A Characterization of the Present-Day Arctic Atmosphere in CCSM4

GIJS DE BOER,* WILLIAM CHAPMAN,[+] JENNIFER E. KAY,[#] BRIAN MEDEIROS,[#]
MATTHEW D. SHUPE,[@] STEVE VAVRUS,[&] AND JOHN WALSH**

* Lawrence Berkeley National Laboratory, Berkeley, California and Cooperative Institute for Research in
Environmental Sciences, and NOAA/ESRL/PSD, Boulder, Colorado
[+] University of Illinois at Urbana–Champaign, Urbana, Illinois
[#] National Center for Atmospheric Research, Boulder, Colorado
[@] Cooperative Institute for Research in Environmental Sciences, and NOAA/ESRL/PSD, Boulder, Colorado
[&] University of Wisconsin—Madison, Madison, Wisconsin
** University of Alaska Fairbanks, Fairbanks, Alaska

ABSTRACT

Simulation of key features of the Arctic atmosphere in the Community Climate System Model, version 4 (CCSM4) is evaluated against observational and reanalysis datasets for the present-day (1981–2005). Surface air temperature, sea level pressure, cloud cover and phase, precipitation and evaporation, the atmospheric energy budget, and lower-tropospheric stability are evaluated. Simulated surface air temperatures are found to be slightly too cold when compared with the 40-yr ECMWF Re-Analysis (ERA-40). Spatial patterns and temporal variability are well simulated. Evaluation of the sea level pressure demonstrates some large biases, most noticeably an under simulation of the Beaufort High during spring and autumn. Monthly Arctic-wide biases of up to 13 mb are reported. Cloud cover is underpredicted for all but summer months, and cloud phase is demonstrated to be different from observations. Despite low cloud cover, simulated all-sky liquid water paths are too high, while ice water path was generally too low. Precipitation is found to be excessive over much of the Arctic compared to ERA-40 and the Global Precipitation Climatology Project (GPCP) estimates. With some exceptions, evaporation is well captured by CCSM4, resulting in $P - E$ estimates that are too high. CCSM4 energy budget terms show promising agreement with estimates from several sources. The most noticeable exception to this is the top of the atmosphere (TOA) fluxes that are found to be too low while surface fluxes are found to be too high during summer months. Finally, the lower troposphere is found to be too stable when compared to ERA-40 during all times of year but particularly during spring and summer months.

1. Introduction

The observed and projected changes in the Arctic region are some of the most striking concerns surrounding climate trends (Solomon et al. 2007). Reduced summer sea ice extent and volume (e.g., Markus et al. 2009), increased permafrost melt (e.g., Hinzman et al. 2005), shifts in ecosystems and habitats (e.g., Prowse et al. 2009), and amplified signals in temperature increases observed at high latitudes (e.g., Serreze et al. 2009) are examples of changes that likely have important consequences both within the Arctic and globally. Earth system models

(ESMs) represent one of the most powerful tools in understanding these changes. A new generation of ESMs has been utilized to prepare climate projections for the fifth phase of the Coupled Model Intercomparison Project (CMIP5). Those results will be used in the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5). The objective of this study is to assess the ability of the Community Climate System Model, version 4 (CCSM4) to simulate various components of the present-day Arctic atmosphere. Six major atmospheric characteristics are evaluated because of their significant implications on regional and global climate. Properties are chosen based in part upon guidance from Walsh et al. (2005) and include surface air temperature (T_{sfc}), sea level pressure (SLP), cloud distribution and phase, precipitation and evaporation ($P - E$), the

Corresponding author address: Gijs de Boer, R/PSD3, 325 Broadway, Boulder, CO 80304.
E-mail: gijs.deboer@noaa.gov

Arctic atmospheric energy budget, and lower-tropospheric stability. To provide a basis for comparison for current results, a brief overview of recent studies involving these properties is included below.

Among the most important atmospheric characteristics, T_{sfc} represents a fundamental near-surface thermal measure of climate. It acts as a governing force in modulation of surface properties including sea ice and land cover. Chapman and Walsh (2007) compared Arctic T_{sfc} as simulated in 14 climate models used in IPCC AR4 and demonstrated them to feature a mean cold bias of 1–2 K, with a maximum regional cold bias over the Barents Sea of 8–12 K in winter and 6–8 K in spring. The previous version of CCSM (CCSM3; Collins et al. 2006) did not demonstrate this cold bias, instead having a Barents Sea warm bias of 4–6 K. On a seasonal basis, CCSM3 seasonal and annual RMSE were among the lowest of the 14 models compared, with seasonal biases of 2–3 K.

Accurate simulation of SLP is vital to prediction of future climate because of its influence on surface wind speed and direction, temperature, and precipitation. Wind speed and direction are particularly important in the Arctic for advection of simulated sea ice (DeWeaver and Bitz 2006) and governance of heat fluxes between the ocean–land surface and atmosphere. Chapman and Walsh (2007) also evaluated simulated Arctic SLP. In general, ESM-simulated storm tracks were demonstrated to be shorter than those observed, with observed storms often reaching the Kara Sea and simulated storm tracks ending in the Barents Sea (for a reference map, see Fig. 1). Together with a shift in the location and strength of the Beaufort High, this resulted in "subtle but significant" differences between observed and simulated sea ice transport. Interestingly, CCSM3 was shown to be an outlier compared to the other 13 models, with negative annual SLP biases generally ranging between −2 and −12 mb throughout the Arctic and seasonal RMSE between 3.5 mb (summer) and 13 mb (spring). These values were the highest of the 14 models compared.

Clouds remain among the largest sources of uncertainty in climate simulation (Solomon et al. 2007). In particular, high-latitude clouds (e.g., mixed-phase clouds, ice clouds) have proven challenging for models to simulate correctly (e.g., Klein et al. 2009; Kay et al. 2011), in part due to how different Arctic atmospheric conditions can be from the rest of the world. Uncertainty in high-latitude cloud simulation is hypothesized to be important because of the ability of clouds to influence changes in Arctic Ocean mixed layer temperature and sea ice melting rates (e.g., Eisenman et al. 2007; Gorodetskaya et al. 2008) and precipitation, both of which regulate surface albedo. Comparisons between GCMs and Arctic observations are challenging to complete due to inconsistencies between
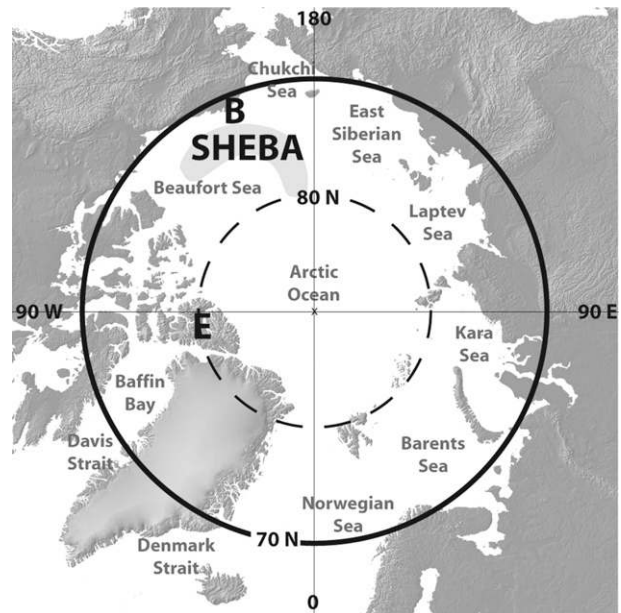

FIG. 1. A map of the Arctic. The bold circle represents the analyzed portion of the domain (70°N and above). Also shown are major bodies of water, and the locations of Eureka, Canada (E), Barrow, Alaska (B), and approximate range of locations for the SHEBA experiment (gray shading labeled SHEBA). Map topographical data courtesy of National Oceanic and Atmospheric Administration (NOAA) National Geophysical Data Center (NGDC).

cloud fraction definitions between simulations and sensors. Despite these challenges, Walsh et al. (2002) demonstrated improvement in simulation of Arctic Ocean cloud cover between Atmospheric Model Intercomparison Project version 1 (AMIP1) and AMIP2 simulations. Recent evaluations still demonstrate problems with cloud simulation in the Arctic, however. Inoue et al. (2006) compared regional climate model results with observations from the Surface Heat Budget of the Arctic Ocean experiment (SHEBA; Uttal et al. 2002), revealing discrepancies between observed and simulated cloud vertical distribution and phase. Vavrus et al. (2009) showed the ensemble mean winter and spring cloud amounts from 20 CMIP3 models to be too high. One bias observed in this comparison was a tendency for models to overpredict wintertime low-level cloudiness. This bias was addressed in CCSM4 by work outlined in Vavrus and Waliser (2008) with a parameterization (FREEZEDRY) designed to reduce cloudiness at low levels during Arctic winter.

Changes in the freshwater budget of the Arctic can have large implications for global climate. As discussed in Aagaard and Carmack (1989), Arctic freshwater plays an important role in stratifying the Arctic Ocean, with Arctic river discharge constituting roughly 10% of the world's freshwater runoff. This results in a low-salinity

and low-temperature mixed layer that impacts sea ice development and lifetime, meaning shifts in precipitation patterns both north and south of 70°N may impact Arctic sea ice extent. In addition to the influence of freshwater on Arctic Ocean structure and sea ice lifetime, Oechel et al. (1993) observed wetter sites to have lower fluxes of CO_2 to the atmosphere than drier ones, and argued that changes in precipitation and temperature patterns could result in localized lowering of water tables, resulting in increased CO_2 surface emissions. Walsh et al. (1998) provided an overview of $P - E$ as simulated by AMIP models and found ensemble mean precipitation and net surface freshwater gain exceeded observational estimates. These patterns were highly correlated between different models, implying potential contributions of increased precipitation to evaporation and vice versa. In a later study, Kattsov and Walsh (2000) demonstrated twentieth-century increases in Arctic precipitation in the ECHAM4 model, a trend also observed in observational estimates provided in the 1996 IPCC report (Trenberth et al. 1996) and a study by Karl (1998).

Understanding of the atmospheric energy budget can provide insight into the behavior of the complex Arctic climate system. Several studies have been completed to better understand how the different components of this system interact. Initially, Nakamura and Oort (1988, hereafter NO88) utilized a combination of surface in situ and satellite data to estimate the four main components governing the energy budget, namely the atmospheric storage of heat ($\partial E/\partial t$), surface fluxes (F_{SFC}), the net top of the atmosphere (TOA) radiative flux (F_{RAD}), and the transport of heat across 70°N (F_{WALL}). These terms are related in the energy budget by

$$\partial E/\partial t = F_{RAD} + F_{WALL} + F_{SFC}. \quad (1)$$

Because of scarce polar surface temperature observations, F_{SFC} was computed as a residual. Using this methodology, Serreze et al. (2007) compared energy budgets from both the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) and the National Centers for Environmental Prediction (NCEP)–National Center for Atmospheric Research (NCAR) reanalysis (NRA) products. They found F_{SFC} significantly impacts the atmospheric (and oceanic) energy budget and showed that in a mean sense the ERA-40 atmosphere demonstrated a net loss of energy between August and February. Additionally, while they considered ERA-40 to provide a valuable description of the atmospheric energy budget, they noted significant issues with top of the atmosphere (TOA) radiation, net surface flux (both due in part to snow and ice albedo parameterization) and the resulting residual transport terms, and

that these terms were different from the NRA. Porter et al. (2010) expanded on this analysis, comparing energy budgets from the NRA and the Japan Meteorological Society (JRA), while using TOA and surface radiation estimates from the Clouds and the Earth's Radiant Energy System (CERES). In this work, the authors utilized CERES data to derive F_{WALL}, resulting in reduced discrepancies between the JRA and NRA. Nevertheless, they concluded that large uncertainties remain, and that even with more advanced estimates of surface and TOA radiative fluxes, differences exist between various reanalyses.

The thermal structure of the lower troposphere controls near-surface humidity, radiation, and heat fluxes. For example, Boé et al. (2009) demonstrated temperature inversions in the lower-Arctic atmosphere regulate heat content of the upper-Arctic Ocean through modification of longwave radiation emitted to space. Because many CMIP3 models overpredict atmospheric inversion strength, it is generally believed that outgoing longwave radiation is too high, resulting in a cooling of the lower atmosphere and an underestimation of Arctic climate change. Medeiros et al. (2011) analyzed spatial distributions of wintertime inversion strengths from the 21 CMIP3 climate models and compared these distributions to those calculated from ERA-40. Comparing temperature differences between 850 and 1000 mb, they found the models to replicate the bimodal structure found in the reanalysis, with a North Atlantic unstable mode, and stable regimes over the Arctic Ocean and land areas. Overall, land regions agreed better with the reanalysis than ocean regions, but models were found to exaggerate stable and underestimate unstable regimes. Finally, the authors also compared the inversion strength from the Atmospheric Infrared Sounder (AIRS) to that found in the reanalysis and found that AIRS 1000-mb temperature tends to be biased warm compared to ERA-40, resulting in a weakened inversion strength between 850 and 1000 mb.

