Snow Mass over North America: Observations and Results from the Second Phase of the Atmospheric Model Intercomparison Project

ALLAN FREI

Department of Geography, Hunter College, Program in Earth and Environmental Sciences, City University of New York, New York, New York

ROSS BROWN

Meteorological Service of Canada, Dorval, Quebec, Canada

JAMES A. MILLER

Department of Geography, Arizona State University, Phoenix, Arizona

DAVID A. ROBINSON

Department of Geography, Rutgers–The State University of New Jersey, Piscataway, New Jersey

(Manuscript received 1 November 2004, in final form 2 March 2005)

ABSTRACT

Eighteen global atmospheric general circulation models (AGCMs) participating in the second phase of the Atmospheric Model Intercomparison Project (AMIP-2) are evaluated for their ability to simulate the observed spatial and temporal variability in snow mass, or water equivalent (SWE), over North America during the AMIP-2 period (1979–95). The evaluation is based on a new gridded SWE dataset developed from objective analysis of daily snow depth observations from Canada and the United States with snow density estimated from a simple snowpack model. Most AMIP-2 models simulate the seasonal timing and the relative spatial patterns of continental-scale SWE fairly well. However, there is a tendency to overestimate the rate of ablation during spring, and significant between-model variability is found in every aspect of the simulations, and at every spatial scale analyzed. For example, on the continental scale, the peak monthly SWE integrated over the North American continent in AMIP-2 models varies between $\pm 50\%$ of the observed value of ~ 1500 km^3. The volume of water in the snowpack, and the magnitudes of model errors, are significant in comparison to major fluxes in the continental water balance. It also appears that the median result from the suite of models tends to do a better job of estimating climatological mean features than any individual model. Year-to-year variations in large-scale SWE are only weakly correlated to observed variations, indicating that sea surface temperatures (specified from observations as boundary conditions) do not drive interannual variations of SWE in these models. These results have implications for simulations of the large-scale hydrologic cycle and for climate change impact assessments.

1. Introduction

The Atmospheric Model Intercomparison Project (AMIP) was initiated in 1989 under the auspices of the World Climate Research Program. Its mission is to systematically compare and evaluate atmospheric general circulation models (AGCMs) that have been developed by an international array of research institutes for investigation of climate change issues (Gates 1992). This study was carried out as diagnostic subproject number 28 with a mandate to evaluate GCM simulations of snow cover. In the first phase of AMIP (AMIP-1) (Gates et al. 1999), Frei and Robinson (1998) evaluated the ability of 27 AGCMs to simulate the spatial extent of snow, or snow-covered area (SCA), over North America. SCA is important for its direct impacts on the

Corresponding author address: Allan Frei, Dept. of Geography, Hunter College, The Graduate Center, Program in Earth and Environmental Sciences, City University of New York, 695 Park Ave., New York, NY 10021.
E-mail: afrei@hunter.cuny.edu

surface energy budget and indirect impacts via feedback mechanisms on temperatures, meridional temperature gradients, atmospheric greenhouse gases, atmospheric and oceanic circulation, and local effects on stability and cloudiness (Cohen 1994; Barry 2002). AMIP-1 models displayed seasonal biases in SCA over both North America and Eurasia, and underestimated interannual variability.

The second phase of AMIP (AMIP-2), initiated in 1999, includes results from a more recent generation of models, with improved parameterizations, a longer integration period, and increased spatial resolution. Frei et al. (2003) noted major improvements in AMIP-2 SCA simulations, including the elimination of some temporal and spatial biases in the seasonal cycle of continental-scale SCA variations, as well as more realistic interannual variability. Over Eurasia, interannual variability of SCA tends to be underestimated; over North America, however, interannual variability tends to be more realistic, although significant between-model variability remains.

In the past, the lack of a high quality, global gridded snow water equivalent (SWE) dataset has been an obstacle for evaluating snow mass in GCMs. Previous evaluations of GCM snow mass (Foster et al. 1996; Yang et al. 1999) used information derived from either passive microwave data—which tends to underestimate snow depth, is unreliable over mountainous and forested regions, and for which the algorithms cannot be applied with confidence to all regions of North America (Armstrong and Brodzik 2002)—or from the U.S. Air Force (USAF) snow depth climatology (Foster and Davy 1988), which was derived from a station observation network that had inadequate coverage to capture the true spatial variability of the snowpack. To address these problems, a major effort was made as part of this AMIP-2 snow evaluation subproject to develop a high quality SWE dataset for GCM evaluation. The resulting gridded dataset (Brown et al. 2003) provides monthly mean values of SWE over the North American continent for model evaluation purposes. Unfortunately, this process could not be extended to Eurasia because of insufficient in situ snow data during the AMIP-2 period.

The purpose of this study is to complement the evaluation of SCA in AMIP-2 models (Frei et al. 2003) by evaluating AMIP-2 SWE simulations. Toward that end, we examine the abilities of the models to simulate continental-scale monthly and seasonal variations in SWE over North America (NA; section 5); simulate broad spatial patterns of SWE across North America (section 6); and simulate regional-scale SWE variations over subregions of NA where the observational network is most dense and reliable for estimating regional values (section 7). First, brief overviews of the data (section 2), models (section 3), and methodology (section 4) are presented. The results sections are followed by discussion and conclusions (section 8).

2. Data

a. Gridded SWE dataset for North America

The primary dataset used for model evaluation is the gridded SWE dataset produced specifically for this project by Brown et al. (2003). The dataset covers the entire NA landmass but for this study the spatial domain was confined to the region between 20° and 70° because of difficulties resolving the Arctic archipelago in the SWE analysis. SWE values were estimated from a hybrid approach based on optimal interpolation of daily snow depth observations from over 8000 stations from the United States and Canada, with snow density estimated from a simple snowpack model. The snow depth analysis is based on the scheme developed by Brasnett (1999) and employed operationally at the Canadian Meteorological Centre (CMC). The analysis was run on a 0.3° Gaussian grid and reinterpolated to a regular 0.25° latitude–longitude for use in applied studies. The first-guess field for the optimal interpolation uses a simple snow accumulation, aging, and melt model driven by 6-hourly values of air temperature and precipitation from the 15-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-analysis (ERA-15) with extensions from the Tropical Ocean Global Atmosphere (TOGA) operational data archive to cover the full AMIP-2 period (1979–96). The optimal-interpolation process includes elevation weighting to maintain the integrity of observed data in mountainous regions. Gridded snow depth and derived SWE values were found to agree well with available independent in situ and satellite data over midlatitudinal regions of the North America, and the snow depth climatology exhibited several improvements over the USAF product of Foster and Davy (1988). While Brown et al. (2003) is considered the currently best available estimate of large-scale variability in SWE over NA for the AMIP-2 period, reliability does vary spatially with the density of the observed snow depth reporting network. The authors have the greatest confidence in the product in "data rich" areas of NA with relatively small orographic effects, which are mainly located south of ~55°N and east of the Rocky Mountains. This dataset is available from the Canadian Cryospheric Information Network (see online at www.ccin.ca).

TABLE 1. Summary of atmospheric general circulation models evaluated in this analysis.

No.	Model acronym	Research institute	Resolution Lat (°)	Resolution Lon (°)
1	CCCMA	Canadian Centre for Climate Modelling and Analysis, Canada	3.75	3.75
2	CCSR	Center for Climate System Research, Japan	2.81	2.81
3	CNRM	Centre National de Recherches Meteorologiques, France	2.81	2.81
4	COLA	Center for Ocean–Land–Atmosphere Studies, United States	1.76	2.81
5	DNM	Department of Numerical Mathematics, Russia	3.91	5.00
6	ECMWF	European Centre for Medium-Range Weather Forecasts, United Kingdom	1.96	2.00
7	GISS	Goddard Institute for Space Studies, United States	3.91	5.00
8	GLA	Goddard Laboratory for Atmospheres, United States	3.91	5.00
9	JMA	Japanese Meteorological Agency, Japan	1.88	1.88
10	MPI	Max-Planck-Institut für Meteorologie, Germany	2.81	2.81
11	MRI	Meteorological Research Institute, Japan	2.81	2.81
12	NCAR	National Center for Atmospheric Research, United States	2.81	2.81
13	PNNL	Pacific Northwest National Laboratory, United States	2.81	2.81
14	SUNYA	The University at Albany, State University of New York, United States	2.81	2.81
15	UGAMP	The U.K. Universities' Global Atmospheric Modelling Programme, United Kingdom	2.50	3.75
16	UIUC	University of Illinois at Urbana–Champaign, United States	3.91	5.00
17	UKMO	U.K. Meteorological Office, United Kingdom	2.50	3.75
18	YONU	Yonsei University, Korea	3.91	5.00

b. *Additional datasets*

The Brown et al. (2003) dataset is complemented by two additional datasets. First, National Oceanic and Atmospheric Administration (NOAA) weekly satellite-derived snow extent observations are used to compute fractional SCA. As the NOAA dataset is derived independently of the SWE dataset, it is used as partial verification. The NOAA data, described in Robinson (1993), consist of digitized weekly charts of snow cover derived from visual interpretation of visible satellite imagery by trained meteorologists. The charts are digitized on an 89 × 89 polar stereographic grid for the Northern Hemisphere (NH) with a grid spacing of 190.5 km at 60°N. The study used monthly snow-cover frequency values computed from the NOAA data by Rutgers University. These contain the corrections recommended by Robinson et al. (1991) and the Rutgers weighting scheme (Robinson 1993) to correctly partition weekly charts into appropriate months.

