

The Canadian Seasonal to Interannual Prediction System. Part I: Models and Initialization

WILLIAM J. MERRYFIELD, WOO-SUNG LEE, GEORGE J. BOER, VIATCHESLAV V. KHARIN,
JOHN F. SCINOCCA, GREGORY M. FLATO, R. S. AJAYAMOHAN, AND JOHN C. FYFE

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

YOUNMIN TANG

*Environmental Science and Engineering, University of Northern British Columbia, Prince George, British
Columbia, Canada*

SAROJA POLAVARAPU

Environment Canada, Downsview, Ontario, Canada

(Manuscript received 31 July 2012, in final form 9 January 2013)

ABSTRACT

The Canadian Seasonal to Interannual Prediction System (CanSIPS) became operational at Environment Canada's Canadian Meteorological Centre (CMC) in December 2011, replacing CMC's previous two-tier system. CanSIPS is a two-model forecasting system that combines ensemble forecasts from the Canadian Centre for Climate Modeling and Analysis (CCCma) Coupled Climate Model, versions 3 and 4 (CanCM3 and CanCM4, respectively). Mean climate as well as climate trends and variability in these models are evaluated in freely running historical simulations. Initial conditions for CanSIPS forecasts are obtained from an ensemble of coupled assimilation runs. These runs assimilate gridded atmospheric analyses by means of a procedure that resembles the incremental analysis update technique, but introduces only a fraction of the analysis increment in order that differences between ensemble members reflect the magnitude of observational uncertainties. The land surface is initialized through its response to the assimilative meteorology, whereas sea ice concentration and sea surface temperature are relaxed toward gridded observational values. The subsurface ocean is initialized through surface forcing provided by the assimilation run, together with an offline variational assimilation of gridded observational temperatures followed by an adjustment of the salinity field to preserve static stability. The performance of CanSIPS historical forecasts initialized every month over the period 1981–2010 is documented in a companion paper. The CanCM4 model and the initialization procedures developed for CanSIPS have been employed as well for decadal forecasts, including those contributing to phase 5 of the Coupled Model Intercomparison Project.

1. Introduction

The potential economic value of monthly, seasonal, and longer-range forecasts having even modest skill has motivated attempts to produce and improve such forecasts for more than half a century (Namias 1968; Livezey

1990). During this period the methods employed to produce these forecasts have evolved considerably. Earlier efforts relied on a combination of objective and subjective, experience-based guidance founded on empirical measures of persistence and available data on surface anomalies (Namias 1964; Wagner 1989). However, as computing power increased and observation-based datasets improved, fully objective methods became the norm.

Two types of objective methods currently in use are statistical forecasts, which employ empirically derived

Corresponding author address: William Merryfield, Canadian Centre for Climate Modelling and Analysis, University of Victoria, P.O. Box 1700 STN CSC, Victoria, BC V8W 2Y2, Canada.
E-mail: bill.merryfield@ec.gc.ca

relations based on past behavior, and dynamical forecasts, in which the equations describing the climate system are integrated in time from initial conditions constrained by observations. Until relatively recently dynamical atmospheric forecasts typically employed a “two tier” methodology whereby offline predictions of future anomalies of sea surface temperature (SST) serve as boundary conditions for an atmospheric general circulation model (AGCM). Increasingly, however, long-range forecasts use a “one tier” approach based on a coupled global climate model (CGCM). In this case, the interactive ocean component of the model provides future SSTs as an integral part of the forecast. Because of the fundamentally probabilistic nature of long-range predictions, dynamical forecasts usually are based on ensembles of predictions, from which probabilistic information may be derived. It has furthermore been shown in numerous studies (e.g., Kharin et al. 2009) that multimodel ensembles tend to offer better performance than single-model ensembles for a given ensemble size.

The evolution of operational seasonal forecasting in Canada has reflected these developments. Beginning in the mid-1990s, Environment Canada’s Canadian Meteorological Centre (CMC) produced one-season forecasts using a two-tier, multimodel dynamical forecasting system in which future SST anomalies were specified by persisting their mean values from the month preceding the forecast. Initially this system employed six forecasts from each of two AGCMs, with bias correction and skill assessments obtained from a set of retrospective forecasts comprising the Historical Forecasting Project (HFP; Derome et al. 2001). In 2007 this system was upgraded to 10 forecasts from each of four AGCMs, with corresponding retrospective forecasts provided by the second phase of the HFP, or HFP2 (Kharin et al. 2009).

Although SST anomaly persistence provides a reasonable forecast of SST at relatively short range, this ceases to be the case beyond a few months when the forecast range approaches the decorrelation time scale for SST anomalies (Goddard and Mason 2002). For this reason, the HFP2 two-tier dynamical forecasts were restricted to a range of four months, providing seasonal (three month) mean surface temperature and precipitation at zero- and one-month lead, as well as first-month forecasts for surface temperature. Longer leads necessitated a different approach, so seasonal forecasts of surface temperature within Canada at leads of three, six, and nine months were obtained statistically through canonical correlation analysis (CCA) as described in Shabbar and Barnston (1996).

A one-tier system eliminates the need to apply different forecast methods at different leads as in the

hybrid two-tier/statistical system described above, and has the potential also to improve the quality of the forecasts. The development of a one-tier multiseasonal forecasting system suitable for operational use in Canada began with a pilot project, the first phase of the Coupled Historical Forecast Project (CHFP1; Merryfield et al. 2010), that adapted the Canadian Centre for Climate Modeling and Analysis (CCCma) CGCM to seasonal forecasting by applying a simple initialization method to the then-current CGCM3.1(T63) model version. This approach, similar to that of Keenlyside et al. (2005), relaxed model SSTs to observation-based time series during a multiyear period preceding each forecast. Although far from optimal, this method did initialize the crucial equatorial Pacific region with some skill, leading to reasonably skillful El Niño–Southern Oscillation (ENSO) forecasts and global skills competitive with those of the HFP2 system despite the smaller ensemble size (one model with 10 ensemble members vs four models with 10 ensemble members each). In addition, the skill of CHFP1 SST forecasts out to 12 months considerably exceeded that of the persistence assumption employed by the HFP2 two-tier forecasts.

CHFP1, although modest in scope and leaving considerable room for improvement, served to develop computational infrastructure and associated verification tools in parallel with efforts to improve the model and its initialization. In addition, CHFP1 provided a baseline against which the skill of subsequent versions of the forecast system could be compared.

The remainder of the paper describes the second phase of the Coupled Historical Forecast Project (CHFP2), which led to the development of the Canadian Seasonal to Interannual Prediction System (CanSIPS). This second-generation system considerably improved on the first-generation CHFP1 system through extensive model development, its use of two model versions, CCCma Coupled Climate Model, versions 3 and 4 (CanCM3 and CanCM4, respectively), and the implementation of a far more comprehensive initialization procedure. It was designed as an operation-ready system, and replaced Canada’s previous operational seasonal prediction system, which was based two-tier dynamical and longer-range statistical forecasts as detailed above, in late 2011. Sections 2 and 3 describe the formulation and properties of CanCM3 and CanCM4, which are used mainly for climate forecasting and have not been described in detail elsewhere. Section 4 focuses on the procedures used to initialize these models, and a summary and discussion are provided in section 5. A companion paper will describe in detail the performance of the CanSIPS historical forecasts.

TABLE 1. CCCma climate model configurations used for CHFP2/CanSIPS in relation to other applications (IPCC AR4, CHFP1, and IPCC AR5).

Application	Coupled model	Atmosphere	Ocean	Flux adjustment	Carbon
IPCC AR4	CGCM3.1(T63)	CanAM3	CanOM3	Yes	No
CHFP1	CGCM3.1(T63)	CanAM3	CanOM3	No	No
CHFP2/CanSIPS	CanCM3	CanAM3	CanOM4	No	No
	CanCM4	CanAM4	CanOM4	No	No
IPCC AR5 short term	CanCM4	CanAM4	CanOM4	No	No
IPCC AR5 long term	CanESM2	CanAM4	CanOM4	No	Yes

2. CanCM3 and CanCM4 climate models

CanSIPS combines ensemble forecasts from two versions of the CCCma climate model, CanCM3 and CanCM4, in order to take advantage of the generally greater skill of multimodel ensembles for a given ensemble size (e.g., Kirtman and Min 2009; Kharin et al. 2009). These two models share a common ocean component, CCCma's fourth-generation ocean model (CanOM4), but differ in their atmospheric components: CanCM3 uses CCCma's third-generation atmospheric general circulation model (CanAM3; also known as AGCM3), whereas CanCM4 employs the fourth-generation version (CanAM4). By comparison, the CHFP1 pilot project described in the previous section used CanAM3 with an earlier ocean model version CanOM3; this was essentially the CGCM3.1(T63) climate model that contributed to the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC AR4), except that flux adjustments were not used in the hindcasts. These model configurations are summarized in Table 1.

a. Model components

The CanAM3 atmospheric component is described in detail by McFarlane et al. (2005) and Scinocca et al. (2008). Briefly, it is a spectral model which uses a T63 truncation and 31 hybrid (sigma pressure) vertical coordinate levels extending from the surface to 1 hPa. Physical tendencies are computed on the linear transform grid yielding a horizontal grid spacing of approximately 2.8°. Land surface processes are represented by version 3 of the Canadian Land Surface Scheme (CLASS; Verseghy 2000), and sea ice dynamics are treated according to the cavitating fluid approach of Flato and Hibler (1992).

CanAM4 operates with the same horizontal resolution and upper boundary as CanAM3, but has 35 hybrid vertical coordinate levels providing more uniform resolution across the tropopause. Upgraded physical parameterizations in CanAM4 include a correlated-k distribution radiative transfer scheme (Li and Barker 2005), a more general treatment of radiative transfer in the presence of clouds using the Monte Carlo independent column approximation (Barker et al. 2008),

a prognostic bulk aerosol scheme with a full sulphur cycle, along with organic and black carbon, mineral dust, and sea salt (Lohmann et al. 1999; Croft et al. 2005), a fully prognostic single-moment cloud microphysics scheme (Lohmann and Roeckner 1996; Rotstayn 1997; Khairoutdinov and Kogan 2000), and new shallow convection scheme (von Salzen et al. 2005). Land surface and sea ice are treated as in CanAM3.

The ocean component CanOM4, common to CanCM3 and CanCM4, differs in several important respects from its predecessor CanOM3 used for IPCC AR4 and CHFP1. Vertical resolution is increased from 29 to 40 levels with vertical spacings ranging from 10 m near the ocean surface to greater than 300 m at abyssal depths. Subsurface heating due to penetration of shortwave radiation beneath the first level is represented as in Zaharieva et al. (2008), with chlorophyll concentrations specified according to daily climatological means derived from the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) satellite data for 1998–2002 (Yoder and Kennelly 2003). Diapycnal mixing in the surface mixed layer and ocean interior are represented according to the *K*-profile parameterization of Large et al. (1994), together with a version of the tidal mixing parameterization of Simmons et al. (2004). Horizontal friction is specified according to the anisotropic viscosity formulation of Large et al. (2001), whereas eddy-induced transport and along-isopycnal diffusion are as in Gent et al. (1995).

b. Coupling

CanCM3 and CanCM4 couple their atmosphere and ocean components once per day. In both models, the ocean component receives daily mean surface heat, freshwater, and momentum fluxes computed by the atmospheric component, and after stepping forward by one day passes updated daily mean SST values back to the atmosphere. Ocean surface velocities are not taken into account in computing the surface momentum flux. The surface beneath each atmospheric cell is entirely ocean or land, with precisely six ocean grid cells under each atmospheric grid cell. While this configuration simplifies model physics and the interpolation of coupling

fields, it enforces the relatively low atmospheric model resolution on the structure of coastlines. In contrast to previous CCCma climate model versions, no flux adjustments are applied in either CanCM3 or CanCM4 (Table 1).

c. Radiative forcing

CanCM3 and CanCM4 both represent anthropogenic influences on radiative forcing, inclusion of which has been shown to improve seasonal forecast skill in predicting global mean and regional temperatures and temperature trends (Doblas-Reyes et al. 2006; Liniger et al. 2007). In CanCM3 this is implemented through equivalent CO₂ forcing representing the effects of all greenhouse gases (GHGs), with compensating influences of anthropogenic aerosols represented crudely by imposing a 40% reduction on increases of this GHG forcing above preindustrial levels. CanCM4 employs a far more comprehensive treatment of radiative forcing. Concentrations of CO₂ and other radiatively important GHGs until 2005 are specified according to the representative concentration pathway (RCP) historical scenario developed for the IPCC Fifth Assessment Report (AR5) Coupled Model Intercomparison Project Phase 5 (CMIP5), whereas after 2005 the RCP4.5 scenario (Paolino et al. 2010) is used. Direct and indirect effects of aerosols on climate are treated through a prognostic bulk aerosol scheme with full sulphur cycle, organic and black carbon, mineral dust, and sea salt (von Salzen et al. 2005; Ma et al. 2010), with anthropogenic emissions specified through the CMIP5 historical and RCP4.5 scenarios. Solar cycle irradiance variations and volcanic stratospheric aerosols are treated according to CMIP5 recommendations. Because explosive volcanic eruptions cannot accurately be predicted, volcanic effects in the CanCM4 component of CanSIPS are represented by exponential decay of the initial volcanic stratospheric aerosols with a time scale of one year.

3. Model performance

This section examines aspects of the behavior of freely running CanCM3 and CanCM4 simulations that bear upon the suitability of these models for subseasonal to multiseasonal forecasting. In particular, the models should be able to reasonably represent (i) global trends attributable to changes in radiative forcing, (ii) climatological averages and the mean seasonal cycle, and (iii) unforced climate variability on time scales relevant to long-range forecasting. These properties of the two models are examined below.

a. Global temperature trends

The gross effect of anthropogenic forcing changes is apparent in time series of global mean temperature.

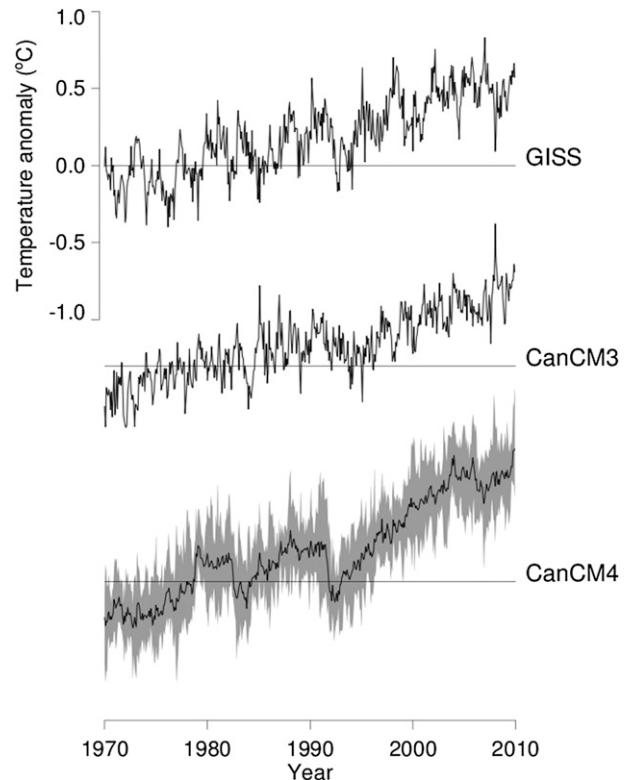


FIG. 1. Monthly global average surface temperatures for 1970–2009, expressed as anomalies relative to the 1970–90 mean. (top) GISS observational dataset; (middle) freely running CanCM3; (bottom) freely running CanCM4 (ensemble mean of 10 runs, with range of values indicated by shading). Values missing from the observations are excluded in constructing the model averages. Temperature scale is the same for each case.

Figure 1 shows such time series for the period 1970–2009 from the Goddard Institute for Space Studies (GISS) observational dataset (Hansen et al. 2010; Fig. 1, top), a freely running, historically forced CanCM3 simulation (Fig. 1, middle), and 10 historically forced CanCM4 simulations (Fig. 1, bottom), where values shown are anomalies relative to 1970–90 averages. In addition to anthropogenically forced trends, these time series exhibit signatures of unforced internal climate variability, uncorrelated between observations and different model runs, which is attributable to ENSO and other influences (Fyfe et al. 2010). The observational and CanCM4 time series furthermore show transient coolings due to the large explosive volcanic events of 1982 and 1991; these are particularly evident in the CanCM4 ensemble mean because the unforced climate variability tends to be filtered out by the ensemble averaging process. These coolings are not present in the CanCM3 time series because that model does not represent volcanic aerosol effects.

TABLE 2. Trends in 1970–2009 time series of monthly global mean surface temperature, from HadCRUT and GISS observations, one CanCM3 historical run, and an ensemble of 10 CanCM4 historical runs. Values missing from the observations are excluded in constructing the respective model trends. Uncertainties represent 95% confidence intervals.

Data	Trend ($^{\circ}\text{C}$ decade $^{-1}$)	Trend ($^{\circ}\text{C}$ decade $^{-1}$)
	HadCRUT masked	GISS masked
HadCRUT	0.15 ± 0.01	—
GISS	—	0.16 ± 0.01
CanCM3	0.14 ± 0.01	0.15 ± 0.01
CanCM4	0.21 ± 0.01	0.23 ± 0.01

To facilitate accurate comparison between the modeled trends and those represented by the GISS and Hadley Centre/Climatic Research Unit temperature (HadCRUT) (not shown) observational datasets, two further processing steps were undertaken. First, the times and grid locations at which GISS and HadCRUT data are missing due to a lack of observations were excluded in computing the modeled global means (the model time series in Fig. 1 exclude data missing from the GISS dataset). Second, the signatures of dynamically induced atmospheric variability, ENSO, and explosive volcanic eruptions were removed using procedures described in Thompson et al. (2009); trends in the residual time series were then computed as in Fyfe et al. (2010), who demonstrated that this procedure substantially increases confidence in the calculated trends. The resulting estimated modeled trends and their statistical uncertainties are compared with those obtained from the GISS and HadCRUT datasets in Table 2. Of note is that while CanCM3 has a trend comparable to that inferred from observations, CanCM4's trend significantly exceeds that estimated from both observational datasets. This tendency is also evident in CMIP5 decadal predictions using CanCM4 (Fyfe et al. 2011; Kharin et al. 2012) as well as some other models (Kim et al. 2012).

b. Mean climate and seasonal cycle

All climate models have imperfections that lead to biases in simulated climate. In a forecasting context where model initial conditions are constrained by the observed climate state at a particular time, this implies that the model will progressively drift from a comparatively realistic simulated climate at the beginning of the forecast period toward one that increasingly becomes imprinted with the biases of the unconstrained model. This drift is commonly removed by subtracting the forecast climatology, dependent on the time of year and forecast lead time and computed by averaging over a large set of historical forecasts, from the forecast fields to obtain predicted anomalies. While such a procedure

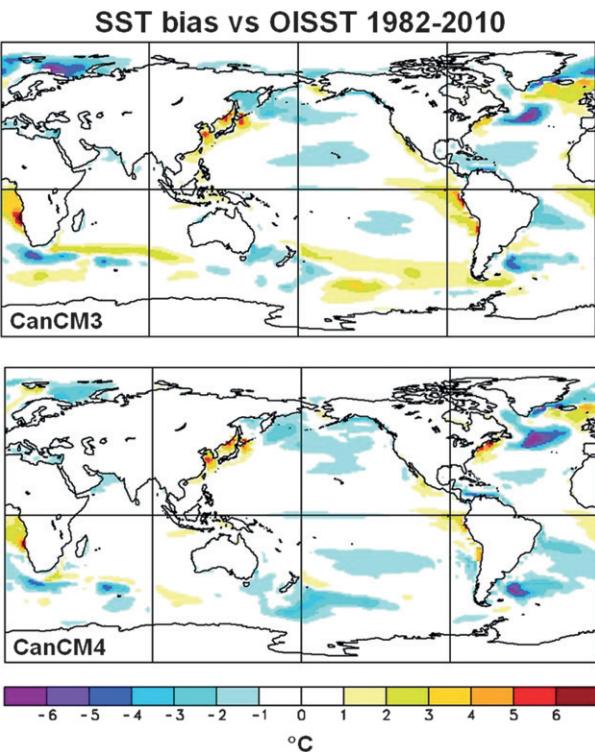


FIG. 2. Annual mean SST biases for 1982–2010 in (top) CanCM3 and (bottom) CanCM4 historical runs, relative to the OISST observational dataset for the same period.

compensates for model biases and consequent drifts, it is clearly desirable that biases be minimized, as errors in the mean state can affect the response of the atmosphere to surface forcing anomalies even when the anomalies are represented accurately (e.g., Balmaseda et al. 2010).

1) OCEAN CLIMATOLOGY

Sea surface temperature is the most crucial ocean attribute for multiseasonal forecasting because it is the main avenue by which the relatively slowly evolving ocean influences the atmosphere. Although one-tier forecasting systems have a significant advantage over their two-tier counterparts in being able to predict the dynamical evolution of future SST anomalies, they also suffer an inherent disadvantage in that the model SST climatologies contain biases, whereas such biases can be eliminated in the specified SSTs used by two-tier forecasts. Having a realistic SST climatology is thus a highly desirable property for a multiseasonal forecast model.

The annual mean SST biases for CanCM3 and CanCM4, obtained by subtracting the 1982–2010 mean of the optimum interpolation SST (OISST, version 2) observational analysis (Reynolds et al. 2002)¹ from like

¹ The first full year of the OISST analysis is 1982.

TABLE 3. Global mean SST bias and mean absolute error for the CanCM3 and CanCM4 models, based on differences between 1982–2010 means for free model runs under historical forcing and the OISST observational analysis as shown in Fig. 2.

Model	Season	Mean bias	Mean absolute error
CanCM3	Annual mean	0.00	0.88
	DJF	0.17	1.00
	MAM	0.07	0.93
	JJA	-0.07	0.99
	SON	-0.17	1.04
CanCM4	Annual mean	-0.29	0.79
	DJF	-0.25	0.85
	MAM	-0.29	0.79
	JJA	-0.25	0.85
	SON	-0.36	0.94

averages for historical model runs, are shown in Fig. 2. The annual global mean bias in CanCM3 is negligible, whereas mean SST in CanCM4 is 0.29°C cooler than observed (Table 3). Local biases are generally modest, less than $\pm 2^\circ\text{C}$ or so in most locations. Notable exceptions include a cold “bull’s-eye” in the North Atlantic, attributable to an excessively zonal North Atlantic Current (Randall et al. 2007), and strong warm biases (more severe in CanCM3 than CanCM4) off the coasts of western tropical South America and Africa that are associated with underrepresented coastal upwelling and marine stratocumulus clouds (de Szoeke et al. 2010; Zheng et al. 2011). These specific localized

strong biases are prevalent in many current-generation climate models.

