Origins and Levels of Seasonal Forecast Skill for Sea Ice in Hudson Bay Using Canonical Correlation Analysis

ADRIENNE TIVY
Foothills Climate Analysis Facility, Centre for Alpine and Arctic Climate Research, Department of Geography, University of Calgary, Calgary, Alberta, Canada, and International Arctic Research Center, University of Alaska Fairbanks, Fairbanks, Alaska

STEPHEN E. L. HOWELL
Climate Processes Section, Atmospheric Science and Technology Directorate, Environment Canada, Toronto, Ontario, Canada

BEA ALT
Balanced Environmental Associates, Ottawa, Ontario, Canada

JOHN J. YACKEL
Foothills Climate Analysis Facility, Centre for Alpine and Arctic Climate Research, Department of Geography, University of Calgary, Calgary, Alberta, Canada

THOMAS CARRIERES
Canadian Ice Service, Meteorological Service of Canada, Environment Canada, Ottawa, Ontario, Canada

(Manuscript received 16 November 2009, in final form 3 August 2010)

ABSTRACT

Canonical correlation analysis (CCA) is used to estimate the levels and sources of seasonal forecast skill for July ice concentration in Hudson Bay over the 1971–2005 period. July is an important transition month in the seasonal cycle of sea ice in Hudson Bay because it is the month when the sea ice clears enough to allow the first passage of ships to the Port of Churchill. Sea surface temperature (quasi global, North Atlantic, and North Pacific), Northern Hemisphere 500-mb geopotential height \( z_{500} \), sea level pressure (SLP), and regional surface air temperature (SAT) are tested as predictors at 3-, 6-, and 9-month lead times. The model with the highest skill has three predictors—fall North Atlantic SST, fall \( z_{500} \), and fall SAT—and significant tercile forecast skill covering 61% of the Hudson Bay region.

The highest skill for a single-predictor model is from fall North Atlantic SST (6-month lead). Fall SST explains 69% of the variance in July ice concentration in Hudson Bay and a possible atmospheric link that accounts for the lagged relationship is presented. CCA diagnostics suggest that changes in the subpolar North Atlantic gyre and the Atlantic multidecadal oscillation (AMO), reflected in sea surface temperature, precede a deepening/weakening of the winter upper-air ridge northwest of Hudson Bay. Changes in the height of the ridge are reflected in the strength of the winter northwesterly winds over Hudson Bay that have a direct impact on the winter ice thickness distribution; anomalies in winter ice severity are later reflected in the pattern and timing of spring breakup. July ice concentration in Hudson Bay has declined by approximately 20% per decade between 1979 and 2007, and the hypothesized link to the AMO may help explain this significant loss of ice.

1. Introduction

September ice extent in the Arctic has declined by approximately 10% decade\(^{-1}\) between 1979 and 2007 (e.g., Comiso et al. 2008). This dramatic loss of ice has fueled
interest in the seasonal predictability of sea ice, both regionally and Arctic wide, as a means to isolate the main internal and external climate forcings. Regionally, with the exception of the Bering Sea (Parkinson and Cavalieri 2008), the greatest reductions have occurred in Hudson Bay (Parkinson and Cavalieri 2008). Over the same period, looking at July, which is the month before the region becomes essentially ice free and the shipping season begins, sea ice declined by approximately 20% decade$^{-1}$ (Stewart et al. 2010). July is an important month in the Hudson Bay region since it is the month before the region becomes essentially ice free and the shipping season begins. Exacerbating the need for sea ice forecasts in this region is the potential of the Arctic Sea Bridge (Murmansk, Russia, to Churchill, Canada) becoming a major commercial shipping route.

Spring breakup in Hudson Bay starts in May when leads, areas of open water within the ice pack, form in the northwestern part of the bay. Dynamic- and thermodynamic-driven melt proceeds from the northwest and the east with the ice along the southwest coast clearing last. In Hudson Strait, breakup is also initiated by the presence of open water as leads form along the northern coast, causing the ice to clear from northwest to southeast, with Ungava Bay clearing last (CIS 2002; Fig. 1a). In a still-relevant review of the ice regime in Hudson Bay, Markham (1986) attributes the first clearing of ice in northwestern Hudson Bay to the mean northerly winds throughout the winter, which keep the ice thin and drifting offshore and which continue into spring, widening the leads and initiating melt. The clearing in western Hudson Strait is also attributed to the winter and spring wind forcing of leads but is combined with tides that ultimately flush the ice out into the Labrador Sea. Clearing in eastern Hudson Bay is initiated by spring runoff and proceeds relatively quickly since this is an area of fast ice that is generally thinner than the heavily deformed ice in the center of the bay (Markham 1986).

Factors that influence the timing and pattern of spring breakup in Hudson Bay include local wind forcing, tides, river runoff, air temperature, snow depth on the ice, and the ocean heat flux. The importance of the local wind forcing on spring sea ice variability has been identified in both observations (Wang et al. 1994b) and modeling studies (Wang et al. 1994a; Saucier and Dionne 1998), where strengthened northerly winds are associated with heavy ice conditions. A positive correlation between river runoff and sea ice extent at a 1-yr lag is identified in Wang et al. (1994a) and is supported by the modeling study of Saucier and Dionne (1998), who show an increase in modeled ice thickness in southeast Hudson Bay due to a freshening of the mixed layer caused by increased river runoff. Saucier and Dionne also show a decrease in the winter ice volume by more than 20% in response to a temperature increase of 2°C. Using the same model, Joly et al. (2011) show a decrease in the winter ice season by 7–9 weeks with a mean warming of approximately 4°C. The sensitivity of sea ice cover to temperature is supported by observations where Etkin (1991) find a significant negative correlation between the timing of breakup and cumulative melting degree days. The importance of snow depth on the sea ice is highlighted in Gough et al. (2004). In their study of ice thickness at coastal stations around Hudson Bay, Gough et al. found that snow depth on the ice contributed more to the interannual variability in ice thickness than air temperature. The influence of the ocean heat flux on ice severity is explored indirectly in Gough and Houser (2005), where a link is identified between the length of the ice-free season and summer air temperature to the timing of freeze up in Hudson Strait.

