



Environnement  
Canada

Environment  
Canada

# **The Canadian Seasonal to Interannual Prediction System version 2 (CanSIPsv2)**

**Canadian Meteorological Centre Technical Note**

H. Lin<sup>1</sup>, W. J. Merryfield<sup>2</sup>, R. Muncaster<sup>1</sup>, G. Smith<sup>1</sup>, M. Markovic<sup>3</sup>,  
A. Erfani<sup>3</sup>, S. Kharin<sup>2</sup>, W.-S. Lee<sup>2</sup>, M. Charron<sup>1</sup>

*1-Meteorological Research Division*

*2-Canadian Centre for Climate Modelling and Analysis (CCCma)*

*3-Canadian Meteorological Centre (CMC)*

7 May 2019

<b>Revisions</b>			
<u>Version</u>	<u>Date</u>	<u>Authors</u>	<u>Remarks</u>
1.0	2019/04/22	Hai Lin	First draft
1.1	2019/04/26	Hai Lin	Corrected the bias figures. Comments from Ryan Muncaster, Bill Merryfield
1.2	2019/05/01	Hai Lin	Figures of CanSIPsv2 uses CanCM4i plus GEM-NEMO
1.3	2019/05/03	Bill Merrifield Hai Lin	Added CanCM4i information, sea ice verification, 6.6 and 9
1.4	2019/05/06	Hai Lin	All figures of CanSIPsv2 with CanCM4i and GEM-NEMO, made available by Slava Kharin

## Table of Contents

1	Introduction .....	4
2	Modifications to models .....	6
2.1	CanCM4i .....	6
2.2	GEM-NEMO .....	6
3	Forecast initialization .....	7
3.1	CanCM4i .....	7
3.2	GEM-NEMO .....	8
4	Reforecasts.....	8
4.1	CanCM4i .....	8
4.2	GEM-NEMO .....	9
5	Systematic errors .....	10
6	Skill evaluation of CanSIPsv2 comparing to CanSIPS.....	14
6.1	Surface air temperature .....	15
6.2	Precipitation .....	17
6.3	500-hPa geopotential height.....	20
6.4	SST and ENSO.....	23
6.5	The PNA and NAO .....	25
6.6	Sea ice .....	27
7	Qualitative evaluation of the parallel run .....	29
8	Details of implementation .....	29
9	Summary.....	30
10	Acknowledgements.....	30
11	References.....	31
Appendix A	List of Acronyms.....	35

## 1 Introduction

In September 1995 CMC began issuing forecasts for seasonal mean near-surface temperatures and accumulated precipitation in Canada using a two-model ensemble dynamical prediction system based on the first Historical Forecast Project, or HFP (Derome et al. 2001). These forecasts for the standard seasons DJF, MAM, JJA and SON were issued four times per year on December 1, March 1, June 1 and September 1. Initially the two component atmospheric circulation models, each with six ensemble members, were CCCma's AGCM2 (McFarlane et al. 1992), and the SEF global spectral model developed at Recherche en prévision numérique (RPN; Ritchie, 1991), whereas the GEM model (Côté et al. 1998) developed at RPN replaced the SEF model in March 2004.

CMC began issuing longer-lead statistical forecasts for seasonal mean near-surface temperatures and accumulated precipitation in Canada in late 1996. These forecasts, at lead times of three, six and nine months, were based on application of Canonical Correlation Analysis (CCA) with seasonal means of sea surface temperature (SST) and 500-mb geopotential height over the preceding 12 months as predictors (Shabbar and Barnston 1996). The seasons forecast by this system were the same as the zero-lead dynamical forecasts (except with longer lead) and the dates of issuance were also the same.

The CMC's dynamical seasonal forecast system underwent a major upgrade in October 2007 to a four-model system using AGCM2, SEF, GEM and CCCma's AGCM3 (Scinocca et al. 2008) with total of 40 ensemble members, 10 from each model. From the beginning, each 10-member model ensemble was constructed from 24-h lagged initial conditions covering the 10 previous days of the forecast initiation. That period was reduced to 5 days in 2008 by using 12h lags. Seasonal forecasts were issued both at zero-month lead (months 1-3) and one-month lead (months 2-4) on the first day of every month, and a first-month forecast for surface temperature only was issued on the 1st and 15<sup>th</sup> of every month. Hindcasts for bias correction and skill assessment were provided by the second Historical Forecasting Project (HFP2; Kharin et al. 2009). The CCA-based longer-lead forecasts continued to be issued four times per year.

The main basis for skill in a dynamical seasonal forecasting system is its ability to model the atmospheric response to relatively slowly varying boundary conditions, primarily anomalies in sea surface temperature (SST). Because all the models comprising the HFP and HFP2 systems are atmospheric circulation models with no dynamical ocean, future SSTs in these forecasts had to be prescribed, based on information available at the beginning of the forecast. Although there are numerous possible choices for such a prescription, in the HFP and HFP2 systems SSTs during the forecast periods were specified to be the mean SST anomalies during the 30 days preceding the forecast, added to the seasonally varying SST climatology. Of course, a serious limitation of this prescription of anomaly persistence is that actual SST anomalies evolve with

typical time scales of a few months, so that the prescribed SST anomalies at the end of the forecast may not resemble the actual SST anomalies at that time. This limits the utility of such “two-tier” forecast systems beyond forecast periods of a few months, and indeed the HFP forecasts were limited to three months and the HFP2 forecasts to four months for this reason.

The two-tier seasonal forecast system was replaced by the Canadian Seasonal to Interannual Prediction System (CanSIPS) in December 2011, which is a global coupled one-tier system. CanSIPS uses two atmosphere-ocean coupled climate models, CanCM3 and CanCM4, that were developed at CCCma, with a total of 20 ensemble members, 10 from each model. From the beginning of each month, 10 ensemble members of forecast for each model were produced for 12 months from initial conditions of small perturbations at the same time, i.e., with a burst start approach. Details on the CanSIPS system are described in Merryfield et al. (2013).

It is noteworthy to point out that CanSIPS has been an integral part of several international activities related to multi-model ensemble seasonal forecasts, including the North American Multi-model Ensemble (NMME) (e.g., Kirtman et al. 2014), the WMO Long-Range Forecast Multi-Model Ensemble (e.g., Kim et al. 2015), and the multi-model climate prediction of the Asian-Pacific Economic Cooperation Climate Center (APCC) (e.g., Min, et al. 2014).

From September 1995 to June 2015, output of the first month from the seasonal forecast system was used to produce the 30-day-averaged temperature anomaly forecast, which was the only monthly forecast product in Canada. Since July 2015, the monthly forecast is produced in CMC based on the Global Ensemble Prediction System (GEPS; Gagnon et al. 2015), which takes advantage of the increased resolution and improved initialization of GEPS (Lin et al. 2016).

The most recent update is that the ECCC numerical weather prediction (NWP) model, GEM, is coupled with the NEMO ocean model to develop an NWP-based global atmosphere-ocean-sea ice coupled model, which is called GEM-NEMO. CanCM4 is upgraded to CanCM4i with improved initialization. GEM-NEMO replaces CanCM3 in CanSIPS and is combined with CanCM4i to form the second version of CanSIPS for multi-model ensemble seasonal prediction. Development of GEM-NEMO seasonal forecast model took place mainly at RPN, exploiting the coupled model and modeling expertise available there, which also benefits from the contribution of the CONCEPTS project.

This document describes the latest modifications to the Canadian Seasonal to Interannual Prediction System, becoming CanSIPS version 2, hereafter referred to as CanSIPSv2. In Sections 2 we present modifications to the prediction components of CanSIPS in support of the operational implementation of July 2019 to replace CanSIPS. Section 3 describes the forecast initialization. The reforecast is introduced in Section 4. Section 5 describes systematic errors in CanCM2 and GEM-NEMO. Skill evaluation results are presented in Section 6. Section 7 gives

evaluation of the parallel run. The detail of implementation is given in Section 8. Finally, we summarize results in Section 9.

The current implementation represents cumulative efforts of many people not only from the CanCM4i and GEM-NEMO teams, but also from other groups such as the modelling and post-processing sections in research and development divisions, and operational sections of Environment and Climate Change Canada (ECCC) co-located at the Canadian Meteorological Centre (CMC) in Dorval, Québec

## **2 Modifications to models**

Like its predecessors (the HFP, HFP2, and CanSIPS systems), CanSIPSv2 is a multi-model system which takes advantage of the generally greater skill of multi-model compared to single-model systems for a given ensemble size (e.g. Kharin et al. 2009).

CanSIPSv2 consists of two global coupled models, CanCM4i and GEM-NEMO. CanCM4i applies an improved initialization to the same version of the CanCM4 model used previously in CanSIPS. This model is described in detail in Merryfield et al. (2013), so only its most basic aspects are outlined here. On the other hand, GEM-NEMO is a new model introduced to CanSIPSv2 and will be described in detail here.

### **2.1 CanCM4i**

CanCM4i, like CanCM4, couples CCCma's fourth-generation atmospheric model CanAM4 to CCCma's CanOM4 ocean component. Atmospheric horizontal resolution is T63, corresponding to a 128×64 Gaussian grid, with 35 vertical levels and a top at 1 hPa. The 256×192 grid of CanOM4 provides horizontal resolution of approximately 1.4 degrees in longitude and 0.94 degrees in latitude, with 40 vertical levels. Version 2.7 of the CLASS land surface scheme and a single-category, cavitating-fluid sea ice model are employed, both formulated on the atmospheric model grid.

### **2.2 GEM-NEMO**

GEM-NEMO, developed at the Recherche en Prévision Numérique (RPN), is a fully coupled model with the atmospheric component of GEM (Côté et al. 1998) and the ocean component of NEMO (Nucleus for European Modelling of the Ocean, <http://www.nemo-ocean.eu>).

The GEM version used is 4.8-LTS.13, with a horizontal resolution of 256×128 grid points (about 155 km) evenly spaced on a latitude-longitude grid, and 79 vertical levels with the top at 0.075 hPa. The land scheme is ISBA (Noilhan and Planton, 1989; Noilhan and Mahfouf, 1996), where each grid point is assumed to be independent (no horizontal exchange). The atmospheric deep

convection parameterization scheme of Kain-Fritsch (Kain and Fritsch 1990) is used. The Kuo-transient scheme (Bélair et al. 2005) is applied for shallow convections. The model allows for vegetation and ozone seasonal evolutions during the integration. The evolution of anthropogenic radiative forcings are treated simply through specification of equivalent CO<sub>2</sub> concentrations with a linear trend. The GEM model is integrated with a time step of one hour. Among the 10 ensemble members in both the forecast and the reforecast, implicit surface flux is used in members 1-5 and explicit surface flux is applied in members 6-10.

The ocean model NEMO is NEMO 3.6 ORCA 1 with a horizontal resolution of 1°×1° (1/3 degree meridionally near the equator) and 50 vertical levels. The CICE 4.0 (Community of Ice CodE, Hunke and Lipscomb 2010) model is used for the sea-ice component with five ice categories. The NEMO model is run with a 30-minute time step.

GEM and NEMO are coupled once an hour through the GOSSIP coupler. No flux corrections are employed in the coupled model.

### **3 Forecast initialization**

Similar to what was done in CanSIPS, each of the two models in CanSIPSv2, CanCM4i and GEM-NEMO, makes its real-time forecast once a month, initialized at the beginning of each month. Each model produces forecasts of 10 ensemble members with all the integrations starting from the same date and lasting for 12 months. However, the two models take different approaches in the initialization for the forecasts.

#### **3.1 CanCM4i**

The only change in CanCM4i forecast initialization compared to CanCM4 initialization in CanSIPS is that initial conditions for Northern Hemisphere (Arctic) sea ice thickness are obtained from the SMv3 statistical model of Dirkson et al. (2017). (Additional changes in reforecast initialization that are intended to improve consistency between reforecasts and real-time forecasts are described in section 4.) Other aspects, common to CanCM4 in CanSIPS, are that each CanCM4i ensemble member is initialized from a separate assimilating model run in which atmospheric temperature, moisture, horizontal winds, SST and sea ice concentration are constrained by values from the Global Deterministic Prediction System (GDPS) analysis. Land surface variables are initialized through the response of the land surface component to atmospheric conditions in the assimilating model runs, whereas subsurface ocean temperatures are constrained by values from the Global Ice Ocean Prediction System (GIOPS) analysis through an offline procedure. Details of the initialization methods are provided in Merryfield et al. (2013).

## 3.2 GEM-NEMO

In the forecast, the atmospheric initial conditions of GEM-NEMO come from those of the Global Ensemble Prediction System (GEPS; Gagnon et al. 2015), that are generated from EnKF with data assimilation. Ten of the GEPS perturbed initial conditions are used for the GEM-NEMO seasonal forecast.