The second additional dataset consists of gridded surface temperature and precipitation estimates obtained from the "Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (V1.02)" (Willmott and Matsuura 2001). This global dataset includes over 7000 stations for air temperature and over 20 000 for precipitation that were interpolated on to a 0.5° by 0.5° latitude–longitude grid using climatologically aided interpolation (Willmott and Robeson 1995), an enhanced version of a traditional distance-weighting method, and spans the time period 1950–99. The precipitation data were corrected for gauge-induced systematic errors caused by wind, wetting on the interior walls of the gauge, and evaporation from the gauge (Legates and Willmott 1990). However, more recent analyses suggest that the necessary gauge corrections were underestimated over North America (Adam and Lettenmaier 2003). This product is available from the Center for Climatic Research, Department of Geography, at the University of Delaware (see online at http://climate.geog.udel.edu/~climate/).

3. Models

Table 1 identifies the 18 modeling groups whose AMIP-2 results are available to the diagnostic subprojects at the time of this writing and are included in this analysis. AMIP modeling groups run experiments for designated years with identically specified boundary conditions, including observed sea surface temperatures, so that discrepancies in model results are attributable to internal differences between atmospheric models. A variety of numerical schemes are employed, including both finite differences and spherical harmonics. The AMIP-2 time domain spans from 1979 through 1995. Therefore, 16 complete winter seasons, from 1979/80 through 1994/95, are available for evaluation. The study presented here was carried out by diagnostic subproject number 28 with a mandate to evaluate GCM simulations of snow cover. For background information on AMIP, and a list of references, see the AMIP Web site (www-pcmdi.llnl.gov/Projects/amip/index.php).

Although models differ somewhat in their treatment of snow, for the most part AMIP-2 models use a relatively simple approach where snow properties (albedo, thermal conductivity) are incorporated into the top surface layer. Models accumulate surface snow during precipitation events when the temperature of the lowest atmospheric level is at or below freezing, and the snowfall rate exceeds the melt rate. Snowmelt typically occurs as a result of the energy balance of the snowpack, including terms for sensible and latent heat fluxes, with some models including a term for the latent heat flux of nonfrozen precipitation. Most models include a term for sublimation, which is then added to the evaporative flux from the surface to the atmosphere. Many models parameterize fractional snow coverage in a grid box using a critical threshold of SWE. Fractional coverage sometimes depends on surface vegetation characteristics as well. Snow cover usually affects the surface albedo and surface thermal properties (i.e., heat conduction and heat capacity). The impact of the albedo parameterization employed can have significant impact on the surface energy balance (Nolin and Frei 2001). In some models the parameterizations for these effects depend on surface roughness, and surface roughness is sometimes parameterized as a function of snow cover. Many of the models employ the Simple Biosphere Model (SiB; Sellers et al. 1986), Biosphere–Atmosphere Transfer Scheme (BATS; Dickinson et al. 1993), or land surface model (LSM; Bonan 1996) land surface biosphere routines, which include modules for handling snow. A comprehensive discussion of the processes affecting the snowpack in each individual model is not possible here for practical reasons. Model output is provided as monthly mean values for all fields. Observed and modeled fields are computed using the same methodologies. The interested reader is referred to the AMIP model documentation Web site (http://www-pcmdi.llnl.gov/projects/modeldoc/amip2/) for more details.

4. Methodology

The spatial domain of this analysis includes North American land areas between 20° and 70°N latitude and 200° and 300°E longitude. The temporal domain covers 16 snow seasons in the AMIP-2 time period from 1979/80 to 1994/95. In a number of figures, box-and-whisker plots are used to display monthly observed and model results. Observed values are shown with asterisks, and results from AMIP-2 GCMs are summarized using box-and-whiskers. Model bias is defined for our purposes as the difference between the median model value and the observed value. Between-model variability is characterized by the spread indicated in the box-and-whisker plots. In addition, bar plots of seasonal mean results are shown for each model to facilitate evaluation of individual models. Three types of analyses are presented: 1) results integrated over the entire study domain; 2) analysis of the spatial patterns of SWE over the study domain; and 3) analysis for selected regions.

a. *Spatially integrated SWE*

Simulated SWE values integrated over the entire spatial domain are compared to values from the gridded dataset in section 5. Monthly mean SWE values in each model are area weighted and integrated over all grid cells designated as "land" whose centers lie in the study spatial domain. SWE is expressed volumetrically as km^3 of liquid water. SCA is calculated using the method of Frei et al. (2003). For each model, SCA is expressed as a fractional coverage by dividing the absolute spatial extent of snow by the total land area in the study domain. These methods for estimating SWE and SCA are designed to remove the effects of the different grid resolutions, and the resulting different total land areas, in the model output and observed datasets. For example, the North American surface areas within the study domain calculated from the models included in this study range from 1.8×10^7 to 2.3×10^7 km^2; the surface area calculated from the SWE dataset is 1.9×10^7 km^2, and from the NOAA dataset is 2.0×10^7 km^2.

b. *Spatial variation in SWE over NA*

Spatial variations in observed and simulated seasonal SWE are investigated in section 6 through comparison of contour maps. The presentation and discussion of results from each model is beyond the scope of this article, so in most cases we investigate the median model response. To facilitate the comparison, results from each model and the gridded SWE dataset are regridded to a common 2.5° × 2.5° latitude–longitude grid. Three different spatial interpolation techniques were performed and assessed: nearest neighbor, linear interpolation, and cubic convolution (which is commonly used in the field of image processing). Results from the three techniques differ only in some details, but not in the basic spatial patterns, and not at all in the conclusions. Linear interpolation was chosen for the figures presented here because the contour lines appear somewhat smoother than the other two methods. Contour lines are plotted using a logarithmic scale in order to capture the large spatial variability in snow mass found over North America and using the Albers equal area projection to allow realistic visual comparisons over a large latitude range. Note that the reinterpolated results are used solely for the purposes of visual presentation; the evaluations of SWE values integrated over the entire continent are performed using the supplied resolution of each grid as described in section 4a.

c. Regional-scale analysis of SWE

Detailed evaluations are performed over four regions that are located in areas with relatively dense snow depth observation networks, where topographic influences on snow cover are modest, and where the derived SWE dataset is considered most reliable for estimating regional mean SWE variations (section 7). These regions are limited to a minimum size constraint (≥8° wide in both latitude and longitude) to ensure that the areas are adequately sampled by all models. In addition to evaluations of SWE, the regional analyses include evaluations of SWE accumulation, SCA, temperature, and precipitation.

5. Spatially integrated results

In this section we examine the mean and variability of monthly and seasonally averaged results for SWE and SCA integrated over the continental landmass. In all cases except one (which is noted in the text), differences between parametric and nonparametric results are insignificant, so that only parametric statistics are reported.

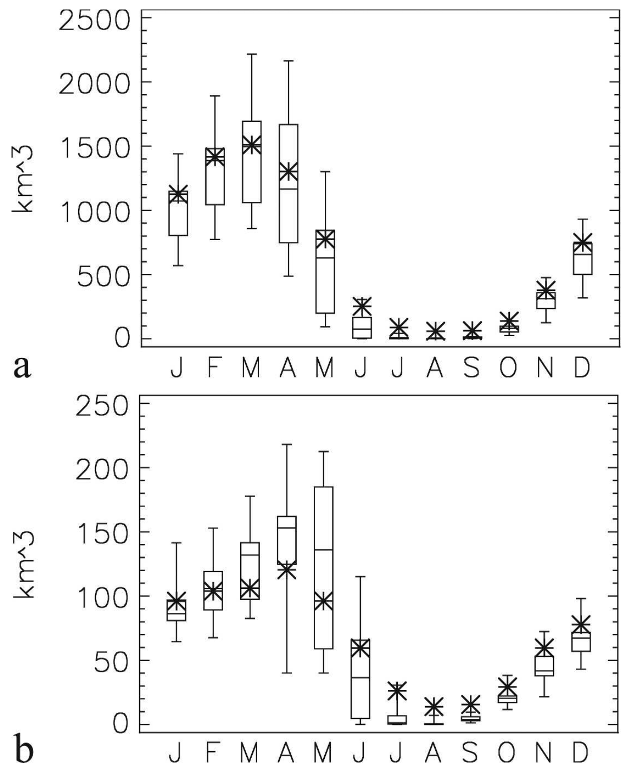


FIG. 1. North American snow water equivalent (km^3). Values from the gridded SWE dataset of Brown et al. (2003) are indicated with asterisks. Box-and-whisker plots indicate model results from 18 AMIP-2 AGCMs and are interpreted as follows: middle line shows the median value; top and bottom of box show the upper and lower quartiles (i.e., 75th and 25th percentile values); and whiskers show the minimum and maximum model values. (a) Monthly mean values; (b) monthly standard deviations.

a. Monthly mean results

The seasonal cycle of monthly mean SWE over the study domain is shown in Fig. 1a (asterisks). SWE varies from a mean summer minimum in July–September of under 200 km^3 to a mean winter maximum in March of ~1500 km^3. Also shown in Fig. 1a are box plots that characterize the monthly model values. While models tend to capture the timing of the seasonal cycle well, they tend to ablate snow too rapidly in the spring. Estimates of the snow mass over North America vary widely between models with peak SWE values exhibiting a range from 700 to 2200 km^3. This between-model variability is large compared to variability (mean ± two standard deviations) from the gridded SWE dataset of 1300 to 1700 km^3: more than half of the models have mean March SWE values outside of this range.