In this work, CCSM4's representation of each of these properties is evaluated. The following section provides an overview of CCSM4 and the datasets utilized for verification. This section is followed by the evaluation of present-day climate, a discussion including comparisons to CCSM3 and an integrated view of observed parameters and a summary of results.

2. CCSM4 and data

a. CCSM4 description

CCSM4 consists of four component models, including ones for the atmosphere [Community Atmosphere Model, version 4 (CAM4)], sea ice [Los Alamos Sea Ice Model

(CICE)], land [Community Land Model (CLM)], and ocean [Parallel Ocean Program, version 2 (POP2)]. Each of these components has seen improvements since the previous release (CCSM3), the highlights of which are outlined below. Information on all aspects of CCSM4 can be found in Gent et al. (2011). While analysis presented here focused on the atmosphere, the coupled nature of the component models results in atmospheric modification due to, for example, changes to the land surface, sea ice, or ocean characteristics. Therefore, relevant adjustments to all component models are briefly outlined.

Changes implemented in CAM4 include handling of deep convection, Arctic cloud fraction ("FREEZEDRY"), radiation, and scalability. The sea ice component of CCSM4 is new, with CICE replacing the former Community Sea Ice Model (CSIM5). CICE provides enhancements to ice physics and computational efficiency, as well as a new shortwave radiative transfer scheme, melt pond physics, and aerosol deposition for both snow and ice. CLM also includes several new capabilities, of which a new frozen soil scheme, and updates to the snow model to include aerosol deposition, grain-size dependent snow aging, vertically resolved snowpack heating, and new snow cover fraction and snow burial fraction parameterizations are most relevant to Arctic study. Potentially relevant changes to POP2 include a new near-surface eddy flux parameterization, new overflow parameterizations for the Denmark Strait, a submesoscale mixing scheme and an increased number of vertical levels.

Analyzed simulations were run featuring $0.9° \times 1.25°$ (f09_g16) resolution for atmosphere (finite volume) and land grids and a displaced pole ocean and sea ice grid (gx1v6). The atmosphere included 26 vertical levels. For each variable, up to six ensemble members were analyzed, with five producing monthly mean output, and the sixth, dubbed "Mother of All Runs" (MOAR) producing higher-frequency output. Here we analyze simulation results between 1981 and 2005, expecting this period to be representative of "present-day" climate.

b. Verification datasets

Because of the variety of atmospheric properties analyzed, no single dataset can provide verification by itself. Therefore, we evaluate model results using a combination of reanalysis products together with satellite and ground-based observations.

To evaluate simulation of T_{sfc} and SLP, CCSM4 was compared to the ERA-40. Liu et al. (2007) evaluated ERA-40 T_{sfc} against buoy and coastal estimates from the International Arctic Buoy Programme (IABP)/Polar Exchange at the Sea Surface (POLES). In that work, ERA-40 estimates were found to have consistent warm biases with a mean value of 1.48 K. This warm bias was

found throughout the Arctic, with the largest values occurring over the Laptev Sea and in the marginal ice zones and was reported to be consistent across the seasonal cycle. Maximum annual temperatures occurred one month later (August) in ERA-40 than in IABP/POLES measurements (July). Generally, biases were larger in winter than summer, which was attributed to surface air temperature over sea ice oscillating around freezing during summer months. As outlined in Bromwich et al. (2007), while ERA-40-derived estimates of surface pressure have inconsistencies in the Southern Hemisphere, the Northern Hemispheric estimates are thought to be reliable. Because ERA-40 only covers until 2002, to assure a comparison between an equal number of years, the periods covered by ERA-40 (1978–2002) and in the CCSM4 simulations (1981–2005) were slightly different. Because the simulations are not expected to represent events in actual time, but rather produce a representative climate, the length of the comparison period is more important than capturing the same exact years.

Comparison of CCSM4 cloud fraction to that from observations is filled with challenges. Two primary concerns stem from differences in threshold-defined cloud fraction and spatial sampling. Because cloud boundaries are defined differently by different sensors, thresholds utilized can result in different cloud fraction estimates. Instrument definitions are also often different from model-defined cloud fraction. A sampling concern arises because GCM cloud fraction is the fractional area of a grid box that is cloudy at a given time. Passive satellite sensors with large footprints may employ a similar definition, but may only sample a specific area at very low temporal resolution. Ground-based sensors may sample a specific location more frequently, but have a narrower field of view and are not able to cover an entire GCM grid box. Because of this, cloud fraction for ground-based sensors is defined as the fraction of time that clouds were observed at a specific location. Therefore sensor location within a GCM grid box will influence reported mean cloud fraction. To mitigate the thresholding concern, a large variety of estimates from different sensors is reported. While reanalyses provide similar sampling to GCM results, Walsh et al. (2008) demonstrated reanalysis cloud fractions to be prone to biases, making their cloud frequency estimates less reliable than those measured. The sampling issue is addressed in part by averaging over large time periods when possible (e.g., monthly or seasonally over multiple years of data).

Several years of continuous measurements are available from ground-based observations and are used for site-specific comparisons to CCSM4 simulations. Included are measurements from the U.S. Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) site at Barrow, Alaska, the Study of Environmental Arctic Change (SEARCH) site in Eureka, Canada, and SHEBA

(B, E, and SHEBA in Fig. 1, respectively). Each site features advanced cloud radars, microwave radiometers, and lidars to measure cloud presence, detect cloud phase and retrieve microphysical properties. Numerous works have utilized these measurements to analyze cloud properties, including but not limited to Shupe et al. (2011, 2008, 2005); Shupe (2011), de Boer et al. (2009), and Verlinde et al. (2007). Methods used for specific cloud property retrievals are outlined in Shupe (2007) and Shupe et al. (2006, 2005). Efforts were made to temporally match model sampling to dataset time scale. Because of the relatively short nature of records used from these sites (1–3 years), interannual variability is a concern. To analyze the impact of this variability, the 25-yr simulation period was subsampled at comparable intervals using a 2-yr running mean. This analysis revealed that for simulated total cloud fraction, the differences between the mean 2-yr mean and extreme 2-yr means were generally under 15%. Similar calculations by phase resulted in lower numbers for ice clouds (generally <8%), comparable values for liquid clouds (generally <15%), and slightly higher for mixed-phase clouds (generally <20%). This implies that, at least in the model, the 2-yr window sampled within the 25-yr present-day period generally introduces differences in cloud fraction of <15%.

Arctic-wide cloud properties are compared with estimates from several sources (similar to Vavrus and Waliser (2008)). Included are estimates from surface-based climatologies including human and instrument observations from Hahn et al. (1995), Huschke (1969), Makshtas et al. (1999), and Serreze and Barry (2005) [hereafter ground based estimates (GRDEST)]. In addition, several satellite-based platforms are utilized, including estimates from the International Cloud Climatology Project (ISCCP), the Television Infrared Observation Satellite (TIROS) Observational Sounder (TOVS) Pathfinder B, the High-Resolution Infrared Sounder (HIRS), the Moderate Resolution Imaging Spectroradiometer (MODIS), the Advanced Very High Resolution Radiometer (AVHRR), CERES (Kato et al. 2006) (hereafter SATEST), and a combined estimate from *CloudSat* and the Cloud–Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument on the Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) as in Kay and Gettelman (2009). *CloudSat*/CALIOP estimates can be split into low cloud and total cloud fractions to match model definitions. For reasons discussed previously, uncertainty in several of these measurements is large. Despite this, all estimates are included, and important differences between them are discussed in the manuscript.

Perhaps more troublesome than clouds is verification of precipitation and evaporation. Precipitation gauge issues,

including contamination by blowing snow and reduced coverage have been reported in the literature (e.g., Serreze et al. 2005). A general lack of gauge coverage over the Arctic Ocean and sea ice, along with measurement errors up to 100% make accurate characterization of precipitation over the Arctic challenging. Nevertheless, estimates have been assembled. The Global Precipitation Climatology Project (GPCP; Adler et al. 2003) combines satellite- and gauge-based estimates to produce a 2.5° × 2.5° global dataset, taking strengths and weaknesses of individual contributors into account. Serreze et al. (2005) evaluated precipitation estimates from GPCP, and ERA-40, ERA-15, and NCEP-1 reanalyses against surface gauge measurements. The analysis performed was for regions north of 45°N (thought to cover primary contributors to the Arctic watershed) and demonstrated negative median biases in most datasets. Perhaps surprisingly, GPCP estimates were demonstrated to have relatively poor skill at high latitudes when compared to reanalyses, with all three reanalysis products outperforming GPCP estimates. Because of this result, the authors recommended use of reanalysis precipitation estimates over high-latitude regions. Based in part on this evaluation, the present study presents results both in relation to GPCP estimates and those from ERA-40, with the comparison to ERA-40 discussed most thoroughly.

Arctic atmospheric energy budget estimates are compared with those from the literature. Included as sources for comparison is the work of Serreze et al. (2007), who derived estimates of individual energy budget components largely from the ERA-40 reanalysis and compared them with the NRA (Kalnay et al. 1996), and TOA radiation with that from the Earth Radiation Budget Experiment (ERBE) satellite. ERA-40 was found to provide valuable insight into the atmospheric energy budget but was labeled as deficient in its representation of TOA and surface radiation. Additionally, we compare CCSM4 values to estimates provided in Porter et al. (2010), who used NRA and JRA (Onogi et al. 2007) reanalyses for their energy budget calculations along with estimates of TOA and surface radiances from CERES. In Porter et al. (2010) it was concluded that there was no evidence of either the JRA or NRA providing better estimates of the energy budget terms.

For analysis of lower-tropospheric temperature structure, results from ERA-40 and ERA-Interim reanalyses are utilized based on findings from Medeiros et al. (2011), Tjernström and Graversen (2009), and Zhang et al. (2011). They found temperature inversions from these products to more accurately portray lower-tropospheric temperatures than those estimated by AIRS. Tjernström and Graversen (2009) found the ERA-40 to demonstrate a surface warm bias when compared to the SHEBA dataset.
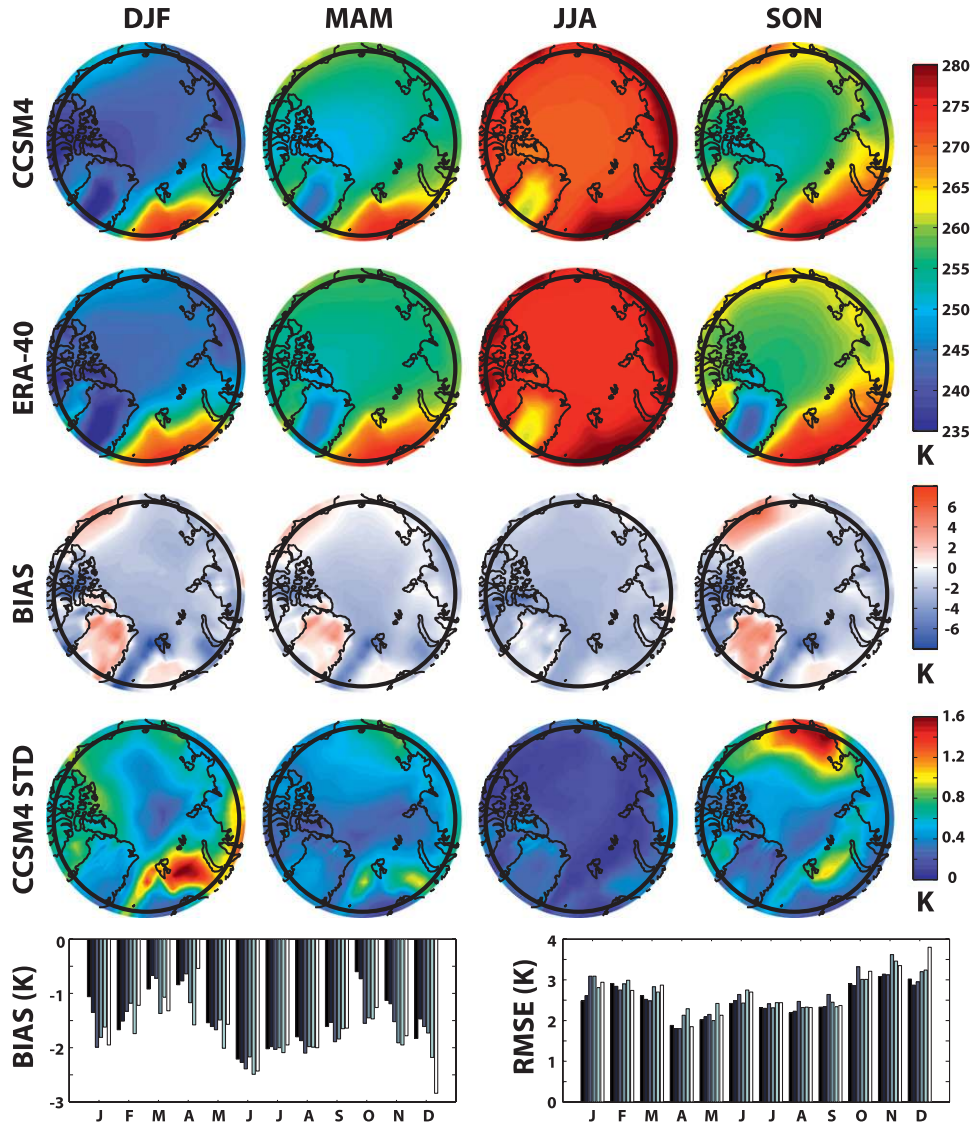
FIG. 2. (left to right)Seasonal mean surface air temperature as simulated by (top to bottom) the first CCSM4 ensemble member, from ERA-40, model bias (CCSM4-ERA-40), the standard deviation between all six CCSM4 ensemble members, and Arctic-wide CCSM4 bias and RMSE for all six ensemble members (different bars) per month. 70°N is indicated by the bold black ring near the perimeter of each map.