For the land surface initial conditions, in addition to taking advantage of the CMC analysis one important consideration is to make it consistent with the reforecast. In the reforecast, which will be introduced in section 4.2, we initialize the land surface by running the Surface Prediction System or SPS (Carrera et al. 2010) forced by the near-surface atmospheric and the precipitation fields of the ERA-interim reanalysis. In the real-time forecast, the land surface initial fields are generated by forcing the SPS with the CMC analysis.

The ocean and sea ice initial conditions in the forecast come from the CMC GIOPS analysis (Smith et al. 2016). The near surface air temperature from GEPS and the snow depth from SPS are used for the initialization of these variables over sea ice in CICE. For the reforecast, the ocean and sea ice initial conditions are described in section 4.2. All ensemble members have the same ocean and sea ice initial conditions.

## 4 Reforecasts

For seasonal predictions, model drift and systematic model errors become a serious problem that contaminates the forecast quality. A common practice in seasonal prediction is to perform a historical reforecast (or hindcast) to estimate the model climatology and statistics, so that calibrated anomaly forecasts can be made. Another purpose of the reforecast is to generate a long record of forecast data to quantify the performance of the forecast system. Therefore, the reforecast is an important component of the seasonal forecast system. As in CanSIPS, the reforecast is made with each individual model, covering 31 years from 1980-2010 with 10 members of 12-month integration starting from the beginning of each month. In the reforecast, each individual model has the same configuration as in the forecast.

### 4.1 CanCM4i

CanCM4i employs several changes to the CanCM4 reforecast initialization procedure described in Merryfield et al. (2013). First, as in the real time forecasts, modelled sea ice thickness is constrained near values from the SMv3 statistical model of Dirkson et al. (2017), which is much more realistic than the stationary model-based climatology employed for CanCM4, in part because it accounts for the long-term thinning of Arctic sea ice. Second, modelled sea ice concentration is constrained near values from the Had2CIS product described in section 4.2. Had2CIS provides improved temporal consistency across the reforecast period, as well as improved consistency with sea ice concentrations from the GDPS analysis used in real time,



compared to the concentration product used to initialize CanCM4 reforecasts. Finally, initial subsurface ocean temperatures are constrained by the ORAP5 ocean reanalysis (Zuo et al. 2015), which is both closer in its formulation to the GIPS analysis used in real time, and has more realistic trends than the ocean reanalysis used to initialize CanCM4 reforecasts.

## **4.2 GEM-NEMO**

In the reforecast, the atmospheric conditions are initialized using the ERA-interim reanalysis (Dee et al, 2011). Random isotropic perturbations are added to the reanalysis fields to create 10 different initial conditions. The atmospheric perturbations are homogeneous and isotropic as in Gauthier et al. (1999). Only the streamfunction and the unbalanced temperatures are perturbed as in EnKF (see Houtekamer et al., 2009). These perturbed fields are then transformed to wind, temperature and surface pressure.

The ocean initial fields come from the ORAP5 ocean reanalysis (Zuo et al. 2015). The monthly values of temperature, salinity, zonal and meridional currents are interpolated to daily values. The sea ice concentration initial fields are the Had2CIS, which was prepared by Woo-Sung Lee of CCCma, who combined the digitized sea ice charts from the Canadian Ice Service (CIS) with the HadISST2.2. The original HadISST2.2 employs an ice chart-based bias correction of the passive microwave record (Titchner and Rayner 2014). The monthly HadISST2.2 data, the digitized CIS weekly sea ice charts over the Arctic region and the weekly CIS "Great Lakes ice charts" over Great Lakes are interpolated to daily data before the combination. The sea ice thickness is interpolated from the monthly ORAP5 data. The near surface air temperature from ERA-interim and the snow depth from SPS are used for the initialization of these variables over sea ice in CICE.

For the land surface initialization, we make use of historical dataset of surface fields produced by SPS (Carrera et al. 2010). It is previously known as GEM-surf and it was used in several studies at high resolution (see Separovic et al. 2014, Ioannidou et al. 2014 and Bernier and Bélair 2012). To generate more realistic land surface conditions, we have run the SPS for the period 1980-2010 to create surface fields compatible with the surface scheme of the model. The SPS is simply the surface schemes (land, sea ice and glacier) of the GEM model used in offline mode. To initialize the land surface fields in the reforecast, we have forced SPS by the near-surface atmospheric and the precipitation fields coming from the ERA-interim reanalyses. The surface pressure, the 2-m temperature and dew-point depression as well as the solar and infra-red downward radiative fluxes at the surface are provided to SPS at a 3-hour interval. To limit the snow accumulation over glaciers, the maximum snow depth is set to 12 meters in the initial condition.

## 5 Systematic errors

In this section, the systematic error is analyzed for the CanCM4i and GEM-NEMO models, which is the difference between the predicted and observed climatology over the reforecast period of 1981-2010. The objective is to demonstrate the general behavior of the two models related to model drift. When doing seasonal prediction, forecasts of anomalies are produced, where the model climate estimated from the reforecast is subtracted, and the systematic error is removed in the forecast. It thus appears that there is no direct impact of the systematic error on the forecast skill, which will be presented in the next section. However, the systematic error does affect the forecast skill through the model dynamics. For example, a biased middle-latitude westerly jet would lead to errors in Rossby wave propagation and teleconnection patterns forced by ENSO, resulting in an inaccurate anomaly forecast. Therefore, it is important to reduce the model bias.

The verification data are the ERA-interim reanalyses for 2-meter air temperature (T2m) and 200-hPa zonal wind (U200), the Global Precipitation Climatology Project (GPCP; Huffman et al. 2009) dataset for precipitation, and the NOAA Optimum Interpolation (OI) Sea Surface Temperature V2 (Reynolds et al. 2002) for SST. The systematic error presented here is for the 1-month lead seasonal mean, averaged over the 12 initialization months.

We first look at the model bias of SST, which is shown in Figure 1. As can be seen, CanCM4i is doing a good job in controlling the model drift, with biases only in limited areas near the west coast of South America, east coast of East Asia and North America, and along the tropical Pacific and high latitude North Atlantic and Southern oceans (Figure 1a). On the other hand, GEM-NEMO systematically has cold biases in the tropical oceans (Figure 1b), indicating that the model drift in GEM-NEMO is relatively strong. This is an area of future improvement for GEM-NEMO.

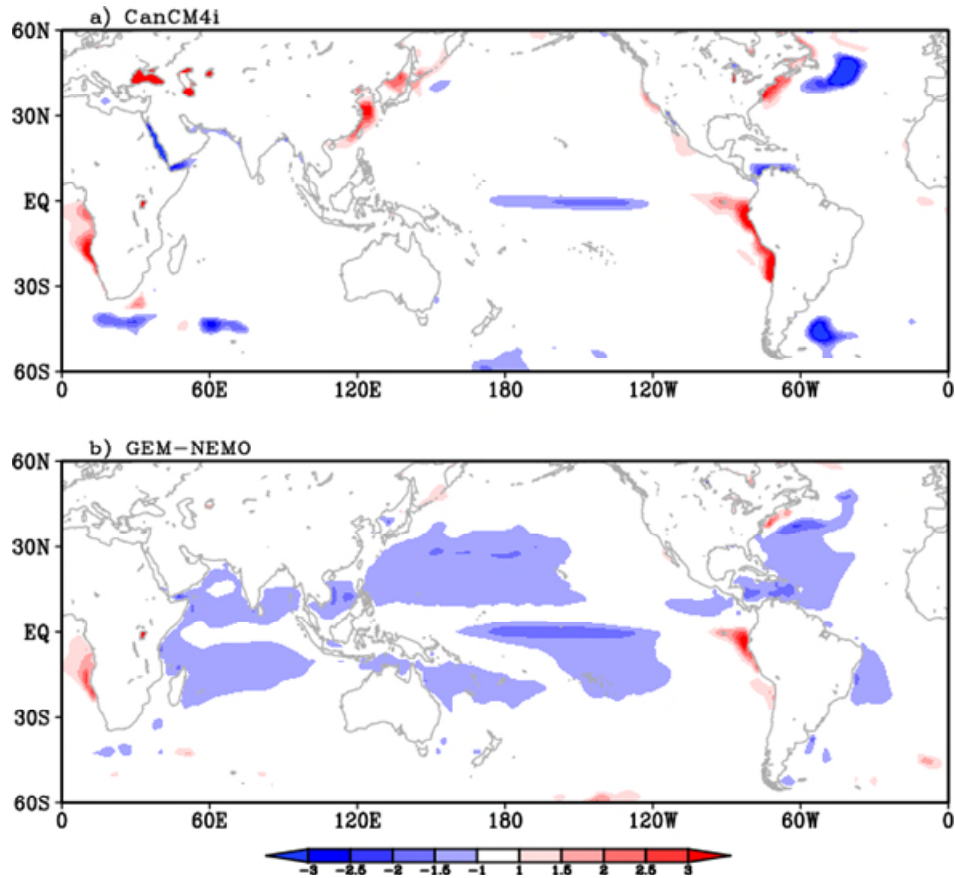


Figure 1. Seasonal mean SST systematic error at one-month lead averaged for all 12 initialization months a) for CanCM4i, and b) for GEM-NEMO.

Figure 2 shows the geographical distribution of systematic error for T2m. The CanCM4i model is seen to have quite strong cold biases over the Northern Hemisphere Polar Region (Figure 2a), that may be related to the treatment of sea ice physics, or atmospheric physics over sea ice. Both models have warm T2m biases over the Northern middle-high latitude continents, and cold biases over the mountains. CanCM4i has warm biases over equatorial South America. Over the ocean, CanCM4i shows some warm biases near the west coast of South America, east coast of East Asia and North America. Cold biases are observed in GEM-NEMO in the tropical Pacific, which are related to the SST biases as discussed above.

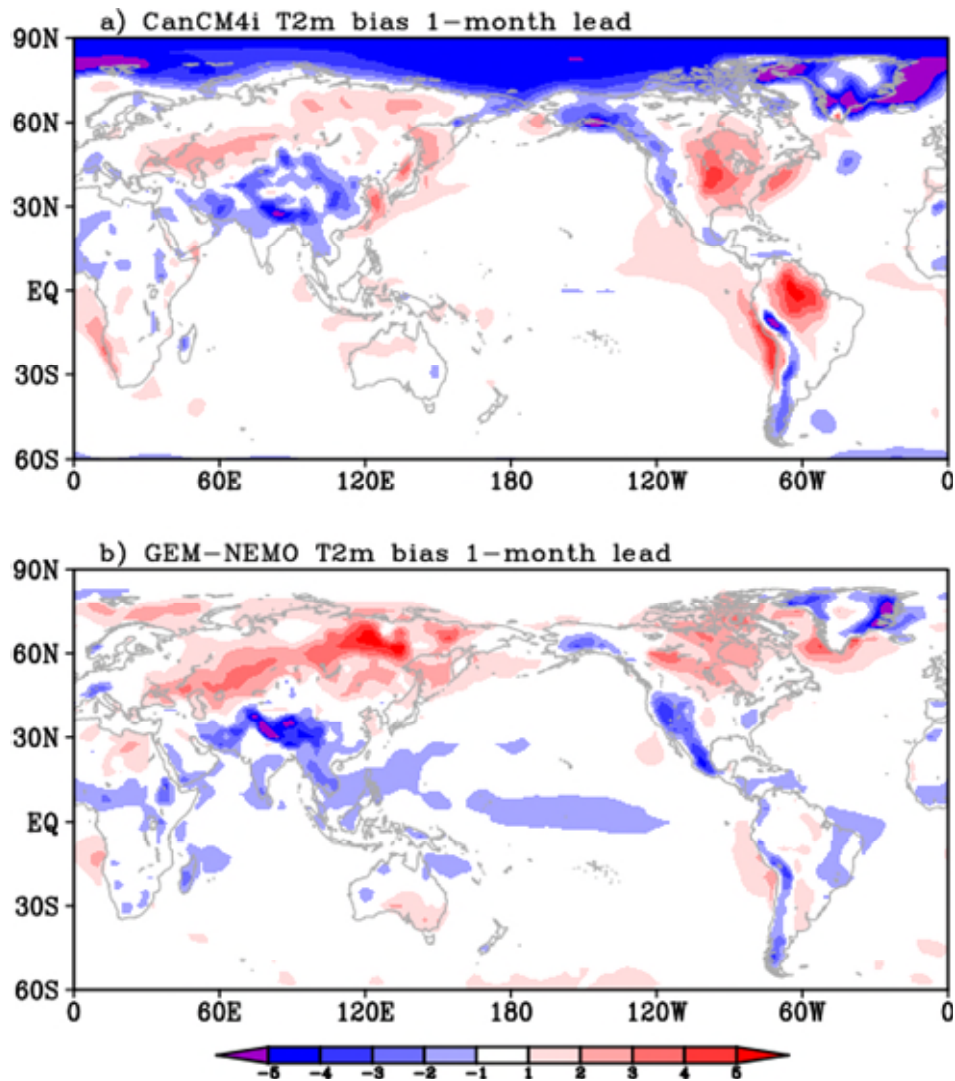


Figure 2. Seasonal mean T2m systematic error at 1-month lead for CanCM4i (top) and GEM-NEMO (bottom). Unit: °C.