Mean absolute error (MAE) provides an overall measure of model errors in SST climatology. For annual mean SST, MAE in CanCM4 is lower than in CanCM3 despite the smaller mean bias of the latter, indicating that CanCM4 tends to have smaller regional SST errors than CanCM3. The mean bias and MAE for individual seasons are not dramatically larger than for the annual mean (Table 3), indicating that the relatively small errors in annual mean SST do not arise from cancellation of much larger seasonally varying errors.

The equatorial Pacific Ocean is a particularly important region to model accurately because of its fundamental role in ENSO (Guilyardi et al. 2009). A common model bias in this region is an excessively cool upwelling or “cold tongue” region stretching across much of the equatorial Pacific (e.g., Reichler and Kim 2008). This bias is present in CanCM3 and CanCM4 although it is not severe as it does not exceed -1.2°C in CanCM3 and -1.5°C in CanCM4.

The accuracy of the modeled seasonal cycle in the equatorial Pacific is important in climate models in part because its amplitude tends to anticorrelate with the modeled level of ENSO variability (Guilyardi 2006), although this is not the case for CanCM3 versus CanCM4. Figure 3 shows the average seasonal cycle about annual mean equatorial Pacific SST for 1982–2010 according to the OISST analysis and historical runs of CanCM3 and CanCM4. The observed seasonal cycle is

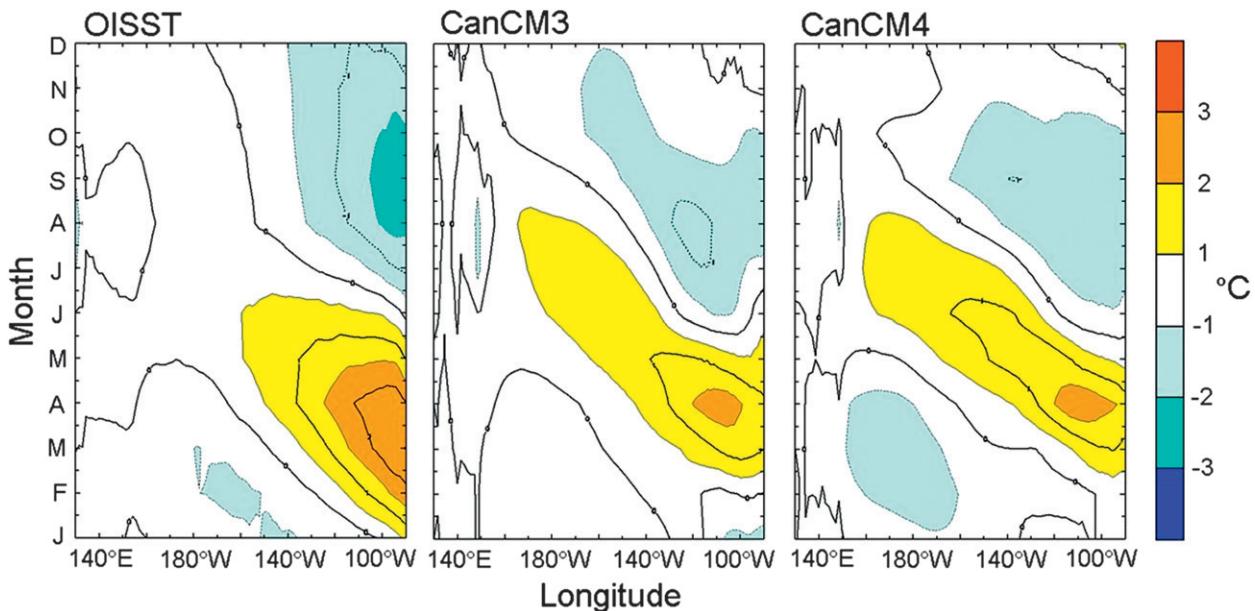


FIG. 3. Climatological mean seasonal cycle of equatorial Pacific SST relative to the annual mean, for the (left) OISST observational dataset, (middle) CanCM3, and (right) CanCM4.

dominated by the annual harmonic despite the strong semiannual component of solar radiation at the top of the atmosphere (Fu and Wang 2001). The seasonal cycle in the models is also primarily annual rather than semiannual, but is somewhat too weak and exhibits phasing errors, with a delayed spring maximum and premature autumn minimum.

The equatorial Pacific thermocline is of fundamental importance because its vertical motions strongly influence ENSO-related SST anomalies particularly in the eastern Pacific. Also, its properties such as zonal slope, which is strongly connected to the mean zonal wind stress, and mean depth strongly influence ENSO amplitude, period, and stability in simplified models of ENSO (e.g., Fedorov and Philander 2001). Physically the thermocline is defined as the depth of the maximum vertical temperature gradient, whereas the depth of the 20°C isotherm serves as a widely used practical definition that is applicable for present-day climate (Yang and Wang 2009). Figure 4 shows cross sections of 1991–2000 means of ocean temperature versus depth and longitude at the equator in the Simple Ocean Data Assimilation (SODA) 1.4.2 ocean analysis (Carton and Giese 2008), CanCM3, and CanCM4. The observed 20°C isotherm is indicated by the dashed red curve in all three panels. Its depth in the western and central Pacific is correctly represented in both models, whereas in the eastern Pacific the modeled 20°C isotherm is slightly too shallow, particularly in CanCM3. Also, the thermocline in CanCM3 tends to be too diffuse (i.e., vertical gradients near the depth of the 20°C are too weak except in the far eastern Pacific), which leads to a significant warm bias beneath the thermocline.

Currents in the equatorial Pacific are primarily zonal and are dominated by the Equatorial Undercurrent (EUC). Figure 5 shows annual mean zonal velocity as a function of depth and longitude along the equator (left-hand panels) and as a function of depth and latitude at 140°W (right-hand panels) based on 1991–2000 averages from SODA 1.4.2,² CanCM3, and CanCM4. Several biases are evident in the models. Although peak current speeds approach those observed, the maxima in the models are too deep (125–130 m in the models as compared to about 105 m observed) and too far west (140°–150°W in the models vs about 130°W observed). In addition, the westward-flowing South Equatorial Current is too strong at the equator in the models, and the eastward North Equatorial

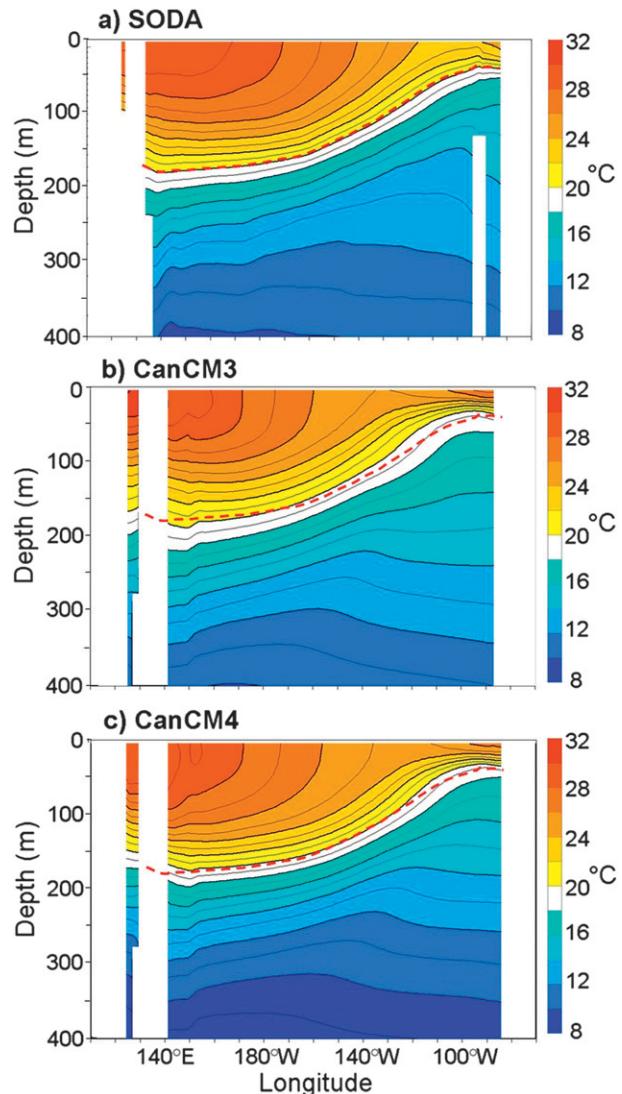


FIG. 4. Annual mean equatorial Pacific temperature as a function of longitude and depth for the period 1991–2000 from (a) the SODA 1.4.2 ocean analysis, (b) CanCM3, and (c) CanCM4. The observed 20°C isotherm is indicated by the dashed red curves.

Counter Current which peaks at about 50-m depth is too weak.

The seasonally varying depth of the ocean surface mixed layer is another important global ocean property in relation to climate variability because it determines the effective ocean heat capacity governing the response of SST to atmospheric forcing (Saravanan and Chang 1999; Yu and Boer 2006). In the extratropics, ocean mixed layer depth (MLD) tends to be deepest in late winter, and shallowest in late summer. This seasonality is illustrated in Fig. 6, where MLD has been computed from the observational Polar Science Center Hydrographic Climatology (PHC)/World Ocean Atlas (WOA)

² A similar picture is obtained from shipboard observations spanning the 1990s (Johnson et al. 2001) and from the Tropical Atmosphere Ocean (TAO) mooring data (Brodeau et al. 2010).

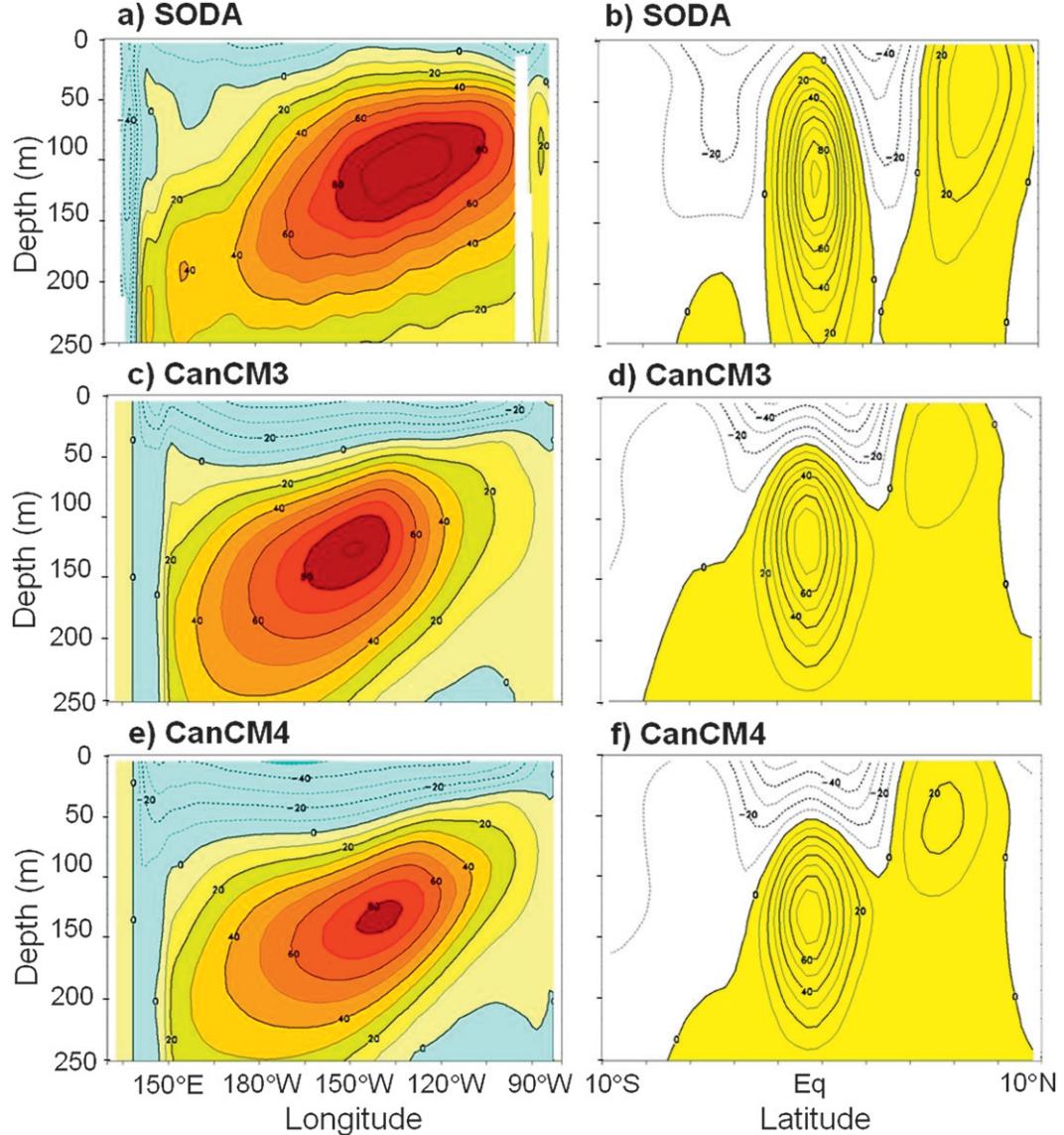


FIG. 5. Annual mean Pacific zonal ocean velocity (left) as a function of depth and longitude at the equator and (right) as a function of depth and latitude at 140°W based on 1991–2000 averages from (a),(b) the SODA 1.4.2 ocean analysis; (c),(d) CanCM3; and (e),(f) CanCM4. Contour intervals are 10 cm s^{-1} , with westward velocities shaded in (b),(d), and (f).

climatology (Steele et al. 2001) and 1991–2000 model climatologies using the algorithm of Kara et al. (2000). In the winter hemispheres, local maxima of MLD corresponding to various mode water formation regions (Talley 1999) are generally well represented, although model MLDs in the Pacific subtropical and central mode water formation regions east of Japan are somewhat too deep, a common tendency in climate models (Lienert et al. 2011). A notable exception to this general agreement is that winter MLDs in the North Atlantic are represented rather poorly in the models. For example,

the models show a zonally elongated region of shallow MLD at about 45°N that is not present in observations; this appears to be associated with the large surface cold bias in this region (Fig. 2) and coincident fresh bias (not shown), which in turn are attributable to errors in the path of the North Atlantic Current. In addition, the very deep winter mixed layers associated with deep water formation in the northwest Atlantic are displaced significantly southeastward in the models. In the summer hemispheres and in the tropics, the relatively shallow mixed layers are mainly wind driven and are represented

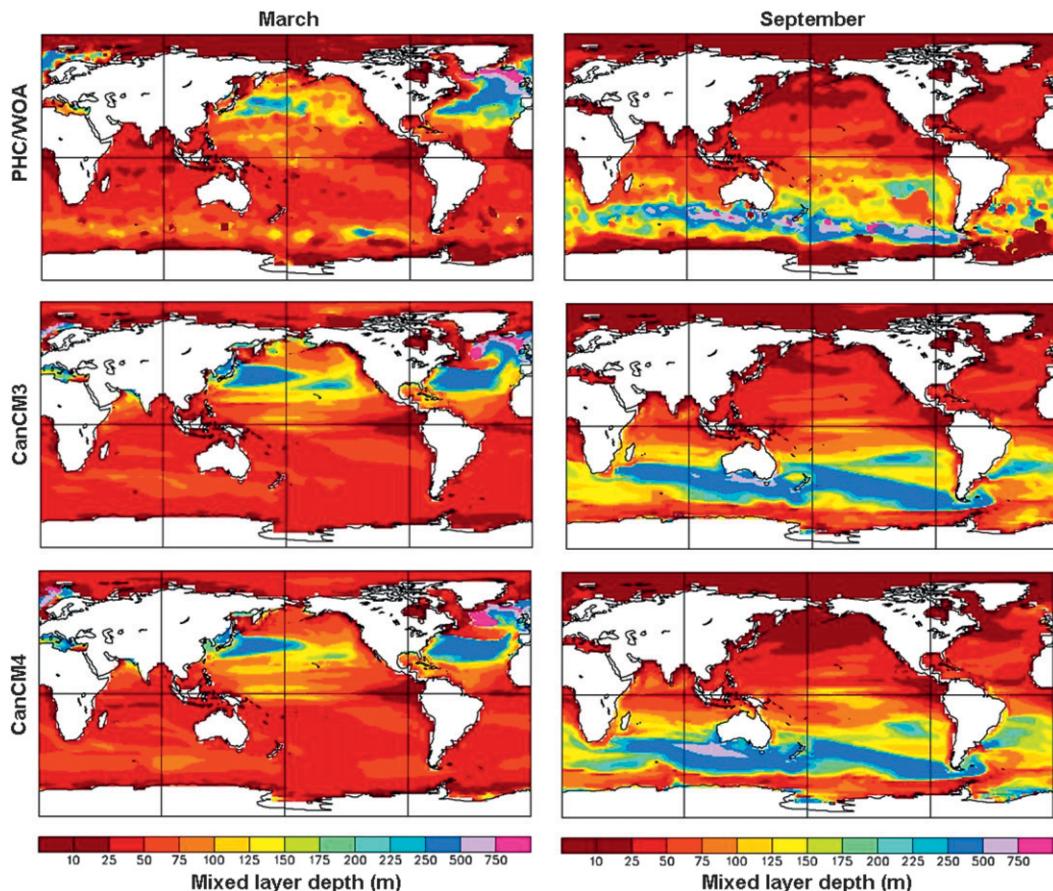


FIG. 6. Ocean mixed layer depths in (left) March and (right) September based on monthly mean temperature and salinity from the (top) PHC/WOA observational climatology and 1991–2000 model climatologies from historical runs of (middle) CanCM3 and (bottom) CanCM4, computed using the algorithm of Kara et al. (2000).

reasonably well, although austral summer MLD in the Southern Ocean at around 60°S tends to be somewhat too shallow in the models.

Volume transports associated with various aspects of ocean circulation provide further metrics for assessing the ocean component of climate models, and a standard

set of such metrics has been developed for CMIP5 (e.g., Griffies et al. 2011). Table 4 lists values for a subset of the CMIP5 ocean transports as determined from 1981–2010 averages from one historically forced CanCM3 simulation and 10 historically forced CanCM4 simulations, in comparison with available observationally based

TABLE 4. Modeled and observationally determined ocean volume transports.

Transport in Sv ^a	CanCM3 ^b	CanCM4 ^b	Observed	Reference
Atlantic meridional overturning (maximum, including eddy transport)	18.4 ± 1.2	18.0 ± 0.8	—	—
Atlantic meridional overturning near 25°N	14.6 ± 1.0	14.9 ± 0.7	18.5 ± 2^c 18.7 ± 4.8^d	Ganachaud (2003) Rayner et al. (2011)
Drake Passage	146 ± 2	159 ± 2	134 ± 15 to 27	Cunningham et al. (2003)
Indonesian Throughflow	-20.0 ± 0.8	-19.7 ± 1.2	-15.0 (-10.7 to -18.7)	Sprintall et al. (2009)
Bering Strait	0.83 ± 0.06	0.95 ± 0.06	0.83 ± 0.1^e	Roach et al. (1995)

^a Eastward and northward values are positive.

^b Means and interannual standard deviations for 1981–2010.

^c Standard deviation uncertainty of inversion estimate.

^d Standard deviation of 10-day low-pass-filtered time series.

^e Estimated interannual variability.

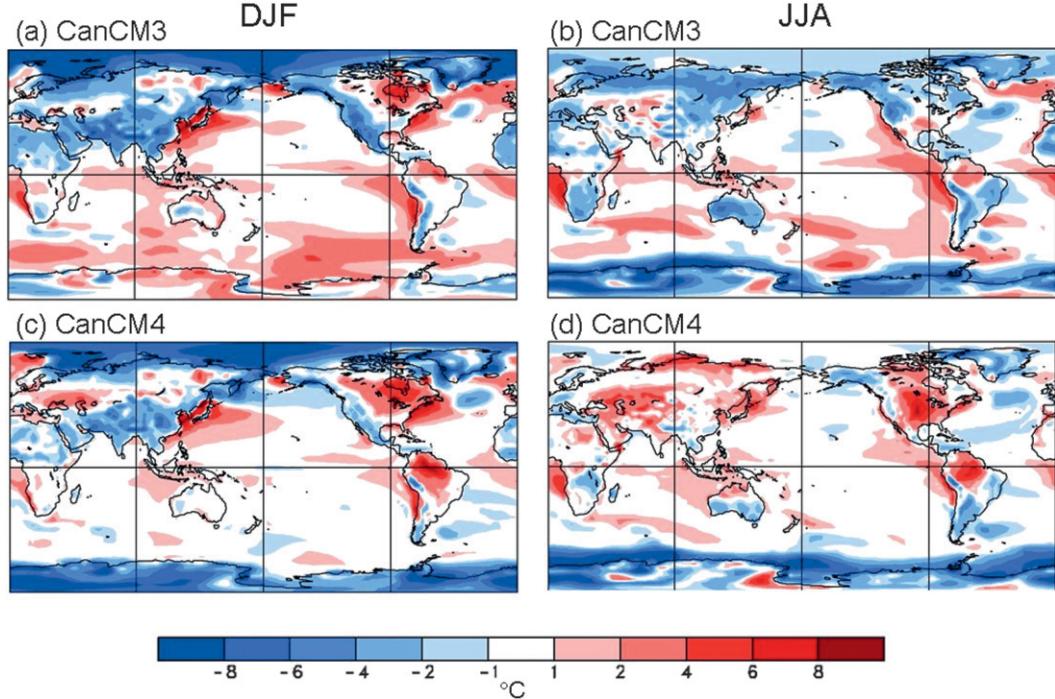


FIG. 7. Near-surface temperature biases in historically forced runs of (a),(b) CanCM3 and (c),(d) CanCM4 averaging over boreal (a),(c) winters (DJF) and (b),(d) summers (JJA) in 1981–2010. Observational values are from ERA-Interim for the same period.

values. The maximum transport associated with the North Atlantic meridional overturning circulation (AMOC), as represented by the streamfunction for zonally averaged flow in the Atlantic and including the parameterized effects of eddies, is about 18 Sv in both models ($1 \text{ Sv} = 10^6 \text{ m}^3 \text{ s}^{-1}$; no specific observation-based comparison is available for this case). A measure of the AMOC that is observationally constrained is transport due to the mean flow (i.e., excluding effective eddy transports) near 25°N. This is about 15 Sv in both models, which is about 3 Sv lower than the inverse model-based estimate of Ganachaud (2003) at 24°N, and also a 4-yr mean value from the Natural Environment Research Council (NERC) Rapid Climate Change (RAPID) monitoring array (Rayner et al. 2011). Both models have Atlantic Circumpolar Current transports through Drake Passage that are toward the high end of the range of mean values estimated by Cunningham et al. (2003). Modeled Indonesian Throughflow transports are somewhat stronger than the upper range of mean values estimated by Sprintall et al. (2009), whereas mean transports through Bering Strait are comparable to those estimated by Roach et al. (1995).

