Assessment of Snow, Sea Ice, and Related Climate Processes in Canada's Earth-System Model and Climate Prediction System

Paul J. Kushner[1], Lawrence R. Mudryk[2], William Merryfield[2], Jaison T. Ambadan[3], Aaron Berg[3], Adéline Bichet[4], Ross Brown[2], Christopher P. Derksen[2], Stephen J. Déry[5], Arlan Dirkson[6], Greg Flato[2], Christopher G. Fletcher[7], John C. Fyfe[2], Nathan Gillett[2], Christian Haas[8,9], Stephen Howell[2], Frédéric Laliberté[2], Kelly McCusker[10], Michael Sigmond[2], Reinel Sospedra-Alfonso[2], Neil F. Tandon[2], Chad Thackeray[7], Bruno Tremblay[11], Francis W. Zwiers[12]

[1]Department of Physics, University of Toronto, M5S 1A7, Canada
[2]Climate Research Division, Environment and Climate Change Canada, M3H 5T4, Canada
[3]Department of Geography, University of Guelph, N1G 2W1, Canada
[4]CNRS-LGGE/MEOM, 38041 Grenoble, France
[5]Department of Environmental Science, University of Northern British Columbia, V2N 4Z9, Canada
[6]School of Earth and Ocean Sciences, University of Victoria, V8W 2Y2, Canada
[7]Department of Geography and Environmental Management, University of Waterloo, N2L 3G1, Canada
[8]Department of Earth and Space Science and Engineering, York University, M3J 1P3, Canada
[9]Climate Sciences Division, Alfred Wegener Institute, 27570, Germany
[10]Department of Atmospheric Sciences, University of Washington, 98195-1640, United States
[11]Department of Atmospheric and Oceanic Sciences, McGill University, H3A 0B9, Canada
[12]Pacific Climate Impacts Consortium, University of Victoria, V8P 5C2, Canada

Correspondence to: Paul J. Kushner (paul.kushner@utoronto.ca)

Abstract. This study assesses the ability of the Canadian Earth-system Model 2 (CanESM2) and the Canadian Seasonal to Interannual Prediction System (CanSIPS) to simulate and predict snow and sea ice from seasonal to multi-decadal timescales, with a focus on the Canadian sector. To account for observational uncertainty, model structural uncertainty, and internal climate variability, the analysis uses multi-source observations, multiple Earth-System Models (ESMs) in Phase 5 of the Coupled Model Intercomparison Project (CMIP5), and large initial condition ensembles of CanESM2 and other models. It is found that the ability of the CanESM2 simulation to capture snow-related climate parameters, such as cold-region surface temperature and precipitation, lies within the range of currently available international models. Accounting for the considerable disagreement among satellite-era observational datasets on the distribution of snow water equivalent, CanESM2 has too much springtime snow mass over the Canada, reflecting a broader Northern Hemisphere positive bias. Biases in seasonal snow cover extent are generally less pronounced. CanESM2 also exhibits retreat of springtime snow generally greater than observational estimates, after accounting for observational uncertainty and internal variability. Sea ice is biased low in the Canadian Arctic, which makes it difficult to assess the realism of long-term sea-ice trends there. The strengths and weaknesses of the modeling system need to be understood as a practical tradeoff: the Canadian models are relatively inexpensive computationally because of their moderate resolution, thus enabling their use in operational seasonal prediction and for generating large ensembles of multidecadal simulations. Improvements in climate prediction systems like CanSIPS rely not just on simulation quality but

also on using novel observational constraints and the ready transfer of research to an operational setting. Improvements in seasonal forecasting practice arising from recent research include accurate initialization of snow and frozen soil, accounting for observational uncertainty in forecast verification, and sea-ice thickness initialization using statistical predictors available in real time.

5 **1 Introduction**

Seasonal snow cover and sea ice are integral to the cultural identity, history, and economy of northern nations like Canada. They also exert an enormous physical influence on the earth system, ranging from local interactions with winds and temperatures in the Arctic and snow-covered regions, to larger-scale interactions with weather systems and ocean circulation, to global-scale influences on the Earth's energy balance. In recent decades, dramatic changes in Canada's snow cover and sea

10 ice have been witnessed and documented (Derksen et al., 2012; Najafi et al., 2015). This has driven the need to better understand and predict these fields for the coming seasons, years and decades. To address this need, Canada has helped lead the global effort to better observe and model snow, sea ice, and related climate parameters (such as northern high latitude land-surface temperature and precipitation). This effort includes Canadian contributions to the International Polar Year (e.g. Kulkarni et al., 2012), to the development of earth system model and climate prediction systems (Merryfield et al., 2013a;

15 Sigmond et al., 2013; van den Hurk et al., 2016), and to leadership of ongoing field and remote sensing efforts (King et al., 2015).

As part of Canada's larger effort in snow and sea-ice research, the focus here is on seasonal and longer timescale prediction of terrestrial snow, sea-ice cover, and related climate variability. The purpose of this paper is to evaluate the ability of Canada's

20 current earth system model (ESM) and climate prediction system to carry out this kind of prediction in the context of the development of new observational products. This work was undertaken by the Canadian Sea Ice and Snow Evolution Network (CanSISE), a core project of the Climate Change and Atmospheric Research Program of the Natural Sciences and Engineering Research Council of Canada (CCAR/NSERC)[1]. Model evaluation, which typically compares a model to observations, needs to account for several sources of uncertainty, including impacts of spatial and temporal sampling in the presence of internal

25 climate variability and observational uncertainty (whether instrumental error or errors related to data processing and retrieval

---

[1] The CanSISE Network was funded for five years starting in 2013. It is a partnership between several Canadian Universities (Toronto, British Columbia, Guelph, McGill, Northern British Columbia, Victoria, Waterloo, and York); Environment and Climate Change Canada [ECCC; research groups include the Canadian Centre for Climate Modelling and Analysis (CCCma) and the Climate Processes Section, both in the Climate Research Division; and the Canadian Ice Service (CIS)], and the Pacific Climate Impacts Consortium (PCIC).

systems). Our evaluation of Canadian models is helped by ready access and comparison with output from internationally available models, to provide a suitable scientific context.

This study focuses on snow, sea ice and related climate parameters and processes relevant to the Canadian land mass and the pan-Arctic region. The Canadian ESM and climate prediction system has been studied in a variety of related settings (e.g. Arora et al. 2011; Merryfield et al. 2013a, b; Gillet et al. 2012; Sigmond et al. 2013, Kirtman et al. 2013; Flato et al. 2013). We here seek to more fully assess simulation and prediction of seasonal snow cover and regional sea-ice variability accompanied by a more complete characterization of observational uncertainty, model structural uncertainty, and internal climate variability. After reviewing the current generation Canadian Seasonal to Interannual Prediction System and Canadian Earth-System Model 2 (CanSIPS and CanESM2; section 2), we characterize climatological behavior and trends for snow and sea ice in these systems (section 3), provide an overview of recent developments in seasonal snow and sea-ice prediction (section 4), and conclude (section 5) with a summary and discussion of new directions for prediction-system development.

2 Models and data used

In section 3, our analysis will focus on CanESM2 (Arora et al., 2011; Scinocca et al., 2016). This is the earth-system model used by the Canadian Centre for Climate Modeling and Analysis (CCCma) of Environment and Climate Change Canada (ECCC) for its contribution to Phase 5 of the Coupled Model Intercomparison Project (CMIP5). CanESM2 combines atmosphere, ocean, land-surface (including snow), sea-ice, and carbon-cycle components in a coupled framework in which all model components interact. The system can simulate the past and projected state of global temperature, circulation, carbon dioxide concentrations, etc. under the influence of external forcing, but independently of assimilated ocean and atmospheric initialization data. As summarized in Arora et al. (2011), the atmospheric and oceanic components are the fourth generation atmospheric and oceanic general circulation models CanAM4 and CanOM4, the prognostic carbon-cycle components are the Canadian Model of Ocean Carbon (CMOC) and the Canadian Terrestrial Ecosystem Model (CTEM), the land-surface component (including the snow scheme) is version 3 of the Canadian Land Surface Scheme (CLASS), and the sea-ice component is the Flato and Hibler (1992) cavitating fluid scheme. As with most other models participating in CMIP5, CanESM2 does not use flux adjustments that artificially constrain the climate system to be in a state of energy and water balance. Following CMIP5 protocols, the model includes concentrations and emissions of greenhouse gases, aerosol and aerosol-precursor emssions, and presriptions for land-cover change (Arora et al. 2011).

CanESM2 has moderate spatial resolution compared to other CMIP5 models (approximately 2.8º horizontal grid spacing and up to 35 vertical levels in the atmosphere; approximately 100 km horizontal grid spacing and up to 40 levels in the ocean). This resolution accounts for constraints on available computing resources. It sufficiently resolves salient features of the global atmosphere-ocean circulation while still permitting the execution of large initial condition ensembles of model simulations to

adequately sample internal variability under different external forcings. We note that ECCC has also made a complementary multi-year investment in regional climate modelling (Scinocca et al., 2016) to provide higher resolution over North America (with versions at 50 km and at 25 km grid resolution) to address the shortcomings of coarse resolution.

5   In section 4, we consider the application over Canada of CanSIPS (Merryfield et al. 2013a), the operational prediction system of the Canadian Metoerological Centre (CMC) for climate variability on seasonal to interannual (several-month to multiple-year) timescales. Like CanESM2, CanSIPS is also a multi-component interactive system. However, unlike CanESM2, when operating as a prediction system CanSIPS starts from an initial state that approximates the real-world state at a given initial time. CanSIPS includes 1) a data assimilation system that estimates realistic initial states of the atmosphere, ocean, land, and
10   sea ice to start the forecasts; 2) two separate coupled climate models (the earlier generation Canadian Coupled Model 3 [CanCM3] and the later generation Canadian Coupled Model [CanCM4]) that advance the simulated system from this initial condition (using an ensemble size of 10 for each model); and 3) diagnostic systems to analyze the output and generate useful forecasts for operational use within ECCC's Meteorological Service of Canada (e.g. Fig. 14 below and the probabilistic seasonal forecast at https://weather.gc.ca/saisons/prob_e.html). Evaluations of CanSIPS need to consider all three parts of the
15   seasonal prediction system. CanCM4 has the same atmosphere, ocean, land, and sea ice components as CanESM2, but does not include CanESM2's carbon cycle components. CanCM3 has the previous-generation atmosphere and ocean components relative to CanCM4 and CanESM2, but the same land-surface and sea-ice components as CanCM4 and CanESM2 (Merryfield et al. 2013a).

20   Merryfield et al. (2013a) summarize the performance of CanCM3 and CanCM4 when the models are run independently of assimilated data. CanCM4 reduces the global mean absolute error of ocean surface temperatures compared to CanCM3, indicating an overall improvement in the coupled ocean atmosphere state that results from improved physical parameterizations and finer resolution. Relative to CanCM3 and observations, CanCM4 tends to warm more rapidly under the effects of anthropogenic radiative forcing over the 1970-2009 period. In CanCM3, the simulation is characterized by excessive pan-
25   Arctic sea-ice cover in summer and winter and a small rate of sea-ice loss compared to observations. In CanCM4, while there is still excessive sea-ice cover in winter, there is too little sea ice in summer (see section 3 below). The rate of sea-ice loss in CanCM4 is more in line with recent observations than that in CanCM3 (Stroeve et al., 2012); however, caution is required to interpret recent sea-ice loss rates in light of the large amount of multidecadal variability expected in these trends (e.g. Notz, 2012; Swart et al., 2015). Because CMIP5 simulations were carried out with CanESM2 but not CanCM4, the simulations
30   required to do a clean comparison of CanCM4 and CanESM2, and thus gauge the impact of carbon-cycle processes on simulation quality, are not available.

When run as a prediction system, CanSIPS, combining CanCM3 and CanCM4, is able to show multi-month skill in seasonal forecasts of detrended sea-ice area anomalies, comparable to that obtained in other modelling systems (Merryfield et al.,

2013b), and generally enhanced skill relative to a statistical persistence forecast (Sigmond et al., 2013). The assessed skill depends on the verification dataset (Sigmond et al. 2013), especially for total (non-detrended) anomalies. Such issues will be revisited in this study.

Our assessment of CanSIPS/CanESM2 is enhanced by two recent research products arising from CanSISE: the Blended-5 snow water equivalent (SWE) dataset of Mudryk et al. (2015) and the CanESM2 Large Ensemble of simulations from CanESM2. The Blended-5 dataset addresses the need for a SWE verification dataset, and, potentially, for initialization of snow-related parameters in CanSIPS and other prediction systems. Blended-5 builds on long-term work of ECCC (e.g. Brown et al., 2010; Derksen and Brown, 2012; Brown and Derksen, 2013) and consists of an ensemble of gridded SWE datasets over 1981-2010 from a variety of sources including remote sensing, land surface assimilation systems, and reanalysis driven snow models. The papers of Mudryk et al. (2015, 2017) detail the components, quality assessment, and characteristics of the Blended-5 dataset.

