Initialization and Potential Predictability of Soil Moisture in the Canadian Seasonal to Interannual Prediction System

REINEL SOSPEDRA-ALFONSO AND WILLIAM J. MERRYFIELD

Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada, Victoria, British Columbia, Canada

(Manuscript received 18 October 2017, in final form 15 February 2018)

ABSTRACT

The initialization and potential predictability of soil moisture in CanCM4 hindcasts during 1981–2010 is assessed. CanCM4 is one of the two global climate models employed by the Canadian Seasonal to Interannual Prediction System (CanSIPS) providing operational multiseasonal forecasts for Environment and Climate Change Canada (ECCC). Soil moisture forecast initialization in CanSIPS is determined by the response of the land component to forcing from data-constrained model atmospheric fields. We evaluate hindcast initial conditions for soil moisture and its atmospheric forcings against observation-based datasets. Although model values of soil moisture variability compare relatively well with a blend of two reanalysis products, there is significant disagreement in the tropics and arid regions linked to biases in precipitation, as well as in snow-covered regions, likely the result of biases in the timing of snow onset and melt. The temporal variance of initial soil moisture anomalies is typically larger in regions of considerable precipitation variability and in cold continental areas of shallow soil depth. Appreciable variance of initial conditions, combined with persistence of the initial anomalies and the model's ability to represent future climate variations, lead to potentially predictable soil moisture variance exceeding 60% of the total variance for up to 3–4 months in the tropics and 6–7 months in the mid- to high latitudes during hemispheric winter. Potential predictability at longer leads is primarily found in the tropics and extratropical areas of ENSO-teleconnected influences. We use lagged partial correlations to show that ENSO-teleconnected precipitation in CanCM4 is a likely source of potential predictability of soil moisture up to 1-yr lead in CanSIPS hindcasts.

1. Introduction

Soil moisture is the water stored between the ground surface and the water table, below which the soil is saturated (Bonan 2008). It plays a central role in the land surface water and energy balances through evapotranspiration, modulating thermal properties at the land–air interface and returning up to 60% of precipitation over land to the atmosphere (Seneviratne et al. 2010). Dry conditions resulting from soil moisture deficit have been

shown to amplify climate extremes, such as heat waves and droughts (Oglesby and Erickson 1989; Hirschi et al. 2011; Granier et al. 2007; Fischer et al. 2007). Soil moisture also contributes to surface and subsurface runoff and thus, the occurrence of floods (Grillakis et al. 2016), and it has a strong influence on vegetation productivity and the terrestrial carbon cycle (Falloon et al. 2011). Accurate estimation of soil moisture is therefore important for weather and climate prediction, water budget analysis, climate extreme predictions, streamflow prediction, predictions of wildfire activity, and so forth.

The temporal behavior of soil moisture is primarily determined by rainfall, evapotranspiration, and runoff. High soil moisture values can alter the properties of the land–atmosphere boundary layer through latent heat flux, creating favorable conditions for precipitation, which in turn feeds back on soil moisture. By contrast, low soil moisture values limit latent heat flux, making more energy available for sensible heating and thus increasing near-surface air temperature, leading to lower

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-17-0707.s1>.

Corresponding author: Reinel Sospedra-Alfonso, reinel.sospedra-alfonso@canada.ca

DOI: 10.1175/JCLI-D-17-0707.1

relative humidity that enhances evaporation and tends to further lower soil moisture. In regions with seasonal snow cover, snowmelt also plays an important role in influencing soil moisture, mainly during spring and into summer. Snow cover furthermore acts as an insulator that can effectively block the water exchange between the atmosphere and soil surface, thus contributing to the persistence of soil moisture anomalies from the early stages of the snow season. These forms of land–atmosphere feedback (Koster et al. 2004; Seneviratne et al. 2010; Miralles et al. 2012), combined with the persistence of deep soil moisture anomalies from weeks to several months (Entin et al. 2000; Koster and Suarez 2001; Seneviratne et al. 2006; Orth and Seneviratne 2012), contribute both regionally and globally to climate variability and predictability from subseasonal to decadal time scales (Wang and Kumar 1998; Schär et al. 1999; Conil et al. 2009; Koster et al. 2010, 2011; van den Hurk et al. 2012; Guo et al. 2012; Materia et al. 2014; Bellucci et al. 2015).

At subseasonal to seasonal time scales, realistic initialization of soil moisture has been shown to enhance forecast skill of near-surface climate over land. With the forecast systems participating in the second phase of the Global Land–Atmosphere Coupling Experiment (GLACE-2), Koster et al. (2010) showed that realistic soil moisture initialization contributed to temperature and precipitation forecast skill over North America out to 2 months and 45 days, respectively. Drewitt et al. (2012) carried out a detailed study of land surface initialization effects on forecast skill by using one of the systems participating in GLACE-2. Kanamitsu et al. (2003) employed the National Centers for Environmental Prediction (NCEP) seasonal forecast system to examine the effect of land surface initial conditions on predictability of soil moisture itself. They found that realistic initialization of soil moisture led to high predictive skill over arid and semiarid regions, partly because of the good representation of land surface evaporation, which is the main driver of soil moisture in these regions. Soil moisture initialization was also shown to impart predictive skill over continental temperate zones, although skill in these regions was significantly lower. Kanamitsu et al. (2003) also emphasized the predictive value of soil moisture initialization for near-surface temperature forecasts. The effects of climate on soil moisture predictability were recently examined by Nicolai-Shaw et al. (2016), who used soil moisture estimates from the European Centre for Medium-Range Weather Forecasts (ECMWF) interim land reanalysis (ERA-Interim/Land) and a blend of remotely sensed data to show that long-term predictability up to 5 months' lead time is linked to variations of the North Atlantic Oscillation (NAO), the Southern

Oscillation index (SOI), and the Antarctic Oscillation (AAO), depending on the regions of influences of these teleconnection indices, whereas persistence is the dominant source of soil moisture predictability at subseasonal time scales.

In this work, we examine the initialization and potential predictability (PP; defined in section 2c below) of the volumetric fraction of soil moisture (VFSM [$\text{m}^3 \text{m}^{-3}$], defined by the fraction of liquid + frozen water in the soil volume) in the Canadian Seasonal to Interannual Prediction System (CanSIPS), which produces operational ensemble multiseasonal forecasts for Environment and Climate Change Canada (ECCC) using two global climate models, CanCM3 and CanCM4 (Merryfield et al. 2013). First, we assess CanSIPS assimilation runs, which provide the forecast initial conditions, described in section 2a. We examine the temporal variability of the initial conditions of soil moisture and its main atmospheric forcing fields [i.e., precipitation (P) and potential evapotranspiration (PET)], evaluated against observation-based datasets. Then, we examine sources and behavior of VFSM potential predictability in CanSIPS retrospective forecasts (hindcasts) over 1981–2010. The role of El Niño–Southern Oscillation (ENSO) teleconnections on the PP of VFSM is discussed. In addition to our global results, we provide a detailed analysis for four grid locations representative of different climate, water content, and soil moisture–atmosphere interaction regimes across North America. The analysis of representative grid locations is useful, as it allows for a clearer description of the mechanisms responsible for predictability. This work is aligned with our previous studies on the representation of hydrological land surface variables in CanSIPS (Sospedra-Alfonso et al. 2016a,b), where we showed that hindcasts' skill of snow water equivalent (SWE) at short lead times is largely determined by reasonable model initialization combined with the persistence of initial anomalies, whereas long lead skill is mainly the result of SWE response to ENSO variability combined with the ability of the forecast to predict ENSO. Here, we show that both persistence of initial anomalies and ENSO–teleconnected effects are sources of potential predictability in CanSIPS soil moisture hindcasts.

The remainder of this paper is organized as follows. In section 2a, we provide a brief overview of CanSIPS. Section 2b describes the observation-based datasets used to evaluate CanSIPS assimilation runs. The ensemble method for calculating potential predictability is described in section 2c. Characterization and evaluations of soil moisture and atmospheric forcing fields' initial conditions are discussed in section 3. The behavior and sources for potential predictability of soil moisture in CanSIPS, including the influences of

ENSO teleconnections on VFSM, are examined locally in [section 4](#) and globally in [section 5](#). A summary and conclusions are presented in [section 6](#).

2. Data and methods

a. Model data and overview of CanSIPS

We provide here a short description of the CanSIPS models and the methods used to initialize them; additional details are provided in [Merryfield et al. \(2013\)](#). CanSIPS currently is based on two global climate models developed at the Canadian Centre for Climate Modeling and Analysis (CCCma), CanCM3 and CanCM4, and it provides Environment and Climate Change Canada's operational seasonal forecasts. These models have the same ocean, land surface, and sea ice components but different atmospheric components. Atmospheric and land variables are represented on a T63 Gaussian grid, equivalent to a horizontal resolution of approximately 2.8° in longitude and latitude.

