Representation of Snow in the Canadian Seasonal to Interannual Prediction System. Part II: Potential Predictability and Hindcast Skill

REINEL SOSPEDRA-ALFONSO, WILLIAM J. MERRYFIELD, AND VIATCHESLAV V. KHARIN

Canadian Centre for Climate Modelling and Analysis, University of Victoria, Victoria, British Columbia, Canada

(Manuscript received 14 January 2016, in final form 20 May 2016)

ABSTRACT

This paper examines potential predictability (PP) and actual skill for snow water equivalent (SWE) in the Canadian Seasonal to Interannual Prediction System (CanSIPS). A significant PP is found for SWE, with potentially predictable variance over 50% of the total variance at up to a 5-month lead in mid- to high latitudes in forecasts initialized after snow onset. Much, though not all, of this PP stems from a tendency for SWE anomalies to persist through the snow season. Although the spring melt acts as a PP barrier regardless of initialization date, in some regions significant PP reemerges in the following snow season. This is due primarily to ENSO teleconnections that are modeled realistically by CanSIPS, particularly in northwestern North America. Actual skill of CanSIPS in forecasting SWE is assessed using several verification datasets. Highest skills are obtained using a blend of five such datasets, consistent with the hypothesis that skill scores are degraded by errors in the verification data as well as by forecast errors, and that observational errors can be reduced by blending multiple datasets, much as forecast errors can be reduced by averaging across different models. Actual skill for SWE is comparable to, though generally lower than, that implied by PP. This is due in part to the similar autocorrelation properties of the forecast and observed SWE anomalies, which provide skill through anomaly persistence, combined with reasonably accurate initialization of SWE by CanSIPS. Long-lead skill across snow seasons is found to be linked to ENSO, particularly in western North America, much as for PP.

1. Introduction

This is the second of two papers focusing on the representation of snow on land in the Canadian Seasonal to Interannual Prediction System (CanSIPS), which produces ensemble multiseasonal forecasts using two climate models, the Third and Fourth Generation Canadian Coupled Global Climate Model (CanCM3 and CanCM4, respectively; Merryfield et al. 2013). In the first paper (Sospedra-Alfonso et al. 2016, hereafter

Part I), we examined the ability of CanSIPS to provide realistic forecast initial conditions for snow cover. We showed that snow water equivalent (SWE) in the assimilating model runs that provide initial conditions for CanCM3 and CanCM4 represent reasonably well the observed daily climatology, the interannual variations in maximum SWE, and the timing of snow onset and melt. This paper is concerned with the potential predictability (PP) of SWE derived under the assumption that CanSIPS models are perfect representations of the real world and compares it to the actual skill of CanSIPS forecasts of SWE verified against several observational datasets.

The Northern Hemisphere land surface snow has long been recognized as a major component of the climate system and is known to be a potential source of climate predictability on seasonal and longer time scales. The strong annual cycle and interannual variability of the Northern Hemisphere snow cover, together with its physical properties such as albedo, thermal conductivity, emissivity, and latent heat flux, have a lasting effect on the land surface energy budget as well as atmospheric

Denotes Open Access content.

Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JHM-D-16-0027.s1>.

Corresponding author address: Reinel Sospedra-Alfonso, Canadian Centre for Climate Modelling and Analysis, University of Victoria, 3800 Finnerty Road, Victoria, BC V8P 5C2, Canada.
E-mail: sospedra@uvic.ca

DOI: 10.1175/JHM-D-16-0027.1

circulations (Peings et al. 2011). For example, changes in the surface energy budget during the snowmelt can modulate near-surface temperature variability (Dewey 1977) and extend the effects of snow mass anomalies from late winter to early summer (Yeh et al. 1983).

Snow anomalies also respond to and amplify the impact of climate variability patterns such as El Niño–Southern Oscillation (ENSO; Groisman et al. 1994; Yang 1996; Ferranti and Molteni 1999; Martineu et al. 1999; Corti et al. 2000; Shaman and Tziperman 2005; Wu et al. 2012), the North Atlantic Oscillation (NAO), and the Arctic Oscillation (AO; Clark et al. 1999; Bojariu and Gimeno 2003; Bamzai 2003; Saito and Cohen 2003; Cohen and Fletcher 2007; Riddle et al. 2013). Sobolowski and Frei (2007) showed that ENSO and NAO together explain about 23% of the total winter SWE variance across North America in 1980–97, with centers of action in the principal component loadings over the U.S. northern plains and in western Canada. They also found strong associations between ENSO and NAO variability with SWE over the Great Lakes region and a relationship between ENSO and SWE across the North American midwest toward the west coast associated with warm El Niño events, thus concluding that ENSO is a likely candidate predictor of SWE in these regions. Yang (1996) showed that ENSO indices are precursors for Eurasian snow variability, and Corti et al. (2000) highlighted the significant relationship between the positive and negative Eurasian snow-depth anomalies with warm and cold ENSO events. Schlosser and Dirmeyer (2001), on the other hand, argued that ENSO variability alone cannot account for predictability of the Eurasian snow cover and highlighted the importance of initialization in seasonal climate forecasts. Winter NAO and AO have also been shown to have a consistent relationship with Eurasian snow cover (Serreze et al. 1997) and to be a driver of the Eurasian spring snow cover extent (Bojariu and Gimeno 2003; Bamzai 2003). Conversely, the Eurasian snow cover in the fall has been shown to influence winter NAO and AO (Cohen and Entekhabi 1999; Bojariu and Gimeno 2003; Cohen and Fletcher 2007; Riddle et al. 2013) and to be a skillful predictor of mean climate conditions in the mid- to high latitudes during winter (Cohen and Entekhabi 1999; Cohen et al. 2001; Saito et al. 2001; Gong et al. 2004). Recently, Lin and Wu (2011) showed that, besides sea surface temperature anomalies linked to ENSO and NAO, Tibetan Plateau snow cover in the fall is a good predictor of North American winter temperatures.

These studies have mainly focused on either the effects of climate variability on the snow cover or the influences of snow cover on climate, and to some extent,

the contribution of snow to subseasonal and seasonal predictability. Fewer studies have addressed the impact of snow initialization on the skill of dynamical subseasonal and seasonal forecast. Among these, Shongwe et al. (2007) showed that the presence of large initial snow anomalies improves the forecast of cold spring seasons over most of Europe, except for western regions, where temperatures are influenced by nearby water bodies. Peings et al. (2011) examined the effects of snow initialization in seasonal hindcasts produced with 3-month ARPEGE global atmospheric model simulations driven by observed sea surface temperature. They found that predictability and anomaly correlation skill of near-surface temperature improve when snow was initialized realistically on 1 March compared to when initial snow was specified by the model simulation, although influences on the atmospheric circulation were less evident. Jeong et al. (2013) examined impacts of snow initialization on seasonal predictability of near-surface air temperature in ensembles of simulations using the National Center for Atmospheric Research (NCAR) Community Atmosphere Model, version 3.0 (CAM3.0), and found appreciable gains in potential predictability up to 2 months, particularly in regions having strong snow–albedo feedback. Orsolini et al. (2013) performed a set of forecast experiments using the coupled European Centre for Medium-Range Weather Forecasts (ECMWF) system and contrasted the performance of forecast sets having realistic and unrealistic snow initializations. The former were shown to have superior skill out to 30 days and also exhibited reduced model biases. More recently, Ambadan et al. (2016) highlighted the importance of realistic snow and soil moisture initialization in CanCM3 forecasts of springtime near-surface air temperature.

The influence of snow cover on climate variability as well as the effects of snow initialization on subseasonal to seasonal forecasts are a motivation to examine the impact of snow initialization on forecasts of snow itself. We examine the potential predictability of SWE in the “perfect model” framework by employing analysis of variance (ANOVA; e.g., von Storch and Zwiers 1999) on the 10 CanCM3 and 10 CanCM4 ensemble members to estimate the potentially predictable variance fraction of SWE as simulated in the CanSIPS models in the 1981–2010 hindcast period. The essence of this approach is to partition the total interannual variability of the climate system into two components, unpredictable chaotic fluctuations and potentially predictable variability associated with internal climate variability modes (e.g., ENSO, NAO, and AO) and/or external forcing (e.g., solar variability, explosive volcano eruptions, and anthropogenic radiative forcing). The potential

predictability of the climate system is thus defined as the ratio of the potentially predictable variance to the total variance.

The sources of SWE potential predictability identified by this method are examined by partitioning PP into a component attributable to persistence of initial SWE anomalies, characterized by the autocorrelation of predicted SWE anomalies, and a remaining component that represents PP arising from climate variations that are potentially predictable at longer lead times. We show that persistence of initial SWE anomalies accounts for most of the potentially predictable variance at relatively short lead times, whereas climate variations associated with ENSO variability are found to be a source of SWE potential predictability in North America for longer lead times.

Potential predictability of SWE is compared to actual skill of CanSIPS SWE forecasts verified against observation-based data. To capitalize on PP due to persistence of initial anomalies, the method used to initialize forecast SWE must itself be reasonably skillful; this was established in [Part I](#). Likewise, in order to capitalize on PP due to predictability of climate variations such as ENSO, these variations must be predicted with some skill. We provide an assessment of actual CanSIPS forecast skill of SWE by computing the temporal anomaly correlation coefficient (ACC) and mean square skill score (MSSS) for CanSIPS ensemble-mean hindcasts, using several SWE data products as verifying observations. These skill scores are found to be dependent on the SWE verification datasets, highlighting the uncertainty in the forecast skill estimation. The highest skill scores are obtained using the multidataset blend described by [Mudryk et al. \(2015\)](#). Although there is not an a priori relationship between potential and actual skill in general ([Kumar et al. 2014](#)), we find an approximately linear dependence between $\sqrt{\text{PP}}$ and ACC over North America and Canada for lead times of up to 5 months and forecast months in October–May, with potential skill estimates typically exceeding actual skill estimates.

