ice quantities.

In quantifying skill for quantities with a nonstationary mean, like most measures of Arctic sea ice (i.e., SIE, SIT, SIC), skill metrics like the anomaly correlation coefficient (ACC) can be substantially affected by the presence of a trend. Thus, it is often useful to separate skill associated with the trend from that associated with interannual variability. By construction, the definition of interannual variability is determined by how one chooses to define the long-term trend. The choice of trend definition should thus be made carefully depending on the variable and time period considered, as a detrending method suited for one sea ice quantity may not be well suited for another sea ice quantity. In particular, if the trend accelerates with time, then a choice other than linear trend removal may be appropriate. Throughout this work, we therefore consider skill metrics after having detrended the time series of interest using both linear and quadratic fits and assess the sensitivity of skill to the use of each trend-fitting method. When differences in skill are evident, an assessment of the time series under consideration is done to inform the choice of the most appropriate fit, bearing in mind the potential for overfitting when applying multiparameter fits to short records.

3. SIT initialization methods

a. Original

The SIT-IM currently used in CanSIPS will be referred to as Original and consists of nudging SIT values toward a monthly SIT climatology (Merryfield et al. 2013a), the CCCma synthetic sea ice thickness climatology, which was developed for use under the Atmospheric Model Intercomparison Project (AMIP). These climatological ice thicknesses were obtained through a sea ice growth relationship similar to that described by Anderson (1961), using prescribed seasonally varying climatological sea ice concentrations and near-surface temperatures as input. Because of the late twentieth-century epoch of these input data, the simulated thicknesses are more reflective of conditions before 2000 than of the more recent period, which has seen a substantial decline in ice volume. Hence, this method does not account for the negative trend in SIT, nor does it represent SIT interannual variability. Consequently, the use of Original leads to an underestimation of the negative SIA trend in hindcasts and potentially limits skill in forecasting SIA interannual variability (Sigmond et al. 2013).

b. PIOMAS

The Pan–Arctic Ice Ocean Modeling and Assimilation System (PIOMAS) is a high-resolution (averaging 9/8°) coupled sea ice–ocean model, in which sea ice evolves based on a multicategory ice thickness and enthalpy distribution (TED) model (Zhang and Rothrock 2003). Several fields are assimilated in PIOMAS through a flow-dependent nudging, including SIC data from the National Snow and Ice Data Center (NSIDC), as well as SSTs and atmospheric variables from the National
Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (Kalnay et al. 1996). We refer the reader to other studies for a detailed assessment of the skill of PIOMAS SIT reconstructions relative to observations (Schweiger et al. 2011; Laxon et al. 2013; Stroeve et al. 2014a) and other reanalyses (Chevallier et al. 2016), but in general, SIT from PIOMAS compares reasonably well with a range of satellite and in situ observations.

Monthly SIT fields from PIOMAS are regularly updated online but are not available in real time. Hindcasts initialized by relaxing SIT to PIOMAS are performed because PIOMAS is used to train the statistical models considered in this study for SIT initialization. Hindcasts initialized with PIOMAS are therefore expected to represent an upper limit for predictability relative to the statistical models developed here.

Comparisons of March and September SIT climatologies—defined over the period 1981–2010 to be consistent with hindcasts—for PIOMAS and Original are presented in Fig. 1. Both PIOMAS and Original show larger ice thickness in the western Arctic compared with the eastern Arctic. However, PIOMAS generally has thinner ice in the central Arctic, extending south into the Laptev and Kara Seas, whereas PIOMAS has thicker ice in the western Kara Sea, Barents Sea, Greenland Sea, and along the Greenland, Canadian, and Alaskan coastlines into the Chukchi Sea. These greater ice thickness values for PIOMAS are more widespread in March compared with September. Differences in SIT between PIOMAS and Original during the spring months resemble those in winter (not shown).

c. Statistical models

Motivated by the potential for improving upon the Original initialization scheme, three statistical models, denoted SMv1 through SMv3, have been developed to estimate monthly mean SIT in real time. Although these three models rely on PIOMAS SIC and SIT data to estimate model parameters, the statistical models only require that PIOMAS data be available up to one year prior to the month in which SIT is to be estimated. Because these statistical models do not require PIOMAS SIT fields in real time, they can be applied in an operational setting.

The statistical models make use of (either or both of) two predictor fields: SIC from PIOMAS and sea level pressure (SLP) from the ERA-Interim (Dee et al. 2011) and ERA-40 (Uppala et al. 2005) reanalyses. The physical basis for these predictors lies in their thermal and dynamical relationships with SIT. SIC is considered because it is correlated with mean ice thickness in most months (Lisæter et al. 2003), as well as locally with SIT in the marginal ice zone (e.g., Tietsche et al. 2013). Furthermore, both SIC and SIT show a strong negative trend over most of the Arctic. SLP is chosen because of the coupling between atmospheric motion and sea ice motion on monthly to multimonthly time scales (Thordike and Colony 1982), potentially influencing SIC through convergent or divergent ice motion (Rigor et al. 2002). Other near-surface atmospheric dynamical predictors such as winds are expected to be well correlated with SLP on the time scales considered.