The models appear to capture the overall seasonal cycle in SWE interannual variability (Fig. 1b) reasonably well, although there is a noticeable increase in differences between models in May. No conclusions are made about internnual variability over the June–September period when the gridded SWE dataset is less reliable.

Figure 2a shows mean monthly continental SCA (expressed as fractional coverage) from the AMIP-2 models compared to SCA from the NOAA dataset (crosses) and SCA derived from the gridded SWE (asterisks). The NOAA and gridded SWE datasets, although derived independently, agree well with respect to mean monthly SCA. We include both datasets here as an independent verification of the Brown et al. (2003) dataset, and also to corroborate the results of Frei et al. (2003). Both show a peak areal coverage of ~0.7 (i.e., 70%) in January and a minimum of <0.1 in August. Significant between-model variability is found, with peak coverage varying from 0.5 to 0.8. In addition, we find a tendency for models to underestimate spring SCA. During months other than April, May, and June, however, the median model value is very close to observed values. The interannual variability of SCA is ≤0.05 during all months, although there is some dis-

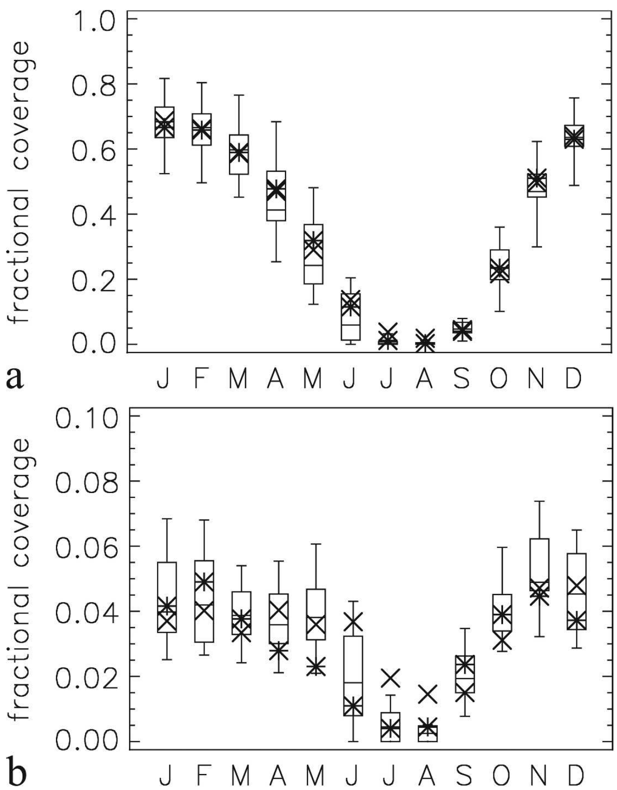


FIG. 2. North American snow-covered area. Units are fractional coverage of the entire spatial domain. Values from the gridded SWE dataset of Brown et al. (2003) are indicated with asterisks; observed values from the NOAA dataset are indicated with crosses. Box-and-whisker plots indicate model results from 18 AMIP-2 AGCMs and are interpreted as follows: middle line shows the median value; top and bottom of box show the upper and lower quartiles (i.e., 75th and 25th percentile values); and whiskers show the minimum and maximum model values. (a) Monthly mean values; (b) monthly standard deviations. The crosses are not visible in some cases because they lie exactly over the asterisks.

agreement between the two observational datasets. The disagreements are particularly apparent during spring and summer (Fig. 2b), the period of greatest uncertainties in both datasets: the gridded SWE dataset is known to be less reliable during the summer months, and a recent study of the NOAA dataset in northern Canada (Wang et al. 2004) shows it was unable to capture the snowmelt transition because of the combined influences of cloud cover and less frequent satellite coverage at high latitudes. Some differences in interannual variability are found between models, although the differences are small compared to the magnitude of mean SCA.

There is little indication of significant temporal correlations between the gridded SWE dataset and modeled SWE (Fig. 3) over the 16 seasons. This indicates that, in AGCMs, the direct influence of SSTs on snow variations is negligible. Note that using the Spearman ranked correlation (not shown), which is generally more robust to non-Gaussian distributions, gives similar results. These results are not surprising, given that other research (Zwiers 1995; Compo and Sardeshmukh 2004; Kang et al. 2004; Liu and Wu 2004) has indicated poor potential predictability of midlatitude climate in AGCMs driven by SST boundary forcing (see discussion section).

b. Seasonal mean results

Differences between models are highlighted by comparing seasonal (October–June) mean SWE in Fig. 4a. Few models fall within $\pm 10\%$ of the mean value from the gridded SWE dataset of ~ 850 km^3, and there is a tendency for models to underestimate SWE (9 out of 17 models). Only three models overestimate SWE. Results from the models differ significantly, varying between 400 and 1200 km^3.

With regards to the interannual variability of SWE, models tend to overestimate the standard deviation (Fig. 4b), yet underestimate the interquartile range (IQR) (Fig. 4c). This indicates that the shape of the distribution in models is different than in the gridded SWE dataset, with the middle of the distribution too tight (high kurtosis), and the tails of the distribution too wide (low kurtosis). These differences are less significant than differences in mean seasonal SWE (Fig. 4a). Note that this is the only statistic for which we show

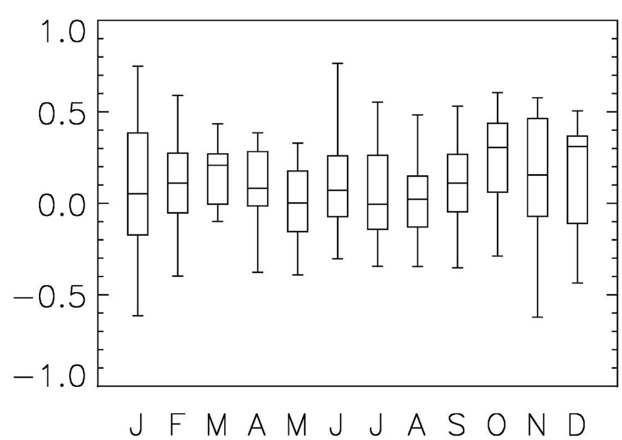


FIG. 3. Pearson correlation coefficients between modeled North American SWE and values from the gridded SWE dataset of Brown et al. (2003). Box-and-whisker plots indicate model results from 18 AMIP-2 AGCMs and are interpreted as follows: middle line shows the median value; top and bottom of box show the upper and lower quartiles (i.e., 75th and 25th percentile values); and whiskers show the minimum and maximum model values.

