# Coupling of the Common Land Model to the NCAR Community Climate Model

XUBIN ZENG

*Institute of Atmospheric Physics, The University of Arizona, Tucson, Arizona*

MUHAMMAD SHAIKH, YONGJIU DAI, AND ROBERT E. DICKINSON

*School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia*

RANGA MYNENI

*Department of Geography, Boston University, Boston, Massachusetts*

(Manuscript received 26 February 2001, in final form 10 October 2001)

## ABSTRACT

The Common Land Model (CLM), which results from a 3-yr joint effort among seven land modeling groups, has been coupled with the National Center for Atmospheric Research (NCAR) Community Climate Model (CCM3). Two 15-yr simulations of CCM3 coupled with CLM and the NCAR Land Surface Model (LSM), respectively, are used to document the relative impact of CLM versus LSM on land surface climate. It is found that CLM significantly reduces the summer cold bias of surface air temperature in LSM, which is associated with higher sensible heat fluxes and lower latent heat fluxes in CLM, and the winter warm bias over seasonally snow-covered regions, especially in Eurasia. CLM also significantly improves the simulation of the annual cycle of runoff in LSM. In addition, CLM simulates the snow mass better than LSM during the snow accumulation stage. These improvements are primarily caused by the improved parameterizations in runoff, snow, and other processes (e.g., turbulence) in CLM. The new land boundary data (e.g., leaf-area index, fractional vegetation cover, albedo) also contribute to the improvement in surface air temperature simulation over some regions. Overall, CLM has little impact on precipitation and surface net radiative fluxes.

## 1. Introduction

The land component of climate models represents many important processes that control the transfers of water and energy to the atmosphere. Its importance for weather forecasting and climate has been increasingly recognized in the past two decades (Betts et al. 1996; Koster et al. 2000, and references therein). For this reason, the original bucket-type land surface model (e.g., Manabe 1969) has been replaced by more physically based representations of the global soil–vegetation–atmosphere transfer system (e.g., Sellers et al. 1997). This paper reports results of our project to couple a recently developed state-of-the-art land surface model; that is, the Common Land Model version 1 (CLM; Dai et al. 2002, manuscript submitted to *Bull. Amer. Meteor. Soc.*), with the National Center for Atmospheric Research (NCAR) Community Climate Model (CCM3).

The CLM project represents a multi-institution multidisciplinary effort that brings together a broader range of expertise than can be accomplished within any one research group. The history, progress, and future plan of the CLM project are discussed in Dai et al. (2002, manuscript submitted to *Bull. Amer. Meteor. Soc.*). The original motivation for CLM came from the desire to have a truly community-developed land surface model as the land component of the NCAR Community Climate System Model (CCSM). In other words, CLM is intended as an improved biophysics package to replace the NCAR Land Surface Model (LSM; Bonan 1996). Since the initial CLM code was completed in October 1998, the FORTRAN program has gone through four iterations for improvement of the coding standard. After finishing the initial CLM test in February 1999, CLM has also gone through four phases of vigorous beta tests. Very comprehensive observational data have been used: a variety of multiyear point observational data over different regions of the world, regional data over the U.S. Red–Arkansas River basin, and global data from the Global Soil Wetness Project. These data include all data in the project for the Intercomparison of Land Parameterization Schemes. CLM has also been tested in the multiagency Land Data Assimilation System over the continental United States. Results from these extensive

---

*Corresponding author address:* Dr. Xubin Zeng, Dept. of Atmospheric Sciences, The University of Arizona, P.O. Box 210081, Tucson, AZ 85721.
E-mail: xubin@atmo.arizona.edu.

tests will be published elsewhere by CLM participants. Here we focus on the coupling of CLM with the NCAR CCM3.

Section 2 briefly describes CLM, LSM, and CCM3, while section 3 discusses our land boundary data, which are derived based on recent satellite remote sensing data, global field survey data, and literature survey. Section 4 evaluates the results from the CCM3–CLM and compares them with those from CCM3–LSM, while section 5 gives some conclusions.