For years 1994 onward, the statistical models use a 15-yr training period \( \tau = \{ t_{15}, t_{14}, \ldots, t_{-1} \} \) preceding the target year for initialization \( t_c \). Over \( \tau \) the predictor(s) and predictand are both known. The predictor and predictand fields for month \( m \) and year \( t \) are denoted respectively by \( x_m(t) \) and \( y_m(t) \). Statistical model parameters are first estimated over \( \tau \) using the predictor(s) \( x_{m}(\tau) \) and predictand \( y_{m}(\tau) \). The parameters are then applied with the real-time predictor(s) \( x_{m}(t_c) \) to make an estimate of \( y_{m}(t_c) \), denoted \( \hat{y}_{m}(t_c) \). For the predictand \( y_{m}(t_c) \), we use PIOMAS SIT, and skill measures for SIT estimated by the statistical models are computed by treating PIOMAS as truth. For years through 1993 (prior to having a 15-yr training period), the statistical models use simpler approaches for estimating SIT (to be described) based on training data spanning a shorter period \( \tau = \{ 1979, \ldots, t_{-1} \} \). The three statistical models are described below and summarized in Table 1.

1) SMV1

The SMV1 statistical model for initializing SIT is described in detail in Dirkson et al. (2015). In brief, SMV1 employs maximum covariance analysis (MCA) over \( \tau \) to identify patterns of covariability between PIOMAS SIT and each of two predictors: PIOMAS SIC and lagged (4-month averaged) sea level pressure (SLPlag). These predictors, denoted as \( \mathbf{x}^{\text{SIC}}(\tau) \) and \( \mathbf{x}^{\text{SLPlag}}(\tau) \), are used separately to construct distinct MCA models. In the formulation of SMV1, separate SIT estimates are made using the leading mode of covariability between SIT and each predictor. This is motivated by the finding in Dirkson et al. (2015) that SLPlag is the more skillful predictor over the first nine years of the verifying period (1995–2003) when the negative trend in SIT is smaller, whereas SIC performs best over the second 9-yr period (2004–12) when the negative trend is larger. It is worthwhile noting the first mode of covariability between SIC and SIT has a strong negative trend circa \( t_c = 2000 \) onward and that this mode is uncorrelated with the first mode of covariability between SLPlag and SIT on an interannual basis. These estimates of SIT are combined to produce the final estimates, weighted by the relative importance of interannual variability and the trend,
where $s^2$ represents the total variance of sea ice volume in the interval $t$ and is the sum of the variances associated with interannual variability and the trend, respectively denoted $s_{I}^2$ and $s_{T}^2$. Prior to 1994, SIT is set to its average value for the month considered over the shorter training period $t$. Finally, from 1994 onward, a 5-yr mean bias correction is applied to SIT estimated by Eq. (1) to reduce a positive SIT bias.

Over 1994–2012 and across all calendar months, SMv1 reduces the areal and temporal mean absolute error (ATMAE) of estimated SIT by 50% compared with Original (Table 2). The primary contributor to this improvement is the more accurate representation of the negative trend in SIT. Although SMv1 skill in modeling SIT interannual variability is improved through the inclusion of $\frac{\sigma^2_{m_{SLP}}(t)}{\sigma^2_{m_{SIC}}(t)}$ in Eq. (1), such skill is still limited. We illustrate these differences in predictability using maps of the ACC for SIT for two cases: with and without the long-term trend in SIT included. Here, we define interannual variability through linear detrending, which was chosen after inspecting local SIT time series at several locations and concluding that a nonlinear component in the trend is not important. The ACCs are calculated separately for each calendar month over the period 1994–2012 (over which time SIT is not obtained from climatology) and then averaged over two 3-month periods: January–March (JFM) and July–September (JAS). Statistical significance at the 95% confidence level is estimated by resampling detrended SIT time series using bootstrapping, with a sample size of 10 000 at each grid point. When the trend is included (Fig. 2a), SMv1 correlations are significant over most of the domain in both 3-month periods. However, areas of statistically significant skill are localized in the central Arctic and Beaufort Sea in JAS when the trend is removed (Fig. 2b).

2) SMv2

SMv2 improves upon SMv1 through an additional step. After computing SMv1 SIT, the sign of SIC and predicted SIT anomalies relative to climatology over $t$ are compared at each grid location. In cases where the SIC and SIT anomalies disagree in sign, the SIT anomaly is set to a value proportional to the local SIC anomaly using

$$\tilde{y}_m(t) = \langle y_m \rangle_t + \alpha [x_m^{SIC}(t) - \langle x_m^{SIC} \rangle_t],$$

where $\alpha$ is a proportionality constant found through sensitivity testing, and the angled brackets $\langle \rangle_t$ denote the mean over the training period. The parameter $\alpha$ has been tested at values in the range $[0.5, 3]$ at increments of 0.5 m. The ATMAE for Eq. (2) varies only slightly for different $\alpha$ values in this range, with a minimum in the ATMAE occurring at $\alpha = 2$ m. Equation (2) resembles that employed by Tietsche et al. (2013), although their method nudges SIT in an assimilation cycle based on differences between modeled and observed SIC. Applying Eq. (2) for 1981–93 with $t = t_s$ instead of the SIT climatology (as in SMv1) further improves skill based on the ATMAE (Table 2). As in SMv1, from 1994 onward, a 5-yr mean bias correction is applied to SIT estimates to reduce a positive SIT bias.