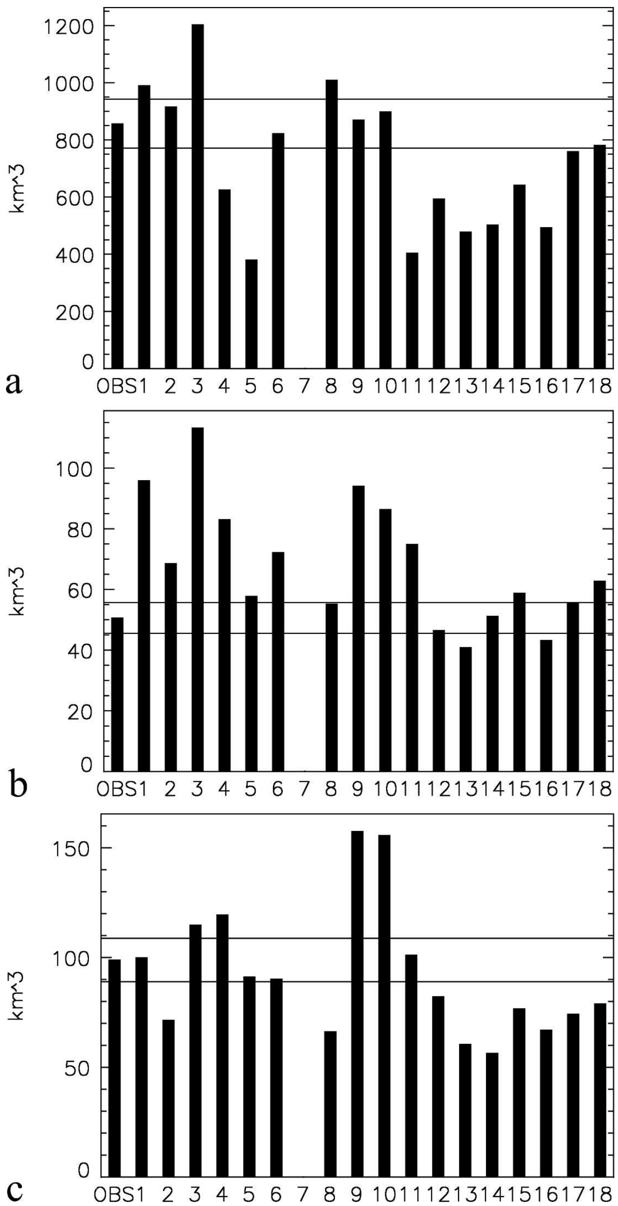


FIG. 4. Seasonal (Oct–Jun) mean North American snow water equivalent (km^3). Observed values are indicated by the leftmost bar. Each numbered bar corresponds to one model, and model numbers correspond to those shown in Table 1. Horizontal lines correspond to observed values ±10%. (a) Seasonal mean values, (b) seasonal standard deviations, and (c) seasonal interquartile ranges. Results from model 7 are omitted because values were much higher than observations or other models; see text for further explanation.

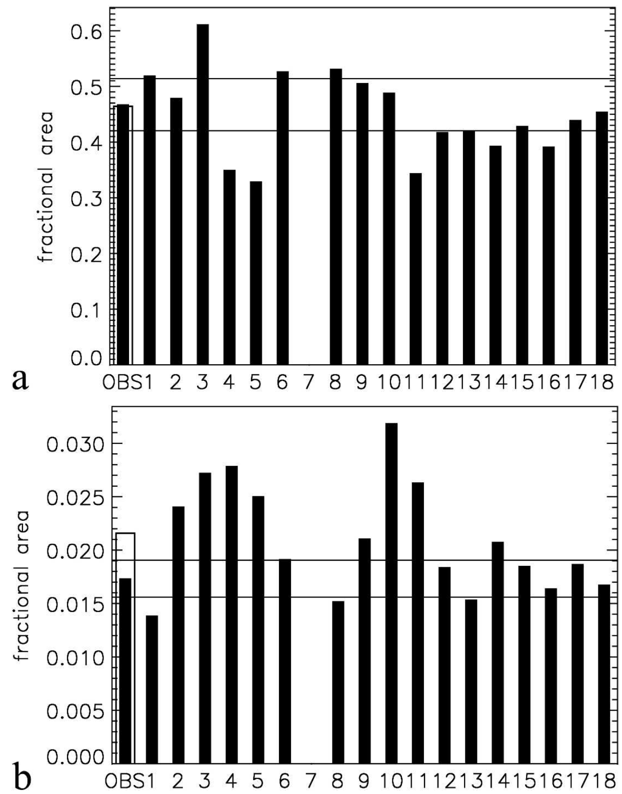


FIG. 5. Seasonal (Oct–Jun) mean North American snow-covered area. Units are fractional coverage of the entire spatial domain. Observed values, shown indicated by leftmost bar, are from the gridded SWE dataset (solid bar) and from the NOAA dataset (open bar). Each numbered bar corresponds to one model, and model numbers correspond to those shown in Table 1. Horizontal lines correspond to gridded SWE values ±10%. (a) Seasonal mean values; (b) seasonal standard deviations.

both parametric and nonparametric results, because in all other cases the differences were insignificant.

With regards to mean seasonal SCA (Fig. 5a), there is a significant spread in model results, varying from between ~0.3 and 0.6 compared to the observed value of ~0.46. The interannual variability of seasonal SCA is approximately an order of magnitude lower than mean values, but there is significant spread between models and some difference between the NOAA and gridded SWE datasets (Fig. 5b).

6. Spatial variation in SWE over NA

The spatial distribution of mean seasonal (October–June) SWE from the gridded SWE dataset is shown in Fig. 6a. Details of the orographic effects on SWE in western North America remain apparent even though SWE values were reinterpolated to a relatively coarse resolution (2.5° × 2.5° latitude–longitude). To facilitate the evaluation of model results a median model field was calculated at each grid point using the mean seasonal SWE from each model over all simulation years. The resulting median model field of mean seasonal SWE (Fig. 6b) captures the basic shape of the observed field: for example, larger values in the mountainous

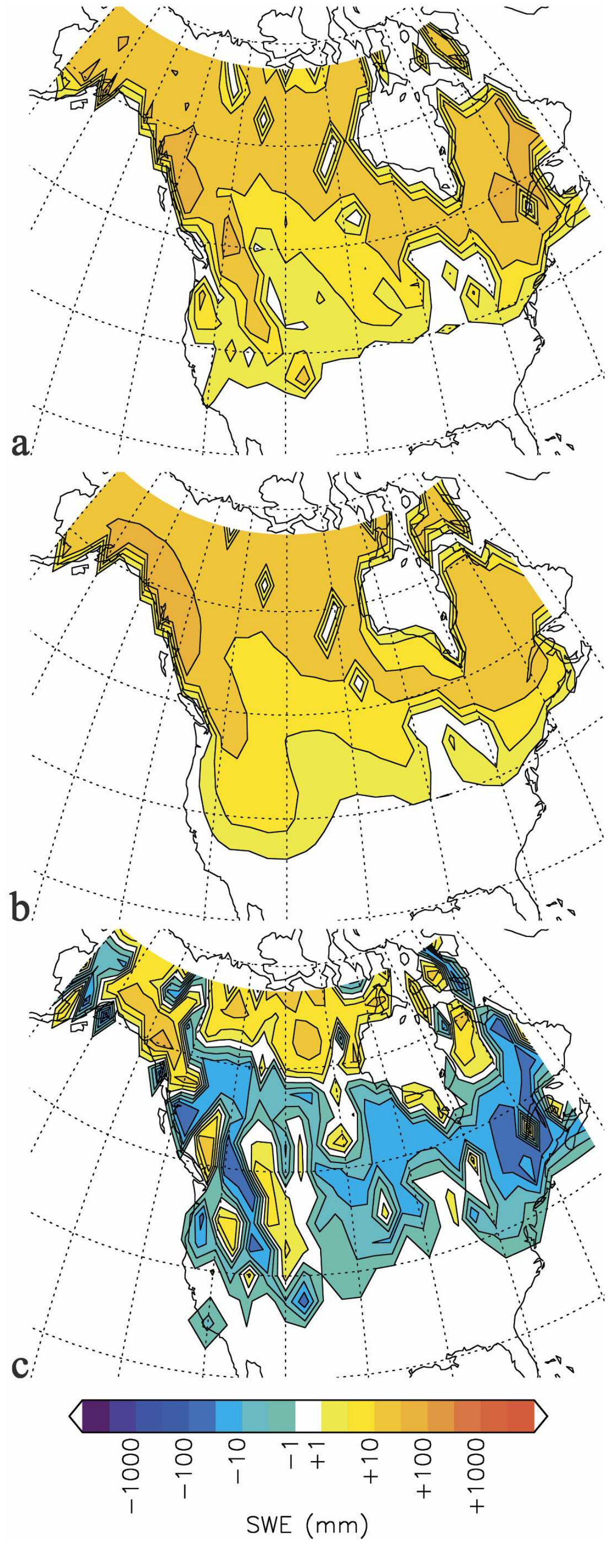


FIG. 6. Seasonal (Oct–Jun) mean SWE (mm) regridded to 2.5° × 2.5° lat–lon resolution. (a) Observed, (b) median model, and (c) model anomaly. Regridding was done using linear interpolation, and results are plotted on an Albers equal area projection using a logarithmic scale for contour lines.

western sections, eastern Canada, and the boreal forest; smaller values in the lee of the Rocky Mountains; and a reasonable estimation of the southern boundary of the snowpack.

The median model anomaly field (median model minus gridded SWE dataset) (Fig. 6c) reveals that models tend to underestimate SWE over much of the midlatitudes of NA, particularly over eastern Canada. Over western North America, the anomaly pattern is most likely related to the smoothed representation of topography in the models, which results in underestimation of SWE at higher elevations, and overestimation in adjacent lower-elevation areas. For example, Liang et al. (2004) found reduced biases in model precipitation patterns over the western United States in a regional model compared to an AGCM due to improved representation of topography. Over high latitudes, where our confidence in the gridded SWE estimates is lower, the models exhibit a tendency to overestimate SWE. Results of the same analysis using *maximum monthly* SWE (not shown) rather than *mean seasonal* SWE are almost identical. The underestimation occurs largely in spring (cf. Fig. 1a).