2) ATMOSPHERIC CLIMATOLOGY

The most widely used products of long-range prediction systems have historically been monthly or seasonal mean anomalies of near-surface (2 m) temperature and

accumulated precipitation. Although model biases in these forecast fields can be corrected for as described at the beginning of this subsection, the magnitude of freely running model biases will affect the rate of forecast drift and possibly the quality of the forecasts themselves.

For atmospheric quantities, two sources of bias are (i) biases intrinsic to the AGCM that appear even when realistic SST and sea ice boundary conditions are prescribed, and (ii) coupled biases that arise from errors in the representation of these boundary conditions by the coupled model. Some intrinsic biases of CanAM3 are described in McFarlane et al. (2005) and Scinocca et al. (2008), whereas von Salzen et al. (2013) discusses biases (mainly relating to clouds and precipitation) of CanAM4.

Near-surface temperature biases in historically forced runs of CanCM3 and CanCM4, based on boreal winter [December–February (DJF)] and summer [June–August (JJA)], hereafter all three-month periods are designated by the first letter of each respective month] averages for 1981–2010 with the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) and ECMWF Re-Analysis Interim (ERA-Interim) (Dee et al. 2011) used to represent observations, are shown in Fig. 7. CanCM3 tends to exhibit cold biases over most landmasses, particularly over western North America and southern Asia in DJF and most extratropical land in JJA. By contrast, CanCM4

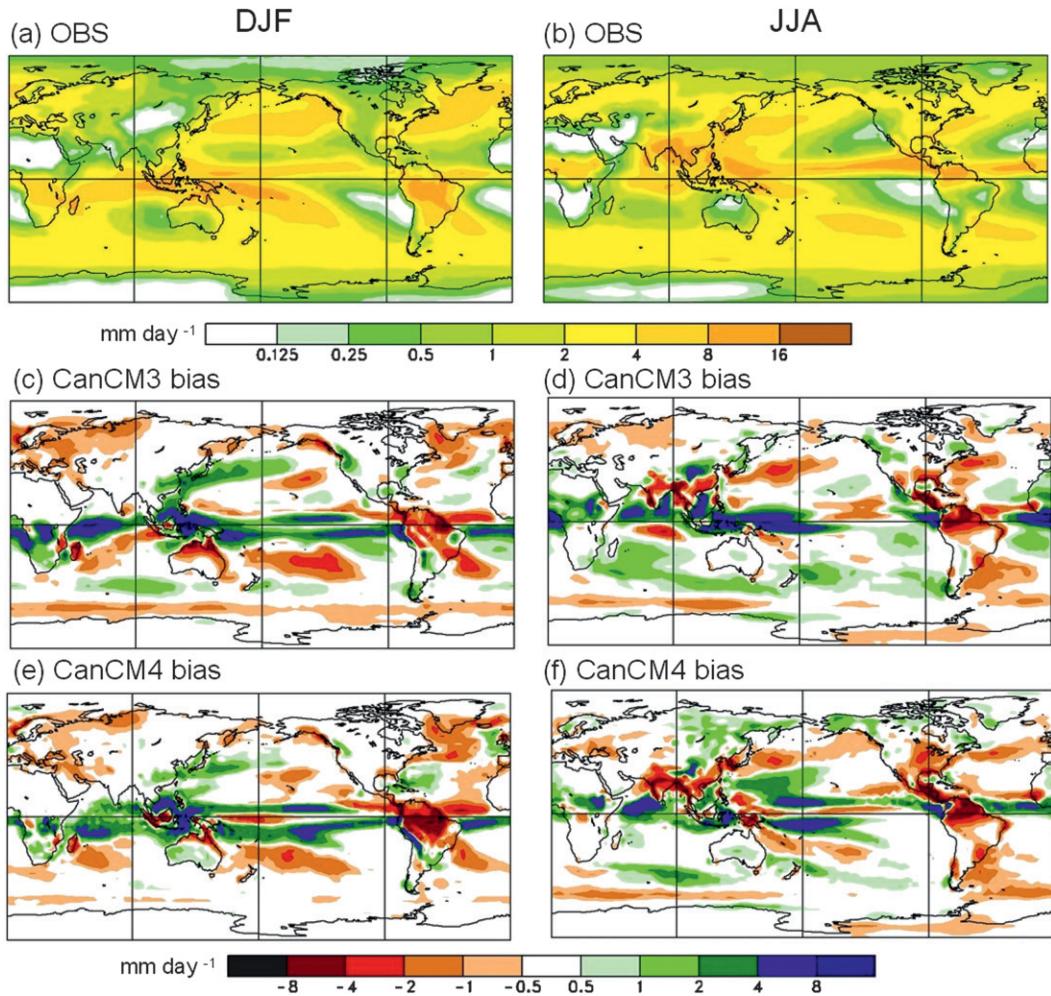


FIG. 8. Observed climatological precipitation for (a) DJF and (b) JJA, based on the GPCP2.1 observational dataset for 1981–2010. Seasonal-mean precipitation bias in a CanCM3 historical run for (c) DJF and (d) JJA, relative to GPCP2.1 for 1981–2010. (e),(f) As in (c),(d), but for CanCM4.

shows strong cold biases over Antarctica in DJF, warm biases over much of Asia in JJA, and warm biases over much of North America and the Amazon basin in both seasons. Biases over land in the two models largely reflect the behavior of their atmospheric components when observation-based SSTs are prescribed (not shown).

Over the oceans, pronounced midlatitude DJF warm biases immediately east of Asia and North America in both models also appear to originate in the atmospheric models. However, the low-latitude warm biases immediately west of Africa and South America are amplified by coupling particularly in CanCM3, and the strong DJF warm biases over the Southern Ocean in CanCM3 are not present in uncoupled CanAM3 simulations.

Modeled seasonal mean precipitation fields are compared with the Global Precipitation Climatology Project, version 2.1 (GPCP2.1), observational dataset

(Adler et al. 2003) in Fig. 8, where the top row shows GPCP2.1 climatological means for DJF (left) and JJA (right) based on a 1981–2010 averaging period. In general, these large scale patterns are represented reasonably well in the models. However, there are some significant regional biases, shown in the remaining panels of Fig. 8. These include

- Too little precipitation in tropical South America;
- Too much precipitation in sub-Saharan Africa, south of the equator in DJF and north of the equator in JJA;
- Too little DJF precipitation in the Middle East, particularly in CanCM4;
- Too little DJF precipitation in northern Australia, particularly in CanCM3;
- Too little monsoon-season (JJA) precipitation in southern and Southeast Asia;

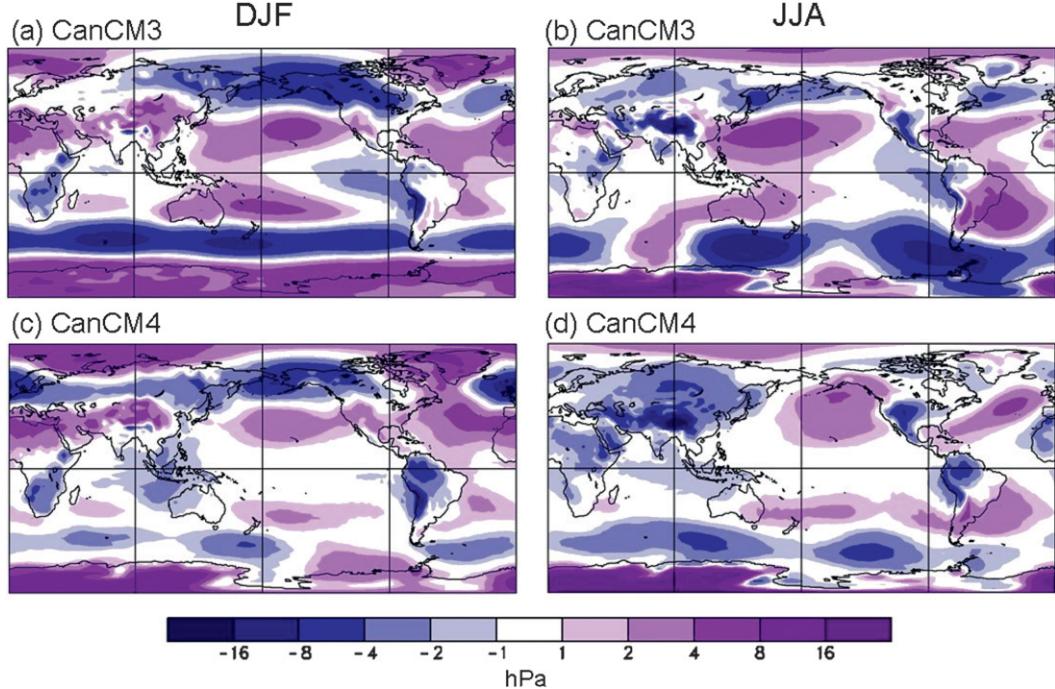


FIG. 9. Mean sea level pressure biases in historically forced runs of (a),(b) CanCM3 and (c),(d) CanCM4 averaging over boreal (a),(c) winters (DJF) and (b),(d) summers (JJA) in 1981–2010. Observational values are from ERA-40 and ERA-Interim for the same period.

- Too little JJA precipitation in southern and eastern North America, particularly in CanCM4.

These biases over land are evident as well in the uncoupled atmospheric models (e.g., von Salzen et al. 2013).

CanCM3 and CanCM4 also exhibit a so-called split intertropical convergence zone (ITCZ) in the tropical Pacific, with a South Pacific convergence zone that is too zonal and extends too far eastward. This bias, which is common in climate models, is particularly severe in CanCM4. The split-ITCZ bias is not as pronounced in uncoupled CanAM3 and CanAM4, and thus appears to be attributable to errors in representing coupled feedbacks, although a general tendency for excessive tropical precipitation in atmospheric models may be a primary cause of such errors (Lin 2007).

Model biases in representing seasonally averaged mean sea level pressure are shown in Fig. 9. The magnitude of such biases is generally reduced in CanCM4 compared to CanCM3, particularly in the Southern Hemisphere. Biases over land are mostly similar to those in atmospheric model runs in which observational SSTs are applied (not shown). However, coupling induces some significant changes in the bias patterns, mainly over ocean regions. For example, the dipolar bias pattern over the northeast Atlantic in DJF, which is of

some significance because of its influence on the North Atlantic storm track and European climate, is common in climate models (e.g., Hazeleger et al. 2011) but is intensified by coupling in both the CanSIPS and other models (e.g., Donner et al. 2011). This may be due to the influence of the strong localized cold bias in North Atlantic SST, evident in Fig. 2, on atmospheric circulation (Keeley et al. 2012). Other features apparently induced by coupling include the DJF high pressure bias in the subtropical North Pacific and the low pressure bias southeast of Australia in CanCM4.

3) SEA ICE CLIMATOLOGY AND TRENDS

Global climate models such as CanCM3 and CanCM4 that simulate sea ice prognostically can potentially generate forecasts of sea ice conditions, although the utility of such forecasts has not yet been explored thoroughly. As is the case for other climate variables, the quality of the sea ice simulation in a freely running model is likely to bear on its predictive capability.

Sea ice extent, defined as the area over which sea ice concentration exceeds 0.15, undergoes large seasonal cycles in both hemispheres, with maximum extent typically occurring in late winter and minimum extent in late summer. Figure 10 compares the observed mean seasonal cycles during 2001–10 with corresponding values

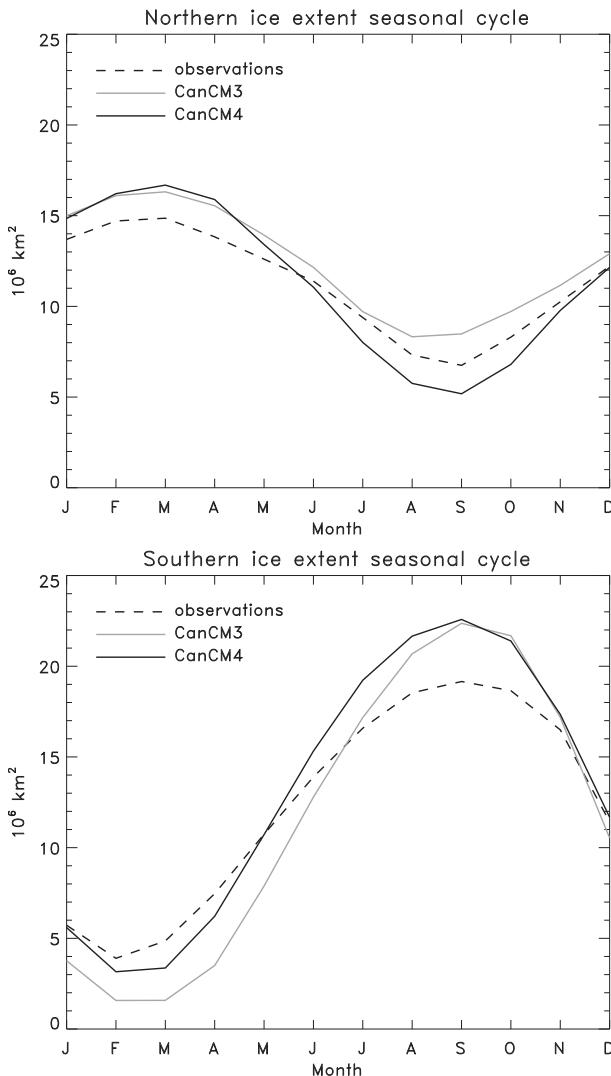


FIG. 10. (top) Climatological mean seasonal cycle for 1981–2000 of Northern Hemisphere sea ice extent from the HadISST1.1 observational dataset and historical runs of CanCM3 and CanCM4. (bottom) Seasonal climatologies for Southern Hemisphere sea ice extent.

for CanCM3 and CanCM4. The winter maximum of Northern Hemisphere ice extent is too large in both models (top panel of Fig. 10), with too much ice in the Labrador, Greenland, Barents, and Okhotsk Seas, and too little in the Bering Sea (Figs. 11a,c). By contrast, late summer ice extent is too high in CanCM3 but too low in CanCM4, with corresponding ice concentration biases evident throughout much of the Arctic Ocean (Figs. 11b,d).

Northern Hemisphere sea ice extent has undergone a steady decline, particularly in its annual minimum, since continuous satellite monitoring began in 1979. Figure 12 shows time series of annual maximum (March)

and minimum (September) ice extent beginning in 1979, from observations and historical runs of the two models (the model time series extend to 2020). Also shown are corresponding linear trends for the 1979–2011 period. March trends for both models are similar to the observed trend, despite the unrealistically high March ice extents in the models. The trend in September extent in CanCM4 is similar to the observed trend, whereas it is markedly weaker in CanCM3.

In the Southern Hemisphere, both models have too much sea ice in austral winter, and too little in austral summer (Fig. 10, bottom panel). The latter bias is more severe in CanCM3 than in CanCM4, reflecting the Southern Ocean warm bias in CanCM3, which is largest in austral summer, whereas in austral winter both models have a cold SST bias in this region (not shown). Such an apparent correspondence between biases in Antarctic sea ice extent and Southern Ocean SST is evident in other current-generation climate models as well (Griffies et al. 2011; Sterl et al. 2012; Landrum et al. 2012).

c. Climate variability

The ability of dynamical models to forecast climate on subseasonal, seasonal, and longer time scales with some skill rests largely on their ability to represent potentially predictable climate anomalies when integrated forward from observationally constrained initial states. It is clearly desirable, therefore, that forecast models be able to represent potentially predictable climate phenomena when not constrained by observational data. This subsection examines CanSIPS model representations of two such phenomena: ENSO, which is a major source of predictability on seasonal to multi-seasonal time scales, and the Madden–Julian oscillation (MJO), which imparts predictability on subseasonal time scales. Although both of these phenomena originate in the tropics they have wider-ranging influences; therefore, associated global teleconnections are considered as well.

1) ENSO VARIABILITY

Versions of the CCCma coupled model preceding those used in CanSIPS exhibited ENSO variability that was much too weak and tended to be concentrated unrealistically in the central equatorial Pacific (e.g., Merryfield 2006). Modeled ENSO variability has, however, become considerably more realistic with the introduction of improved ocean and atmospheric model components in CanCM3 and CanCM4.

ENSO variability is commonly described through anomalies of the Niño-3.4 index, defined as average SST in a region bounded by 5°S–5°N, 120°–170°W. Positive Niño-3.4 anomalies of sufficient magnitude and duration

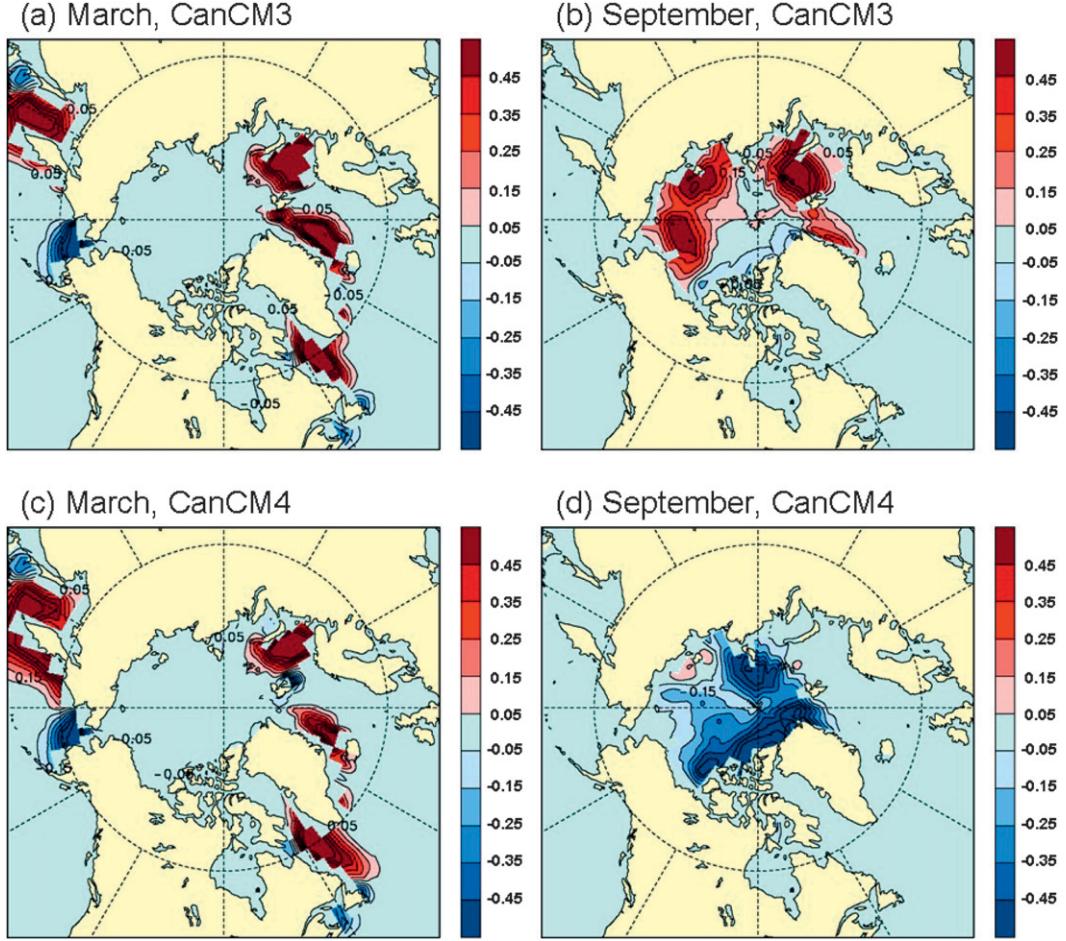


FIG. 11. Climatological sea ice concentration biases for CanCM3 in (a) March and (b) September, relative to HadISST1.1 for 2001–10. Similarly computed biases for CanCM4 in (c) March and (d) September.

are indicative of El Niño events, whereas negative anomalies are similarly indicative of La Niña events.

Figure 13 (top) shows a 50-yr time series (1961–2010) of detrended monthly Niño-3.4 anomalies as determined from the Hadley Center Sea Ice and SST, version 1.1 (HadISST1.1), observational dataset, considered here because of its relatively long temporal coverage and the fact that ENSO variability undergoes considerable interdecadal variation. Evident in this record are El Niño and La Niña events that occur 2–3 times per decade on average, with occasional particularly large events such as the El Niños of 1982/83 and 1997/98. Comparable time series for CanCM3 and CanCM4 are shown in the middle and bottom panels, respectively. Corresponding power spectra plotted in a variance preserving format are shown in Fig. 14, where shading indicates the ensemble standard deviation for the 10 CanCM4 runs. Some statistics of these time series are summarized in Table 5; these include a characteristic period defined as

the period at which the power spectral variance is bisected (Merryfield 2006).