**Fig. 1.** (a) July mean ice concentration in tenths ice coverage and (b) standard deviation in July mean ice concentration in tenths ice coverage. The mean opening date for the shipping route to Churchill, 1972–2004, is 25 Jul, and the four choke points where ice clears last along the route are outlined in black.
Interannual variability in spring ice conditions has also been linked to large-scale climate variability. Mysak et al. (1996) and Wang et al. (1994b) show greater than normal spring ice extent in Hudson Bay in response to coincident positive North Atlantic Oscillation and El Niño events. A deepening of the Icelandic low (positive NAO) during winter is associated with negative winter temperature anomalies and enhanced northerly winds over eastern Canada, including Hudson Bay (Agnew 1993; Wang et al. 1994b). The Pacific–North America (PNA) teleconnection pattern (Wallace and Gutzler 1981), which is often well developed during strong El Niño episodes, has an intensified center of high pressure over western Canada, resulting in enhanced northerly winds over Hudson Bay (Wang et al. 1994b). Colder than normal air temperatures and stronger than normal northerly winds promote ice thermodynamic and dynamic ice growth in Hudson Bay. Tivy et al. (2007) link variability in the opening date of the shipping route to Churchill with low frequency variability in the North Atlantic. Evidence of both decadal and multidecadal variability in ice anomalies are shown in Mysak and Venegas (1998) and Venegas and Mysak (2000), although Hudson Bay is not discussed directly in these studies.

Canonical correlation analysis (CCA) is a multivariate statistical technique that measures the linear relationship between two or more multidimensional variables. The technique is slightly different from the more commonly used singular value decomposition (SVD) since, instead of maximizing covariance, cross-correlation is maximized (Von Storch and Zwiers 1999). In climate research, CCA is used to identify coupled patterns in climate data (e.g., Bretherton et al. 1992), and a variant of CCA developed by Barnett and Preisendorfer (1987, hereafter BP) is the most reliable and most commonly used statistical seasonal forecasting technique (Zwiers and Von Storch 2004). Operationally, CCA models based on the BP methodology are used for generating seasonal precipitation and temperature forecasts in North America (Barnston and Barnston 1996) and other parts of the world—Africa (Thiaw et al. 1999; Landman and Mason 1999), Korea (Hwang et al. 2001), Europe (Johansson et al. 1998), the tropical Pacific Islands (He and Barnston 1996)—and worldwide (Barnston and Smith 1996). In all of these studies predictability stems from low frequency variability in the ocean. While dynamic models are state of the art for short- and medium-range predictions of the state of the atmosphere, at lead times greater than 3 months, statistical models generally outperform dynamic models (Zwiers and Von Storch 2004).

Compared to the atmosphere, seasonal sea ice forecasting is in its infancy. While the benefits of coupling sea ice and ocean models to operational numerical weather prediction models are starting to be realized (e.g., Pellerin et al. 2004) and while this is an area of active research, the focus to date has been on short-term forecasting. Climate models that include atmosphere–sea ice–ocean coupling are also starting to be used for seasonal predictions, but their skill for sea ice forecasting is quite limited to date. Statistical models have been developed for seasonal sea ice forecasting, but they have focused solely on the prediction of scalar sea ice parameters: seasonal and regional ice extent (Barnett 1980; Johnson et al. 1985; Chapman and Walsh 1991; Droboth et al. 2006, Droboth 2007; Lindsay et al. 2008), the Beaufort severity index (Droboth and Maslanik 2002; Droboth 2003), the opening date of the shipping season in Hudson Bay (Tivy et al. 2007), and the length of the ice-free and ice-covered seasons in Hudson Strait (Gough and Houser 2005). These models all rely on multiple linear regression techniques that relate sea ice severity to antecedent climate, and the skill of the models is most sensitive to lead time and whether the long-term trend is removed from the data prior to model development.

With a focus on Hudson Bay, this study expands on previous work by assessing the empirical predictability of both spatial and interannual variability in early-summer sea ice concentration at 3-, 6-, and 9-month lead times. CCA is used both as a forecasting tool and a diagnostic tool to explore potential dynamic pathways in the climate system that link variability in atmospheric and oceanic predictors to variability in sea ice. It is hoped that results from this analysis will offer insight into the recent dramatic decline in early summer sea ice concentration in Hudson Bay. In the next two sections the datasets used in this study are presented, the rationale for the predictors chosen for the experiment are discussed, and the CCA methodology is described. Results from each predictor experiment are presented and discussed in section 4, followed by a summary of the main findings in section 5.

2. Data

Sea ice data on a 0.25° grid is obtained from the Canadian Ice Service Digital Archive (CISDA) regional charts database (CIS 2007a,b). The dataset spans from 1969 to present and is a compilation of the weekly regional ice charts that are an integration of ice information from various satellite sensors, aerial reconnaissance, ship observations, and other in situ measurements. This dataset is selected over the passive-microwave-derived record since, as Agnew and Howell (2002) have shown, passive-microwave-derived ice concentration can underestimate ice area by approximately 30% during spring melt in marginal sea ice regions such as Hudson Bay.
Monthly sea surface temperature (SST) data are from the Met Office Hadley Centre Sea Ice and Sea Surface Temperature version 1.1 (HadISST1.1) dataset (Rayner et al. 2003) on a 1° grid for the period 1870 to present. Monthly 500-mb geopotential height ($z_{500}$), surface air temperature (SAT), sea level pressure (SLP), and surface vectors winds all on a 2.5° grid are from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis dataset (Kalnay et al. 1996). Monthly sea level fields on a 1° grid for the period 1993–2007 are combined Ocean Topography Experiment (TOPEX)/Poseidon and Jason-1 data with the inverse barometric effect removed.