Despite the combination of this with a slight cool bias in midtropospheric temperatures, ERA-40 was found to represent the overall vertical structure of the SHEBA atmosphere quite well. A notable challenge in using these datasets is the coarse vertical resolution. The ERA-40 dataset used assumes standard pressure levels (1000, 925, 850, 775, 700, 600, 500 mb in the lower atmosphere) and is used to evaluate CCSM4 output interpolated to the same levels. This spacing likely fails to capture some of the more complex inversion structures that have been observed in the Arctic using radiosondes and other measurement devices.

3. Analysis of present-day climate

a. Surface air temperature

Comparisons of monthly, seasonal, and annual mean surface air temperature (T_{sfc}) are presented in Fig. 2. Illustrated are seasonal mean T_{sfc} (TREFHT in CCSM4) from the first ensemble member (top row), from ERA-40 (second row), and the difference between the two (CCSM4 − ERA-40, third row). As noted above, to compare 25-yr periods, the results presented are between the years 1978–2002 for the ERA-40 surface temperatures and 1981–2005 for values from CCSM4. As was found in

Chapman and Walsh (2007) for other GCMs, CCSM4 generally simulates Arctic T_{sfc} accurately across all seasons. Spatial patterns in T_{sfc} from ERA-40 are matched by CCSM4. With a few exceptions, biases are negative but small (generally <2 K). Some of these apparent negative biases are likely the result of the positive bias in ERA-40 temperatures discussed in Section 2b. The most notable biases include a cold bias of up to 9 K along eastern Greenland during winter. This cold bias is present year-round, but the magnitude is decreased during the rest of the year. The north shore of Alaska is shown to be too warm for much of the year, with a maximum warm bias during autumn of up to 6 K. Much of the Arctic Ocean is too cold, with cold biases of 1–3 K, while Greenland is too warm for all seasons except summer.

The fourth row of Fig. 2 illustrates the standard deviation in mean T_{sfc} between CCSM4 ensemble members. This standard deviation provides information on how different results from individual simulation ensemble members are. Low standard deviation values provide evidence that the patterns observed were predictable by the model and that the comparison of the first ensemble member discussed above is valid. High values indicate that there was a lot of disagreement between the ensemble members, likely implying that the temperature differences are due to elements that the model struggles to capture, reducing confidence in the comparison provided above. In general, differences are small (standard deviations <0.8 K). On an Arctic-wide basis, winter has the most variability between simulations, with many areas having standard deviations of 0.4–0.9 K. Summer demonstrates the smallest differences between simulations, with standard deviations under 0.4 K. Extreme intersimulation variability is found in winter over the Barents Sea (~1.6 K). There is a comparable amount of deviation over the East Siberian–Chukchi Seas during autumn (~1.4 K) and a small signature over the Barents Sea and Fram Straight during spring (0.8 ~1.0 K).

The last row of Fig. 2 shows biases (left) and root-mean-squared differences (right) as calculated for the six CCSM4 ensemble members relative to ERA-40 for the entire Arctic (70°–90°N), with individual bars representing the six ensemble members. On average, CCSM4 T_{sfc} is lower than that from ERA-40, with the largest differences occurring in summer and winter. The large summer bias is due mainly to underprediction of T_{sfc} across most of the Arctic Ocean, the Canadian Archipelago, eastern Siberia, and Alaska. RMSE ranges between roughly 2 K in April to 3 K in November and December. Biases range between −0.75 K in March and April to −2.5 K in June. Seasonal RMSE and bias are generally consistent between ensemble members, with the largest differences occurring in October and April.

In addition to monthly means, MOAR output and ERA-40 estimates are analyzed to examine the ability of CCSM4 to capture the distribution of 6-hourly mean T_{sfc}. Comparing these distributions measures CCSM4's ability to represent not only seasonal and monthly mean T_{sfc}, but also extreme values within a given month. While not necessarily capturing a full diurnal cycle, 6-hourly sampling provides diurnal and day to day variability occurring with events having submonthly time scales (e.g., synoptic frontal passages). Figure 3 illustrates this analysis. The small cold bias is evident, with CCSM4 T_{sfc} distributions (black) shifted toward colder temperatures compared with those from ERA-40 (gray). Generally, CCSM4 accurately captures distribution widths; meaning that the range of T_{sfc} covered with 6-hourly sampling is close to that from the reanalysis. Months with larger open water fractions tend to have a primary or secondary distribution peak near 273 K. During colder months (November–March), most temperatures sampled fall between 235 and 250 K. For a small number of cases, there are differences in the extremes of the distributions. From May–September, for example, CCSM4 tends to include more warm cases, with T_{sfc} up to 312 K in July, when the ERA-40 distribution only extends to 308 K. Similarly, in December and January the left tail of the CCSM4 distribution reaches ~210 K, roughly 5 K colder than ERA-40. These instances only make up a very small fraction of the total distribution, however, and likely have little influence on mean Arctic climate.

b. Sea level pressure

While simulated T_{sfc} agrees well with ERA-40, SLP demonstrates large and significant differences. Similar to Fig. 2, Fig. 4 illustrates seasonal mean SLP biases of up to 14 mb or greater for the first CCSM4 ensemble member. Most strikingly, CCSM4 severely underpredicts the Beaufort High, a main driver for advection of sea ice throughout the Arctic Ocean. This underprediction is largest during spring and autumn, with biases exceeding −14 mb. This bias is reduced during winter, but even then the bias approaches −10 mb. In addition, CCSM4 demonstrates a wintertime negative bias in SLP over the Norwegian Sea of ~10 mb, indicating that the storm track in the model includes deeper cyclones than those in the reanalysis. Summer demonstrates the closest comparison to reanalysis surface pressures, with some small (4 mb) negative biases over the Arctic Ocean.

The fourth row of Fig. 4 illustrates ensemble standard deviation. In general, summer and autumn feature smaller SLP differences than winter and spring, with standard deviations generally <1 mb. Extreme standard deviation of 2.5 mb and above is found over the Kara Sea during winter, possibly representing differences in the length of
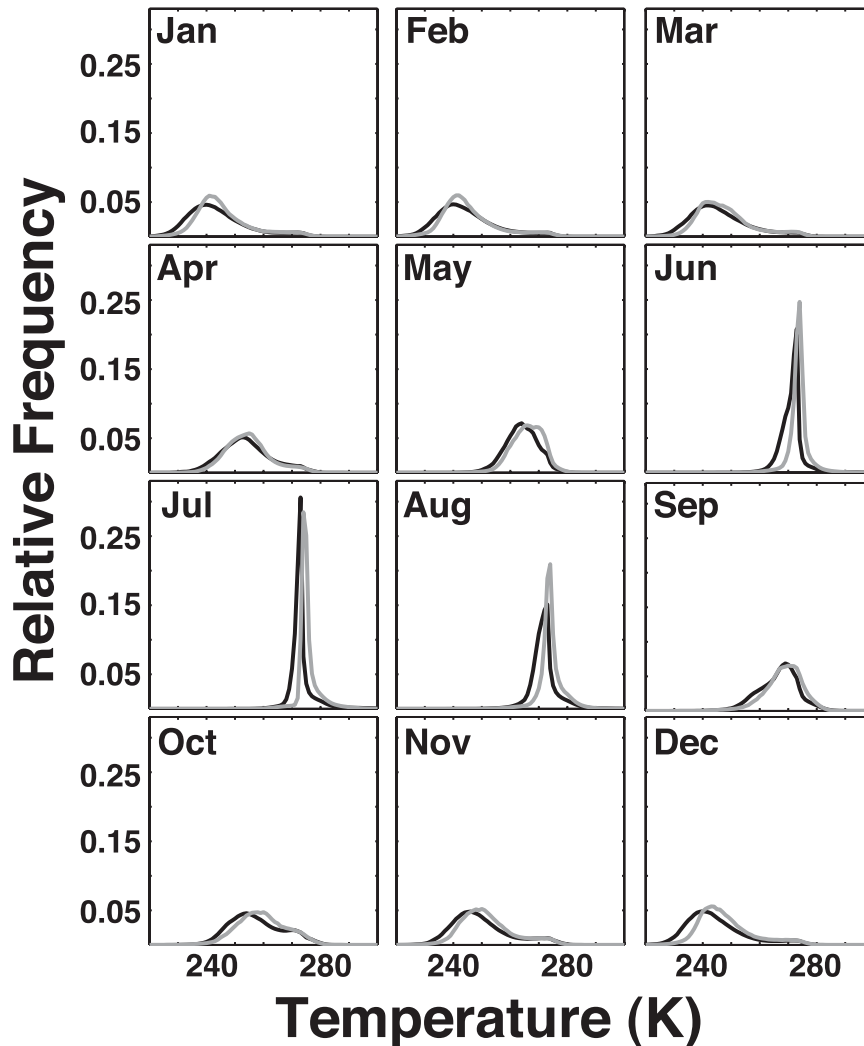
FIG. 3. Monthly 6-hourly mean 2-m temperature distributions for 70°–90°N from CCSM4 MOAR (black) and ERA-40 (gray). CCSM4 was sampled between 1981 and 2005, while ERA-40 was sampled from 1981 to 2002.

simulated storm tracks coming out of the North Atlantic Ocean. There is also some variability in the spring SLP field extending into the Kara and Laptev Seas (~2 mb) and a smaller amount of variability over the East Siberian Sea during autumn (~1.25 mb). These differences do not appear to be significant when compared to the large differences between CCSM4 ensemble member 1 and the ERA-40 reanalysis, in particular because the locations of maximum standard deviations do not correlate with the largest differences between CCSM4 and ERA-40.

The bottom row of Fig. 4 shows Arctic wide RMSE and bias values for all ensemble members, relative to ERA-40. This clearly demonstrates the generally negative SLP bias in CCSM4. With the exception of June, every month has negative biases, with the largest occurring in March and April (6–13 mb). RMSE is largest in

spring, with values of up to 13 mb and smallest in summer (2–4 mb). While different ensemble members generally show the same patterns on a month-to-month basis, differences between them can be quite large. In a given month, ensemble to ensemble differences in bias and RMSE reach 8 mb during winter and spring, the times of largest intersimulation variability.