ENSO variability in CanCM3, although stronger than in previous model versions, is still somewhat weaker than observed, with a standard deviation of 0.54 for this period as compared to 0.84 for HadISST. CanCM3 variability is also too rapid, with a characteristic period of slightly over 2 yr as compared with the observed value of 3.27 yr. By contrast, ENSO in CanCM4 is more realistic both in terms of its amplitude, which is slightly too large on average, and its characteristic period, which is approximately correct. Also evident from the power spectra are that CanCM3 has far too little Niño-3.4 variability on decadal to multidecadal time scales, whereas CanCM4 has slightly too much. Another statistic in Table 5 is skewness: observed ENSO SST anomalies tend to have positive skewness in the Niño-3.4 region, with El Niño anomalies stronger than La Niña anomalies (Burgers and Stephenson 1999). However,

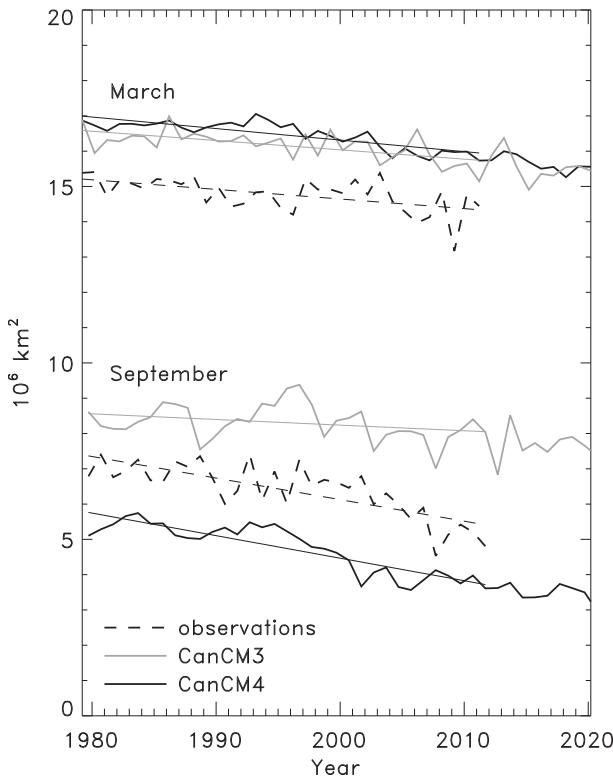


FIG. 12. Time series, dating from the start of the satellite record in 1979, of monthly mean Northern Hemisphere sea ice extent in March and September from HadISST1.1 (dashed line), as well as historical runs of CanCM3 (light solid line) and CanCM4 (dark solid line). Corresponding linear trends for 1979–2011 are indicated by the straight lines.

Niño-3.4 skewness is negative in both models. Modeled skewness remains negative farther east in the Niño-3 region bounded by 5°S–5°N, 90°–150°W (not shown), in contrast to observed positive values, which increase monotonically toward South America (e.g., Monahan and Dai 2004).

An important aspect of ENSO variability is its seasonality, indicated in Fig. 15 by the standard deviation of the Niño-3.4 index as a function of calendar month. Observed ENSO variability (dashed) tends to be strongest in boreal winter, peaking in December, and weakest in late spring. These tendencies are not always represented realistically in climate models (Joseph and Nigam 2006). However, ENSO seasonality is approximately correct in CanCM3 and CanCM4 (light and dark solid lines, respectively), with the amplitude biases persisting in all calendar months, although the seasonal minimum in ENSO activity occurs about one month too late (in May instead of April) in both models.

The spatial pattern and regional intensity of ENSO SST variability are indicated in Fig. 16 by the standard

deviation of monthly SST anomalies throughout the tropical Pacific. The model biases in ENSO amplitude (too weak in CanCM3 and too strong in CanCM4) are again evident. In addition, there are further biases in the patterns of near-equatorial SST variability. For example, the meridional extent of modeled variability is too narrow, particularly in CanCM3, and the SST anomalies extend too far westward, nearly to the Maritime Continent, in both models. These biases have each been found to occur frequently in other coupled models (Joseph and Nigam 2006).

2) ENSO TELECONNECTIONS

An important attribute of seasonal forecast models is their ability to represent ENSO teleconnections, since the global influence of ENSO plays a substantial role in shaping seasonal climate anomalies. In North America, these influences tend to be strongest in boreal winter (e.g., Ropelewski and Halpert 1986; Shabbar et al. 1997), although ENSO influences on western Canadian temperatures persist into early spring (Shabbar 2006). In addition, precipitation over the Great Basin and possibly the high plains of the United States may be influenced in summer (Ropelewski and Halpert 1986), although the strength of this signal varies across different analyses (Wang et al. 2007).

Observed and modeled ENSO teleconnections are assessed here through linear regressions of climate variables with the Niño-3.4 index as for example in Yang and DelSole (2012).³ This approach assumes linearity implying opposite responses to El Niño and La Niña, whereas that approximation is known to be violated to varying degrees in certain regions (e.g., Wu and Hsieh 2004; Wang et al. 2007). Observed teleconnections are diagnosed from the ERA-40 and ERA-Interim reanalyses for surface temperature and pressure, and the Global Precipitation Climatology Project, version 2.2 (GPCP2.2), dataset for precipitation, for the years 1981–2010, which corresponds to the CanSIPS hindcast period, whereas modeled teleconnections are diagnosed for the same period in CanCM3 and CanCM4 historical runs. Field significance is assessed very simply by requiring that the local correlation coefficient exceed 0.3. Assuming temporal independence in the 30-yr sample this implies $p \leq 0.1$ or $>90\%$ confidence for a two-tailed test. Such a local significance test does not account for the finite extent and hence spatial correlation of teleconnection patterns, and so may tend both to underestimate

³ Construction of composite differences between El Niño and La Niña phases serves as an alternative approach (e.g., Smith et al. 2012).

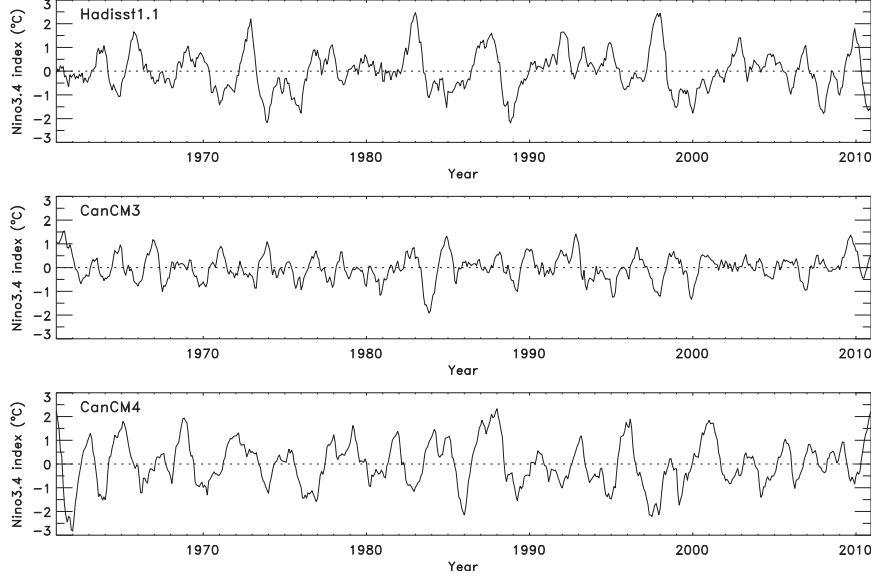


FIG. 13. Monthly mean Niño-3.4 index anomalies for 1961–2010 from (top) HadISST1.1 and historical runs of (middle) CanCM3 and (bottom) CanCM4.

the spatial coherence and extent of true teleconnections and produce occasional small-scale “false positives.” A more sophisticated field significance test that takes spatial correlations into account is described in DelSole and Yang (2011) and Yang and DelSole (2012).

ENSO-teleconnected regression patterns for boreal winter (DJF) are shown in Fig. 17. Surface temperature and mean sea level pressure (MSLP) anomalies are indicated by colors and contours, respectively, in the

left-hand panels, and standardized precipitation anomalies are shown in the right-hand panels. Values where the local correlation with the Niño-3.4 index are <0.3 are not shown for reasons indicated above. For CanCM4, regression and correlation values averaged over the ensemble of 10 historical runs are used. Although the 0.3 correlation threshold is maintained despite the larger sample size to maintain similarity with the observational and CanCM3 maps, confidence in the CanCM4 patterns can be considered higher than for the other cases.

Focusing first on the tropical oceans, the equatorial Pacific warm anomalies associated with El Niño in both models extend too far to the west as implied by the SST variability patterns in Fig. 16, whereas modeled warm anomalies in the tropical Indian and Atlantic Oceans occur more or less as in observations. The precipitation response in the tropical Pacific and eastern Indian Oceans is particularly important because its accuracy in models is a major factor determining the fidelity of extratropical teleconnections which are forced largely by tropical diabatic heating anomalies (AchutaRao and

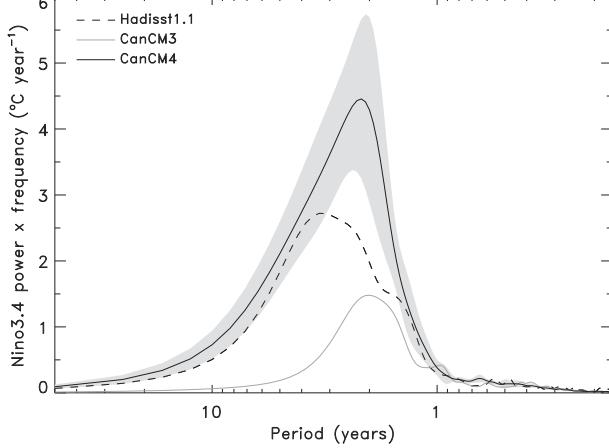


FIG. 14. Power spectra in variance preserving format of the 1961–2010 monthly Niño-3.4 index from HadISST1.1 (dashed line), one CanCM3 historical run (light solid line), and 10 CanCM4 historical runs (dark solid line, with shading representing the ensemble standard deviation). Split-cosine tapering was applied to the first and last 10% of time series, and spectra were smoothed using a 12-bin Parzen window.

TABLE 5. Observational and model statistics of detrended monthly Niño-3.4 index for 1961–2010.

Niño-3.4 data	Std dev ($^{\circ}$ C)	Skewness	Period (yr)
HadISST1.1	0.84	0.23	3.27
CanCM3	0.54	-0.12	2.04
CanCM4*	1.03 ± 0.07	-0.25 ± 0.18	2.98 ± 0.31

* With ensemble standard deviations.

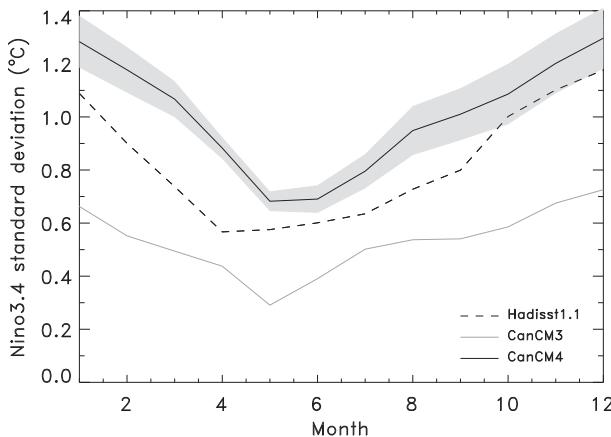


FIG. 15. Seasonal cycle of monthly Niño-3.4 index standard deviation, based on 1961–2010 time series from HadisST1.1 (dashed line), one CanCM3 historical run (light solid line), and 10 CanCM4 historical runs (dark solid line, with shading representing the ensemble standard deviation).

Sperber 2006, and references therein). Both models correctly show that a strong precipitation increase in the central and eastern equatorial Pacific occurs with El Niño events. In observations, associated diabatic heating anomalies occurring near the date line are largely responsible for extratropical teleconnections (e.g., Trenberth et al. 1998). However, in CanCM3 and CanCM4 this region of increased precipitation extends too far westward, a common model bias that may in turn lead to westward displacements in extratropical teleconnections (Wittenberg et al. 2006).

The dominant feature in the MSLP regression patterns, other than the tropical east–west dipole associated with the Southern Oscillation (present but not clearly evident in Fig. 17 as a result of the choice of contour levels), is a pronounced deepening of the Aleutian low associated with El Niño. This feature, which is connected to shifts in circulation including the North Pacific storm track (Ren et al. 2007), is modeled reasonably realistically, although its center occurs somewhat too far west on average in the 10 CanCM4 historical runs.

Pronounced and extensive features occur in the DJF surface temperature regression patterns over all continental landmasses except Eurasia and Antarctica. [A cold feature in northeastern Asia identified by Yang and DelSole (2012) in observations is weakly evident in the present analysis.] An observed warm anomaly in northwestern North America associated with El Niño is evident in both models, although its extension into central North America in observations and CanCM3 is not evident in CanCM4. The extensive cold anomaly in the southeastern United States and northern Mexico in CanCM4 is in good agreement with that deduced from observations by

DelSole and Yang (2011), although it is less extensive in the current analyses of observations and CanCM3. This feature does, however, become more prominent in March–May (MAM), while shifting westward in observations and both models; similarly the aforementioned warm anomalies remain in MAM but become confined to northwestern North America (not shown).

The primary North American signals in the DJF precipitation regression patterns, most evident in observations and CanCM4, are an El Niño-related precipitation increase across southern North America, accompanied by decreases in the Pacific Northwest and Great Lakes regions. These signals are in general agreement with analyses by Shabbar (2006) and Yang and DelSole (2012).

Other DJF teleconnections over land common to differing degrees in observations and both models are a general El Niño-linked warming over tropical landmasses and Australia that tends to be stronger in the models, reduced precipitation in southern Central America and northern South America accompanied by wetter conditions to the southeast, dry conditions over the Maritime Continent and Southeast Asia, and enhanced precipitation in southwest Asia, the Middle East, and the Horn of Africa. A strong El Niño-linked tendency toward drought in southern Africa is only weakly evident in the models, whereas some features in the models such as cooling in eastern China in CanCM3 and enhanced precipitation in West Africa in both models are not evident in observations.

Teleconnection patterns in the boreal summer months of JJA (Fig. 18) tend to be less distinct than in DJF, likely because ENSO itself is weaker in these months (Fig. 15). (The regression patterns for CanCM3 are particularly “noisy” because of its comparatively weak ENSO.) For temperature there is little commonality between observed and modeled patterns other than warming tendencies in parts of northern Africa, southern Asia, and equatorial South America, much as found by Yang and DelSole (2012) who compared observed patterns with those from coupled hindcasts.⁴

JJA precipitation signals over land include dry conditions over the Maritime Continent, the Horn of Africa, Central America, and northern South America (not clearly evident in CanCM3), along with decreased rainfall in the Indian summer monsoon region. There is also increased rainfall in the western United States in all three cases although in differing regions. Such a signal,

⁴ The pronounced cooling in the Ural region of Eurasia is not seen clearly in studies that consider different analysis periods and so may be a sampling artifact.

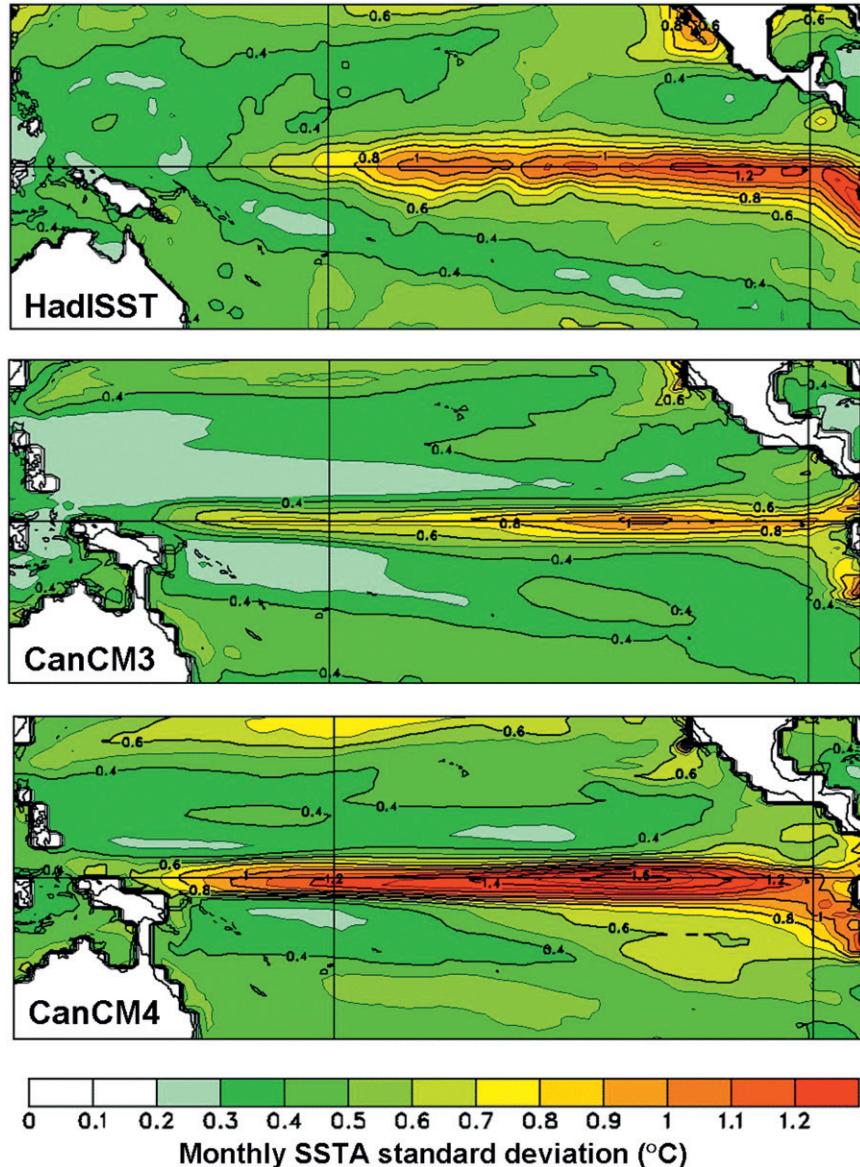


FIG. 16. Standard deviations of 1961–2010 monthly SST anomalies from (top) HadISST1.1 and historical runs of (middle) CanCM3 and (bottom) CanCM4.

though identified in previous analyses (Ropelewski and Halpert 1986, 1987) was found not to be statistically significant by Yang and DelSole (2012). Some ENSO impacts on boreal summer rainfall that are evident in Fig. 18b and other analyses are not captured by CanCM3 or CanCM4. These include El Niño-linked dry conditions in the Sahel region of Africa and eastern Australia.

3) INTERANNUAL VARIABILITY OF SEASONAL MEANS

A further potential bias in climate prediction models, in addition to biases in representing the climatological

mean and seasonal cycle, is erroneous representation of the level of interannual variability as has been noted above for ENSO-related SST variability in CanCM3 and CanCM4. Such biases should be addressed for example when combining forecasts from different models, because models that have unrealistically large interannual climate anomalies will otherwise tend to dominate multimodel ensemble averages unless some correction is applied (e.g., Kharin et al. 2009).

Figure 19 shows the ratio of modeled to observed standard deviations of seasonal mean (DJF and JJA) near-surface temperature anomalies for the 1981–2010

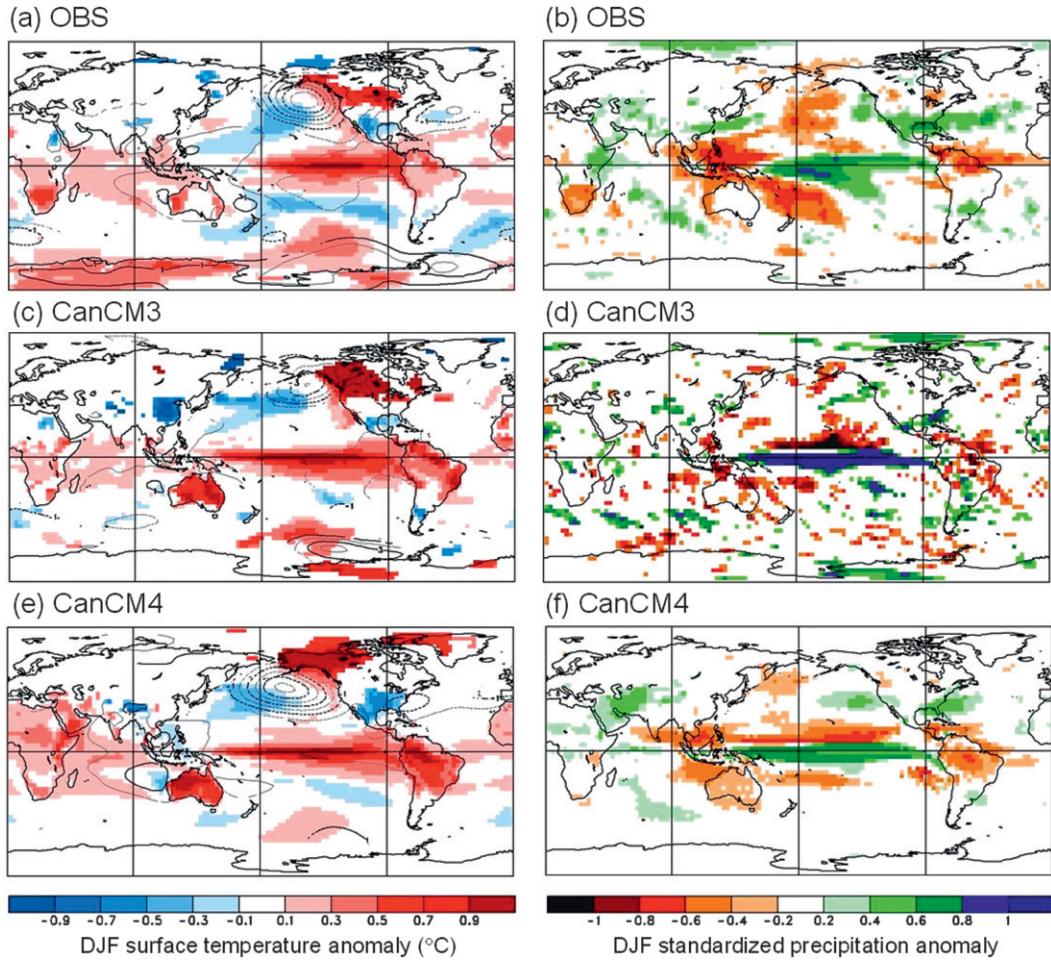


FIG. 17. Regressions against DJF averages of the Niño-3.4 index of DJF averaged mean sea level pressure (contours), (a),(c),(e) near-surface temperature (colors) and (b),(d),(f) standardized precipitation anomalies (colors) for 1981–2010 from (a),(b) observations; (c),(d) one CanCM3 historical run; and (e),(f) 10 CanCM4 historical runs. Values for which the correlations are <0.3 (two-tailed $p \geq 0.1$) are not plotted. Observational data sources are ERA-40 and ERA-Interim for pressure and temperature, and GPCP2.2 for precipitation.