Interannual variability in model simulations is depicted spatially in a similar fashion. For each model, at each grid cell, the IQR of seasonal mean SWE for all simulation years is calculated; then, the median IQR value of all models is calculated. Observations indicate, again not surprisingly, that interannual variability tends to be highest over the snowiest regions (Fig. 7a). The models capture the general spatial pattern of variability reasonably well, but miss many of the smaller-scale features (Fig. 7b). As measured by IQR, models tend to underestimate interannual variability of the seasonal mean snowpack (Fig. 7c) almost everywhere, particularly the deeper snowpacks of the high elevation west and eastern Canada. Note that using standard deviation rather than IQR provides different results (not shown). Although we choose to use IQR for this figure because nonparametric statistics are generally more robust than parametric statistics, the results regarding interannual variability are considered less robust as they depend on the metric.

Between-model variability is characterized using the IQR of model values of seasonal mean SWE at each grid point. This is calculated by taking, at each grid point, the mean seasonal snowpack for each model, and calculating the IQR value over all models. Figure 8a shows that the largest between-model variability is found over the regions with the deepest snowpacks, where interannual variability is also highest: the mountainous west and eastern Canada. The minimum and maximum seasonal mean SWE fields are also shown

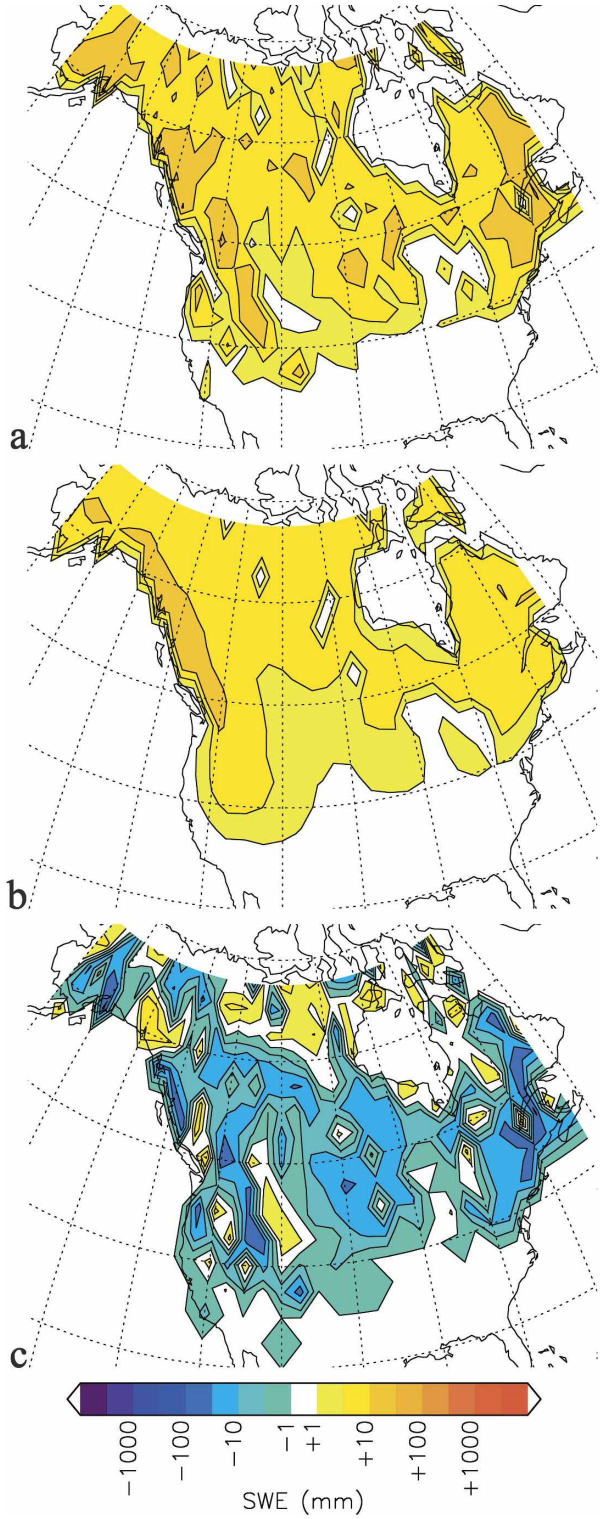


FIG. 7. Interannual variability of seasonal (Oct–Jun) mean SWE (mm). The IQR of seasonal mean SWE regridded to 2.5° × 2.5° lat–lon resolution. (a) Observed, (b) median model, and (c) model anomaly. Regridding was done using linear interpolation, and results are plotted on an Albers equal area projection using a logarithmic scale for contour lines.

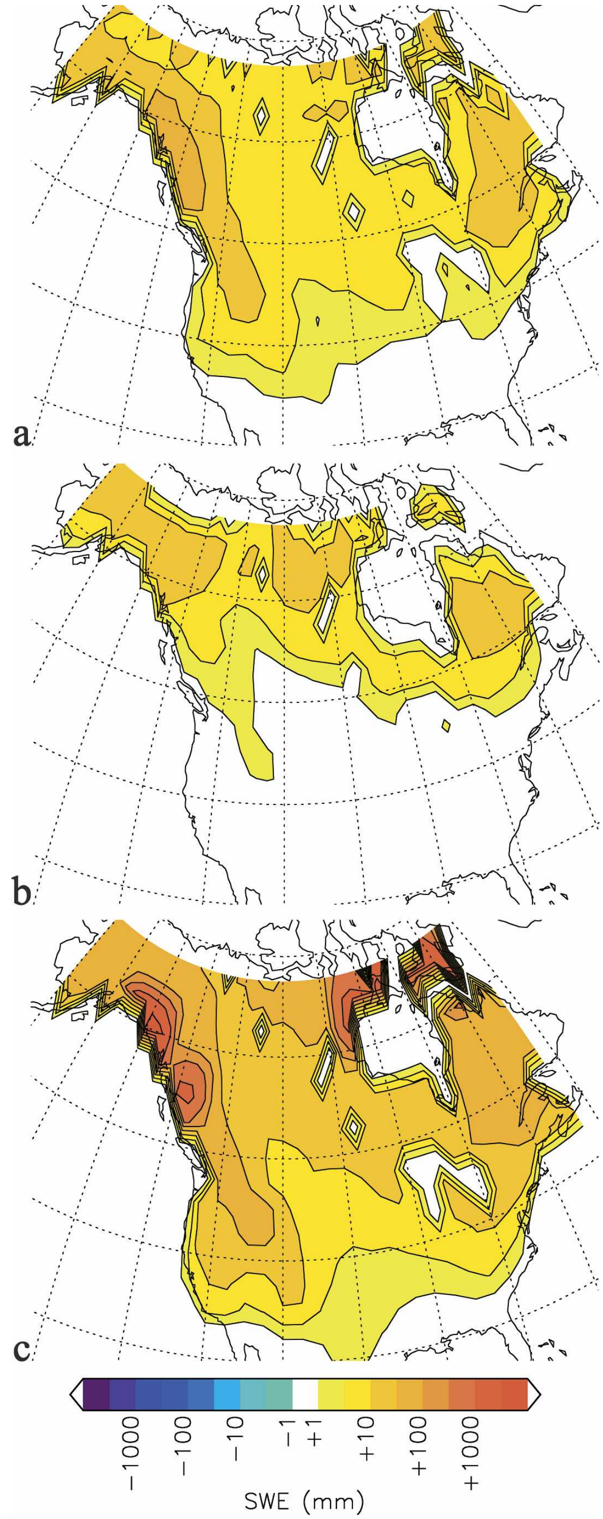


FIG. 8. Between-model variability of seasonal (Oct–Jun) mean SWE (mm) regridded to 2.5° × 2.5° lat–lon resolution. (a) IQR of seasonal mean SWE between models; (b) minimum seasonal mean SWE; and (c) maximum seasonal mean SWE. Regridding was done using linear interpolation, and results are plotted on an Albers equal area projection using a logarithmic scale for contour lines.

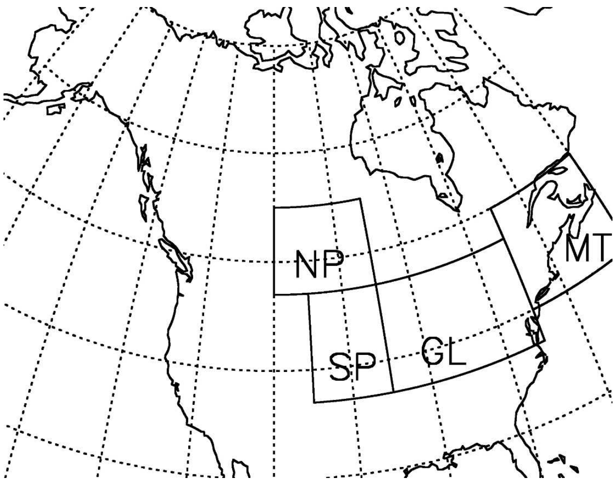


FIG. 9. Boundaries of regions examined in regional analysis in section 7. Regions are referred to in the text as follows: southern plains (SP); northern plains (NP); Great Lakes (GL); and Maritime (MT). Boundaries are defined in Table 2.