3. Methods

a. Choice of predictand

The predictand is July average ice concentration from 1971 to 2005 in Hudson Bay and Hudson Strait (Fig. 1). July is the main transition month between complete ice coverage in winter and open water in summer (Fig. 2), and ice conditions in July largely determine the start of the shipping season. The mean opening date for the shipping route to Churchill is 25 July (Tivy et al. 2007), and an approximation of the shipping route through Hudson Strait and Hudson Bay is shown (Fig. 1a) along with the four regions where ice clears last along the route (Tivy et al. 2007). The mean pattern in July ice concentration (Fig. 1a) shows that the ice clears first in northwestern Hudson Bay, east Hudson Bay, and Hudson Strait; the largest variability (Fig. 1b) is along the seaward margin.

b. Choice of predictors

Four climate variables are tested as potential predictors: 3-month-averaged SST, $z_{500}$, SAT, and SLP. SST is chosen because it varies over time much more slowly than the atmosphere and is the main source of predictability in seasonal forecast models for the atmosphere (e.g., Barnston 1994; Shabbar and Barnston 1996) and sea ice (Lindsay et al. 2008). North Atlantic (SST-NA; 20°–70°N, 75°W–10°E), North Pacific (SST-NP; 20°–70°N, 105°E–120°W), and quasi-global (SST-LG; 50°S–70°N, 360° azimuth) SST are all tested as potential predictors. The SST predictor domain is restricted to the North Atlantic and North Pacific to ensure that, after prefiltering (see section 3c), variability in tropical Pacific SST does not mask the main modes of low frequency variability in the North Atlantic (e.g., North Atlantic tripole) and North Pacific (e.g., Pacific decadal oscillation), which may influence weather patterns over the Hudson Bay region. Northern Hemisphere SLP and $z_{500}$ are included since they are the variables for which preferred modes of low-frequency atmospheric variability and teleconnection patterns are identified (e.g., Wallace and Gutzler 1981) and monitored (e.g., available online at http://www.cpc.noaa.gov/data/teledeo/telecontents.shtml). Local SAT over Hudson Bay (50°–60°N, 265°–295°E) is included since it is an important factor controlling sea ice thickness. Other variables that likely influence interannual variability—such as local wind forcing, snow depth on the ice, and the ocean heat flux—are not included owing to the lack of comprehensive datasets.

c. CCA methodology

CCA is a variant of a multiple linear regression analysis that is used in climate research to determine the linear relationship between variables that vary in space as well as in time. The objective of CCA analysis is to find correlated patterns and lead–lag relationships between the predictor and predictand fields. The end result of CCA is a series of modes for which each mode is a pair of predictor and predictand patterns (canonical component vectors), each with a time series (canonical component time series) that describes how each pattern varies in time. The mathematics is similar to singular value decomposition (SVD) analysis with the exception that, instead of maximizing the covariance between the two variables, the correlation is maximized. The correlation between the two canonical component time series decreases with each successive mode and like SVD or EOF analysis, the modes are uncorrelated. Because CCA is a regression technique, the standard assumptions of linear regression, that is, independence and constant variance of the errors, normality of the distribution of errors, and linearity, all apply.
Prior to CCA a linear trend is removed at each grid point in the predictand and predictor data. This is done to reduce the likelihood that the CCA-modeled predictor–predictand relationship is due to a common external forcing, namely, the change in radiative forcing due to increased greenhouse gases in the atmosphere. It is acknowledged that there may be some leakage, as the true global warming signal is nonlinear (Solomon et al. 2007) and this is not addressed.

After detrending, the general CCA methodology follows BP; the mathematics of the methodology is presented in the appendix of BP. The predictor and predictand data are standardized and prefiltered using principal component (PC) analysis. The PCs are truncated to reduce the amount of noise in the predictor and predictand datasets prior to CCA. The number of PC and CCA modes retained in the final model is chosen by maximizing the field-averaged “leave one out” cross-validated correlation coefficient between the observed and modeled data. A leave-one-out cross-validation scheme is commonly used to prevent overfitting (Michaelson 1987; Elsner and Schmertmann 1994) and, as in BP, the use of the technique is justified since after detrending only 3% of the grid points have significant autocorrelations at 1-yr lag.

The PC modes are truncated to reduce the amount of noise in the predictor and predictand datasets prior to CCA. The choice of PC truncation points in this study differs from BP and other studies (e.g., Barnston 1994; Johansson et al. 1998) in which the delineation between signal and noise was determined using Monte Carlo techniques. Livezey and Smith (1999) caution that the choice of PC truncation in the BP methodology can have a profound impact on the CCA results and that often sample sizes are too small to determine the optimum truncation technique a priori. Results from their sensitivity study of the BP methodology to different PC truncation rules suggests that the optimal number of predictor modes is much larger, and the optimal number of predictand modes is much smaller than the result in BP. As a result, two simple methods for determining PC truncation points are used in this study. For the predictand, where Livezey and Smith (1999) found that the BP methodology retained too many modes, the very conservative North et al. (1982) criterion is used. For the predictor fields that Livezey and Smith (1999) found the BP methodology retained too few modes, a more liberal approach that uses screen plots to separate signal from noise is used (O’Lenic and Livezey 1988).

To generate the final model the linear trend that was removed is added back. For simplicity, CCA + trend refers to the detrended CCA result with the linear trend added back at each grid point. Skill is evaluated for both the detrended CCA result and the CCA + trend result so as to isolate the relative contribution to predictive skill of interannual covariability in the predictor and predictand and a persistent long-term trend in the predictand.