Shown in Figure 3 is the systematic error in precipitation rate. The two models seem to have a similar precipitation bias distribution. Excessive precipitation is found over a large area of the tropics with insufficient precipitation over the equatorial Pacific, a feature of the double-intertropical convergence zone (ITCZ) which is common to many coupled general circulation models (e.g., Lin 2007). The magnitude of the wet bias in CanCM4i is stronger than GEM-NEMO. CanCM4i also has a strong wet bias in the tropical Indian ocean, which is absent in GEM-NEMO. Over the tropical South American continent, CanCM4i has a stronger dry bias than GEM-NEMO.

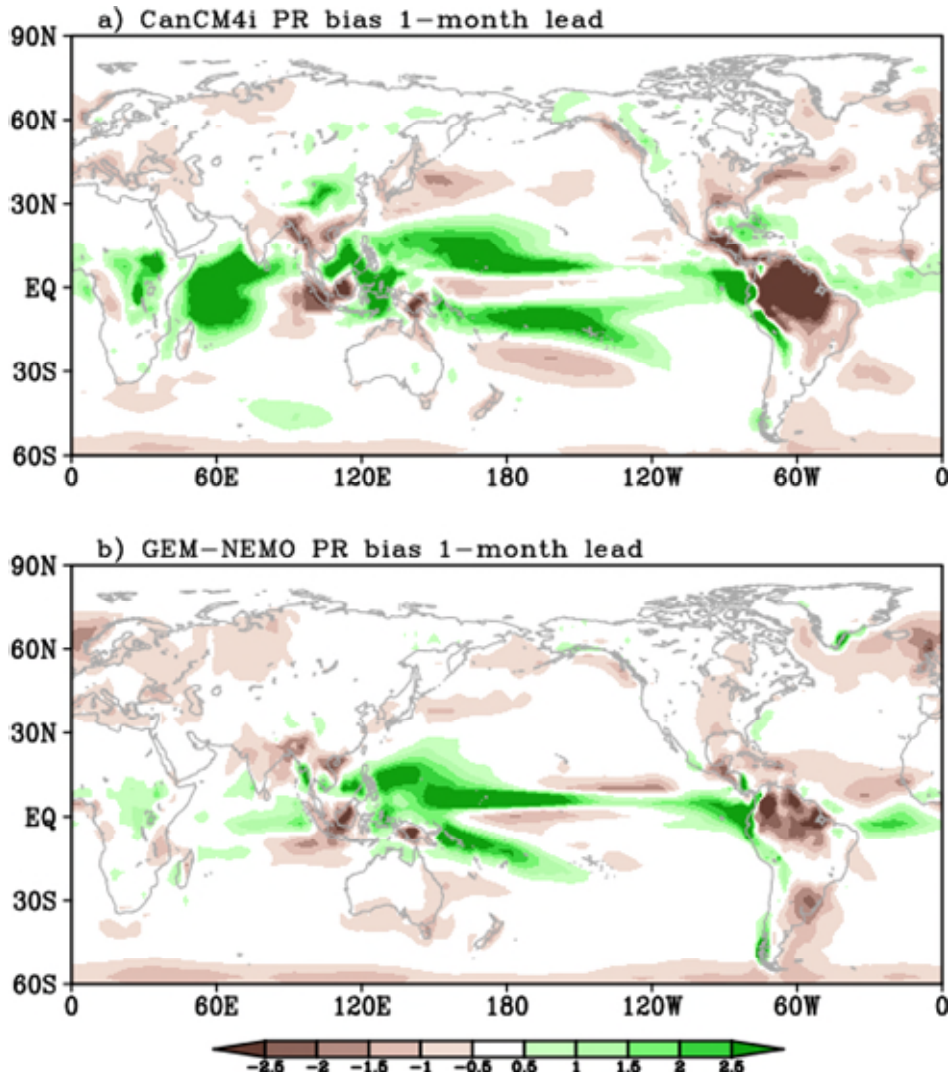


Figure 3. Precipitation systematic error at 1-month lead for CanCM4i (top) and GEM-NEMO (bottom). Unit: mm/day.

Many previous studies have discussed the importance of the extratropical zonal westerly flow for Rossby wave propagation (e.g., Hoskins and Ambrizzi, 1993). The systematic error of U200 is illustrated in Figure 4. The two models have quite different distributions of U200 bias in the extratropical regions. CanCM4i tends to have overestimated westerlies in the Northern Hemisphere middle latitudes, whereas GEM-NEMO has easterly zonal wind systematic errors.

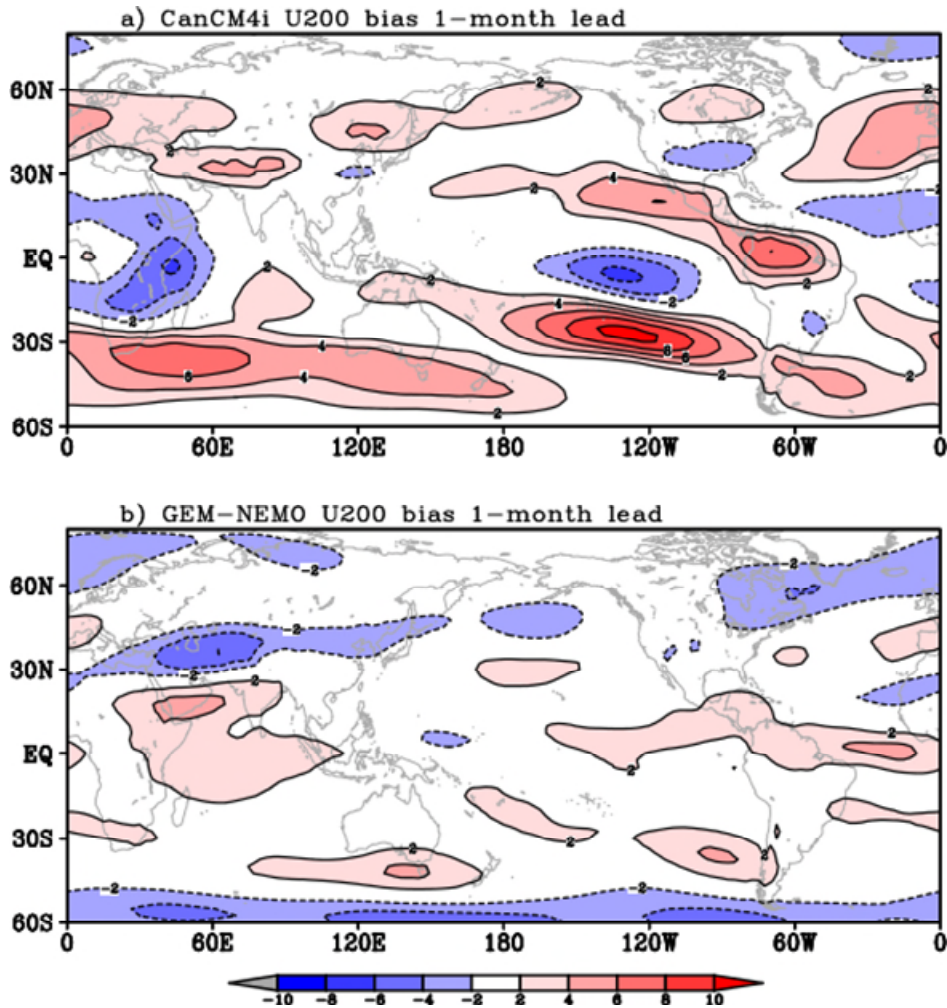


Figure 4. Systematic errors of 200 hPa zonal wind at 1-month lead for CanCM4i (top) and GEM-NEMO (bottom). Unit: m/s.

## 6 Skill evaluation of CanSIPsv2 comparing to CanSIPS

The performance of CanSIPsv2 is evaluated in comparison with CanSIPS for the hindcast period 1981-2010. The verification data used are the ERA-interim reanalysis for air temperatures and geopotential height, and the GPCP V2.3 dataset for precipitation.

The skill measures considered are for deterministic forecasts, determined from the ensemble mean of the predicted anomalies. In addition to the global maps for winter and summer forecasts, we consider area averages (over the globe, global land, North America and Canada for temperature and precipitation, over the globe, the Northern Hemisphere extratropics, tropics and Southern Hemisphere extratropics for 500-hPa geopotential height) of anomaly correlation, i.e. the correlation coefficient between predicted and observed anomalies. In addition, in assessing the ensemble forecast skill over Canada, we also discuss the continuous ranked probability skill



score (CRPSS; e.g., Bradley and Schwartz 2011) which measures differences between the forecast and observed distributions. A comparison of the skills of the two systems follows.

### 6.1 Surface air temperature

Figure 5 shows the global geographical distribution of anomaly correlation skill for predictions of seasonal mean surface air temperature from the CanSIPS hindcasts (left) and the new CanSIPSv2 hindcasts (right), for the DJF season with zero-month lead. As is usually the case the highest skill for both systems is in the tropics and especially the tropical Pacific, where SST anomalies are most persistent and ENSO imparts relatively high predictability. Skills are also appreciable in many locations over land, including most of North America where much of the seasonal predictability particularly in winter and early spring is attributable to the teleconnected influence of ENSO. In general, the skill distributions of the two systems are comparable. The global average skill of CanSIPSv2 is higher than CanSIPS. Similar conclusion can be obtained for the JJA season (Figure 6), as well as other seasons (not shown).

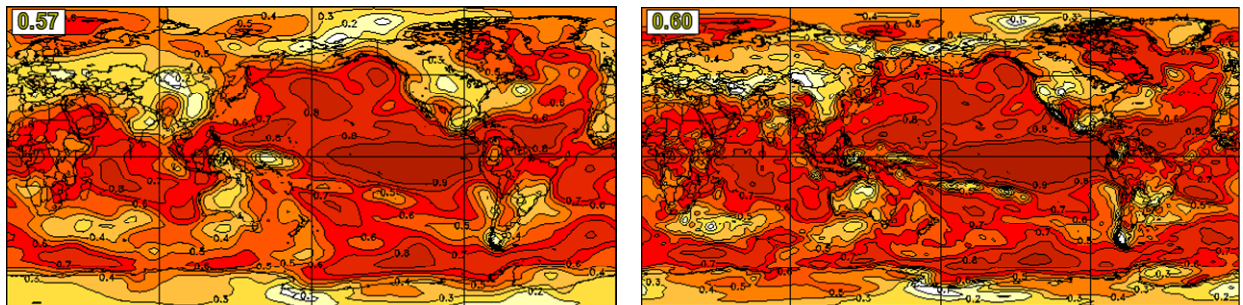


Figure 5. Geographical distribution of anomaly correlation skill for predictions of seasonal mean DJF near-surface air temperature at zero month lead, for the CanSIPS (left) and CanSIPSv2 (right), based on historical forecasts for 1981-2010. Verification data is ERA-interim. Globally averaged skills are marked in the upper-left corner of each panel.

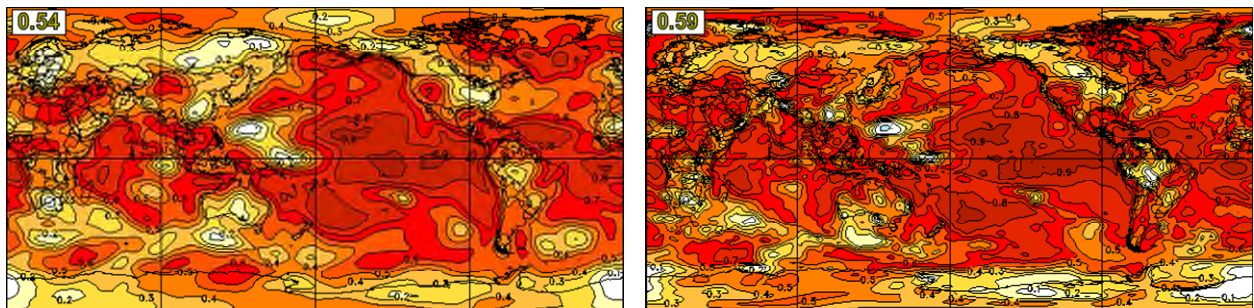


Figure 6. Same as Figure 5, but for JJA.

Shown in Figure 7 are T2m anomaly correlation skills for different regions and for lead times of zero month (top) and one month (bottom) averaged over all the 12 initialization months. It is

evident that GEM-NEMO alone has a performance comparable to or better than that of CanSIPS. By combining GEM-NEMO with CanCM4i, CanSIPSv2 skill is significantly improved over that of the CanSIPS hindcasts over all the regions.