The additional step in SMv2 given by Eq. (2) relies on the assumption that SIC and SIT are positively correlated on a year-to-year basis. This assumption is more robust in the marginal ice zone but is less valid in locations where SIC is consistently near 100%. As measured by the ATMAE, SMv2 outperforms SMv1 (Table 2). Furthermore, SMv2 improves over SMv1 in terms of the ACC for SIT. This improvement is seen when the trend is included in the Bering Sea and Sea of Okhotsk in JFM and in Fram Strait and Davis Strait in JAS (Fig. 2c). SMv2 is seen to greatly improve upon SMv1 in terms of its representation of interannual variability (Fig. 2d) in that the use
of SMv2 results in larger values of skill and a larger spatial extent of significant correlation in both JFM and JAS.

3) SMv3

The improvement of SMv2 relative to SMv1, particularly with respect to interannual variability, further demonstrates that a large fraction of skill using the MCA-based approach (SMv1) is a result of capturing the negative trend in SIT. To assess the skill of a model that represents this trend simply through extrapolation, we introduce the third statistical model given by

\[
\tilde{y}_m(t_e) = \tilde{y}_m(t_e) + \alpha[x_m^{SIC}(t_e) - \tilde{x}_m^{SIC}(t_e)].
\]

In Eq. (3), \(\tilde{y}_m(t_e)\) and \(\tilde{x}_m^{SIC}(t_e)\) respectively represent the extrapolation of the local linear SIT and SIC trends calculated over \(\tau\), and the quantity \(x_m^{SIC}(t_e) - \tilde{x}_m^{SIC}(t_e)\) represents detrended SIC anomalies. Like Eq. (2), Eq. (3) assumes that SIC and SIT are positively correlated on an interannual basis and is therefore subject to the same limitations stated previously.

While SMv3 shows the same skill values as SMv2 over 1994–2012 based on the ATMAE (Table 2), Eq. (2) remains the model with the lowest ATMAE over 1981–93. Skill based on the ACC is generally better for SMv3 in regions where both SMv2 and SMv3 have positive skill (Figs. 2e,f). However, detrended SIT fields produced by SMv3 show poorer skill relative to SMv2 in the near-polar region of the ice pack where SIC varies relatively little. In the marginal ice zone, SMv3 generally performs better than SMv2. The largest improvements relative to SMv2 in the detrended case are in JAS.

4. Hindcast results

a. Verification data

To assess the sea ice hindcasts performed in this study, we use the NSIDC merged SIC dataset (Meier et al. 2014). This product uses the climate data record (CDR) SIC merging algorithm, which combines estimates of SIC derived from NASA bootstrap (Comiso 1986) and NASA Team (Gloersen 1984) retrieval algorithms. The CDR algorithm uses the more accurate estimate of the sea ice edge from NASA bootstrap based on a 10% SIC coverage threshold. Within this region, the merging algorithm assigns SIC on a gridpoint-by-gridpoint basis according to the larger value between NASA bootstrap and NASA Team. This is done because both NASA bootstrap and NASA Team tend to underestimate SIC but from different sources of bias (Meier et al. 2014). Prior to all calculations, this product is interpolated onto the CanCM3 model grid. Afterward, both the verification product and hindcast SIC values below 10% are set to 0% to be consistent with the original verification dataset.

b. Sea ice area

The dependence of Arctic sea ice prediction skill on the five SIT-IMs is first assessed by considering hindcast SIA. SIA is defined as the area integral of SIC in the NH. SIA anomalies for individual hindcast ensemble members are calculated relative to the 1981–2010 baseline climatology for ensemble mean SIA. The final deterministic SIA anomaly hindcast is then defined as the mean of the SIA anomalies across all ensemble members.

We first consider hindcasts of September SIA anomalies initialized in May. The hindcast SIA anomalies (both the ensemble mean and ensemble spread, defined as \(\pm 2\) standard deviations) are indicated along with observed SIA anomalies in Fig. 3. A second-order polynomial fit is shown for both the observed and hindcast SIA anomalies to provide a visual comparison of trends. The accelerating decline of SIA motivates the use of a quadratic fit. Compared against the fitting statistics of the widely used linear fit, the quadratic fit reduces the root-mean-square error (RMSE) between the fit and observed SIA from 0.49 \(\times 10^6\) km\(^2\) to 0.07 \(\times 10^6\) km\(^2\) and increases \(r^2\) from 0.73 to 0.83. The RMSE between the ensemble mean hindcast anomalies and the observed anomalies in