All models are trained on the 35-year record from 1971 to 2005. Models are generated for each predictor at lead times of 3, 6, and 9 months; seasonal averages for each predictor are January–March (JFM), October–December (OND), and July–September (JAS). To determine whether sources of skill identified in the single predictor experiments are unique to a specific predictor field or common to two or more predictor fields, mixed-predictor experiments are run. It is assumed that, if the skill of the mixed predictor experiment is higher than the skill of the individual predictor experiment, then the signals are unique.

d. Evaluation of skill

The field-averaged (global) cross-validated correlation between the observed and modeled data is used to evaluate model skill; it is the standard metric used in the literature to evaluate the skill of CCA-based forecasting models (e.g., Barnston 1994; Barnston and He 1996; Shabbar and Barnston 1996; Johansson et al. 1998; Hwang et al. 2001; Mo and Thiaw 2002; Cannon and Hsieh 2008). The significance of local skill (skill at a single grid point) is determined using a one-tailed$t$test of the null hypothesis that there is no skill. For a normally distributed dataset with a mean of 0 and standard deviations of 1, the$t$threshold for significance at the 95% confidence level is 1.645. For a normally distributed dataset with a mean of 0 and standard deviations of 5, the$t$threshold for significance at the 95% confidence level is 1.96. The$t$threshold for significance at the 95% confidence level using the Monte Carlo-based (1000 trials) field significance test outlined in Livezey and Chen (1983). For detrended July ice concentration the 95% threshold determined from the Monte Carlo run is 5.4%.

As a final check, significance of experiments that pass field significance are tested using a Monte Carlo simulation for which the predictand field is randomly shuffled 100 times and the percent of grid points with significant local skill recorded to determine a 95% confidence level. Although the field-averaged correlation coefficient is used to determine PC and CCA truncation points, it is a measure that is less intuitive than the percentage of grid points with significant local correlation. In the discussion, the percentage of grid points with significant local correlation is used to compare models.

Forecasts will be expressed categorically as above normal, near normal, and below normal. This is the standard in seasonal forecasting since there is an appreciable amount of uncertainty. Instead of ranking the data to determine the tercile boundaries (BP), the threshold is set at $\pm 0.43$ standard deviations from the mean, which is the convention used at the Canadian Meteorological Centre (Serfvanckx et al. 1999). In theory, both methods would yield the same limits for a normally distributed population.
Climatology and persistence are used as benchmark models. Persistence is the monthly averaged ice concentration anomaly at each grid from both the preceding July and the preceding November added to the mean July ice concentration. Complete ice coverage in winter prevents the use of anomaly persistence at the time of the forecast.

4. Results

Depending on the variable and lead time, initial PC truncation for the climate variables ranged between 2 and 14 and represented between 69% and 86% of the total variability in the original data. The climate truncation points are conservative compared to the results of the sensitivity analysis in Livezey and Smith (1999). PC truncation for July sea ice concentration was chosen at four modes, representing 60% of the original variance in the data. Tables 1–4 summarize the skills scores for the benchmark models and all predictor experiments at 3-, 6-, and 9-month lead times.

a. Forecast skill

1) PERSISTENCE AND CLIMATOLOGY

Table 1 shows skill results for the three benchmark models. Persistence skill scores were calculated using both raw and detrended monthly ice concentration data. With the long-term trend removed, there was no significant skill in forecasts made from July anomaly persistence, and the skill from November anomaly persistence is barely significant (5.9%). For global tercile skill (Table 1, column 3), persistence did not outperform climatology. The skill for persistence was not calculated in a cross-validated framework and is likely inflated. As a result, the low skill for November persistence is not considered significant, so the only benchmark model available is climatology (long-term mean), which has a tercile skill score of 33%.

2) NINE-MONTH LEAD TIME (JAS)

The results from the single predictor experiments run at a 9-month lead time, predicting detrended July ice concentration from the preceding summer (JAS), yielded no CCA models with a cross-validated correlation skill that passed the field significance test. For all models, the percentage of grid points with a significant correlation between the observed and modeled data was less than 5.4%.

3) SIX-MONTH LEAD TIME (OND)

The results from the single predictor experiments run at a 6-month lead time, predicting detrended July ice concentration from the preceding fall (OND), are summarized in Table 2. The fall forecasts had significant skill covering 9.4%, 15.6%, and 21.6% of the total area for SAT, $z_{500}$, and North Atlantic SST, respectively. Forecast skill for SLP, global SST, and North Pacific SST is not significant. Figure 3 shows the geographic distribution of skill for each predictor with significant forecast skill. All three predictors, particularly SST-NA, show some skill in Hudson Strait. In Hudson Bay, the skill for SST-NA and $z_{500}$ is in eastern Hudson Bay whereas the skill for SAT is in northwestern Hudson Bay. SST-NA has the greatest

### Table 1. Forecast skill for the three benchmark models.

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>Global cross-validated correlation; percentage grid points $r &gt; 0.28$</th>
<th>Global tercile skill score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>Detrended</td>
<td>Raw</td>
</tr>
<tr>
<td>Persistence</td>
<td>Not significant</td>
<td>0.100; 16.3</td>
</tr>
</tbody>
</table>

### Table 2. Forecast skill for 6-month lead (OND).

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Global cross-validated correlation; percentage grid points $r &gt; 0.28$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_{500}$</td>
<td>0.012; 15.6</td>
</tr>
<tr>
<td>SLP</td>
<td>Not significant</td>
</tr>
<tr>
<td>SAT</td>
<td>−0.026; 9.4</td>
</tr>
<tr>
<td>SST-GL</td>
<td>Not significant</td>
</tr>
<tr>
<td>SST-NA</td>
<td>0.105; 21.6</td>
</tr>
<tr>
<td>SST-NP</td>
<td>Not significant</td>
</tr>
</tbody>
</table>