CanSIPS assimilation runs, 10 for each model, are constrained by observational values of atmospheric temperature, specific humidity, and horizontal winds on spatial scales larger than about 1000 km, as well as by sea surface temperature and sea ice concentration. Different assimilation runs, begun from different model states and spun up for more than 30 years in the case of CanCM3 and more than 300 years in the case of CanCM4, provide slightly differing hindcast initial conditions for each CanSIPS ensemble member. Land initial conditions are obtained from the response of the land component, the Canadian Land Surface Scheme (CLASS) version 2.7, to the data-constrained atmospheric model fields. Because the atmospheric fields are unconstrained on horizontal scales smaller than approximately 1000 km, the differences between ensemble members represent observational uncertainties and provide ensemble spread in the land initial conditions. CanSIPS hindcasts run freely for 12 months from the initial conditions provided by the assimilation runs at the start of each month during the period 1981–2010. For simplicity, we examine the 10 ensemble assimilation and hindcast runs from CanCM4 only, which is known to predict ENSO skillfully at seasonal to annual lead times ([Barnston et al. 2018](#)).

To characterize soil moisture in CanSIPS, we employ CanCM4 monthly and daily averages from assimilation and hindcast runs, respectively, of liquid + frozen volumetric soil moisture fraction in the three soil layers of CLASS 2.7, having a combined depth of up to 410 cm. Also, we consider daily (monthly) precipitation and near-surface air temperature from hindcast (assimilation) runs,

as well as monthly Niño-3.4 index hindcasts, to evaluate soil moisture initialization and to examine sources and mechanisms of soil moisture potential predictability in CanCM4. Near-surface air temperature is used to compute potential evapotranspiration by means of the Thornthwaite method ([Thornthwaite 1948](#)), corrected with the sunshine hours at a given latitude ([Rosenberg et al. 1983](#)). The precipitation evapotranspiration index $PEI = P - PET$ (e.g., [Vicente-Serrano et al. 2010](#)) is used to characterize the geographic distribution of water surplus or deficit at the start of the hindcasts.

Although Thornthwaite PET estimates are known to be biased over arid or semiarid regions, this method has been widely used for the Arctic, Boreal regions, and tropical monsoon regions largely for its simplicity (e.g., [Kumar et al. 1987](#); [Fisher et al. 2011](#); [McMahon et al. 2013](#)), as it requires only air temperature as input data. Other methods may require net radiation, atmospheric pressure, wind speed, humidity, and so forth, and are also prone to inaccuracies depending on climate, landscape, and vegetation cover. Despite its simplicity, the Thornthwaite method has been shown to perform relatively well over subtropical, temperate, and temperate–continental regions when compared to field measurements, lysimeter observations, or other baseline methods ([Fisher et al. 2011](#); [McMahon et al. 2013](#)). Here, we do not attempt to provide a detailed account of PET in CanSIPS, but rather gain insight into soil moisture potential predictability and its possible connection with the main atmospheric variables: precipitation and temperature. For a regional and more detailed analysis, and particularly for short time scales, other methods for computing potential evapotranspiration may be preferred ([Kumar et al. 1987](#); [Grace and Quick 1988](#); [Fisher et al. 2011](#)). Refer to [McMahon et al. \(2013\)](#) for a comprehensive summary of different methods to estimate actual and potential evaporation and evapotranspiration using meteorological data.

b. Observation-based datasets for evaluation of CanSIPS initial conditions

We employ several observation-based datasets for evaluation of CanCM4 assimilation runs. Monthly VFSM is derived from a blend (done by a simple arithmetic average) of two satellite-era reanalysis products (hereafter called Blended-2): ERA-Interim/Land ([Balsamo et al. 2015](#)) and version 2 of the National Aeronautics and Space Administration (NASA) Modern-Era Retrospective Analysis for Research and Application (MERRA-2; [Gelaro et al. 2017](#)). We use a blended dataset to reduce observational uncertainty and structural uncertainty in the land surface models employed by these reanalyses.

ERA-Interim/Land uses a land surface model driven with meteorological data from ERA-Interim (Dee et al. 2011) that is corrected with monthly precipitation from NASA's Global Precipitation Climatology Project (GPCP; Adler et al. 2003) and assimilates observations of near-surface air temperature and humidity. MERRA-2 constrains meteorological fields from version 5 of the Goddard Earth Observing System (GEOS-5) atmospheric general circulation model through assimilation of observed pressure, temperature, wind components, and height from satellites and other, more direct measurement sources, such as ships, buoys, balloons, and aircraft. Unlike its predecessor, MERRA (Rienecker et al. 2011), MERRA-2 includes the assimilation of gauge- and satellite-based precipitation observations. Neither ERA-Interim/Land nor MERRA-2 assimilates land surface data.

Although there are clear limitations in the use of reanalysis data to evaluate model soil moisture outputs, there is the advantage of using spatially and temporally complete gridded datasets that have been previously validated. Reichle et al. (2017) validated surface and root soil moisture from both MERRA-2 and ERA-Interim/Land for a variety of land surface characteristics and climatological conditions by using over 550 in situ measurement sites from several networks. They showed that regardless of the many differences in both reanalysis products, their performances against the in situ observations are mutually consistent and generally good, with relatively low unbiased root-mean-square error surface soil moisture ($0.053 \text{ m}^3 \text{ m}^{-3}$ for MERRA-2 and $0.056 \text{ m}^3 \text{ m}^{-3}$ for ERA-Interim/Land) and relatively high correlations (0.68 for both) and anomaly correlations (0.62 for both) averaged over all sites. The two products do differ slightly in their root-zone soil moisture biases, with MERRA-2 having $0.016 \text{ m}^3 \text{ m}^{-3}$ and ERA-Interim/Land having $0.048 \text{ m}^3 \text{ m}^{-3}$. Because the number and thickness of soil layers of these data products and CLASS are different, here, VFSM is derived with different maximum soil column depths, though actual values are computed with soil depths that vary spatially.

Monthly values of near-surface air temperature over land, which are used to derive PET with the Thornthwaite method, are from the Berkeley Earth dataset (<http://berkeleyearth.org/data/>), specifically its $1^\circ \times 1^\circ$ gridded reconstruction of land surface air temperature records from approximately 37 000 stations. Monthly precipitation is from version 2.3 of NASA's GPCP (Adler et al. 2003; data available at <http://gpcp.umd.edu/>), which combines observations and satellite data on a $2.5^\circ \times 2.5^\circ$ global grid. All observation-based datasets were re-gridded to the $\approx 2.8^\circ$ CanSIPS resolution.

c. Ensemble method for potential predictability

We follow the methodology described by Sospedra-Alfonso et al. (2016a) and employ one-way ANOVA (von Storch and Zwiers 1999) with the ensemble method (DelSole et al. 2013) to estimate the potentially predictable variance fraction in CanSIPS soil moisture hindcasts. We consider the ensemble of $E = 10$ soil moisture linearly detrended hindcasts $X_{ey}(t)$ over a $Y = 30$ -yr period (1981–2010), with $e = 1, \dots, E$ and $y = 1, \dots, Y$, and having lead time $t = 0, \dots, 364$ in days or $t = 0, \dots, 11$ in months. At a given lead time t , the total variance of $X_{ey}(t)$ can be decomposed as $\sigma_T^2(t) = \sigma_P^2(t) + \sigma_N^2(t)$, where the noise variance

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

represents the unpredictable weather fluctuations characterized by the ensemble spread about the means $\bar{X}_y(t) = \sum_{e=1}^E X_{ey}(t)/E$, and the potentially predictable variance $\sigma_P^2(t) = \sigma_S^2(t) - \sigma_N^2(t)/E$ is related to the variance of the ensemble means

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

about the climatological grand ensemble mean $\bar{X}(t) = \sum_{e=1}^E \sum_{y=1}^E X_{ey}(t)/EY$. We characterize the potential predictability of soil moisture by the variance fraction

$$\text{PP}(t) = \sigma_P^2(t)/\sigma_T^2(t), \quad (3)$$

which, for 10 ensemble members and a 30-yr period, is statistically significant at the 95% confidence level when $\text{PP}(t) > 0.048$ (Sospedra-Alfonso et al. 2016a; Rowell 1998). A variance fraction $\text{PP}(t) = 0$ indicates that the total variance is determined by the unpredictable weather noise [Eq. (1)], whereas $\text{PP}(t)$ values close to 1 indicate that the total variance cannot be explained by the unpredictable component alone, implying potential for predictability.