### Table 1. The algorithms for all statistical models used to initialize SIT: SMv1, SMv2, and SMv3.

<table>
<thead>
<tr>
<th>Period</th>
<th>Step</th>
<th>SMv1</th>
<th>SMv2</th>
<th>SMv3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1981–93</td>
<td>1</td>
<td>Set SIT to ((y_m)_\tau)</td>
<td>Set SIT to Eq. (2) with (\tau = \tau_i)</td>
<td>Set SIT to Eq. (2) with (\tau = \tau_i)</td>
</tr>
<tr>
<td>1994–2012</td>
<td>2</td>
<td>Set SIT to Eq. (1)</td>
<td>Set SIT to Eq. (1)</td>
<td>Set SIT to Eq. (3)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Subtract preceding 5-yr mean error</td>
<td>Where SIC and SIT anomalies disagree in sign, set SIT to Eq. (2)</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>—</td>
<td>Subtract preceding 5-yr mean error</td>
<td>—</td>
</tr>
</tbody>
</table>

### Table 2. ATMAE across calendar months for two periods: 1981–93 and 1994–2012. The percentage that the ATMAE improves relative to Original (IRO) is displayed for SMv1, SMv2, and SMv3.

<table>
<thead>
<tr>
<th></th>
<th>ATMAE (m)</th>
<th>IRO</th>
<th>ATMAE (m)</th>
<th>IRO</th>
</tr>
</thead>
<tbody>
<tr>
<td>1981–93</td>
<td>Original</td>
<td>0.57</td>
<td>—</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>SMv1</td>
<td>0.32</td>
<td>44%</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>SMv2</td>
<td>0.29</td>
<td>49%</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>SMv3</td>
<td>0.29</td>
<td>49%</td>
<td>0.26</td>
</tr>
<tr>
<td>1994–2012</td>
<td>Original</td>
<td>—</td>
<td>0.56</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>SMv1</td>
<td>0.32</td>
<td>44%</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>SMv2</td>
<td>0.29</td>
<td>49%</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>SMv3</td>
<td>0.29</td>
<td>49%</td>
<td>0.26</td>
</tr>
</tbody>
</table>
SIA is indicated on each panel in addition to the ACC with the long-term trend included \((r)\), linearly detrended \((r_l)\), and quadratically detrended \((r_q)\).

The RMSE for September hindcasts of SIA decreases from 0.89 \(\times 10^6\) km\(^2\) using Original to between 0.55 \(\times 10^6\) km\(^2\) and 0.61 \(\times 10^6\) km\(^2\) when using the improved SIT-IMs. Additionally, skill as measured by the ACC is improved considerably using all SIT-IMs other than Original when the trend is included, with ACC values increasing from 0.38 to 0.87–0.91. When the linear trend is removed, SMv1-SMv3 and PIOMAS hindcasts show similar skill (ACCs between 0.52 and 0.65); however, skill is sensitive to the detrending method chosen, as can be seen by the further reduction in skill using quadratic detrending, with ACC values ranging from 0.24 to 0.49.

Although the negative trend in SIA is represented in hindcasts initialized with SIT-IMs other than Original, the long-term trend is still underestimated relative to the observed trend. Beyond an inherently small SIA/SIE trend in uninitialized historical forecasts using CanCM3 (Merryfield et al. 2013a), one of the possible reasons for the underestimation of the predicted negative trend in summer posited in Sigmond et al. (2013) is the fact that the SIC dataset HadISST1.1 used to initialize hindcasts has smaller trends in SIA compared with the NSIDC product, particularly in winter and spring and including May when these hindcasts are initialized. We test this hypothesis directly with a separate set of hindcasts initialized with SIC from NSIDC and SIT from SMv3. A comparison with hindcasts initialized with HadISST1.1 SIC and SMv3 SIT reveals that, although hindcasts initialized in May start out with a more negative linear trend in SIA when initializing with NSIDC SIC (by 0.23 \(\times 10^6\) km\(^2\) decade\(^{-1}\)) over the month of April prior to a May initialization, the difference reduces to only 0.07 \(\times 10^6\) km\(^2\) decade\(^{-1}\) for a lead time of zero months (averaged over May). By September, the initialized SIC field has no statistically significant influence on the trend. Other factors must therefore be inhibiting sea ice decline in CanCM3, and the exact source of bias is beyond the scope of this study.

To assess the skill of each of the 6-month hindcasts initialized in the four months considered, we consider the ACC for SIA over the period 1981–2012 (Fig. 4). Skills that are statistically significant at the 95% confidence level are determined by resampling detrended time series using bootstrapping with a sample size of 10,000. The ACCs shown in Fig. 4a are for the full SIA
anomaly time series, including the long-term trend. We see that hindcast skills for each target month are statistically significant for nearly all SIT-IMs, excluding August SIA when initialized in March using Original. Hindcast SIA skills obtained using all other SIT-IMs improve substantially over Original. This improvement is especially evident for hindcasts initialized in winter or spring. The improvements in skill for hindcasts initialized with SMv1-SMv3 and PIOMAS result primarily from their improved representation of the SIA trend.