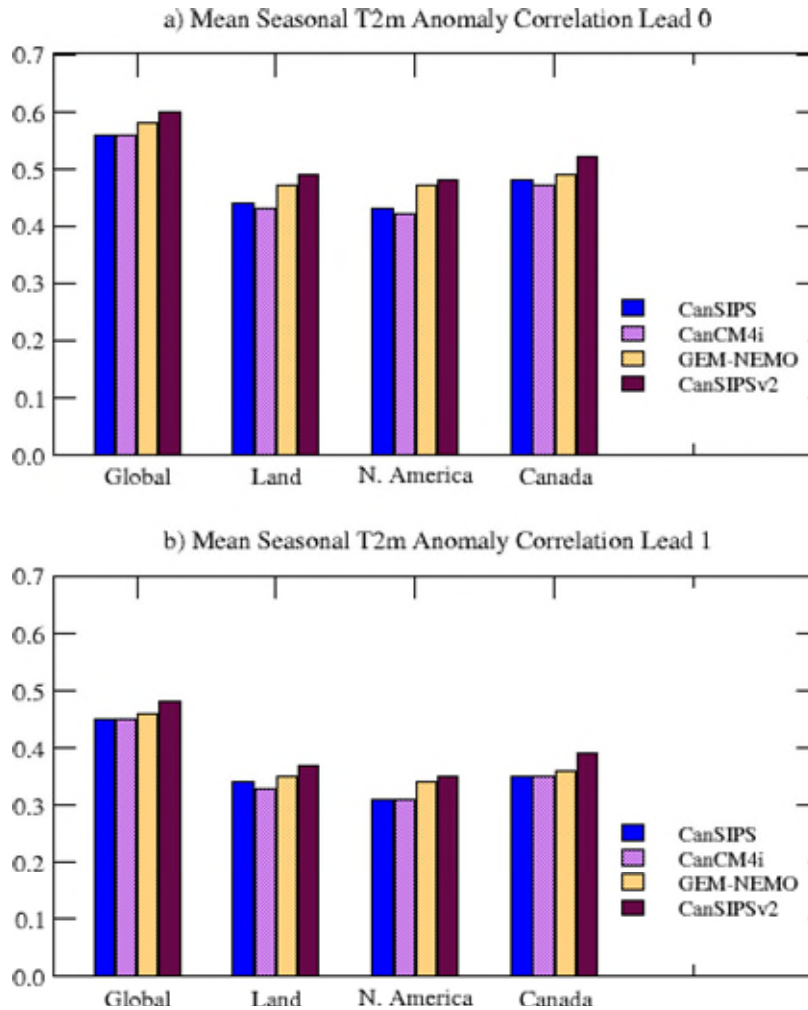


Figure 7. Anomaly correlation skills for hindcasts of seasonal mean near-surface air temperature at lead times of zero month (top) and one month (bottom). The skills shown are averages for all 12 initialization months. The CanSIPSv2 hindcast skills are shown separately for the individual component models CanCM4i and GEM-NEMO as well as for the combined CanSIPSv2 hindcasts. The scores are compared with CanSIPS.

To measure the quality of probability forecasts of the ensemble system, we present in Figure 8 the continuous ranked probability skill score (CRPSS) for DJF seasonal mean T2m forecast at zero month lead in Canada. In both CanSIPS and CanSIPSv2, high skill scores are observed in a large part of northeast Canada and the west coastal regions. Relatively low skill is seen in the Prairies. The CRPSS skill distribution is consistent with the anomaly correlation of the ensemble mean forecast as shown in Figure 5. Comparing the two systems, it is clear that CanSIPSv2



performs better than the previous CanSIPS, especially in eastern Canada. Similar feature of the CRPSS skill scores can be seen for JJA (Figure 9). It is also observed that the skill is improved in CanSIPSv2 comparing to CanSIPS.

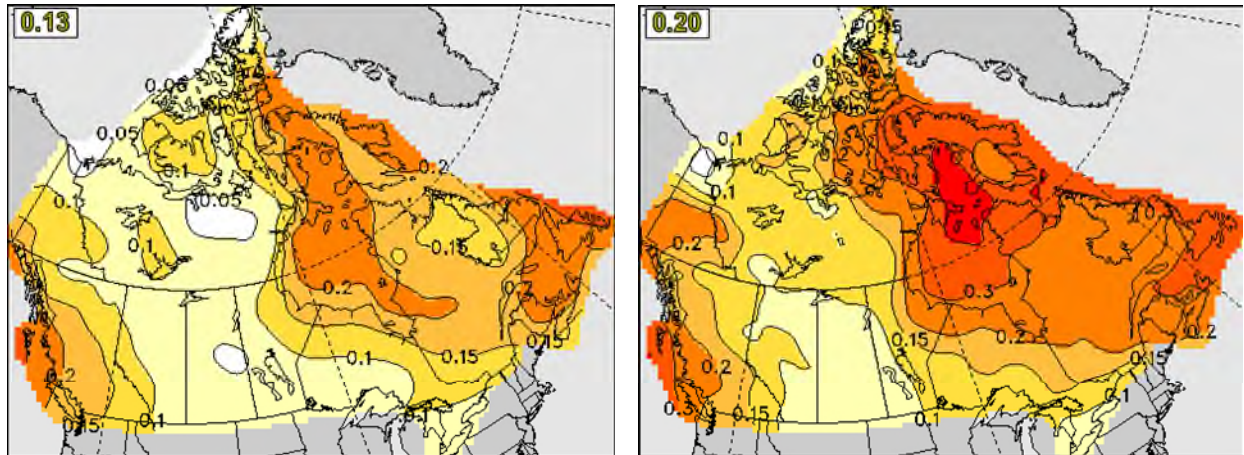


Figure 8. CRPSS skill for seasonal mean DJF near-surface air temperature at zero month lead, for the CanSIPS (left) and CanSIPSv2 (right), based on historical forecasts for 1981-2010. Verification data is ERA-interim. Globally averaged skills are marked in the upper-left corner of each panel.

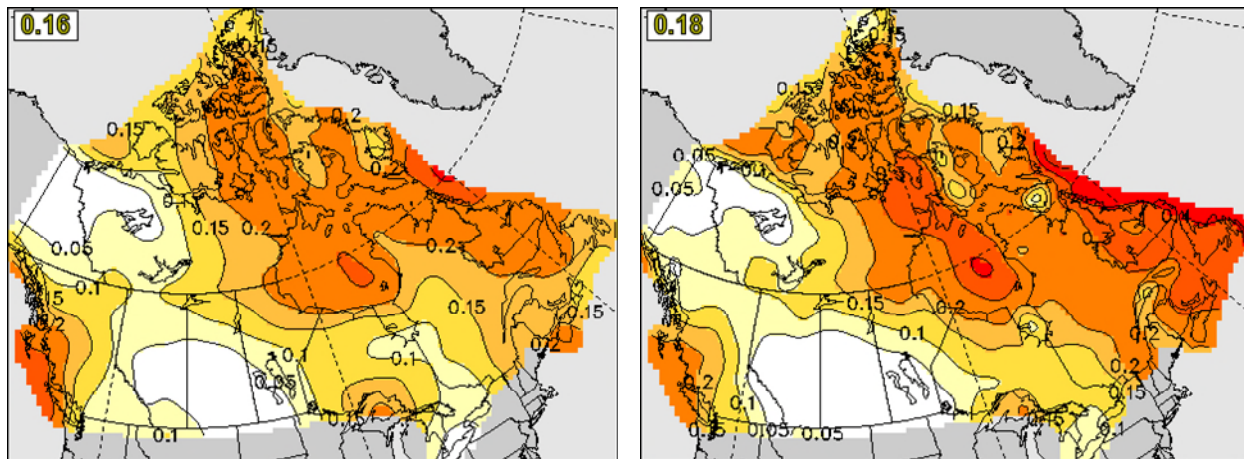


Figure 9. Same as Figure 8, but for JJA.

## 6.2 Precipitation

The global distribution of anomaly correlation skill for seasonal mean precipitation forecast in DJF with zero-month lead is shown in Figure 10 for CanSIPS (left) and CanSIPSv2 (right) hindcasts. Again, high skill is mainly observed in the tropical regions, reflecting the contribution of ENSO. Some skill of precipitation is seen in the western and eastern coastal regions of North

America. The global average skill of CanSIPsv2 is higher than CanSIPS, although the distributions of the two systems are very similar. For the JJA forecast, the precipitation skill has a similar distribution as DJF but with a weaker magnitude (Figure 11). In this season, both systems produce skillful seasonal precipitation forecasts over western North America. Again, the global average skill of precipitation in CanSIPsv2 is higher than that of CanSIPS.

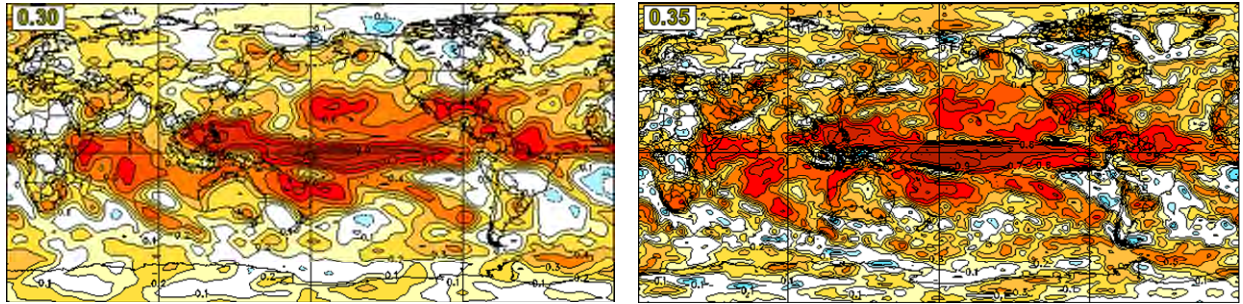


Figure 10. Geographical distribution of anomaly correlation skill for predictions of seasonal mean DJF precipitation at zero month lead, for the CanSIPS (left) and CanSIPsv2 (right), based on historical forecasts for 1981-2010. Verification data is GPCP V2.3. Globally averaged skills are marked in the upper-left corner of each panel.

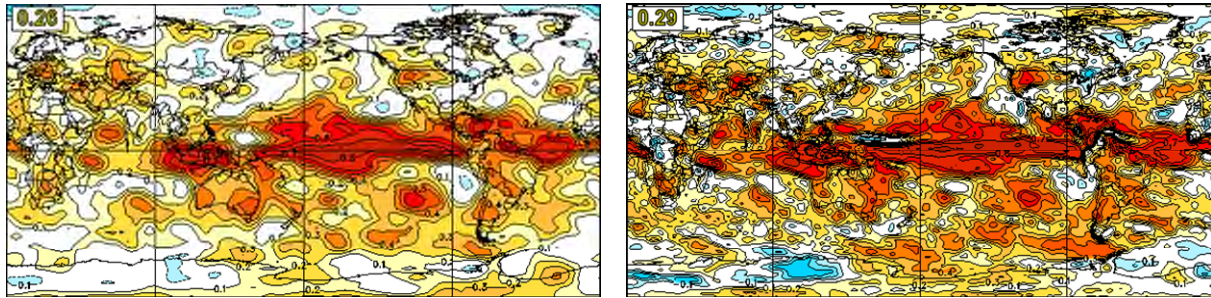


Figure 11. Same as Figure 10, but for JJA.

The anomaly correlation skills of precipitation averaged over all 12 initialization months for different regions are illustrated in Figure 12 for lead times of zero month (top) and one month (bottom). In general, the precipitation skill is low, especially for lead-time of one month. It is seen that CanSIPsv2 outperforms CanSIPS in precipitation forecast in all regions for the 0-month lead seasonal forecast. For the 1-month lead, we see improvement of precipitation skill over the globe and land but some degradation over North America and Canada, although the skill is very low.