Also following Sospedra-Alfonso et al. (2016a), we estimate the persistence of the initial soil moisture anomalies with the temporal autocorrelation

$$\text{AC}(t) = \frac{1}{E} \sum_{e=1}^E \frac{\sum_{y=1}^Y \bar{X}'_y(0) X'_{ey}(t)}{\sqrt{\sum_{y=1}^Y \bar{X}_y^2(0) \sum_{y=1}^Y X_{ey}^2(t)}}, \quad (4)$$

where $X'_{ey}(t) = X_{ey}(t) - \bar{X}(t)$ denotes the hindcasts' anomalies at lead time t , and $\bar{X}'_y(0) = \bar{X}_y(0) - \bar{X}(0)$ denotes the initial ensemble means anomalies computed

with CanCM4 data-constrained runs. The squared autocorrelation $AC^2(t)$ provides the soil moisture variance fraction at lead time t that can be linearly attributed to the initial soil moisture anomalies. A strong and statistically significant autocorrelation between hindcast anomalies at lead time t and the initial ensemble mean anomalies indicates that soil moisture has a commensurable degree of persistence that imparts potential predictability to that lead time.

d. Regression analysis

For the evaluations of CanCM4 assimilation runs (section 3), we employ ensemble means of VFSM and PEI annual averages to compute the temporal anomaly correlation coefficient (ACC) relative to the various observation-based datasets described in section 2b. To investigate the effects of precipitation and potential evapotranspiration on VFSM interannual variability in CanCM4 hindcasts (section 5), we employ lagged partial correlations between interannual time series of monthly VFSM and precipitation or PET for different initialization and lead times. We use partial correlations to remove the covariance that might exist between precipitation and PET as explanatory variables of VFSM variability. When computing the lagged partial correlations between monthly VFSM and, say, precipitation, we control for all covariates of lagged PET, and vice versa. We then examine the regression patterns for the largest partial correlation over the lags considered, as well as the corresponding lags (in months), which are generally different among grid cells. Lagged values are considered because the land surface tends to integrate atmospheric forcing; therefore, interannual variations of VFSM in a given month may be better explained by variations of the atmospheric forcing in previous months. To assess the relationship between ENSO teleconnections and VFSM, we examine anomaly regression patterns of VFSM, precipitation, and PET against lagged Niño-3.4 index obtained with temporal Pearson correlations. Because we quantify the strength of the linear relationship between soil moisture and its potential drivers, weak correlations should be interpreted with caution, as there may be nonlinear relationships among these quantities that are not examined in this work. Throughout the text, we will refer to statistical significance for both partial and full correlations based on two-tailed $p < 0.1$ probabilities, unless stated otherwise. We employ model ensemble means for all these computations.

3. Initialization of precipitation, potential evapotranspiration, and soil moisture

In this section, we evaluate the ability of CanCM4 assimilation runs to provide initial conditions for soil

moisture with realistic climatological means and interannual variability, as well as realistic phasing of interannual variations. We rely on the reanalysis datasets described in section 2b to facilitate these comparisons. Because persistence is an important source of soil moisture predictability (section 5a below), a realistic initialization should contribute appreciably to soil moisture actual skill on subseasonal to seasonal time scales.

a. Climatology of the initial states

As a measure of aridity, we show in Fig. 1a the climatology of annual mean PEI to characterize soil water surplus or deficit (Vicente-Serrano et al. 2010) in CanCM4 assimilation runs. The geographic patterns of PEI are largely consistent with known distributions of aridity indices (e.g., Zomer et al. 2008; Maliva and Missimer 2012), except in the Amazon basin, where temperature is biased high and precipitation is biased low in CanCM4 historical climate simulations (Merryfield et al. 2013). These biases, combined with the overestimation of PET by the Thornthwaite method in the presence of warm biases (Kumar et al. 1987), may explain the inconsistencies. A comparison with PEI values obtained with GPCP2.3 and Berkeley Earth datasets for precipitation and temperature, respectively (Fig. 1b), reveals that except for the tropical regions of South America, the annual mean soil water surplus or deficit is generally well represented in CanCM4 assimilation runs. Climatological means and interannual standard deviations of seasonal averages of modeled and observed P and PET are given in Figs. S1 and S2 of the supplementary material as references.

Interannual standard deviation of annual mean PEI in CanCM4 assimilation runs is shown in Fig. 1c. The largest interannual variations are mostly in the tropics and humid subtropical regions, which are characterized by large interannual precipitation variances (Fig. S1). The large standard deviations in the Sahara, the Arabian Peninsula, northwestern India, and northern Australia are mostly the result of pronounced interannual variance of potential evapotranspiration during hemispheric summer (Fig. S1). Standard deviations in Fig. 1c are mostly within a factor of 0.2 of those derived with GPCP2.3 and Berkeley Earth datasets (Fig. 1d), although the CanCM4 assimilation runs tend to overestimate PEI variances over arid regions in Asia and Africa and in east Africa, resulting from overestimations of not only potential evapotranspiration, but also precipitation variance (Fig. S3).

Climatological means and interannual standard deviations of simulated VFSM annual averages are shown in Figs. 2a and 2b. The climatological mean patterns

FIG. 1. (top) Climatological mean in 1981–2010 of PEI annual averages from (a) CanCM4 assimilation runs and (b) GPCP2.3 and Berkeley Earth datasets for precipitation and temperature, respectively. (bottom) Climatological interannual standard deviation of PEI annual averages from (c) CanCM4 assimilation runs and (d) its ratio in CanCM4 assimilation runs to that in the observation-based dataset. Model and observation-based quantities are indicated with subscripts m and o , respectively. Grid locations nearest to 1) GBY-NL, 2) MRD-BC, 3) ONW-MS, and 4) LEH-TX are indicated in (a). See Table 1.

(Fig. 2a) are mostly consistent with those of PEI (Fig. 1a), with typically larger soil moisture fractions in regions of water surplus, most notably in the tropics and humid subtropical regions, although soil porosity also

influences VFSM. Arid and semiarid regions have relatively low soil water content, as expected, and low VFSM is also found in some northern cold continental regions. There is, however, a relatively high VFSM in

FIG. 2. Climatological (a) mean and (b) interannual standard deviation of VFSM annual averages from CanCM4 assimilation runs in 1981–2010; model (c) bias relative to Blended-2 and (d) standard deviation are shown in units of the Blended-2 standard deviation. Model and observation-based quantities are indicated with subscripts m and o , respectively.

TABLE 1. Location and climate characteristics of grid cells used in this study.

Name	Province/state	Label	Grid lat °N/lon °W	Characteristics
Goose Bay	Newfoundland and Labrador, Canada	GBY-NL	54.42/59.06	Taiga
Mission Ridge	British Columbia, Canada	MRD-BC	51.63/120.94	Cordillera
Onward	Mississippi, United States	ONW-MS	32.09/90.00	Humid subtropical
Lehman	Texas, United States	LHM-TX	34.88/104.06	Cold semiarid

the northern Sahara ($0.2\text{--}0.25\text{ m}^3\text{ m}^{-3}$) mostly the result of the water content in the model's deepest soil layer (not shown). VFSM standard deviations are pronounced in the tropics and humid subtropical regions, although the strongest variability is found in Siberia, the Himalayas, and the Tibetan Plateau, where total soil depths (not shown) are shallow. The strong VFSM variability in these regions is likely the result of snow cover interannual variations and resulting variability of soil water input from snowmelt. Overall, VFSM annual averages in CanCM4 assimilation runs tend to have considerable variability in the regions of typically larger precipitation variability, although it is less seasonally dependent than precipitation (Fig. S1), likely a result of its anomaly persistence.

b. Evaluations of VFSM against reanalysis

VFSM biases in CanCM4 assimilation runs normalized by Blended-2 interannual standard deviation are shown in Fig. 2c. Relative to Blended-2, CanCM4 tends to underestimate soil moisture over most of the globe, most notably in cold continental regions of eastern Canada and northwestern Eurasia and in the Siberian tundra. Regions in which assimilating CanCM4 overestimates VFSM relative to Blended-2 are generally arid regions that include the Sahara, much of the Arabian Peninsula, east-central Asia, and interior eastern Australia. These patterns are largely consistent with analogous precipitation biases relative to GPCP2.3 (not shown), which are most strongly negative in tropical South America and most strongly positive in the Sahara. VFSM interannual standard deviation in the model shows significant departure from those in Blended-2 (Fig. 2d), with a tendency for overestimations in the arid regions of Africa and Asia mentioned above, in the areas of Siberia, Himalayas, and the Tibetan Plateau that exhibit high standard deviations in Fig. 2b, and in central east Africa. A comparison of Figs. 1d and 2d reveals that regions with excessive PEI interannual variations tend to have an excess of VFSM interannual variance.

ACC of simulated PEI and VFSM annual averages during 1981–2010 against the corresponding observation-based datasets is shown in Fig. 3. Overall, ACC for PEI

is statistically significant over 90% of the land surface (excluding Antarctica and Greenland) with global area mean value of approximately 0.6 (Fig. 3a), whereas ACC for VFSM is statistically significant over 70% of the land surface with global area mean value of approximately 0.4 (Fig. 3b, Table 2). The analogous quantities for P and PET computed separately are 80% and 0.5, and 97% and 0.7, respectively (Table 2). Seasonally, there is a tendency for both P and PET to have the lowest global area mean ACC and land area fraction of statistically significant ACC during boreal summer (JJA), whereas VFSM has the lowest values during the following season (Table 2). This behavior of lowest skill for VFSM in the season following that of lowest skill for the climatic variables also occurs for the hemispheric computations (not shown). For VFSM, regions with $\text{ACC} > 0.6$ include the western continental United States, eastern South America, western Europe, some areas of Siberia, Indonesia, and surrounding coastal areas, most of Australia, and much of southern and southeastern Africa. ACC patterns for VFSM resemble

FIG. 3. Anomaly correlation coefficient of (a) PEI and (b) VFSM annual averages during 1981–2010 from CanCM4 assimilation runs and the observation-based datasets of Figs. 1 and 2.