The long-term trend is a prominent aspect of the SIA record, and it is important that a forecast system should capture it. However, trend extrapolation could be captured by a simple regression-based statistical model and without the need for a comprehensive AOGCM. Of arguably greater importance in the assessment of the skill of a seasonal forecast system is quantifying how well interannual variability is captured. To estimate interannual skill for each set of hindcasts, we calculate the ACC using both linearly detrended SIA anomalies (Fig. 4b) and quadratically detrended SIA anomalies (Fig. 4c). Linear detrending is shown so that these results may be compared against other studies that have also represented interannual skill using linear detrending. However, we argue that...
quadratic detrending is the more appropriate technique for SIA over the time period considered here because linear detrending leaves a residual long-term quadratic signal in the observed time series, particularly in target months during the late spring and summer when the trend and its curvature are largest, as evidenced by the lower RMSE and higher \( r^2 \) for the quadratic fit to September SIA discussed above. Others have used different approaches for detrending SIA/SIE time series, such as linear regression onto observed CO\(_2\) concentrations (Germe et al. 2014). This approach has also been attempted here, but the results are very similar to those obtained using linear detrending and so are not shown.

The skill of the detrended hindcasts is lower than that of hindcasts including the trend, as has been found in other studies (e.g., Chevallier et al. 2013; Sigmond et al. 2013). Furthermore, skill is generally higher when using linear detrending as compared with quadratic detrending. Hindcasts generated with SMv1 show relatively minor improvements relative to Original with linear

![ACCs for SIA over the period 1981–2012](image)

**FIG. 4.** ACCs for SIA over the period 1981–2012, shown as a function of target month (horizontal axis) and lead month (vertical axis). The ACCs measure (a) overall skill based on the original SIA time series, (b) interannual skill based on linearly detrended SIA time series, and (c) interannual skill based on quadratically detrended SIA time series. Stippling indicates statistical significance at the 95% confidence level.
detrending, whereas there is a reduction in skill when using quadratic detrending. Hindcasts initialized using SMv2, SMv3, and PIOMAS, on the other hand, show generally greater skill both with the trend included and with the two types of detrending. These results demonstrate that a better representation of interannual variability in the initialized SIT field improves interannual hindcast skill of SIA.

Focusing on hindcasts initialized in March, detrended SIA skill falls below significant values within one–two months after initialization for Original, SMv1, and PIOMAS (Figs. 4b,c). Skill using SMv2 and SMv3 is larger than that for SMv1, Original, and PIOMAS. However, the relatively small differences in skill for target months April through August between these SIT-IMs is likely attributable to sampling rather than robust differences in skill, particularly when ACC values lie near the threshold for statistical significance. The reemergence of skill in August when using linear detrending for SMv1-SMv3 and PIOMAS is less obvious with quadratic detrending and is likely associated with residual skill resulting from the incomplete removal of the trend.

For hindcasts initialized in May, a barrier in predictive skill is seen when Original is used to initialize SIT compared with hindcasts initialized in June, as was found in Sigmond et al. (2013) based on both CanSIPS models. Hindcasts initialized with SMv1 also see this barrier in predictive skill when using quadratic detrending. However, skillful hindcasts initialized in May are produced using SMv2, SMv3, and PIOMAS for all target months with the use of either linear or quadratic detrending. The skill of hindcasts initialized in May in these cases is similar to the skill for those initialized in June, when statistically significant skill predicting detrended SIA through October is produced using all SIT-IMs.

Hindcasts initialized in September show significant skill for interannual variability only through October for all SIT-IMs when quadratic detrending is used, whereas hindcasts initialized with SMv2, SMv3, and PIOMAS show significant skill throughout all six target months using linear detrending. Skill decreases with increasing lead time through December regardless of the detrending method used and increases again in January for all SIT-IMs. This reemergence of skill is similar to that found in previous predictability studies (Holland et al. 2011; Day et al. 2014; Tietsche et al. 2014), in which perfect model experiments showed reemergence of skill peaking in winter. Furthermore, Sigmond et al. (2013) found that significant detrended skill using CanSIPS can be achieved in January–February at a lead time of 11 months. This high level of skill in winter has been attributed to the predictability of the location of the ice edge (Holland et al. 2011), resulting from heat transport variations and/or the persistence of SST anomalies (Bitz et al. 2005; Blanchard-Wrigglesworth et al. 2011).

c. Regional skill

Total Arctic SIA is an integrated measure of sea ice cover that is often useful to describe the conditions of the ice pack as a whole and is commonly used when estimating sea ice prediction skill. While useful for comparing skill between models, integrated measures like SIA and SIE provide little information regarding regional sea ice conditions of interest to potential forecast users. The practical utility of sea ice forecasts is increased by considering spatially distributed quantities like regional SIC.

1) SIC SKILL: ACC

Maps of ACC for predicted SIC initialized in May are shown in Fig. 5 for time series that include the trend (overall skill) and in Fig. 6 for time series that have been linearly detrended (interannual skill). Detrending using a second-order polynomial fit was also performed, but differences are generally small between the two trend definitions. In regions where skill differences are larger, examination of the SIC time series reveals that a quadratic fit misrepresents the trend because of overfitting resulting from a relatively small signal-to-noise ratio compared with SIA. At each model grid point, the ACC 95% confidence level is computed by resampling detrended SIC anomalies using bootstrapping with a sample size of 10,000.