With the exception of the Beaufort High, CCSM4 does appear to correctly simulate semipermanent SLP features impacting Arctic circulation. These circulations do not necessarily occur north of 70° (and therefore are not shown in Fig. 4) but impact Arctic circulation sufficiently to include in this discussion. The Siberian High is simulated to be in place for much of the year, with a weakening during the summer months. This center of circulation does appear to be displaced slightly to the
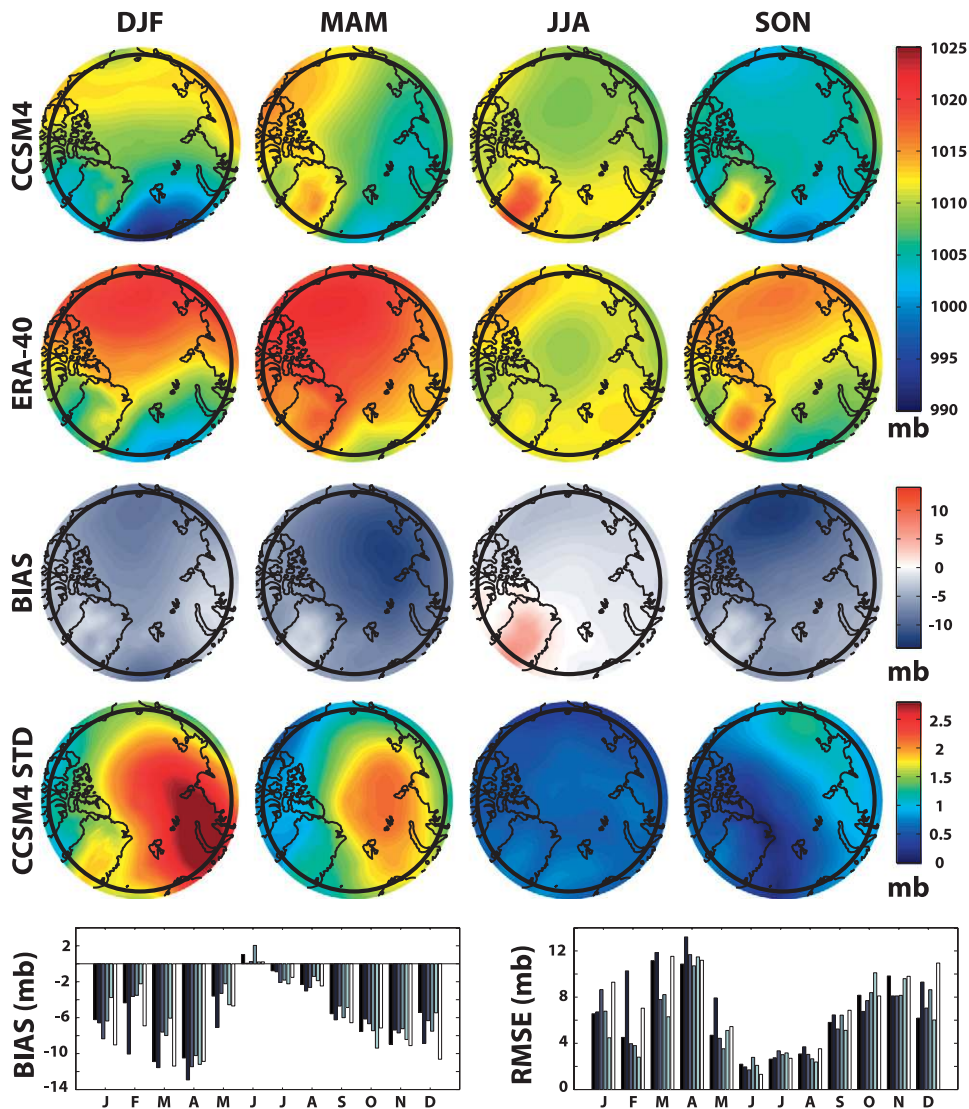
FIG. 4. As in Fig. 2, but for sea level pressure.

south, particularly during winter, resulting in a negative bias over northern Siberia during that time period, and a positive bias further to the south. The Aleutian low, generally found off of the southern coast of Alaska except during summer, is also present in CCSM4. During spring, it appears to be displaced slightly westward, resulting in negative SLP biases over eastern Siberia, and positive SLP biases over southern Alaska. Finally, the Icelandic Low is also present in CCSM4, however, it is generally too strong. This is evident from negative SLP biases over the Nordic, Barents, and Kara Seas during autumn, winter, and spring. Together, these SLP features influence modes of variability like the Arctic and North Atlantic Oscillations (AO and NAO, respectively). Evaluation of the model's ability to simulate these modes is reviewed in J. Hurrell et al. (2011, personal communication).

c. Clouds

As discussed earlier, we present a wide range of methods and measurements for cloud evaluation. Arctic-wide, CCSM4 appears to underestimate total cloud cover for much of the year compared to satellite and ground-based observations. Figure 5 illustrates Arctic cloud cover estimates from a suite of satellite sensors (SATEST, medium-gray bars), surface observations (GRDEST, dark gray bars), seasonal *CloudSat*/CALIPSO estimates (dashed lines), observations from the three ground-based sites used in the current study (light gray bars), and the CCSM4 ensemble spread (black bars). As touched upon above, individual bar widths are due to different mechanisms. For example, the spread in GRDEST is due in part to different sensors employed in different studies
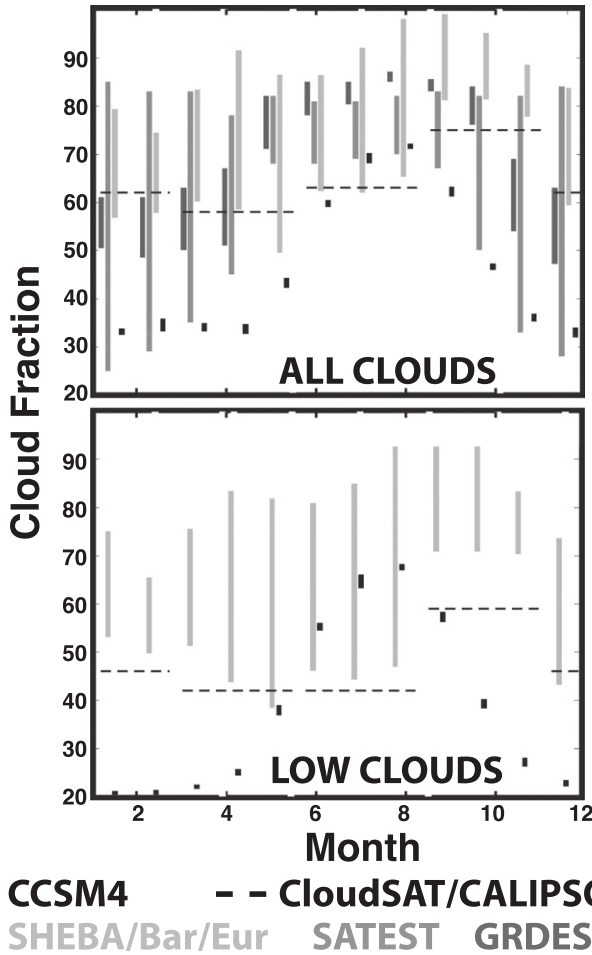
FIG. 5. CCSM cloud fraction for the entire Arctic region (70°–90°N) plotted with estimates of cloud fraction from several satellite and ground-based sources (see text for details). Comparisons are included for (top) all clouds and (bottom) low clouds only.

and the locations included. SATEST differences are largely due to instrument specifications, sampling, and thresholds used in cloud-detection algorithms. The idea is that combined presentation of various estimates from different locations and with different thresholds employed begins to capture the true variability in Arctic cloud cover and that despite dataset differences, patterns emerge. When compared to all other estimates, CCSM4 provides the lowest cloud fraction for all but summer months. Differences reported in winter and spring appear to be larger than would be expected because of the sampling and thresholding issues alone. In general, the seasonal pattern of cloudiness is captured in simulations, with more cloudiness during summer and early autumn and minimum cloud coverage during winter. The lower half of Fig. 5 shows a similar comparison for low cloud fraction. Again, there are large differences between *CloudSat*/CALIPSO estimates (dashed lines),
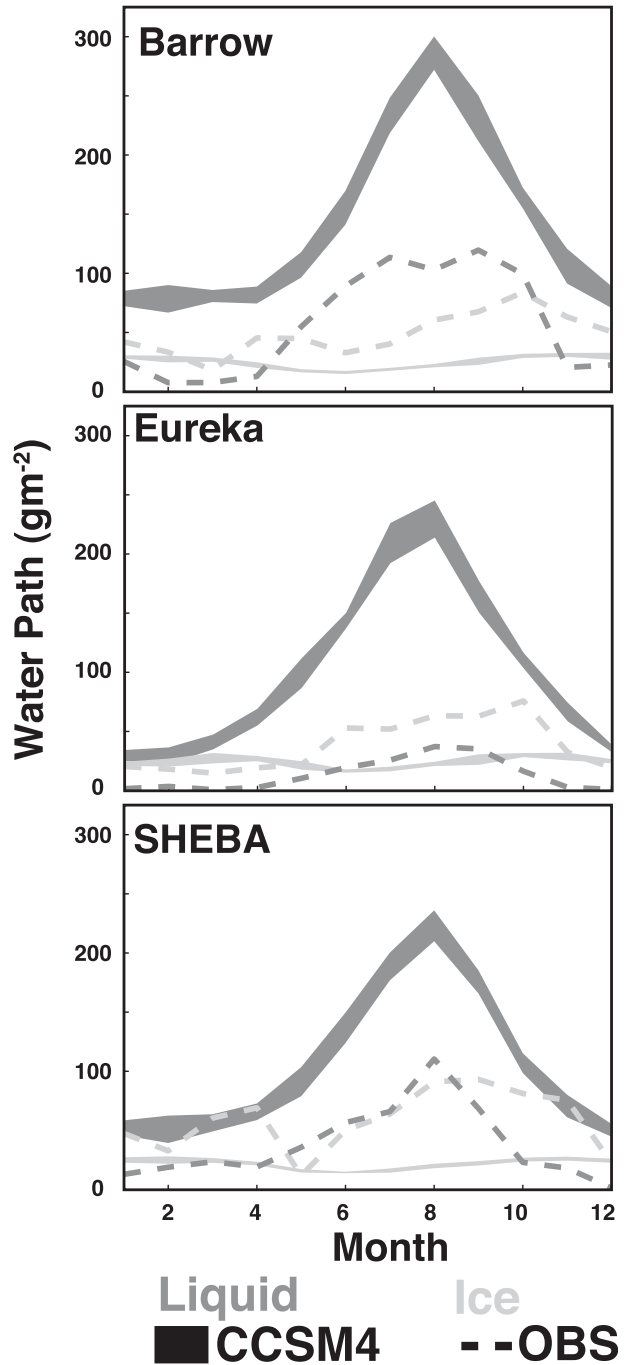

FIG. 6. CCSM4 (shading) and observationally derived (dashed-lines) all-sky liquid (darker) and ice (lighter) water paths for three observation sites.

surface measurements from SHEBA, Barrow and Eureka (light gray lines), and the CCSM4 simulations (black bars). CCSM4 appears to grossly underestimate low cloud occurrence during all but summer months.

A more direct comparison of cloud properties is illustrated in Fig. 6. Here CCSM4 total grid box liquid (LWP)
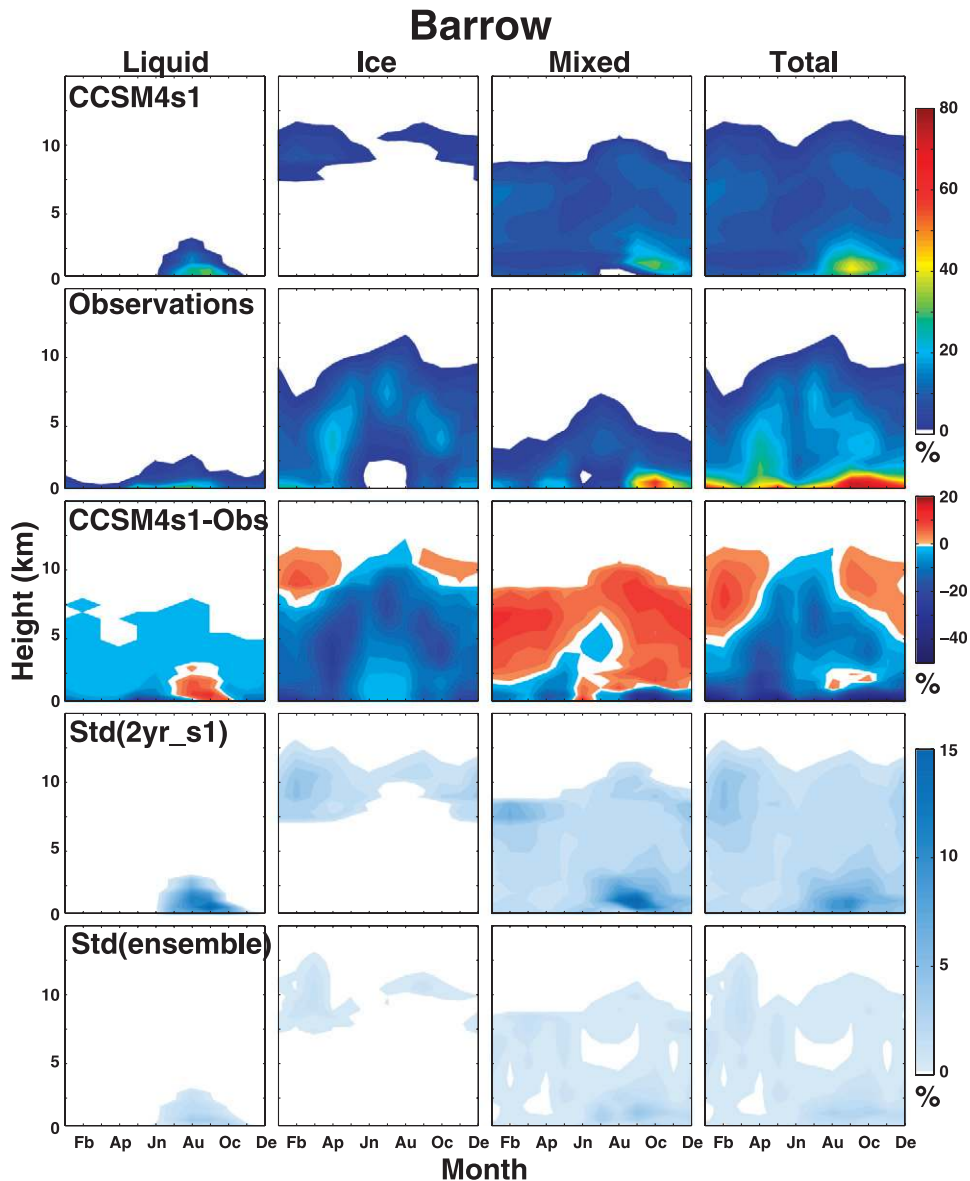
Barrow