This paper is organized as follows. In [section 2a](#), we provide a brief overview of CanSIPS and describe the SWE observation-based products used as verifying observations. The ensemble method for calculating potential predictability and measures for actual skill are detailed in [sections 2b](#) and [2c](#), respectively. The behavior and sources of SWE potential predictability in CanSIPS, including the influences of ENSO teleconnections on SWE across North America, are examined in [section 3](#). The actual forecast skill of SWE in CanSIPS is discussed in [section 4](#). A summary and conclusions are presented in [section 5](#).

2. Data and methods

a. Data and overview of CanSIPS

The CanSIPS models and their initialization are described in detail by [Merryfield et al. \(2013\)](#) and also discussed in [Part I](#); thus, we provide only a brief summary here. CanSIPS is based on two global climate models developed at the Canadian Centre for Climate Modelling and Analysis (CCCma), CanCM3 and CanCM4, and provides Environment Canada's (now Environment and Climate Change Canada) operational seasonal forecasts. Both CanCM3 and CanCM4 share common ocean, land surface, and sea ice components but differ in their atmospheric component. Each forecast ensemble member (10 each for CanCM3 and CanCM4) is initialized from a separate assimilating run in which atmospheric winds, temperature, and humidity as well as sea surface temperature and sea ice concentration are constrained near observed values. Forecast initial conditions for the land component including snow cover are determined by the response of the Canadian Land Surface Scheme (CLASS), version 2.7, to forcing from model atmospheric fields constrained by 6-hourly reanalysis data. Because the atmospheric variables are constrained only on horizontal scales larger than about 1000 km with a relaxational time scale of 24 h, differences are present on smaller scales that represent observational uncertainties, as shown in Fig. 23 of [Merryfield et al. \(2013\)](#). SWE initial conditions for each ensemble member thus differ from each other, as illustrated in Fig. 2 of [Part I](#).

CanSIPS retrospective forecasts (hindcasts) are initialized at the beginning of each month during a multi-decadal hindcast period and have a 12-month range. We examine daily mean SWE on the $\approx 2.8^\circ$ atmospheric–land surface model grid for CanCM3 and CanCM4 and consider hindcasts initialized at the beginning of October, December, February, and April in 1981–2010. October and April are representative of the snow onset and melt periods in the mid- to high latitudes, respectively, whereas December and February provide two different stages in the snow accumulation period. Computations are performed on land grid cells only, with lakes and sea ice excluded. When evaluating the influence of ENSO teleconnections, we consider monthly mean values of SWE, surface temperature, and precipitation, as well as the Niño-3.4 index (defined as the averaged sea surface temperature anomaly over the Pacific Ocean region 5°S – 5°N , 120° – 170°W).

We also employ monthly averages of ensemble-mean SWE in CanCM3 and CanCM4 hindcasts to compute the temporal ACC and MSSS based on several SWE observation-based products for every calendar month

and lead time. Here, lead time is defined as the number of months between initialization and the beginning of the verification period, with lead 0 corresponding to the first forecast month, lead 1 to the second forecast month, etc. Spatially averaged ACC and MSSS across the Northern Hemisphere, North America, and Canada are also computed. The ACC and MSSS are computed for the two-model ensemble-mean values on which CanSIPS forecasts are based. The PP of SWE is computed separately for CanCM3 and CanCM4 for reasons explained in the next subsection, and the mean of the two values is reported for the spatial averages in North America and Canada examined in [section 4](#).

The SWE datasets employed as verifying observations are 1) the National Aeronautics and Space Administration (NASA) Modern-Era Retrospective Analysis for Research and Applications (MERRA; [Rienecker et al. 2011](#)); 2) the ECMWF interim land reanalysis (ERA-Interim/Land; [Balsamo et al. 2013](#)); 3) the GlobSnow analysis, version 2, developed through the European Space Agency GlobSnow project and produced by the Finnish Meteorological Institute ([Takala et al. 2011](#)); 4) the Global Land Data Assimilation System, version 2 (GLDAS-2; [Rodell et al. 2004](#)); 5) the Crocus snow scheme driven by ERA-Interim ([Brun et al. 2013](#)); 6) MERRA-Land ([Reichle et al. 2011](#)); and 7) ERA-Interim ([Dee et al. 2011](#)). Among these, MERRA and GlobSnow were previously employed in [Part I](#) to assess the initialization of snow cover in CanSIPS. Data products 1–5 were analyzed in detail by [Mudryk et al. \(2015\)](#) and combined into a blended multidataset SWE product, denoted here as Blended-5. [Mudryk et al. \(2015\)](#) found the constituent SWE products of Blended-5 to be the most mutually consistent. Therefore, we employ Blended-5 regridded to CanSIPS resolution as our primary verifying dataset and find that CanSIPS predictive skill for SWE is higher when verified against Blended-5 than any individual SWE product listed above.

b. Ensemble method for potential predictability

Previous studies of potential predictability based on observations and/or climate models include [Madden \(1976\)](#), [Zwiers \(1996\)](#), [Rowell \(1998\)](#), [Boer \(2004\)](#), and [DelSole et al. \(2013\)](#). Here, we estimate the potential predictability in CanSIPS SWE hindcasts by means of one-way ANOVA via the ensemble method ([DelSole et al. 2013](#)). In ANOVA, it is assumed that the state $\eta_y(t)$ to be predicted in year y at lead time t can be modeled in the form $\eta_y(t) = \pi_y(t) + \nu_y(t)$, where $\pi_y(t)$ is the potentially predictable “signal” and $\nu_y(t)$ is an unpredictable chaotic fluctuation superimposed on $\pi_y(t)$. The goal is to estimate the fraction of the total variance of $\eta_y(t)$ at lead time t that is explained by the variance of $\pi_y(t)$.

For CanCM3 and CanCM4 individually, we consider the ensemble of $E = 10$ SWE forecasts $X_{ey}(t)$, $e = 1, \dots, E$, where the subindex $y = 1, \dots, Y$ counts the forecast years (here, $Y = 30$), and $t = 0, \dots, 364$ is the lead time in days or $t = 0, \dots, 11$ is the lead time in months. An unbiased estimator of the internally generated variability due to unpredictable chaotic fluctuations or “noise” is given by ([von Storch and Zwiers 1999](#); [DelSole et al. 2013](#))

$$\sigma_N^2(t) = \frac{1}{Y(E-1)} \sum_{y=1}^Y \sum_{e=1}^E [X_{ey}(t) - \bar{X}_y(t)]^2, \quad (1)$$

where $\bar{X}_y(t) = \sum_{e=1}^E X_{ey}(t)/E$ is the ensemble mean in year y at lead time t . An unbiased estimator of the variance of the ensemble means is given by

$$\sigma_S^2(t) = \frac{1}{Y-1} \sum_{y=1}^Y [\bar{X}_y(t) - X(t)]^2, \quad (2)$$

where $X(t) = \sum_{e=1}^E \sum_{y=1}^Y X_{ey}(t)/EY$ is the climatological grand ensemble mean at lead time t .

Because of the finite number of ensemble members, σ_S^2 is a biased estimator of the potentially predictable “signal” variance as it contains the residual “noise” variance remaining in the ensemble means after averaging over a finite number of ensemble members ([von Storch and Zwiers 1999](#)). An unbiased estimator of the potentially predictable “signal” is obtained by subtracting this residual “noise” variance from $\sigma_S^2(t)$ and is given by

$$\sigma_P^2(t) = \sigma_S^2(t) - \sigma_N^2(t)/E. \quad (3)$$

Note that the sampling variability may result in negative estimates of $\sigma_P^2(t)$. An unbiased estimator of the total interannual variance is then obtained as the sum of the two unbiased estimates of the “signal” and “noise” variability, that is,

$$\sigma_T^2(t) = \sigma_P^2(t) + \sigma_N^2(t).$$

Potential predictability is defined here in terms of the fraction of the potentially predictable variance to the total variance, also referred to as the potential predictability variance fraction (e.g., [Boer et al. 2013](#))

$$\text{PP}(t) = \sigma_P^2(t)/\sigma_T^2(t) = \frac{\sigma_S^2(t) - \sigma_N^2(t)/E}{\sigma_S^2(t) + \sigma_N^2(t)/A}, \quad (4)$$

where $A = E/(E-1)$, and in the last equality PP has been written in terms of Eqs. (1) and (2). The statistical significance of the potential predictability estimate PP(t) can be assessed by computing the F -ratio statistic

$$F(t) = \frac{E\sigma_s^2(t)}{\sigma_N^2(t)} = \frac{1 + (E-1)PP(t)}{1 - PP(t)} \quad (5)$$

and comparing it against critical values of the F distribution $F[Y-1, (E-1)Y]$ with $Y-1$ and $(E-1)Y$ degrees of freedom (von Storch and Zwiers 1999). For a 10-member ensemble and 30 forecast years, the null hypothesis that the potential predictability is zero is rejected at the 5% significance level if $F(t) > 1.51$, or equivalently, if $PP(t) > 0.048$.

To assess persistence of initial SWE anomalies throughout the forecast, we compute the temporal autocorrelation (AC) of SWE forecast anomalies $X'_{ey}(t) = X_{ey}(t) - X(t)$ against the initial ensemble-mean anomaly $\bar{X}'_y(0) = \bar{X}_y(0) - X(0)$ averaged across the ensemble as

$$AC(t) = \frac{1}{E} \sum_{e=1}^E \frac{\sum_{y=1}^Y \bar{X}'_y(0) X'_{ey}(t)}{\sqrt{\sum_{y=1}^Y \bar{X}'_y(0)^2} \sqrt{\sum_{y=1}^Y X'^2_{ey}(t)}}. \quad (6)$$

We use the initial ensemble-mean anomaly instead of initial anomalies of individual ensemble members to account for uncertainty in the initial conditions of SWE anomalies, which results in a degradation of autocorrelation at the zero lead time. The variable $AC^2(t)$ is a measure of the SWE variability fraction at lead time t that can be linearly attributed to the initial SWE anomalies and can be compared with PP of SWE to assess the contribution of the memory of snowpack initial conditions to the potential predictability. The squared autocorrelation estimated as above is slightly positively biased for small to moderate autocorrelation values. This bias is largest in the absence of autocorrelation and is about 0.035 for a sample size of $Y = 30$. The leading-order bias is estimated and subtracted from the $AC^2(t)$ estimates by employing the jackknife resampling procedure (e.g., von Storch and Zwiers 1999).