CanSIPS hindcast period. As in other instances these results are based on a single CanCM3 historical run and 10 CanCM4 historical runs. (Standard deviations for CanCM4 are computed as the mean of the standard deviations for the individual runs.) It is evident that interannual variability in CanCM4 is generally larger than in CanCM3, likely as a result of the difference in ENSO amplitudes. In particular, CanCM3 variability tends to be somewhat weaker than observed in most ocean regions whereas such variability in CanCM4 is slightly too large. A location of excessive ocean variability common to both models is the far western equatorial Pacific, a result of the unrealistic westward extension of ENSO SST anomalies illustrated in Fig. 16.

Other biases in variability include CanCM4 having far too much variability in tropical South America, which is likely a consequence of the severe dry bias evident in

Fig. 8 combined with the modest excess in ENSO variability. Variability in South Asia, central Asia, and much of North America is also too high in both models, particularly in JJA.

Biases in the interannual variability of precipitation as measured by the ratio of standard deviations tend to scale with biases in the mean, which are considerable in some locations in fractional terms (Fig. 8). Considering instead the ratio of modeled to observed standard deviations scaled by their respective means [i.e., $(\sigma_m/\bar{m})/(\sigma_o/\bar{o})$, where m denotes model values, o denotes observed values, and the overbars denote mean values], provides a measure of differences in *relative* variability at each location. This measure (Fig. 20) indicates that relative variability over land in the models tends to exceed that deduced from observations, particularly in seasons and regions where the models exhibit

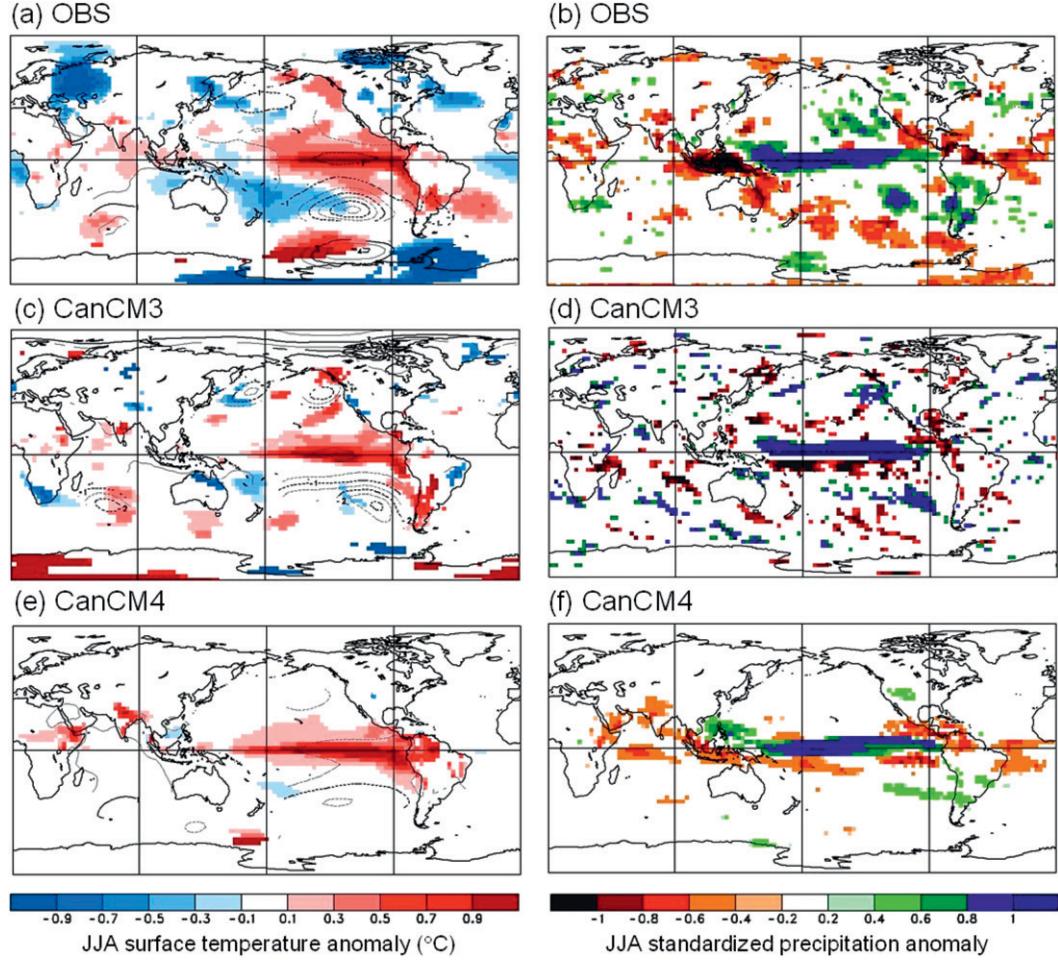


FIG. 18. As in Fig. 17, but for JJA.

dry biases, such as North America and South Asia in JJA and Central and tropical South America in all seasons. Instances where modeled relative variability is weaker than observed over land such as southern Africa tend to correspond to regions where an observed ENSO teleconnection is not clearly evident in one or both models.

4) MJO VARIABILITY

The MJO is characterized by a coherent, eastward-propagating concentration of tropical deep convection that is most prominent in the Indo-Pacific and has a characteristic period of 30–80 days (Wheeler and Kiladis 1999). In addition to its considerable direct influence on tropical climate, diabatic heating anomalies associated with the MJO give rise to Rossby wave trains that influence extratropical climate in both hemispheres (e.g., Matthews et al. 2004; Donald et al. 2006). These influences include the lagged modulation of North American temperature and precipitation particularly in

the boreal winter months (Mo and Higgins 1998; Jones 2000; Bond and Vecchi 2003; Vecchi and Bond 2004; Lin and Brunet 2009; Lin et al. 2010; Zhou et al. 2012) and a two-way connection with the North Atlantic Oscillation (Lin et al. 2009). In addition, the MJO is associated with westerly wind bursts that can contribute to the growth or termination of ENSO events (Seiki and Takayabu 2007).

The MJO is thus a potential source of subseasonal and possibly seasonal (through its influence on ENSO) climate predictability. Such predictability stems in part from the lagged influence of the initial state of the tropical Indo-Pacific atmosphere on other regions, but potentially also from any forecast skill in predicting the future evolution of the MJO itself. As for ENSO, the ability of a model to represent MJO variability when not constrained by observations is likely to bear on its ability to predict the MJO. However, the MJO has not been found to be simulated realistically in most climate models (Lin et al. 2006), suggesting a need for improved

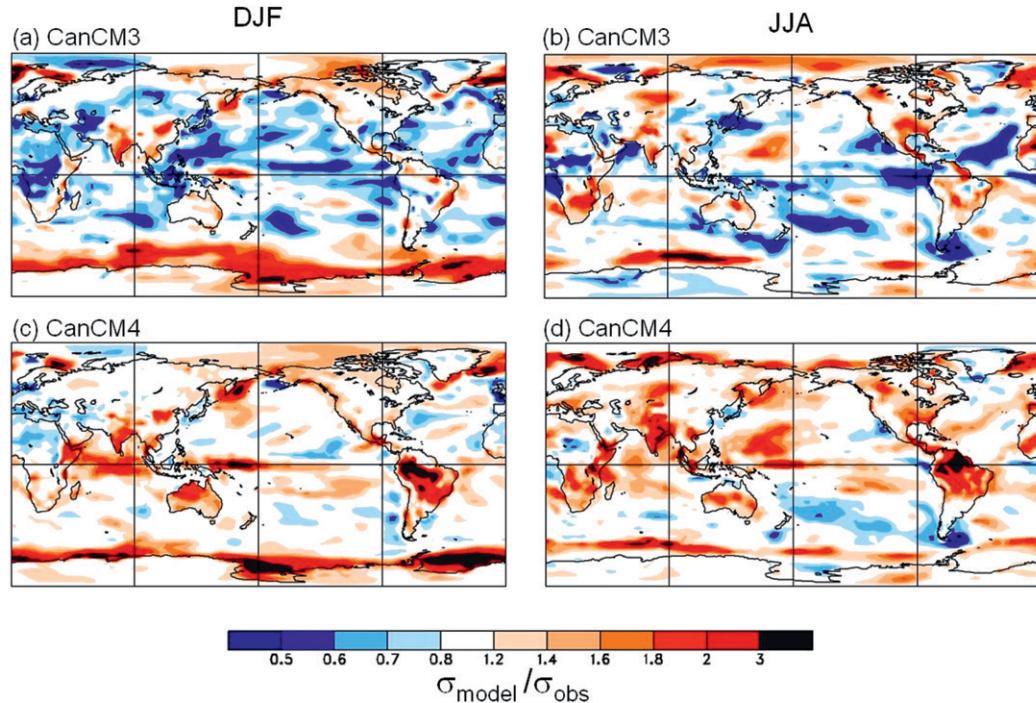


FIG. 19. Ratios of modeled to observed standard deviations of seasonal mean near-surface temperatures in 1981–2010 for (a) CanCM3 in DJF, (b) CanCM3 in JJA, (c) CanCM4 in DJF, and (d) CanCM4 in JJA. Observed temperatures are from the ERA-40 and ERA-Interim.

parameterizations of deep convection and perhaps increased resolution in such models.

A comprehensive set of diagnostics for describing MJO variability in climate simulations has been developed by the Climate Variability and Predictability (CLIVAR) MJO Working Group (Waliser et al. 2009) and applied to a variety of climate models (Kim et al. 2009). These diagnostics have been applied to 50-yr time series from freely running CanCM3 and CanCM4 simulations, enabling aspects of MJO variability to be quantified relative to the observed MJO and other models. A small subset of these diagnostics that provides a general view of the modeled MJO variability is described here. Figure 21 shows the 20–100-day filtered variance of wintertime (November–April) tropical outgoing longwave radiation (OLR), which provides a measure of the intraseasonal variability of deep convection. The top panel shows the variability of OLR measurements from the Advanced Very High Resolution Radiometer (AVHRR). By comparison, intraseasonal OLR variability in CanCM3 (middle panel, note difference in scale) is much too weak, and there are significant errors in its distribution with too much variability in the western Indian Ocean and western Pacific relative to the central Indian Ocean and Maritime Continent (pattern correlation 0.71). These biases may be related to a tendency for large-scale stratiform

precipitation to supplant deep convective tropical precipitation in CanCM3 (Scinocca and McFarlane 2004). The magnitude of intraseasonal OLR variability is more realistic in CanCM4 although still somewhat too weak (bottom panel), whereas its spatial distribution, with a pattern correlation of 0.86, is considerably more realistic than in CanCM3.

While Fig. 21 describes the amount and spatial distribution of tropical intraseasonal variability, it does not provide information about the spatiotemporal organization of that variability. One way of determining such properties, including multivariate phase relationships, is through the combined empirical orthogonal function (EOF) analysis of meridionally averaged (15°S – 15°N) anomalies of 850- and 200-hPa zonal winds and OLR developed by Wheeler and Hendon (2004). Observationally, the first two EOF modes form a pair that describe the eastward-propagating MJO; the corresponding principal components (PC) are often used to describe MJO evolution (e.g., Fig. 7 of Wheeler and Hendon 2004). Several properties of these leading modes as observed and simulated in CanCM3 and CanCM4 are summarized in Table 6. Whereas the two leading EOF modes explain 0.43 of the overall observed variance, the explained variance fraction in both models is only about half this value. Furthermore, whereas approximately 0.6 of the variance associated with the observed leading

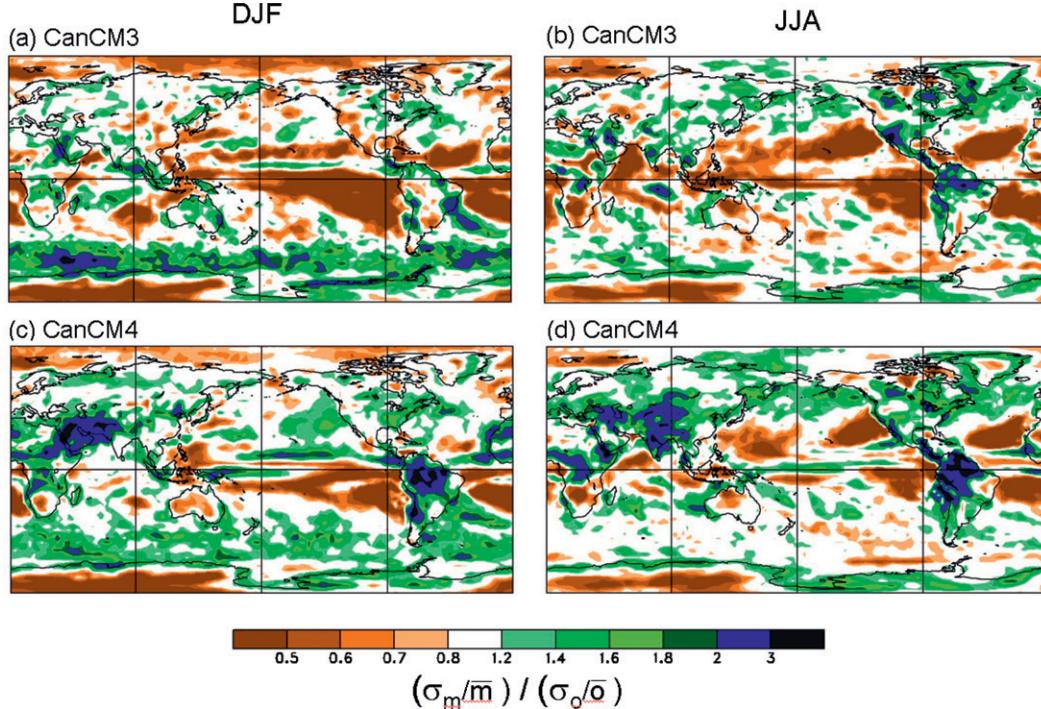


FIG. 20. Ratio of modeled coefficient of variation of seasonal mean precipitation in 1981–2010 (i.e., standard deviation scaled by the mean, σ_m/\bar{m} , to corresponding observed values σ_o/\bar{o}), for (a) CanCM3 in DJF, (b) CanCM3 in JJA, (c) CanCM4 in DJF, and (d) CanCM4 in JJA. Observed precipitation values are from the GPCP2.2 dataset.

modes lies in the 30–80-day band, this fraction is considerably lower for the models. Finally, the squared coherence of the two modes in this intraseasonal band is again lower in the models than observed, especially in CanCM4. Taken together, these results indicate that the MJO is not well simulated by either model, and that while the amount and distribution of tropical deep convective variability is more realistic in CanCM4 than in CanCM3, the MJO-like component of that variability remains rather weak.

4. Initialization procedure

The objective of climate forecast initialization procedures is to impart observational information about the state of the climate system at the forecast start time into the prediction model. This requires assimilation procedures whereby measurements that are local in space and time act to constrain the full model state in a way that is consistent with the governing dynamics. While the state of the ocean is usually viewed as being the primary source of seasonal to multiseasonal predictability, the initial states of other climate system components (atmosphere, land, sea ice) as well as radiative forcings (e.g., volcanic aerosols) can also contribute to predictability on differing time scales.

Operational climate prediction systems have typically been initialized by combining separate assimilative analyses for different climate system components (e.g., Saha et al. 2006). However, techniques for coupled data

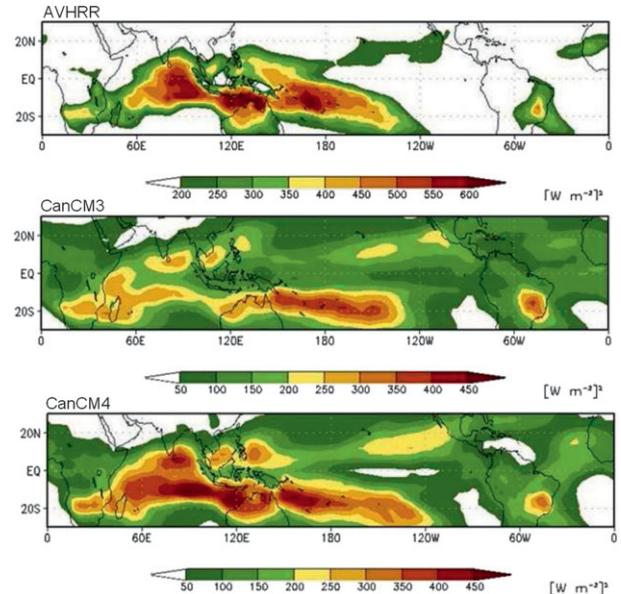


FIG. 21. Winter (November–April) 20–100-day-filtered OLR variance from (top) AVHRR observations, and 50-yr time series from (middle) CanCM3 and (bottom) CanCM4.

TABLE 6. Properties of the first two combined EOFs of near-equatorial U_{850} , U_{200} , and OLR in observations and the CanCM3 and CanCM4 models.

Source	Explained variance of EOF1 and EOF2	Fraction of PC1 variance in 30–80-day band	Fraction of PC2 variance in 30–80-day band	(Coherence) ² of PC1 and PC2 in 30–80-day band
Observed	0.43	0.58	0.60	0.77
CanCM3	0.22	0.41	0.34	0.48
CanCM4	0.18	0.27	0.31	0.33

assimilation are being developed (e.g., Chang et al. 2012) which offer the prospect of improved compatibility and balance between the different components of climate models.

A major consideration guiding the development of initialization methodologies for CanSIPS was that, once CanSIPS was implemented operationally, gridded CMC analyses for global atmospheric fields, SST, and sea ice concentration would be available in near-real time. Furthermore, a CMC analysis for the subsurface ocean, although not available immediately, was under development. Therefore each of these climate system components is initialized through the assimilation of gridded analyses.

A further design choice was that assimilation of these data products would be undertaken while running the coupled model, using methods that are relatively simple to implement. This was a natural approach because assimilation methodologies were being developed essentially ab initio and were no more difficult to implement in a coupled environment than for separately running model components; also, choosing relatively simple methodologies was an option because of the availability of global gridded products and was desirable in order to expedite system development.

This general approach to forecast initialization is illustrated in the top portion of Fig. 22: atmospheric variables, SST, and sea ice concentration are assimilated continuously during coupled model runs, whereas the subsurface ocean is treated through a separate offline procedure described below. In each case full field values are assimilated as is common practice in seasonal forecasting, and identical initialization procedures are used for the two models except where noted below. Assimilation constraints are released upon initiating the forecast. Details regarding generation of the forecast ensemble and the initialization of specific climate system components are described below.

a. Assimilation run ensemble

Several strategies have been employed to generate climate forecast ensembles for a particular model. For example, the use of stochastic or perturbed physics will cause each ensemble member to evolve differently even

if the initial conditions are identical (Luo et al. 2005; Doblas-Reyes et al. 2009). Alternatively one can run each forecast ensemble member from different initial conditions. These can be obtained from sets of atmospheric or other analyses valid at differing times preceding the forecast start, an approach that was employed in CMC's former two-tier system (Kharin et al. 2009). Multiple initial conditions can also be produced by introducing optimally rapidly growing perturbations computed from singular or bred vectors (e.g., Palmer 1999; Yang et al. 2006; Kug et al. 2010). Yet another approach is to draw from an ensemble of different assimilation calculations; this is the approach adopted here.

CanSIPS produces ensembles of forecast initial conditions, 10 for each model, through separate assimilation runs for each ensemble member. Each such run, though assimilating the same data, originated from different initial conditions so that the simulated climate states (constrained by observations) differ between ensemble members. This setup, illustrated schematically in the lower part of Fig. 22, produces sets of forecast initial conditions that are statistically equivalent, and ensures at least a modicum of consistency between the climate system components due to its coupled nature.

The initial conditions from which the assimilation runs originate were specified somewhat differently for CanCM3 and CanCM4. For CanCM3, two control runs with ca. 1990 radiative forcing were first initialized from PHC/WOA climatology (Steele et al. 2001) and run for 50 yr. Ten initial states for the assimilation runs were then drawn from model states at the ends of years 30 to 50 in these two runs at intervals of 5 yr. Assimilation was then begun in 1948, providing some three decades of assimilative input prior to the start of the first historical forecasts initialized in 1979. For the CanCM4 assimilation runs, which provide initial conditions for CMIP5 decadal forecasts beginning at the end of 1960 in addition to the CanSIPS hindcasts and forecasts, the initialization procedure began with a 350-yr spinup from PHC/WOA initial conditions, during which ERA-40 and other data for the decade of the 1960s was assimilated repeatedly using ca. 1958 radiative forcing. This procedure was intended to bring the deep ocean into

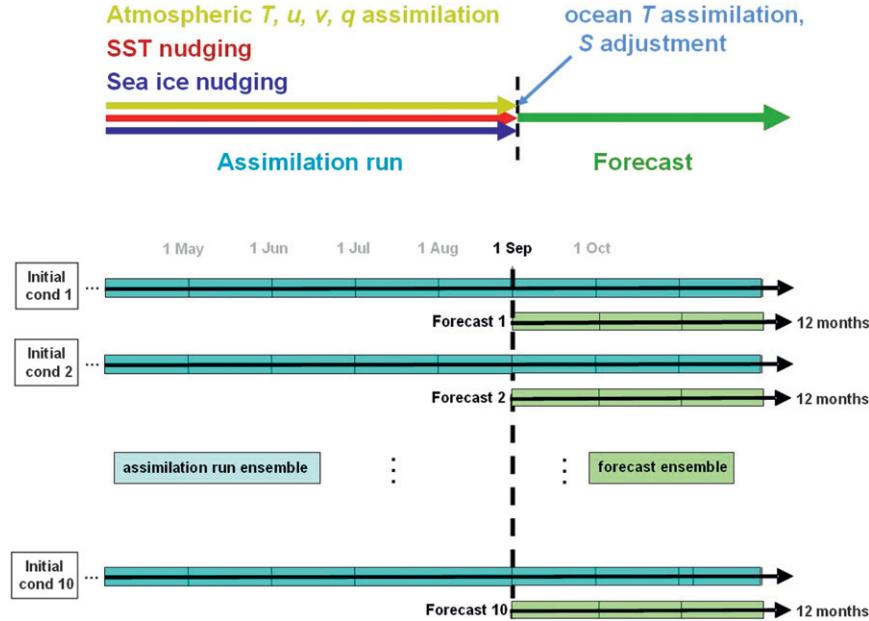


FIG. 22. Schematic illustration of the (top) assimilation and (bottom) ensemble generation procedures used to initialize CanSIPS forecasts.

approximate thermal equilibrium with climate forcing from this epoch in order to minimize model drifts unrelated to climate forcing over the multidecadal period, beginning in late 1960, spanned by the CMIP5 decadal hindcasts and forecasts. Ten model states were then drawn at 1-yr intervals from this run to initialize an ensemble of runs assimilating data from 1958 until the present. For both models, radiative forcings for the assimilation runs are the same as those for the freely running climate simulations as described in section 2c.

b. Atmospheric initialization

Observation-based initialization of the model atmospheric state provides directly, through the atmospheric initial conditions, enhanced prediction skill during the first 10 days or so of a forecast. This is an asset to climate predictions of the first month or season following initialization. In addition, assimilation of atmospheric data over a period leading up to the forecast can improve the initialization of other components of the climate system as discussed in the following subsections.