The use of large initial condition ensembles has afforded a renewed assessment of the impacts of natural climate variability on recent and projected climatic variability and trends (e.g. Deser et al. 2012, Kay et al. 2015). The CanSISE project team designed the CanESM2 Large Ensemble (e.g. Sigmond and Fyfe, 2016), which consists of four sets of 50 simulations each of CanESM2 that examine the impact of natural and anthropogenic forcings over the period 1950-2100 in the presence of internal climate variability. For each of the five realizations run by CCCma for CMIP5, a new set of ten simulations is generated by slightly perturbing the atmospheric state at the beginning of 1950, with ten different perturbations. These 50 realizations are then integrated forward until 2005 with CMIP5 historical forcings (Taylor et al., 2012); from 2006-2100, the RCP8.5 CMIP5 scenario is used. The first ensemble set, which applies all available external forcings, will be the one used here. Additional sets of attribution integrations not analyzed here include just historic natural external forcings (solar and volcanic), just historic anthropogenic aerosol forcings, and just stratospheric ozone forcing. Each realization in each set is identical apart from its initial conditions. Thus, the ensemble mean of a given 50 member set is characterized by about a factor seven less internal variability than a single realization, and therefore provides a relatively robust estimate of that set's externally forced signal. The distinctively forced ensembles permit attribution of observed climate signals to different external forcings. The CanESM2 Large Ensemble has been used in several current and ongoing studies (Sigmond and Fyfe, 2016; McKusker et al., 2016; Gagné et al., 2016; Fyfe et al., 2016; Mudryk et al., 2017; Kirchmeier-Young et al., 2016). We also use similar initial condition ensembles of the National Center for Atmospheric Research Community Earth-System Model 1 (NCAR CESM1; Kay et al., 2015) and the NCAR Community Climate System Model 4 (CCSM4; Mudryk et al., 2013). Other observational sources and modelling results used in this study will be described in the text and figure captions. In what follows, our primary focus is on Canada and the pan-Arctic, placed in the context of Northern Hemisphere climate.

3 CanESM2 climatology and trends

3.1 Observed and simulated terrestrial snow climatology

We first evaluate the climatological characteristics of CanESM2's land-surface temperature, precipitation, and snow water equivalent (SWE) for the Canadian land mass. In winter and spring, the distribution of land-surface temperature over Canada is well reproduced in CanESM2, although a warm bias is evident in both seasons (left and central panels of the top rows of Figs. 1-2). The Taylor (2001) diagram for land-surface temperature (top row, right panels of Figs. 1-2) shows that CanESM2 compares well to other CMIP5 models (the same models and realizations are used in Mudryk et al., 2017) in capturing the spatial pattern and correlation with observations, although the spatial gradients are somewhat stronger than observed for winter (as shown by the distance of the CanESM2 point from the origin in the Taylor diagram), associated with a stronger south-to-north temperature gradient than observed. In JFM precipitation (central row of Figs. 1-2), the general pattern and spatial gradient strength are captured in the model, but there is excessive wintertime precipitation over most of Canada, including the Western Cordillera, sub-Arctic and Arctic, in both seasons. This excessive precipitation contributes towards a bias of excessive SWE over much of Western Canada and the Canadian Sub-Arctic that is particularly pronounced in spring (lower row of Figs. 1-2). CanESM2 SWE has greater spatial variance than the Blended-5 SWE ensemble mean and most of the individual component datasets of the Blended 5 (lower right panel of Figs. 1-2). Generally speaking, the Taylor diagrams in Figs. 1-2 suggest that CanESM2 is well within the state of the art of current models for the climate parameters related to seasonal snow cover.

Observed SWE climatology, variability and trends are relatively non-robust compared to variables such as land-surface temperature (Mudryk et al., 2015, 2017) and for this reason we assess some aspects of the spread across the Blended-5 SWE datasets. Individual observational datasets contributing to Blended-5 also show stronger spatial gradients than the Blended-5 mean (circles filled with light brown in the Taylor diagram in the lower right panels of Figs. 1-2). This is in part expected because the observational mean will cancel random errors. However, this also suggests that there is considerable uncertainty in the spatial variance, and so it is difficult to assess how realistically spatial variance is captured in CanESM2 and the other CMIP5 models. This observational uncertainty is also evident in the seasonal cycle of total snow mass aggregated for Canada and the Northern Hemisphere, as well as geographic subregions (top two rows of Fig. 3, gray shading). For example, for the Northern Hemisphere (second row, first column), the range in Blended-5 estimates of peak snow mass in February is over 50% of the average, and is driven mainly by uncertainty in Arctic (second row, third column) and alpine regions (second row, fourth column). The individual datasets in the Blended-5 product are not shown in Fig. 3, but their characteristics are discussed in Mudryk et al. 2015. The NASA Global Land Data Assimilation System (GLDAS) provides an estimate well below the

multi-dataset mean, the MERRA reanalysis dataset typically provides a central estimate, and the maximum estimate varies with region among the remaining three datasets.

After accounting for the considerable observational uncertainty in total snow mass, it is nevertheless possible to assess the realism of CanESM2's simulation. The CanESM2 snow mass over Canada is plotted as originally available on the model's land grid (light teal points, shown only in the first row and first column of Fig. 3) and as adjusted to reflect the observational mask which is on a finer scale (dark teal points). The adjustment is downward because some of the model's snow mass is located in grid cells that, in reality, are only partially covered by land. The positive bias of CanESM2 relative to the observational mean (Figs. 1-2) is evident in the seasonal cycle in snow mass over Canada (first row of Fig. 3), especially in spring in midlatitudes, and reflects a broader Northern Hemisphere positive bias (second row of Fig. 3). Sospedra-Alfonso et al. (2016b) also find that CanCM4 model that contribute to CanSIPS features a positive springtime SWE bias. For comparison, the CMIP5 multi-model mean over Canada (red x symbols) does not feature as pronounced a bias. The CMIP5 model range (not shown) spans from the lowest observational estimate to above the CanESM2, but CanESM2 is on the high end, especially during spring in the midlatitudes. Our assessment is that, especially in midlatitudes, CanESM2 simulates excessive springtime snow associated with excessive wintertime precipitation building up throughout winter and into spring (middle rows of Figs. 1-2).

The seasonal cycle of snow-cover extent (SCE) is shown in the third and fourth rows of Fig. 3. Here, observed SCE is derived from the Blended-5 dataset by converting SWE to SCE using a threshold of 4 mm; this threshold was tested in Mudryk et al., 2017. For the observational products in the Blended-5 dataset, the relative uncertainty in SCE is generally less than for snow mass. For example, the observational range in peak Northern Hemisphere SCE in January is about 15%, and is dominated by uncertainty in midlatitude and Arctic regions. A modest positive springtime excess of SCE is evident for CanESM2 for Canada and the Northern Hemisphere. On the whole, observed SCE is better constrained than observed snow mass, and simulated SCE is generally more realistic than simulated snow mass for CanESM2, as well as for the average over the CMIP5 models.

3.2 Observed and simulated trends in terrestrial snow
A standard target for snow process analysis in climate models is trends of SCE, which are strongly temperature controlled (e.g. Brutel-Vuilmet et al., 2013; Mudryk et al., 2017). Assessing the ability of models to capture these trends needs to account for natural variability, forced variability, observational uncertainty, and intermodel differences. We show in Fig. 4 the trends in snow cover extent (SCE) derived from the Blended-5 dataset for the Northern Hemisphere in January-March and April-June. In both seasons, there is a spread of observed seasonal snow cover reduction estimates from 0.0 to -0.5 million km^2 per decade in winter (based on a simple interquartile range for this small number of observational datasets), and from -0.1 to -0.6 million km^2 in spring. The red horizontal line in the box plot represents the median over the Blended-5 datasets. Over the 25 year observational period, the trends correspond to an approximate snow loss that is at most 4% of JFM SCE and in the range of 5-

40% of AMJ SCE. The range of trends from the CanESM2 Large Ensemble, NCAR CESM1, and NCAR CCSM4 suggests that internal variability alone provides an uncertainty range of about 0.5 million km^2 per decade. Assuming internal variability is realistic in the models, this is the limit of precision we can expect in assessing recent trends. CanESM2 consistently produces greater snow loss than NCAR CCSM4 and CESM1, especially in AMJ. We conclude that all the models displayed fall within wintertime snow retreat estimates, that NCAR CCSM4 and CESM1 overlap with estimates of observed snow retreat in spring, but that CanESM2 exhibits more spring snow retreat than our best estimate of the observations. This excessive snow retreat is associated in part with excessive global warming in the model mentioned in section 2 (Mudryk et al., 2017).

A more challenging target for the purpose of simulation and attribution of climate change on a regional scale is the spatial pattern of observed climate fluctuations in recent decades. Acknowledging the overall biases noted above, we concentrate our analysis on the response patterns with spatial means removed. We first show the wintertime 1981-2005 land-surface temperature trend pattern with the spatial mean removed in the upper left panel of Fig. 5. This represents a predominantly positive meridional gradient of land-surface temperature change from southern to northern Canada, reflecting wintertime Arctic amplification of warming. The same field in individual realizations of CESM1 and CanESM2 has a spatial correlation in the range of -0.6 to +0.6 (Taylor diagram in lower left panel of Fig. 5), suggesting that these patterns are affected by significant internal variability (Deser et al., 2012). There are more realizations with positive than negative spatial correlation in winter land-surface temperature trend patterns, which is consistent with the anticipated effect of anthropogenic forcing. Wintertime land-surface temperature trends systematically show greater spatial variance than the estimated warming pattern from the single observational land-surface temperature dataset employed here. This could be related to stronger (more negative) meridional gradients in land-surface temperature and its trends in the models compared to the observational dataset. Spring time land-surface temperature trend patterns (top left panel of Fig. 6) feature anomalously negative changes in the Canadian prairie regions and positive changes around the coastal regions. It is harder to find realizations in the spring that correspond to the observed pattern (left Taylor diagram in bottom row of Fig. 6), and spatial variance of the land-surface temperature trends appear to be biased high as in winter. For precipitation, there is little evidence of consistent pattern matching between the observations and individual realizations, for winter or spring (middle column, top and bottom panels of Figs. 5-6). For SWE spatial patterns (right column, top and bottom panels of Figs. 5-6), the structural details of the trend maps are also not readily found in the models compared to the mean of the Blended-5 SWE. Compared to the ensemble mean of the Blended-5 observations, the simulations show greater spatial variance in SWE trends, but this is partially owing to smoothing of spatial structure in observational errors, as is shown by the scatter of SWE trends by individual contributors to the Blended-5 dataset (light brown circles in Taylor diagrams).

It is possible to find individual realizations in the CanESM2 large ensemble that either match fairly well or fairly poorly with observed trends. In the second row of Figs. 5 and 6, we show for the two seasons of JFM and AMJ respectively the land-surface temperature, precipitation, and SWE trends for the *best all-round match realization*, which is a single realization having

the greatest average pattern correlation across the three fields (land-surface temperature, precipitation, and SWE) and the two seasons of JFM and AMJ. The spatial pattern correlation coefficient of each field with its observational counterpart is labelled. Plotted in the third row of Figs. 5 and 6 is the *worst all-round match realization*, which is the single realization having the least (most negative) pattern correlation. The best-match realization exhibits tradeoffs across fields, for example in the ability to represent the structured pattern of springtime precipitation change (r=0.38 for the middle panel of the second row of Fig. 6) versus wintertime land-surface temperature change (r=0.01 for the left panel of the second row of Fig. 5). The worst-match case exhibits a similar range of correlations, on the negative side, and generally looks quite different from the best-match case. This preliminary analysis of intraensemble variability suggests limits on how much regional-scale information about changes for snow cover and related climate variables can be extracted from ESMs. The key point is that caution is needed in judging a model on its ability to reproduce spatial patterns of trends in SWE and related climate parameters, even on these multidecadal timescales.

The spatial pattern of CanESM2 land-surface temperature and precipitation trends is generally representative of that found in individual realizations of the CMIP5 datasets, in the sense that the individual realizations of CanESM2 and other CMIP5 models have positive pattern correlations with the CMIP5 multi-model mean (Taylor diagrams not shown). Consistently, the CMIP5 multimodel mean of the land-surface temperature and precipitation trends are generally similar to the CanESM2 ensemble mean (winter example shown in the top two rows of Fig. 7; note that in Fig. 7 the spatial mean of the patterns is not removed in order to allow comparison of the overall responses in CanESM2 to CMIP5). However, for SWE, we find CanESM2's pattern is typically opposite that of individual realizations from other models in CMIP5 (not shown) as is also evident in the ensemble mean (bottom row of Fig. 7). In particular, CanESM2 shows a strong positive trend in the Western Cordillera and a weaker positive trend in Southern Ontario and eastern Canada in both winter (Fig. 7) and spring (not shown), whereas a reduction of SWE is found in these regions and seasons in CMIP5.