Mixed predictor

<table>
<thead>
<tr>
<th>Predictor</th>
<th>CCA</th>
<th>CCA + trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>SST-NA, $z_{500}$</td>
<td>0.228; 38.3</td>
<td>0.344; 68.5</td>
</tr>
<tr>
<td>SST-NA, SAT</td>
<td>0.196; 28.7</td>
<td>0.307; 56.2</td>
</tr>
<tr>
<td>$z_{500}$, SAT</td>
<td>0.050; 14.8</td>
<td>0.231; 41.0</td>
</tr>
<tr>
<td>SST-NA, $z_{500}$, SAT</td>
<td>0.251; 44.6</td>
<td>0.346; 69.0</td>
</tr>
</tbody>
</table>

### Table 3. Forecast skill for 3-month lead (JFM).

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Global cross-validated correlation; percentage grid points $r &gt; 0.28$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_{500}$</td>
<td>Not significant</td>
</tr>
<tr>
<td>SLP</td>
<td>Not significant</td>
</tr>
<tr>
<td>SAT</td>
<td>0.056; 14.3</td>
</tr>
<tr>
<td>SST-GL</td>
<td>Not significant</td>
</tr>
<tr>
<td>SST-NA</td>
<td>−0.020; 5.9</td>
</tr>
<tr>
<td>SST-NP</td>
<td>Not significant</td>
</tr>
</tbody>
</table>

Mixed predictor

<table>
<thead>
<tr>
<th>Predictor</th>
<th>CCA</th>
<th>CCA + trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAT, SST-NA</td>
<td>0.118; 20.3</td>
<td>0.268; 45.0</td>
</tr>
<tr>
<td>SAT, SST-NA (OND)</td>
<td>0.133; 24.9</td>
<td>0.278; 49.5</td>
</tr>
<tr>
<td>SAT, $z_{500}$ (OND)</td>
<td>0.164; 23.4</td>
<td>0.300; 55.8</td>
</tr>
<tr>
<td>SAT, SST-NA (OND), $z_{500}$ (OND)</td>
<td>0.232; 39.0</td>
<td>0.335; 62.1</td>
</tr>
</tbody>
</table>
regional coverage of significant skill and the highest global cross-validated correlation, $r = 0.105$. Although there is significant regional overlap in skill between the three predictors, skill in northwestern Hudson Bay is unique to SAT and skill in southeastern Hudson Bay is unique to $z_{500}$. Mixed predictor experiment results are listed at the bottom of Table 2. The addition of $z_{500}$ and SAT to SST-NA forecasts raises the coverage of skill from 21.6% to 38.3% and 28.7%, respectively. Although the addition of SAT to $z_{500}$ forecasts did not raise the percent coverage, 15.6% to 14.8%, the cross-validated correlation skill increased from 0.012 to 0.05. The highest skill, 44.6% (0.251), is achieved by combining all three predictors. The geographic distribution of skill for the three predictor model (Fig. 3d) compared to the single predictor models highlights the fact that each predictor provides unique predictive information and that the best model is the combination of all three predictors.

4) THREE-MONTH LEAD TIME (JFM)

The results from the single predictor experiments run at a 3-month lead time—predicting detrended July ice concentration from the preceding winter—are summarized in Table 3. No significant skill was found in forecasts from $z_{500}$, SLP, near-global SST, and North Pacific SST. The skill from North Atlantic SST, which was the highest skill for a single predictor in fall, is barely significant (5.9%) and is the only significant model that does not

<table>
<thead>
<tr>
<th>Truncation points (% var. expl.)</th>
<th>Canonical correlation(^2) (% predictand var. expl.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC—July ice concentration</td>
<td>PC—predictor</td>
</tr>
<tr>
<td>SAT (OND)</td>
<td>4 (60)</td>
</tr>
<tr>
<td>SAT (JFM)</td>
<td>4 (60)</td>
</tr>
<tr>
<td>SST-NA (OND)</td>
<td>4 (60)</td>
</tr>
<tr>
<td>$z_{500}$ (OND)</td>
<td>4 (60)</td>
</tr>
</tbody>
</table>

**Fig. 3.** Geographic distribution of CCA skill (cross-validated correlation) for fall forecasts from (a) SST-NA, (b) local surface air temperature, (c) Northern Hemisphere $z_{500}$, and (d) mixed: SST-NA, $z_{500}$, and SAT.
outperform November persistence. The skill for SAT forecasts increased slightly from 9.4% \((\pm 0.026)\) to 14.3% \((\pm 0.056)\). The geographic distribution of skill (Fig. 4b) for SAT differs from fall (Fig. 3b). There is almost no skill in Hudson Strait, and in Hudson Bay the skill shifts from northwestern Hudson Bay to central Hudson Bay. For SST some skill remains in Hudson Strait, but there is no skill in Hudson Bay (Fig. 4a). An increase in skill from combining both predictors (Table 3) suggests each predictor adds unique predictive information.

To determine whether there is value in including winter information in the fall model, the winter values of SST-NA and \(z_{500}\) are replaced with the fall values. The geographic distribution of skill for the new three predictor model—SST-NA (OND), \(z_{500}\) (OND), and SAT (JFM)—is shown in Fig. 4c. Compared to the original fall model, skill coverage is reduced in northwestern Hudson Bay and eastern Hudson Bay. Also, forecast skill (Table 3) for the new model, 39% \((\pm 0.234)\), is less than the original fall model, 44.6% \((\pm 0.251)\). It is concluded that there is little value in replacing fall SAT with winter SAT.