FIG. 7. (left to right) Mean 2-yr mean cloud fraction plotted by phase for Barrow, Alaska, from the (top) first CCSM4 ensemble member and (second from top) ground-based observations. (middle) The difference between the two fields is plotted. (second from bottom) The calculated standard deviation between simulated 2-yr means (over the 25-yr period of interest) is shown, and (bottom) standard deviation of mean 2-yr mean values between ensemble members is plotted.

and ice water paths (IWP) for the grid point closest to the observational sites are compared with retrieved estimates of all-sky LWP and IWP from surface sensors. Simulated LWP and IWP demonstrate similar patterns between different locations, with LWP peaking during summer and exceeding IWP year-round. IWP is demonstrated to be relatively constant throughout the year, with a slight decrease during late spring and summer months. Comparison with observations reveals that despite underpredicting cloud fraction, CCSM4 overpredicts LWP.

The opposite is true for ice, as CCSM4 underpredicts all-sky IWP, particularly during summer and autumn months. Winter is generally an exception, with simulated IWP comparable to that observed.

Comparison of location-specific cloud cover for three ground-based observational sites provides insight into these differences in water path. Figure 7 compares mean 2-yr cloud fraction by phase for Barrow, Alaska, between the first ensemble member (first row) and ground-based observations (second row). This evaluation compares

measurements from a single point to grid-average quantities from the simulation. While some differences may be expected as a result, in general the clouds observed at Barrow are stratiform in nature and cover large areas. Therefore when comparing monthly averages, these differences are likely to be small. Liquid-only clouds are slightly overestimated during late summer and early autumn but underpredicted the rest of the year. They are also underpredicted at higher altitudes, likely the result of temperatures below the 268 K threshold required in CCSM4 for liquid-only clouds. Ice-only clouds are severely underrepresented at low altitudes, again likely due to temperature-dependent phase partitioning. Also evident in the ice-only clouds is an overprediction of clouds at higher altitudes, particularly during winter months. Mixed phase clouds are underpredicted at low levels and overpredicted at higher altitudes. Because temperature-dependent partitioning follows a linear relationship, many high-level simulated mixed-phase clouds have unrealistic ratios of water to ice mass. Comparisons with SHEBA and Eureka (not shown) yielded similar results. It seems that most high-altitude mixed-phase clouds in CCSM4 should in fact be ice-only clouds, particularly during summer. This would increase IWP during this time and decrease unrealistically high LWP. A reduction of high-level mixed-phase clouds during winter would also reduce winter LWP, bringing the simulated LWP closer to that observed.

To understand potential impacts of interannual variability, standard deviation in simulated 2-yr mean cloud fraction at Barrow between 1981 and 2005 was calculated (4th row, Fig. 7). Assuming variability in observations is comparable to that simulated, large standard deviations between 2-yr means would indicate that perceived biases may be due only to the period of available observations. In general, standard deviations are found to be small (<5%), with the larger variability occurring in low liquid and mixed-phase clouds in late summer and autumn. Also analyzed was the variability between ensemble members (bottom row, Fig. 7). Differences between individual simulations were small, with maximum differences occurring for low-level mixed-phase and liquid clouds (roughly 5%).

d. Precipitation–evaporation

Prior to discussing validation of CCSM4 precipitation, it should be noted that differences between primary validation datasets (ERA-40 and GPCP) are comparable to apparent biases of CCSM4 relative to either dataset. As mentioned, observations of Arctic precipitation are subject to large uncertainties, arising in part from gauge undercatch of snow and preferential siting of gauges at low elevations. Serreze et al. (2005) concluded ERA-40 was the most successful global reanalysis in capturing Arctic

precipitation and its variations in space and time. While ERA-40 appears to be the best validation product, it is important to recognize that its clouds and radiation are modeled, and that these simulations are not subject to constraints on radiation components making large errors possible.

Evaluation of CCSM4 surface moisture flux ($P - E$) is illustrated in Fig. 8. In general, CCSM4 recreates spatial patterns demonstrated in ERA-40 (Fig. 8, top two rows). There are high $P - E$ values over coastal Greenland, the Kara Sea, and western Russia during all seasons in both CCSM4 and the reanalysis. Both indicate increased Arctic Ocean $P - E$ during summer and autumn, and demonstrate minimum $P - E$ over the Nordic and Barents Seas during autumn, winter, and spring. During summer, both have minimum $P - E$ over Siberia, but CCSM4 $P - E$ is excessive over the Arctic throughout the year (Fig. 8, third row). This is particularly true at lower latitudes and along coastal Greenland. Based on evaluation of both precipitation and evaporation separately, this seems to be due to excessive model precipitation. Evaluation of the standard deviation between CCSM4 simulations indicates very little variation between the six runs. The main exceptions to this occur along coastal Greenland, where seasonal mean standard deviations of up to 0.35 mm day^{-1} occur.

CCSM4 precipitation amounts show a positive bias relative to ERA-40 (1978–2002) over most of the Arctic (Fig. 9, top). Expressed as a percentage of the total, biases are largest and positive over the Greenland and Alaskan mountains but also strongly positive over much of the Arctic Ocean in all seasons except summer and are negative over the Barents Sea during winter. This implies that CCSM4 oversimulates Arctic Ocean precipitation relative to ERA-40. South of 70°N, CCSM4 captures precipitation maxima known to exist in the North Atlantic and North Pacific storm tracks, and in areas of coastal orography (southeast Alaska, Canada's western coast, southern Greenland, and western Norway). These patterns are potentially important for correctly capturing runoff amounts into the Arctic Ocean. Biases in CCSM4's precipitation amounts relative to GPCP (1978–2002, Fig. 9, bottom) are generally smaller than those relative to ERA-40, especially over mountainous areas. As was the case with ERA-40, CCSM4's biases are largest and positive (+40% to +150%) over the Arctic Ocean, northern Greenland, and northern North America. Positive biases over northern Alaska and Greenland are apparent year-round, while the Arctic Ocean bias is less apparent in summer. The Arctic Ocean seems to be a region of oversimulated precipitation no matter which database is used for validation.

With some notable exceptions, evaporation patterns are captured accurately by CCSM4 (not shown). Exceptions
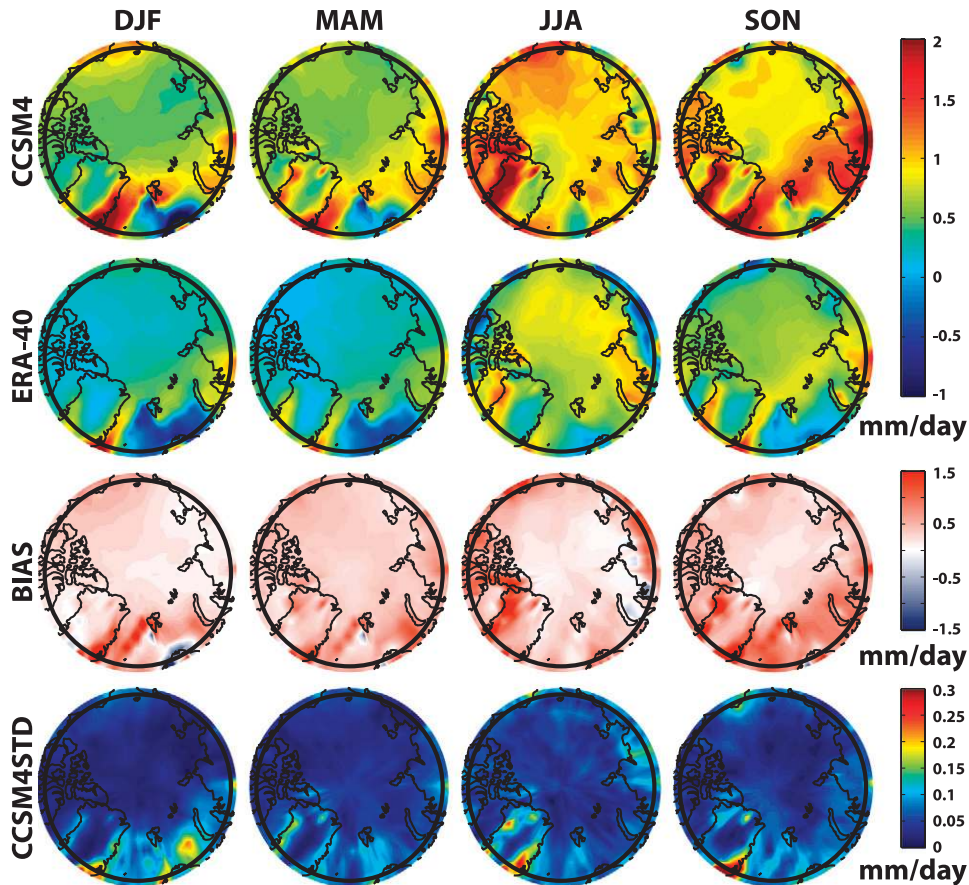
FIG. 8. As in Fig. 2, but for $P - E$.

occur southeast of Greenland and over the Norwegian Sea, where CCSM4 has much higher levels of surface evaporation than ERA-40. This is particularly true during winter and autumn, when model evaporation exceeds that in ERA-40 by up to 3.5 mm day^{-1}. In addition, there are two main regions of reduced model evaporation. These include the Greenland and Barents Seas. The magnitude of decreased evaporation in these regions is not as extreme as the Norwegian Sea increase, with maximum negative biases of 2 mm day^{-1}. When combined with the precipitation comparison above, there is a general over-prediction of Arctic Ocean $P - E$ in CCSM4 compared to ERA-40. This implies that CCSM4 has more moisture collection at the surface than the reanalysis product does. These values are generally small (0–1 mm day^{-1}), with the largest discrepancies occurring during winter over the Greenland Sea (1.5 mm day^{-1}).

e. Energy budget

Monthly ensemble-mean estimates of atmospheric energy budget components are presented in Table 1. Evaluation of these numbers becomes tricky due to the temporal evolution of these terms. In their analysis, Porter et al. (2010) only included a 5-yr climatology of JRA and NRA budgets, covering November 2000 through October 2005. Serreze et al. (2007) on the other hand, included a longer time period, covering 1979–2001. Therefore, Table 1 includes estimates covering the 1981–2005 period and estimates for 2000–05 (in parentheses). Differences between the two time periods are generally small.

In comparing CCSM4 estimates to those from reanalyses there are some encouraging similarities. The CCSM4 annual mean net surface flux of 14 W m^{-2} matches the JRA estimate and is within 3 W m^{-2} of ERA-40. The 5-yr NRA analysis provides a slightly lower estimate of 5 W m^{-2}. Interestingly, analysis of the Arctic Ocean energy budget by Serreze et al. (2007) indicated that the 11 W m^{-2} result from ERA-40 is likely too large, implying the CCSM4 derived estimate is also too large.