For hindcast anomalies that include the trend (Fig. 5), we see large improvements in skill in both the western and eastern Arctic when SMv1–SMv3 and PIOMAS are used to initialize ice thickness compared with when Original is used. Skill is greater in the western Arctic in target months June and July for hindcasts with improved thickness initializations compared with Original hindcasts, whereas hindcasts show significant skill through July in a large portion of the eastern Arctic regardless of the SIT-IM used. In August through October, significant skill is reduced to small localized areas for Original, whereas SMv1–SMv3 and PIOMAS hindcasts show significant skill over a large and coherent portion of both the western and eastern Arctic.

Skill for predicting interannual SIC variability in hindcasts initialized in May is shown in Fig. 6. In general, we see lower skill for the improved SIT-IMs than when including the trend, whereas skill in Original hindcasts is similar to when the long-term trend is included. Overall, some significant regional skill is evident for all initialization methods throughout all six target months, and
ACC values throughout the Arctic are generally positive independent of the initialization method used. Similar skill is seen for all SIT-IMs in the first two target months (through June). Improved skill relative to Original is seen in the Beaufort Sea in August and September using SMv2, SMv3, and PIOMAS. In October, significant skill is found in the Kara Sea when SIT is initialized with SMv2, SMv3, and PIOMAS. For target months July through October, hindcasts initialized with PIOMAS and SMv2 are more skillful than with other SIT-IMs. However, apparently confounding results such as improved skill using SMv2 relative to PIOMAS in August through October, or significant skill in the Laptev Sea in October only seen in SMv3 hindcasts, suggest that some skill features may be the artifacts of sampling.

As a scalar measure of regional skill for all initialization months and target months, we calculate the fraction of area within the relevant domain that contains statistically significant SIC ACC values (Fig. 7). The relevant domain is defined as the region where observed SIC standard deviation is at least 1% over the time period considered. This restriction excludes locations that are almost always ice covered or ice free. The advantage over considering skill relative to the total domain is that this fractional quantity is less dependent on the proportional area of the marginal ice zone to the total ice extent, which changes considerably from month to month.

We first consider the areal fraction of significant ACC skill (AFSS) when the trend is included (solid lines in Fig. 7). In all initialization months, SMv1–SMv3 and PIOMAS hindcasts show a larger AFSS than Original throughout nearly all target months. Differences between Original and other SIT-IMs become more pronounced three months after initialization for hindcasts initialized...
in March, and between one and two months after initialization for hindcasts initialized in May, June, or September.

When initialized in March, the AFSS decreases relatively rapidly through May for all SIT-IMs. From May through August, skill remains significant across approximately 30%–50% of the relevant domain for hindcasts initialized with the improved SIT-IMs but continues to decrease down to 10% of the domain by August for hindcasts initialized with Original. Hindcasts initialized in May with the improved SIT-IMs show an AFSS averaging near 60% for target months June through October, while Original hindcast skill gradually decreases to near 10% of the relevant domain. For the target month of September, hindcasts initialized with SMv1–SMv3 and PIOMAS have an AFSS 5–6 times greater than that for Original hindcasts. When initialized in June, the AFSS is near or exceeds 60% in all target months using the improved SIT-IMs, whereas it decreases to below 30% in September using Original. When initialized in September, the AFSS decreases similarly for all SIT-IMs down to 35%–40% in February.

The AFSS values after the trend has been removed (indicated by dashed lines in Fig. 7) generally vary less between different SIT-IMs than when the AFSS values are calculated with the trend. When initialized in March, the AFSS decreases nearly identically for hindcasts other than those initialized with SMv1 through all target months. Over this time, the AFSS exceeds 20% through May but decreases below 20% for later target months. For May initialized hindcasts, the largest AFSS differences between different SIT-IMs is seen from June through August, with greatest skill seen using SMv2 and PIOMAS. For September and October target months, the AFSS grows again for SIT-IMs other than Original and covers over 20% of the relevant domain. When initialized in June, all SIT-IMs yield predictions with an AFSS exceeding 30% through all six target months. Skill using PIOMAS is higher than for other SIT-IMs through
all target months, with the other SIT-IMs performing similarly to each other. For hindcasts initialized in September, the AFSS remains above 50% through November using all SIT-IMs, at which time, a larger separation in skill emerges between SIT-IMs with greatest skill seen using PIOMAS and SMv3. Skills converge again between initialization methods in January and February.

2) SIC SKILL: RMSE

Skill measures based on correlation coefficients are insensitive to errors in the magnitude of the hindcast anomalies. To provide a more complete assessment of hindcast performance in predicting SIC, we consider the spatial distribution of RMSE for hindcasts initialized in May. We focus on hindcasts initialized in May to target summer months. Maps of the difference between the RMSE of SIC anomalies for SMv3 and Original hindcasts (SMv3 minus Original) are shown in Fig. 8. We focus on SMv3 because differences in this skill metric are small between SMv2, SMv3, and PIOMAS. We do not show SMv1 because of its poor interannual skill in SIA and SIC as measured by the ACC. RMSE differences are defined such that better skill for SMv3 is indicated by negative values. The RMSE is calculated separately over two 16-yr periods from 1981 to 1996 and from 1997 to 2012 to highlight differences in skill during periods when the magnitude of the negative trend is small compared with when it is larger.