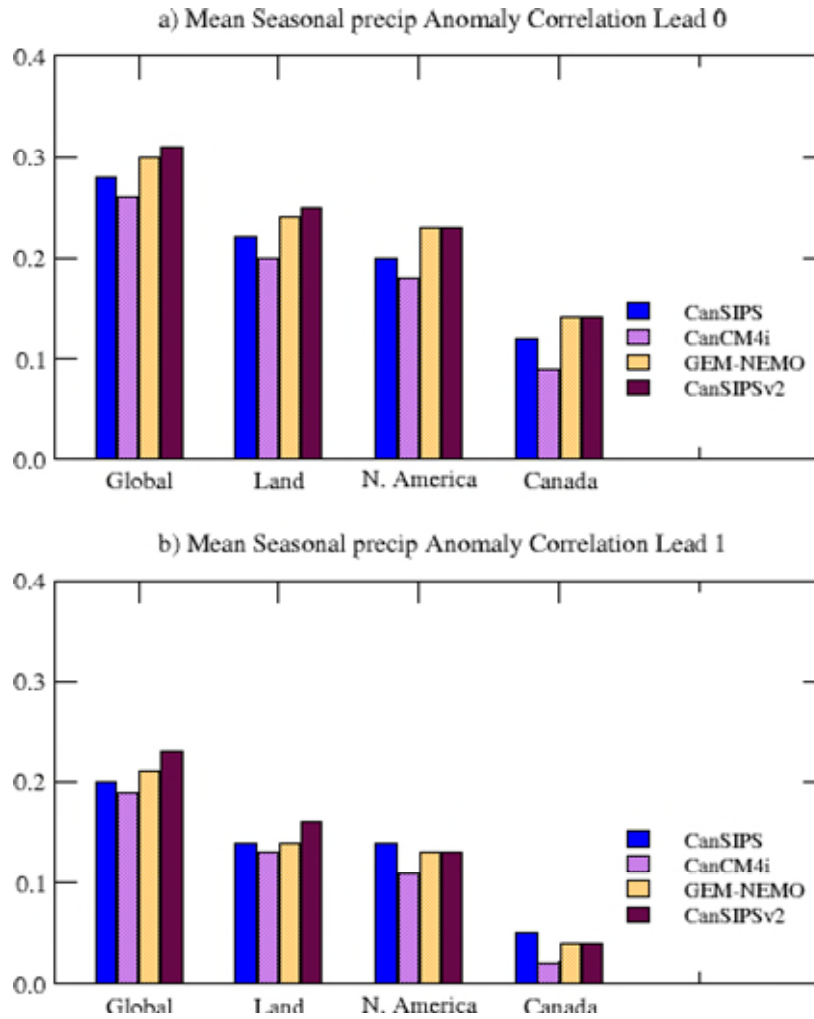


Figure 12. Same as Figure 7, except for precipitation verified against GPCP V2.3, at lead times of zero month (top) and one month (bottom).

The CRPSS skills for DJF and JJA seasonal mean precipitation forecasts at zero month lead in Canada are shown in Figure 13 and Figure 14, respectively. Although the precipitation forecast skill is low, appreciable improvement in CanSIPSv2 can be identified for DJF. The two systems have comparable CRPSS skill in JJA.



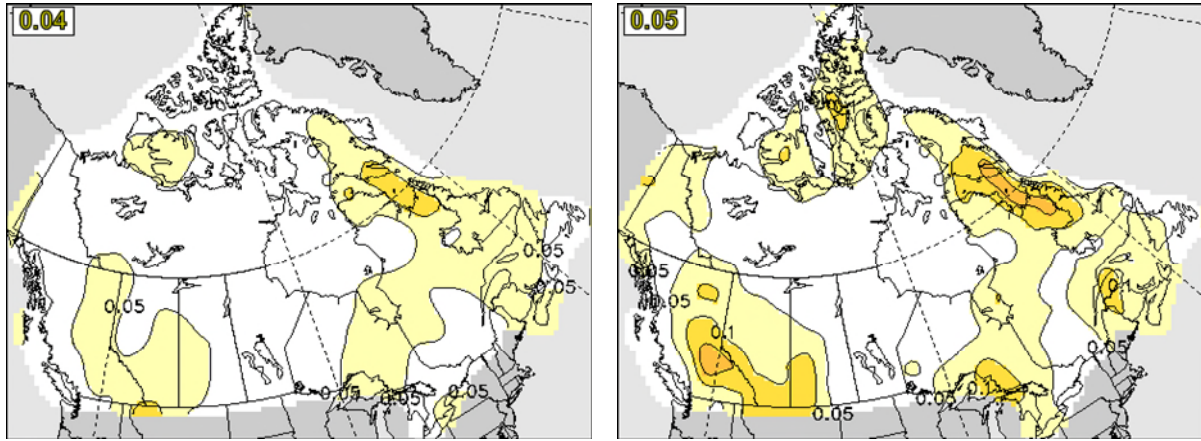


Figure 13. CRPSS skill for seasonal mean DJF precipitation at zero month lead, for the CanSIPS (left) and CanSIPSv2 (right), based on historical forecasts for 1981-2010. Verification data is GPCP V2.3. Globally averaged skills are marked in the upper-left corner of each panel.

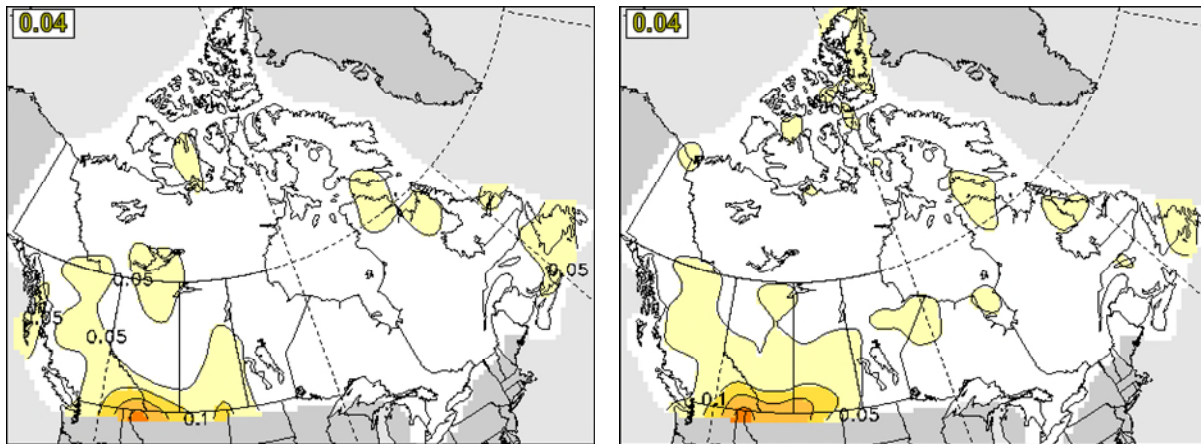


Figure 14. Same as Figure 13, but for JJA.

### 6.3 500-hPa geopotential height

In this section, we evaluate the seasonal forecast skill of 500-hPa geopotential height (Z500). Spatial distributions of the forecast skill for zero lead for the winter season (DJF) and summer season (JJA) are presented.

The anomaly correlation skills for DJF seasonal mean Z500 at zero month lead, i.e., initialized on December 1, are compared between CanSIPS and CanSIPSv2 in Figure 15. Similar to T2m and precipitation, significant skill of Z500 is found in the tropics. In the Northern Hemisphere extratropics, the correlation skill is high over the North Pacific and most part of Canada, resulting from variability of the PNA associated with ENSO (e.g., Wallace and Gutzler 1981). Relatively high correlation skills are also observed over Greenland and the middle latitude North

Atlantic, which appears related to the NAO. In general, the two systems have a similar skill distribution. However, the skills over the North Pacific and Greenland region as well as the global average are higher in CanSIPsv2 than in CanSIPS.

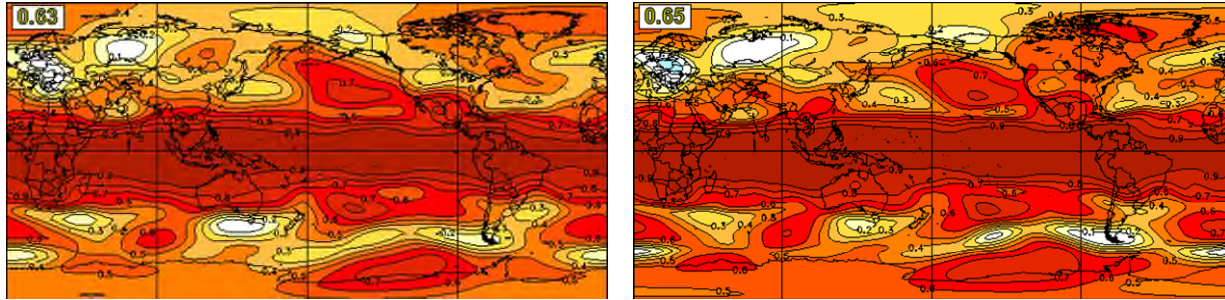


Figure 15. Geographical distribution of anomaly correlation skill for seasonal mean DJF 500-hPa geopotential height at zero month lead, for the CanSIPS (left) and CanSIPsv2 (right), based on historical forecasts for 1981-2010. Verification data is ERA-interim reanalysis. Globally averaged skills are marked in the upper-left corner of each panel.

For the season of JJA, as can be seen in Figure 16, the forecast skill of seasonal mean Z500 is also high in the tropics. In the Northern Hemisphere extratropics, besides the Greenland and polar regions, four centres of high skill are found in a wave pattern along the middle latitudes, that are eastern Europe, China, the North Pacific and western North America. The last is likely responsible for the high temperature and precipitation skills in western North America as seen in Figure 6 and Figure 11. CanSIPsv2 has a higher global average of skill value than CanSIPS, although the two systems have a very similar skill distribution.

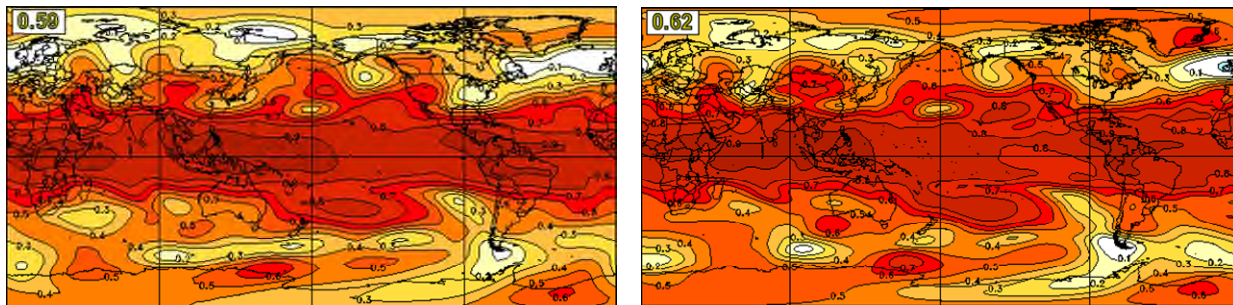


Figure 16. Same as Figure 15, but for JJA.

The anomaly correlation skill averaged over all 12 initialization months for different regions, i.e., the globe, the Northern Hemisphere extratropics (30N-90N), the tropics (30S-30N) and the Southern Hemisphere extratropics (30S-90S), are presented in Figure 17 for lead times of zero month (top) and one month (bottom). Again, the CanSIPsv2 system outperforms CanSIPS in all the regions at both the lead times.

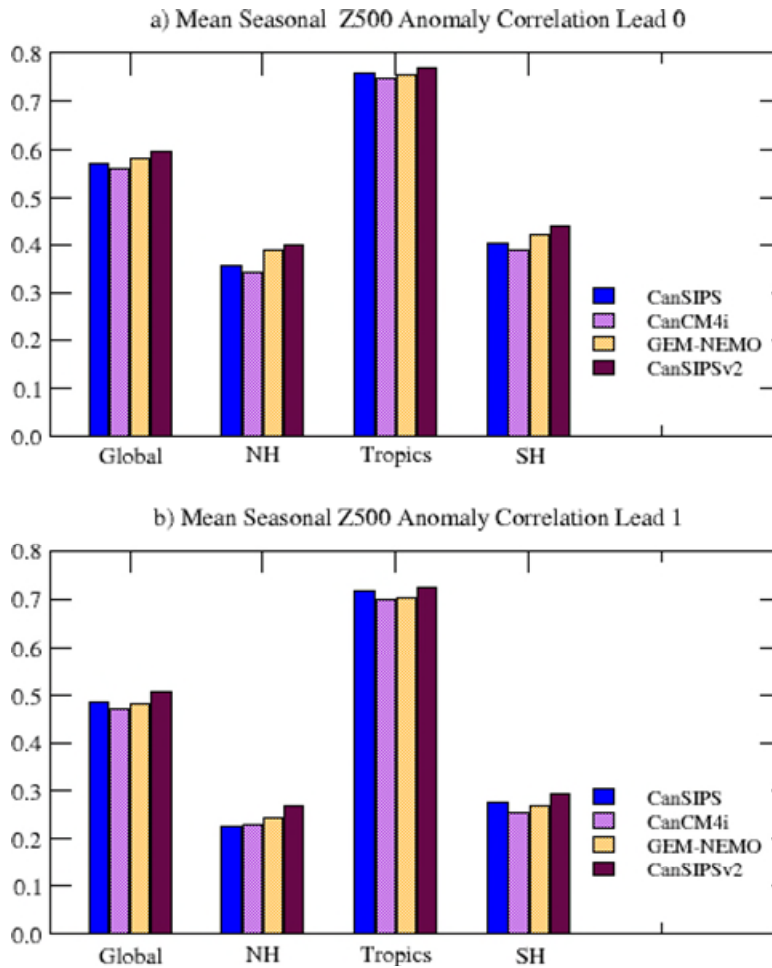


Figure 17. Anomaly correlation skills for hindcasts of seasonal mean Z500 at lead times of zero month (top) and one month (bottom). The skills shown are averages for all 12 initialization months over the globe, the Northern Hemisphere extratropics (30°-90°N), the tropics (30°S-30°N), and the Southern Hemisphere extratropics (90°S-30°S). The CanSIPSv2 hindcast skills are shown separately for the individual component models CanCM4i and GEM-NEMO as well as for the combined CanSIPSv2 hindcasts. The scores are compared with CanSIPS.