Because PP is a model-specific property that can differ markedly between models, calculations of PP as detailed above are carried out separately for CanCM3 and CanCM4. For the sake of brevity, in sections 3a–c we focus mainly on the results obtained for CanCM4, whereas parallel results obtained for CanCM3 are discussed in the supplemental material.

c. Measures of actual forecast skill

As detailed above, CanSIPS is a two-model forecasting system that combines ensemble forecasts from global climate models CanCM3 and CanCM4. Forecast anomalies are constructed as simple averages of

ensemble and monthly mean anomalies from the two models, and we follow this methodology when evaluating CanSIPS skill. As measures of actual skill of the CanSIPS forecasts of SWE anomalies, we consider ACC and MSSS computed using the verifying observations described in section 2a. These two skill measures are somewhat complementary in that ACC is sensitive to the sign and relative magnitudes of the predicted anomalies, whereas MSSS penalizes errors in forecasting anomaly magnitudes even when the sign is correct. The ACC between the forecast ensemble means $\bar{X}_y(t)$ and observations $O_y(t)$ at lead time t is computed as

$$ACC(t) = \frac{\sum_{y=1}^Y \bar{X}'_y(t) O'_y(t)}{\sqrt{\sum_{y=1}^Y \bar{X}'_y(t)^2} \sqrt{\sum_{y=1}^Y O'^2_y(t)}}, \quad (7)$$

where $O'_y(t) = O_y(t) - O(t)$ denotes observational anomalies from climatology $O(t)$. MSSS is computed using the observed climatology as a reference forecast,

$$MSSS(t) = 1 - \frac{\sum_{y=1}^Y [\bar{X}'_y(t) - O'_y(t)]^2}{\sum_{y=1}^Y O'^2_y(t)}. \quad (8)$$

The MSSS values for the raw ensemble-mean forecast anomalies $\bar{X}'_y(t)$ are often negative as the error variance of the raw forecast ensemble-mean anomalies exceeds that of the climatology forecast. In the following, we therefore report MSSS for the linearly rescaled forecast anomalies $\hat{X}'_y(t) = s(t)\bar{X}'_y(t)$ where the rescaling coefficient $s(t)$ is determined by minimizing the mean-square error using a leave-one-out cross validation procedure. Spatial mean ACC and MSSS values are obtained through area-weighted averaging in regions where the model and observed interannual standard deviation of SWE is greater than 0.1 mm.

3. Behavior and sources of SWE potential predictability

In section 3a, we first compute PP of CanCM4 SWE forecasts at two of the Canadian weather stations considered in Part I, Goose Bay in Newfoundland and Labrador (GBY-NL), and Mission Ridge in British Columbia (MRD-BC), which are representative of different snow-climate zones. GBY-NL is in the taiga region of eastern Canada whereas MRD-BC is in the

FIG. 1. The SWE daily mean forecasts for CanCM4 at the grid locations nearest to (left) Goose Bay and (right) Mission Ridge in eastern and western Canada, respectively, initialized at the start of (a),(b) October; (c),(d) December; (e),(f) February; and (g),(h) April 1981–2010 showing SWE predictions from each ensemble member and forecast year (light gray), ensemble mean in each forecast year (dark gray), climatological ensemble mean (black), noise variance (red), total variance (blue), and potentially predictable variance (green). Variances are scaled by a factor of 0.02 to fit the panels.

cordilleras of western Canada. Details about these two sites and the performance of CanCM3 and CanCM4 in initializing SWE at these locations are presented in Part I. We then extend the analysis to the entire Northern Hemisphere. The role of SWE anomaly persistence on PP is discussed in section 3b. The influence of ENSO teleconnections on PP of SWE in western North America is discussed in section 3c. The behavior of CanCM3 SWE forecasts is summarized in the supplemental material.

a. Potential predictability of SWE

Figure 1 shows the daily time evolution of SWE forecasts for each of the 10 CanCM4 ensemble members initialized at the beginning of October, December, February, and April in 1981–2010 for the model grid cells containing GBY-NL and MRD-BC. Also shown are the ensemble means for each forecast year and their

climatological mean as well as the noise and the potentially predictable and total variances. Figure 1a shows the forecasts for GBY-NL initialized in October, which is approximately the time of snow onset at this site. The interannual SWE variability is low early in the snow season because of the low average snow accumulation. As the snow accumulation period progresses, the spread of the ensemble means increases mostly because of the increasing spread of the individual ensemble members (i.e., noise). The total variance $\sigma_T^2(t)$ is approximately equal to the noise variance $\sigma_N^2(t)$, implying low PP for forecasts initialized in October. These PP values computed from Eq. (4) are indicated by the black curve in Fig. 2a. The PP drops quickly at the start of the forecast and becomes statistically indistinguishable from zero within 2 weeks.

Daily evolution of SWE forecasts for GBY-NL initialized later in the snow season is depicted in Figs. 1c,

FIG. 2. Black curves show SWE potential predictability [Eq. (4)] for CanCM4 daily mean forecasts at the grid points nearest to (left) Goose Bay and (right) Mission Ridge in eastern and western Canada, respectively, for (a),(b) October; (c),(d) December; (e),(f) February; and (g),(h) April 1981–2010. Gray curves show temporal squared autocorrelation of forecast to initial SWE anomalies. Dashed line indicates the 95% confidence level.

1e, and 1g. In these cases, the potentially predictable variance explains a much larger fraction of the total interannual variability during the snow accumulation period (until approximately 2–4 weeks after the snow peak) compared to when the forecast is initialized in October (Fig. 1a). This implies higher PP when initialization is done at this stage of the snow season, as shown in Figs. 2c, 2e, and 2g. At this location, $\sigma_p^2(t)$ tends to remain relatively unchanged throughout the forecast, approximately until both σ_T^2 and σ_N^2 reach a maximum (typically during the springtime), to then drop quickly to zero. The timing of maximum σ_T^2 thus corresponds to a temporal barrier for potential predictability at this location regardless of the forecast initialization date. After this time, and into the following snow season, SWE at this station in northeastern Canada is no longer potentially predictable.

Corresponding results for the grid cell containing MRD-BC are shown in Fig. 1 (right) and Fig. 2 (right). By contrast with GBY-NL, SWE forecasts for MRD-BC initialized in October show significant PP during most of the snow season (Fig. 2b). After a relatively rapid decrease during the first month, PP remains low, although statistically significant, throughout the next 3 months, and then increases again leading up to the climatological SWE peak, reaching PP values >0.2 for about 2 months until the end of the snow season. This reemergence of PP several months into the forecast suggests that predictable large-scale climate variations and/or external forcing may be contributing appreciably to interannual SWE variations at this location.

Appreciable PP is found in SWE forecasts at MRD-BC beyond about 2 months for initializations in December and February, and beyond about 1 month after

initialization in April (Figs. 1d,f,h and 2d,f,h). In each case PP decreases rapidly initially, then reaches a plateau at the peak of the snow season and decreases rapidly again toward the end of the snow season. This behavior is particularly evident for the forecast initialized in December, when the plateau extends over 2 months with $PP > 0.4$ (Fig. 2d). This is an indication of a relatively large $\sigma_p^2(t)$ that does not necessarily originate from the SWE initial conditions, consistent with the increasing values of $\sigma_p^2(t)$ about the time of PP reemergence, as shown in Fig. 1 (right). Also noteworthy is the significant PP in the subsequent snow season for the forecasts initialized in April (Fig. 2h). In this case, all memory of the initial SWE anomalies has been erased by the spring melt, but there is nonetheless a predictable component of interannual SWE variation in the following snow season leading to $PP \approx 0.3$ at the longest lead time in late March. Causes for the re-emergent PP at MRD-BC are investigated in section 3c, whereas the remainder of this section and section 3b consider the contribution to PP originating from SWE initial conditions.

Figure 3 shows PP of CanCM4 SWE forecasts at lead times $t = 0, 15, 30$, and 60 days for the four initial months considered. Overall, PP tends to decrease with lead time at a rate that depends on location and initialization date. Initially, PP is generally highest in northern Siberia, the Pacific Northwest, the Tibetan Plateau and the Karakoram, and to some extent eastern Canada and Scandinavia, with values depending on initialization month. The PP is typically very high (>0.9) at the start of the forecast in December, February, and April, and remains high (>0.8) in the first 2 weeks. For October, PP remains relatively low throughout the forecast, with the highest values found north of 60°N (~ 0.5 after 15 days). For April, there is appreciable PP (>0.4) in the regions mentioned above 30 days after initialization, but much lower values after 60 days (<0.2), except in the Tibetan Plateau, eastern Siberia, and the Pacific Northwest (~ 0.3). Low values are found in the midlatitudes where there is a relatively large noise variance characteristic of the snowmelt. The PP of SWE initialized in February is typically highest at every lead time among all initial months considered.