## 2. Model description

The Common Land Model has one vegetation layer with a realistic photosynthesis-conductance model based on Bonan (1996). It has 10 unevenly spaced vertical soil layers with the bottom layer at 3.43-m depth. A thin top layer of 1.75 cm is specified to realistically simulate surface soil fluxes and subsequently the diurnal cycle of surface soil temperature. CLM also contains up to five snow layers (depending on the total snow depth) with water flow within snow considered. Model state variables include snow and soil temperature, ice lens mass and liquid mass in each layer, leaf temperature, canopy water storage, nondimensional snow age, snow-layer number, and snow-layer thickness.

Radiative transfer through a canopy is based on a simplified two-stream approach under the constraint that canopy albedo approaches the prescribed albedo for bare soil (or thick canopy) as leaf-area index (LAI) approaches zero [or infinity (i.e., 2 or higher values, depending on the solar zenith angle)]. Soil and snow albedos are both based on Dickinson et al. (1993). Soil albedo depends on soil color type for each grid box and soil moisture in the top soil layer. Snow fraction depends on snow depth and surface (bare soil or canopy) roughness length, while snow albedo depends on solar zenith angle and snow age, which considers the effects of grain growth, meltwater, and dirt and soot.

The turbulence scheme in CLM considers turbulence above, within, and below canopy. It also includes the consistent treatment of the laminar layer over canopy and bare soil (Zeng and Dickinson 1998), and includes an improved treatment of turbulence under free convective conditions (Zeng et al. 1998a).

Leaf temperature is computed based on the energy balance between net radiative flux and latent and sensible heat fluxes (Dickinson et al. 1993). Soil and snow temperatures are predicted using a heat diffusion equation in 10 soil layers and up to five snow layers. The volumetric heat capacity of soil (or snow) considers the contributions of solids, water, and ice in soil (or water and ice in snow). The soil (or snow) thermal conductivity depends on soil water and ice (or snow density). Lake temperature is computed based on the six-layer lake model of Bonan (1996).

Surface evapotranspiration consists of evaporation of precipitation intercepted by stems and leaves, canopy transpiration as controlled by photosynthesis, and bare soil evaporation. For sunlit and shaded leaves, the photosynthesis-conductance submodel separately computes the leaf assimilation rate as the minimum among three rates that are limited by the efficiency of the photosynthetic enzyme system, the amount of photosynthetically active radiation captured by the leaf chlorophyll, and the capacity of leaf to export or utilize the products of photosynthesis, respectively (Bonan 1996). The canopy conductance is the combination of those over sublit and shaded leaves.

Model runoff includes surface runoff and base flow, both of which are computed over saturated and unsaturated areas separately. The fraction of saturated area depends on topographic features and the nondimensional mean water table depth.

Water and ice in soil and snow and those intercepted by canopy are predicted based on mass conservation equations. The hydralic conductivity and soil matrix potential are parameterized following Clapp and Hornberger (1978) and Cosby et al. (1984). Water flow in both unfrozen and frozen soils is considered. Snow water that is over the water-holding capacity of snow can percolate into the underlying snow or soil layer. Snow compaction considers the contributions from snow metamorphism, overburden, and melt (Jordan 1991).

Because the NCAR Land Surface Model has been fully documented in Bonan (1996), much less detail is given here. LSM contains one vegetation layer, six soil layers, and one snow layer. Its photosynthesis-conductance model and its coupling strategy with CCM3 are largely followed by CLM. Its surface hydrology submodel considers the fractional coverage of convective precipitation as well as the spatial distribution of precipitation throughfall and soil water in precipitation regions. While CLM has a more detailed treatment of snow through its multilayer snow parameterization, LSM has a more detailed treatment of surface hydrology (particularly runoff). The complexity level of other land surface processes is similar between LSM and CLM.

Because CLM has been developed primarily based on the Biosphere–Atmosphere Transfer Scheme (Dickinson et al. 1993), LSM, and the snow model of Dai and Zeng (1996), results from Biosphere–Atmosphere Transfer Scheme (BATS) coupled with CCM3 will also be briefly compared in this paper. BATS contains one vegetation layer, three soil layers, and one snow layer. Overall, BATS is the simplest land model among the three considered here.