Differences in RMSE between SMv3 and Original in May are generally small over both periods. From June through September, we see a broad pattern of higher skill for SMv3 hindcasts in the western Arctic and lower skill in the eastern Arctic. The lower skill in the eastern Arctic in both periods is mainly confined to the northern Greenland and Barents Seas. Higher skill using SMv3 is evident over 1997–2012 throughout much of the Arctic Basin from July through October. This improvement is most noticeable from August through September over the Beaufort, northern Chukchi, East Siberian, Laptev, and northern Kara Seas. In September and October, improvements relative to Original span the Alaskan and Russian coasts and the Canadian Archipelago. The improvement in predicting SIC in the Kara Sea is evident in June–July and again in October when the ice edge advances south into this region. The improvement in fall in the Kara Sea was seen previously using the ACC of detrended SIC in Fig. 6, indicating a better representation of interannual refreezing in this region using SMv3.

The next subsection examines reasons for the markedly greater skill of SMv3 hindcasts compared with Original hindcasts in the Kara Sea, as well as the lower skill in the adjacent northern Greenland and Barents Seas, in order to better understand how the SIT initialization together with model biases impacts skill in forecasting SIC.

3) SIT INFLUENCE ON HINDCAST ERRORS

The Nordic seas region extending from the northern Greenland Sea into the Barents and Kara Seas shows a cold SST bias and positive SIC bias in freely running
simulations of CanCM3 (Merryfield et al. 2013a). To understand two features in Fig. 8—the poorer skill in SMv3 hindcasts in the northern Greenland and Barents Seas from June through October and the improved skill in the Kara Sea in spring and again in fall—we consider mean differences between SMv3 and Original hindcasts of SIT and SST in this region, presented in Fig. 9. To compare against the differences in skill using the RMSE, we consider hindcasts initialized in May and mean differences calculated separately over the same two periods, 1981–96 and 1997–2012. The reader should bear in mind that as lead time increases, differences in SIT and SST fields diminish as the model drifts closer to its own climatology (i.e., as the memory of initial conditions is lost).

We first focus on the region of poorer skill in SMv3 hindcasts throughout northern Greenland Sea and Barents Sea (Fig. 8) and consider the corresponding differences in SIT and SST in Fig. 9. We see that in this region, SIT is greater in the SMv3 hindcasts throughout all six target months in the first period and throughout the first two–three target months in the second period. Additionally, we see cooler SSTs in this region (to the south of the SIT differences), which are coldest and most expansive in the first period. Because SIT initialization is the only difference in these hindcasts, these cooler SSTs are evidently due to the thicker SITs in this region. The SIT bias and corresponding SIC bias in this region in freely running simulations is thus amplified by the thicker SITs at initialization and therefore degrades skill predicting SIC. Furthermore, the fact that SIC skill in the second period is as poor as in the first period, despite the SIT differences being smaller, implies that the sensitivity of SIC skill to these SIT and SST differences is high.

We now focus on the region of improved skill for SIC RMSE in SMv3 hindcasts in the Kara Sea (Fig. 8). Improved skill in June through July and again in October is largest in the second period. During this time, SIT is thinner in the eastern Kara Sea in the first three target months. By July, the ice has retreated north, and warmer SSTs emerge throughout all of the Kara Sea, peak in August, and wane through October. The timing of the improvement in skill implies that the thinner SITs at the time of initialization improve skill in the Kara Sea in June and July (during melting) and that the subsequent warmer SSTs lead to enhanced skill (during freezing) in October.

5. Discussion and conclusions

In this study, we employed hindcasts from CanCM3 to assess the influence of different methods for initializing SIT on the skill of Arctic sea ice area and concentration. The climatological ice thickness initialization employed in CanSIPS, denoted Original, and several more accurate methods were considered, including statistical models SMv1–SMv3 designed to be usable in a real-time operational setting as well as PIOMAS. These statistical models differ in their ability to capture interannual variability in the PIOMAS SIT time series used here as a benchmark, but are similar with respect to being able to capture the long-term trend. Using only one predictor field (SIC) and one model parameter, the statistical model SMv3 performs comparably to the more complicated model SMv2, with both outperforming SMv1.

Evaluating hindcast skill over 1981–2012, improvement in Arctic sea ice prediction was found using all improved SIT-IMs for nearly all measures of skill. When skill is
assessed including the long-term trend, both integrated SIA and regional SIC hindcast skills were shown to strongly rely on the accuracy of the SIT initialization. We found a large sensitivity to interannual SIA skill to the detrending method used, and we argue that detrending using a quadratic fit is more appropriate than using a linear fit over the time period considered. Arctic SIA was found to be more sensitive to the ice thickness initialization than regional SIC when considering the skill of interannual SIA variations around the trend based on linear detrending. However, assessing interannual skill of SIA based on quadratic detrending produced skill for SIA that is more qualitatively similar to regional SIC; that is, greater interannual skill in SIA generally corresponded to greater interannual skill in the areal fraction of significant skill in SIC. This result provides further support for the choice of assessing interannual skill for SIA based on quadratic detrending.