Figure 4 shows the number of consecutive days following the start of CanCM4 forecasts that PP exceeds 0.5 and 0.9. Overall, the longest such intervals (with actual values depending on the time of initialization and location) are found in the high latitudes ($>60^\circ\text{N}$), Pacific Northwest, Tibetan Plateau, and the Karakoram. There are typically more days with potential predictability when the forecast is initialized well into the snow accumulation period (e.g., December–February in the

mid- to high latitudes). The duration of potential predictability is reduced significantly for forecasts initialized earlier during the snow onset. For example, $PP > 0.5$ for up to 3–4 weeks for the forecasts initialized in October and up to 4–5 months for the forecasts initialized in December. Because of the snowmelt, forecasts initialized in April have $PP > 0.5$ for up to 1–2 months in the high latitudes and for slightly more time in the Tibetan Plateau. The longest time period for which PP remains greater than 0.9 corresponds to the forecasts initialized in February, ranging up to approximately 6 weeks.

b. Potential predictability due to anomaly persistence

We showed in section 3a that SWE PP is high at the start of the forecast and can remain relatively high for some time depending on the initialization date. Because SWE is a cumulative quantity, it contains information about previous snowfalls that can last for relatively long times, particularly at higher latitudes and/or elevations where temperature remains sufficiently low for longer periods. This “memory” of previous snow accumulation should therefore contribute to PP.

Evidence of this is depicted in Fig. 2, which, in addition to PP, shows the squared temporal autocorrelation of forecast SWE anomalies defined in Eq. (6). For GBY-NL (Fig. 2, left), AC^2 remains close to PP for most of the forecasts, with the exception, to some extent, of the forecasts initialized in April. This suggests that PP of SWE at GBY-NL can largely be attributed to the persistence of initial SWE anomalies, particularly for the forecasts initialized in December and February (Figs. 2c,e).

By contrast, AC^2 for MRD-BC (Figs. 2b,d,f,h) initially decays at approximately the same rate as PP, but then continues to decay while PP stabilizes or even increases as described in section 3a. This indicates that there is a component of PP that cannot be explained in terms of persistence of initial anomalies. Such behavior is particularly evident for the forecasts initialized in October and December (Figs. 2b,d), which retain significant PP about 120 and 60 days, respectively, after initialization (in about February), despite much lower values of AC^2 .

Figure 5 shows the geographical distribution of AC^2 for CanCM4 SWE forecasts at $t = 15, 30$, and 60 days initialized in October, December, February, and April. The geographical patterns are generally similar to those for PP (Fig. 3), although AC^2 tends to be nonsignificant at some peripheral snow regions where PP remains significant. The AC^2 is particularly low relative to PP for the forecast initialized in October, with little persistence of initial SWE anomalies after 15 days. Geographical patterns of AC^2 for the forecasts initialized in December, February, and April are similar to those for PP,

FIG. 3. Potential predictability of daily SWE from CanCM4 forecast runs at lead times [from (top) to (bottom)] of 0, 15, 30, and 60 days. Rows are for equal lead time and columns for initial conditions at the start of the indicated month.

although AC^2 values again are somewhat lower than for PP.

The autocorrelation of SWE in CanCM4 forecasts initialized in December, February, and April essentially vanishes in springtime, typically 1–3 weeks after the total variance σ_T^2 reaches its maximum during the spring melt. This is shown in Fig. 6 (top). The PP of SWE, on

the other hand, tends to remain statistically significant for a slightly longer period. This eradication of AC and PP occurs well before climatological SWE itself vanishes, as evidenced by Fig. 6 (bottom), which shows the lag following peak σ_T^2 at which climatological SWE decreases below 1 mm. Thus, the date of maximum σ_T^2 largely coincides with a temporal barrier for SWE

FIG. 4. Consecutive days with potential predictability of daily mean SWE larger than PP_c equal to (top) 0.5 and (bottom) 0.9 from CanCM4 forecast runs. Rows are for equal PP_c and columns for initial conditions at the start of the indicated month.

potential predictability. There are, however, regions where SWE is potentially predictable even at the subsequent snow season, as we show next.

c. Potential predictability due to ENSO teleconnections in North America

While much of PP at shorter lead times is attributable to the persistence of initial SWE anomalies as shown above, there is an additional long-range component of PP that is associated with predictability of future snow accumulation and is therefore not directly associated with the initial SWE anomalies. This PP component is more prominent at certain locations such as the Mission Ridge observation station, as illustrated in Fig. 2, and can remain appreciable at very long lead times, extending even into the next snow season, as indicated in Fig. 2h. This long-range component of PP is particularly pronounced over western North America. In this subsection, we present evidence that this long-range PP is mainly a result of potentially predictable temperature and precipitation variations connected with ENSO.

The long-range behavior of PP is illustrated in Fig. 7, which shows PP of March-averaged SWE in North America for CanCM4 forecasts initialized at the beginning of February (lead time $t = 1$ month), December ($t = 3$ months), October ($t = 5$ months), and April ($t = 11$ months). For $t = 1$ month, PP is greater than 0.8 in

northwestern Canada, the Canadian Arctic Archipelago, and a portion of northeastern Canada, whereas PP is greater than 0.5 in the rest of Canada and in the U.S. Rocky Mountains. These relatively high PP values are at least partly attributable to the persistence of initial SWE anomalies, as discussed in section 3a above. For $t = 3$ months, the pattern of SWE PP is similar to that for $t = 1$ month, although PP values are somewhat reduced. By $t = 5$ months, which corresponds to the forecast initialized at the start of October, prior to snow onset in the midlatitudes, PP of SWE is mostly lost, except in the Pacific Northwest and the southern Rocky Mountains as well as the far northern latitudes. For $t = 11$ months, which corresponds to the forecasts initialized in April of the previous snow season, any influence of SWE initialization has been erased by the spring melt and summer season. Nonetheless, PP in March is significant in the Pacific Northwest and much of the Rocky Mountains. Since the linear trend over the 30-yr hindcast period in CanCM4 March-averaged SWE hindcasts is typically $<5\%$ of the signal standard deviation in North America regardless of lead time (not shown), and since PP values when linear trends are removed are almost indistinguishable from those in Fig. 7, we argue that such significant long-lead PP is the result of SWE response to other climate variables most likely linked to ENSO variability, combined with the ability of the forecasts to predict ENSO, as we show next.

FIG. 5. Temporal squared autocorrelation for CanCM4 forecast daily mean SWE anomalies relative to initial SWE anomalies at lead times of (top) 15, (middle) 30, and (bottom) 60 days. Rows are for equal lead time and columns for initial conditions at the start of the indicated month.

Historical simulations of CanCM4 represent the amplitude, frequency content, and seasonality of observed ENSO variability reasonably well (Merryfield et al. 2013). This model fidelity likely is one reason why CanCM4's ENSO prediction skill is competitive with that of other current-generation climate prediction models (Merryfield et al. 2011; DelSole et al. 2014). The strongest ENSO influences on North America occur in winter through atmospheric teleconnections, and these influences on near-surface temperature and precipitation in December–February (DJF) are also represented reasonably well by CanCM4, although temperature influences of opposite polarity tend to be somewhat more concentrated in northwestern and southeastern North America, respectively, than is observed (Merryfield et al. 2013). We characterize ENSO influences on North America in CanCM4 forecasts by regressing forecast

fields at a given lead time on forecast values of the Niño-3.4 index, frequently used to characterize ENSO, at the same lead time. Despite the ENSO spring predictability barrier (e.g., Webster 1995), CanCM4 has considerable skill at forecasting the Niño-3.4 index up to 1-yr lead, including for predictions initialized in April, for which the anomaly correlation skill score verified against the Optimum Interpolation Sea Surface Temperature (OISST) observational dataset remains over 0.70 throughout the forecast (not shown).

Figures 8a and 8b show the regression patterns for December- and March-averaged SWE forecasts initialized at the beginning of April, corresponding to 8- and 11-month lead times, respectively. Statistical significance at $p = 0.10$ according to a two-tailed Student's t test is indicated by cross hatching where correlation values exceed 0.3. At both lead times, the SWE response

FIG. 6. Lag following the date of peak SWE total variance σ_7^2 at which SWE anomaly autocorrelation, potential predictability, and snowmelt in CanCM4 forecasts cease to be significant (i.e., $AC < 0.3$, $PP < 0.1$, and $SWE < 1$ mm, respectively). Columns are for initial conditions at the start of the indicated month. This shows that the date of maximum total variance largely coincides with a temporal barrier of SWE potential predictability.

to positive Niño-3.4 anomalies is negative for most of the Pacific Northwest, the northern Rocky Mountains, and a large portion of central-to-eastern Canada, particularly near the Great Lakes and Hudson Bay regions. A positive SWE response occurs in parts of northwestern Canada and Alaska, the southern Rocky Mountains, and northwestern portions of the U.S. Great Plains. These features are largely consistent with the SWE response to ENSO in observation-based studies (Sobolowski and Frei 2007; Seager et al. 2010). The negative SWE anomalies in Canada have larger absolute values in March relative to December, whereas the positive anomaly pattern in the U.S. Great Plains has decreased in extent. For both target months, the negative SWE response in the Pacific Northwest is significant, and the positive response in the southern Rocky

Mountains is significant in December and still present but somewhat reduced in March. These regions correspond to where March-averaged SWE is found to be potentially predictable even across snow seasons (Fig. 7d).

The SWE response to ENSO is likely driven by ENSO influences on temperature T and precipitation P . Figures 8c–f show regression patterns for December- and March-averaged T and P forecasts initialized in April as for SWE (i.e., 8- and 11-month lead times). Forecast T anomalies are positive and P anomalies are negative in the Pacific Northwest when El Niño conditions are forecast, with both influences contributing to the observed reduction in SWE. The opposite behavior is found in the southern Rocky Mountains and U.S. Great Plains, with lower T and higher P under forecast El Niño conditions both contributing to the

FIG. 7. The PP of March-averaged SWE for CanCM4 forecasts in North America at lead times of (a) 1, (b) 3, (c) 5, and (d) 11 months indicating the time from forecast initial values.

forecast increase in SWE. The T and P regression patterns are similar for both target months and largely independent of forecast lead time (not shown), but they tend to be weaker in March than in December, except for the increase of negative P anomalies in eastern Canada. These regression patterns are similar to those for DJF in CanCM4 shown in [Merryfield et al. \(2013\)](#) and include the biases in these patterns described above. Regression patterns computed with linear trends removed are very similar to those in [Fig. 8](#) (not shown).