The NCAR Community Climate Model version 3 is a spectral atmospheric model with T42 truncation (or about 2.8° horizontal resolution) and 18 levels in the vertical. It employs comprehensive parameterizations of deep convection, shallow and nonprecipitating convection, shortwave and longwave radiation, and atmospheric boundary layer turbulence. Observed monthly sea surface temperature and sea ice are prescribed. Additional model details are provided in Kiehl et al. (1998).

The output of CLM (LSM or BATS) to the atmospheric model (CCM3) at every time step includes surface albedos (direct beam and diffuse for visible and near-infrared wave bands), upward longwave radiation, sensible and latent heat fluxes, and surface wind stress. The input from CCM3 to CLM (LSM or BATS) at every time step includes incident solar radiation (direct beam and diffuse for visible and near-infrared wave bands), incident longwave radiation, convective and large-scale precipitation, lowest model-level temperature, horizontal wind components, specific humidity, pressure, and height above surface. All three land models are coupled to CCM3 explicitly. In other words, equations in the land and atmospheric models are solved separately. We maintain LSM's explicit scheme, since we or others have not noticed any serious problems. Further, the explicit scheme maintains the modular structure of CLM, which is important for its potentially wide applications in regional and global models.

TABLE 1. The IGBP land cover classification.

| | |
|---|---|
| | Natural vegetation |
| 1 | Evergreen needleleaf forests |
| 2 | Evergreen broadleaf forests |
| 3 | Deciduous needleleaf forests |
| 4 | Deciduous broadleaf forests |
| 5 | Mixed forests |
| 6 | Closed shrublands |
| 7 | Open shrublands |
| 8 | Woody savannas |
| 9 | Savannas |
| 10 | Grasslands |
| 11 | Permanent wetlands |
| | Developed and mosaic lands |
| 12 | Croplands |
| 13 | Urban and builtup lands |
| 14 | Cropland/natural vegetation mosaic |
| | Nonvegetation lands |
| 15 | Snow and ice |
| 16 | Barren |
| 17 | Water bodies |

## 3. Comprehensive land data

### *a. Fractional vegetation cover and leaf-area index*

Both fractional vegetation cover ($\sigma_v$) and green leaf-area index are needed in land models to determine surface energy, water, and trace gas exchanges. Although many climate models have included a seasonal variation of LAI, the limited information available from the satellite normalized difference vegetation index (NDVI) product precludes deriving seasonal variations of both $\sigma_v$ and LAI. Hence, Gutman and Ignatov (1998) prescribed a fixed value for LAI. Alternatively, Sellers et al. (1996) and other authors assumed that spatial and seasonal variations of NDVI were all given by variations of LAI. Any of the four possible assumptions (i.e., whether fractional vegetation and/or LAI vary seasonally or are kept constant) could be used in a climate model, but Zeng et al. (2000) argued that fixing fractional vegetation and allowing LAI to vary is a realistic assumption from a modeling viewpoint and can be supported by current observational data. In this context, fractional vegetation is viewed as determined by types of vegetation and long-term edaphic and climatic controls, whereas LAI, as provided by model simulations (e.g., Dickinson et al. 1998) or remote sensing (e.g., Myneni et al. 1997), includes all the seasonality of canopies, ranging from near-zero values in the more extreme cases of annual or harvested vegetation to full canopy values of 5–10. Therefore, the global 1-km $\sigma_v$ data from Zeng et al. (2000) and the global 8-km LAI data from Myneni et al. (1997) are used here for CLM.

The global 1-km $\sigma_v$ data were derived from the Advanced Very High Resolution Radiometer (AVHRR) NDVI data for 1 yr (April 1992–March 1993). Recently, we have also used the method of Zeng et al. (2000) to derive the global 8-km $\sigma_v$ for each year from 1982 to 2000 based on the AVHRR Land Pathfinder dataset (James and Kalluri 1994). Results averaged over the 19-yr period are found to be consistent with the 1-km data. The monthly LAI data (Myneni et al. 1997) were derived from July 1981 to June 1991 based on the global 8-km Land Pathfinder NDVI dataset (James and Kalluri 1994). The multiyear averaged monthly LAI data are used in our study. Leaf-area index here refers to the total one-sided area of all green canopy elements over unit ground area. Loveland et al. (1999) provide global 1-km land cover data for six different land cover classification schemes. For this study, we adopt the International Geosphere–Biosphere Program (IGBP) land cover classification (see Table 1), which will be provided as a standard product by the MODIS land team of the National Aeronautics and Space Administration (NASA) Earth Observing System (EOS) project.