3.2 Arctic sea ice in CanESM2

Turning to sea ice, we recall that it is well established that summertime Arctic sea-ice area or extent is biased low in CanESM2 (Stroeve et al., 2012; Merryfield et al. 2013a; Laliberté et al. 2016). We thus focus a limited amount of additional analysis on sea ice in the Canadian sector. The established low bias is borne out in the Canadian Arctic sector (top two panels of Fig. 8), where CanESM2 has less than half of the observed sea-ice coverage in the Beaufort Sea-Arctic Ocean sector. Further limiting the utility of regional sea-ice analysis with this model is the moderate spatial resolution of the model and its associated land-sea distribution, particularly in the Canadian Arctic Archipelago (central panel of Fig. 8). The summertime sea-ice extent is among the lowest of all CMIP5 models in the Canadian Arctic as a whole. In Canadian Arctic regions, summertime sea-ice extent is biased low in the Beaufort Sea and is practically zero in Hudson Bay and Baffin Bay (Fig. 9, left column). This bias contributes to the outcome that the sea ice reaches nominally ice-free summertime conditions at times comparable to present-day in CanESM2. The bias is evident throughout the seasonal cycle in most regions (Fig. 9, right column), with the exception

of Baffin Bay, although not as extreme relative to other models in other seasons as it is in summer. In this respect, the quality of simulation in CanESM2 is not as good as that of other ESMs such as NCAR CESM1 (lower panel of Fig. 8), which provide a better baseline for regional sea-ice studies both in terms of climatology and land-sea distribution.

5   Process investigations of sea ice by CanSISE include a focus on the relationship between sea-ice drift and Arctic winds, since realistic sea-ice dynamics are crucial for accurate representation of sea ice (Notz, 2012). International Arctic Buoy (IABP) Programme measurements (Tschudi et al., 2016) show that sea-ice drift speed peaks in September, when sea ice is thinnest (lower panel of Fig. 10), and not at the time of peak wind speed in December. However, in CanESM2, the peak sea-ice drift speed occurs in November and is more in phase with the seasonal cycle of near-surface wind speed. Other models in the CMIP5
10   archive that have more modern sea-ice components are able to reproduce more closely the observed seasonal cycle of sea-ice drift speed (Tandon et. al., submitted manuscript). These results provide strong motivation to transition to a modeling system with improved sea-ice and related processes in the Arctic.

4 Snow and sea-ice related forecast performance and development of CanSIPS

Operational seasonal forecasts based on coupled global ocean-atmosphere models have been produced for about two decades
15   internationally (Stockdale et al., 1998) and in Canada (by CanSIPS) since 2011. Over this period the main emphasis has been on predicting seasonal meteorological variables describing near-surface temperature, atmospheric circulation and precipitation, as well as sea-surface temperatures, since these are a major driver of seasonal climate variations. Potential has also existed for such systems to usefully predict additional variables, including snow and sea ice, particularly as the sophistication of the models and the methods used to initialize them have increased. With respect to the cryosphere, however, such capabilities have
20   received little attention until relatively recently compared to other areas of focus in seasonal prediction (e.g. Blanchard-Wrigglesworth et al. 2011, Sigmond et al. 2013, Guémas et al. 2016).

4.1 Characteristics of CanSIPS related to seasonal forecasts of terrestrial snow
Research carried out under CanSISE examined the ability of CanSIPS both to realistically initialize SWE and to predict future
25   SWE variations (Sospedra-Alfonso et al., 2016a, b, c). This was the first study of snow in an operational seasonal forecast system. Regarding seasonal prediction of snow by CanSIPS, anomaly correlation skill for wintertime SWE is high at short lead times, and remains statistically significant (greater than 0.3) at lead times of at least 6 months for certain regions (Fig. 11), which suggests potential for practical utilization of such forecasts. Two primary sources of potential predictability (PP, defined as the ratio of 'signal' variance describing interannual variability of ensemble means to total variance consisting of the
30   sum of 'signal' and 'noise' components) and skill in CanSIPS forecasts of SWE have been identified (Sospedra-Alfonso et al., 2016b, c). The first, which is most important at short lead times, is the demonstrated ability of CanSIPS to provide realistic initial values for SWE (Sospedra-Alfonso et al., 2016a), combined with the natural tendency for SWE anomalies to persist

throughout the snow season in regions where winter land-surface temperatures remain below freezing, so that the snowpack accumulates until the onset of spring melt. The second main source of PP and skill, which becomes increasingly prevalent at longer lead times, is the ability of CanSIPS to predict future climatic conditions such as land-surface temperature and precipitation anomalies which influence snow accumulation and melt. A large part of this type of predictability and skill arises
5   from ENSO, which strongly influences winter climate in North America and is skillfully forecast by CanSIPS up to a year in advance.

The value of skillful seasonal forecasting of snow in turn depends on process representation and initialization at the land-surface. For example, Ambadan et al. (2015) have investigated the impact of initialization of SWE, soil liquid water, and soil
10   frozen water on potential predictability of springtime surface air temperature in the CanSIPS system (Fig. 12). Realistic initialization of these variables enhances potential predictability by as much as 30% in terms of variance explained within the potential predictability framework. This shows that it is important to regard snow initialization in the broader setting of land-surface initialization, and that there is evidence for quantitative improvement in regional predictability as more observational information on the state of the land surface is brought into the prediction system. Current operational practice in CanSIPS uses
15   observed atmospheric forcing to bring the land surface (including soil moisture and snow cover) into a realistic state. Although this procedure performs reasonably well for snow (within observational uncertainty), potential remains for improving the initialization and forecasting of snow and other land variables by assimilating observation-based land data directly in real time.

Blending different sources of data from highly uncertain observations has led to improved characterization of the forecast skill
20   of the CanSIPS system. Fig. 13 shows the degree of agreement between SWE forecasts from CanSIPS and several SWE products over Canada (similar results are found for other regions). The degree of agreement is measured as the anomaly correlation coefficient for a one-month forecast (with lead 0 from initialization). The five datasets are the Blended-5 dataset (blue) and four individual datasets including two components of the Blended-5 dataset. Even though all observational datasets are being compared to the same forecast, it is the Blended-5 dataset, capturing the mean of several observational datasets, that
25   agrees best with the forecast. It is clear that improving verification datasets through blending, which can be reasonably expected to lead to cancellation of independent errors in observational estimates, impacts assessed agreement with the forecast. To reiterate, in this case, improved calculated skill is derived from an apparent improvement in the quality of the verification data and not an improvement of the forecast (Sospedra-Alfonso et al., 2016b). Whether or not such improved consistency might be found for the prediction of other quantities, the broader point is that there is a need to ensure that verification data is continually
30   updated in order to fairly compare predictions to the best available data (Massonet et al., 2016).

Recent research in snow analysis and observational datasets is expected to support operational improvements in CanSIPS and hence in ECCC's operational prediction capacity. For example, CanSISE work has led to new efforts to develop an operational real-time snow amount forecast for the coming months, which could be used in several impacted sectors such as outdoor

recreation, water resource planning, and agriculture (Fig. 14, snow amount forecast shown as above and below normal SWE amounts). In this successful proof of concept, we note satisfactory general agreement with the MERRA analysis, which is independent of CanSIPS and is itself subject to some uncertainty. This indicates promise for this new forecast product, while highlighting issues of observational uncertainty addressed in part by our recent research.

4.2 Sea-ice forecasting with CanSIPS

Much as for snow, the ability of global climate model-based seasonal forecasting systems to predict sea ice has also received little attention until recently, although such assessments have now been carried out for several systems (Guémas et al. 2016). In the area of sea-ice prediction, CanSIPS hindcasts, despite some of the simulation deficiencies described above, have

10   demonstrated skill in seasonal predictions of sea ice (e.g. Sigmond et al., 2013; Merryfield et al., 2013b). While these prior studies have focused on forecast skill of area-integrated quantities such as sea-ice area, recent work (Sigmond et al., 2016) has also shown significant forecast skill of more user relevant sea-ice metrics such as the first calendar date that sea ice melts (retreat date; Figs. 15a-c) or freezes up (advance date; Figs. 15d-f). Advance dates are skillfully predicted at lead times of 5 months on average (3.3 months for detrended anomalies), and retreat dates at lead times of 3 months (2.2 months for detrended

15   anomalies). For retreat dates, the main source of forecast skill is persistence, while advance date predictions benefit from predictable ocean temperatures.

Sea ice predictability is also assisted by persistence of sea-ice thickness (e.g. Chevallier and Salas Y Mélia, 2014), but CanSIPS does not take advantage of this in that it currently employs an initialization method that uses only climatological sea-ice

20   thickness (SIT) information. Since real time SIT observations are limited, Dirkson et al. (2015, 2017) have developed several statistical models of varying complexity for initializing SIT in operational predictions. These are based on predictors available in real time together with historical SIT values represented by the pan-Arctic Ice and Ocean Modelling System, or PIOMAS (Zhang and Rothrock, 2003), which is frequently used as a reference dataset for SIT due to the sparseness of historical SIT observations. The first such model (known as "SMv1"), described in Dirkson et al. (2015), uses a statistical approach to find

25   an optimal combination of sea ice concentration and sea level pressure information to provide useful sea ice thickness information. While this model reduces temporal- and spatial- mean absolute errors in the SIT initial conditions by 48% relative to the original CanSIPS initialization (when validated against PIOMAS SIT values), and shows consistent skill estimating ice volume in all months, much of this improvement in skill emerges from a more accurate representation of local negative trends in SIT. Two additional statistical models, "SMv2" and "SMv3", that improve on SMv1 with respect to interannual variations

30   in SIT anomalies are described in Dirkson et al. (2017), and seasonal sea ice volume from SMv3 is compared to that from CanSIPS initial conditions in Fig. 16. Seasonal forecasting experiments using these SIT initial conditions demonstrate general improvement forecasting both pan-Arctic sea-ice extent and local sea-ice concentration compared to the current operational system, with most significant improvements afforded by initializing with either SMv2 or SMv3 (Dirkson et al., 2017).

5 Conclusion

We have assessed characteristics of snow, sea ice, and related climate parameters in Environment and Climate Change Canada's (ECCC) earth system model (ESM) CanESM2 and seasonal to interannual climate prediction system CanSIPS, with a focus on the Canadian sector of the Northern Hemisphere. This assessment is intended to provide a baseline for future versions of the models with respect to these important societally-relevant climate parameters. It has highlighted the application of the Blended-5 multisource snow water equivalent (SWE) (Mudryk et al., 2015) and the CanESM2 Large Ensemble of climate simulations. In addition, it has highlighted new developments in sea ice, snow, and related climate parameter prediction on seasonal timescales. We summarize our key findings:

- The CanESM2 simulation of climate parameters over the Canadian land mass closely tied to snow – land surface temperature and precipitation on land in cold regions – lies well within the range of currently available international models. There is considerable disagreement among observational datasets on the amount and geographical structure of snow water equivalent (SWE) in the satellite era. The CanESM2 simulation of SWE performs as well as available international models in this area. Even accounting for this observational uncertainty, however, there is a bias towards excessive seasonal snow cover and unrealistic spatial distribution of SWE in the spring over the Canadian land mass and over the Northern Hemisphere as a whole. Excessive precipitation over the Canadian land mass contributes to this bias.

- Accounting for observational uncertainty, CanESM2 simulates a greater retreat of springtime snow over the satellite era than most of the available observations assessed here and other models that include large initial condition ensembles. The spatial pattern of the observed temperature, precipitation, and SWE trends is strongly influenced by internal variability. This makes it difficult to assess the model-simulated patterns of change in the variables we have examined. Nevertheless, Western Cordillera trends in SWE in CanESM2 represent a recent increase that is opposite to those found in typical CMIP5 models.

- Previously identified biases towards low Arctic sea ice extent are also reflected in regional biases: in Hudson's Bay and the Canadian Beaufort Sea sector, the sea ice extent is biased low and this undermines projections of when regional sectors of the Arctic will be ice free. In the current system, there are tradeoffs related to the resolution of geographical features in the CanESM and CanSIPS systems that impact both the snow and sea ice simulations. This provides an urgent area of improvement for future model development.

- Recent work suggests promising potential for seasonal forecasting of snow, sea ice and related climate parameters using CanSIPS. For example, accurate initialization of frozen and liquid soil water, in addition to improved SWE representation, might lead to significantly improved seasonal temperature forecasts. Furthermore, the Blended-5 example shows that accounting for observational uncertainty can lead to better understanding of forecast quality. This result suggests initialization could also be improved in this manner. This and related work has stimulated the development of ECCC's first experimental seasonal snow amount forecast product.

- Despite biases in the sea ice simulation, it is possible to develop potentially useful new seasonal forecast products for sea ice advance and retreat. In addition, implementing sea ice thickness initialization using indirect statistical predictors of thickness can improve sea ice forecasts compared to the current methodology. Motivated by the promising research results, improved sea ice thickness initialization (as initially explored by Lindsay et al., 2012 and Day et al. 2014) is being considered for implementation in the CanSIPS system.

Further improvements in the CanSIPS and CanESM climate prediction and projection capacity for snow, sea ice and related climate variables also hinge on assessing model process representation in more depth. For example, critical to capture accurately is the snow albedo feedback process, which governs the seasonality of snow cover and land surface temperature and hydroclimatic responses to climate change (Qu and Hall, 2007, 2014; Hall et al., 2008; Thackeray et al., 2015; Thackeray and Fletcher, 2016). Thackeray et al. (2015) show that CanESM2 places among the best CMIP5 models for all regions in terms of the overall simulation of snow cover fraction and snow-covered surface albedo. Further progress in this kind of process representation will be achieved in part through internationally coordinated inter-comparison efforts associated with CMIP6, including the Land Surface, Snow and Soil moisture Model Intercomparison Program (LS3MIP; van den Hurk et al 2016) and the Earth System Model Snow Inter-comparison Project (ESM-SnowMIP). Besides ongoing work on sea-ice mechanics and its relationship, through wind driving, to sea ice drift, CanSISE research is also currently characterizing snow cover on sea ice in models and observations, which also serves as a potential source of model error in the timing and amplitude of sea-ice growth and melt.