Because SWE is a cumulative product of previous snowfall and snowmelt events, the influence of T and P

on SWE variability is not limited to contemporary months, but also has contributions from previous months in the snow season (e.g., [Sospedra-Alfonso et al. 2015](#)). This may explain, at least in part, the intensification between December and March of the negative Niño-3.4-regressed SWE anomalies in western and east-central Canada ([Figs. 8a,b](#)), where peak climatological SWE tends to occur in March or later ([Part I](#)). Also, this may explain why PP in the U.S. southern Rocky Mountains is relatively high in March at 11-month lead ([Fig. 7d](#)), even though the ENSO-regressed SWE anomalies are not statistically significant at this month ([Fig. 8b](#)). Conversely,

FIG. 8. Regressions of CanCM4-predicted (left) December- and (right) March-averaged (a),(b) SWE; (c),(d) surface temperature; and (e),(f) precipitation against CanCM4-predicted Niño-3.4 index for the forecasts initialized in April. Lead times are 8 (left) and 11 months (right). Cross-hatched regions correspond to correlations >0.3 .

regions with significant SWE forecast anomalies associated with ENSO but low PP may be the result of relatively large noise variances that overcome signal variability.

4. CanSIPS SWE hindcast skill

The previous section quantified the potential predictability of SWE for each of the two CanSIPS models separately, referring to daily values in order to clearly resolve the strong seasonality of SWE, whereas in this section we assess the actual skill of CanSIPS hindcasts of SWE. In doing so, we apply the methodology used in constructing deterministic CanSIPS forecasts; namely, monthly and seasonal mean forecast anomalies are described by unweighted averages of the ensemble-mean anomalies from the two models. Here we consider monthly mean forecasts, although comparable results are obtained when seasonal means are considered instead.

We first examine data products for SWE forecast verification. The seasonal and lead time dependence of the ACC and MSSS and their geographical distribution are then examined, after which these actual skills are compared to the PP-based “potential skills” evaluated in [section 3](#).

a. Choice of verification product

Assessment of climate forecast skill over large regions or globally requires a gridded verification data product for the variable of interest whose spatiotemporal coverage is sufficient for the desired analysis to be performed. Station-based and remotely sensed measurements are inevitably subject to spatial and temporal gaps as well as discontinuities because of observing system changes, and so generally do not suit this purpose. Climate forecast verification datasets therefore most often consist either of reanalyses obtained from data-constrained models or of syntheses of observations in which various procedures are applied to fill gaps and ensure that physical constraints are obeyed.

Although several gridded SWE products are available that potentially could be applied to SWE forecast verification, these datasets have not previously been considered in that context. Therefore, we evaluate skills for these multiple SWE data products covering the 1981–2010 verification period that were described in [section 2a](#), including the blend of five such products (Blended-5) described in [Mudryk et al. \(2015\)](#).

Forecast skills obtained using these various products for verification are summarized in [Fig. 9](#), which shows Northern Hemisphere averages of ACC and MSSS computed as described in [section 2c](#) for every target month at 0- and 1-month lead times for five different

verification products. These skills are consistently highest in winter and early spring, after the snowpack has accumulated and before the spring melt, although there are substantial differences in skill scores obtained using the different verification products. Highest scores are consistently obtained using Blended-5, followed closely by MERRA and ERA-Interim/Land, whereas scores obtained using ERA-Interim and MERRA-Land are much lower, reflecting known issues affecting SWE values in those products ([Reichle 2012](#)).

Although it has not been explicitly demonstrated, an attractive hypothesis is that the higher skills obtained using the Blended-5 product result from a partial cancellation of observational errors in different data products, in much the same way that combining multiple forecast models tends to reduce errors in the aggregated forecast (e.g., [Kirtman et al. 2014](#)). These results also serve as a reminder that skill can reflect errors in the verification product, in addition to the accuracy of the forecast itself. Because CanSIPS SWE initial conditions are set indirectly through the response of the CLASS land surface model to ERA-Interim-based forcing of the model atmosphere, the initialization procedure does not explicitly favor Blended-5 over the other products, indicating that this is not the reason for the higher scores for Blended-5. Based on these considerations and the results in [Fig. 9](#), the Blended-5 product is selected for verification in the remainder of this section.

b. Regional, seasonal, and lead time dependence

The geographic distribution of ACC for CanSIPS verified against Blended-5 at lead times $t = 0, 2$, and 4 months is shown in [Fig. 10](#). The highest ACC values are generally found in the mid- to high latitudes, being typically >0.6 at lead 0 for December, February, and to some extent April. In October, CanSIPS has lower but still significant ACC with values >0.6 only in some high-latitude regions. At 2-month lead time, CanSIPS has significant skill (>0.3) northward of 45°N for the forecasts initialized in December and February, and northward of 60°N for the forecast initialized in October. There is little skill at 2-month lead for the forecasts initialized in April as this corresponds to June when much of the snow cover has receded to high latitudes. At 4-month lead, CanSIPS has significant skill for forecasts initialized in December, and to a lesser degree, for forecasts initialized in October. Little skill is found for the forecasts initialized in February and April since 4-month lead time corresponds to June and August, respectively, when the snow cover has receded considerably.

The geographic distribution of MSSS for CanSIPS verified against Blended-5 at lead times $t = 0, 2$, and

FIG. 9. Spatial averages of (top) ACC and (bottom) MSSS in the Northern Hemisphere for CanSIPS vs verifying observations at lead times (a),(c) 0 and (b),(d) 1 month. Verifying observations are Blended-5 (B5), ERA-Interim/Land (EI-L), ERA-Interim (EI), MERRA-Land (M-L), and MERRA (M). See text for details.

4 months is shown in Fig. 11. As expected, the MSSS values are smaller than the corresponding ACC values as MSSS is bounded from above by ACC^2 [e.g., Murphy 1988, his Eq. (12)]. The spatial pattern of MSSS for 0- and 2-month lead generally follows that of ACC, with elevated skill values in most of eastern Europe, western Russia, and large portions of the 45° – 60° N latitude band in North America for forecasts initialized in December and February. The MSSS values at 4-month lead are generally very low, with the exception, to some extent, of the forecasts initialized in December, at which time the initial anomalies tend to persist through April and seasonal climate variability is relatively predictable.

c. Comparison of potential and actual skill

Because PP represents the fraction of total variance that is potentially predictable, $\sqrt{PP}$ corresponds to the correlation between the signal component of the forecasts and the total forecast values (signal plus noise).

Therefore, $\sqrt{PP}$ is an appropriate measure to compare with the actual ACC skill obtained by verifying forecasts against observations. To make these comparisons somewhat more region-specific than the skill scores for the entire Northern Hemisphere shown in Fig. 9, we consider averages over North America representing a wide range of latitudes, including regions characterized by ephemeral winter snow, and also averages for Canada representing primarily northern latitudes having a permanent winter snowpack.

Figure 12 compares spatial means of $\sqrt{PP}$ and ACC for CanSIPS SWE verified against Blended-5 in North America and Canada as a function of month and lead time. The $\sqrt{PP}$ values are generally higher than ACC for both regions, and both measures are typically higher for Canada than for the whole of North America. Higher forecast skill in Canada relative to North America is likely, at least in part, the result of excluding the southern regions with shallow or ephemeral snow cover that are characterized by a relatively large noise variance and

FIG. 10. The ACC for CanSIPS vs Blended-5 at lead times of (top) 0, (middle) 2, and (bottom) 4 months. Rows are for equal lead times and columns for initial conditions at the start of the indicated month.

hence are less predictable. Higher values occur during the snow accumulation period, between snow onset and the start of the snowmelt (e.g., November–April for 0-month lead). Area-averaged ACC exceed 0.3 for up to 3 months lead time for North America and up to 4–5 months for Canada (Figs. 12c,d), whereas area-averaged $\sqrt{\text{PP}}$ exceed 0.3 for up to 6 months for North America and 7 months for Canada (Figs. 12a,b).

Figures 12a–d also indicate that there is a temporal barrier for SWE predictability and skill in late spring. For example, for leads 1 month or greater, area-averaged $\sqrt{\text{PP}}$ and ACC in North America drop to much lower values in May and the following months. This degradation implies that forecasts initialized in April or earlier have relatively low potential predictability and actual skill by May. A corollary is that forecasts initialized earlier in the snow season tend to remain skillful for longer times. For example, North

America-averaged $\sqrt{\text{PP}}$ remains >0.3 for 7 months for the forecasts initialized in November, but only 3 months for forecasts initialized in March (Fig. 12a). A likely cause for the rapid decline in PP of SWE in springtime is the relatively large noise variance during snowmelt (Fig. 1), when relatively small differences in the timing of the melt can give rise to large differences in SWE across the ensemble.

Scatterplots of these area-averaged ACC and $\sqrt{\text{PP}}$ over North America and Canada are shown in Figs. 12e and 12f. Potential skill given by $\sqrt{\text{PP}}$ is generally higher than the actual forecast ACC. The ACC is found to increase approximately linearly with $\sqrt{\text{PP}}$ with a coefficient of determination R^2 of 0.94 for both North America and Canada. The deficit in ACC relative to $\sqrt{\text{PP}}$ (with the best linear fit lines having slopes and intercepts of 0.73 and -0.10 for North America and 0.82 and -0.13 for Canada, respectively) is commonly

FIG. 11. The MSSS for CanSIPS vs Blended-5 at lead times of (top) 0, (middle) 2, and (bottom) 4 months. Rows are for equal lead times and columns for initial conditions at the start of the indicated month.

interpreted as being due to model biases, although potential skill is not necessarily constrained to be greater than actual skill (Kumar et al. 2014), and error in the model initialization of SWE and in the verification data may also play a role.