Hindcast SIA anomalies including the trend were found to have statistically significant skill out to six months using all improved SIT-IMs, with largest improvements relative to Original seen in the summer months when initialized in winter and early spring. Using SMv2, SMv3, or PIOMAS, significant skill of quadratically detrended September SIA anomalies can be achieved initializing in May, extending the time that significant skill can be obtained by at least one month relative to using Original. Additionally, unlike hindcasts using Original, linearly detrended August SIA anomalies were found to be predictable with significant skill from March using all improved SIT-IMs. Although this latter result is likely a consequence of the incomplete removal of the trend, it is relevant for comparison against other studies that have used linear detrending to show that statistically significant predictions of September SIA/SIE can be made from March. With these improvements in skill during the summer months, CanCM3 skill for detrended SIA anomalies becomes more comparable to that found in other studies (Chevallier et al. 2013; Wang et al. 2013; Msadek et al. 2014; Collow et al. 2015). Although a direct comparison to these studies is not possible because of differing temporal coverage, the ACC values 0.52–0.65 obtained here for linearly detrended September SIA anomalies when initializing in May are comparable to those ranging from 0.4 to 0.6 found in Chevallier et al. (2013), Wang et al. (2013), and Msadek et al. (2014) for the same initialization and target month. Skill reemergence for detrended SIA anomalies was found for winter target months for hindcasts initialized in late summer, independent of the SIT-IM used (as in Day et al. 2014) as well as the detrending technique used.

ACC-based regional skill of SIC (with the trend included) was found to be substantially improved using the more accurate SIT-IMs. Large improvements across both the eastern and western Arctic in the summer
months result from initializing with the improved SIT-IMs compared with Original. Although the predictive skill of SIC for hindcasts initialized by the improved SIT-IMs is reduced substantially when the trend is removed, hindcasts of September SIC are more skillful than Original using SMv2, SMv3, and PIOMAS, specifically in the Beaufort Sea (where some skill is also present using Original) and the eastern central Arctic extending into the Kara Sea. In October, the region of improved skill grows to include all of the Kara Sea.

Given these results, it is clear that similar hindcast skill to that which is obtained initializing with PIOMAS can also be achieved initializing with the statistical models SMv2 or SMv3. Although the SIC ACC-based skill results for May initialization show that SMv2 hindcasts outperform SMv3 hindcasts, these results should not be overinterpreted since skill for SMv2 initialized hindcasts is also greater than PIOMAS initialized hindcasts. Because the statistical models are trained with PIOMAS as the predictand, any improvement in hindcast skill relative to PIOMAS hindcasts seems likely to be the result of sampling. Additionally, SMv3 is substantially simpler than SMv2. Not only does this simplicity hold practical value in terms of parsimony, the simpler statistical model reduces the chance of gaining skill from overfitting. Further, because the only information that SMv3 requires from PIOMAS is the estimate of the local SIT trend over the training period, if PIOMAS were to become unavailable at some point in the future, SMv3 could potentially be constructed with another reanalysis product that adequately represents these trends. For operational purposes, we therefore recommend the use of SMv3 for SIT initialization.

Considering the RMSE as an error metric of SIC brought attention to an unexpected result of using a more accurate SIT initialization within CanCM3. In the region extending from the northern Greenland Sea into the Barents Sea, negative SST biases in CanCM3 were found to be amplified by the thicker initial ice cover specified by the improved SIT-IMs in this region. It should be noted that CanCM4 shows a SIC bias of the opposite sign in this region in freely running September hindcasts (Merryfield et al. 2013a), suggesting that the use of the multimodel approach used in CanSIPS could mitigate this source of error. Finally, as more accurate initial conditions should always be preferred, this result further motivates efforts to reduce such model biases. As another example, Collow et al. (2015) found that considerable improvement in seasonal sea ice predictions could be achieved not only by improving SIT initial conditions but also by reducing a SST bias within the Climate Forecast System, version 2 (CFSv2).

Limitations on observational knowledge of sea ice thickness (both in real time and historically) pose a challenge to operational forecasting centers that wish to produce seasonal forecasts of sea ice. However, as demand is growing for such forecasts, efforts are needed to circumvent this limitation. By applying relatively simple statistical model techniques, SIT initialization can be improved substantially compared with a climatological initialization. Robust improvements in predictive skills of SIA and regional SIC result from using improved ice thickness initializations in CanCM3.

Acknowledgments. We thank Michael Sigmond and Slava Kharin, in addition to two anonymous reviewers, for their helpful comments on earlier versions of this manuscript. AD would like to thank Woo-Sung Lee for producing the hindcasts used in this work. AD and WM acknowledge funding from the Canadian Sea Ice and Snow Evolution Network (CanSISE). AM acknowledges support from the Natural Sciences and Engineering Research Council of Canada. PIOMAS data used in this study are available at ftp://pscftp.apl.washington.edu/zhang/PIOMAS/data/v2.1. NSIDC data used in this study are available at ftp://sidads.colorado.edu/pub/DATASETS/NOAA/G02202_v2/north.

REFERENCES