CanSISE demonstrates the utility of entraining a network of researchers bridging observational and modelling communities to focus on a related set of processes in evaluation of earth system models and climate prediction systems. The results suggest that there can be several benefits to updated multi-source observational datasets for climate prediction, monitoring, and assessment. Our focus in this paper has been on recently produced multi-source snow observational datasets, but our results suggest that there are benefits of multi-source temperature, precipitation, and sea ice datasets that follow a similar approach. We have articulated the tradeoffs involved in constraints on CanESM2's resolution in light of limitations of available advanced computing resources. Running the model at two degrees latitude/longitude permits the creation of the CanESM2 Large Ensemble set, but can entail under-resolution of key features of interest in applications, such as the Canadian Arctic Archipelago's channels and islands. We suggest that similar large ensembles be considered based on future model versions, accounting for these tradeoffs, and being complemented by ECCC regional climate model simulations (e.g. Scinocca et al., 2016).

References

Ambadan, J. T., Berg, A. and Merryfield, W. J.: Influence of snow and soil moisture initialization on sub-seasonal predictability and forecast skill in boreal spring, Climate Dynamics, 1–17, doi:10.1007/s00382-015-2821-9, 2015.

Arora, V. K., Scinocca, J. F., Boer, G. J., Christian, J. R., Denman, K. L., Flato, G. M., Kharin, V. V., Lee, W. G. and

5 Merryfield, W. J.: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, Geophys. Res. Lett., 38(5), doi:10.1029/2010GL046270, 2011.

Blanchard-Wrigglesworth, E., Armour, K.C., Bitz, C.M., and DeWeaver,E.: Persistence and Inherent Predictability of Arctic Sea Ice in a GCM Ensemble and Observations. J. Climate, 24, 231–250, doi:10.1175/2010JCLI3775.1, 2011.

Blazey, B. A., Holland, M. M. and Hunke, E. C.: Arctic Ocean sea ice snow depth evaluation and bias sensitivity in CCSM,

10 The Cryosphere, 7(6), 1887–1900, doi:10.5194/tc-7-1887-2013, 2013.

Brown, R. and Derksen, C.: Is Eurasian October snow cover extent increasing?, Environmental Research Letters, 8(2), 024006, 2013.

Brown, R., Derksen, C. and Wang, L.: A multi-data set analysis of variability and change in Arctic spring snow cover extent, 1967–2008, J. Geophys. Res., 115(D16), doi:10.1029/2010JD013975, 2010.

15 Brutel-Vuilmet, C., Ménégoz, M. and Krinner, G.: An analysis of present and future seasonal Northern Hemisphere land snow cover simulated by CMIP5 coupled climate models, The Cryosphere, 7(1), 67–80, doi:10.5194/tc-7-67-2013, 2013.

Chevallier, M., and Salas Y Mélia, D.: The role of sea ice thickness distribution in the Arctic sea ice potential predictability: a diagnostic approach with a coupled GCM. J. Climate, 25, 3025-3038, doi:10.1175/JCLI-D-11-00209.1, 2014.

Day, J. J., Hawkins, E., and Tietsche, S.: Will Arctic sea ice thickness initialization improve seasonal forecast skill?, Geophys.

20 Res. Lett., 41, 7566–7575, doi:10.1002/2014GL061694, 2014.

Derksen, C. and Brown, R.: Spring snow cover extent reductions in the 2008–2012 period exceeding climate model projections, Geophys. Res. Lett., 39(19), doi:10.1029/2012GL053387, 2012.

Derksen, C., Smith, S. L., Sharp, M., Brown, L., Howell, S., Copland, L., Mueller, D. R., Gauthier, Y., Fletcher, C. G., Tivy, A., Bernier, M., Bourgeois, J., Brown, R., Burn, C. R., Duguay, C., Kushner, P., Langlois, A., Lewkowicz, A. G., Royer, A.

25 and Walker, A.: Variability and change in the Canadian cryosphere, Climatic Change, 115(1), 59–88, doi:10.1007/s10584-012-0470-0, 2012.

Déry, S., Knudsvig, H., Hernández-Henríquez, M. and Coxson, D.: Net Snowpack Accumulation and Ablation Characteristics in the Inland Temperate Rainforest of the Upper Fraser River Basin, Canada, Hydrology, 1(1), 1–19, doi:10.3390/hydrology1010001, 2014.

30 Deser, C., Knutti, R., Solomon, S. and Phillips, A. S.: Communication of the role of natural variability in future North American climate, Nature Clim. Change, 2(11), 775–779, doi:10.1038/nclimate1562, 2012.

Dirkson, A., Merryfield, W. J. and Monahan, A.: Real-time estimation of Arctic sea ice thickness through maximum covariance analysis: Statistical sea ice thickness, Geophysical Research Letters, 42(12), 4869–4877, doi:10.1002/2015GL063930, 2015.

Dirkson, A., Merryfield, W. J. and Monahan, A.: Impacts of Sea Ice Thickness Initialization on Seasonal Arctic Sea Ice Predictions, J. Climate, 30(3), 1001–1017, doi:10.1175/JCLI-D-16-0437.1, 2016.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, S., Emori, S., Eyring, V., Forest, C., Gleckler, E., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C. and Rummukainen, M.: Evaluation of Climate Models, in: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Stocker et al., Cambridge University Press, 741-866, 2013.

Fletcher, C. G., Kushner, P. J., Hall, A. and Qu, X.: Circulation responses to snow albedo feedback in climate change, Geophys. Res. Lett., 36(9), doi:10.1029/2009GL038011, 2009.

Fyfe, J. C. and Flato, G. M.: Enhanced Climate Change and Its Detection over the Rocky Mountains, J. Climate, 12(1), 230–243, doi:10.1175/1520-0442(1999)012<0230:ECCAID>2.0.CO;2, 1999.

Fyfe, J. C., Derksen, C., Mudryk, L., Flato, G. M., Santer, B. D., Swart, N. C., Molotch, N. P., Zhang, X., Wan, H., Arora, V. K., Scinocca, J. and Jiao, Y.: Large near-term projected snowpack loss over the western United States, Nature Communications, 8, 14996, doi:10.1038/ncomms14996, 2017.

Gagné, M.-È., Fyfe, J. C., Gillett, N. P., Polyakov, I. V. and Flato, G. M.: Aerosol-driven increase in Arctic sea ice over the middle of the 20th Century, Geophys. Res. Lett., doi:10.1002/2016GL071941, 2017a.

Gagné, M.-È., Kirchmeier-Young, M. C., Gillett, N. P. and Fyfe, J. C.: Arctic sea ice response to the eruptions of Agung, El Chichón and Pinatubo, J. Geophys. Res. Atmos., doi:10.1002/2017JD027038, 2017b.

Gillett, N. P., Arora, V. K., Flato, G. M., Scinocca, J. F. and von Salzen, K.: Improved constraints on 21st-century warming derived using 160 years of temperature observations, Geophys. Res. Lett., 39(1), doi:10.1029/2011GL050226, 2012.

Hall, A., Qu, X. and Neelin, J. D.: Improving predictions of summer climate change in the United States, Geophys. Res. Lett., 35(1), doi:10.1029/2007GL032012, 2008.

Guémas, V., Blanchard-Wrigglesworth, E., Chevallier, M., Day, J., Déqué, M., Doblas-Reyes, F., Fuckar, N., Germe, A., Hawkins, E., Keeley, S., Koenigk, T., Salas y Mélia, D., Tietsche, S.: A review on Arctic sea ice predictability and prediction on seasonalto-decadal timescales. Q. J. Royal Meteorol. Soc.,142,546–561. doi:10.1002/qj.2401, 2016.

Hernández-Henríquez, M. A., Déry, S. J. and Derksen, C.: Polar amplification and elevation-dependence in trends of Northern Hemisphere snow cover extent, 1971–2014, Environmental Research Letters, 10(4), 044010, doi:10.1088/1748-9326/10/4/044010, 2015.

Hori, M., Sugiura, K., Kobayashi, K., Aoki, T., Tanikawa, T., Kuchiki, K., Niwano, M. and Enomoto, H.: A 38-year (1978–2015) Northern Hemisphere daily snow cover extent product derived using consistent objective criteria from satellite-borne optical sensors, Remote Sensing of Environment, 191, 402–418, doi:10.1016/j.rse.2017.01.023, 2017.

Howell, S. E. L., Laliberté, F., Kwok, R., Derksen, C. and King, J.: Landfast ice thickness in the Canadian Arctic Archipelago from Observations and Models, The Cryosphere Discuss., 2016, 1–39, doi:10.5194/tc-2016-71, 2016.

van den Hurk, B., Kim, H., Krinner, G., Seneviratne, S. I., Derksen, C., Oki, T., Douville, H., Colin, J., Ducharne, A., Cheruy, F., Viovy, N., Puma, M. J., Wada, Y., Li, W., Jia, B., Alessandri, A., Lawrence, D. M., Weedon, G. P., Ellis, R., Hagemann,

S., Mao, J., Flanner, M. G., Zampieri, M., Materia, S., Law, R. M. and Sheffield, J.: LS3MIP (v1.0) contribution to CMIP6: the Land Surface, Snow and Soil moisture Model Intercomparison Project – aims, setup and expected outcome, Geosci. Model Dev., 9(8), 2809–2832, doi:10.5194/gmd-9-2809-2016, 2016.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S. C., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J.-F., Lawrence, D., Lindsay, K., Middleton, A., Munoz, E., Neale, R., Oleson, K., Polvani, L. and Vertenstein, M.: The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability, Bull. Amer. Meteor. Soc., 96(8), 1333–1349, doi:10.1175/BAMS-D-13-00255.1, 2014.

Kharin, V. V., Zwiers, F. W., Teng, Q., Boer, G. J., Derome, J. and Fontecilla, J.: Skill assessment of seasonal hindcasts from the Canadian historical forecast project, Atmosphere-Ocean, 47(3), 204–223, doi:10.3137/AO1101.2009, 2009.

King, J., Howell, S., Derksen, C., Rutter, N., Toose, P., Beckers, J. F., Haas, C., Kurtz, N. and Richter-Menge, J.: Evaluation of Operation IceBridge quick-look snow depth estimates on sea ice, Geophys. Res. Lett., 42(21), 9302–9310, doi:10.1002/2015GL066389, 2015.

Kirchmeirer-Young, M., Zwiers, F. and Gillett, N. P.: Attribution of Extreme Events in Arctic Sea Ice Extent, J. Climate, 30(2), 553–571, doi:10.1175/JCLI-D-16-0412.1, 2016.

Kirtman, B., Power, B. and coauthors: Near-term Climate Change: Projections and Predictability, in: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Stocker et al., Cambridge University Press, 953–1028, 2013.

Kulkarni, T., Watkins, J. M., Nickels, S. and Lemmen, D. S.: Canadian International Polar Year (2007–2008): an introduction, Climatic Change, 115(1), 1–11, doi:10.1007/s10584-012-0583-5, 2012.

Laliberté, F., Howell, S. E. L. and Kushner, P. J.: Regional variability of a projected sea ice-free Arctic during the summer months, Geophys. Res. Lett., 43(1), 2015GL066855, doi:10.1002/2015GL066855, 2016.

Lindsay, R., Haas, C., Hendricks, S., Hunkeler, P., Kurtz, N., Paden, J., Panzer, B., Sonntag, J., Yungel, J., and Zhang, J.: Seasonal forecasts of Arctic sea ice initialized with observations of ice thickness, Geophys. Res. Lett., 39, L21502, 2012.

Massonnet, F., Bellprat, O., Guemas, V. and Doblas-Reyes, F. J.: Using climate models to estimate the quality of global observational data sets, Science, 354(6311), 452, doi:10.1126/science.aaf6369, 2016.

McCusker, K. E., Fyfe, J. C. and Sigmond, M.: Twenty-five winters of unexpected Eurasian cooling unlikely due to Arctic sea-ice loss, Nature Geosci, 9(11), 838–842, 2016.

Merryfield, W. J., Lee, W.-S., Boer, G. J., Kharin, V. V., Scinocca, J. F., Flato, G. M., Ajayamohan, R. S., Fyfe, J. C., Tang, Y. and Polavarapu, S.: The Canadian Seasonal to Interannual Prediction System. Part I: Models and Initialization, Mon. Wea. Rev., 141(8), 2910–2945, doi:10.1175/MWR-D-12-00216.1, 2013a.

Merryfield, W. J., Lee, W.-S., Wang, W., Chen, M. and Kumar, A.: Multi-system seasonal predictions of Arctic sea ice, Geophys. Res. Lett., 40(8), 1551–1556, doi:10.1002/grl.50317, 2013b.