5) CCA + TREND: OPERATIONAL MODEL

To generate the final forecast, the long-term trend is added to the CCA prediction. Skill in all models improves dramatically with the addition of the long-term trend (Tables 2 and 3). The relative ranking of the models does not change, however, and all of the model skill scores outperform the persistence benchmark models.

The model proposed to be used operationally is the fall three-predictor model. With the addition of the long-term trend forecast skill increases from 44.6% \((\pm 0.251)\) to 69% \((\pm 0.346)\). Figure 5 is a comparison of the geographical distribution of skill for CCA and CCA + trend. The regions with greatest skill coincide with the regions with the most variability (Fig. 1b). The main operational consideration is the shipping route to Churchill, and the model has significant skill covering most choke point areas along the route (Fig. 1a). The geographic distribution of tercile skill (not shown) is similar to the correlation skill shown in Fig. 5b. The model outperforms climatology (Table 1), there is significant tercile skill covering 61% of the Hudson Bay region, and the tercile skill at a given location ranges from 40% to 80%.

b. Source of predictability: Interannual variability

The final forecast model was a combination of three predictors: SST-NA (OND), SAT (OND or JFM), and \(z_{500}\) (OND). Canonical correlation maps and time series are used to explore sources of predictive skill for each predictor. Table 4 summarizes final PC and CCA truncation points for each predictor. The squared canonical correlations are also listed for each CCA mode in each experiment along with the percentage of the original variance...
in the predictand explained by each CCA mode. The canonical correlation is the correlation between the predictor and predictand canonical correlation time series for a given CCA mode. Because the squared canonical correlations are the eigenvalues of the cross-correlation matrix, they can be used to calculate the percentage of the original variance in the predictand that is explained by each CCA mode. For example, the first CCA mode (CCA1) from the fall SAT predictor experiment explains 38% of the original variance in the predictand \([60\% \times 0.47 / (0.47 + 0.28)]\), and CCA1 from the fall SST predictor experiment explains 31% of the original variance in the predictand \([60\% \times 0.67 / (0.67 + 0.48 + 0.12 + 0.02)]\). Recall that only 60% of the original variance in the July ice concentration data was used in the CCA analysis since 40% was removed when the number of principal components retained for the analysis was truncated at four.

1) FALL AND WINTER SAT EXPERIMENTS

In the winter SAT experiment CCA truncation was at two modes. CCA1 and CCA2 explain 39% and 21%, respectively, of the original variance in the predictand. The canonical correlation maps and time series for each mode are shown in Figs. 6 and 7. Correlations are significant throughout the SAT and ice concentration fields for CCA1; the signs are opposite between the two fields, suggesting a simple thermodynamic link. Warmer air temperatures in winter will slow ice growth, resulting in a thinner ice cover. In spring, less energy will be required to transition to open water. CCA2 is similar—correlations are opposite between the two fields, suggesting a thermodynamic link. The canonical correlation time series (Fig. 7c) exhibits strong interannual variability whereas the time series for CCA1 is much smoother, suggesting that a decadal or multidecadal component to the variability may be present. Significance in both SAT and ice concentration are for the most part restricted to Hudson Bay.

Results for fall are similar to winter. CCA truncation was at two modes; CCA1 and CCA2 explain 38% and 22%, respectively, of the original variance in the predictand. Canonical correlation maps and correlation time series are similar to winter, as is their interpretation.

2) FALL Z500 EXPERIMENTS

For the fall \(z_{500}\) predictor CCA truncation was at three modes; CCA1, CCA2, and CCA3 explain 39%, 12%, and 8%, respectively, of the original variance in the predictand. CCA2 and CCA3 are omitted from the discussion; because they explain so little of the original variance in July ice concentration, it is assumed that their contribution to forecast skill is minimal compared to CCA1.

The CCA1 predictor and predictand patterns are shown in Fig. 8. The pattern in \(z_{500}\) resembles the Pacific–North American teleconnection pattern. The PNA along with the NAO (or Arctic Oscillation) are the dominant modes of low frequency variability in the extratropical Northern Hemisphere, and both have been linked to extreme ice conditions in Hudson Bay (Wang et al.1994b; Mysak et al. 1996). The region of significant positive correlation over Hudson Bay is accompanied by a second region of significant positive correlation over the tropical Pacific and two regions of significant negative correlation over the North Pacific and the tropical Atlantic. The centers of action are shifted compared to the fall and winter PNA patterns. To test the relationship, the \(z_{500}\) canonical correlation time series is correlated with the fall PNA index. The correlation between the two time series was not significant at 95%. The \(z_{500}\) predictor pattern has six regions of significant correlation; to test if there is any relationship to large-scale climate, the canonical correlation time series is correlated with fall indices of the NAO, the east
Atlantic pattern (EA), the east Atlantic/North Pacific pattern (EP/NP), the west Pacific pattern (WP), the east Atlantic/west Russia pattern (EA/WR), the Scandinavian pattern (SCA), and the tropical/Northern Hemisphere pattern (TNH). Monthly time series for these teleconnection patterns are available from the National Oceanic and Atmospheric Administration Climate Prediction Center (CPC) at their Web site (available at http://www.cpc.noaa.gov/data/teledoc/telecontents.shtml). The patterns are calculated using rotated principle component analysis (RPCA); therefore, the times series associated with the patterns in any given month are independent of each other. Significant correlations were found with the WP ($r = 0.45$), the EP/NP ($r = 0.36$), the EA/WR ($r = 0.58$), and the TNH ($r = 0.43$). Of the four modes, only the TNH has a strong center of action in the Hudson Bay region. The relative dominance of these four teleconnections changes year to year, which may explain why there is a weak relationship to all four patterns and not a strong relationship to one. The lack of a significant correlation with the PNA or NAO suggests that only in certain cases—specifically, coincident positive NAO and strong El Niño (enhanced PNA) events as identified in Wang et al. (1994b) and Mysak et al. (1996)—is there a link between the phase and strength of these two teleconnections and sea ice variability.