TABLE 2. Global area mean ACC and land area fraction with ACC > 0.3 (in %) for seasonal and annual averages of volumetric soil moisture fraction, precipitation, and PET in CanCM4 assimilation runs relative to Blended-2, GPCP2.3, and Earth Berkeley datasets, respectively, in 1981–2010. Both Greenland and Antarctica are excluded from the computations. Quantities in bold denote lowest skill.

		DJF	MAM	JJA	SON	Annual
Global mean ACC	VFSM	0.40	0.39	0.39	0.37	0.41
	<i>P</i>	0.55	0.54	0.48	0.55	0.50
	PET	0.71	0.77	0.70	0.72	0.73
Land fraction (%)	VFSM	66	67	69	65	70
	<i>P</i>	83	84	78	87	80
	PET	95	97	93	94	97

those of PEI, except for the arid regions (desert and steppe), the high-latitude cold continental areas, and the tundra, where biases relative to Blended-2 are also large (Figs. 2c,d). This resemblance is to a large extent determined by precipitation, for which ACC patterns both for annual (Fig. S4) and seasonal averages (not shown) largely coincide with those of VFSM, except at latitudes northward of approximately 55°N, where soil frozen water content and snow cover strongly affect VFSM. This suggests that by improving the representation of precipitation in CanCM4 assimilation runs, the biases in the initialization of soil moisture in CanSIPS, which occurs indirectly through the response of the CLASS land scheme to the behavior of the assimilating atmospheric model, should also improve.

4. Soil moisture potential predictability: Local results

Before we discuss the global results, we first examine the behavior and sources of potential predictability of VFSM from CanCM4 hindcast runs in the four grid locations shown in Fig. 1a and described in Table 1 that are representative of different climate, water content, and soil moisture–atmosphere interaction regimes across North America. The examination of representative grid locations is useful for a clearer description of the mechanisms responsible for predictability. In the following, we adopt the names of the measurement sites from the British Columbia Automated Snow Pillow in Canada (www.brfc.env.gov.bc.ca) and the Soil Climate Analysis Network (SCAN) in the United States (<http://www.wcc.nrcs.usda.gov/scan/>) that are nearest to the centers of these grid cells, though we do not employ data from those stations in this analysis. These locations are Goose Bay, Labrador (GBY-NL), and Mission Ridge, British Columbia (MRD-BC), in Canada, and Onward, Mississippi (ONW-MS), and Lehman, Texas (LEH-TX),

in the United States. The climatological values described in section 4a are obtained from the CanCM4 assimilation runs and are compared against the observation-based datasets (section 2b). The predictability results shown in sections 4b and 4c are obtained from CanCM4 hindcast runs, and comparisons of soil moisture persistence against reanalyses are provided in section 4b.

a. Precipitation, potential evapotranspiration, and soil moisture climatology

Monthly climatologies of VFSM, precipitation, and temperature at the four grid locations are shown in Fig. 4. Overall, there is a good agreement between the modeled and observed seasonality, although there are significant biases as well. There are generally negative model biases in VFSM and significant biases in precipitation that are positive or negative, depending on grid location. Biases in both VFSM and precipitation are largest at ONW-MS and smallest at LEH-TX. Temperature climatologies tend to agree well, although there is a tendency for positive biases in the model, except for the negative bias at MRD-BC (Fig. 4a).

GBY-NL (Fig. 4a) in eastern Canada is representative of the taiga, with estimated mean temperature ranging from below −15°C in winter to almost 15°C in summer and seasonal snow cover typically from late October to May, with maximum SWE typically about 250 mm (not shown). Annual precipitation is variable throughout the year and tends to peak in summer and fall in both the model ($\approx 90 \text{ mm month}^{-1}$) and observations ($\approx 110 \text{ mm month}^{-1}$), whereas VFSM remains mostly unchanged throughout the year in the model (reanalysis) at about 15% (25%), with slight maxima during the shoulder seasons. MRD-BC (Fig. 4b) in western Canada is representative of a cordillera, having mean temperature ranging from −10°C in winter to 10°C in summer and seasonal snow cover that typically lasts from early October to June, with maximum SWE above 350 mm (not shown). Precipitation peaks during fall and early summer, although modeled values are highest in the fall ($\approx 125 \text{ mm month}^{-1}$), and VFSM, which is above 20% throughout the year, has a noticeable seasonality with maxima during the shoulder seasons that are higher in spring ($\approx 25\%$ in the model). Modeled frozen VFSM in GBY-NL (not shown) ranges from 5% to 10% during winter and early spring, which is considerably higher than that in MRD-BC ($<2\%$).

Soil moisture in ONW-MS (Fig. 4c) has the strongest seasonality among all sites, with modeled volumetric fractions ranging from about 20% in summer to approximately 30% in winter. Values are significantly higher in the blended reanalysis product ($\approx 35\%$ – 45%). Observed precipitation peaks in winter and late spring at over 150 mm month^{−1}, whereas modeled values peak in summer at

FIG. 4. Monthly climatological mean of (green) VFSM, (blue) precipitation, and (red) temperature from (solid curves and dark green) CanCM4 hindcast runs and (dashed curves and light green) observation-based datasets during 1981–2010 at grid locations nearest to (a) GBY-NL, (b) MRD-BC, (c) ONW-MS, and (d) LEH-TX. Note the different scales for precipitation.

about $130 \text{ mm month}^{-1}$. Average monthly temperature ranges from 10°C in winter to over 25°C in summer for both the model and observations, characteristic of a humid subtropical climate. LEH-TX (Fig. 4d) has a cold semiarid climate, with average winter temperature just above freezing and average summer temperature at about 25°C . Precipitation is low and peaks at about 90 mm month^{-1} in summer, resulting in low VFSM ($\approx 15\%$) that remains relatively unchanged throughout the year. For all sites, PET (not shown) peaks in summer at values that decrease with latitude, ranging from about $100 \text{ mm month}^{-1}$ in GBY-NL and MRD-BC to 200 and $250 \text{ mm month}^{-1}$ in LEH-TX and ONW-MS, respectively.

b. Seasonal dependence of soil moisture anomalies

The different climates and soil moisture–atmosphere interactions of these locations affect the persistence of

soil moisture anomalies. In particular, seasonal variations of soil moisture climate drivers should affect the seasonal cycle of soil moisture persistence. We use autocorrelation (AC) as defined in Eq. (4) to describe persistence of initial soil moisture anomalies.

The seasonality of modeled initial soil moisture anomaly persistence in GBY-NL, MRD-BC, ONW-MS, and LEH-TX is shown in Fig. 5: the horizontal axis indicates the initial month of the forecast, and the vertical axis represents the progression of the corresponding forecast through its 12-month range. Overall, soil moisture persistence tends to be longest during late fall and winter, except for ONW-MS, which has longest persistence in winter and spring, and tends to be shortest during late spring and summer. The rate of AC decay varies considerably with location, with AC remaining significant for up to 7 months at GBY-NL and MRD-BC

FIG. 5. Autocorrelation of daily VFSM in CanCM4 hindcast runs (1981–2010) at grid locations nearest to (a) GBY-NL, (b) MRD-BC, (c) ONW-MS, and (d) LEH-TX. Initial anomalies are from the hindcast initial conditions.

(Figs. 5a,b) and up to 5 months at ONW-MS (Fig. 5c), but only to <3 months at LEH-TX (Fig. 5d). Soil moisture persistence in late fall and winter in GBY-NL and MRD-BC is largely the result of the insulating properties of the snow cover that helps preserve the anomalies from the early stages of the snow season until spring. The seasonality and duration of soil anomaly persistence in ONW-MS and LEH-TX are largely consistent with the findings of Wang and Kumar (1998) for the North American domain 105° – 85° W and 34° – 49° W. They attribute these seasonal differences to the seasonal variation in PET resulting from the seasonal cycle of surface insolation.