The long-range behavior of CanSIPS actual skill is illustrated in Fig. 13, which shows March-averaged ACC^2 between Blended-5 and CanSIPS SWE forecasts initialized at the beginning of February (lead time $t = 1$ month), December (lead $t = 3$ months), October (lead $t = 5$ months), and April (lead $t = 11$ months). The geographical distribution of ACC^2 is similar to that of PP in Fig. 7, but with somewhat reduced values. Overall, ACC^2 tends to decrease with lead time in much the same manner as PP and remains significant in large portions of Canada and the western United States at 3-month lead. For longer lead times, there is significant skill in the Canadian Arctic Archipelago (at 5-month lead), and in some portions of the western United States (at 5- and 11-month lead) where ENSO-teleconnected SWE anomalies occur (Fig. 8b) and PP is high (Figs. 7c,d). This suggests that CanSIPS is able to capitalize on this source of potential predictability, resulting in actual skill in

forecasting SWE anomalies in this region at long lead times.

d. Comparison of skill with anomaly persistence

A common practice when evaluating forecast skill is to compare the performance of model-based forecasts to that of simple reference forecasts. An example of such a reference forecast is the persistence “forecast,” which is obtained by persisting in time the anomalies that are present at the forecast initialization time. We persist the SWE anomalies from the forecast initial conditions, as this represents a source of information that corresponds well with uncertain available observations (Part I) and is readily available in real time, in contrast to the gridded verification products considered here that are generally not available in real time. More precisely, we employ ensemble-mean SWE averaged over the first day of the forecast as representing the initial conditions. The Blended-5 dataset is employed for verification as in the preceding subsections.

Figures 14a and 14b compare spatially averaged ACC based on the CanSIPS hindcasts (vertical axis) and on persistence of initial SWE anomalies (ACC_p) as

FIG. 12. (a)–(d) Annual cycle and (e), (f) scatterplots of spatial means for $\sqrt{PP}$ of CanSIPS and ACC of CanSIPS vs Blended-5 in (left) North America and (right) Canada as a function of lead time. The PP is the average of that in CanCM3 and CanCM4 computed separately. Black squares in (a)–(d) correspond to the forecast initialized on 1 Nov.

described above (horizontal axis). CanSIPS scores tend to exceed those obtained from persistence at all lead times, and although the differences are usually not large, they are consistently higher in cases where the area-averaged skills are the largest, primarily in

winter and early spring at short lead times. By contrast, the instances where persistence skills are slightly higher mainly occur for seasons and lead times where area-averaged skill is low, which suggests that sampling noise may affect the signs of these differences.

FIG. 13. The ACC^2 of March-averaged SWE for CanSIPS forecasts vs Blended-5 in North America at lead times of (a) 1, (b) 3, (c) 5, and (d) 11 months indicating the time from forecast initial values.

Maps of ACC minus ACC_p in North America are shown in Fig. 15. The tendency for CanSIPS scores to exceed persistence is clearly evident, although little difference is seen at short lead times at higher latitudes in the core of the snow season (December and February), where anomaly persistence accounts for much of the potential predictability according to the results of section 3. By contrast, CanSIPS scores tend to exceed those of persistence at lower latitudes, and at short lead time in the periphery of the snow season (October and April), where snowpacks are thinner and forecast snowfall and snowmelt have relatively

more influence. Negative values, indicating higher scores for persistence, occur mainly in the periphery of the snow season at longer leads when skill generally is relatively low, consistent with Fig. 14.

e. Comparison of SWE anomaly persistence in CanSIPS and observations

Persistence of initial SWE anomalies has been shown to be an important source of SWE potential predictability and actual skill in CanSIPS (sections 3b, 4d). Large autocorrelation in ensemble forecasts is known to be indicative of low dispersion among the ensemble

FIG. 14. Scatterplots of spatial means of (a),(b) ACC for CanSIPS vs Blended-5 (vertical axis) and ACC_p for CanSIPS initial anomaly vs Blended-5 (horizontal axis); and (c),(d) AC for CanSIPS (vertical axis) and Blended-5 (horizontal axis) in (left) North America and (right) Canada. Dots correspond to target months and colors denote lead times.

members (Kumar et al. 2014), which can be desirable or not depending on whether this indicates real predictability or artificial overconfidence in the forecast. It is thus worthwhile to compare SWE anomaly persistence in the forecasts with that in the observations. The relatively high skill up to 3 months of the persistent forecasts discussed in the previous section suggests that SWE anomalies in the observations are strongly autocorrelated. Here we compare temporal autocorrelation of SWE forecast anomalies in CanSIPS [Eq. (6)] with temporal autocorrelation of SWE anomalies in Blended-5 [Eq. (6) for one ensemble member being the observations].

Figures 14c and 14d compare spatially averaged ACC in North America and Canada, respectively, based on

CanSIPS hindcasts (vertical axis) and observations (horizontal axis). Considerable persistence is found in both CanSIPS and observations at up to 5 months lead time, depending on forecast date. Averaged AC in CanSIPS tends to be about the same as that of the verifying observations, with values as high as 0.8 in North America and 0.95 in Canada. In the whole of North America, averaged AC in CanSIPS tends to be slightly lower than that of the verifying observations at 0-month lead time, most likely because of relatively larger ensemble spread in regions of shallow and/or ephemeral snow in the midlatitudes. The similar degree of persistence of CanSIPS initial SWE anomalies to that of verifying observations suggests that the larger values of averaged $\sqrt{PP}$ relative to ACC in North America

FIG. 15. Difference of ACC for CanSIPS vs Blended-5 minus ACC_p for Blended-5 anomalies and CanSIPS initial anomalies at lead times of (top) 0, (middle) 1, and (bottom) 2 months. Rows are for equal lead times and columns for initial conditions at the start of the indicated month.

(discussed in [section 4c](#)) are because of a relatively large SWE signal variance and not the result of underdispersion of forecast SWE anomalies, which also would tend to increase PP ([Kumar et al. 2014](#)).

5. Summary and conclusions

We have evaluated potential and actual skill in CanSIPS multiseasonal forecasts of SWE in 1981–2010 over the Northern Hemisphere. The PP, defined as a potentially predictable variance fraction according to Eq. (4), was employed as a measure of potential skill, and ACC and MSSS based on verification against several SWE gridded observation-based products were employed as measures of actual skill.

The PP was found to vary strongly with region, initialization date, and lead time, as illustrated in [Fig. 3](#). Geographically, PP tends to be highest in the northern latitudes and regions of higher elevation, typically

having longer duration of snow cover and colder temperatures to preserve the snowpack. With respect to initialization date and lead time, highest PP values tend to occur early in forecasts that start well after snow onset, particularly in the core of the snow season (e.g., December and February). Although tending to decrease with lead time, these initially high PP values can remain appreciable throughout much of the snow season ([Fig. 4](#)). The persistence of initial SWE anomalies, quantified by the squared autocorrelation, was found to contribute substantially to these high PP values. By contrast, in forecasts initialized early in the snow season, PP is initially relatively low and decays more rapidly.

A temporal barrier for potential predictability occurs during springtime, typically around 1–4 weeks (depending on location and initialization date) after the total variance σ_T^2 reaches a maximum. Persistence of initial SWE anomalies has generally vanished within 3 weeks of maximum σ_T^2 and climatological SWE falls

below 1 mm after approximately 3–6 weeks, with little variation with respect to location and initialization date. However, there are regions such as the Rockies and western Canada where SWE is potentially predictable even across snow seasons. Such predictability cannot be attributed to anomaly persistence and must be a result of CanSIPS skill in predicting other climate variables, primarily temperature and precipitation, which influence snowfall and snowmelt. ENSO teleconnections likely play a role since the modeled influence of ENSO on temperature, precipitation, and SWE in these regions resembles that in observations and is largely independent of forecast lead time. Moreover, comparisons against verifying observations show that CanSIPS capitalizes in this source of potential predictability and has actual long-lead skill in forecasting SWE anomalies in this region.

Although we have only described long-lead skill at forecasting SWE in North America, there are also indications of long-lead skill, likely associated with ENSO, in the Karakoram and adjoining Tibetan Plateau (Sospedra-Alfonso and Merryfield 2016). Climate modes of variability other than ENSO may also contribute to SWE predictability and skill; for example, the combination of warm ENSO and negative NAO regimes is known to favor snowfall on the eastern coast of North America (e.g., Notaro et al. 2006; Seager et al. 2010). Such contributions are not examined in this work except to note that the short-lead skill of the CanSIPS models in predicting the winter NAO is modest at best (Butler et al. 2016), and the longer-lead skill for NAO prediction is likely lower still.

Comparisons of SWE forecasts from the two CanSIPS models show that, even though CanCM3 has initially larger noise variance than CanCM4, CanCM4 is generally more dispersive than CanCM3, with a faster increase in noise variance and a slightly faster potential predictability decay. Also, CanCM4 typically has larger signal-to-noise variance than CanCM3 at the beginning of the forecast (not shown), mostly due to the larger initial noise variance (i.e., ensemble spread) in CanCM3. However, the difference of PP between the two models is generally within 0.1 at the beginning of the forecasts initialized in December, February, and April.

The actual skill of CanSIPS forecasts of SWE was assessed using several observation-based verification products. The highest skill values were obtained using a blend of five SWE datasets developed by Mudryk et al. (2015) and denoted here as Blended-5. This could be a result of errors in the individual datasets tending to offset each other in the blended product, much as multimodel-averaged climate simulations and forecasts tend to outperform those of individual models. CanSIPS

was found to be reasonably skillful at forecasting SWE, with ACC values based on verification against Blended-5 averaged over the Northern Hemisphere exceeding 0.5 at 0-month lead in forecasts initialized from December through June and 0.3 at 1-month lead in forecast initialized from December through May. The MSSS exhibits similar behavior but with lesser values.