Mudryk, L. R., Kushner, P. J. and Derksen, C.: Interpreting observed northern hemisphere snow trends with large ensembles of climate simulations, Climate Dynamics, 43(1), 345–359, doi:10.1007/s00382-013-1954-y, 2013.

Mudryk, L. R., Derksen, C., Kushner, P. J. and Brown, R.: Characterization of Northern Hemisphere Snow Water Equivalent Datasets, 1981–2010, J. Climate, 28(20), 8037–8051, doi:10.1175/JCLI-D-15-0229.1, 2015.

5   Mudryk, L. R., Kushner, P. J., Derksen, C. and Thackeray, C.: Snow cover response to temperature in observational and climate model ensembles, Geophys. Res. Lett., 44(2), 919–926, doi:10.1002/2016GL071789, 2017.

Najafi, M. R., Zwiers, F. and Gillett, N. P.: Attribution of the spring snow cover extent decline in the Northern Hemisphere, Eurasia and North America to anthropogenic influence, Climatic Change, 1–16, doi:10.1007/s10584-016-1632-2, 2016.

Notz, D.: Challenges in simulating sea ice in Earth System Models, WIREs Clim Change, 3(6), 509–526, doi:10.1002/wcc.189,

10 2012.

Qu, X. and Hall, A.: What Controls the Strength of Snow-Albedo Feedback?, J. Climate, 20(15), 3971–3981, doi:10.1175/JCLI4186.1, 2007.

Qu, X. and Hall, A.: On the persistent spread in snow-albedo feedback, Climate Dynamics, 42(1), 69–81, doi:10.1007/s00382-013-1774-0, 2013.

15 Scinocca, J. F., Kharin, V. V., Jiao, Y., Qian, M. W., Lazare, M., Solheim, L., Flato, G. M., Biner, S., Desgagne, M. and Dugas, B.: Coordinated Global and Regional Climate Modeling, J. Climate, 29(1), 17–35, doi:10.1175/JCLI-D-15-0161.1, 2015.

Sigmond, M. and Fyfe, J. C.: Tropical Pacific impacts on cooling North American winters, Nature Clim. Change, 6(10), 970–974, doi:10.1038/nclimate3069, 2016.

Sigmond, M., Fyfe, J. C., Flato, G. M., Kharin, V. V. and Merryfield, W. J.: Seasonal forecast skill of Arctic sea ice area in a

20 dynamical forecast system, Geophys. Res. Lett., 40(3), 529–534, doi:10.1002/grl.50129, 2013.

Sigmond, M., Reader, M. C., Flato, G. M., Merryfield, W. J. and Tivy, A.: Skillful seasonal forecasts of Arctic sea ice retreat and advance dates in a dynamical forecast system, Geophys. Res. Lett., 43(24), 12,457-12,465, doi:10.1002/2016GL071396, 2016.

Sospedra-Alfonso, R. and Merryfield, W. J.: Potential and actual predictability of snow water equivalent in historical forecasts

25 of the Canadian Fourth Generation Coupled Climate Model (CanCM4)., in Climate Prediction S&T Digest., 2016a.

Sospedra-Alfonso, R., Mudryk, L., Merryfield, W. and Derksen, C.: Representation of snow in the Canadian Seasonal to Interannual Prediction System: Part I. Initialization, J. Hydrometeor, doi:10.1175/JHM-D-14-0223.1, 2016b.

Sospedra-Alfonso, R., Merryfield, W. J. and Kharin, V. V.: Representation of Snow in the Canadian Seasonal to Interannual Prediction System. Part II: Potential Predictability and Hindcast Skill, J. Hydrometeor., 17(9), 2511–2535, doi:10.1175/JHM-

30 D-16-0027.1, 2016c.

Stockdale, T. N., Anderson, D. L. T., Alves, J. O. S. and Balmaseda, M. A.: Global seasonal rainfall forecasts using a coupled ocean-atmosphere model, Nature, 392(6674), 370–373, doi:10.1038/32861, 1998.

Stroeve, J. C., Kattsov, V., Barrett, A., Serreze, M., Pavlova, T., Holland, M. and Meier, W. N.: Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations, Geophys. Res. Lett., 39(16), doi:10.1029/2012GL052676, 2012.

Swart, N. C., Fyfe, J. C., Hawkins, E., Kay, J. E. and Jahn, A.: Influence of internal variability on Arctic sea-ice trends, Nature Clim. Change, 5(2), 86–89, doi:10.1038/nclimate2483, 2015.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res., 106(D7), 7183–7192, doi:10.1029/2000JD900719, 2001.

5   Taylor, K. E., Stouffer, R. J. and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, Bull. Amer. Meteor. Soc., 93(4), 485–498, doi:10.1175/BAMS-D-11-00094.1, 2011.

Thackeray, C. W. and Fletcher, C. G.: Snow albedo feedback: Current knowledge, importance, outstanding issues and future directions, Progress in Physical Geography, 1–17, doi:10.1177/0309133315620999, 2016.

Thackeray, C. W., Fletcher, C. G. and Derksen, C.: Quantifying the skill of CMIP5 models in simulating seasonal albedo and

10   snow cover evolution: CMIP5-simulated albedo and SCF skill, Journal of Geophysical Research: Atmospheres, 120(12), 5831–5849, doi:10.1002/2015JD023325, 2015.

Tschudi, M., Fowler, C., Maslanik, J., Stewart, J. and Asher, W.: Polar Pathfinder Daily 25 km EASE-Grid Sea Ice Motion Vectors, Version 3, NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado USA., doi:10.5067/O57VAIT2AYYY, 2016.

15   van den Hurk, B., Kim, H., Krinner, G., Seneviratne, S. I., Derksen, C., Oki, T., Douville, H., Colin, J., Ducharne, A., Cheruy, F., Viovy, N., Puma, M. J., Wada, Y., Li, W., Jia, B., Alessandri, A., Lawrence, D. M., Weedon, G. P., Ellis, R., Hagemann, S., Mao, J., Flanner, M. G., Zampieri, M., Materia, S., Law, R. M., and Sheffield, J.: LS3MIP (v1.0) contribution to CMIP6: the Land Surface, Snow and Soil moisture Model Intercomparison Project – aims, setup and expected outcome, Geosci. Model Dev., 9, 2809-2832, https://doi.org/10.5194/gmd-9-2809-2016, 2016.

20   Warren, S. G., Rigor, I. G., Untersteiner, N., Radionov, V. F., Bryazgin, N. N., Aleksandrov, Y. I. and Colony, R.: Snow Depth on Arctic Sea Ice, J. Climate, 12(6), 1814–1829, doi:10.1175/1520-0442(1999)012<1814:SDOASI>2.0.CO;2, 1999.