Looking at the annual cycle (Fig. 10), the CCSM4 seasonal amplitudes of F_{RAD} and F_{SFC} are too small when compared to observations. These budgets are especially deficient in summer, when CCSM4 has a negative bias. Both 5- and 25-yr CCSM4 F_{RAD} terms are higher than
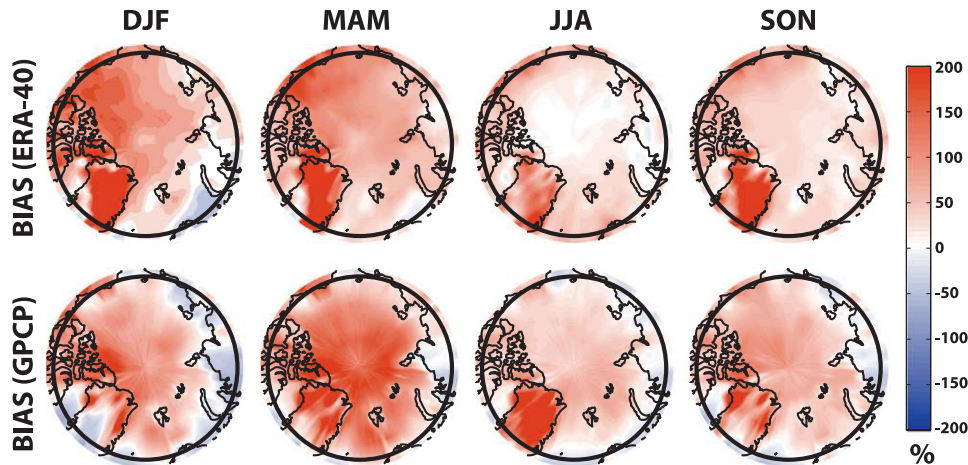
FIG. 9. (left to right) Seasonal mean CCSM4 precipitation biases as compared to (top) ERA-40 and (bottom) GPCP. 70°N is indicated by the bold black ring near the perimeter of each map.

measured and reanalysis estimates in January and February, are roughly equal for October, November, December, and March and are too low for the rest of the year. The negative bias in summertime F_{RAD} likely results from errors in cloud simulation and associated impacts on shortwave radiation.

Calculated as a residual, the annual mean contribution from atmospheric transport (F_{WALL}) of 102 W m^{-2} is similar to estimates from ERA-40 (100 W m^{-2}) and NRA (103 W m^{-2} for 1981–2005 and 104 W m^{-2} for 2001–05). The JRA estimate is smaller, (94 W m^{-2}). This difference between NRA and JRA estimates from Porter et al. (2010) was hypothesized to result from differences in F_{SFC} terms. Specifically, differences in summertime surface sensible heating was indicated as a major contributor to derived differences. Seasonally, F_{WALL} acts as expected, with energy convergence into the pole resulting from differential TOA solar radiation with latitude. As discussed in Serreze et al. (2007), these values are largest in winter due to a maximum in the meridional temperature gradient. Despite correctly representing seasonal cycle phase, amplitude is quite different at times from those derived from reanalyses. For example, May's 5-yr F_{WALL} of 104 W m^{-2} compares with estimates of 99 and 106 W m^{-2} for JRA and NRA, respectively. At the same time, the 25-yr value of 109 W m^{-2} compares less favorably with ERA-40 and NRA estimates of 66 and 77 W m^{-2}.

f. Boundary layer structure

Medeiros et al. (2011) and Boé et al. (2009) note that models often feature excessive surface inversions when compared with observations and CCSM4 is no exception. The first evaluation used is similar to that employed in Medeiros et al. (2011), in which differences in monthly

mean T_{sfc} and T_{850mb} are used as an indicator of mean lower-tropospheric stability. Shown in Fig. 11, analysis of the frequency of occurrence of different temperature gradients demonstrates CCSM4 ensemble members (black lines) are generally too stable, with larger increases (or smaller decreases) in temperature between 850 mb and the surface than those reported by ERA-40 and ERA-Interim. There is little difference between distributions for the five included ensemble members (MOAR not included due to issues with output files). Similarly, despite covering different time periods (1981–2002 for ERA-40 and 1989–2005 for ERA-Interim), differences between ERA-40 and ERA-Interim are generally small. The contrast between open Arctic Ocean and sea ice surfaces shows up as a bimodal distribution in the ocean distributions (top), with a secondary peak on the far left end of the distributions for all seasons besides summer when the lower atmosphere

TABLE 1. Monthly ensemble-mean atmospheric energy budget estimates from CCSM4 in W m^{-2}. Values in parentheses are for the period between 2001 and 2005 only, while values outside of the parentheses are for 1981–2005.

| Month	$\partial E/\partial t$	F_{RAD}	F_{SFC}	F_{WALL}
January	−6 (−4)	−166 (−167)	50 (50)	110 (112)
February	3 (1)	−162 (−163)	46 (46)	119 (118)
March	6 (6)	−141 (−142)	37 (37)	111 (110)
April	23 (21)	−100 (−99)	17 (18)	105 (101)
May	42 (45)	−51 (−50)	−12 (−14)	104 (109)
June	29 (34)	−8 (−4)	−51 (−56)	89 (94)
July	11 (9)	−9 (−5)	−66 (−68)	86 (83)
August	−11 (−11)	−78 (−77)	−29 (−30)	96 (97)
September	−32 (−32)	−148 (−148)	20 (20)	97 (96)
October	−31 (−32)	−180 (−181)	51 (53)	99 (96)
November	−21 (−15)	−179 (−182)	55 (55)	104 (112)
December	−13 (−16)	−171 (−172)	52 (54)	106 (102)
Mean	0 (0)	−116 (−116)	14 (14)	102 (102)
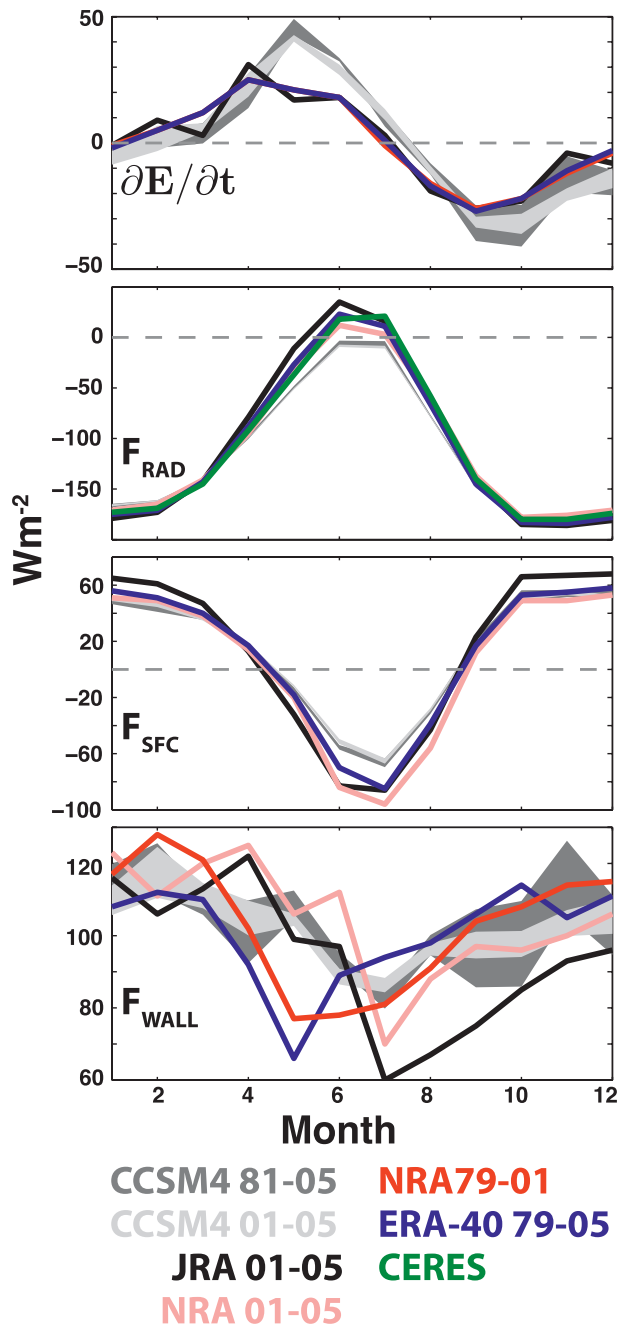
FIG. 10. A comparison of (top to bottom) CCSM4 energy budget terms for both 25- (darker shading) and 5-yr (lighter shading) comparison periods. The shading represents the ensemble spread. Included in the comparison are estimates from the JRA and NRA reanalysis for 2001–05 (black and pink lines, from Porter et al. 2010), from the NRA reanalysis for 1979–2001, and the ERA-40 reanalysis from 1979–2005 (red and blue lines, from Serreze et al. 2007) and from CERES for F_{RAD} (green line, from Porter et al. 2010).

appears well mixed. During all seasons, CCSM4 tends to have a more stable mean atmospheric state, with changes in temperature between the surface and 850 mb of up to 10–20 K.

The pattern of excessive surface inversions holds true over land as well (Fig. 11, bottom panels), with CCSM4 temperature gradients more positive than those from ERA-40 and ERA-Interim. Again a noticeable secondary peak forms at the left edge of the distribution, likely due to regions where snow and ice melt thereby leaving soil or vegetated surfaces behind. Both ocean and land samples demonstrate the largest positive temperature gradients during spring, possibly resulting from snow-covered surfaces having warmer air advected over them.

Spatial distributions of Arctic inversion frequency (not shown) are similar between different CCSM4 ensemble members. During winter and spring, CCSM4 has its most unstable air located over the Norwegian and Barents Seas. The rest of the Arctic, particularly areas over the Canadian Arctic Archipelago, Siberia and Arctic sea ice demonstrate surface inversions. Summer and autumn feature a more muted pattern with the Norwegian and Barents Seas along with Siberia and Alaska having the most unstable air, and only portions of extreme northern Canada (e.g., Ellesmere Island) and the perennial pack ice to the north of Greenland and the Archipelago featuring weak surface inversions. This pattern is generally similar to that from ERA-40, with some notable exceptions being decreased inversion frequency Arctic wide in ERA-40 during summer months and decreased inversion frequency along northern Eurasia during autumn.

Another analysis method involves collecting high-frequency samples from the MOAR and deriving surface inversion strength using 6-hourly mean atmospheric temperatures at the coarse vertical resolutions of CCSM4 and ERA-40. Monthly distributions of these strengths along with similar calculations from ERA-40 are illustrated in Fig. 12. For all months, CCSM4 has stronger inversions than ERA-40. While ERA-40 inversions generally remain at 20 K km^{-1} or weaker, CCSM4 produces inversion that can be twice as strong or more. The percentages shown in Fig. 12 indicate the percentage of samples that featured an inversion. CCSM4 is shown to produce more frequent inversions in all months. This is particularly true from April through October, when CCSM4 inversion frequency exceeds that from ERA-40 by 30% or more. Both CCSM4 and ERA-40 capture increased inversions during July, likely due to warm air temperatures being advected over a relatively cool ocean surface.

4. Discussion

a. Comparison to CCSM3

To provide insight into differences between model versions, comparisons between CCSM4 and CCSM3
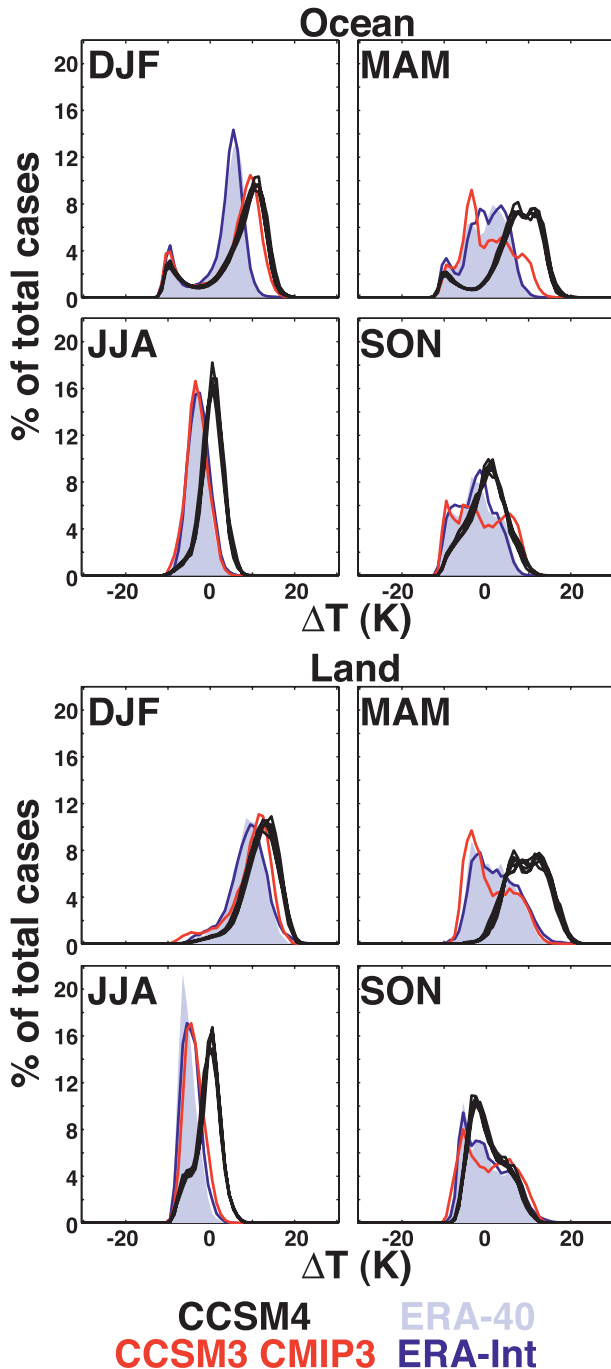
FIG. 11. Seasonal distributions of atmospheric temperature differences between 850 mb and the surface for (top) ocean and (bottom) land areas. Shown are the six CCSM4 ensemble members (black lines), the CCSM3 CMIP3 twentieth-century simulation (red), the ERA-Interim reanalysis product for 1989–2005 (darker blue line), and the ERA-40 reanalysis product for 1981–2002 (light blue shading).

simulations are made for variables evaluated for CCSM3 in the literature.