Potential and actual skills were related by comparing $\sqrt{\text{PP}}$ and ACC averaged over North America and Canada. These values were found to be approximately linearly related, with $\sqrt{\text{PP}}$ generally exceeding ACC. Based on these areal averages, SWE forecasts initialized relatively early in the snow season (e.g., November) remain skillful over a longer forecast range than forecasts initialized later in the snow season (e.g., March) because of a prediction barrier in the spring melt. The steep decline in potential and actual skill in May is a consequence of a large increase in noise variance that occurs during the snowmelt.

The skill of CanSIPS forecasts of SWE was found on average to exceed that of anomaly persistence, with such “added value” tending to be most pronounced near the periphery of the seasonal snowpack, as well as across consecutive snow years, in which case initial anomalies are erased by the spring melt so that persistence cannot contribute. The autocorrelation properties of forecast and observed SWE anomalies are similar, showing that CanSIPS captures this important property of the snowpack. However, CanSIPS actual skill falls somewhat short of the potential predictability, as is often (though not always) the case for climate forecasts. Possible reasons for this include systematic model errors, as well as differences between initial SWE anomalies in the forecasts and those in the verifying analysis; although CanSIPS initializes SWE reasonably realistically based on comparison with independent SWE analyses as demonstrated in Part I, skill stemming from the initial conditions will nonetheless be degraded by the differences that do exist between the SWE initial conditions and verifying analysis. This is in part because of the significant uncertainties that exist in all observation-based SWE analyses. Having an accurate SWE analysis available in real time would enable accurate forecast initialization and verification against a consistent data product and would likely increase skill scores and decrease the gap between actual skill and PP.

Acknowledgments. We are grateful to two anonymous reviewers for their comments and suggestions that helped improve the exposition of this work. This research is supported in part by the Natural Sciences and Engineering Research Council of Canada’s (NSERC) Climate Change and Atmospheric Research (CCAR)

program through the Canadian Sea Ice and Snow Evolution (CanSISE) Network (www.cansise.ca). R.S.A. is funded by an NSERC Visiting Fellowship.

REFERENCES

- Ambadan, J., A. A. Berg, and W. J. Merryfield, 2016: Influence of snow and soil moisture initialization on sub-seasonal predictability and forecast skill in boreal spring. *Climate Dyn.*, **47**, 49–65, doi:[10.1007/s00382-015-2821-9](https://doi.org/10.1007/s00382-015-2821-9).
- Balsamo, G., and Coauthors, 2015: ERA-Interim/Land: A global land water resources dataset. *Hydrol. Earth Syst. Sci.*, **19**, 389–407, doi:[10.5194/hess-19-389-2015](https://doi.org/10.5194/hess-19-389-2015).
- Bamzai, A. S., 2003: Relationship of snow cover variability and Arctic Oscillation index on a hierarchy of time scales. *Int. J. Climatol.*, **23**, 131–142, doi:[10.1002/joc.854](https://doi.org/10.1002/joc.854).
- Boer, G. J., 2004: Long time-scale potential predictability in an ensemble of coupled climate models. *Climate Dyn.*, **23**, 29–44, doi:[10.1007/s00382-004-0419-8](https://doi.org/10.1007/s00382-004-0419-8).
- , V. V. Kharin, and W. J. Merryfield, 2013: Decadal predictability and forecast skill. *Climate Dyn.*, **41**, 1817–1833, doi:[10.1007/s00382-013-1705-0](https://doi.org/10.1007/s00382-013-1705-0).
- Bojariu, R., and L. Gimeno, 2003: The role of snow cover fluctuations in multiannual NAO persistence. *Geophys. Res. Lett.*, **30**, 1156, doi:[10.1029/2002GL015651](https://doi.org/10.1029/2002GL015651).
- Brun, E., V. Vionnet, A. Boone, B. Decharme, Y. Peings, R. Valette, F. Karbou, and S. Morin, 2013: Simulation of northern Eurasian local snow depth, mass, and density using a detailed snowpack model and meteorological reanalyses. *J. Hydrometeorol.*, **14**, 203–219, doi:[10.1175/JHM-D-12-012.1](https://doi.org/10.1175/JHM-D-12-012.1).
- Butler, A. H., and Coauthors, 2016: The Climate-system Historical Forecast Project: Do stratosphere-resolving models make better seasonal climate predictions in boreal winter? *Quart. J. Roy. Meteor. Soc.*, **142**, 1413–1427, doi:[10.1002/qj.2743](https://doi.org/10.1002/qj.2743).
- Clark, M. P., M. C. Serreze, and D. A. Robinson, 1999: Atmospheric controls on Eurasian snow extent. *Int. J. Climatol.*, **19**, 27–40, doi:[10.1002/\(SICI\)1097-0088\(199901\)19:1<27::AID-JOC346>3.0.CO;2-N](https://doi.org/10.1002/(SICI)1097-0088(199901)19:1<27::AID-JOC346>3.0.CO;2-N).
- Cohen, J. L., and D. Entekhabi, 1999: Eurasian snow cover variability and Northern Hemisphere climate predictability. *Geophys. Res. Lett.*, **26**, 345–348, doi:[10.1029/1998GL900321](https://doi.org/10.1029/1998GL900321).
- , and C. Fletcher, 2007: Improved skill of Northern Hemisphere winter surface temperature predictions based on land-atmosphere fall anomalies. *J. Climate*, **20**, 4118–4132, doi:[10.1175/JCLI4241.1](https://doi.org/10.1175/JCLI4241.1).
- , K. Saito, and D. Entekhabi, 2001: The role of the Siberian High in Northern Hemisphere climate variability. *Geophys. Res. Lett.*, **28**, 299–302, doi:[10.1029/2000GL011927](https://doi.org/10.1029/2000GL011927).
- Corti, S., F. Molteni, and C. Branković, 2000: Predictability of snow-depth anomalies over Eurasia and associated circulation patterns. *Quart. J. Roy. Meteor. Soc.*, **126**, 241–262, doi:[10.1002/qj.49712656212](https://doi.org/10.1002/qj.49712656212).
- Dee, P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi:[10.1002/qj.828](https://doi.org/10.1002/qj.828).
- DelSole, T., A. Kumar, and J. Bhashkar, 2013: Potential seasonal predictability: Comparison between empirical and dynamical model estimates. *Geophys. Res. Lett.*, **40**, 3200–3206, doi:[10.1002/grl.50581](https://doi.org/10.1002/grl.50581).
- , J. Nattala, and M. K. Tippett, 2014: Skill improvement from increased ensemble size and model diversity. *Geophys. Res. Lett.*, **41**, 7331–7342, doi:[10.1002/2014GL060133](https://doi.org/10.1002/2014GL060133).
- Dewey, K. F., 1977: Daily maximum and minimum temperature forecasts and the influence of snow cover. *Mon. Wea. Rev.*, **105**, 1594–1597, doi:[10.1175/1520-0493\(1977\)105<1594:DMAMTF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1977)105<1594:DMAMTF>2.0.CO;2).
- Ferranti, L., and F. Molteni, 1999: Ensemble simulations of Eurasian snow-depth anomalies and their influence on the summer Asian monsoon. *Quart. J. Roy. Meteor. Soc.*, **125**, 2597–2610, doi:[10.1002/qj.49712555913](https://doi.org/10.1002/qj.49712555913).
- Gong, G., D. Entekhabi, J. Cohen, and D. Robinson, 2004: Sensitivity of atmospheric response to modeled snow anomaly characteristics. *J. Geophys. Res.*, **109**, D06107, doi:[10.1029/2003JD004160](https://doi.org/10.1029/2003JD004160).
- Groisman, P. Y., T. R. Karl, and R. W. Knight, 1994: Changes of snow cover, temperature, and radiative heat balance over the Northern Hemisphere. *J. Climate*, **7**, 1633–1656, doi:[10.1175/1520-0442\(1994\)007<1633:COSCTA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1994)007<1633:COSCTA>2.0.CO;2).
- Jeong, J.-H., H. W. Linderholm, S.-H. Woo, C. Folland, B.-M. Kim, S.-J. Kim, and D. Chen, 2013: Impacts of snow initialization on subseasonal forecasts of surface air temperature for the cold season. *J. Climate*, **26**, 1956–1972, doi:[10.1175/JCLI-D-12-00159.1](https://doi.org/10.1175/JCLI-D-12-00159.1).
- Kirtman, B., and Coauthors, 2014: The North American Multi-model Ensemble: Phase-1 seasonal-to-interannual prediction; Phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc.*, **95**, 585–601, doi:[10.1175/BAMS-D-12-00050.1](https://doi.org/10.1175/BAMS-D-12-00050.1).
- Kumar, A., P. Peng, and M. Chen, 2014: Is there a relationship between potential and actual skill? *Mon. Wea. Rev.*, **142**, 2220–2227, doi:[10.1175/MWR-D-13-00287.1](https://doi.org/10.1175/MWR-D-13-00287.1).
- Lin, H., and Z. Wu, 2011: Contribution of the autumn Tibetan Plateau snow cover to seasonal prediction of North America winter temperature. *J. Climate*, **24**, 2801–2813, doi:[10.1175/2010JCLI3889.1](https://doi.org/10.1175/2010JCLI3889.1).
- Madden, R. A., 1976: Estimates of natural variability of time-averaged sea level pressure. *Mon. Wea. Rev.*, **104**, 942–952, doi:[10.1175/1520-0493\(1976\)104<0942:EOTNVO>2.0.CO;2](https://doi.org/10.1175/1520-0493(1976)104<0942:EOTNVO>2.0.CO;2).
- Martineu, C., J.-Y. Caneill, and R. Sadourny, 1999: Potential predictability of European winters from the analysis of seasonal simulations with an AGCM. *J. Climate*, **12**, 3033–3061, doi:[10.1175/1520-0442\(1999\)012<3033:PPOEWF>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<3033:PPOEWF>2.0.CO;2).
- Merryfield, W. J., B. Denis, J.-S. Fontecilla, W.-S. Lee, V. V. Kharin, J. Hodgson, and B. Archambault, 2011: The Canadian Seasonal to Interannual Prediction System (CanSIPS). Tech. Note, Environment Canada, 51 pp. [Available online at http://collaboration.cmc.ec.gc.ca/cmc/cmoei/product_guide/docs/lib/op_systems/doc_opchanges/technote_cansips_20111124_e.pdf].
- , W.-S. Lee, G. J. Boer, V. V. Kharin, J. F. Scinocca, G. M. Flato, R. S. Ajayamohan, and J. C. Fyfe, 2013: The Canadian Seasonal to Interannual Prediction System. Part I: Models and initialization. *Mon. Wea. Rev.*, **141**, 2910–2945, doi:[10.1175/MWR-D-12-00216.1](https://doi.org/10.1175/MWR-D-12-00216.1).
- Mudryk, L. R., C. Derksen, P. J. Kushner, and R. Brown, 2015: Characterization of Northern Hemisphere snow water equivalent datasets, 1981–2010. *J. Climate*, **28**, 8037–8051, doi:[10.1175/JCLI-D-15-0229.1](https://doi.org/10.1175/JCLI-D-15-0229.1).
- Murphy, A. H., 1988: Skill scores based on the mean square error and their relationships to the correlation coefficient. *Mon. Wea. Rev.*, **116**, 2417, doi:[10.1175/1520-0493\(1988\)116<2417:SSBOTM>2.0.CO;2](https://doi.org/10.1175/1520-0493(1988)116<2417:SSBOTM>2.0.CO;2).
- Notaro, M., W. C. Wang, and W. Gong, 2006: Model and observational analysis of northeast U. S. regional climate and its relationship to the PNA and NAO patterns during early winter. *Mon. Wea. Rev.*, **134**, 3479–3505, doi:[10.1175/MWR3234.1](https://doi.org/10.1175/MWR3234.1).