The seasonality of Blended-2 soil moisture anomaly persistence, shown in Fig. S5, is largely consistent with the modeled values (Fig. 5), though the persistence time scales are somewhat different. Soil moisture anomalies are significantly more persistent in the reanalysis

product than in the model, particularly at MRD-BC in summer and LEH-TX in most seasons. At LEH-TX, the differences are large, despite the good agreement between both modeled and observed climatologies of soil moisture and the meteorological data at the start of the forecast (Fig. 4d), although modeled soil moisture interannual variance is much smaller than observed and negligible during spring to early fall (not shown). This implies an initially weak soil moisture signal variance that can be obscured by weather noise (as shown in Fig. 6d and discussed below), which in turn impacts persistence.

c. Potential predictability of soil moisture

PP of daily soil moisture as defined in Eq. (3) at GBY-NL, MRD-BC, ONW-MS, and LEH-TX, obtained from 1-yr hindcasts initialized at the start of every month, is shown in Fig. 6. At GBY-NL (Fig. 6a), PP is initially high

FIG. 6. Potential predictability of daily VFSM in CanCM4 hindcast runs (1981–2010) at (a) GBY-NL, (b) MRD-BC, (c) ONW-MS, and (d) LEH-TX. Dashed line indicates the 95% confidence level.

and decays rapidly, except from late autumn until spring (NDJFM) when PP plateaus, apparently because soil moisture anomalies are then “frozen in” due to the insulating effects of the seasonal snow cover. These PP plateaus are consistent with the anomaly persistence patterns of Fig. 5a. A slight reemergence of PP is apparent in April and May, most likely from the spring snowmelt, although its amplitude depends on the initialization time. Soil moisture PP in MRD-BC (Fig. 6b) shares some of these features but shows interesting differences with GBY-NL as well, including higher but slowly decaying PP during winter, longer-lived predictability for most initialization times with reemergences during spring and summer, and reemergence of predictability at up to 1-yr lead time during spring. Likely causes of these differences are 1) higher precipitation interannual variance at MRD-BC during fall, as well as 2) higher mean precipitation during fall and winter,

3) the influence of ENSO-teleconnected precipitation and temperature anomalies that are potentially predictable at long leads, and possibly 4) higher surface and/or subsurface runoff. These factors are discussed in turn below.

Higher soil moisture predictability at MRD-BC relative to GBY-NL during winter is likely the result of stronger soil moisture anomalies (not shown) that are driven by stronger interannual precipitation variance in the fall and are preserved by the snow cover (Fig. 5b). The relatively lower persistence and thus faster decaying predictability during winter in MRD-BC (Figs. 5a,b, Figs. 6a,b) may be the result of higher mean winter precipitation (Fig. 4b); this leads to deeper snowpack and higher insulation, which preserves heat stored in the ground during summer, resulting in lower frozen soil moisture fraction and higher subsurface runoff (not shown) enhanced by MRD-BC’s surrounding

orography. In addition, the SWE excess in MRD-BC ($>40\%$) and its interannual variability may significantly impact soil moisture anomalies through spring snowmelt, thus contributing to predictability.

At long leads, soil moisture predictability at MRD-BC appears to derive from the effect of ENSO-teleconnected temperature and precipitation anomalies on SWE. Such teleconnections are reasonably realistic over North America in CanCM4, particularly in winter (cf. Fig. 17 of Merryfield et al. 2013), although the warm anomalies that accompany El Niño tend to be concentrated farther north and west than observed. ENSO-teleconnected soil moisture anomalies during spring are found to be statistically significant at up to 12-month lead (Fig. S6b). These are most likely driven by the effects of temperature anomalies on SWE during late winter and early spring. We note that ENSO-teleconnected PET anomalies during spring tend to be positive for all leads (Fig. S7b); thus, temperature effects on soil moisture, if any, should be primarily through snowmelt and possibly any covariability with precipitation. This is consistent with the positive anomalies of spring soil moisture regressed against temperature (not shown).

PP of soil moisture at ONW-MS (Fig. 6c) initially decays quickly, regardless of initialization month, toward an envelope that exhibits a strong seasonality. This “envelope of predictability” is likely determined by ENSO-teleconnected anomalies that occur mainly from December to June (Fig. S6c). Positive ENSO-related soil moisture anomalies are linked to positive ENSO-teleconnected precipitation anomalies (Fig. S8c) and negative PET anomalies (Fig. S7c) during those months and, to some extent, anomaly persistence during spring (Fig. 5c), despite relatively strong P - and PET-induced VFSM variability (Figs. S9c, S10c). Soil moisture persistence may in part be the result of the relatively high VFSM during spring (Fig. 4c), leading to soil saturation and possibly runoff. Soil moisture in LEH-TX is mostly characterized by short-lived predictability (Fig. 6d) resulting from initial anomaly persistence (Fig. 5) that depends on initialization time. Persistence and PP tend to be longest for hindcasts initialized in November–January (Figs. 5d, 6d), when potential evapotranspiration is low and precipitation is typically about 20–30 mm month⁻¹ (Fig. 4d). PP is slightly enhanced during late winter and spring (Fig. 6d), when positive ENSO-teleconnected soil moisture anomalies are statistically significant (Fig. S6d), driven by positive ENSO-teleconnected precipitation anomalies (Fig. S8d, S9d) and negative ENSO-teleconnected PET anomalies (Fig. S7d, S10d) that are statistically significant but much weaker than at ONW-MS.

5. Soil moisture potential predictability: Global results

Having discussed the local results, we now examine the potential predictability of soil moisture in 1981–2010 CanCM4 hindcasts globally. We show results for initializations at the start of each principal season (i.e., MAM, JJA, SON, and DJF) and different lead times. We first examine persistence of initial soil moisture anomalies as a source of predictability and then focus on long-lead potential predictability, its geographic distribution, its likely causes and behavior, and its link to ENSO teleconnections.

a. Anomaly persistence and long-lead predictability

The persistence of initial soil moisture anomalies in CanCM4 hindcasts is represented in Fig. 7 as the number of consecutive months following initialization for which AC [Eq. (4)] is statistically significant with one-tailed $p < 0.1$. Soil moisture persistence depends on climate zone and initialization date, with longest persistence typically found in the higher latitudes during boreal winter, mostly because of the insulating effect of the snow cover and the relatively low precipitation variability, compared to that in the tropics (e.g., Fig. S1). Arid and semiarid regions tend to have longest soil moisture persistence during the colder months, when potential evapotranspiration is relatively low. This is consistent with the results of section 4b detailing the seasonal dependence of soil moisture anomalies at four representative North American grid locations (Fig. 5). Temperate zones and cold continental regions (without a dry season) tend to have persistent soil moisture anomalies with AC > 0.3 over 4 months in all seasons, most notably in the eastern United States, northern Europe, southeastern Asia, and subtropical regions of eastern South America. In the tropics, soil moisture persists over 4 months at some areas of eastern South America, western Africa, and the monsoon regions of South Asia, depending on the season. However, it is relatively short lived in the continental rain forest (i.e., northwestern South America and central Africa), likely due to short residence time resulting from frequent rainfalls throughout the year.

PP of soil moisture is statistically significant (Fig. S11) in the regions and seasons where AC is also significant (Fig. 7), reflecting that persistence is a source of soil moisture predictability. There are regions, however, where potential predictability of soil moisture is not linked to the persistence of initial anomalies. This is clear in Fig. 8, which shows monthly PP of CanCM4 soil moisture hindcasts for long leads (7, 9, and 11 months)

FIG. 7. Number of consecutive months from initial date with autocorrelation of monthly volumetric soil moisture fraction statistically significant at one-tailed $p < 0.1$. Initial conditions from CanCM4 assimilation runs correspond to the dates indicated in the panels.

and 4 initialization months, with regions having $AC > 0.3$ masked out. Several regions of long-lived soil moisture PP are shown, most notably in the tropics, and also in the extratropics, particularly in southern North America (April to June) and the cold arid regions of west-central Asia (December to June). The single largest region with the strongest PP for all target months is tropical South America, with an area that changes slightly with seasons and has higher values in the

rainforest. Likely causes of such predictability are discussed in the following sections.

b. Variability of soil moisture linked to precipitation and potential evapotranspiration

VFSM regressed against lagged precipitation and potential evapotranspiration in long-lead (7, 9, and 11 month) hindcasts is shown in Fig. 9 and Fig. S12, respectively. The regression values are at the lags of

FIG. 8. Potential predictability of monthly volumetric soil moisture fraction in CanCM4 hindcast runs (1981–2010) at (from top to bottom) 7, 9, and 11 months after initialization at the start of (from left to right) June, September, December, and March. The few regions where autocorrelation is > 0.3 are shown in gray. This suggests that regions in color have sources for potential predictability of soil moisture other than persistence of initial anomalies.

FIG. 9. Lagged partial correlation of monthly VFSM and precipitation in CanCM4 hindcast runs (1981–2010) for several initialization and target months. Values shown correspond to the maximum absolute partial correlation over lags from 0 to 7 months.

maximum partial correlation absolute value, with the maximum taken over lags ranging from 0 to 7 months. Generally, the lags of maximum absolute partial correlation vary with location, initialization time, and lead time (Fig. S13), though they tend to be within 3 months, except in some cold regions during late winter and early spring. Strong positive correlation (>0.8) of VFSM with precipitation is typically found in regions of significant PP (Fig. 8), regardless of initialization and lead times. Lags of maximum correlation at these regions are typically 0–2 months, depending on the target month (Fig. S13), possibly because precipitation variations from previous weeks to months tend to drive column soil moisture interannual variability and contribute to VFSM potential predictability. By contrast, lagged correlations against PET are generally weaker (>-0.6) for all initialization and lead times (Fig. S12), suggesting that the effect of PET on interannual variations of VFSM is comparatively small, though possible nonlinear relationships are not considered here. Interestingly, there are regions where VFSM and antecedent PET tend to be in phase (positive correlations), which may be in part the result of increased atmospheric moisture-enhancing precipitation during the lagged months (e.g., tropical South America) or the result of snowmelt in the snow-covered regions.