In order to see the skill evolution with the increase of lead time, shown in Figure 18 is the Z500 anomaly correlation skill averaged over the globe as a function of lead time. The skills shown are averages for all 12 initialization months. As can be seen, the improvement of forecast skill of CanSIPSv2 over that of CanSIPS occurs for all lead times.



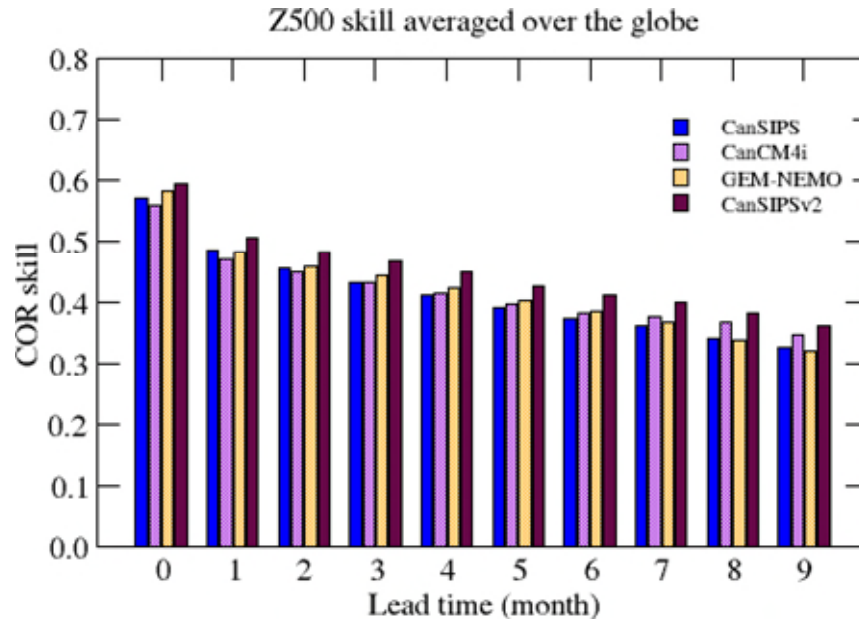


Figure 18. Globally averaged anomaly correlation reforecast skill for seasonal mean Z500 as a function of lead time. The skills shown are averages for all 12 initialization months.

### 6.4 SST and ENSO

We first show the anomaly correlation skill of DJF seasonal mean SST at one-month lead, i.e., initialized on November 1, in Figure 19. The two systems have a very similar skill distribution. High skill of SST forecast can be found over the global ocean, with the maximum skill in the ENSO region of the equatorial eastern Pacific. The global average skill of CanSIPsv2 (0.61) is higher than that of CanSIPS (0.55). The skill for the one-month lead SST forecast of JJA has a similar distribution as in DJF, but weaker and less concentrated in the ENSO region (Figure 20).

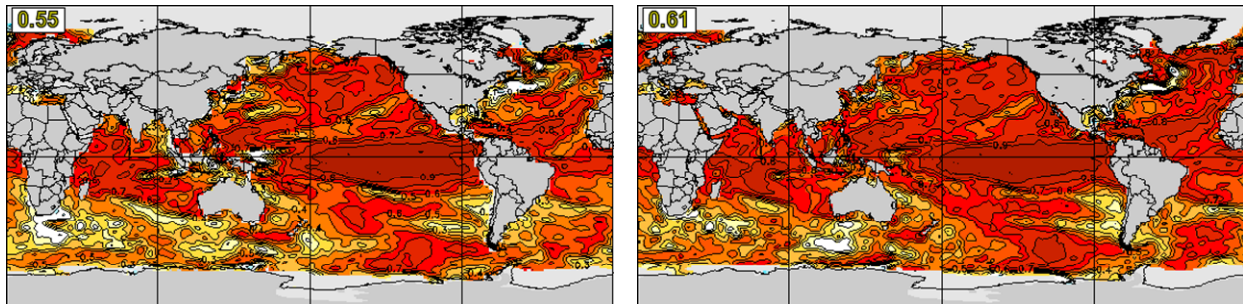


Figure 19. Anomaly correlation reforecast skill of DJF seasonal mean SST at one-month lead. Verified against OISST. Left: CanSIPS, Right CanSIPsv2. The globally averaged skill is marked in the upper-left corner of each panel.

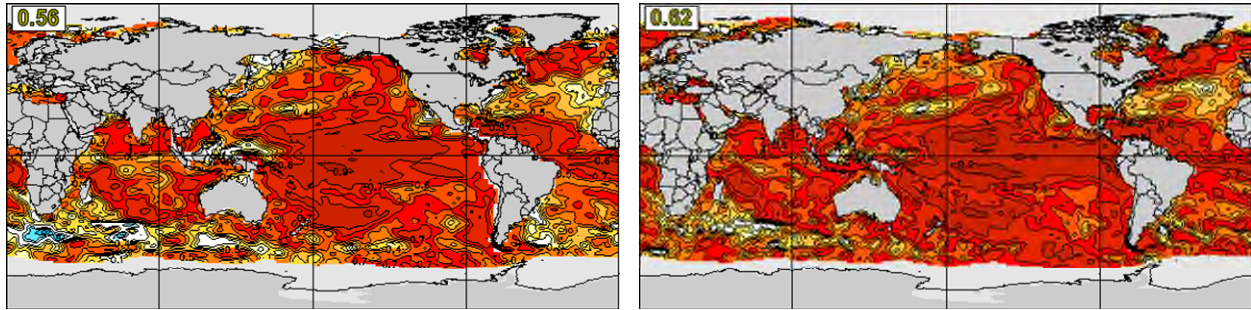


Figure 20. Same as Figure 19, but for JJA reforecast initialized on May 1.

To demonstrate the forecast skill of ENSO, the Nino3.4 index is used which is defined as the SST anomaly averaged in the area of 5°N-5°S, 170°W-120°W. Figure 21 provides a summary of the correlation skill of the seasonal mean Nino3.4 index as a function of target season (vertical axis) and lead time in month (horizontal axis). The skill scores averaged among all the target seasons as a function of lead are printed along the top of the figure, while those averaged among all the lead times as a function of the target season are shown along the right-hand side. The grand average for all the target seasons and lead times is printed at the top-right corner inside the figure. In general, the two systems have comparable ENSO forecast skill. CanSIPSv2 has a slightly better overall skill. For the winter target seasons, CanSIPSv2 shows a better performance. For example, for the DJF and JFM forecasts, the skill of 0.90 for CanSIPSv2 reaches seven-month lead, which is about two months longer than CanSIPS.

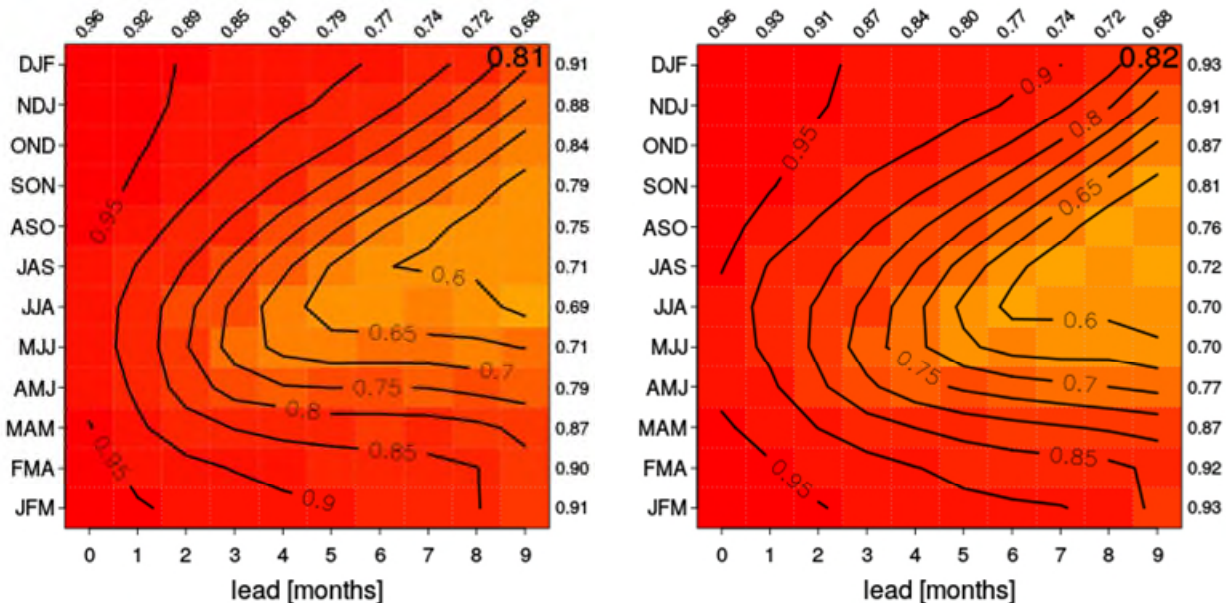


Figure 21. Anomaly correlation reforecast skill of seasonal mean Nino3.4 index as a function of target season (vertical axis) and lead time (horizontal axis). Verified against OISST. Left: CanSIPS, Right CanSIPSv2.



## 6.5 The PNA and NAO

The Pacific-North American pattern and the North Atlantic Oscillation (NAO) are two leading modes of interannual variability of seasonal mean circulation in the Northern Hemisphere extratropics (e.g., Wallace and Gutzler 1981), that greatly influence the weather and climate. Here we analyze the forecast skill of these two patterns.

The PNA and NAO are defined as the first and second rotated empirical orthogonal function (REOF) modes of the monthly mean 500-hPa geopotential height over the Northern Hemisphere using the NCEP/NCAR reanalyses, following Barnston and Livezey (1987). Monthly mean anomalies for all 12 calendar months are used. Since the PNA and NAO variability has the largest variance in the cold season, the patterns are thus dominated by characteristics in winter. The PNA and NAO indices are calculated as the projections of the seasonal mean 500 hPa geopotential height anomalies onto the respective REOF patterns. In the following discussion, comparisons of PNA and NAO skills are made between CanSIPS and CanSIPSv2. The verification indices are obtained by projecting the corresponding ERA-interim seasonal mean Z500 anomaly onto the PNA and NAO patterns.

The correlation skills of the PNA index as a function of target season are presented in Figure 22 for lead times of zero month (top) and one month (bottom). The PNA skill changes with target season, with high skill in winter, spring to early summer. In late summer and autumn when the PNA is not well defined, the skill is low. In general, the PNA skill of CanSIPSv2 is similar to that of CanSIPS. The wintertime PNA skill is improved in CanSIPSv2 comparing to CanSIPS at the 0-month lead time (Figure 22a). At the 1-month lead, however, there is some degradation of wintertime PNA skill, especially for DJF and JFM (Figure 22b).