- Orsolini, Y. J., R. Senan, G. Balsamo, F. J. Doblas-Reyes, F. Vitart, A. Weisheimer, A. Carrasco, and R. E. Benestad, 2013: Impact of snow initialization on sub-seasonal forecasts. *Climate Dyn.*, **41**, 1969–1982, doi:[10.1007/s00382-013-1782-0](https://doi.org/10.1007/s00382-013-1782-0).
- Peings, Y., H. Douville, R. Alkama, and B. Decharme, 2011: Snow contribution to springtime atmospheric predictability over the second half of the twentieth century. *Climate Dyn.*, **37**, 985–1004, doi:[10.1007/s00382-010-0884-1](https://doi.org/10.1007/s00382-010-0884-1).
- Reichle, R. H., 2012: The MERRA-Land Data Product. GMAO Office Note 3 (Version 1.2), NASA, 38 pp. [Available online at <https://gmao.gsfc.nasa.gov/pubs/docs/Reichle541.pdf>.]
- , R. Koster, G. D. Lannoy, B. Forman, Q. Liu, S. Mahanam, and A. Toure, 2011: Assessment and enhancement of MERRA land surface hydrology estimates. *J. Climate*, **24**, 6322–6338, doi:[10.1175/JCLI-D-10-05033.1](https://doi.org/10.1175/JCLI-D-10-05033.1).
- Riddle, E. E., A. H. Butler, J. C. Furtado, J. L. Cohen, and A. Kumar, 2013: CFSv2 ensemble prediction of the wintertime Arctic Oscillation. *Climate Dyn.*, **41**, 1099–1116, doi:[10.1007/s00382-013-1850-5](https://doi.org/10.1007/s00382-013-1850-5).
- Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648, doi:[10.1175/JCLI-D-11-00015.1](https://doi.org/10.1175/JCLI-D-11-00015.1).
- Rodell, M., and Coauthors, 2004: The Global Land Data Assimilation System. *Bull. Amer. Meteor. Soc.*, **85**, 381–394, doi:[10.1175/BAMS-85-3-381](https://doi.org/10.1175/BAMS-85-3-381).
- Rowell, D. P., 1998: Assessing potential seasonal predictability with an ensemble of multidecadal GCM simulations. *J. Climate*, **11**, 109–120, doi:[10.1175/1520-0442\(1998\)011<0109:APSPA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1998)011<0109:APSPA>2.0.CO;2).
- Saito, K., and J. Cohen, 2003: The potential role of snow cover in forcing interannual variability of the major Northern Hemisphere mode. *Geophys. Res. Lett.*, **30**, 1302, doi:[10.1029/2002GL016341](https://doi.org/10.1029/2002GL016341).
- , —, and D. Entekhabi, 2001: Evolution of atmospheric response to early-season Eurasian snow cover anomalies. *Mon. Wea. Rev.*, **129**, 2746–2760, doi:[10.1175/1520-0493\(2001\)129<2746:EOARTE>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2746:EOARTE>2.0.CO;2).
- Schlosser, C. A., and P. A. Dirmeyer, 2001: Potential predictability of Eurasia snow cover. *Atmos. Sci. Lett.*, **2**, 1–8, doi:[10.1006/asle.2001.0037](https://doi.org/10.1006/asle.2001.0037).
- Seager, R., Y. Kushnir, J. Nakamura, M. Ting, and N. Naik, 2010: Northern Hemisphere winter snow anomalies: ENSO, NAO and the winter of 2009/10. *Geophys. Res. Lett.*, **37**, L14703, doi:[10.1029/2010GL043830](https://doi.org/10.1029/2010GL043830).
- Serreze, M. C., F. Carse, R. G. Barry, and J. C. Rogers, 1997: Icelandic low cyclone activity: Climatological features, linkages with the NAO, and relationships with recent changes in the Northern Hemispheric circulation. *J. Climate*, **10**, 453–464, doi:[10.1175/1520-0442\(1997\)010<0453:ILCACF>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<0453:ILCACF>2.0.CO;2).
- Shaman, J., and E. Tziperman, 2005: The effect of ENSO on Tibetan Plateau snow depth: A stationary wave teleconnection mechanism and implications for the South Asian Monsoons. *J. Climate*, **18**, 2067–2079, doi:[10.1175/JCLI3391.1](https://doi.org/10.1175/JCLI3391.1).
- Shongwe, M. E., C. A. T. Ferro, C. A. S. Coelho, and G. J. van Oldenborgh, 2007: Predictability of cold spring seasons in Europe. *Mon. Wea. Rev.*, **135**, 4185–4201, doi:[10.1175/2007MWR2094.1](https://doi.org/10.1175/2007MWR2094.1).
- Sobolowski, S., and A. Frei, 2007: Lagged relationships between North American snow mass and atmospheric teleconnections indices. *Int. J. Climatol.*, **27**, 221–231, doi:[10.1002/joc.1395](https://doi.org/10.1002/joc.1395).
- Sospedra-Alfonso, R., and W. J. Merryfield, 2016: Potential and actual predictability of snow water equivalent in historical forecasts of the Canadian Fourth Generation Coupled Climate Model (CanCM4). *40th NOAA Annual Climate Diagnostics and Prediction Workshop*, Denver, CO, NOAA, 129–134. [Available online at <http://www.nws.noaa.gov/ost/climate/STIP/40CDPW/40cdpw-RSospedra-Alfonso.pdf>.]
- , J. R. Melton, and W. J. Merryfield, 2015: Effects of temperature and precipitation on snowpack variability in the central Rocky Mountains as a function of elevation. *Geophys. Res. Lett.*, **42**, 4429–4438, doi:[10.1002/2015GL063898](https://doi.org/10.1002/2015GL063898).
- , L. Mudryk, W. J. Merryfield, and C. Derksen, 2016: Representation of snow in the Canadian Seasonal to Interannual Prediction System. Part I: Initialization. *J. Hydrometeorol.*, **17**, 1467–1488, doi:[10.1175/JHM-D-14-0223.1](https://doi.org/10.1175/JHM-D-14-0223.1).
- Takala, M., K. Luojus, J. Pulliainen, C. Derksen, J. Lemmetyinen, J.-P. Karna, J. Koskinen, and B. Bojkov, 2011: Estimating Northern Hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements. *Remote Sens. Environ.*, **115**, 3517–3529, doi:[10.1016/j.rse.2011.08.014](https://doi.org/10.1016/j.rse.2011.08.014).
- von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, 484 pp.
- Webster, P. J., 1995: The annual cycle and the predictability of the tropical coupled ocean–atmosphere system. *Meteor. Atmos. Phys.*, **56**, 33–55, doi:[10.1007/BF01022520](https://doi.org/10.1007/BF01022520).
- Wu, Z., J. Li, Z. Jiang, and T. Ma, 2012: Modulation of the Tibetan Plateau snow cover on the ENSO teleconnections: From the East Asian summer monsoon perspective. *J. Climate*, **25**, 2481–2489, doi:[10.1175/JCLI-D-11-00135.1](https://doi.org/10.1175/JCLI-D-11-00135.1).
- Yang, S., 1996: ENSO–snow–monsoon associations and seasonal–interannual predictions. *Int. J. Climatol.*, **16**, 125–134, doi:[10.1002/\(SICI\)1097-0088\(199602\)16:2<125::AID-JOC999>3.0.CO;2-V](https://doi.org/10.1002/(SICI)1097-0088(199602)16:2<125::AID-JOC999>3.0.CO;2-V).
- Yeh, T. C., R. T. Wetherald, and S. Manabe, 1983: A model study of the short term climatic and hydrologic effects of sudden snow-cover removal. *Mon. Wea. Rev.*, **111**, 1013–1024, doi:[10.1175/1520-0493\(1983\)111<1013:AMSOTS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1983)111<1013:AMSOTS>2.0.CO;2).
- Zwiers, F. W., 1996: Interannual variability and predictability in an ensemble of AMIP climate simulations conducted with the CCC GCM2. *Climate Dyn.*, **12**, 825–847, doi:[10.1007/s003820050146](https://doi.org/10.1007/s003820050146).