While the 17 IGBP types are used in the derivation of $\sigma_v$, six alternative biomes defined based on vegetation structure are used in the derivation of LAI. Therefore, the $\sigma_v$ and LAI data are not necessarily consistent in each pixel. Furthermore, some of the LAI data in winter months need to be adjusted due to the poor quality of the NDVI data. In contrast, $\sigma_v$ data are derived based on the maximum NDVI in each pixel and hence are not affected by the poor quality of winter NDVI data. Therefore, for global modeling studies, the fractional vegetation cover is individually determined for each model grid box, while a mean seasonal variation of LAI is used for each IGBP land cover type within each 10° latitude zone.

Furthermore, the archived LAI data are defined with respect to *unit ground area*. For the mosaic treatment of the subgrid vegetation variation in the global model (see section 3b), these LAI values are divided by $\sigma_v$ to represent the green leaf-area index with respect to *vegetated area only* (denoted as $L_{gv}$). For each 10° latitude zone and for the land cover types of evergreen need-

TABLE 2. Roughness length $z_o$ and displacement height $d$ discussed in section 3c, parameters $\alpha$ and $L_{s,\min}$ used to compute stem and dead LAI in Eq. (2), coefficients $\alpha$ and $b$ used to compute root distribution in Eq. (3), and vegetation albedo ($A_s$ for wavelengths $<0.7$ $\mu$m and $A_l$ for wavelengths $\geq 0.7$ $\mu$m) under thick canopy conditions as discussed in section 3c.

| Type | $z_o$ (m) | $d$ (m) | $\alpha$ | $L_{s,\min}$ | $\alpha$ (m$^{-1}$) | $b$ (m$^{-1}$) | $A_s$ | $A_l$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1.0 | 11 | 0.5 | 1.0 | 6.706 | 2.175 | 0.04 | 0.20 |
| 2 | 2.2 | 23 | 0.5 | 1.0 | 7.344 | 1.303 | 0.04 | 0.20 |
| 3 | 1.0 | 11 | 0.5 | 1.0 | 7.066 | 1.953 | 0.05 | 0.23 |
| 4 | 0.8 | 13 | 0.5 | 1.0 | 5.990 | 1.955 | 0.07 | 0.24 |
| 5 | 0.8 | 13 | 0.5 | 1.0 | 4.453 | 1.631 | 0.06 | 0.24 |
| 6 | 0.1 | 0.3 | 0.5 | 1.0 | 6.326 | 1.567 | 0.07 | 0.26 |
| 7 | 0.1 | 0.3 | 0.5 | 1.0 | 7.718 | 1.262 | 0.14 | 0.32 |
| 8 | 0.7 | 6.5 | 0.5 | 1.0 | 7.604 | 2.300 | 0.06 | 0.21 |
| 9 | 0.1 | 0.7 | 0.5 | 1.0 | 8.235 | 1.627 | 0.07 | 0.26 |
| 10 | 0.03 | 0.3 | 0.5 | 1.0 | 10.74 | 2.608 | 0.07 | 0.25 |
| 11 | 0.03 | 0.3 | 0.5 | 1.0 |  |  | 0.06 | 0.18 |
| 12 | 0.06 | 0.3 | 0.0 | 0.1 | 5.558 | 2.614 | 0.06 | 0.24 |
| 13 | 0.5 | 3.0 | 0.0 | 0.1 | 5.558 | 2.614 | 0.06 | 0.22 |
| 14 | 0.06 | 0.3 | 0.25 | 0.5 | 5.558 | 2.614 | 0.06 | 0.22 |
| 16 | 0.05 | 0.1 | 0.5 | 1.0 | 4.372 | 0.978 | 0.19 | 0.38 |

leleaf and broadleaf forests (IGBP types 1 and 2, respectively), $L_{\mathrm{gv}}$ values in winter months are adjusted based on

$$L_{\mathrm{gv}} = \max(L_{\mathrm{gv}}, cL_{\mathrm{gv,max}}), \tag{1}$$

where $c = 0.7$ and 0.8 for the above two land cover types, respectively, and $L_{\mathrm{gv,max}}$ is the maximum monthly $L_{\mathrm{gv}}$.