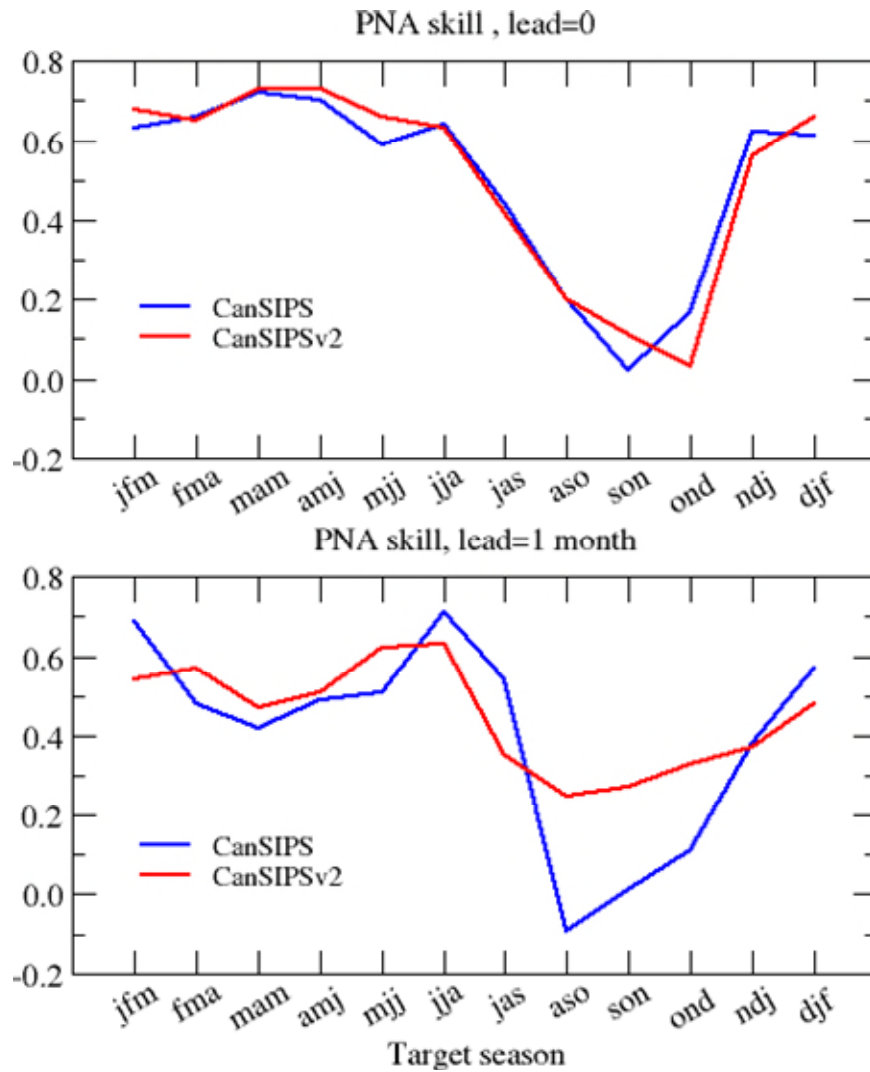


Figure 22. Correlation skill of seasonal mean PNA index as a function of target season lead times of zero month (top) and one month (bottom).

The forecast skill of the NAO index is shown in Figure 23. Relatively high NAO skill is also observed for the winter and spring target seasons. For the forecasts of most target seasons, the CanSIPSv2 system outperforms CanSIPS, especially for the forecasts of one-month lead.

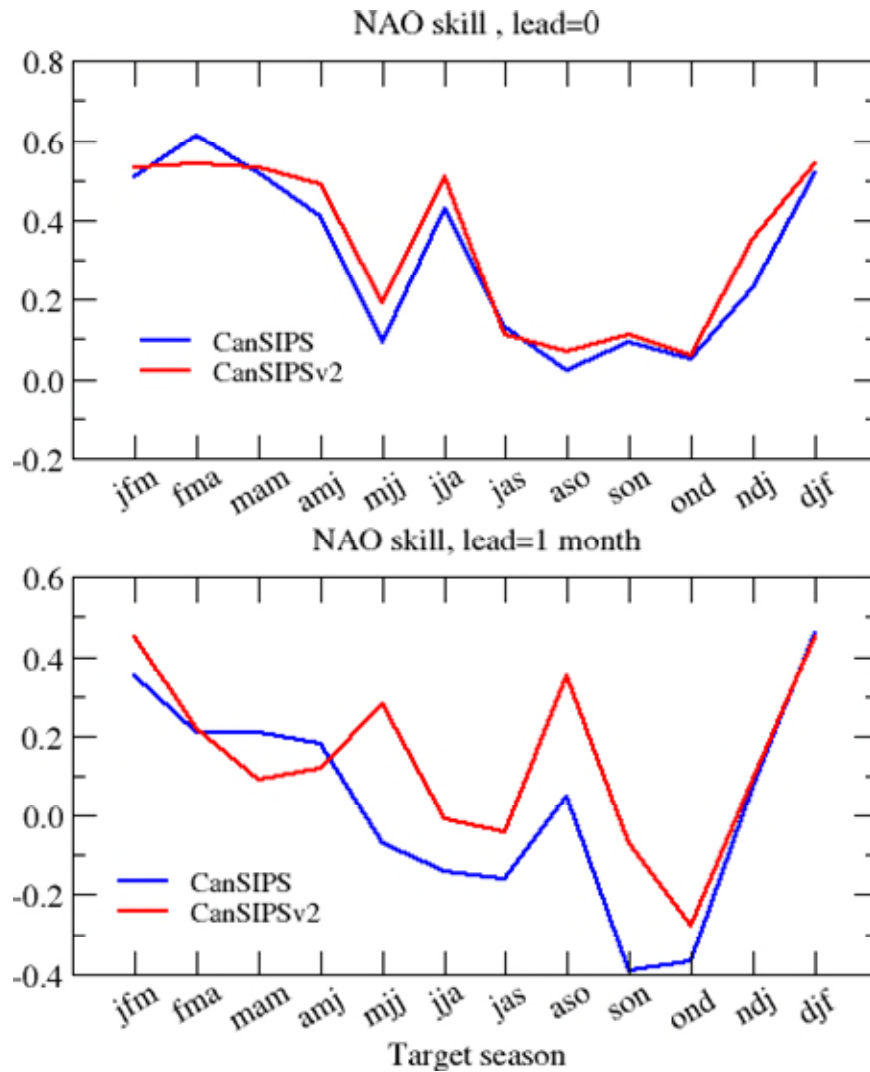


Figure 23. Same as Figure 22, but for the NAO index.

## 6.6 Sea ice

CanSIPS was one of the first seasonal prediction systems to have an interactive sea ice model component that in principle enables seasonal predictions of sea ice. In practice CanSIPS Arctic sea ice prediction skill, though appreciable, is attributable mainly to the long-term trend which however tends to be underrepresented in the reforecasts (Sigmond et al., 2013). This suggests considerable room for improvement, and in fact deficiencies in CanSIPS initialization of sea ice have been identified that likely degrade skill. These include unrealistic Arctic sea ice extent trends in the data product used to initialize sea ice concentration in the reforecasts, as well as the use of a seasonally varying model-based climatology having no long-term thinning trend to initialize sea ice thickness.

A further deficiency of sea ice initialization in the CanSIPS reforecasts is that the sea ice concentration (SIC) product that was used is biased low compared to the GDPS analysis, especially during the summer melt season. This contributed to a notable high bias in real time CanSIPS forecasts of Arctic sea ice extent, which were not useful largely for that reason.

The modifications to the initialization of sea ice in the CanCM4i forecasts and reforecasts described in sections 3.1 and 4.1 are intended to address the above issues and to provide higher forecast skill for sea ice in CanCM4i compared to CanCM4, and to reduce the large biases and resulting errors in real time forecasts. As an example, Figure 24 compares anomaly correlation reforecast skills for prediction of September SIC at four-month lead (initialized at the beginning of May) in CanCM4 and CanCM4i.

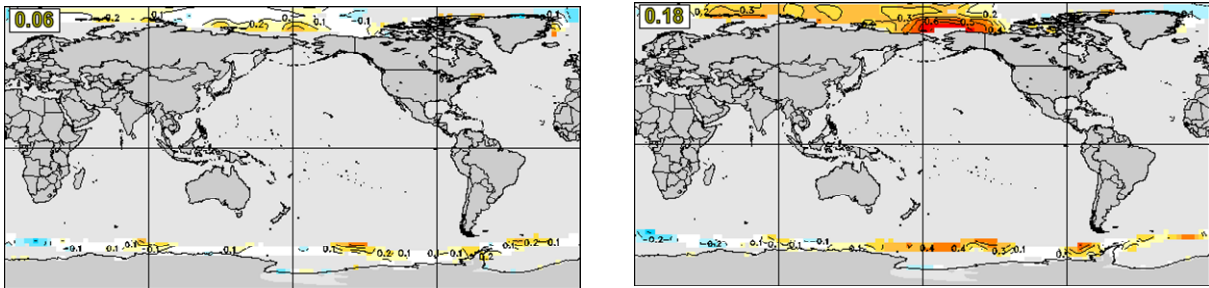


Figure 24. Anomaly correlation reforecast skill of September monthly mean SIC at four-month lead (initialized at the beginning of May), verified against Had2CIS. Left: CanCM4 in CanSIPS, Right: CanCM4i in CanSIPSv2. The globally averaged skill over ice covered regions is marked in the upper-left corner of each panel.

Further improvement in sea ice skill in CanSIPSv2 is provided by GEM-NEMO, which is even more skillful than CanCM4i (e.g. global mean anomaly correlation over ice-covered regions for September SIC at four-month lead near 0.30). This overall improvement is illustrated in Figure 25, which compares anomaly correlation reforecast skills for prediction of September SIC at four-month lead in CanSIPS and CanSIPSv2.

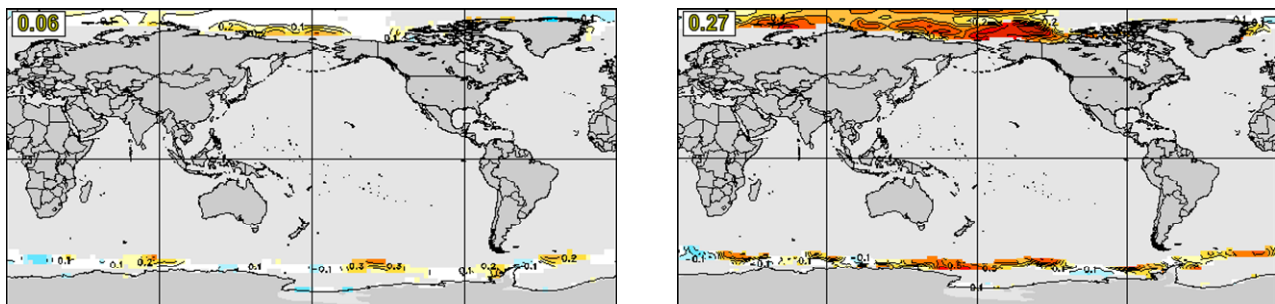


Figure 25. Same as

*Figure 24, but comparing skill for CanSIPS (left) and CanSIPSv2 (right).*

## 7 Qualitative evaluation of the parallel run

The parallel run of May and June 2019 produced consistent seasonal forecasts with the operational CanSIPS.

## 8 Details of implementation

An important aspect of any operational seasonal prediction system is that the same model versions and initialization methods are used for producing the forecasts and reforecasts, and both CanSIPSv2 models fulfill this requirement. However, because the MSC analyses employed for initializing the forecasts in real time are not available for the reforecast period, there are inevitably some differences between the data products used for initializing the forecasts and reforecasts, as described in sections 3 and 4. In general, the choice of data products for initializing the reforecasts is guided by considerations of compatibility with the corresponding analyses used in real time. For example, the ORAP5 ocean reanalysis used by both CanSIPSv2 models is formulated using the same modelling system (NEMO) and basic grid (ORCA025) as the GIOPS ocean analysis used for initialization in real time. Also, the Had2CIS sea ice concentration product was developed and chosen for initialization of the reforecasts for both models because it exhibits lower biases relative to the GPDS analysis employed in real time than other products such as ORAP5, which is substantially biased toward lower concentrations than the GPDS product especially in the summer melt season.

An additional aspect of operational CanSIPSv2 that differs from the reforecasts is that, whereas reforecasts are initialized once per month at 00Z on the first day of each month, operational forecasts are initialized and run every day according to the following schedule, with initialization at 00Z in each case:

- 15<sup>th</sup> day of month: forecast through the remainder of the current month plus six additional months (i.e. six-month forecast with 0.5 month lead), which serves as a mid-month “preview” forecast
- 4<sup>th</sup> to last day of month: 12-month forecast with 4-day lead, which serves as a backup in the event of an issue with the official forecast
- last day of month: 12-month forecast with 1-day lead, which serves as the official forecast
- every other day: forecast through the remainder of the current month plus one additional month, to track system functioning

Note in particular the 1-day lead time for the official forecasts (as compared to 0-day lead for the reforecasts), which ensures their timely delivery by the start of each month but is unlikely to have any measurable impact on seasonal forecast skill.

Probabilistic official forecast products are provided at

- [https://weather.gc.ca/saisons/index\\_e.html](https://weather.gc.ca/saisons/index_e.html)
- <http://climate-scenarios.canada.ca/?page=seasonal-forecasts>

As for CanSIPS, these calibrated probabilistic forecasts are post-processed using the methodology described in Kharin et al. (2017). This improves the reliability of the forecasts, meaning that the forecast probabilities provide a fair measure of the chance of the forecast conditions actually occurring, rather than being over- or under-estimated, as determined from the reforecasts.

## 9 Summary

The main aspects of the CanSIPSv2 system are summarized as follows.