Looking first at T_{sfc}, unlike many IPCC AR4 GCMs, CCSM3 featured a strong (up to 7 K in the annual mean)

warm bias over the Barents Sea and the marginal ice zone north of Scandinavia. This warm bias appears to be reduced in CCSM4. Though still present, summer cold biases are also decreased over Alaska and eastern Siberia. Winter cold biases along the eastern coast of Greenland have increased in magnitude from CCSM3, speculated to be a result of the change from CISM to CICE. In addition, negative summer biases over the central Arctic Ocean have increased. Finally, positive autumn T_{sfc} biases over the Beaufort sea have increased from roughly 1–2 to 6 K. Comparing RMSE between CCSM3 and CCSM4, CCSM3 featured T_{sfc} RMSE of 3, 2.8, 2, and 2.5 K for winter, spring, summer, and autumn, respectively (Chapman and Walsh 2007), which seems to have been improved upon slightly in CCSM4, particularly during spring.

A comparison of the absolute value of SLP biases of CCSM3 and CCSM4 relative to ERA-40 for 1978–2002 (not shown) shows that on an annual basis, biases are reduced in CCSM4. The largest improvements occur over the Canadian Archipelago and Norwegian Sea, with bias reductions of 3–5 mb. The only region within the Arctic that sees increased biases is the Laptev Sea, though these changes are small at roughly 1 mb. On a seasonal basis, differences between model versions are more dramatic. In winter, improvements over the Western Hemisphere are still clear, but the Eastern Hemisphere from Svalbard to the Laptev Sea shows increased biases of 1–2 mb. Spring biases over the North Slope of Alaska, Canadian Archipelago, and Norwegian Sea are reduced by up to 7 mb. This period also has the largest bias increase over the Laptev Sea, with CCSM4 biases roughly 3 mb larger than those for the same region in CCSM3. Summer and autumn both demonstrate bias decreases of up to 3 mb over the Norwegian Sea region. Improvements in the Arctic for all seasons appear balanced by deteriorated performance at slightly lower latitudes, with a ring of higher biases present at sub-Arctic latitudes. Interestingly, analysis of CCSM3 completed by Chapman and Walsh (2007) revealed a negative correlation between Barents Sea SLP biases and temperature biases over the same region. CCSM4 does not demonstrate this same correlation, as despite having negative SLP bias in autumn, winter, and spring, the Barents Sea temperature bias is near zero during those seasons.

Comparison between CCSM3- and CCSM4-simulated lower-tropospheric stability demonstrates that over the ocean, CCSM3 more accurately captures observed temperature gradients in the lower Arctic atmosphere, particularly during summer. Both model versions are too stable during winter, and both seem to accurately represent the fraction of the Arctic with an open ocean surface during winter based on agreement between the far left side of model and reanalysis distributions shown
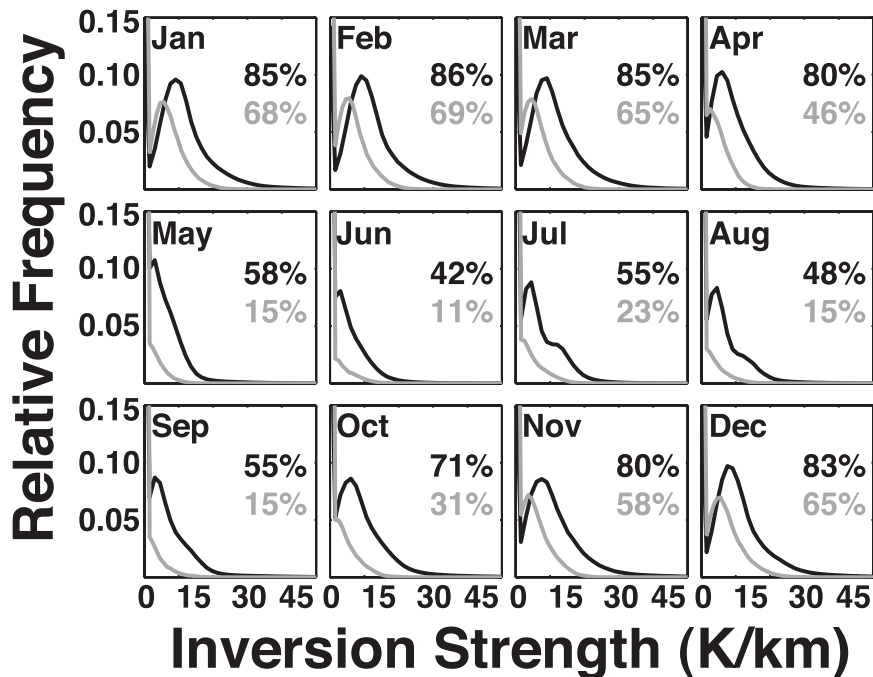
FIG. 12. Monthly inversion strength distributions for 70°–90°N, sampled at 6-hourly intervals from CCSM4 MOAR (black) and ERA-40 (gray). CCSM4 was sampled between 1981 and 2005, while ERA-40 was sampled from 1981 to 2002. The percentages reported indicate the frequency of occurrence of any inversion for the entire Arctic for any given month.

in Fig. 11. Transition seasons appear to give both versions of the model trouble, though CCSM4 clearly overpredicts lower-atmospheric stability, particularly during spring. Over land, similar patterns appear. CCSM3 matches the reanalysis products more closely during summer, with CCSM4 more stable during that time. Both versions are too stable during winter months. Spring and autumn stability are more accurately simulated over land, particularly by CCSM3. CCSM4 has an atmosphere that is generally far too stable during the spring.

Similar to CCSM4 clouds, Gorodetskaya et al. (2008) demonstrated CCSM3 to feature clouds with excessive LWP when compared to the SHEBA and the National Aeronautics and Space Administration (NASA) Water Vapor Project (NVAP) datasets. In the same study, cloud fraction over oceanic regions north of 70°N was shown to be lower than estimates from TOVS for most of the year, excluding summer. Vavrus and Waliser (2008) also looked at cloud amounts over the Arctic and illustrated CCSM3 to significantly overpredict low and total cloud amounts compared to ground-based observations. Both CCSM3 evaluations demonstrated higher cloud fraction than found in the present study for CCSM4. The seasonal cycle, with more clouds in summer, is present in both models, but winter cloud fraction estimates have dropped from 60%–70% to 30% while summer estimates are comparable in all studies (70%–80%). While this is likely

in part due to the FREEZEDRY parameterization, it is important to note that preliminary evaluations of global cloud cover in CCSM4 also demonstrate underestimation of cloud fraction outside of the Arctic, making the relative contribution of FREEZEDRY uncertain.

b. Earth system integration

While evaluation of each variable is presented individually in this work, they are undoubtedly interconnected, as well as connected to different parts of the earth system. Certainly different surface states (ice, water, snow, and soil) will impact surface temperature. For example, both Baffin Bay and the Greenland sea, where CCSM4 demonstrates strong negative temperature biases during autumn, are shown by Jahn et al. (2012) to have excessive sea ice coverage. Similarly, the coastal Beaufort Sea, where CCSM4 T_{sfc} was too large, is shown to have insufficient coverage during autumn. In addition, this coastal Beaufort Sea temperature bias could be the result of decreased cold air advection associated with reduced anticyclonic circulation over the Beaufort Sea, a pattern evident in the SLP analysis. In a broader sense, it is not surprising that elevated standard deviations in simulated T_{sfc} appear to follow the marginal sea ice zone since ice extent between ensemble members is likely to vary. Interactions between atmosphere and surface work in both directions and changes in the atmospheric circulation

brought on by SLP biases, most notably in the form of a weakened Beaufort High, directly impact the advection of sea ice. These modifications to sea ice motion are also discussed in more detail by Jahn et al. (2012). Similar interactions also occur within the atmospheric system. For example, excessive cloud LWP reported is very likely playing a major role in discrepancies noted in the energy budget's F_{RAD} and F_{SFC} terms. Overly thick clouds during summer reduce the amount of solar radiation reaching the surface while increasing the amount of sunlight reflected to the top of the atmosphere. These overly thick clouds could also increase the outgoing longwave radiation, further altering F_{RAD}. Along these lines, Kay et al. (2012) showed that large cloud LWPs in CAM4, similar to those illustrated for CCSM4 in the present analysis, result in a strong negative Arctic cloud feedback with increased greenhouse gases, impacting simulations of future climate. A final example of the complexity of these interactions can be found in the simulation of permafrost in the land model. While CCSM4 demonstrates apparent improvements in the general simulation of permafrost and in simulated T_{sfc} across eastern Alaska and Western Canada, permafrost does not appear in this region despite observational evidence that it should. Initial analysis suggests that this lack of permafrost is the result of insulation from excessive simulated winter snow depths (Lawrence et al. 2012). Errors in any individual component, along with tuning of cloud fraction in climate models to accurately simulate global radiative balance, make diagnosing feedback loops featuring several components using earth system models a challenging task. Additionally, gauging the impact of individual errors on the prediction of future climate is an important component of understanding model results. An example of how results from the present evaluation impact analysis of SLP in future climate simulations is discussed in Vavrus et al. (2012).

5. Summary

An evaluation of the ability of CCSM4 to represent the present-day Arctic atmospheric climate has been performed using a six-member ensemble of CCSM4 simulations for the late twentieth century. Included are evaluations of T_{sfc}, SLP, cloud frequency, thickness and phase, precipitation and evaporation, the large-scale energy budget, and lower-tropospheric stability. Both spatial and temporal patterns are evaluated, along with Arctic-wide variability using high temporal resolution sampling from the MOAR.

- Simulated seasonal spatial patterns in T_{sfc} are similar to those from ERA-40 and Arctic-wide mean biases

are small (generally <2 K). The largest differences between ensemble members and between CCSM4 and ERA-40 occur in the marginal ice zone. CCSM4 generally features colder mean temperatures than ERA-40 but otherwise accurately simulates distributions of 6-hourly mean temperature.

- The Beaufort High is virtually missing from simulated SLP fields. This results in negative Arctic-wide mean biases of up to 12 mb during spring. Overall, CCSM4 presents a negative sea level pressure bias when compared to ERA-40.

- Arctic-wide cloud cover is too low in CCSM4 compared to several observational estimates. Despite this, all-sky LWP is found to be significantly too high in CCSM4. Generally, CCSM4 features lower IWP than observed. Much of this seems to be attributable to the temperature-based phase partitioning employed in CAM4. Comparison with surface observations shows underestimation of liquid clouds for all seasons except summer. Mixed-phase clouds are underpredicted at low altitudes and overpredicted at higher altitudes, and ice clouds are underpredicted at all but the highest altitudes. This pattern is similar at three different observational sites.

- Arctic precipitation is oversimulated when compared to ERA-40 or GPCP. Although these biases are difficult to fully dissect because of the potential for large errors in verification datasets, biases of 40%–150% are calculated over northern North America, northern Greenland, and the Arctic Ocean. Evaporation in the Arctic is simulated with increased accuracy, with the largest differences occurring over the Norwegian Sea, where evaporation is oversimulated by up to 3.5 mm day^{-1}. Combined, CCSM4 $P - E$ is generally too high over much of the Arctic, particularly over coastal Greenland.

- CCSM4 overpredicts surface energy fluxes during summer months and underpredicts it during winter. Here, F_{RAD} is underpredicted during summer months, with CCSM4 featuring negative flux year-round. Seasonal horizontal transport of energy into the Arctic seems to follow expected patterns, and the annual mean transport is similar to that calculated from ERA-40 and NRA.