The strongest correlations in the predictor pattern are over Hudson Bay, and it is concluded from correlation analysis that the relationship to large-scale teleconnection patterns is mixed. Because there was no predictive skill in forecasts made from winter $z_{500}$ and because persistence in the atmosphere is low, it is assumed that the relationship between height anomalies and sea ice is coincident in
fall and that the fall-to-summer memory is in the ice. In general, positive height anomalies are associated with a warmer atmospheric column. The opposite sign in correlation over Hudson Bay between the predictor and predictand could be a thermodynamic response of the ice to a warmer or colder atmosphere. However, the greater predictive skill of $z_{500}$ over SAT in fall and the dipole in the predictand correlation pattern between northwestern Hudson Bay and eastern Hudson Bay (Fig. 8) both suggest a dynamic component to the forcing. A dominant feature in fall Northern Hemisphere 500-mb geopotential heights (Fig. 9) is a ridge west of Hudson Bay. This ridge is the cause of the climatologically dominant northwesterly winds over Hudson Bay. The most extreme year in both canonical correlation time series, 1982, is characterized by high heights over Hudson Bay and less (more) ice concentration in eastern (northwestern) Hudson Bay. Anomalously high heights over Hudson Bay will reduce the gradient of the ridge and weaken the northwesterly winds. The result is more ice in northwestern Hudson Bay because the shore leads will not be maintained by the winds and less ice in eastern Hudson Bay because of reduced advection of cold air.

3) FALL NORTH ATLANTIC SST EXPERIMENTS

In the fall North Atlantic SST experiment PC truncation is at two modes: CCA1 represents 31% of the original variance in the sea ice data and CCA2 represents 22%. The discussion focuses on CCA1, which is responsible for the most predictive skill.

The leading CCA mode is shown in Fig. 10. The canonical correlation pattern in SST is characterized by a strong center south of Greenland over the subpolar gyre and a weaker center of opposite sign along the path of the Gulf Stream. This pattern strongly resembles the first EOF of monthly North Atlantic sea surface height (SSH) (Häkkinen and Rhines 2004) that is used as an index to represent the strength of the subpolar gyre (Häkkinen and Rhines 2004; Häutun et al. 2005). The EOF analysis in Häkkinen and Rhines is repeated using combined TOPEX/Poseidon and Jason-1 data. The first EOF (not shown) is similar to Häkkinen and Rhines over the period in common, 1993–2004, and the correlation coefficient between PC1 and CCA1 SST is 0.62

---

1 No significant correlation was found between CCA1 and the annual or seasonal index of the Atlantic tripole (e.g., Deser et al. 1997).
available online at www.cdc.noaa.gov/Correlation/atltri.data). This suggests that positive (negative) SST anomalies south of Greenland in CCA1 may be related to a weakening (strengthening) of the subpolar gyre. The sea ice pattern is dominated by a dipole between northwestern Hudson Bay/Hudson Strait and eastern Hudson Bay.

Is there a physical link that might explain these patterns of covariability? The potential of a coincident atmospheric forcing, direct ocean forcing, or teleconnection through the atmosphere is addressed. Because of the low skill in the November anomaly persistence, it is suspected that the contribution of a coincident atmospheric or oceanic forcing to the overall skill is small. This is supported by the lower skill of fall $z_{500}$ and SAT predictor models (Table 2). The potential for a delayed ocean forcing in winter and spring is difficult to explore owing to the lack of observations of ocean temperature below the ice, but its influence is also suspected to be small.

To explore the potential of a teleconnection through the atmosphere, the predictor canonical correlation time series is correlated with fall, winter, and spring SAT, $z_{500}$, 850-mb geopotential height ($z_{850}$), and SLP. A summary of the results is shown in Figs. 11 and 12. Moderate correlations ($0.3 < r < 0.5$) with SAT over Hudson Bay and to the northwest of the bay (Fig. 11) in fall and winter suggest a potential thermodynamic forcing that could be enhanced by advection due to the dominant northwesterly winds. No significant correlations were found with SLP in any season, and the only significant correlations with $z_{500}$ and $z_{850}$ were in winter. The apparent atmospheric response in winter is worth exploring further since the pattern in $z_{500}$ (Fig. 12), season, and lead time are all consistent with a series of empirical (e.g., Wen et al. 2005) and modeling (e.g., Kushnir et al. 2002) studies that suggest an early winter NAO-like atmospheric response to a late summer horseshoelike pattern of SST anomalies in the North Atlantic. The question addressed here is how this winter atmospheric response could impact ice concentration in Hudson Bay. In Fig. 12 where the contoured mean winter pattern in $z_{850}$ is superimposed on the correlation map, a region of significant correlations northwest of Hudson Bay sits directly underneath a ridge. A positive (negative) anomaly in the ridge height would
strengthen (weaken) the predominant northerly flow over Hudson Bay. This is supported by composites of meridional surface wind anomalies during winter (Fig. 13) for extreme negative and positive years in the predictor time series. The composites clearly show anomalous southerly and northerly flows. In the context of the main winter and spring processes in Hudson Bay, enhanced (weakened) northerlies would keep ice to the north thin by maintaining the flaw lead and the increased advection of cold air would thicken the fast ice in eastern Hudson Bay. In spring thinner ice to the north (thicker ice to the southeast) would promote early (later) clearing (Fig. 10).