- Like CanSIPS, CanSIPSv2 employs two global atmosphere-ocean-sea ice coupled models to produce multi-model ensemble seasonal predictions. One of the climate models in CanSIPS, CanCM3, was replaced by GEM-NEMO which is an NWP-based coupled model. The remaining climate model in CanSIPS, CanCM4, was upgraded to CanCM4i with improved initialization.
- CanSIPSv2 outperforms CanSIPS in forecasting surface air temperature and 500 hPa geopotential height in all regions for all lead times.
- Although the forecast skill of precipitation is low, CanSIPSv2 has better skill for seasonal mean precipitation forecast at zero-month lead time than CanSIPS.
- The SST and ENSO skill is also improved in CanSIPSv2 comparing to CanSIPS, especially for the Northern Hemisphere winter seasons.
- The NAO skill in CanSIPSv2 is in general higher than CanSIPS. The PNA skill is equivalent in these two systems with slight degradation of skill in winter.
- Significant improvement of sea ice forecast quality is achieved in CanSIPSv2 comparing to the previous system.

## 10 Acknowledgements

Members of the Seasonal Forum and many colleagues at RPN, CCCma and CMC contributed to the development and implementation of CanSIPSv2. In particular, we would like to thank the following colleagues for their various contributions to this project:

Frédéric Dupont, François Roy, Jean- François Lemieux, Jean-Marc Bélanger, Normand Gagnon, Peter Houtekamer, Ron McTaggart-Cowan, Ayrton Zadra, Paul Vaillancourt, Michel Roch, Stéphane Chamberland, Vivian Lee, Michel Desgagné, Stéphane Bélair, Maria Abrahamowicz, Marco Carrera, Nicola Gasset, Katja Winger.

## 11 References

- Barnston, A. G., and R. E. Livezey, 1987: Classification, seasonality and persistence of low-frequency atmospheric circulation patterns, *Mon. Wea. Rev.*, **82**, 1083-1126.
- Bélair, S., J. Mailhot, C. Girard, and P. Vaillancourt, 2005: Boundary layer and shallow cumulus clouds in a medium-range forecast of a large-scale winter system. *Mon. Wea. Rev.*, **133**, 1938-1960.
- Bernier, N. B., and S. Bélair, 2012: High horizontal and vertical resolution limited-area model: Near-surface and wind energy forecast applications, *J. Appl. Meteorol. Climatol.*, **51**, 1061–1078.
- Bradley, A. A., and S. S. Schwartz, 2011: Summary verification measures and their interpretation for ensemble forecasts, *Mon. Wea. Rev.*, **139**, 3075-3089.
- Carrera, M. L., S. Bélair, V. Fortin, B. Bilodeau, D. Charpentier, and I. Doré (2010), Evaluation of snowpack simulations over the Canadian Rockies with an experimental hydrometeorological modeling system, *J. Hydrometeorol.*, **11**, 1123–1140. Charron, M., R. Frenette and N. Gagnon, 2011: First Operational Implementation of the Regional Ensemble Prediction System at CMC (REPS 1.0.0).
- Côté, J., S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. Staniforth, 1998: The operational CMC-MRB Global Environmental Multiscale (GEM) model. Part I: Design considerations and formulation. *Mon. Wea. Rev.*, **126**, 1373–1395.
- Dee, D. P., and Co-authors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, **137**, 553-597.
- Derome, J., Brunet, G., Plante, A., Gagnon, N., Boer, G. J., Zwiers, F. W., Lambert, S. and Ritchie, H. 2001. Seasonal predictions based on two dynamical models. *Atmosphere-Ocean*, **39**: 485–501.
- Dirkson, A., W. J. Merryfield and A. H. Monahan, 2017: Impacts of sea ice thickness initialization on seasonal Arctic sea ice predictions. *J. Climate*, **30**, 1001-1017, doi:10.1175/JCLI-D-16-0437.1.
- Gagnon, N., and Co-authors, 2015: Improvements to the Global Ensemble Prediction System (GEPS) from version 4.0.1 to version 4.1.1. Canadian Meteorological Centre Technical Note. [Available on request from Environment Canada, Centre Météorologique Canadien,

division du développement, 2121 route Transcanadienne, 4e étage, Dorval, Québec, H9P1J3  
or via the following web site :

[http://collaboration.cmc.ec.gc.ca/cmc/cmoe/product\\_guide/docs/lib/technote\\_geps-411\\_20151215\\_e.pdf](http://collaboration.cmc.ec.gc.ca/cmc/cmoe/product_guide/docs/lib/technote_geps-411_20151215_e.pdf)

- Gauthier, P., M. Buehner, and L. Fillion, 1999: Background-error statistics modelling in a 3D variational data assimilation scheme: Estimation and impact on the analyses. Proc. ECMWF Workshop on Diagnosis of Data Assimilation Systems, Reading, United Kingdom, ECMWF, 131–145.
- Hoskins, B. J., and T. Ambrizzi, 1993: Rossby wave propagation on a realistic longitudinally varying flow. *J. Atmos. Sci.*, **50**, 1661-1671.
- Houtekamer, P.L., Mitchell H. L. and Deng X. 2009: Model Error Representation in an Operational Ensemble Kalman Filter, *Mon. Wea. Rev.*, **137**, 2126-2143.
- Hunke, E.C. and Lipscomb, W.H.. 2010. CICE: the Los Alamos sea ice model, documentation and software user's manual, Version 4.1. Los Alamos, NM, Los Alamos National Laboratory. (Tech. Rep. LA-CC-06-012.)
- Huffman, G. J., R. F. Adler, D. T. Bolvin, and G. Gu, 2009: Improving the global precipitation record: GPCP Version 2.1, *Geophys. Res. Lett.*, **36**, L17808, doi:10.1029/2009GL040000.
- Ioannidou, L., W. Yu, and S. Bélair, 2014: Forecasting of surface winds over Eastern Canada using the Canadian offline land surface modeling system, *J. Appl. Meteorol. Climatol.*, **53**, 1760–1774, doi:10.1175/JAMC-D-12-0284.1.
- Kain, J. S., and J. M. Fritsch, 1990: A one-dimensional entraining detraining plume model and its application in convective parameterization. *J. Atmos. Sci.*, **47**, 2784-2802.
- Kim, G., and Coauthors, 2016: Global and regional skill of the seasonal predictions by WMO Lead Centre for Long-Range Forecast Multi-Model Ensemble. *Int. J. Climatol.*, **36**, 1657–1675, <https://doi.org/10.1002/joc.4449>.
- Kharin, V. V., Q. Teng, F. W. Zwiers, G. J. Boer, J. Derome, J. S. Fontecilla, 2009 : Skill assessment of seasonal hindcasts from the Canadian Historical Forecast Project. *Atmos. Ocean.*, **47**, 204-223.
- Kharin, V. V., W. J. Merryfield, G. J. Boer, W.-S. Lee, 2017: A postprocessing method for seasonal forecasts using temporally and spatially smoothed statistics. *Mon. Wea. Rev.*, **145**, 3545-3561, doi:10.1175/MWR-D-16-0337.1.
- Kirtman, B. P., and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc.*, **95**, 585–601, doi:<https://doi.org/10.1175/BAMS-D-12-00050.1>.
- Lin, H., N. Gagnon, S. Beauregard, R. Muncaster, M. Markovic, B. Denis, and M. Charron, 2016: GEPS based Monthly Prediction at the Canadian Meteorological Centre. *Mon. Wea. Rev.*, DOI: <http://dx.doi.org/10.1175/MWR-D-16-0138.1>.
- Lin, J.-L., 2007: The double-ITCZ problem in IPCCAR4 coupled GCMs: ocean-atmosphere feedback analysis. *J. Climate*, **20**, 4497-4525.



- McFarlane, N.A., G.J. Boer, J.-P. Blanchet, and M. Lazare, 1992: The Canadian Climate Centre second-generation general circulation model and its equilibrium climate. *J. Climate*, **5**, 1013-1044.
- Merryfield, W. J., W.-S. Lee, G. J. Boer, V. V. Kharin, J. F. Scinocca, G. M. Flato, R. S. Ajayamohan, J. C. Fyfe, Y. Tang, and S. Polavarapu, 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.
- Min, Y.-M., V. N. Kryjov, and S. M. Oh, 2014: Assessment of APCC multimodel ensemble prediction in seasonal climate forecasting: Retrospective (1983–2003) and real-time forecasts (2008–2013), *JGR Atmosphere*, **119**, 12132-12150.
- Noilhan, J. and S. Planton, 1989: A Simple Parameterization of Land Surface Processes for Meteorological Models. *Mon. Wea. Rev.*, **117**, 536–549.
- Noilhan, J., and J. F. Mahfouf, 1996: The ISBA land surface parameterisation scheme, *Global Planet. Change*, **13**, 145–159.
- Reynolds, R.W., N.A. Rayner, T.M. Smith, D.C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate*, **15**, 1609-1625.
- Ritchie, H., 1991: Application of the semi-Lagrangian method to a multilevel spectral primitive-equations model. *Quart. J. Roy. Meteor. Soc.*, **117**, 91-106.
- Scinocca, J.F., N.A. McFarlane, M. Lazare, J. Li, 2008: The CCCma third generation AGCM and its Extension into the Middle Atmosphere. *Atmospheric Chemistry and Physics*, **8**, 7055-7074.
- Separovic, L., S. Z. Husain, W. Yu, and D. Fernig (2014), High-resolution surface analysis for extended-range downscaling with limited-area atmospheric models, *J. Geophys. Res. Atmos.*, **119**, 13,651–13,682, doi:10.1002/2014JD022387.
- Shabbar, A. and A. G. Barnston, 1996: Skill of seasonal climate forecasts in Canada using canonical correlation analysis. *Mon. Wea. Rev.*, **124**, 2370-2385.
- Sigmond, M., J.C. Fyfe, G.M. Flato, V.V. Kharin and W. J. Merryfield, 2013: Seasonal forecast skill of Arctic sea ice area in a dynamical forecast system, *Geophysical Research Letters*, **40**, 529-534, doi:10.1002/grl.50129.
- Smith, G.C., F. Roy, M. Reszka, D. Surcel Colan, Z. He, D. Deacu, J.-M. Belanger, S. Skachko, Y. Liu, F. Dupont, J.-F. Lemieux, C. Beaudoin, B. Tranchant, M. Drévillon, G. Garric, C.-E. Testut, J.-M. Lellouche, P. Pellerin, H. Ritchie, Y. Lu, F. Davidson, M. Buehner, M. Lajoie and A. Caya, 2016: Sea ice Forecast Verification in the Canadian Global Ice Ocean Prediction System. *Quart. J. Roy. Met. Soc.*, **142**, 659–671, doi: 10.1002/qj.2555.
- Titchner, H. A., and N. A. Rayner (2014), The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2: 1. Sea ice concentrations, *J. Geophys. Res. Atmos.*, **119**, 2864-2889, doi: 10.1002/2013JD020316.
- Wallace, J. M. and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.*, **109**, 784-812.

Zuo, H., Balmaseda, M. A. and Mogensen, K., 2015. The new eddy-permitting ORAP5 ocean reanalysis: description, evaluation and uncertainties in climate signals. *Climate Dynamics*, 10.1007/s00382-015-2675-1.

## Appendix A List of Acronyms

CanCM3	CCCma Coupled Climate Model, versions 3
CanCM4	CCCma Coupled Climate Model, versions 4
CanCM4i	CCCma Coupled Climate Model, versions 4, with improved initialization
CanSIPS	Canadian Seasonal to Interannual Prediction System
CanSIPsv2	CanSIPS version 2
CCA	Canonical Correlation Analysis
CCCma	Canadian Centre for Climate Modeling and Analysis
CIS	Canadian Ice Service
CICE	Community of Ice CodE
CMC	Canadian Meteorological Centre
CONCEPTS	The Canadian Operational Network of Coupled Environmental Prediction Systems
CRPSS	Continuous Rank probability Skill Score
ECCC	Environment and Climate Change Canada
EnKF	Ensemble Kalman Filter
ENSO	El Niño/Southern Oscillation
GDPS	Global Deterministic Prediction System
GEM	Global Environmental Multiscale model
GEM-NEMO	GEM and NEMO coupled model
GEPS	Global Ensemble Prediction System
GIOPS	Global Ice Ocean Prediction System
GOSSIP	Globally Organized System for Simulation Information Passing
GPCP	Global Precipitation Climatology Project
Had2CIS	HadISST2 combined with CIS digitized sea ice charts
HadISST2	Hadley Centre Sea Ice and Sea Surface Temperature Version 2
HFP	Historical Forecast Project
NEMO	Nucleus for European Modelling of the Ocean
MSC	Meteorological Service of Canada (of Environment Canada)
NAO	North Atlantic Oscillation
NCAR	National Center for Atmospheric Research
NCEP	(United States) National Centers for Environmental Prediction
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
PNA	Pacific/North American pattern
OISST	Optimum Interpolation Sea Surface Temperature of NOAA
ORAP5	Ocean Reanalyses Prototype 5
ORAS5	Ocean Reanalysis System 5
RPN	Recherche en Prévision Numérique
REOF	Rotated Empirical Orthogonal Function
SIC	Sea Ice Concentration
SPS	Surface Prediction System

SST	Sea Surface Temperature
T2m	2-meter air temperature
Z500	500-hPa geopotential height