As an example, Fig. 1 shows the $L_{\mathrm{gv}}$ data for deciduous broadleaf forests (IGBP type 4) that exist primarily over Northern Hemisphere mid- and high latitudes with a strong seasonal cycle. For instance, $L_{\mathrm{gv}}$ varies from nearly 0 in winter to 5 in summer over 40°–60°N. Deciduous broadleaf forest also exists over other latitudes and the annual range of $L_{\mathrm{gv}}$ varies significantly with latitudes. The month when $L_{\mathrm{gv}}$ reaches its peak value is also significantly different over different latitude zones. The use of the same (global mean) LAI seasonal variation for each vegetation type would ignore the significant latitudinal dependence of LAI within the same vegetation type.

In addition to green leaf-area index, the land model needs stem and dead leaf-area index with respect to vegetated area ($L_s$). For each 10° latitude zone and each land cover type, it is computed from

$$L_s^n = \max\{[\alpha L_s^{n-1} + \max(L_{\mathrm{gv}}^{n-1} - L_{\mathrm{gv}}^n, 0)], L_{s,\min}\}, \tag{2}$$

where $n$ denotes the $n$th month, $L_{s,\min}$ denotes the prescribed minimum value of $L_s$, and $(1 - \alpha)$ denotes the monthly removal rate of dead leaves. Both $\alpha$ and $L_{s,\min}$ are given in Table 2. Equation (2) differs from that in Sellers et al. (1996) in two aspects. They assumed $\alpha$ to be zero (i.e., immediate removal of all dead leaves each month), while here it is assumed to be 0.5 for most of the vegetation types (Table 2, i.e., 50% of the dead leaves are removed each month). In addition, they assumed $L_{s,\min}$ to be 0.08 or 0.05 (representing the stem area index), while here it is assumed to be 1.0 for most of the vegetation types (Table 2) (representing the stem and dead leaf area index). For deciduous broadleaf forests as shown in Fig. 1, the corresponding $L_s$ values are about 1.0 for most months over most of the latitudinal zones. Here $L_s$ values only increase north of 30°N during and after September, typically adding as much as 2 at its peak in October or November (figures not shown).

### b. Subgrid mosaic tiles

Most of the land–atmosphere coupled global models consider only the dominant vegetation type in each grid box, partly because of a lack of data. The NCAR CCM3–LSM goes one step further by assigning one of the 26 biomes to each grid box. Each biome contains up to three plant functional types (PFTs) with specified percentage area coverage for each. For instance, the biome ''tundra'' is composed of 30% arctic deciduous shrub, 30% arctic grass, and 40% bare soil. Similar to LSM, each grid box over land in CLM contains up to five tiles: lake, wetland, the first two dominant vegetation types (excluding water, snow–ice, or wetland), and bare soil. However, because of the availability of global 1-km land cover and $\sigma_v$ data as discussed in section 3a, different tiles are computed for each grid box, which is different from the PFT approach in LSM. Note that the IGBP type 16 of ''barren'' (see Table 1) still contains about 11% of vegetation cover on average (Zeng et al. 2000) and is treated as a vegetation type different from ''bare soil'' that does not contain any vegetation. All quantities are computed over each tile at each time step, and each tile maintains its own prognostic variables. The tiles in a grid box respond to the mean conditions in the overlying atmospheric grid box, and this grid box, in turn, responds to the areally weighted surface fluxes from the tiles. The tiles within a grid box do not interact with each other directly.