A mechanism consistent with significant PP of VFSM and strong lagged correlations with precipitation is that of precipitation interannual variance, leading to soil moisture anomalies that may persist from days to months. For example, in a region that encompasses west-central Asia (characterized by a cold, semiarid

climate), where there is statistically significant long-range soil moisture PP during winter to early summer (December–June; Fig. 8), the strongest lagged correlation between soil moisture and precipitation occurs during late fall and winter (November–February; Fig. 9), with lags typically from November to December (Fig. S13) when precipitation has the largest interannual variance (not shown). This suggests that partly because of the significant soil moisture persistence in the region (Fig. 7), the precipitation variance during late fall and early winter imprints the soil moisture anomalies that are potentially predictable up to early summer. A similar mechanism may explain the long-lived PP of soil moisture in the southern United States during boreal spring and early summer, where the lagged regression patterns against precipitation are relatively strong during boreal fall and winter (Fig. 9).

In tropical South America, where potential predictability of soil moisture is relatively high throughout the year (Fig. 8), partial correlations against precipitation are also high for most initialization and lead times (Fig. 9). The mechanisms for soil moisture PP in this region may be different to that described above, in that the forcing fields may modulate soil moisture anomalies somewhat continuously, particularly in the rain forest climate regions where a year-round low-pressure system associated with the intertropical convergence zone leads to heavy and frequent rainfall. Thus, a possible cause for potential predictability of soil moisture in such conditions is precipitation-driven variability with soil moisture anomalies modulated by subseasonal to year-round

FIG. 10. Lagged regression anomaly patterns of monthly VFSM against Niño-3.4 index in CanCM4 hindcast runs (1981–2010) for several initialization and target months. Values shown correspond to the maximum absolute Pearson correlation over lags from 0 to 7 months. Values with correlations <0.3 are not plotted.

precipitation. The effects of PET interannual variability may also be relevant to modulating soil moisture anomalies in a near-saturated soil (Figs. 2a, S11), which may compensate at times the precipitation variance, leading to soil moisture anomaly persistence. By contrast, in tropical savanna and monsoonal regions, which are characterized by a dry season, soil moisture PP is likely the result of strong seasonal rainfall with significant interannual variance (Fig. S1) combined with some degree of soil moisture anomaly persistence.

c. ENSO teleconnections and soil moisture

The mechanisms for soil moisture potential predictability described in section 5b are based primarily on precipitation and, to a lesser degree, potential evapotranspiration as drivers of interannual soil moisture variability. In this section, we argue that long-lead soil moisture potential predictability in CanCM4 hindcasts is driven, at least in part, by ENSO-teleconnected precipitation and potential evapotranspiration.

Figure 10 shows statistically significant lagged anomaly correlation patterns of VFSM regressed on Niño-3.4 index in CanCM4 hindcast runs for the same initialization months and lead times as in previous figures. For these plots, the regression values are from the lags with maximum absolute value of Pearson correlation over lags from 0 to 7 months. The corresponding lags are shown in Fig. S14. Overall, ENSO-teleconnected anomalies are roughly of the same magnitude as the interannual standard deviation of VFSM initial conditions' seasonal (Fig. S1) and annual means (Fig. 2b), suggesting that

ENSO plays a significant role on soil moisture interannual variability. The stronger ENSO-teleconnected VFSM anomalies are located (not exclusively) in regions of significant soil moisture potential predictability (Fig. 8) and, more generally, in regions of significant partial correlation between soil moisture and precipitation or potential evapotranspiration (Figs. 9, S12). This is evident in the tropics (Fig. 10), where there is a tendency for relatively strong negative ENSO-teleconnected VFSM anomalies throughout the year, with absolute values and extent that depend on the season. A notable exception is central Africa, where anomalies tend to be positive, although much weaker, both in absolute value and extent. Positive ENSO-teleconnected VFSM anomalies are also found in southern North America (mostly during winter to early summer), in southeastern South America (with variable strengths throughout the year), and in cold arid regions of west-central Asia (mostly during winter and spring).

The negative ENSO-teleconnected VFSM anomalies in the tropics, particularly in South America, correspond to below-normal VFSM during El Niño events, consistent with the negative ENSO-teleconnected precipitation and positive potential evapotranspiration anomalies depicted in Figs. 11 and 12, respectively, with associated lags shown in Fig. S15. These patterns are largely in agreement with the known ENSO-related dry conditions in the region (Fig. 16.10 of Peixoto and Oort 1992; Liebmann and Marengo 2001; Yang and DelSole 2012). The opposite occurs in central Africa (Figs. 10–12), where ENSO-driven precipitation anomalies are positive and PET anomalies, although much weaker, tend to

FIG. 11. As in Fig. 10, but for precipitation.

have the same sign as those for soil moisture. The positive precipitation and soil moisture anomalies are also mostly in agreement with known ENSO-related wet patterns in this region (Fig. 16.10 of Peixoto and Oort 1992), although differences exist in the location of the anomaly patterns (Yang and DelSole 2012).

Positive ENSO-teleconnected VFSM anomalies occur over west-central Asia during winter and spring, where soil moisture has a strong lagged partial correlation with precipitation (Fig. 9) and is potentially predictable during winter up to early summer (Fig. 8). This is consistent with the positive ENSO-teleconnected precipitation anomalies in this region during late fall, winter, and spring (see Fig. 11 and associated lags in Fig. S15). Nazemosadat and Ghasemi (2004) showed that this is indeed a region of ENSO-related precipitation anomalies, with typically wet conditions in the fall over most of Iran during El Niño events, which can extend to winter in the southeastern side of the Caspian Sea coast. Positive ENSO-related soil moisture anomaly patterns are present in southern North America as well, where PP of VFSM is statistically significant during winter up to early summer, although relatively weaker. This region has an ENSO-related wet pattern roughly during October–March (Fig. 16.10 of Peixoto and Oort 1992; Yang and DelSole 2012).

6. Summary and conclusions

CanCM4 is one of two global climate models currently employed by CanSIPS to provide operational ensemble multiseasonal forecasts for ECCO. We assessed the initialization and potential predictability of soil moisture in CanCM4 hindcasts during 1981–2010. Because the initial conditions of soil moisture forecasts in

CanSIPS are determined from the data-constrained model atmospheric fields, we also examined the initialization of precipitation and potential evapotranspiration, particularly of the aridity index $PEI = P - PET$ describing the model's water surplus or deficit (Fig. 1).

Evaluations of CanCM4 assimilation runs showed that annual averages of PEI initial conditions are relatively well represented in CanSIPS, the most notable exception being the Amazon basin, which tends to be unrealistically dry. Annually averaged VFSM initial conditions are biased high in most arid regions and biased low elsewhere, most notably in cold continental climate and tundra (Fig. 2c). The amplitude of the interannual variations of VFSM annual averages were shown to have significant departures from the observation-based datasets, with overestimations in arid regions, where PEI interannual variability is also large relative to observations (Fig. 1d), and in cold continental areas of Asia, characterized by shallow soil depths. The phase of VFSM annual averages in CanCM4 assimilation runs was shown to be in agreement with reanalyses in the regions where PEI (Fig. 3) and particularly precipitation have the highest anomaly correlation with observations, except for arid regions, cold continental areas, and tundra. Seasonally, the lowest global area mean anomaly correlation for VFSM tends to occur during boreal fall, preceded by the lowest value for precipitation in boreal summer (Table 2). Despite these biases, VFSM initial conditions were found to have considerable interannual variability, a necessary ingredient for potential predictability, particularly in regions of considerable precipitation variance (Fig. S1).

Soil moisture anomalies at four grid locations in North America were found to be most persistent during fall

FIG. 12. As in Fig. 10, but for PET. Note that the legend colors are reversed.

and winter and least persistent during spring and summer, with persistence time scales that depend on location. This seasonality of soil moisture persistence is likely the result of the seasonal variations of potential evapotranspiration, which are more strongly related to soil moisture variations during spring and summer (Fig. S10). Persistence tends to be longest in the two locations at higher latitudes, partly because of the insulating effect of the snow cover and possibly because of the weaker effect of PET. Reanalysis soil moisture was found to be generally more persistent than the model, particularly at the most arid of these grid locations (Figs. 5d, S5d), possibly as the result of low interannual variability in relation to weather noise. For two grid locations (Figs. 6b,c), potential predictability of soil moisture at long leads was linked to ENSO teleconnections driving soil moisture variability (Fig. S6), possibly through precipitation (Figs. S8, S9b,c) and temperature effects on the spring freshet (not shown).