- The atmospheric boundary layer in CCSM4 is too stable as compared to ERA-40 and ERA-Interim. This is particularly true during spring over both land and ocean surfaces. The strengths of surface inversions were found to be too great in CCSM4 when compared to ERA-40, with distributions showing a near-doubling of strength. CCSM4 is found to have more inversions than ERA-40 for all months.

CCSM4 provides a consistent representation of present-day Arctic climate. In doing so it represents individual

components of the Arctic atmosphere with respectable accuracy. Yet shortcomings presented illustrate there is room for improvement. Work on the next generation of CCSM [Community Earth System Model, version 1 (CESM1)] has been completed and most aspects of the atmospheric model (CAM5) have been updated. Because of these updates, an evaluation of CESM1 simulations will likely have very different results than those from the current evaluation of CCSM4. It remains to be seen whether the changes implemented in CESM1 will result in improved simulation of elements demonstrating major deficiencies in this analysis, such as errors in the sea level pressure field, cloud phase, and boundary layer stability. Additional work needs to be completed to better understand causes of these and other errors discovered here, as well as to evaluate the performance of CCSM4 relative to other earth system models. In addition, to assure a fair comparison, continued efforts toward improving measurement datasets for model evaluation are required.

REFERENCES

Aagaard, K., and E. Carmack, 1989: The role of sea ice and other freshwater in the Arctic Mediterranean seas. *J. Geophys. Res.,* **94,** 14 485–14 498.

Adler, R., and Coauthors, 2003: The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeor.,* **4,** 1147–1167.

Boé, J., A. Hall, and X. Qu, 2009: Current GCMs' unrealistic negative feedback in the Arctic. *J. Climate,* **22,** 4682–4695.

Bromwich, D., R. Fogt, K. Hodges, and J. Walsh, 2007: A tropospheric assessment of the ERA-40, NCEP, and JRA-25 global reanalyses in the polar regions. *J. Geophys. Res.,* **112,** D10111, doi:10.1029/2006JD007.

Chapman, W., and J. Walsh, 2007: Simulations of Arctic temperature and pressure by global coupled models. *J. Climate,* **20,** 609–632.

Collins, W., and Coauthors, 2006: The Community Climate System Model, version 3 (CCSM3). *J. Climate,* **19,** 2122–2143.

de Boer, G., E. Eloranta, and M. Shupe, 2009: Arctic mixed-phase stratiform cloud properties from multiple years of surface-based measurements at two high-latitude locations. *J. Atmos. Sci.,* **66,** 2874–2887.

DeWeaver, E., and C. Bitz, 2006: Atmospheric circulation and its effect on Arctic sea ice in CCSM3 simulations at medium and high resolution. *J. Climate,* **19,** 2415–2436.

Eisenman, I., N. Untersteiner, and J. Wettlaufer, 2007: On the reliability of simulated Arctic sea ice in global climate models. *Geophys. Res. Lett.,* **34,** L10501, doi:10.1029/2007GL029914.

Gent, P., and Coauthors, 2011: The Community Climate System Model, version 4. *J. Climate,* **24,** 4973–4991.

Gorodetskaya, I., L.-B. Tremblay, B. Liepert, M. Cane, and R. Cullather, 2008: The influence of cloud and surface properties on the Arctic Ocean shortwave radiation budget in coupled models. *J. Climate,* **21,** 866–882.

Hahn, C., S. Warren, and J. London, 1995: The effect of moonlight on observation of cloud cover at night, and application to cloud climatology. *J. Climate,* **8,** 1429–1446.

Hinzman, L., and Coauthors, 2005: Evidence and implications of recent climate change in northern Alaska and other Arctic regions. *Climatic Change,* **72,** 251–298.

Huschke, R., 1969: Arctic cloud statistics from "air-calibrated" surface weather observations. Rand Corporation Memo. Tech. Rep. RM-6173-PR, 79 pp.

Inoue, J., J. Liu, J. Pinto, and J. Curry, 2006: Intercomparison of Arctic regional climate models: Modeling clouds and radiation for SHEBA in May 1998. *J. Climate,* **19,** 4167–4178.

Jahn, A., and Coauthors, 2012: Late-twentieth-century simulation of Arctic sea ice and ocean properties in the CCSM4. *J. Climate,* **25,** 1431–1452.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Karl, T., 1998: Regional trends and variations of temperature and precipitation. *The Regional Impacts of Climate Change: An Assessment of Vulnerability,* R. T. Watson, M. C. Zinyowera, and R. H. Moss, Eds., Cambridge University Press, 411–426.

Kato, S., N. Loeb, P. Minnis, J. Francis, T. Charlock, D. Rutan, E. Clothiaux, and S. Sun-Mack, 2006: Seasonal and interannual variations of top-of-atmosphere irradiance and cloud cover over polar regions derived from the CERES dataset. *Geophys. Res. Lett.,* **33,** L19804, doi:10.1029/2006GL026685.

Kattsov, V., and J. Walsh, 2000: Twentieth-century trends of Arctic precipitation from observational data and a climate model simulation. *J. Climate,* **13,** 1362–1370.

Kay, J., and A. Gettelman, 2009: Cloud influence on and response to seasonal Arctic sea ice loss. *J. Geophys. Res.,* **114,** D18204, doi:10.1029/2009JD011773.

——, K. Raeder, A. Gettelman, and J. Anderson, 2011: The boundary layer response to recent Arctic sea ice loss and implications for high-latitude climate feedbacks. *J. Climate,* **24,** 428–447.

——, M. Holland, C. Bitz, A. Gettelman, E. Blanchard-Wrigglesworth, A. Conley, and D. Bailey, 2012: The influence

of local feedbacks and northward heat transport on the equi-librium Arctic climate response to increased greenhouse gas forcing. *J. Climate,* in press.

Klein, S., and Coauthors, 2009: Intercomparison of model simula-tions of mixed-phase clouds observed during the ARM mixed-phase arctic cloud experiment. Part I: Single-layer cloud. *Quart. J. Roy. Meteor. Soc.,* **135,** 979–1002.

Lawrence, D., A. Slater, and S. Swenson, 2012: Simulation of present-day and future permafrost and seasonally frozen ground conditions in CCSM4. *J. Climate,* **25,** 2207–2225.

Liu, J., Z. Zhang, Y. Hu, L. Chen, Y. Dai, and X. Ren, 2007: As-sessment of surface air temperature over the Arctic Ocean in reanalysis and IPCC AR4 model simulations with IABP/POLES observations. *J. Geophys. Res.,* **113,** D10105, doi:10.1029/2007JD009380.

Makshtas, A., E. L Andreas, P. Snvashchennikov, and V. Timachev, 1999: Accounting for clouds in sea ice models. *Atmos. Res.,* **52,** 77–113.

Markus, T., J. Stroeve, and J. Miller, 2009: Recent changes in Arctic sea ice melt onset, freezeup and melt season length. *J. Geophys. Res.,* **114,** C12024, doi:10.1029/2009JC005436.

Medeiros, B., C. Deser, R. Tomas, and J. Kay, 2011: Arctic in-version strength in climate models. *J. Climate,* **24,** 4733–4740.

Nakamura, N., and A. Oort, 1988: Atmospheric heat budgets of the polar regions. *J. Geophys. Res.,* **93** (D8), 9510–9524.

Oechel, W., S. Hastings, G. Courlitis, M. Jenkins, G. Riechers, and N. Grulke, 1993: Recent changes in Arctic tundra ecosystems from a new carbon sink to a source. *Nature,* **361,** 520–523.

Onogi, K., and Coauthors, 2007: The JRA-25 Reanalysis. *J. Meteor. Soc. Japan,* **85,** 369–432.

Porter, D., J. Cassano, M. Serreze, and D. Kindig, 2010: New es-timates of the large-scale Arctic atmospheric energy budget. *J. Geophys. Res.,* **115,** D08108, doi:10.1029/2009JD012653.

Prowse, T., C. Furgal, F. Wrona, and J. Reist, 2009: Implications of climate change for northern Canada: Freshwater, marine, and terrestrial ecosystems. *Ambio,* **38,** 282–289.

Serreze, M., and R. Barry, 2005: *The Arctic Climate System.* Cambridge University Press, 385 pp.

——, A. Barrett, and F. Lo, 2005: Northern high-latitude pre-cipitation as depicted by atmospheric reanalyses and satellite retrievals. *Mon. Wea. Rev.,* **133,** 3407–3430.

——, ——, A. Slater, M. Steele, J. Zhang, and K. Trenberth, 2007: The large-scale energy budget of the Arctic. *J. Geophys. Res.,* **112,** D11122, doi:10.1029/2006JD008230.

——, ——, J. Stroeve, D. Kindig, and M. Holland, 2009: The emergence of a surface-based Arctic amplification. *Cryosphere,* **3,** 11–19.

Shupe, M., 2007: A ground-based multisensor cloud phase classifier. *Geophys. Res. Lett.,* **34,** L22809, doi:10.1029/2007GL031008.

——, 2011: Clouds at Arctic atmospheric observatories. Part II: Thermodynamic phase characteristics. *J. Appl. Meteor. Cli-matol.,* **50,** 645–661.

——, T. Uttal, and S. Matrosov, 2005: Arctic cloud microphysics retrievals from surface-based remote sensors at SHEBA. *J. Appl. Meteor.,* **44,** 1544–1562.

——, S. Matrosov, and T. Uttal, 2006: Arctic mixed-phase cloud properties derived from surface-based sensors at SHEBA. *J. Atmos. Sci.,* **63,** 697–711.

——, P. Kollias, P. Persson, and G. McFarquhar, 2008: Vertical motions in Arctic mixed-phase stratiform clouds. *J. Atmos. Sci.,* **65,** 1304–1322.

——, V. Walden, E. Eloranta, T. Uttal, J. Campbell, S. Starkweather, and M. Shiobara, 2011: Clouds at Arctic atmospheric obser-vatories. Part I: Occurrence and macrophysical properties. *J. Appl. Meteor. Cliamatol.,* **50,** 626–644.

Solomon, S., D. Qin, M. Manning, M. Marquis, K. Averyt, M. M. B. Tignor, H. L. Miller Jr., and Z. Chen, Eds., 2007: *Climate Change 2007: The Physical Science Basis.* Cambridge Uni-versity Press, 996 pp.

Tjernström, M., and R. Graversen, 2009: The vertical structure of the lower Arctic troposphere analyzed from observations and the ERA-40 reanalysis. *Quart. J. Roy. Meteor. Soc.,* **135,** 431–443.

Trenberth, K. E., J. T. Houghton, and L. G. Meira Filho, 1996: *Climate Change 1995: The Science of Climate Change.* Cam-bridge University Press, 584 pp.

Uttal, T., and Coauthors, 2002: Surface heat budget of the Arctic Ocean. *Bull. Amer. Meteor. Soc.,* **83,** 255–275.

Vavrus, S., and D. Waliser, 2008: An improved parameterization for simulating Arctic cloud amount in the CCSM3 climate model. *J. Climate,* **21,** 5673–5687.

——, ——, A. Schweiger, and J. Francis, 2009: Simulations of 20th and 21st century Arctic cloud amount in the global climate models assesed in the IPCC AR4. *Climate Dyn.,* **33,** 1099–1115.

——, M. Holland, A. Jahn, D. Bailey, and B. Blazey, 2012: Twenty-first-century Arctic climate change in CCSM4. *J. Climate,* **25,** 2696–2710.

Verlinde, J., and Coauthors, 2007: The mixed-phase arctic cloud experiment. *Bull. Amer. Meteor. Soc.,* **88,** 205–221.

Walsh, J., V. Kattsov, D. Portis, and V. Meleshko, 1998: Arctic precipitation and evaporation: Model results and observa-tional estimates. *J. Climate,* **11,** 72–87.

——, ——, W. Chapman, V. Govokova, and T. Pavlova, 2002: Comparison of Arctic climate simulations by uncoupled and coupled global models. *J. Climate,* **15,** 1429–1446.

——, S. Vavrus, and W. Chapman, 2005: Workshop on modeling of the Arctic atmosphere. *Bull. Amer. Meteor. Soc.,* **86,** 845–852.

——, W. Chapman, and D. Portis, 2008: Arctic cloud fraction and radiative fluxes in atmospheric reanalyses. *J. Climate,* **22,** 2316–2334.

Zhang, Y., D. Seidel, J.-C. Golaz, C. Deser, and R. Tomas, 2011: Climatological characteristics of Arctic and Antarctic surface-based inversions. *J. Climate,* **24,** 5167–5186.