Zhang, J. and Rothrock, D. A.: Modeling Global Sea Ice with a Thickness and Enthalpy Distribution Model in Generalized Curvilinear Coordinates, Mon. Wea. Rev., 131(5), 845–861, doi:10.1175/1520-0493(2003)131<0845:MGSIWA>2.0.CO;2, 2003.
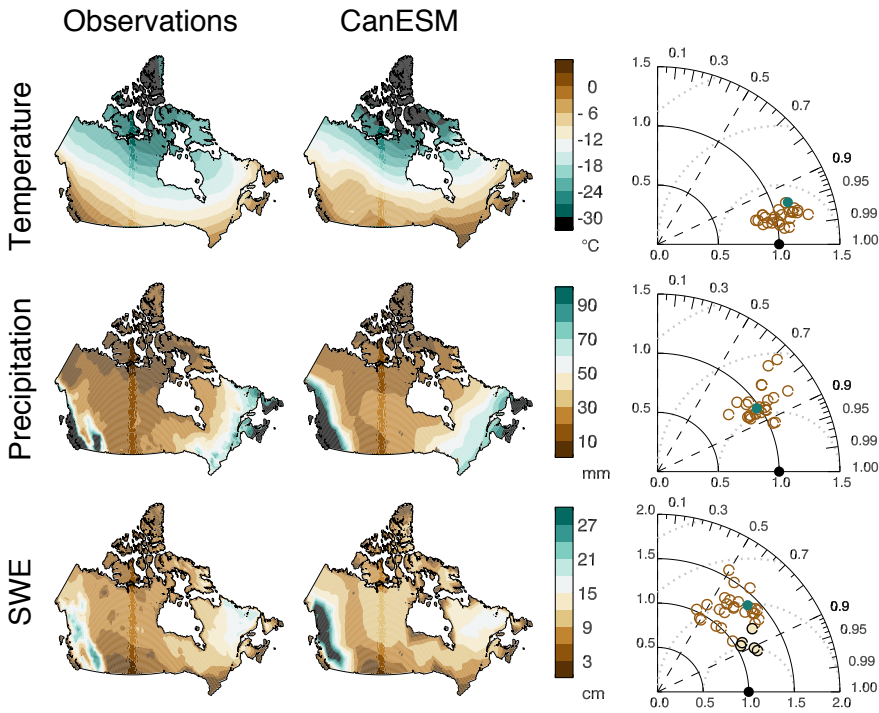
Figure 1: Comparison of simulated Canadian climate in CanESM2 with observations and other climate models for 1981-2005. The left column shows observed January-March (JFM) mean land-surface temperature (top, HadCRUT4), precipitation (centre, CRU TS3.21) and snow water equivalent (bottom, Blended 5 data). The central column shows the same fields as simulated by the ensemble mean of the CanESM2 Large Ensemble. Observations and the CanESM2 output have been mapped to a common grid that represents the model grid spacing. The right column plots Taylor (2001) diagrams showing the correlation (related to the polar angle) and standard deviation relative to observations (distance from origin) of the patterns of these variables in observations (black dot), CanESM2 (green dot) and other CMIP5 climate models (brown circles). Models that are closer to the black dot representing the observations have smaller errors (standard error represented by dotted semicircles at intervals of 0.5). In the SWE Taylor diagram (bottom row), the filled light brown dots compare the individual Blended-5 datasets to the Blended-5 mean.
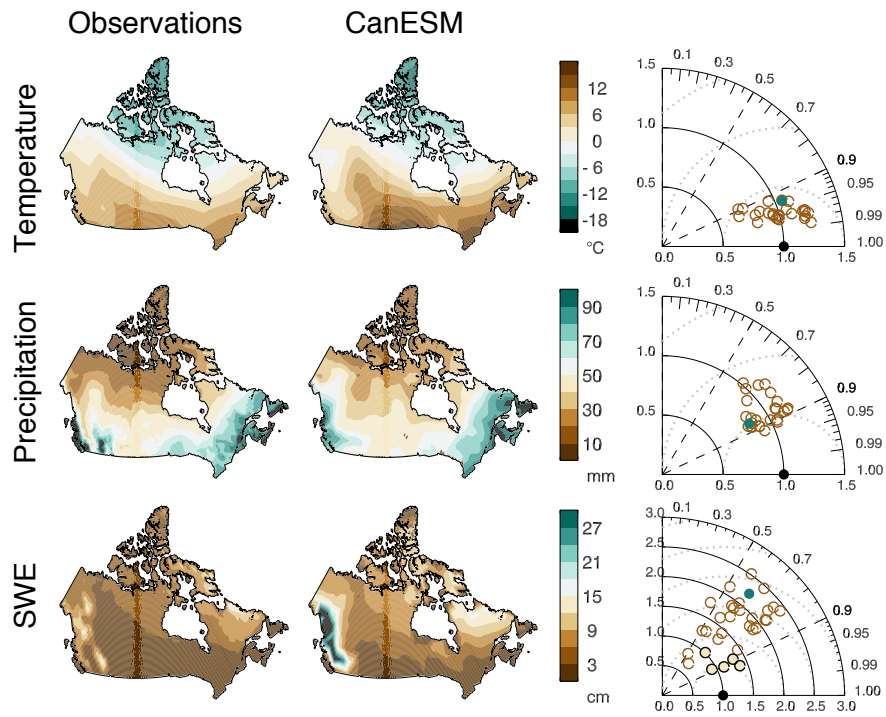
Figure 2: As in Fig. 1, for April-June (AMJ).
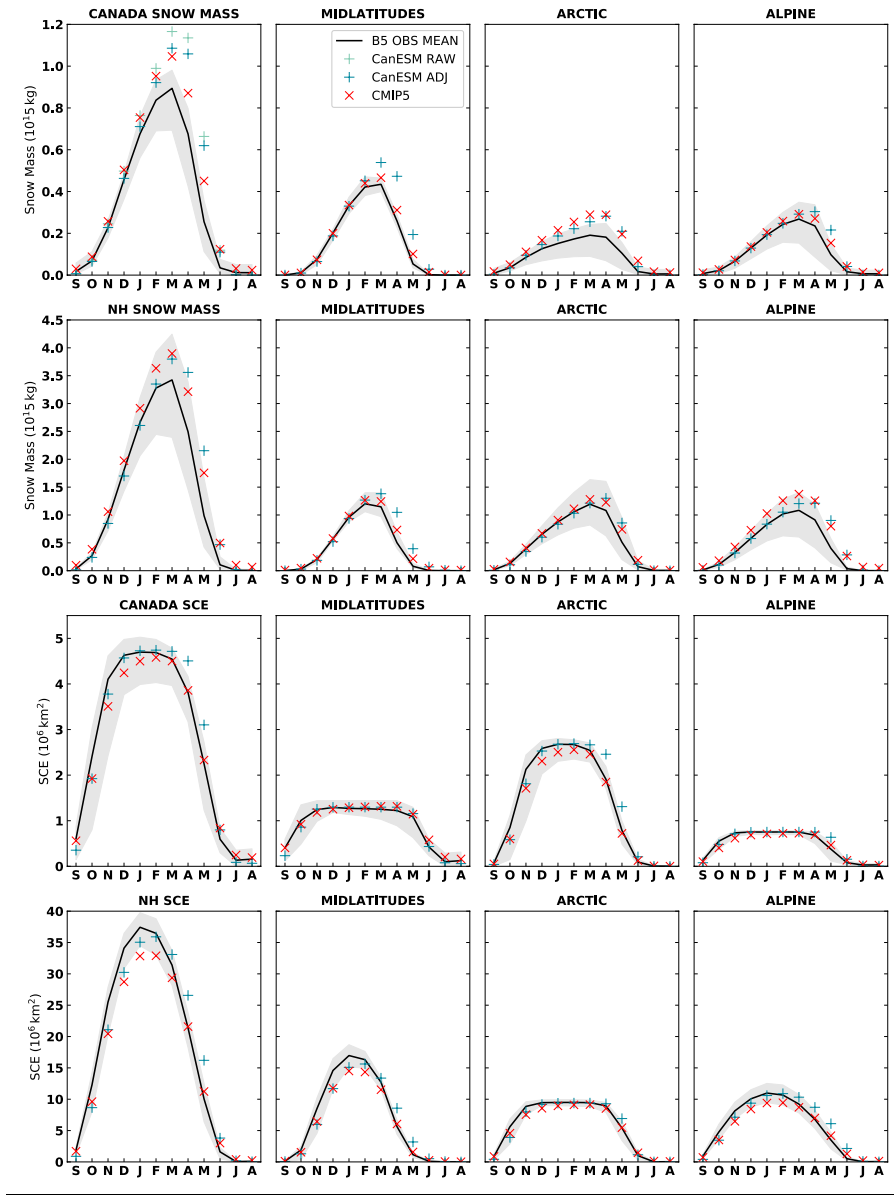
Figure 3: First row: seasonal cycle of Northern Hemisphere 1981-2005 snow mass (in 10^{15} kg) for regions defined in Mudryk et al. 2015: midlatitudes – Northern Hemisphere nonalpine land regions south of 60^0N, Arctic - nonalpine land regions north of 60^0N, and alpine. Gray shading represents the range of Blended-5 datasets, the black curve represents the Blended-5 mean, the light teal points (in the first column of the first row only) mark the ensemble mean of CanESM2 using its land mask, and the dark teal points mark CanESM2 adjusted to represent the same land fractions as the observational mask from the Blended-5 dataset. The CMIP multi-model mean, adjusted to the observational mask, is show with red x symbols. The legend in the second column top row applies to the figure as a whole. Second row: as in top row, but for Canadian land mass only. Third and fourth rows are similar to the first and second rows, but for snow cover extent in 10^6 km^2. The estimate of observed snow cover extent is derived from the Blended-5 SWE dataset using the approach of Mudryk et al. 2017, and is based on a 4 mm SWE threshold for the presence of snow cover; the simulated snow cover extent is based on snow cover fraction directly produced by the models.
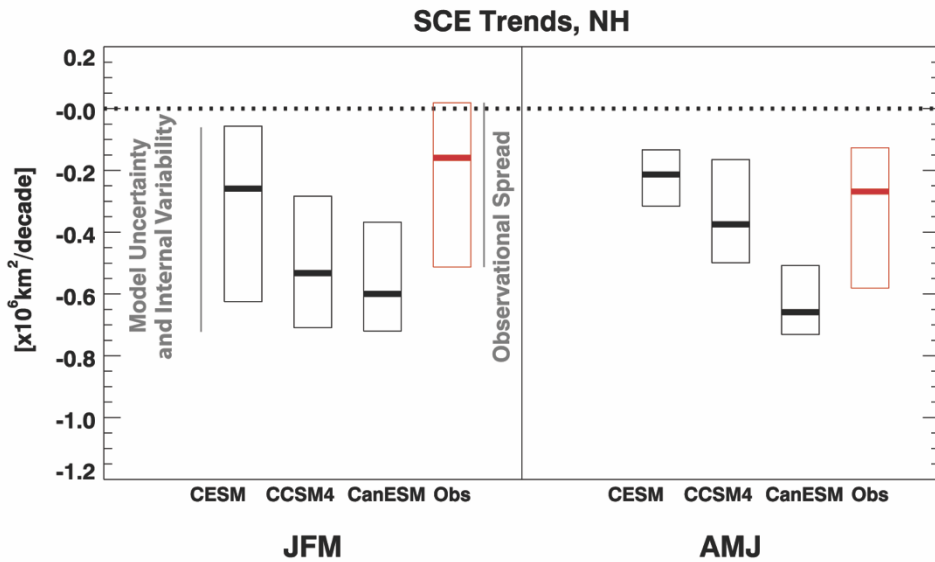
Figure 4: Horizontal bars and boxes show the median and interquartile range of Northern Hemisphere snow cover extent trends calculated over 1981-2005 period for January-March (left) and April-June (right). Black boxes are used for trends calculated from models (large initial condition ensembles of NCAR CCSM4, NCAR CESM1 and CanESM2 as labelled), and red boxes are used for trends derived from the Blended-5 SWE dataset (Mudryk et al. 2017). For the models, the boxes indicate the interquartile range (IQR) of trends captured in individual realizations of the ensembles. For the observations, the boxes indicate the IQR of observed trend estimates. Unlike the observational uncertainty, the uncertainty represented for the models is the impact of internal variability on estimated trends. In addition, the IQR of the observations is obtained from only five datasets, which represents a less robust estimate of uncertainty than that from the 30-50 simulated realizations in the large ensembles. Spread from these distinctive sources of uncertainty are indicated schematically by the extent of the vertical grey lines.