In principle, a full data assimilation package could be used to initialize the atmospheric model. For example, three-dimensional variational data assimilation (3D-Var) has been implemented in a middle atmosphere version of CanAM3, CMAM-DAS (Polavarapu et al. 2005). In practice, however, implementation of such complex data assimilation in both CanSIPS models is difficult to justify since incremental improvement to the quality of the

initial atmospheric state offers limited value on seasonal to interannual time scales. For CanSIPS, a simpler approach is taken in which gridded analysis or reanalysis products from another model are employed in place of states derived from application of 3D-Var (or another complex method) within the CanSIPS models. In addition, the incremental analysis update (IAU) assimilation scheme employed in CMAM-DAS (Polavarapu et al. 2004) has been modified as follows.

As for standard IAU, each 6-h assimilation cycle is begun by first running the model freely for 3 h from the midpoint between analysis states (0300, 0900, 1500, or 2100 UTC) to produce a forecast at a time coinciding with that of the gridded analysis data (0600, 1200, 1800, or 2400 UTC). The difference between the analysis and the forecast model state at this time t_a is taken to be the analysis increment, $\Delta\mathbf{x}_a \equiv \mathbf{x}_{t_a}^{\text{forecast}} - \mathbf{x}_{t_a}^{\text{analysis}}$. The model is then rerun from the start of the forecast for the entire 6-h cycle while applying a constant forcing equal to a fraction of the analysis increment $\Delta\mathbf{x}_a$ at every time step. In the standard implementation of IAU this constant fraction is chosen such that the full analysis increment is applied by the end of the 6-h integration. Schematically, this integration can be represented as

$$\frac{d\mathbf{x}}{dt} = M(\mathbf{x}) - \tau^{-1} \Delta\mathbf{x}_a, \quad (1)$$

where $M(\mathbf{x})$ represents the freely running model equations and $\tau = 6$ h in standard IAU. The gradual insertion

TABLE 7. CanSIPS assimilation data sources.

Field	Frequency	Hindcast data sources	Forecast data sources
Air temperature	6-hourly	ERA-Interim (1989–present)	CMC operational analysis
Zonal wind		ERA-40 (1958–88)	
Meridional wind		NCEP–NCAR reanalysis 1 (1948–57) ^a	
Specific humidity			
Sea ice concentration	Daily	Monthly HadISST ^b	CMC operational analysis
SST	Daily	Weekly OISST (November 1981–2010) ^b monthly ERSST (before November 1981) ^b	CMC operational analysis
Interior ocean temperature	Offline prior to forecast start	Monthly GODAS (1980–2010) ^b Monthly SODA (before 1980) ^{b,d}	Daily GODAS ^c

^a CanCM3 only.^b Interpolated to daily.^c Until daily CMC Mercator analysis available.^d SODA anomalies are superimposed on GODAS climatology.

of the analysis increment acts to temporally filter the model response, primarily on time scales shorter than the 6-h analysis window, in contrast to relaxational nudging, which filters both the response and the evolving model state (Bloom et al. 1996).

Whereas standard IAU is applied to all spatial scales resolved by the model, CanSIPS employs a modified version of IAU in which the increment $\Delta\mathbf{x}_a$ in (1) is subjected to a low-pass spatial filter and the value of τ is allowed to exceed 6 h. We refer to this modified version of IAU as constant increment nudging (CIN). While CIN represents a weaker constraint on the model than IAU, CIN acts mainly to constrain the synoptic and planetary-scale evolution of the climate system, which is most reliably represented by the analysis data.

In CanSIPS, CIN is applied to atmospheric temperature T , zonal and meridional wind, and specific humidity q , employing as input 6-hourly gridded values of these quantities obtained by interpolating analysis data from pressure levels to the model grids. Because the CMC analysis used operationally does not cover the hindcast period, the atmospheric reanalyses listed in Table 7 were utilized in initializing the hindcasts. Specifically these were the ERA-Interim (Dee et al. 2011), which at the time the assimilation runs were undertaken was available from 1989, and ERA-40 (Uppala et al. 2005) to cover the preceding period to 1958. In addition, the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis 1 (Kalnay et al. 1996) was used prior to 1958 for CanCM3 assimilation runs only.

The implementation of CIN in CanSIPS uses $\tau = 24$ h in (1) and low-pass filters the increment $\Delta\mathbf{x}_a$ to include spherical harmonics with total wavenumbers up to 21 in a triangularly truncated spectral representation (T21) of the model and analysis states. These temporal and spatial scales were selected so that the root-mean-square

differences (rmsd) between the nudged model temperature fields and reanalysis fields are comparable to rmsd between pairs of different reanalysis datasets, where the latter provides an estimate of the uncertainty in the input atmospheric reanalysis.

Sensitivity to τ is illustrated in Fig. 23a for a set of cases where the input data are spectrally truncated at T35 and specific humidity is not assimilated (indicated by “-q” in the legend). The various curves represent rmsd of T between a CanCM3 assimilation run and the ERA-40 input data for different τ (colors), as well as between pairs of reanalyses (gray and black) that include ERA-40, ERA-Interim, NCEP–NCAR, and NCEP–Department of Energy (Kanamitsu et al. 2002, labeled as NCEP2), for a period 1979–88 when all four reanalyses are available. Without assimilation (magenta), rmsd between the model and ERA-40 are much larger than the observational uncertainties as expected. By contrast, with $\tau = 6$ h as in standard IAU (red), the model versus ERA-40 rmsd are much smaller than the observational uncertainties even though the ERA-40 input data has been spatially truncated. Higher values of τ yield increasing rmsd, with $\tau = 24$ and 48 h producing rmsd that lie generally within the range of observational uncertainties (green and cyan).

The remaining panels in Fig. 23 illustrate sensitivity to spectral cutoff and to assimilation of specific humidity (indicated by “+q” in the legend), for temperature, vorticity and q with $\tau = 24$ h. Not surprisingly, assimilation of q reduces rmsd relative to the input data for this variable (yellow vs cyan curves in Fig. 23d), although rmsd for T and vorticity are also slightly reduced (Figs. 23b,c). In addition, reducing the truncation wavenumber leads to increased rmsd for all variables. The CIN configuration selected, with $\tau = 24$ h and T21 truncation, leads to the atmospheric state variables lying generally within the range of observational uncertainties as

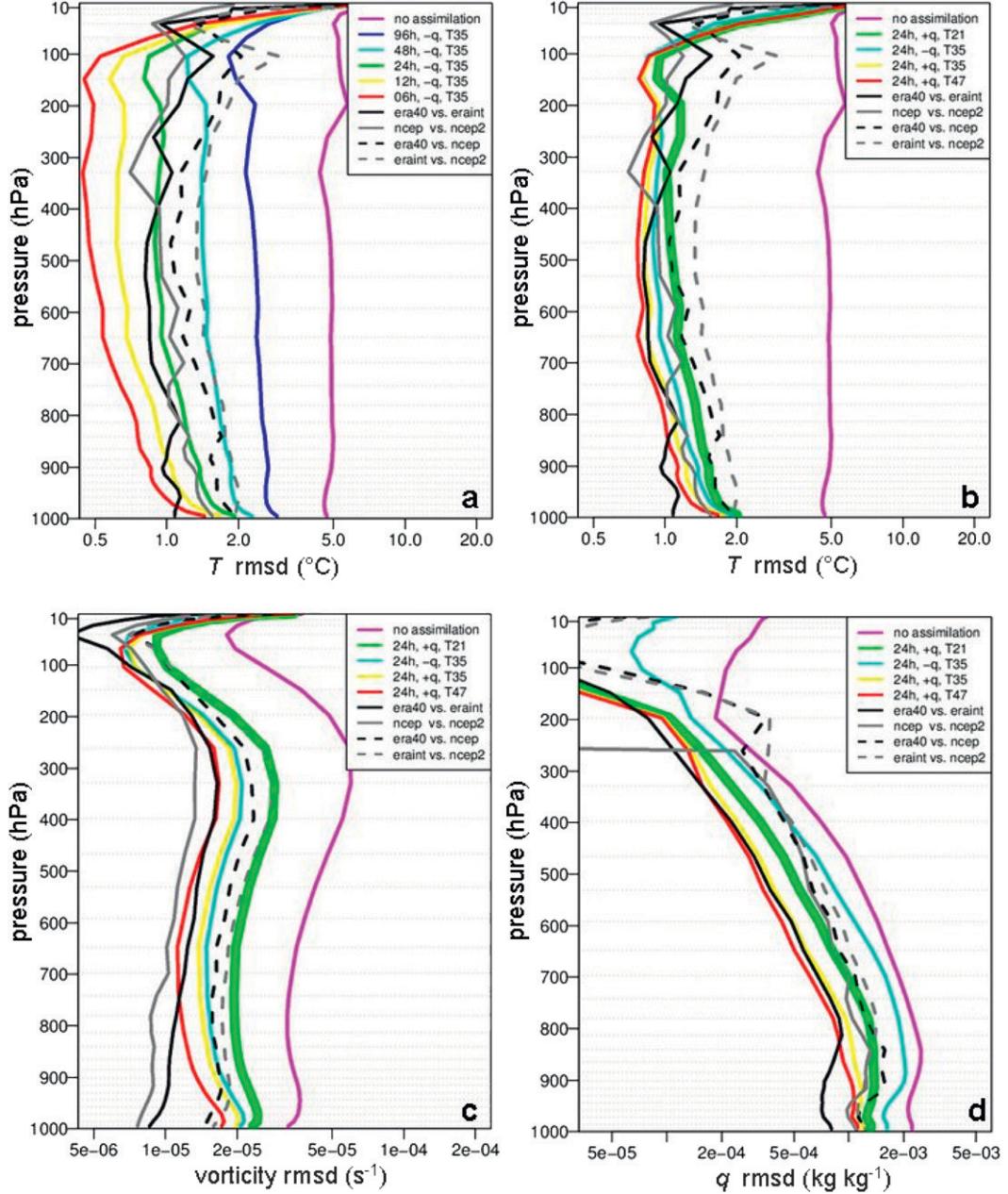


FIG. 23. Root-mean-square differences (rmsd) between CanCM3 assimilation runs and ERA-40 values for different assimilation parameters (colors), and between pairs of reanalyses (gray and black) as indicated in the legends. Shown are (a) dependence on the τ parameter in (1) for temperature with T35 truncation and no assimilation of q , and dependence on truncation wavenumber and q assimilation with τ fixed at 24 h for (b) temperature, (c) vorticity, and (d) q . Thick green curves indicate the assimilation parameters used for CanSIPS, and horizontal dotted lines indicate the pressure levels used in the model.

indicated by the thick green curves. (Different τ could potentially have been assigned for different variables, e.g., smaller τ for vorticity might be motivated by the relatively larger rmsd of that variable. However, a single τ was adopted, mainly for simplicity.) Additional considerations entering into this choice of CIN parameters

are that the spatiotemporal correlation between precipitation produced by the assimilating model and the GPCP2.1 and Xie–Arkin observational datasets (Adler et al. 2003; Xie and Arkin 1997) is optimized at values of about 0.5 near $\tau = 24$ h, and the ensemble spread as measured by rmsd between pairs of ensemble members

(not shown), while somewhat smaller than the model versus ERA-40 rmsd, remains similar in magnitude to the observational uncertainties.

Finally, it is useful to relate CIN, with $\tau = 24$ h and T21 low-pass spatial filtering of the increment $\Delta\mathbf{x}_a$ in (1), to standard relaxation, which may be expressed as

$$\frac{d\mathbf{x}}{dt} = M(\mathbf{x}) - \tau^{-1}[\mathbf{x}(t) - \mathbf{x}^R(t)], \quad (2)$$

where $\mathbf{x}(t)$ is the evolving model state and $\mathbf{x}^R(t)$ is the reference target state obtained by linearly interpolating 6-hourly reanalysis states to the time step of the model. If $\tau = 24$ h and T21 low-pass spatial filtering of the difference $\mathbf{x}(t) - \mathbf{x}^R(t)$ are employed in (2) as for CIN, then one finds little detectable difference between CIN and standard relaxation, unlike for standard IAU (Bloom et al. 1996). This has an obvious physical explanation: since CIN constrains only synoptic- to planetary-scale variations, and since these large-scale variations typically have temporal scales $\gg 6$ h, CIN can be viewed effectively as relaxation because it adequately discretizes these variations with 6-h increments. Subsequent to the development of CanSIPS, CIN and equivalent relaxation schemes have been used and compared in a variety of applications (e.g., Kharin and Scinocca 2012). In all instances, when the time scale τ and the low-pass filter have identical settings there are no discernible differences between the two approaches. For purposes of efficiency, this has led us to consider using relaxation in place of CIN in future versions of CanSIPS.

c. Land initialization

Land surface conditions including soil moisture and snow cover can act as persistent climatic influences, primarily on subseasonal time scales, that potentially can be exploited for climate prediction (e.g., Schlosser and Mocko 2003; Koster et al. 2011; Guo et al. 2011).

Under the CanSIPS assimilation procedure, the model land variables evolve according to meteorological influences provided by the assimilation of observations into the atmospheric component of the model. Thus, although the modeled land surface is not constrained directly by observational measurements, its evolution should bear some resemblance to reality.

An indication of this resemblance is provided by comparing land surface states in the assimilation run to a land surface analysis obtained by forcing CLASS, the land component of the CanSIPS models, offline with historical hydrometeorological forcings consisting of NCEP–NCAR reanalysis 1 fields adjusted to correct for biases relative to observational datasets as described in Alavi et al. (2011) and references therein. Figure 24

shows correlations between this analysis and a CanCM3 assimilation run for time series of monthly mean temperature (top left) and moisture (top right) anomalies in the uppermost soil layer when atmospheric assimilation is not applied. These correlations are very weak in this instance. By contrast, when atmospheric assimilation is applied the correlations are much larger (and higher for temperature than for moisture), indicating reasonable correspondence between CanSIPS land initial conditions and the offline analysis (middle panels of Fig. 24). Furthermore, root-mean-square errors relative to the offline analysis are considerably reduced when atmospheric assimilation is applied (bottom panels of Fig. 24). For CanCM4 assimilation runs, correlations are slightly higher (and RMSE lower) than for CanCM3 in the extratropics, whereas the opposite is true in the tropics (not shown).

d. Sea ice initialization

Sea ice concentration in CanSIPS assimilation runs is treated rather simply by relaxing⁵ modeled concentrations with a 3-day time scale toward daily observational values, which for the historical forecast period were obtained by linearly interpolating between monthly values from the HadISST1.1 dataset (Rayner et al. 2003).⁶ Differences between the modeled concentrations and input concentration data are much smaller than the freely running model biases shown in Fig. 11, typically exceeding ± 0.1 only in locations near the ice edge, and are almost identical in CanCM3 and CanCM4. This indicates that such differences likely arise as a result of inconsistencies between the assimilated sea ice and atmospheric forcing data rather than from the influence of model biases, which are quite different between the two models especially in summer and fall.

Mean sea ice thickness over a grid cell (expressed as mass per unit area) is another prognostic variable in the sea ice component of the CanSIPS models. This variable, for which observational information during the historical period is extremely sparse, is treated through relaxation, again with a 3-day time scale, to input values

⁵ While some form of CIN potentially could be applied to sea ice, implementation would necessarily differ from the atmospheric case due to the different assimilation window (24 vs 6 h), dynamics, and the fact that sea ice couples to the atmosphere only once per day. In addition, sea ice is not global field making spatial filtering problematic. Similar considerations apply to SST assimilation, described in section 4e(1).

⁶ These values tend to be less reliable before the start of satellite observations in 1979, particularly in the Antarctic. Although all CanSIPS historical seasonal forecasts are initialized in 1979 and afterward, earlier CanCM4 decadal predictions are initialized using less reliable pre-1979 sea ice time series.

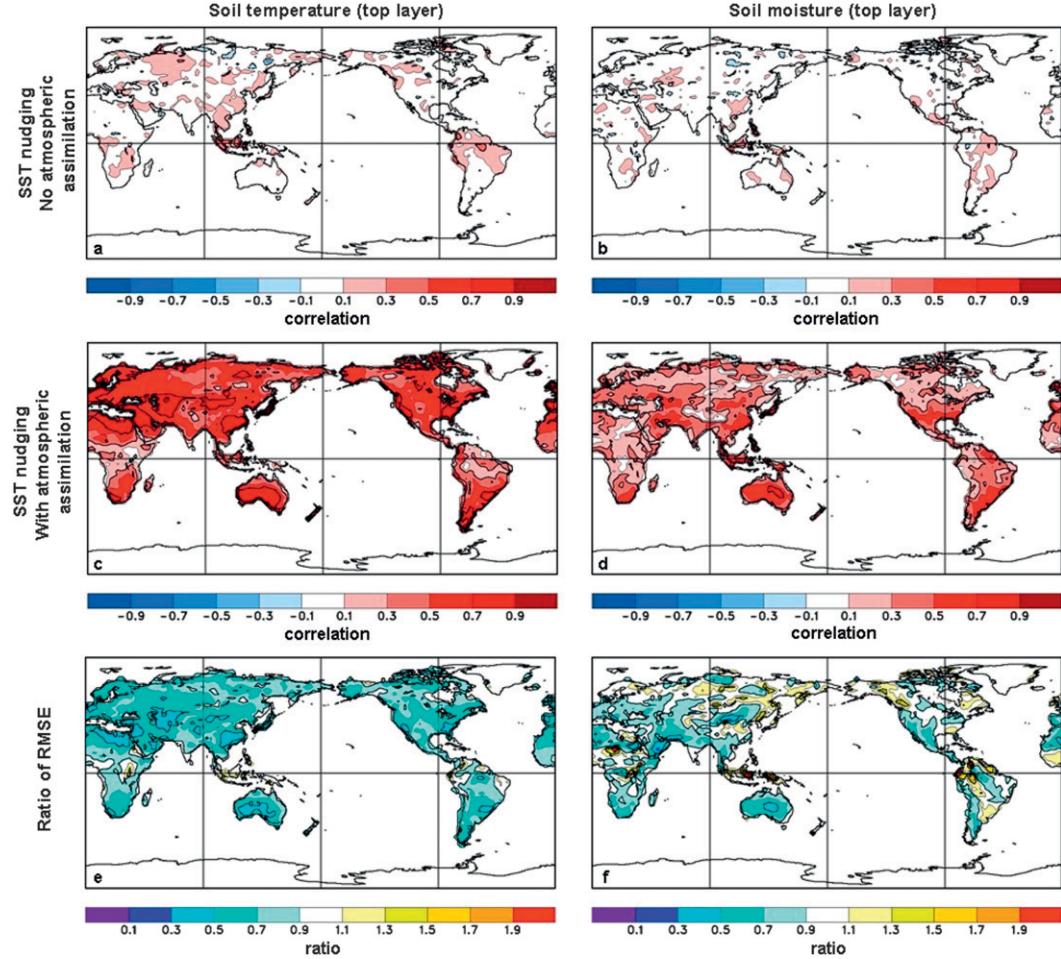


FIG. 24. Relations between an observation-based offline analysis and assimilation run values of monthly mean top-layer soil (a),(c),(e) temperature and (b),(d),(f) moisture. (a),(b) Correlations for a CanCM3 assimilation run having SST but no atmospheric assimilation. (c),(d) Correlations for a CanCM3 assimilation run having SST with atmospheric assimilation. (e),(f) Ratio of RMSE for the run with atmospheric assimilation to the run without.

determined very simply as follows. Where sea ice concentration exceeds a threshold value, mass per unit area is set to input values drawn from a seasonally varying climatology produced by a previous CCCma model version. This climatology reproduces some known features of the thickness distribution, including generally thicker ice in the Arctic than the Antarctic, and generally greater thicknesses, exceeding 3 m, in the western Arctic as compared to the eastern Arctic. This procedure has obvious shortcomings, for example it does not account for the expected reduction of ice mass as polar climate warms except through the reduction of ice area, and will be improved upon in future CanSIPS versions.

As for concentration, deviations between modeled and assimilated ice mass are relatively small, seldom exceeding $\pm 100 \text{ kg m}^{-3}$ (equivalent to $\approx 10\text{-cm}$ ice

thickness) except near the ice edge, and differ little between CanCM3 and CanCM4.

e. Ocean initialization

CanSIPS employs separate data assimilation procedures for the surface and interior ocean as detailed below.

1) SURFACE ASSIMILATION

During the assimilation runs the modeled SST field is relaxed with a 3-day time scale toward daily observational values, which for the historical period are obtained through interpolation of the monthly extended reconstructed SST (ERSST) analysis (Smith and Reynolds 2004), transitioning to the weekly OISST analysis (Reynolds et al. 2002) when the latter record begins in November 1982. (The two products are

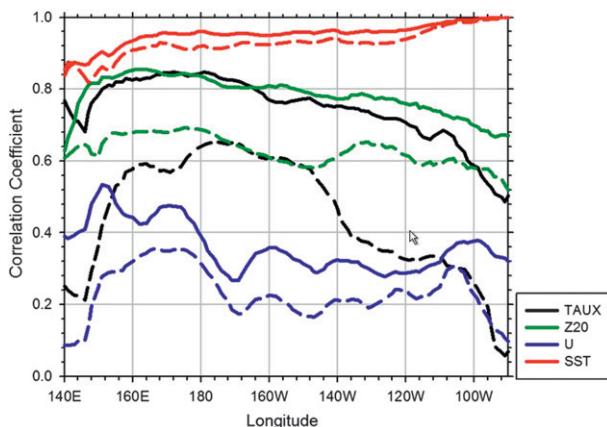


FIG. 25. Correlations of observational values of monthly mean equatorial Pacific zonal wind stress τ_x (black), 20°C isotherm depth Z20 (green), zonal surface current U (blue) and SST (red), and a CanCM3 assimilation run with SST and atmospheric assimilation (solid line), and with SST but no atmospheric assimilation (dashed line).

progressively blended at the transition point over the course of one week.)