Persistence of initial soil moisture anomalies in CanCM4 hindcasts was shown to be seasonally dependent (Figs. 5, 7), with longer durations typically during hemispheric winter and shorter durations during hemispheric summer. This seasonality is likely the result of variations in potential evapotranspiration resulting from the annual cycle of surface insolation (Delworth and Manabe 1988; Wang and Kumar 1998). In snow-covered regions, long anomaly persistence is also the result of cold soil temperature that helps preserve the soil moisture anomalies beneath the snowpack. Geographically, winter freezing and low potential evapotranspiration tend to cause longer soil moisture persistence in the mid- to high latitudes than in the tropics (Fig. 7). By contrast, larger precipitation

variance leads to typically larger soil moisture anomalies in the tropics and subtropics (Fig. S1) that are dissipated on shorter time scales, partly because of frequent rainfalls and because of the higher potential evapotranspiration. These results are consistent with the findings of Arora and Boer (2006). However, regions of long-lived persistence were also found in the Amazon basin, possibly the result of averaging out the daily, high-frequency component of soil moisture anomalies (largely affected by precipitation) and the tendency for decaying anomalies (mostly determined by the dissipative processes of potential evapotranspiration and runoff), combined with the balanced influence of precipitation and potential evapotranspiration on this region's near-saturated soil (Figs. 9, S12, 2a).

Potentially predictable soil moisture ($PP > 0.1$) at lead times up to 4–6 months (or longer in snow-covered regions) was linked to persistence of initial anomalies, with actual duration depending on season and location (Figs. 8, S11). Durations for higher predictability ($PP > 0.6$) are typically much reduced, but can be as long as 3–4 months in the tropics and 6–7 months in the mid- to high latitudes during hemispheric winter (Fig. S11). Although persistence of initial anomalies was found to be key for soil moisture predictability at monthly to subseasonal time scales, the effects of teleconnection indices enhanced potential predictability at such time leads. For longer leads (> 7 months), statistically significant PP of soil moisture was primarily found in tropical regions and in extratropical areas of ENSO-teleconnected influences (Figs. 6, 8, 10), including the southern United States, Australia, and west-central Asia. These PP values are seasonally dependent and generally much lower than for shorter lead times, partly as a

consequence of the growing error variance during the forecast. However, a year-round soil moisture potential predictability is present in the Amazon basin independently of initialization month, tending to be strongest (at about 40%–70% of the total variance) during boreal winter and spring. Lagged ENSO-teleconnected precipitation (possibly combined with some sort of soil moisture anomaly persistence) was shown to be the likely driver of long-lead soil moisture potential predictability in CanSIPS hindcasts, with seasonally varying lags between Niño-3.4 and precipitation depending on the time of ENSO teleconnection influences (Fig. S15) and the lag between ENSO-teleconnected precipitation and soil moisture (Fig. S13) typically within 3 months. Lagged partial regressions of soil moisture and PET were found to be much weaker than for precipitation, except in the Amazon basin (Fig. S12).

This work has shown that potential predictability of soil moisture in CanSIPS hindcasts is highly region and season dependent. The persistence of initial soil moisture anomalies was shown to drive potential predictability at monthly to subseasonal time scales, whereas ENSO-teleconnected precipitation was found to play a key role in soil moisture potential predictability at longer leads up to 1 year. Although actual skill in CanSIPS soil moisture hindcasts will be described elsewhere, the present results point to ways in which such skill can be improved. These include improved skill for predicting precipitation, particularly in tropical and subtropical regions, and in other regions of ENSO-teleconnected influence. For example, the dry bias in the Amazon basin may degrade actual predictability of soil moisture despite the year-round potential predictability found in the region. Also, improved initialization of soil moisture should enhance forecast skill in arid and semiarid, as well as cold, continental regions, which are characterized by persistence of initial anomalies. In cold regions, improved representation of the timing of snow onset and melt should also contribute to better soil moisture forecast skill. Whether CanSIPS is able to capitalize on soil moisture potential predictability and to what extent the model biases described here may degrade this capability are subjects of a future investigation.

Acknowledgments. The authors are thankful to Vivek Arora and Michael Sigmund for suggestions made on an early version of this paper and to three anonymous reviewers for valuable comments. This research is partly supported by the Natural Sciences and Engineering Research Council of Canada's (NSERC) Climate Change and Atmospheric Research (CCAR) program

through the Canadian Sea Ice and Snow Evolution (CanSISE) Network (www.cansise.ca).

REFERENCES

- Adler, R. F., and Coauthors, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeorol.*, **4**, 1147–1167, [https://doi.org/10.1175/1525-7541\(2003\)004<1147:TVGPCP>2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2).
- Arora, V. K., and G. J. Boer, 2006: The temporal variability of soil moisture and surface hydrological quantities in a climate model. *J. Climate*, **19**, 5875–5888, <https://doi.org/10.1175/JCLI3926.1>.
- Balsamo, G., and Coauthors, 2015: ERA-Interim/Land: A global land surface reanalysis data set. *Hydrol. Earth Syst. Sci.*, **19**, 389–407, <https://doi.org/10.5194/hess-19-389-2015>.
- Barnston, A. G., M. K. Tippett, M. Ranganathan, and M. L. L'Heureux, 2018: Deterministic skill of ENSO predictions from the North American multimodel ensemble. *Climate Dyn.*, <https://doi.org/10.1007/s00382-017-3603-3>, in press.
- Bellucci, A., and Coauthors, 2015: Advancements in decadal climate predictability: The role of nonoceanic drivers. *Rev. Geophys.*, **53**, 165–202, <https://doi.org/10.1002/2014RG000473>.
- Bonan, G. B., 2008: *Ecological Climatology: Concepts and Applications*. Cambridge University Press, 550 pp.
- Conil, S., H. Douville, and S. Tyteca, 2009: Contribution of realistic soil moisture initial conditions to boreal summer climate predictability. *Climate Dyn.*, **32**, 75–93, <https://doi.org/10.1007/s00382-008-0375-9>.
- Dee, P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- DelSole, T., A. Kumar, and B. Jha, 2013: Potential seasonal predictability: Comparison between empirical and dynamical model estimates. *Geophys. Res. Lett.*, **40**, 3200–3206, <https://doi.org/10.1002/grl.50581>.
- Delworth, T. L., and S. Manabe, 1988: The influence of potential evaporation on the variabilities of simulated soil wetness and climate. *J. Climate*, **1**, 523–547, [https://doi.org/10.1175/1520-0442\(1988\)001<0523:TIOPEO>2.0.CO;2](https://doi.org/10.1175/1520-0442(1988)001<0523:TIOPEO>2.0.CO;2).
- Drewitt, G., A. A. Berg, W. J. Merryfield, and W.-S. Lee, 2012: Effect of realistic soil moisture initialization on the Canadian CanCM3 seasonal forecast model. *Atmos.–Ocean*, **50**, 466–474, <https://doi.org/10.1080/07055900.2012.722910>.
- Entin, J. K., A. Robock, K. Y. Vinnikov, S. E. Hollinger, S. Liu, and A. Namkhair, 2000: Temporal and spatial scales of observed soil moisture variations in the extratropics. *J. Geophys. Res.*, **105**, 11 865–11 877, <https://doi.org/10.1029/2000JD900051>.
- Falloon, P., C. D. Jones, M. Ades, and K. Paul, 2011: Direct soil moisture controls of future global soil carbon changes: An important source of uncertainty. *Global Biogeochem. Cycles*, **25**, GB3010, <https://doi.org/10.1029/2010GB003938>.
- Fischer, E. M., S. I. Seneviratne, P. L. Vidale, D. Lüthi, and C. Schär, 2007: Soil moisture–atmosphere interactions during the 2003 European summer heat wave. *J. Climate*, **20**, 5081–5099, <https://doi.org/10.1175/JCLI4288.1>.
- Fisher, J. B., R. J. Whittaker, and Y. Malhi, 2011: ET come home: Potential evapotranspiration in geographical ecology. *Global Ecol. Biogeogr.*, **20**, 1–18, <https://doi.org/10.1111/j.1466-8238.2010.00578.x>.

- Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- Grace, B., and B. Quick, 1988: A comparison of methods for the calculation of potential evapotranspiration under windy semi-arid conditions of southern Alberta. *Can. Water Resour. J.*, **13**, 9–19, <https://doi.org/10.4296/cwrj1301009>.
- Granier, A., and Coauthors, 2007: Evidence for soil water control on carbon and water dynamics in European forests during the extremely dry year: 2003. *Agric. For. Meteorol.*, **143**, 123–145, <https://doi.org/10.1016/j.agrformet.2006.12.004>.
- Grillakis, M. G., A. G. Koutroulis, J. Komma, I. K. Tsanis, W. Wagner, and G. Blöschl, 2016: Initial soil moisture effects on flash flood generation—A comparison between basins of contrasting hydro-climatic conditions. *J. Hydrol.*, **541**, 206–217, <https://doi.org/10.1016/j.jhydrol.2016.03.007>.
- Guo, Z., P. A. Dirmeyer, T. DelSole, and R. D. Koster, 2012: Rebound in atmospheric predictability and the role of the land surface. *J. Climate*, **25**, 4744–4749, <https://doi.org/10.1175/JCLI-D-11-00651.1>.
- Hirschi, M., and Coauthors, 2011: Observational evidence for soil-moisture impact on hot extremes in southeastern Europe. *Nat. Geosci.*, **4**, 17–21, <https://doi.org/10.1038/ngeo1032>.
- Kanamitsu, M., C.-H. Lu, J. Schemm, and W. Ebisuzaki, 2003: The predictability of soil moisture and near-surface temperature in hindcasts of NCEP seasonal forecast model. *J. Climate*, **16**, 510–521, [https://doi.org/10.1175/1520-0442\(2003\)016<0510:TPO SMA>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<0510:TPO SMA>2.0.CO;2).
- Koster, R. D., and M. J. Suarez, 2001: Soil moisture memory in climate models. *J. Hydrometeorol.*, **2**, 558–570, [https://doi.org/10.1175/1525-7541\(2001\)002<0558:SMMICM>2.0.CO;2](https://doi.org/10.1175/1525-7541(2001)002<0558:SMMICM>2.0.CO;2).
- , and Coauthors, 2004: Regions of strong coupling between soil moisture and precipitation. *Science*, **305**, 1138–1140, <https://doi.org/10.1126/science.1100217>.
- , and Coauthors, 2010: Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. *Geophys. Res. Lett.*, **37**, L02402, <https://doi.org/10.1029/2009GL041677>.
- , and Coauthors, 2011: The second phase of the global land-atmosphere coupling experiment: Soil moisture contributions to subseasonal forecast skill. *J. Hydrometeorol.*, **12**, 805–822, <https://doi.org/10.1175/2011JHM1365.1>.
- Kumar, K. K., K. R. Kumar, and P. R. Rakhecha, 1987: Comparison of Penman and Thornthwaite methods of estimating potential evapotranspiration for Indian conditions. *Theor. Appl. Climatol.*, **38**, 140–146, <https://doi.org/10.1007/BF00868097>.
- Liebmann, B., and J. A. Marengo, 2001: Interannual variability of the rainy season and rainfall in the Brazilian Amazon basin. *J. Climate*, **14**, 4308–4318, [https://doi.org/10.1175/1520-0442\(2001\)014<4308:IVOTRS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<4308:IVOTRS>2.0.CO;2).
- Maliva, R. G., and T. M. Missimer, 2012: *Arid Lands Water Evolution and Management*. Springer, 1076 pp.
- Materia, S. A., A. Borrelli, A. Bellucci, A. Alessandri, P. Di Pietro, P. Athanasiadis, A. Navarra, and S. Gualdi, 2014: Impact of atmosphere and land surface initial conditions on seasonal forecast of global surface temperature. *J. Climate*, **27**, 9253–9271, <https://doi.org/10.1175/JCLI-D-14-00163.1>.
- McMahon, T. A., M. C. Peel, L. Lowe, R. Srikanthan, and T. R. McVicar, 2013: Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: A pragmatic synthesis. *Hydrol. Earth Syst. Sci.*, **17**, 1331–1363, <https://doi.org/10.5194/hess-17-1331-2013>.
- Merryfield, W. J., and Coauthors, 2013: The Canadian seasonal to interannual prediction system. Part I: Models and initialization. *Mon. Wea. Rev.*, **141**, 2910–2945, <https://doi.org/10.1175/MWR-D-12-00216.1>.
- Miralles, D. G., M. J. van den Berg, A. J. Teuling, and R. A. M. de Jeu, 2012: Soil moisture-temperature coupling: A multiscale observational analysis. *Geophys. Res. Lett.*, **39**, L21707, <https://doi.org/10.1029/2012GL053703>.
- Nazemosadat, M. J., and A. R. Ghasemi, 2004: Quantifying the ENSO-related shifts in the intensity and probability of drought and wet periods in Iran. *J. Climate*, **17**, 4005–4018, [https://doi.org/10.1175/1520-0442\(2004\)017<4005:QTESIT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<4005:QTESIT>2.0.CO;2).
- Nicolai-Shaw, N., L. Gudmundsson, M. Hirschi, and S. I. Seneviratne, 2016: Long-term predictability of soil moisture dynamics at the global scale: Persistence versus large-scale drivers. *Geophys. Res. Lett.*, **43**, 8554–8562, <https://doi.org/10.1002/2016GL069847>.
- Oglesby, R. J., and D. J. Erickson III, 1989: Soil moisture and the persistence of North America drought. *J. Climate*, **2**, 1362–1380, [https://doi.org/10.1175/1520-0442\(1989\)002<1362:SMATPO>2.0.CO;2](https://doi.org/10.1175/1520-0442(1989)002<1362:SMATPO>2.0.CO;2).
- Orth, R., and S. I. Seneviratne, 2012: Analysis of soil moisture memory from observations in Europe. *J. Geophys. Res.*, **117**, D15115, <https://doi.org/10.1029/2011JD017366>.
- Peixoto, J. P., and A. H. Oort, 1992: *Physics of Climate*. American Institute of Physics, 520 pp.
- Reichle, R. H., C. S. Draper, Q. Liu, M. Girotto, S. P. P. Mahanama, R. D. Koster, and G. J. M. De Lannoy, 2017: Assessment of MERRA-2 land surface hydrology estimates. *J. Climate*, **30**, 2937–2960, <https://doi.org/10.1175/JCLI-D-16-0720.1>.
- Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648, <https://doi.org/10.1175/JCLI-D-11-00015.1>.
- Rosenberg, N. J., B. L. Blad, and S. B. Verma, 1983: *Microclimate: The Biological Environment*. John Wiley & Sons, 528 pp.
- Rowell, D. P., 1998: Assessing potential seasonal predictability with an ensemble of multidecadal GCM simulations. *J. Climate*, **11**, 109–120, [https://doi.org/10.1175/1520-0442\(1998\)011<0109:APSPWA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1998)011<0109:APSPWA>2.0.CO;2).
- Schär, C., D. Lüthi, U. Beyerle, and E. Heise, 1999: The soil-precipitation feedback: A process study with a regional climate model. *J. Climate*, **12**, 722–741, [https://doi.org/10.1175/1520-0442\(1999\)012<0722:TSPFAP>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<0722:TSPFAP>2.0.CO;2).
- Seneviratne, S. I., and Coauthors, 2006: Soil moisture memory in AGCM simulations: Analysis of global land-atmosphere coupling experiment (GLACE) data. *J. Hydrometeorol.*, **7**, 1090–1112, <https://doi.org/10.1175/JHM533.1>.
- , T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, 2010: Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Sci. Rev.*, **99**, 125–161, <https://doi.org/10.1016/j.earscirev.2010.02.004>.
- Sospedra-Alfonso, R., W. J. Merryfield, and V. V. Kharin, 2016a: Representation of snow in the Canadian seasonal to interannual prediction system. Part II: Potential predictability and hindcast skill. *J. Hydrometeorol.*, **17**, 2511–2535, <https://doi.org/10.1175/JHM-D-16-0027.1>.
- , L. Mudryk, W. J. Merryfield, and C. Derksen, 2016b: Representation of snow in the Canadian seasonal to interannual prediction system: Part I. Initialization. *J. Hydrometeorol.*, **17**, 1467–1488, <https://doi.org/10.1175/JHM-D-14-0223.1>.

- Thornthwaite, C. W., 1948: An approach toward a rational classification of climate. *Geogr. Rev.*, **38**, 55–94, <https://doi.org/10.2307/210739>.
- van den Hurk, B., F. Doblas-Reyes, G. Balsamo, R. D. Koster, S. I. Seneviratne, and H. Camargo Jr., 2012: Soil moisture effects on seasonal temperature and precipitation forecast scores in Europe. *Climate Dyn.*, **38**, 349–362, <https://doi.org/10.1007/s00382-010-0956-2>.
- Vicente-Serrano, S. M., S. Berguería, and J. I. López-Moreno, 2010: A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *J. Climate*, **23**, 1696–1718, <https://doi.org/10.1175/2009JCLI2909.1>.
- von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, 484 pp.
- Wang, W., and A. Kumar, 1998: A GCM assessment of atmospheric seasonal predictability associated with soil moisture anomalies over North America. *J. Geophys. Res.*, **103**, 28 637–28 646, <https://doi.org/10.1029/1998JD200010>.
- Yang, X., and T. DelSole, 2012: Systematic comparison of ENSO teleconnection patterns between models and observations. *J. Climate*, **25**, 425–446, <https://doi.org/10.1175/JCLI-D-11-00175.1>.
- Zomer, R. J., A. Trabucco, D. A. Bossio, and L. V. Verchot, 2008: Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agric. Ecosyst. Environ.*, **126**, 67–80, <https://doi.org/10.1016/j.agee.2008.01.014>.