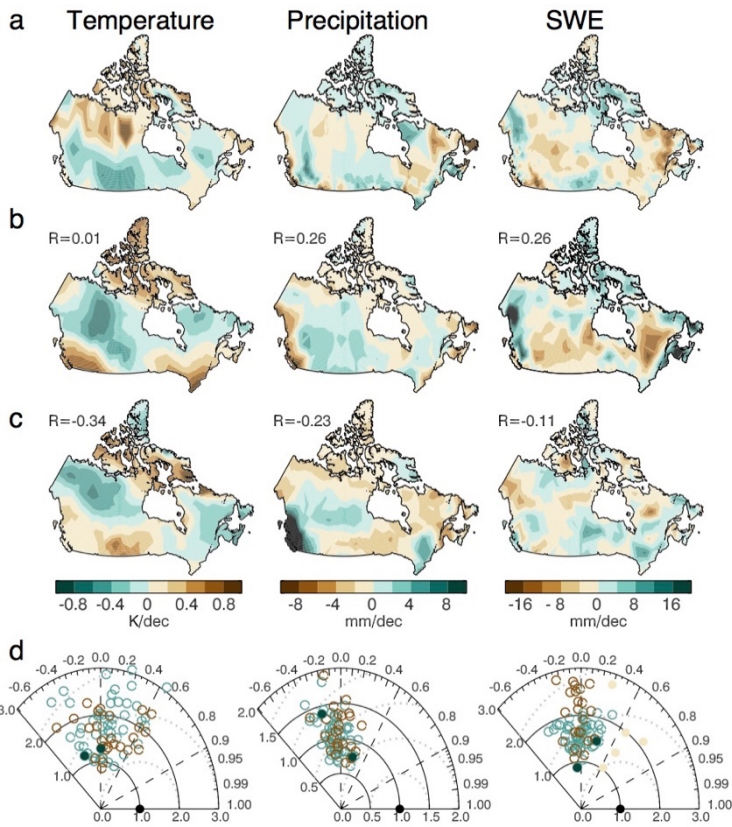
Figure 5: The top row shows the observed 1981-2005 trend over Canada in JFM land-surface temperature, precipitation, and SWE, with the spatial mean of the trend removed, based on the same observational datasets represented in Figs. 1-2. The bottom row shows the Taylor diagram of individual realizations of CanESM2 (teal) and CMIP5 (brown). The Taylor diagram for SWE also includes individual Blended-5 observations (light brown). The second and third rows show the land-surface temperature, precipitation, and SWE trends for the *best match* and *worst match* realizations as defined in the text. The best and worst match realizations are shown as filled circles in the Taylor diagrams. Note that the realizations represented are the same in this figure and in Fig. 6.
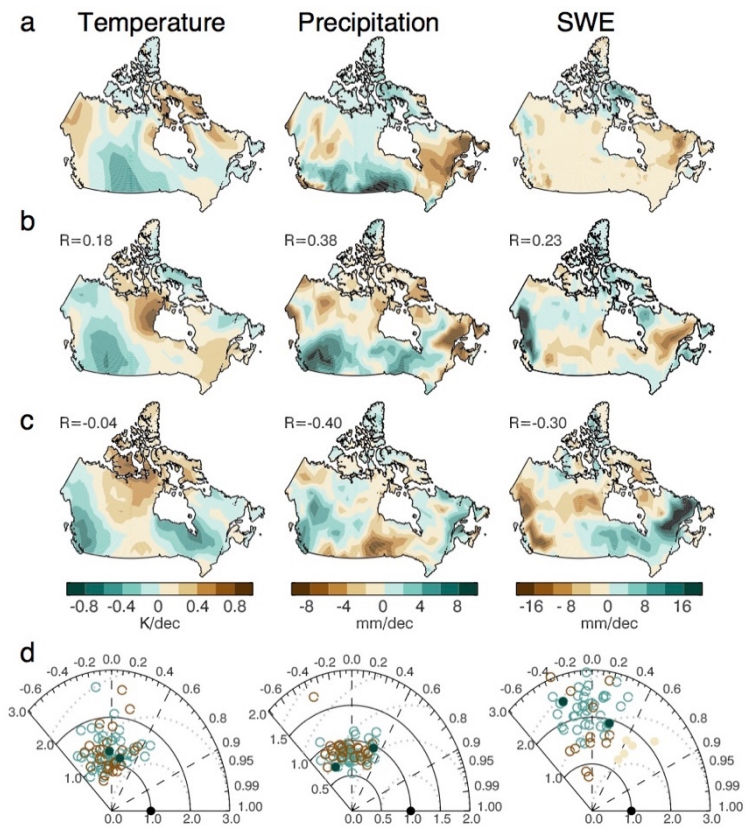
Figure 6: A similar analysis to Fig. 5 but for AMJ. Left column and right column represent AMJ trends, and second and third columns represent the same best-match and worst-match realizations from Fig. 5, but with AMJ trends displayed.
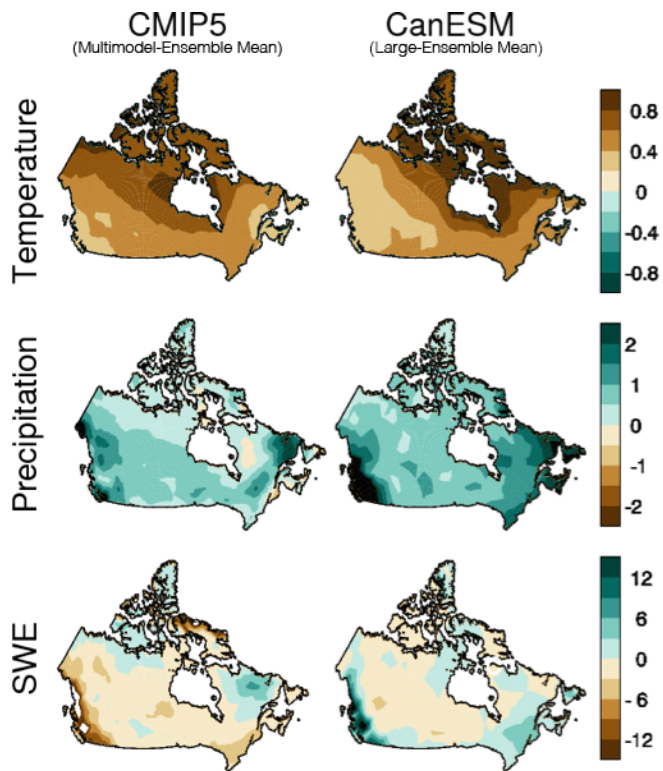
Figure 7: The left column shows the multimodel mean of CMIP5 trends over the 1981-2005 historical period for JFM land-surface temperature, precipitation, and SWE. The right-hand column shows the corresponding multi-realization mean for the CanESM2 Large Ensemble. Unlike Figs. 5-6, the spatial mean of the trend has not been removed.
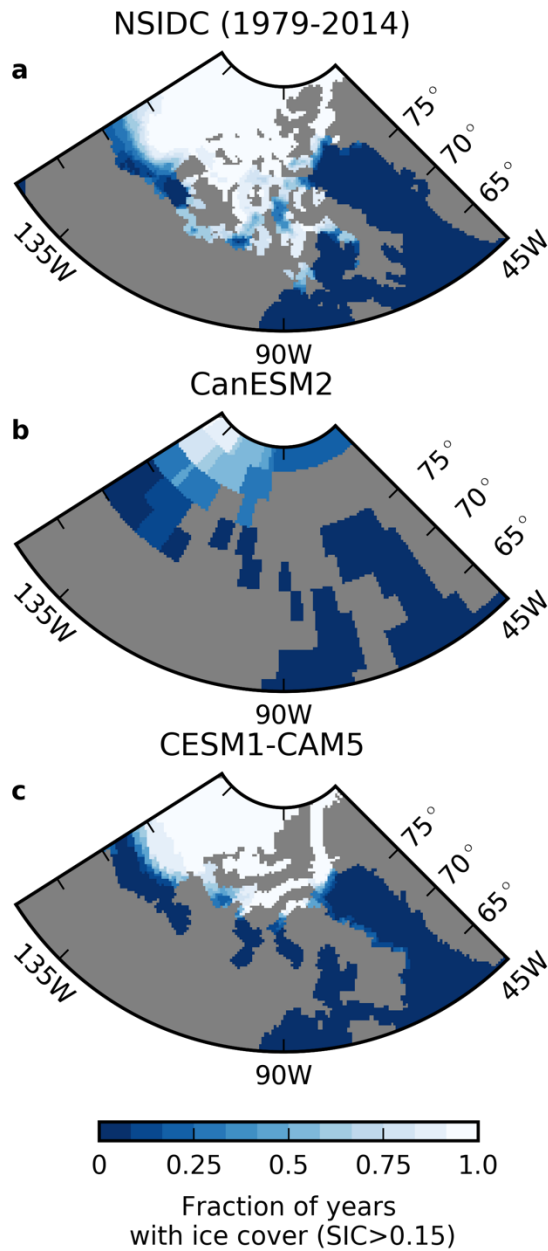
Figure 8: a) Fraction of years (1979-2014) with September sea ice cover (sea ice concentration > 0.15) for the NSIDC passive microwave product on an EASE 25km grid. b) As in a) but for the CanESM2 model, remapped using a nearest neighbour remapping. c) As in b) but for the CESM1-CAM5 model. The grey shading indicates the land-sea mask for each dataset. The NCAR CESM1 CICE grid provide a rough indication of how geographical features of the Canadian Arctic, such as the Canadian Arctic Archipelago, might be resolved in future configurations of CanESM2. Also evident is the low September sea ice extent bias in CanESM2.
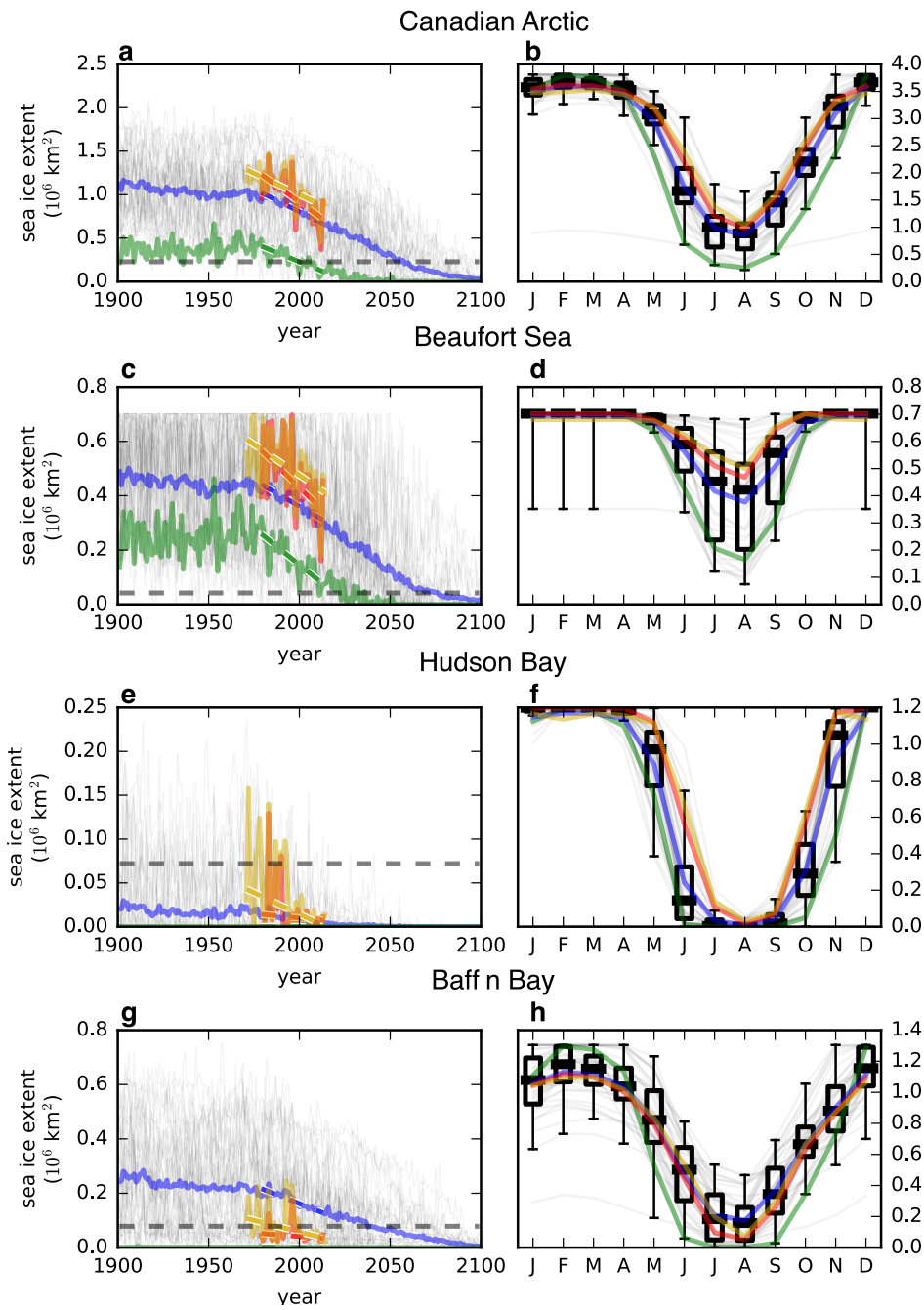
Figure 9: a): September sea ice extent (sea ice concentration larger than 0.15) for the Canadian Arctic [defined by the Canadian Ice Service Data Archive (CISDA) domain]. The CanESM2 model (green), CISDA (yellow), National Snow and Ice Data Center (NSIDC, red) and the multi-model mean (blue) are shown with their 1979-2013 trends. Individual models are shown in light gray. b: The Canadian Arctic seasonal cycle for 1979-2013. Box plots were added to describe the inter model spread (whiskers are 5th and 9th percentiles). c-d, e-f, g-h: same as a, b but for the Beaufort Sea, Baffin Bay and Hudson Bay, respectively. In panels e and g, the CanESM2 (green) Hudson Sea ice amounts are scaled to account for the fractional of ocean present in the CanESM2 land sea model. Individual models are shown in light gray.