Because reliable global time series for surface salinity over the historical period are not available, surface salinity in the assimilation runs is relaxed to annual mean PHC/WOA climatological values with a 180-day time scale, except beneath sea ice where this time scale is shortened to 3 days to prevent spurious salinity changes due to sea ice and SST being continually modified by assimilation.

The atmospheric assimilation procedure detailed above also plays a prominent role in the initialization of the ocean because the surface winds, as well as heat and freshwater fluxes, influence both the surface and interior state of the ocean. The improvement in the ocean state so realized is illustrated for the crucial equatorial Pacific region in Fig. 25, which shows as a function of longitude the correlations of 1972–2001 monthly mean SST, zonal wind stress τ_x , 20°C isotherm depth Z20, and zonal surface current U with corresponding observation-based time series [Atmospheric Model Intercomparison Project (AMIP) SST; Hurrell et al. (2008); SODA 1.4.2/ERA-40 τ_x , Carton and Giese (2008); SODA Global Ocean Analysis (GOA) beta 7 Z20 and U , Carton et al. (2000a,b)]. The solid curves represent correlations for a CanCM3 run that includes SST and atmospheric assimilation, and the dashed curves correlations from a comparable run that includes SST but not atmospheric assimilation. Although the run without atmospheric assimilation shows an appreciable ability to represent variations in τ_x , Z20 and U through the influence of SST on zonal winds as discussed by Keenlyside et al. (2005),

substantial improvements are realized when the atmospheric variables are assimilated as well. Even SST, which is assimilated in both runs, is improved by atmospheric assimilation. That initialization improvements realized through atmospheric assimilation lead to improved multiseasonal prediction skill in the equatorial Pacific and elsewhere is demonstrated by Balmaseda and Anderson (2009) and Dunstone and Smith (2010).

A further aspect of the atmospheric assimilation affecting the ocean is that differences in the atmospheric states between different ensemble members lead also to differences in the simulated ocean states due to variations in ocean surface forcings across the ensemble. Figure 26 shows the rmsd between two CanCM3 assimilation runs of 1972–2001 monthly mean potential temperatures at depths of 56 (top panel) and 510 m (bottom panel). The relatively large rmsd at shallower depths are a result of vertical temperature gradients being generally larger in the upper ocean (beneath the mixed layer) than at greater depths. This is because differences in wind stress lead, through differences in Ekman pumping and wind-generated planetary waves, to differing vertical displacements in the interior ocean. The largest resulting temperature rmsd thus occur where the vertical temperature gradients are large. This interpretation is supported by the occurrence of large rmsd in locations known to have particularly high upper-ocean vertical temperature gradients due to local upwelling. These include the eastern Pacific equatorial thermocline, the upper ocean beneath the intertropical convergence zone west of Central America, and the thermocline ridge of the southern tropical Indian Ocean. The North Atlantic is an additional region of high upper-ocean rmsd, and is essentially the only such region in the deeper ocean. Vertical gradients are not large there, but this region is characterized by strong wintertime deep convection. Because oceanic deep convection is a threshold process triggered by negative surface buoyancy fluxes, the large rmsd across a range of depths in this region are likely caused by differences in convective triggering, which can potentially arise from even small surface flux differences between ensemble members.

The ocean states in the assimilation run, although constrained by data only at the surface, compose a type of historical ocean analysis. One particularly interesting aspect of such analyses is the history of the simulated AMOC, as this bears upon the understanding and prediction of decadal climate variability, which is thought to be strongly influenced by variations in the AMOC and the associated meridional heat transports (e.g., Balmaseda et al. 2007). Figure 27 shows 1960–2010 annual mean AMOC transports at 26°N in the CanCM4

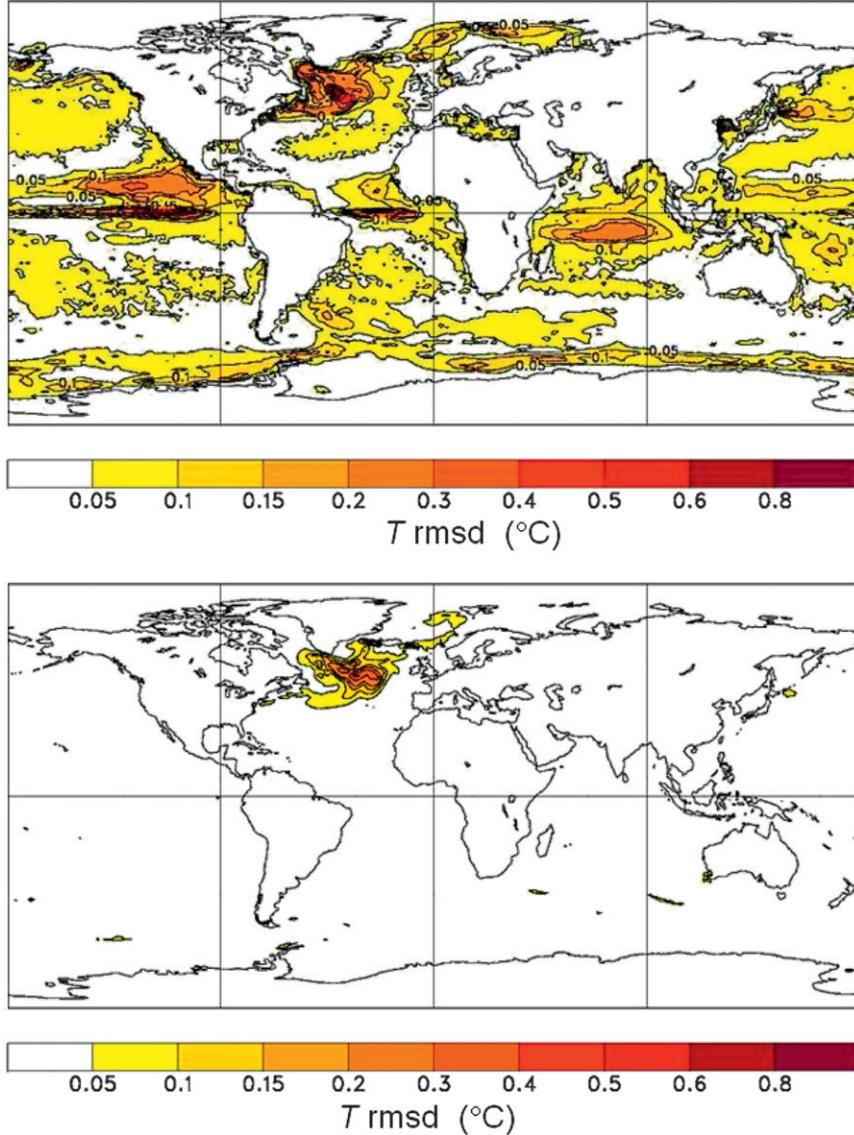


FIG. 26. Root-mean-square differences (rmsd) between monthly mean ocean potential temperatures in two CanCM3 assimilation runs, at depths of (top) 56 and (bottom) 510 m.

assimilation runs, which were more thoroughly spun up than for CanCM3 and serve to initialize CCCma's CMIP5 decadal forecasts. Ensemble mean transports are indicated by the solid line, and the ensemble spread (extreme values) by the shading. The main features are a pronounced increase in AMOC transport in the 1980s and early 1990s, followed by an even more pronounced decline between the late 1990s and 2010. This general pattern has been noted in other ocean analyses (Pohlmann et al. 2013), although the short observational record (e.g., Johns et al. 2011) does not conclusively indicate such a recent decline.

2) INTERIOR ASSIMILATION

The coupled assimilation procedures described above directly constrain the surface ocean but not its interior. Ocean interior assimilation in CanSIPS is through a simple offline procedure that is performed on the ocean state produced by the assimilation run, prior to the start of the forecasts, as indicated in the top panel of Fig. 22. The input data for this procedure are ocean temperature analyses, which for 1981–2010 CanSIPS historical forecasts are provided by monthly mean Global Ocean Data Assimilation System (GODAS; Behringer

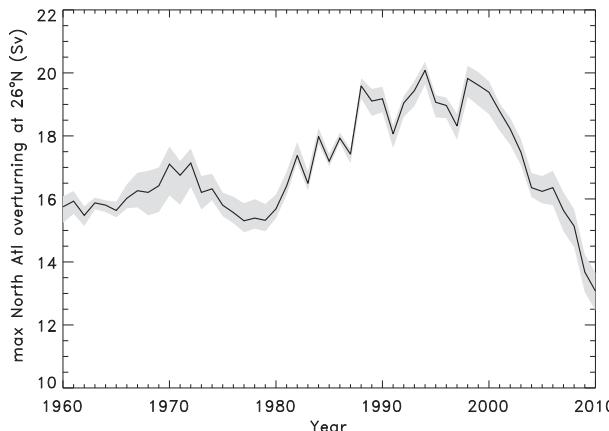


FIG. 27. Time series of annual mean AMOC transport at 26°N in Sv ($1 \text{ Sv} \equiv 10^6 \text{ m}^3 \text{ s}^{-1}$) in the ensemble of 10 CanCM4 assimilation runs. The solid line indicates the ensemble mean, and the shading the range of values across the ensemble.

et al. 1998) temperatures interpolated to the forecast start date.⁷

These observational temperatures are assimilated by means of the variational procedure described by Tang et al. (2004), which uses the background error covariances of Derber and Rosati (1989).⁸ Although variational schemes similar to this one can be used for online assimilation of individual measurements, this simplified application to gridded analyses avoids the complexities of acquiring and processing such measurements, while still being applicable in near-real time if such analyses from another source are available. Because assimilation of temperature alone can cause density inversions leading to spurious convective overturning, salinities in each grid column are adjusted to preserve the column temperature–salinity relationship and hence static stability as described in Troccoli et al. (2002). Despite their simplicity, these procedures were found to substantially increase ENSO predictive skill in ensembles of test forecasts performed during system development, with the Tang et al. (2004) and Troccoli et al. (2002) procedures contributing about equally to the increase in skill. ENSO prediction skill in the CanSIPS hindcasts has been found to be comparable with that of climate forecasts from other operational centers as will be reported in the second paper in this series.

⁷ For seasonal and decadal forecasts initialized at earlier dates before GODAS becomes available, anomalies from the SODA analysis, which dates from 1958, are added to GODAS climatological values.

⁸ These covariances are used over the full range of depths in CanCM3, whereas in CanCM4 they are tapered to zero below 400-m depth according to a Gaussian profile having a 500-m-scale length in order to reduce initialization-induced transients in CanCM4 decadal predictions.

f. Decadal forecast initialization

As mentioned above, the CanCM4 component of CanSIPS has been used to produce decadal hindcasts and forecasts for CMIP5, results from which are detailed elsewhere (Fyfe et al. 2011; Kharin et al. 2012; Kim et al. 2012; Goddard et al. 2012). Because the effectiveness of different approaches for initializing decadal forecasts is still an area of active research, three sets of decadal predictions were produced following the initial tier-1 CMIP5 experimental design (Taylor et al. 2012), with initialization every five years from the start of 1961 until the start of 2006 as summarized in Table 8.

The first set of forecasts, designated set A, was initialized as described above except that the interior ocean assimilation detailed in the previous subsection was not applied. This forecast set is labeled initialization method i2 in the “rip” identifier of the CMIP5 dataset.

The second set of forecasts, designated set B, was subject to the full CanSIPS initialization procedure including interior ocean assimilation. These predictions are thus similar to the CanCM4-based CanSIPS historical forecasts except that volcanic aerosol forcing during the forecast period is specified observationally (even though such forcing would not be known in the case of a true forecast) in accordance with the CMIP5 experimental design. This forecast set, labeled initialization method i1 in the rip identifier of the CMIP5 dataset, has been extended to initializations in all years from 1961 until the present and composes CCCma’s primary contribution to the CMIP5 decadal prediction experiments. In addition, it is contributing to a community effort, described in Smith et al. (2012), to produce multimodel decadal climate predictions in near-real time.

The third set of forecasts, designated set C, is similar to B except that interior ocean temperature anomalies rather than full field values were assimilated. This forecast set has not been submitted to CMIP5.

Comparisons between these three sets of predictions, not shown here, indicate global surface temperature skill over the first several seasons is highest in forecast set B (considering only the start dates available for sets A and C), whereas C becomes slightly more skillful than B after the first two years. Set A is the least skillful across the decadal forecast range, providing a further indication of the effectiveness of the interior ocean assimilation procedure described above.

5. Discussion and conclusions

This paper has described the models and initialization methods used by CanSIPS, Canada’s first coupled model-based operational climate prediction system,

TABLE 8. CanCM4 decadal prediction experiments (10 member ensembles) using different initialization methods as indicated.

CCCma designation	CMIP5 designation	Interior ocean assimilation	Start dates (1 Jan of specified years)
A	i2	None	1961–2005 (every 5 years)
B	i1	Full field	1961–2012 (every year)
C	—*	Anomaly	1961–2005 (every 5 years)

* Not submitted to CMIP5.

which was implemented in December 2011 at CMC. CanSIPS employs predictions from two versions of CCCma's coupled climate model, CanCM3 and CanCM4, in order to take advantage of the generally higher skill of multimodel predictions for a given ensemble size. [Although two is the minimum size for a multimodel ensemble, particularly large gains in skill can be realized in going from one to two models as demonstrated by Kharin et al. (2009).] In this respect, CanSIPS resembles previous multimodel two-tier dynamical prediction systems used at CMC, but is able to extend the dynamical forecast range from 4 to 12 months because of its ability to predict future ocean conditions.

Properties of freely running CanCM3 and CanCM4 including climate trends, biases in annual mean and seasonally varying climatology, and the representation of climate variability have been described. Importantly, the representation of ENSO, considered to be the largest source of multiseasonal predictability, is substantially improved in both models compared to previous CCCma model versions, although both model representations of the MJO are rather poor as in many other climate models. In general, biases and other model properties differ substantially between CanCM3 and CanCM4 and are often complementary. For example, CanCM3 has an ENSO that is somewhat too weak whereas ENSO is slightly too strong in CanCM4. Such differences serve to sample the space of model errors, a consideration sometimes cited heuristically to explain the empirical benefits of the multimodel approach (e.g., Stockdale et al. 2006).

The initialization methods employed by CanSIPS are relatively simple in relation to the range of assimilation methods that have been developed and are in use by prediction systems elsewhere. This is enabled in part by the availability of near-real-time analyses in the CMC forecasting environment, which allows CanSIPS to assimilate gridded data products rather than individual observations, itself a substantial simplification. The relatively simple means by which these products are assimilated into the CanSIPS models offer advantages in terms of computational efficiency, and their

ease of implementation has served to expedite system development.

Although CanSIPS performs reasonably well by current standards as indicated by its historical prediction skill (the subject of the next paper in this series), there remains substantial room for improvement. Avenues of improvement that may factor into future versions of CanSIPS include

- Improved ocean interior assimilation methodology, making use of the Nucleus for European Modeling of the Ocean (NEMO)-based operational analysis under development at CMC;
- A more direct initialization of land surface variables through assimilation of newly developed land surface analyses;
- Improved sea ice assimilation, particularly with respect to ice thickness;
- Model improvements leading to reduced biases and improved internal variability;
- One or more new forecast models including CMC's GEM global forecast model coupled to NEMO.

One promising aspect of NEMO-based ocean initialization of CanSIPS is that a NEMO-based ocean initialization has been shown to perform particularly well in relation to initializations with other ocean analyses in NCEP Climate Forecast System, version 2 (CFSv2), ENSO forecasts (Zhu et al. 2012).

In addition, recently developed methodologies promise to substantially reduce biases in the ocean (Thompson et al. 2006) and atmospheric (Kharin and Scinocca 2012) components of forecast models without degrading other aspects of forecast simulations such as the modeled variability.

A further potential avenue for system improvement is offered by CCCma's Canadian Earth System Model, version 2 (CanESM2), which employs the same physics as CanCM4 but includes terrestrial and ocean ecosystem components and an interactive carbon cycle. The use of CanESM2 for climate prediction could for example enable predictions of changes in ocean productivity associated with ENSO and other influences. Such a broadening of forecast products and applications is certain to occur as climate forecast models become increasingly comprehensive.

Acknowledgments. Valuable comments and advice were provided by Vivek Arora, Mike Lazare, and two anonymous reviewers. This work was supported in part by Canadian Foundation for Climate and Atmospheric Sciences funding to the Global Ocean–Atmosphere Prediction and Predictability Network (GOAPP).

REFERENCES

- AchutaRao, K., and K. R. Sperber, 2006: ENSO simulation in coupled ocean–atmosphere models: Are the current models better? *Climate Dyn.*, **27**, 1–15.
- Adler, R. A., and Coauthors, 2003: The version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeor.*, **4**, 1147–1167.
- Alavi, N., G. Drewitt, A. A. Berg, and W. J. Merryfield, 2011: Soil moisture initialization effects in the CCCMA GCM3: Relationship of seasonal climate forecast error to uncertainty in soil moisture initializations. *Atmos.–Ocean*, **49**, 179–188.
- Balmaseda, M. A., and D. Anderson, 2009: Impact of initialization strategies and observations on seasonal forecast skill. *Geophys. Res. Lett.*, **36**, L01701, doi:10.1029/2008GL035561.
- , G. C. Smith, K. Haines, D. Anderson, T. N. Palmer, and A. Vidard, 2007: Historical reconstruction of the Atlantic meridional overturning circulation from the ECMWF operational ocean reanalysis. *Geophys. Res. Lett.*, **34**, L23615, doi:10.1029/2007GL031645.
- , L. Ferranti, F. Molteni, and T. N. Palmer, 2010: Impact of 2007 and 2008 Arctic ice anomalies on the atmospheric circulation: Implications for long-range predictions. *Quart. J. Roy. Meteor. Soc.*, **136**, 1655–1664.
- Barker, H. W., J. N. S. Cole, J.-J. Morcrette, R. Pincus, P. Raisanen, K. von Salzen, and P. A. Vaillancourt, 2008: The Monte Carlo independent column approximation: An assessment using several global atmospheric models. *Quart. J. Roy. Meteor. Soc.*, **134**, 1463–1478.
- Behringer, D. W., M. Ji, and A. Leetmaa, 1998: An improved coupled model for ENSO prediction and implications for ocean initialization. Part I: The ocean data assimilation system. *Mon. Wea. Rev.*, **126**, 1013–1021.
- Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.*, **124**, 1256–1271.
- Bond, N. A., and G. A. Vecchi, 2003: The influence of the Madden–Julian oscillation on precipitation in Oregon and Washington. *Wea. Forecasting*, **18**, 600–613.
- Brodeau, L., B. Barnier, A.-M. Treguier, T. Penduff, and S. Gulev, 2010: An ERA40-based atmospheric forcing for global ocean circulation models. *Ocean Model.*, **31**, 88–104.
- Burgers, G., and D. B. Stephenson, 1999: The “normality” of El Niño. *Geophys. Res. Lett.*, **26**, 1027–1030.
- Carton, J. A., and B. S. Giese, 2008: A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). *Mon. Wea. Rev.*, **136**, 2999–3017.
- , G. Chepurin, X. Cao, and B. Giese, 2000a: A Simple Ocean Data Assimilation analysis of the global upper ocean 1950–1995. Part I: Methodology. *J. Phys. Oceanogr.*, **30**, 294–309.
- , —, and —, 2000b: A Simple Ocean Data Assimilation analysis of the global upper ocean 1950–1995. Part II: Results. *J. Phys. Oceanogr.*, **30**, 311–326.
- Chang, Y.-S., S. Zhang, A. Rosati, T. L. Delworth, and W. F. Stern, 2012: An assessment of oceanic variability for 1960–2010 from the GFDL ensemble coupled data assimilation. *Climate Dyn.*, **40**, 775–803, doi:10.1007/s00382-012-1412-2.
- Croft, B., U. Lohmann, and K. von Salzen, 2005: Black carbon ageing in the Canadian Centre for Climate Modelling and Analysis atmospheric general circulation model. *Atmos. Chem. Phys.*, **5**, 1931–1949.
- Cunningham, S. A., S. G. Alderson, B. A. King, and M. A. Brandon, 2003: Transport and variability of the Antarctic circumpolar current in Drake Passage. *J. Geophys. Res.*, **108**, 8084, doi:10.1029/2001JC001147.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim Reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597.
- DelSole, T., and X. Yang, 2011: Field significance of regression patterns. *J. Climate*, **24**, 5094–5107.
- Derber, J. D., and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr.*, **19**, 1333–1347.
- Derome, J., and Coauthors, 2001: Seasonal predictions based on two dynamical models. *Atmos.–Ocean*, **39**, 485–501.
- de Szoeke, S. P., C. W. Fairall, D. E. Wolfe, L. Bariteau, and P. Zuidema, 2010: Surface flux observations on the southeastern tropical Pacific Ocean and attribution of SST errors in coupled ocean–atmosphere models. *J. Climate*, **23**, 4152–4174.
- Doblas-Reyes, F. J., R. Hagedorn, T. N. Palmer, and J.-J. Morcrette, 2006: Impact of increasing greenhouse gas concentrations in seasonal ensemble forecasts. *Geophys. Res. Lett.*, **33**, L07708, doi:10.1029/2005GL025061.
- , and Coauthors, 2009: Addressing model uncertainty in seasonal and annual dynamical seasonal forecasts. *Quart. J. Roy. Meteor. Soc.*, **135**, 1538–1559.
- Donald, A., and Coauthors, 2006: Near-global impact of the Madden–Julian oscillation on rainfall. *Geophys. Res. Lett.*, **33**, L09704, doi:10.1029/2005GL025155.
- Donner, L. J., and Coauthors, 2011: The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *J. Climate*, **24**, 3484–3519.
- Dunstone, N. J., and D. M. Smith, 2010: Impact of atmosphere and sub-surface ocean data on decadal climate prediction. *Geophys. Res. Lett.*, **37**, L02709, doi:10.1029/2009GL041609.
- Fedorov, A. V., and S. G. Philander, 2001: A stability analysis of tropical ocean–atmosphere interactions: Bridging measurements and theory for El Niño. *J. Climate*, **14**, 3086–3101.
- Flato, G. M., and W. D. I. Hibler, 1992: Modelling pack ice as a cavitating fluid. *J. Phys. Oceanogr.*, **22**, 626–651.
- Fu, X., and B. Wang, 2001: A coupled modeling study of the seasonal cycle of the Pacific cold tongue. Part I: Simulation and sensitivity experiments. *J. Climate*, **14**, 765–779.
- Fyfe, J. C., N. P. Gillett, and D. W. J. Thompson, 2010: Comparing variability and trends in observed and modelled global-mean surface temperature. *Geophys. Res. Lett.*, **37**, L16802, doi:10.1029/2010GL044255.
- , W. J. Merryfield, V. Kharin, G. J. Boer, W.-S. Lee, and K. von Salzen, 2011: Skillful predictions of decadal trends in global mean surface temperature. *Geophys. Res. Lett.*, **38**, L22801, doi:10.1029/2011GL049508.
- Ganachaud, A., 2003: Large-scale mass transports, water mass formation, and diffusivities estimated from World Ocean Circulation Experiment (WOCE) hydrographic data. *J. Geophys. Res.*, **108**, 3213, doi:10.1029/2002JC001565.
- Gent, P. R., J. Willebrand, T. J. McDougall, and J. C. McWilliams, 1995: Parameterizing eddy-induced tracer transports in ocean circulation models. *J. Phys. Oceanogr.*, **25**, 463–474.
- Goddard, L., and S. Mason, 2002: Sensitivity of seasonal climate forecasts to persisted SST anomalies. *Climate Dyn.*, **19**, 619–631.
- , and Coauthors, 2012: A verification framework for interannual-to-decadal predictions experiments. *Climate Dyn.*, **40**, 245–272, doi:10.1007/s00382-012-1481-2.
- Griffies, S. M., and Coauthors, 2011: The GFDL CM3 coupled climate model: Characteristics of the ocean and sea ice simulations. *J. Climate*, **24**, 3520–3544.