(Figs. 8b and 8c). All models maintain a mean seasonal snowpack of at least 1 mm of SWE in the Rocky Mountains as far south as ~42°N latitude, and the minimum snowpack field has the snow line well to the north of the observed snow line. This is a significant underestimation of the actual snowpack (cf. Fig. 8b to Fig. 6a). The maximum model field (Fig. 8c), on the other hand, maintains a seasonal mean snowpack of at least 1 mm of SWE south of 35°N latitude across the continent, and south of 30°N in northeastern Mexico. This represents a significant overestimation of the seasonal mean SWE field, as the observed mean seasonal snowpack is greater than 1 mm of SWE only as far south as ~40°N except in the Rocky Mountains where the 1-mm SWE line drops to ~35°N (Fig. 6a). The maximum field indicates that at least one model has unusually large anomalies over the Pacific coastal mountains between ~50° and 65°N latitude (this corresponds to model number 7, whose results in some of the figures are omitted due to this unusually large anomaly).

7. Regional-scale analysis

In this section we examine the seasonal climatic cycles over four regions shown in Fig. 9: the southern plains, northern plains, Great Lakes, and Maritimes. The rationale for choosing these regions is discussed in section 4c. The spatial domains of the regions, and a comparison of their mean seasonal (October through April; there is typically little snow cover in these regions between May and September) climates, are provided in Table 2. Note that the discussion about precipitation is based on the Willmott and Matsuura (2001) dataset, which is likely to include underestimates of precipitation gauge corrections (Adam and Lettenmaier 2003).

We evaluate the annual cycles of SWE, SWE accumulation, SCA, precipitation, and temperature in each region (Figs. 10a–d). *SWE accumulation* is defined as the monthly mean SWE value minus the previous month's value. It is evident from the figure that over all regions the timing of the seasonal cycles is captured reasonably well, but that the spread of model results can be large. While the median model result is close to observations, a large variation between models is apparent. As with continental-scale results reported in section 5, no temporal correlations between observed and modeled SWE in these regions are found. Therefore, this report focuses on climatology, and no time series are shown.

The plains are much drier than the other two regions. The southern plains is relatively warm and dry, with a mean seasonal snowpack of only 7-mm SWE, a peak snowpack in January and February of <20 mm, and a mean seasonal snow covered area of less than 30%, making it the least snowy of the four regions. The northern plains are even drier, receiving only 140 mm of precipitation during the winter season compared to ~200 mm in the southern plains. However, because of lower temperatures, a larger snowpack is typically maintained in the northern plains.

Over the southern plains, SCA peaks in January at

TABLE 2. Summary of mean seasonal (Oct–Apr) climate conditions in the four regions. Columns show the region name; region boundary; seasonal mean SWE; month of maximum mean SWE; mean seasonal snow-covered area; mean seasonal total precipitation; and mean seasonal temperature. Values calculated using the AMIP-2 time domain from 1979 through 1995. Snow values calculated from Brown et al. (2003); temperature and precipitation values calculated from Willmott and Matsuura (2001).

Region	Region boundaries lat (°N)/lon (°E)	SWE (mm)	Month of max SWE	SCA (%)	Precipitation (mm)	Temperature (°C)
1) Southern Plains	37–47/255–265	7	Jan	29	198	1.8
2) Northern Plains	47–55/250–265	23	Feb	64	140	−6.2
3) Great Lakes	37–47/265–285	16	Feb	31	485	1.9
4) Maritimes	40–50/285–300	43	Mar	56	594	−3.2

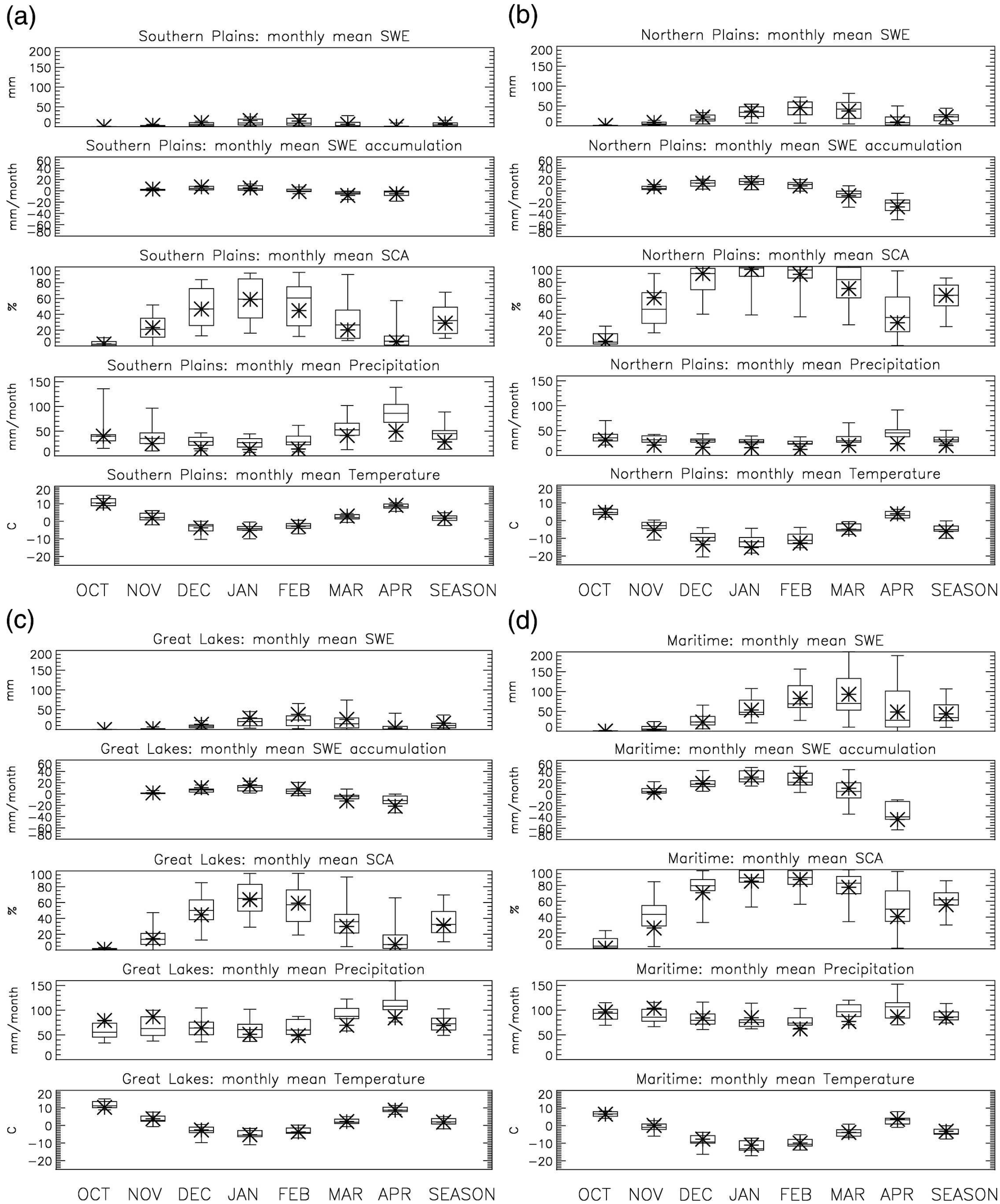


FIG. 10. (a) Regional monthly and seasonal climatology for southern plains region. Observed values are indicated with asterisks. Box-and-whisker plots indicate model results from 18 AMIP-2 AGCMs, interpreted as follows: middle line shows the median value; top and bottom of box show the upper and lower quartiles (i.e., 75th and 25th percentile values); and whiskers show the minimum and maximum model values. Variables shown are SWE, SWE accumulation, SCA, precipitation, and temperature. Snow variables taken from Brown et al. (2003); temperature and precipitation taken from Willmott and Matsuura (2001). (b) Same as (a) except northern plains region. (c). Same as (a) except Great Lakes region. (d) Same as (a) except Maritime region.

60%, and the snowpack has typically ablated by early April (Fig. 10a). Models exhibit a wet bias, with a great deal of between-model variability. No significant temperature bias is observed, and between-model variability in temperature is modest. Biases in SWE and SCA are minimal. While between-model variability in SWE is small, between-model variability in SCA is significant.

In the northern plains, biases in precipitation rates and early season temperatures are apparent (Fig. 10b). Models have a warm bias of 5°C between November and February, resulting in a seasonal mean warm bias of 2°C. Yet, 25% of models have seasonal mean cold biases. Only small biases, but significant between-model variability, are observed in SWE and SCA.