a: September sea ice extent (sea ice concentration larger than 0.15) for the Canadian Arctic (defined as the September CISDA domain). The CanESM2 model (green), CISDA (yellow) NSIDC (red) and the multi-model mean (blue) are shown with their 1979-2013 trends. Individual models are shown in light gray. b: The Canadian Arctic seasonal cycle. Box plots were added to describe the inter model spread (whiskers are 5th and 9th percentiles). c-d, e-f, g-h: same as a,b but for the Beaufort Sea, Baffin Bay and Hudson Bay, respectively.

sea ice drift speed [km day^{-1}]

near–surface wind speed [m s^{-1}]

Figure 10: (a) Seasonal cycle of Arctic average sea ice drift speed (solid, in units of km day^{-1}) and near-surface wind speed (dashed, in units of m s^{-1}) from a historical run of CanESM2 averaged over 1979-2005. The spatial domain used for the calculation is the region north of 68°N for longitudes east of 103°E and west of 124°W and north of 79°N at all other longitudes, excluding gridpoints within 150 km of a coastline. This focuses on regions of year-round drifting ice, and excludes landfast ice. (b) As in (a) but using non-gridded drift speed measurements from the International Arctic Buoy Programme (solid; Tschudi et al., 2016) and near-surface wind data from ERA-Interim (dashed; Dee et al., 2011).
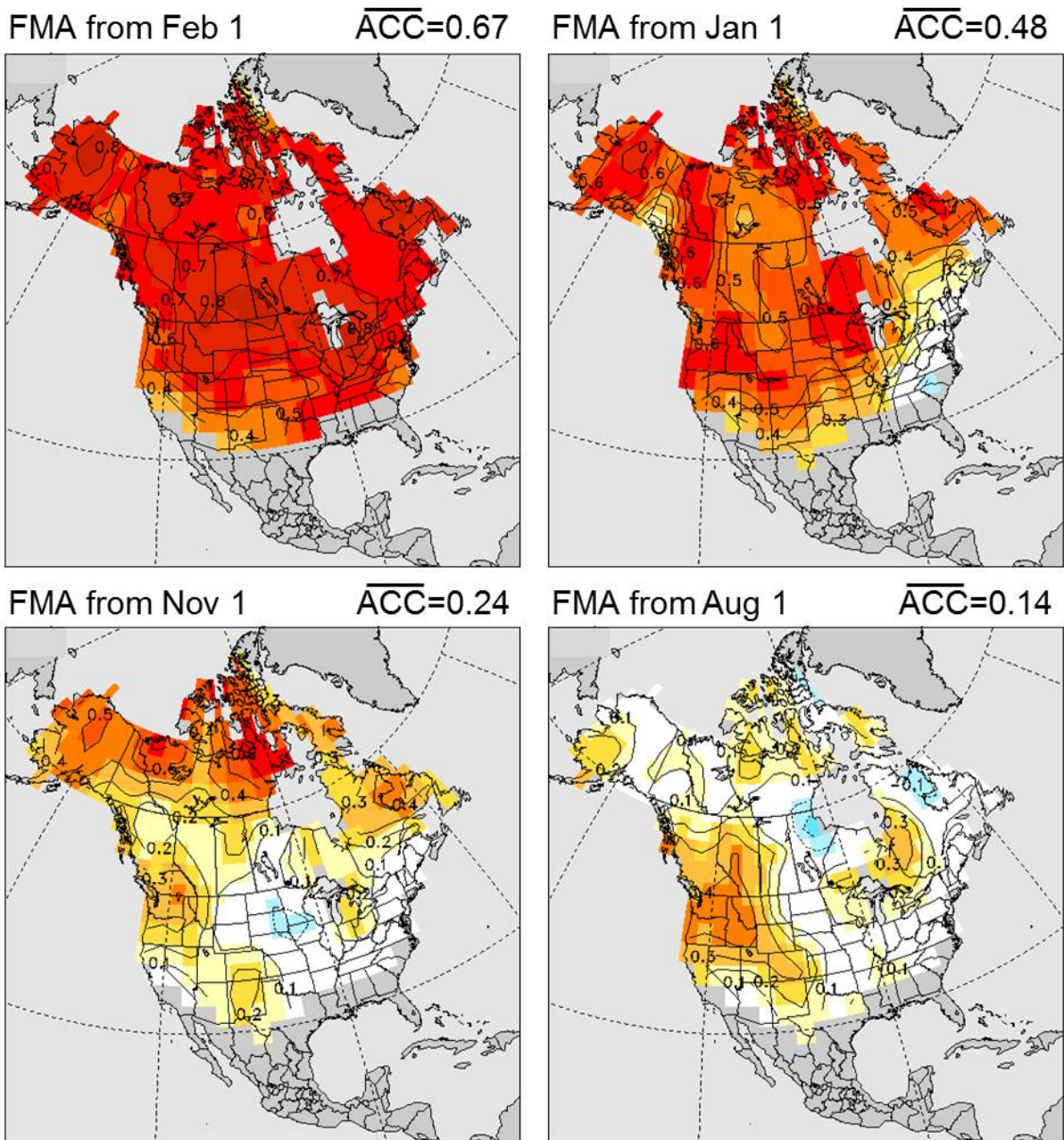
Figure 11: Anomaly correlation coefficient (ACC) skill for February-April (FMA) SWE of 1981-2010 CanSIPS hindcasts initialized at the start of February and the preceding January, November and August (lead times of 0, 1, 3 and 6 months respectively). The Blended-5 dataset of Mudryk et al. (2015) is used for observational verification. Contour interval 0.1, and the overbars denote spatial averages of ACC over areas of North America having seasonal snow cover.
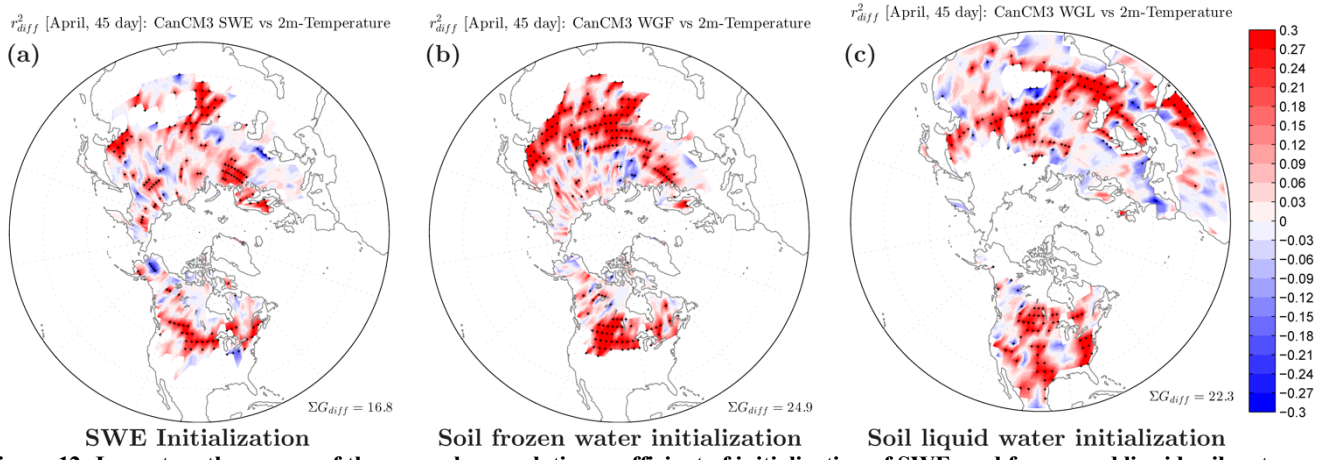
Figure 12: Impact on the square of the anomaly correlation coefficient of initialization of SWE, and frozen and liquid soil water on springtime forecasts of land-surface temperature at 45 day lead. Red colours indicate increase and blue colours decrease of the potential predictability (an idealized model based estimate of potential forecast skill). Based on Ambadan et al. 2015.
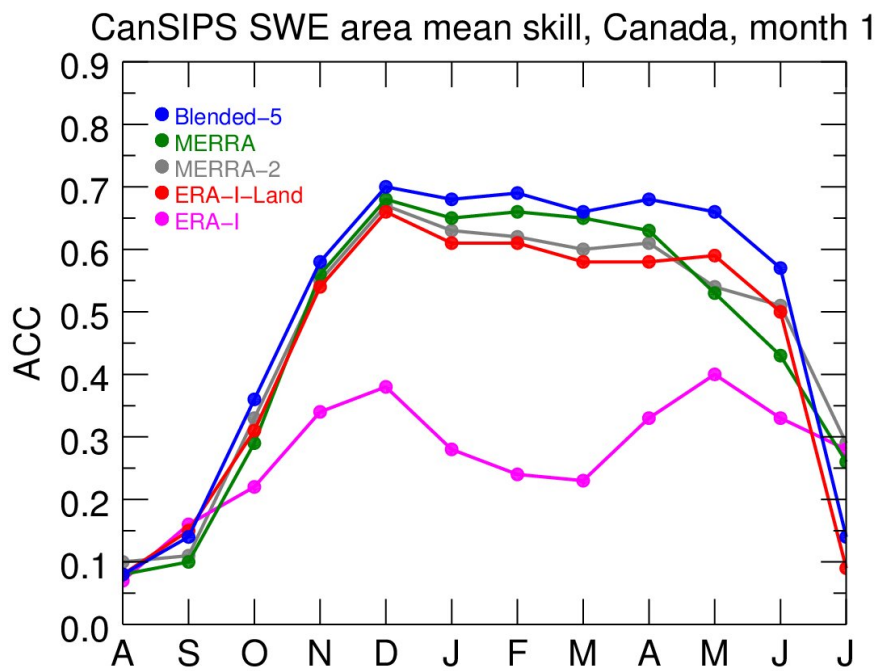
Figure 13: Anomaly correlation coefficient (ACC) averaged over Canada for first-month CanSIPS forecasts of SWE, verified against several observation-based datasets including Blended-5 (blue), MERRA (green), MERRA-2 (grey), ERA-Interim Land (red), and ERA-Interim (magenta); the MERRA and ERA-Interim Land are components of Blended-5. Blended-5 shows the best agreement with the forecast, suggesting a strong influence of observational dataset on evaluation of forecast performance.
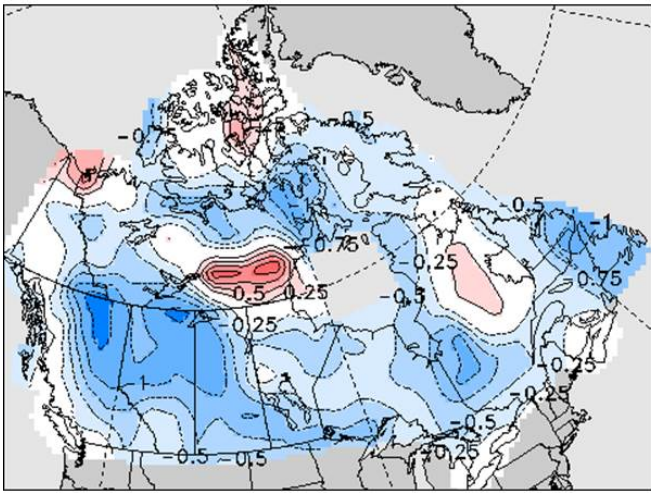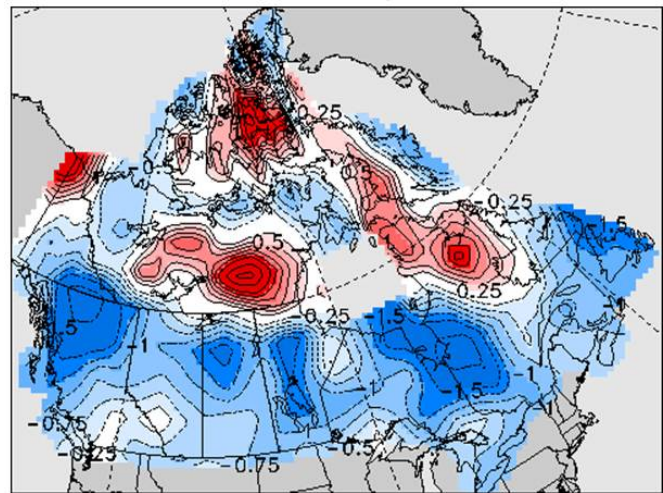
Figure 14: Real time CanSIPS forecast of standardized anomalies of monthly-mean SWE for January 2016 (left), initialized at the end of the preceding month, compared to the MERRA analysis for the same period (right). Contour interval is 0.25, and anomalies are relative to a 1981-2010 base period.
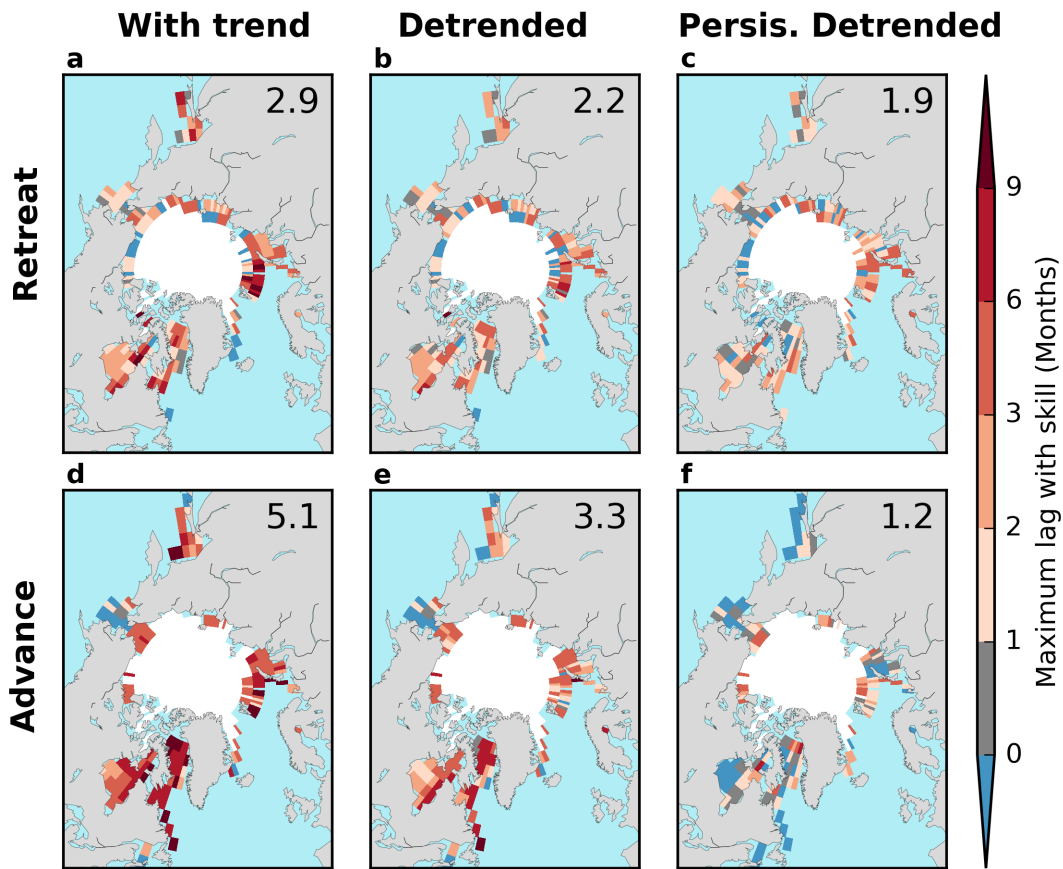
Figure 15: Maximum lead time at which CanSIPS skillfully predicts retreat/advance dates (defined as the calendar date at which sea ice concentration first drops below/exceeds 50%), for total anomalies (first column), detrended anomalies (second column), and for a detrended persistence forecast based on persisting the observed initial sea-ice concentration anomaly. The numbers in the top right corner of each panel indicate the Arctic average maximum lead time (in months). From Sigmond et al. (2016).
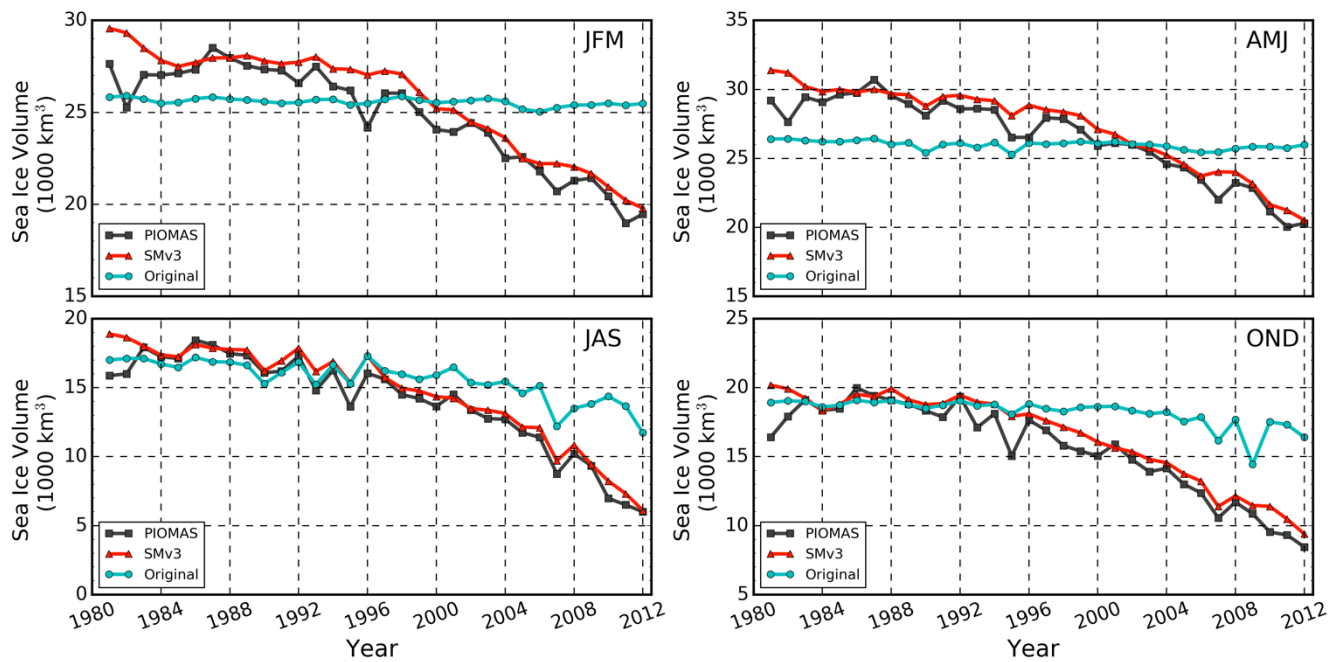
Figure 16: Time series of seasonally averaged sea ice volume over the period 1981–2012, in units of 10^3 km^3 for the PIOMAS sea ice thickness analysis (which assimilates observations) (black), the SMv3 statistical model of Dirkson et al. (2017) (red), and CanSIPS initial conditions (cyan). The CanSIPS system as originally configured incorporates no information about recent sea ice thinning and thus misses the recent downward trends of sea ice volume.