- Guilyardi, E., 2006: El Niño–mean state–seasonal cycle interactions in a multi-model ensemble. *Climate Dyn.*, **26**, 329–348.
- , A. Wittenberg, A. Fedorov, M. Collins, C. Wang, A. Capotondi, and G. van Oldenborgh, 2009: Understanding El Niño in ocean–atmosphere general circulation models: Progress and challenges. *Bull. Amer. Meteor. Soc.*, **90**, 325–340.
- Guo, Z., P. A. Dirmeyer, and T. DelSole, 2011: Land surface impacts on subseasonal and seasonal predictability. *Geophys. Res. Lett.*, **38**, L24812, doi:10.1029/2011GL049945.
- Hansen, J. R., R. Ruedy, M. Sato, and K. Lo, 2010: Global surface temperature change. *Rev. Geophys.*, **48**, RG4004, doi:10.1029/2010RG000345.
- Hazeleger, W., and Coauthors, 2011: EC-Earth V2.2: Description and validation of a new seamless earth system prediction model. *Climate Dyn.*, **39**, 2611–2629, doi:10.1007/s00382-011-1228-5.
- Hurrell, J. W., J. Hack, D. Shea, J. Caron, and J. Rosinski, 2008: A new sea surface temperature and sea ice boundary dataset for the Community Atmosphere Model. *J. Climate*, **21**, 5145–5153.
- Johns, W. E., and Coauthors, 2011: Continuous, array-based estimates of Atlantic Ocean heat transport at 26.5°N. *J. Climate*, **24**, 2429–2449.
- Johnson, G. C., M. J. McPhaden, and E. Firing, 2001: Equatorial Pacific Ocean horizontal velocity, divergence, and upwelling. *J. Phys. Oceanogr.*, **31**, 839–849.
- Jones, C., 2000: Occurrence of extreme precipitation events in California and relationships with the Madden–Julian oscillation. *J. Climate*, **13**, 3576–3587.
- Joseph, R., and S. Nigam, 2006: ENSO evolution and teleconnections in IPCC's twentieth-century climate simulations: Realistic representation? *J. Climate*, **19**, 4360–4377.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.
- Kanamitsu, M., W. Ebisuzaki, J. Woolen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP-DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631–1643.
- Kara, A. B., P. A. Rochford, and H. E. Hurlburt, 2000: An optimal definition for ocean mixed layer depth. *J. Geophys. Res.*, **105**, 16 803–16 821.
- Keeley, S. P. E., R. T. Sutton, and L. C. Shaffrey, 2012: The impact of North Atlantic sea surface temperature errors on the simulation of North Atlantic European region climate. *Quart. J. Roy. Meteor. Soc.*, **138**, 1774–1783, doi:10.1002/qj.1912.
- Keenlyside, N. S., M. Latif, M. Botzet, J. Jungclaus, and U. Schulzweida, 2005: A coupled method for initializing El Niño Southern Oscillation forecasts using sea surface temperature. *Tellus*, **57A**, 340–356.
- Khairoutdinov, M., and Y. Kogan, 2000: A new cloud physics parameterization in a large-eddy simulation model of marine stratocumulus. *Mon. Wea. Rev.*, **128**, 229–243.
- Kharin, V. V., and J. F. Scinocca, 2012: The impact of model fidelity on seasonal predictive skill. *Geophys. Res. Lett.*, **39**, L18803, doi:10.1029/2012GL052815.
- , F. W. Zwiers, Q. Teng, G. J. Boer, J. Derome, and J. S. Fontecilla, 2009: Skill assessment of seasonal hindcasts from the Canadian Historical Forecast Project. *Atmos.–Ocean*, **47**, 204–223.
- , G. J. Boer, W. J. Merryfield, J. F. Scinocca, and W.-S. Lee, 2012: Statistical adjustment of decadal predictions in a changing climate. *Geophys. Res. Lett.*, **39**, L19705, doi:10.1029/2012GL052647.
- Kim, D., and Coauthors, 2009: Application of MJO simulation diagnostics to climate models. *J. Climate*, **22**, 6413–6436.
- Kim, H.-M., P. J. Webster, and J. A. Curry, 2012: Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts. *Geophys. Res. Lett.*, **39**, L10701, doi:10.1029/2012GL051644.
- Kirtman, B. P., and D. Min, 2009: Multimodel ensemble ENSO prediction with CCSM and CFS. *Mon. Wea. Rev.*, **137**, 2908–2930.
- Koster, R., and Coauthors, 2011: The second phase of the Global Land–Atmosphere Coupling Experiment: Soil moisture contributions to subseasonal forecast skill. *J. Hydrometeor.*, **12**, 805–822.
- Kug, J.-S., Y.-G. Ham, M. Kimoto, F.-F. Jin, and I.-S. Kang, 2010: New approach for optimal perturbation method in ensemble climate prediction with empirical singular vector. *Climate Dyn.*, **35**, 331–340.
- Landrum, L., M. M. Holland, D. Schneider, and E. Hunke, 2012: Antarctic sea ice climatology, variability, and late twentieth-century change in CCSM4. *J. Climate*, **25**, 4817–4838.
- Large, W. G., J. C. Williams, and S. C. Doney, 1994: Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Rev. Geophys.*, **32**, 363–403.
- , G. Danabasoglu, J. C. McWilliams, P. R. Gent, and F. O. Bryan, 2001: Equatorial circulation of a global ocean climate model with anisotropic horizontal viscosity. *J. Phys. Oceanogr.*, **31**, 518–536.
- Li, J., and H. W. Barker, 2005: A radiation algorithm with correlated k-distribution. Part I: Local thermal equilibrium. *J. Atmos. Sci.*, **62**, 286–309.
- Lienert, F., J. C. Fyfe, and W. J. Merryfield, 2011: Do climate models capture the tropical influences on North Pacific sea surface temperature variability? *J. Climate*, **24**, 6203–6209.
- Lin, H., and G. Brunet, 2009: The influence of the Madden–Julian oscillation on Canadian wintertime surface air temperature. *Mon. Wea. Rev.*, **137**, 2250–2262.
- , —, and J. Derome, 2009: An observed connection between the North Atlantic oscillation and the Madden–Julian oscillation. *J. Climate*, **22**, 364–380.
- , —, and R. Mo, 2010: Impact of the Madden–Julian oscillation on wintertime precipitation in Canada. *Mon. Wea. Rev.*, **138**, 3822–3839.
- Lin, J.-L., 2007: The double-ITCZ problem in IPCC AR4 coupled GCMs: Ocean–atmosphere feedback analysis. *J. Climate*, **20**, 4497–4525.
- , and Coauthors, 2006: Tropical intraseasonal variability in 14 IPCC AR4 climate models. Part I: Convective signals. *J. Climate*, **19**, 2665–2690.
- Liniger, M. A., H. Mathis, C. Appenzeller, and F. J. Doblas-Reyes, 2007: Realistic greenhouse gas forcing and seasonal forecasts. *Geophys. Res. Lett.*, **34**, L02402, doi:10.1029/2006GL028335.
- Livezey, R. E., 1990: Variability of skill of long-range forecasts and implications for their use and value. *Bull. Amer. Meteor. Soc.*, **71**, 300–309.
- Lohmann, U. K., and E. Roeckner, 1996: Design and performance of a new cloud microphysics scheme developed for the ECHAM general circulation model. *Climate Dyn.*, **12**, 557–572.
- , K. von Salzen, N. McFarlane, H. G. Leighton, and J. Feichter, 1999: Tropospheric sulfur cycle in the Canadian general circulation model. *J. Geophys. Res.*, **104**, 26 833–26 858.
- Luo, J.-J., S. Masson, S. Behera, S. Shingu, and T. Yamagata, 2005: Seasonal climate predictability in a coupled OAGCM using a different approach for ensemble forecasts. *J. Climate*, **18**, 4474–4497.

- Ma, X., K. von Salzen, and J. Cole, 2010: Constraints on interactions between aerosols and clouds on a global scale from a combination of MODIS-CERES satellite data and climate simulations. *Atmos. Chem. Phys.*, **10**, 9851–9861.
- Matthews, A. J., B. J. Hoskins, and M. Masutani, 2004: The global response to tropical heating in the Madden-Julian oscillation during northern winter. *Quart. J. Roy. Meteor. Soc.*, **130**, 1991–2011.
- McFarlane, N. A., J. F. Scinocca, M. Lazare, R. Harvey, D. Verseghy, and J. Li, 2005: The CCCma third generation atmospheric general circulation model. CCCma Internal Rep., 25 pp.
- Merryfield, W. J., 2006: Changes to ENSO under CO₂ doubling in a multi-model ensemble. *J. Climate*, **19**, 4009–4027.
- , W.-S. Lee, G. J. Boer, V. V. Kharin, B. Pal, J. F. Scinocca, and G. M. Flato, 2010: The first Coupled Historical Forecasting Project (CHFP1). *Atmos.–Ocean*, **48**, 263–283.
- Mo, K. C., and R. W. Higgins, 1998: Tropical convection and precipitation regimes in the western United States. *J. Climate*, **11**, 2404–2423.
- Monahan, A. H., and A. Dai, 2004: The spatial and temporal structure of ENSO nonlinearity. *J. Climate*, **17**, 3026–3036.
- Namias, J., 1964: A 5-yr experiment in the preparation of seasonal outlooks. *Mon. Wea. Rev.*, **92**, 449–464.
- , 1968: Long range weather forecasting—History, current status and outlook. *Bull. Amer. Meteor. Soc.*, **49**, 438–470.
- Palmer, T. N., 1999: Predicting uncertainty in forecasts of weather and climate. ECMWF Tech. Rep. 294, 51 pp. [Available online at http://www.ecmwf.int/news/events/training/rcourse_notes/pdf_files/Uncertainty_prediction.pdf.]
- Paolino, R. H., and Coauthors, 2010: The next generation of scenarios for climate change research and assessment. *Nature*, **463**, 747–756.
- Pohlmann, H., D. M. Smith, M. A. Balmaseda, N. S. Keenlyside, S. Masina, D. Matei, W. A. Müller, and P. Rogel, 2013: Predictability of the mid-latitude Atlantic meridional overturning circulation in a multi-model system. *Climate Dyn.*, doi:10.1007/s00382-013-1663-6, in press.
- Polavarapu, S., S. Ren, A. M. Clayton, D. Sankey, and Y. Rochon, 2004: On the relationship between incremental analysis updating and incremental digital filtering. *Mon. Wea. Rev.*, **132**, 2495–2502.
- , T. G. Shepherd, Y. Rochon, and S. Ren, 2005: Some challenges of middle atmosphere data assimilation. *Quart. J. Roy. Meteor. Soc.*, **131**, 3513–3527.
- Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 589–662.
- Rayner, D., and Coauthors, 2011: Monitoring the Atlantic meridional overturning circulation. *Deep-Sea Res.*, **58**, 1744–1753.
- Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, doi:10.1029/2002JD002670.
- Reichler, T., and J. Kim, 2008: How well do coupled models simulate today's climate? *Bull. Amer. Meteor. Soc.*, **89**, 303–311.
- Ren, X., X.-Q. Yang, B. Han, and G.-Y. Xu, 2007: North Pacific storm track variations in winter season and the coupled pattern with the mid-latitude atmosphere–ocean system. *Chin. J. Geophys.*, **50**, 94–103.
- 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.
- Roach, A. T., K. Aagaard, C. H. Pease, S. A. Salo, T. Weingartner, V. Pavlov, and M. Kulakov, 1995: Direct measurements of transport and water properties through the Bering Strait. *J. Geophys. Res.*, **100**, 18 443–18 457.
- Ropelewski, C. F., and M. S. Halpert, 1986: North American precipitation and temperature patterns associated with the El Niño/Southern Oscillation (ENSO). *Mon. Wea. Rev.*, **114**, 2352–2362.
- , and —, 1987: Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Mon. Wea. Rev.*, **115**, 1606–1626.
- Rotstayn, L. D., 1997: A physically based scheme for the treatment of stratiform clouds and precipitation in large-scale models. I: Description and evaluation of the microphysical processes. *Quart. J. Roy. Meteor. Soc.*, **123**, 1227–1282.
- Saha, S., and Coauthors, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 3483–3517.
- Saravanan, R., and P. Chang, 1999: Oceanic mixed layer feedback and tropical Atlantic variability. *Geophys. Res. Lett.*, **26**, 3629–3632.
- Schlosser, C. A., and D. M. Mocko, 2003: Impact of snow conditions in spring dynamical seasonal predictions. *J. Geophys. Res.*, **108**, 8616–8629.
- Scinocca, J. F., and N. McFarlane, 2004: The variability of modeled tropical precipitation. *J. Atmos. Sci.*, **61**, 1993–2015.
- , —, M. Lazare, and J. Li, 2008: The CCCma third generation AGCM and its extension into the middle atmosphere. *Atmos. Chem. Phys.*, **8**, 7055–7074.
- Seiki, A., and Y. N. Takayabu, 2007: Westerly wind bursts and their relationship with intraseasonal variations and ENSO. Part I: Statistics. *Mon. Wea. Rev.*, **135**, 3325–3345.
- Shabbar, A., 2006: The impact of El Niño-Southern Oscillation on the Canadian climate. *Advances in Geophysics*, Vol. 6, Academic Press, 149–153.
- , and A. G. Barnston, 1996: Skill of seasonal climate forecasts in Canada using canonical correlation analysis. *Mon. Wea. Rev.*, **124**, 2370–2385.
- , B. Bonsal, and M. Khandekar, 1997: Canadian precipitation patterns associated with the Southern Oscillation. *J. Climate*, **10**, 3016–3027.
- Simmons, H. L., S. R. Jayne, L. C. S. Laurent, and A. J. Weaver, 2004: Tidally driven mixing in a numerical model of the ocean general circulation. *Ocean Model.*, **6**, 245–263.
- Smith, D. M., A. A. Scaife, and B. P. Kirtman, 2012: What is the current state of scientific knowledge with regard to seasonal and decadal forecasting? *Environ. Res. Lett.*, **7**, 015602, doi:10.1088/1748-9326/7/1/015602.
- Smith, T. M., and R. Reynolds, 2004: Improved extended reconstruction of SST (1854–1997). *J. Climate*, **17**, 2466–2477.
- Sprintall, J., S. E. Wijffels, R. Molcard, and I. Jaya, 2009: Direct estimates of the Indonesian throughflow entering the Indian Ocean: 2004–2006. *J. Geophys. Res.*, **114**, C07001, doi:10.1029/2008JC005257.
- Steele, M., R. Morley, and W. Ermold, 2001: PHC: A global ocean hydrography with a high-quality Arctic Ocean. *J. Climate*, **14**, 2079–2087.
- Sterl, A., and Coauthors, 2012: A look at the ocean in the EC-Earth climate model. *Climate Dyn.*, **39**, 2631–2657.
- Stockdale, T. N., M. A. Balmaseda, and A. Vidard, 2006: Tropical Atlantic SST prediction with coupled ocean–atmosphere GCMs. *J. Climate*, **19**, 6047–6061.
- Talley, L. D., 1999: Some aspects of ocean heat transport by the shallow, intermediate and deep overturning circulations.

- Mechanisms of Global Climate Change at Millennial Time Scales*, *Geophys. Monogr.*, Vol. 112, Amer. Geophys. Union, 1–22.
- Tang, Y., R. Kleeman, A. M. Moore, J. Vialard, and A. Weaver, 2004: An off-line, numerically efficient initialization scheme in an oceanic general circulation model for El Niño-Southern Oscillation prediction. *J. Geophys. Res.*, **109**, C05014, doi:10.1029/2003JC002159.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498.
- Thompson, D. W. J., J. M. Wallace, P. D. Jones, and J. J. Kennedy, 2009: Identifying signatures of natural climate variability in time series of global-mean surface temperature: Methodology and insights. *J. Climate*, **22**, 6120–6141.
- Thompson, K. R., D. G. Wright, Y. Lu, and E. Demirov, 2006: A simple method for reducing seasonal bias and drift in eddy resolving ocean models. *Ocean Model.*, **8**, 109–125.
- Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski, 1998: Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *J. Geophys. Res.*, **103**, 14 291–14 324.
- Troppoli, A., and Coauthors, 2002: Salinity adjustments in the presence of temperature data assimilation. *Mon. Wea. Rev.*, **130**, 89–102.
- Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012.
- Vecchi, G. A., and N. A. Bond, 2004: The Madden-Julian oscillation (MJO) and northern high latitude wintertime surface air temperatures. *Geophys. Res. Lett.*, **31**, L04104, doi:10.1029/2003GL018645.
- Verseghy, D. L., 2000: The Canadian Land Surface Scheme (CLASS): Its history and future. *Atmos.–Ocean*, **38**, 1–13.
- von Salzen, K., N. A. McFarlane, and M. Lazare, 2005: The role of shallow convection in the water and energy cycles of the atmosphere. *Climate Dyn.*, **25**, 671–688.
- , and Coauthors, 2013: The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4). Part I: Physical processes. *Atmos.–Ocean*, **51**, 104–125, doi:10.1080/07055900.2012.755610.
- Wagner, A. J., 1989: Medium- and long-range forecasting. *Wea. Forecasting*, **4**, 413–426.
- Waliser, D., and Coauthors, 2009: MJO simulation diagnostics. *J. Climate*, **22**, 3006–3030.
- Wang, Z., C.-P. Chang, and B. Wang, 2007: Impacts of El Niño and La Niña on the U.S. climate during northern summer. *J. Climate*, **20**, 2165–2177.
- Wheeler, M., and G. N. Kiladis, 1999: Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber-frequency domain. *J. Atmos. Sci.*, **56**, 374–399.
- , and H. H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, **132**, 1917–1932.
- Wittenberg, A. T., A. Rosati, N.-C. Lau, and J. J. Ploshay, 2006: GFDL's CM2 global coupled climate models. Part III: Tropical Pacific climate and ENSO. *J. Climate*, **19**, 698–722.
- Wu, A., and W. W. Hsieh, 2004: The nonlinear Northern Hemisphere winter atmospheric response to ENSO. *Geophys. Res. Lett.*, **31**, L02203, doi:10.1029/2003GL018885.
- Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, **78**, 2539–2558.
- Yang, H., and F. Wang, 2009: Revisiting the thermocline depth in the equatorial Pacific. *J. Climate*, **22**, 3856–3863.
- Yang, S.-C., E. Kalnay, M. Cai, M. Rienecker, G. Yuan, and Z. Toth, 2006: ENSO bred vectors in coupled ocean–atmosphere general circulation models. *J. Climate*, **19**, 1422–1436.
- Yang, X., and T. DelSole, 2012: Systematic comparison of ENSO teleconnection patterns between models and observations. *J. Climate*, **25**, 425–446.
- Yoder, J. A., and M. A. Kennelly, 2003: Seasonal and ENSO variability in global ocean phytoplankton chlorophyll derived from 4 years of SeaWiFS measurements. *Global Biogeochem. Cycles*, **17**, 1112, doi:10.1029/2002GB001942.
- Yu, B., and G. J. Boer, 2006: The variance of sea surface temperature and projected changes with global warming. *Climate Dyn.*, **26**, 801–821.
- Zaharieva, K., J. R. Christian, and K. L. Denman, 2008: Pre-industrial, historical, and fertilization simulations using a global ocean carbon model with new parameterizations of iron limitation, calcification, and N2 fixation. *Prog. Oceanogr.*, **77**, 56–82.
- Zheng, Y., T. Shinoda, J.-L. Lin, and G. N. Kiladis, 2011: Sea surface temperature biases under the stratus cloud deck in the southeast Pacific Ocean in 19 IPCC AR4 coupled general circulation models. *J. Climate*, **24**, 4139–4164.
- Zhou, S., M. L'Heureux, S. Weaver, and A. Kumar, 2012: A composite study of the MJO influence on the surface air temperature and precipitation over the continental United States. *Climate Dyn.*, **38**, 1459–1471.
- Zhu, J., B. Huang, L. Marx, J. L. Kinter III, M. A. Balmaseda, R.-H. Zhang, and Z.-Z. Hu, 2012: Ensemble ENSO hindcasts initialized from multiple ocean analyses. *Geophys. Res. Lett.*, **39**, L09602, doi:10.1029/2012GL051503.