c. Source of predictability: Multidecadal variability

In light of a potential physical mechanism explaining the skill of the fall North Atlantic SST model, CCA is rerun without removing the long-term trend to explore a potential decadal or multidecadal link to the North Atlantic SST. Principal component truncation and the number of PC and CCA modes retained in the final model were the same as for the detrended experiment. The global cross-validated correlation is similar to the detrended experiment ($r = 0.263$), and the first two CCA modes explain 39% and 21%, respectively, of the original variance. The first canonical correlation pattern in sea ice
(Fig. 14) is very similar to the detrended experiment (Fig. 10). The pattern in SST (Fig. 14) is all one sign but strongest south of Greenland and in the midlatitudes, similar to the detrended experiment (Fig. 10). The SST canonical correlation time series (Fig. 14) is dominated by a “jump” in the mid-1990s, the timing of which coincides with a change in phase of the Atlantic multidecadal oscillation (AMO) (Kerr 2000) from cold to warm (e.g., Enfield et al. 2001). The AMO time series is plotted with the July averaged ice concentration and SST canonical correlation time series (Fig. 15). The main feature in all three time series is a jump in the mid-1990s. The Rodionov regime shift detector (Rodionov 2004), which requires no a priori assumptions about the timing of a shift in a time series, is run on all three time series using cutoff lengths ranging between 15 and 25 yr. A Huber’s weight parameter of two and prewhitening are selected (Rodionov 2006). A significant shift in 1995 is detected in the SST canonical correlation time series and the AMO. Depending on the cutoff length, significant shifts in either 1993 or 1997 are detected for July ice concentration.

5. Discussion and conclusions

The origins and levels of seasonal forecast skill for July averaged ice concentration in Hudson Bay were determined using canonical correlation analysis at lead times of 3, 6, and 9 months. Seasonal forecasts made at a 6-month lead time had the highest skill, and there was little to no skill in 9-month, persistence, and climatology forecasts. At a 3-month lead time, only the skill of SAT
forecasts improved; the skill of all other predictors dropped considerably. The best model has three predictors—fall surface air temperature over Hudson Bay, fall North Atlantic sea surface temperatures, and fall Northern Hemisphere 500-mb geopotential height—and explains 53% of the interannual variability in July ice concentration. The model has significant tercile forecast skill covering 61% of the Hudson Bay region, and the tercile skill at a given location ranges from 40% to 80%.

CCA diagnostics were used to investigate the sources of predictive skill. The single predictor that yields the greatest skill is fall North Atlantic SST. A new teleconnection pathway is identified linking variability in the fall subpolar North Atlantic gyre and Atlantic multidecadal oscillation reflected in sea surface temperature anomalies to fall and winter atmospheric thermodynamic and dynamic forcing of sea ice that manifests as sea ice concentration anomalies in July. Specifically, fall SST anomalies are shown to precede a deepening or weakening of the winter upper-air ridge to the northwest of Hudson Bay; this has an effect on the strength of the northwesterly winds, which are known to influence the ice thickness distribution (e.g., Markham 1986; Mysak et al. 1996). Similarly, the physical mechanism leading to predictive skill from fall $z_{500}$ is a deepening of the winter trough over Hudson Bay, which again enhances the northwesterly winds that keep ice thin along the northern coast of Hudson Bay and Hudson Strait and advects cold air over the rest of the bay, promoting ice growth. Predictive skill from fall and winter SAT is likely due to a coincident thermodynamic forcing of sea ice: for the fall $z_{500}$ and SST predictors, the forced sea ice thickness anomalies in fall and winter are suspected to be carried over to July in the memory of the ice. In general, the higher skill of North Atlantic SST and $z_{500}$ predictors compared to SAT suggests that fall and winter sea ice dynamic processes may be more important than sea ice thermodynamic processes in preconditioning the state of the sea ice before breakup.

Predictive skill could potentially improve with the addition of new predictors. An estimate of the ice thickness would be the ultimate predictor. In addition to snow depth on the sea ice, the ocean heat flux, and river runoff, other variables highlighted in modeling studies (e.g., Saucier and Dionne 1998), such as fall mixed layer depth and spring cloud cover, may also influence the severity of the spring ice season. Unfortunately, these are all difficult parameters to measure and monitor, even with remote sensing techniques, and reliable datasets are not yet available for CCA model development. The lead time of sea ice forecasts could potentially be extended by using output from regional or climate models. This technique is called perfect prognostic and is also commonly used in weather forecasting. For example, model forecasts of fall North Atlantic SST could increase the lead time of July ice concentration forecasts beyond six months.

An interesting implication of the link identified in this study between multidecadal variability in ice coverage in Hudson Bay and the Atlantic multidecadal oscillation is that, if the AMO returns to a cool phase as predicted in a recent modeling study (Keenlyside et al. 2008), there could be some recovery in early summer ice concentration in Hudson Bay. Further exploration of the teleconnection pathway identified in this study, particularly in the context of hypothesized large-scale ice–ocean–atmosphere feedbacks, is an avenue for future research.

**Acknowledgments.** The authors are especially grateful to Simon Mason for providing the source code for the climate prediction tool developed at the International Research Institute for Climate and Society. We would also thank John Walsh, Sheldon Drobot, Trudy Wohlleben, and four anonymous reviewers for their very constructive
comments and suggestions that greatly improved this manuscript. Funding for this project was provided by the Canadian Ice Service of Environment Canada through the Canadian Long-Range Ice Forecasting (CLIF) research program. Dr. John Yackel was supported by an NSERC Discovery Grant, and revisions to this manuscript were made by the lead author under a postdoctoral fellowship with the International Arctic Research Center in Fairbanks, Alaska.

REFERENCES

——, 2007a: Regional Charts: Canadian Ice Service Ice Regime Regions (CISIRR) and Sub-Regions with Associated Data Quality Indices. Canadian Ice Service, Environment Canada, 102 pp.