The Great Lakes and Maritime regions receive abundant precipitation: ~500 and ~600 mm per season, respectively (Figs. 10c and 10d). The Great Lakes region experiences relatively warm temperatures (comparable to those observed in the southern plains). As a result of being wet and warm, the mean snowpack in the Great Lakes is comparable to the northern plains. As over the plains, model simulations vary significantly. For example, while 25% of the models maintain a mean monthly snowpack as high as 5–40 mm well into April, other models have peak midwinter monthly values as low as 5 mm.

The Maritime region, in contrast, is relatively cold (only the northern plains are colder) and wet (no other region receives as much precipitation). The mean seasonal snowpack of >40 mm SWE and peak snowpack of 100 mm SWE in March are almost double the values found in any other region. As in the other regions, between-model variability is significant, resulting in peak monthly SWE values ranging from <50 to 200 mm.

Seasonal (October–April) mean values of SWE, SCA, total precipitation, and temperature from observations and from each model are shown in Figs. 11a–d. "Seasonal" is defined differently than in earlier sections of this article because over these regions the snow season is shorter than over higher latitudes. Perusal of these bar plots reveals that a simple diagnosis of problems from a suite of models is impossible. For example, in one region one model may be very wet and warm, yet only moderately underestimate SWE and actually overestimate SCA; another model might be cold and dry, yet produce similar mean SWE and SCA values in the same region!

8. Discussion and conclusions

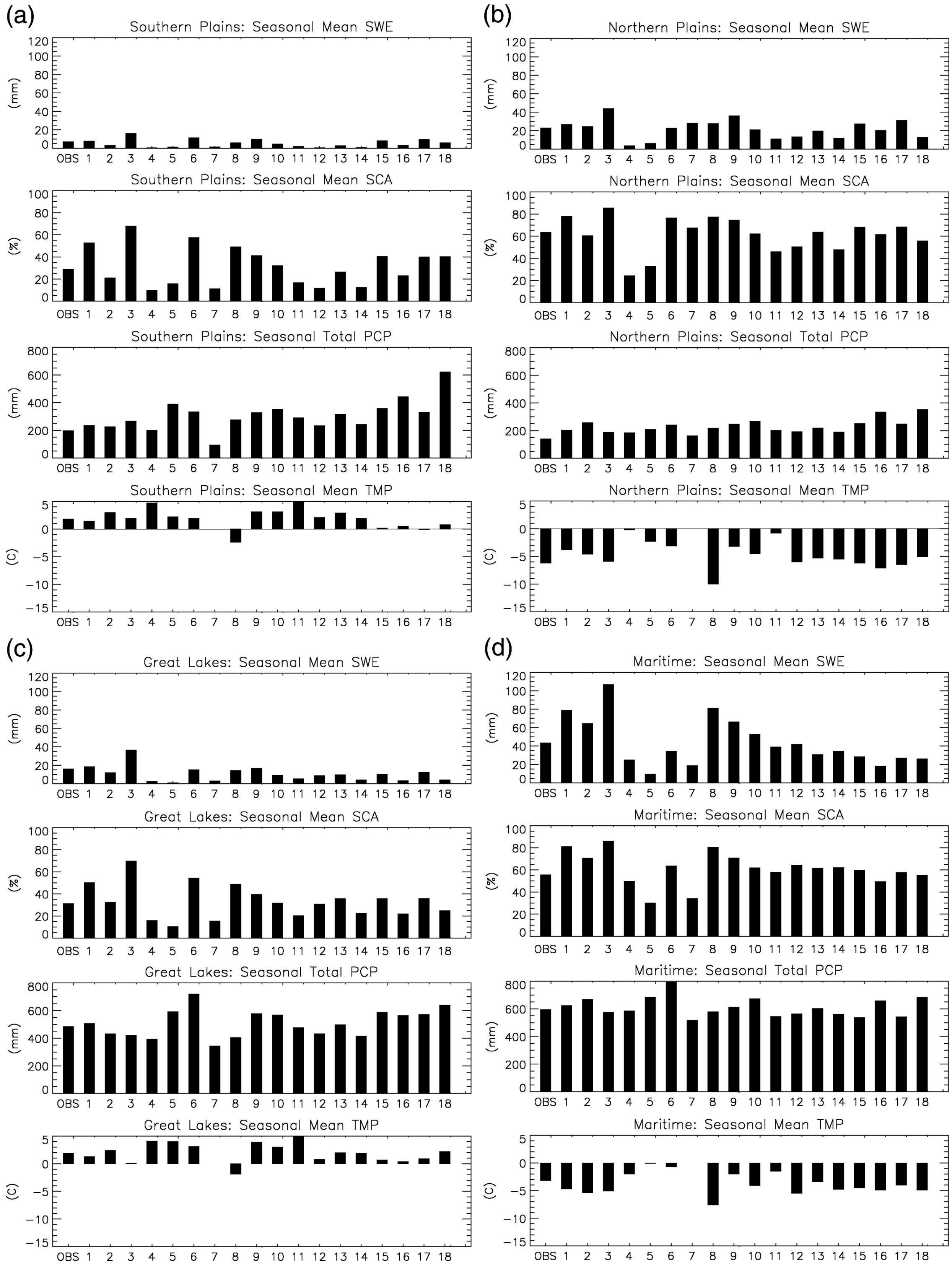
The large-scale spatial distribution of observed snow mass, or SWE, over North America reflects the interaction of a number of influences: proximity to moisture sources, orography, predominant storm-track locations, and the location of the 0°C isotherm. The largest SWE values are found over the western cordillera, with a secondary maximum over eastern Canada. SWE values are typically low over much of the continental interior, with a well-defined band of higher SWE values over the boreal forest. The total estimated snow mass integrated over the entire study domain (between 20° and 70°N latitude and 200° and 300°E longitude), which covers most of North America but excludes Greenland and sections of the Arctic archipelago, peaks in March at ~1500 km^3. The snow covered area (SCA) peaks at around 70% in January.

Most AMIP-2 models simulate the seasonal timing and the relative spatial patterns of continental-scale SWE reasonably well. Analyses of continental- to regional-scale variations show that simulated SWE values typically peak in the appropriate month. In most models, the deepest snowpacks are found over the western cordillera and over eastern Canada, just as indicated by observations. The median model result, however, tends to overestimate the rate of snow ablation during spring. This is likely related to the exclusion of subgrid-scale treatments for terrain and land cover, and to the difficulty of modeling all the subgrid-scale processes that contribute to snow ablation. However, a detailed evaluation of this problem is beyond the scope of this analysis.

Significant between-model variability is found in every aspect of the simulations. On the continental scale, the peak monthly snow mass integrated over the North American continent in AMIP-2 models varies between ±50% of the observed value of ~1500 km^3. The volume of water in the snowpack, as well as the magnitude of model errors, is significant from the perspective of the large-scale water balance. For example, the mean

→

FIG. 11. (a) Regional (Oct–Apr) seasonal climatology for southern plains region. Observed values are indicated by the leftmost bar. Each numbered bar corresponds to one model, and model numbers correspond to those shown in Table 1. Temperature values for model 7 are omitted because they were unavailable at the time of this writing. "Seasonal" is defined differently than in earlier sections of this article because over these regions the snow season is shorter than over higher latitudes. Snow variables taken from Brown et al. (2003); temperature and precipitation taken from Willmott and Matsuura (2001). (b) Same as (a) except northern plains. (c) Same as (a) except Great Lakes. (d) Same as (a) except Maritime.


(a)
Southern Plains: Seasonal Mean SWE
(mm)
Southern Plains: Seasonal Mean SCA
(%)
Southern Plains: Seasonal Total PCP
(mm)
Southern Plains: Seasonal Mean TMP
(C)
(b)
Northern Plains: Seasonal Mean SWE
(mm)
Northern Plains: Seasonal Mean SCA
(%)
Northern Plains: Seasonal Total PCP
(mm)
Northern Plains: Seasonal Mean TMP
(C)
(c)
Great Lakes: Seasonal Mean SWE
(mm)
Great Lakes: Seasonal Mean SCA
(%)
Great Lakes: Seasonal Total PCP
(mm)
Great Lakes: Seasonal Mean TMP
(C)
(d)
Maritime: Seasonal Mean SWE
(mm)
Maritime: Seasonal Mean SCA
(%)
Maritime: Seasonal Total PCP
(mm)
Maritime: Seasonal Mean TMP
(C)
OBS 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

total precipitation over North America is ~18 000 km^3 yr^{-1}; the mean total runoff from North America to the oceans is ~8000 km^3 yr^{-1}; the global volume of water in all rivers is ~2000 km^3 yr^{-1}; the total volume of water incorporated in the earth's biota is ~1000 km^3 yr^{-1}; and the average annual availability of freshwater for the United States and Canada is ~3000 km^3 yr^{-1} in each country (Gleick 2000). Clearly, adequate simulation of the continental-scale hydrologic cycle requires a reasonably accurate depiction of the magnitude and timing of snow accumulation and ablation.

Over smaller regions, similar between-model variability is observed for SWE, as well as for other variables including SCA, temperature, and precipitation. It is hoped that the bar plots showing individual model results (Fig. 11) will be helpful to some of the modeling groups. Most models exhibit biases in one area or another.

On the other hand, it is of interest that the median result from the suite of models tends to do a better job of estimating climatological mean features than any individual model. These two characteristics of modeled climate—large between-model variability, and the tendency for the median or mean of a suite of models to better approximate observed climate than any individual model—has been found for many climatological variables and suggests that experiments using ensemble results from several models might be valuable. The reasons for this have not been fully explained, but see Krishnamurti et al. (2000), Kharin and Zwiers (2002), and Palmer et al. (2004) for further discussion.

Year-to-year variations in large-scale SWE appear to be only weakly correlated to observed variations. Such weak correlations, which were found for SCA as well (Frei et al. 2003; Frei and Robinson 1998), indicate that in AMIP-2 models, sea surface temperatures, which are the primary boundary conditions imposed on models in the AMIP experiments, are not significant factors in determining the interannual variations in the large-scale snowpack. These results are consistent with other studies that show that the potential predictability of midlatitude climate is poor in AGCM experiments driven by SST boundary conditions. In AMIP-1 models, potential predictability of midlatitude climate was generally poor (Zwiers 1995). More recent work corroborates that predictability of precipitation over midlatitude Northern Hemisphere landmasses is poor during both winter (Compo and Sardeshmukh 2004) and summer (Kang et al. 2004). It is likely that the correct simulation of circulation patterns requires the inclusion of feedbacks between the atmosphere and ocean. For example, Liu and Wu (2004) show that the atmospheric response to winter SSTs depends on the inclusion of a coupled atmosphere–ocean model.

The results of this study have implications for simulations of the large-scale hydrologic cycle. Over much of North America a significant portion of annual precipitation falls in frozen form, and the snowpack plays a major role in hydrological processes. Even in many arid areas that do not receive much snow, such as parts of the western United States, a significant portion of river water originates as snowfall at higher elevations. The magnitude and timing of the seasonal snowpack therefore have great impacts on local and regional hydrology, affecting riparian ecology, wetlands, nutrient cycling, water quality, as well as water supply for human consumption. As GCMs are increasingly being used to generate scenarios of climate change at regional to global scales, and are including more detailed submodels for biospheric processes (with particular regards to carbon cycling), the choice of model (or models) may be a critical factor in determining the types and magnitudes of expected impacts, and may influence subsequent policy decisions.

Acknowledgments. This research was supported by grants from the National Science Foundation (NSF ATM-9818098) and the National Aeronautics and Space Administration (NASA NAG-7546). RB acknowledges the important contribution of the Canadian Meteorological Centre in developing the gridded SWE dataset for North America, and funding support from the CRYSYS project of the Meteorological Service of Canada.

REFERENCES

Adam, J. C., and D. P. Lettenmaier, 2003: Adjustment of global gridded precipitation for systematic bias. *J. Geophys. Res.,* **108,** 4257, doi:10.1029/2002JD002499.

Armstrong, R. L., and M. J. Brodzik, 2002: Hemispheric-scale comparison and evaluation of passive microwave snow algorithms. *Ann. Glaciol.,* **34,** 38–44.

Barry, R., 2002: The role of snow and ice in the global climate system: A review. *Polar Geogr.,* **26** (3), 235–246.

Bonan, G. B., 1996: A land surface model (LSM version 1.0) for ecological, hydrological, and atmospheric studies. Tech. Note NCAR/TN-417+STR, National Center for Atmospheric Research, 150 pp.

Brasnett, B., 1999: A global analysis of snow depth for numerical weather prediction. *J. Appl. Meteor.,* **38,** 726–740.

Brown, R. D., B. Brasnett, and D. A. Robinson, 2003: Gridded North American monthly snow depth and snow water equivalent for GCM evaluation. *Atmos.–Ocean,* **41,** 1–14.

Cohen, J., 1994: Snow cover and climate. *Weather,* **49** (5), 150–156.

Compo, G. P., and P. D. Sardeshmukh, 2004: Storm track predictability on seasonal and decadal scales. *J. Climate,* **17,** 3701–3720.

Dickinson, R. E., A. Henderson-Sellers, and P. J. Kennedy, 1993: Biosphere–Atmosphere Transfer Scheme version 1e as coupled to the NCAR Community Climate Model. Tech. Note NCAR/TN-387+STR, National Center for Atmospheric Research, 72 pp.

Foster, D. J., and R. D. Davy, 1988: Global snow depth climatology. USAF Publication USAFETAC/TN-88/006, Scott Air Force Base, IL, 48 pp.

Foster, J., and Coauthors, 1996: Snow cover and snow mass intercomparison of general circulation models and remotely sensed datasets. *J. Climate,* **9,** 409–426.

Frei, A., and D. A. Robinson, 1998: Evaluation of snow extent and its variability in the Atmospheric Model Intercomparison Project. *J. Geophys. Res.,* **103,** 8859–8871.

——, J. A. Miller, and D. A. Robinson, 2003: Improved simulations of snow extent in the second phase of the Atmospheric Model Intercomparison Project (AMIP-2). *J. Geophys. Res.,* **108,** 4369, doi:10.1029/2002JD003030.

Gates, W. L., 1992: AMIP: The atmospheric model intercomparison project. *Bull. Amer. Meteor. Soc.,* **73,** 1962–1970.

——, and Coauthors, 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Amer. Meteor. Soc.,* **80,** 29–56.

Gleick, P. H., 2000: *The Worlds Water 2000–2001: The Biennial Report on Freshwater Resources*. Island Press, 315 pp.

Kang, I.-S., J.-Y. Lee, and C.-K. Park, 2004: Potential predictability of summer mean precipitation in a dynamical seasonal prediction system with systematic error correction. *J. Climate,* **17,** 834–845.

Kharin, V. V., and F. W. Zwiers, 2002: Climate predictions with multimodel ensembles. *J. Climate,* **15,** 793–799.

Krishnamurti, T. N., C. M. Kishtawal, Z. Shang, T. Larow, D. Bachiochi, E. Williford, S. Gadgil, and S. Surendran, 2000: Mulitmodel ensemble forecasts for weather and seasonal climate. *J. Climate,* **13,** 4196–4216.

Legates, D. R., and C. J. Willmott, 1990: Mean seasonal and spatial variability in gauge-corrected global precipitation. *Int. J. Climatol.,* **10,** 111–127.

Liang, X.-Z., L. Li, K. E. Kunkel, M. Ting, and J. X. L. Wang, 2004: Regional climate model simulation of U.S. precipitation during 1982–2002. Part 1: Annual cycle. *J. Climate,* **17,** 3510–3529.

Liu, Z., and L. Wu, 2004: Atmospheric response to North Pacific SST: The role of ocean–atmosphere coupling. *J. Climate,* **17,** 1859–1882.

Nolin, A. W., and A. Frei, 2001: Remote sensing of snow and characterization of snow albedo for climate simulations. *Remote Sensing and Climate Modeling: Synergies and Limitations*, M. Beniston and M. M. Verstraete, Eds., Kluwer Academic, 159–180.

Palmer, T. N., and Coauthors, 2004: Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bull. Amer. Meteor. Soc.,* **85,** 683–702.

Robinson, D. A., 1993: Monitoring Northern Hemisphere snow cover. Proceedings of Snow Watch '92: Detection Strategies for Snow and Ice, Glaciological Data Rep. GD-25, National Snow and Ice Data Center, 1–25.

——, F. T. Keimig, and K. F. Dewey, 1991: Recent variations in Northern Hemisphere snow cover. *Proc. 15th Annual Climate Diagnostics Workshop,* Asheville, NC, National Oceanic and Atmospheric Administration, 219–224.

Sellers, P. J., Y. Mintz, Y. C. Sud, and A. Dalcher, 1986: A simple biosphere model (SiB) for use within general circulation models. *J. Atmos. Sci.,* **43,** 505–531.

Wang, L., M. Sharp, R. Brown, C. Derksen, and B. Rivard, 2004: Evaluation of spring snow covered area depletion in the Canadian Arctic from NOAA snow charts. *Remote Sens. Environ.,* **95,** 453–463.

Willmott, C. J., and S. M. Robeson, 1995: Climatologically aided interpolation (CAI) of terrestrial air temperatures. *Int. J. Climatol.,* **15,** 221–229.

——, and K. Matsuura, cited 2001: Terrestrial air temperature and precipitation: Monthly and annual time series (1950–1999). Center for Climatic Research, Department of Geography, University of Delaware. [Available online at http://climate.geog.udel.edu/~climate/.]

Yang, Z.-L., and Coauthors, 1999: Simulation of snow mass and extent in general circulation models. *Hydrol. Processes,* **13,** 2097–2113.

Zwiers, F. W., 1995. Inter-comparison of inter-annual variability and potential predictability. *Proc. AMIP Scientific Conf.,* Monterey, CA, World Climate Research Programme, World Meteorological Organization, 245–251.