The first two dominant vegetation types (denoted as v1 and v2, respectively), their fractional vegetation coverages with respect to grid box (denoted as p1 and p2,

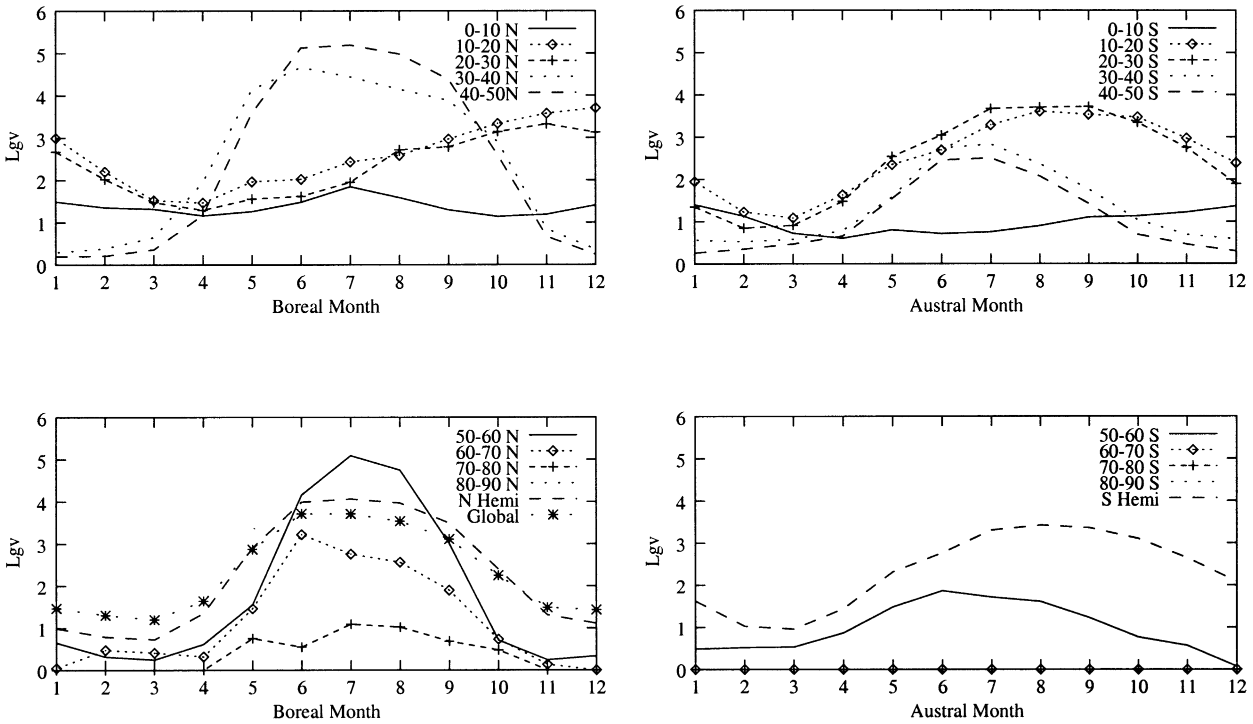


FIG. 1. The annual cycle of green leaf-area index with respect to vegetated area ($L_{gv}$) of deciduous broadleaf forests (IGBP type 4) for each 10° lat zone. Boreal and austral months are used for Northern and Southern Hemisphere zones, respectively.

respectively), as well as the percentage of wetland, lake, and bare soil with respect to grid box can be directly obtained from the global 1-km land cover and $\sigma_v$ data. The values of p1 and p2 are then adjusted to conserve the total percentage of vegetated area in the grid box while setting fractions of wetland, lake, and total vegetation less than 5% to zero. Because one of the purposes of this paper is to compare CCM3–CLM and CCM3–LSM results, these data are further adjusted to be fully consistent with the CCM3–LSM T42 land/ocean mask data in which fractional land/ocean are not considered. If fractional land/ocean are considered (e.g., in the NCAR land–atmosphere–ocean coupled Climate System Model), the above adjustment is unnecessary.

Figure 2 shows the global distribution of the dominant vegetation type and its area coverage (v1 and p1, respectively), while Fig. 3 shows the distribution of v2 and p2, respectively. The dominant type is barren with area coverage of vegetation from 0% to 25% over North Africa, while it is evergreen broadleaf forest over the Amazon basin with area coverage from 75% to 100% (Fig. 2). Over part of these two regions, p2 is zero; that is, the second dominant type is not considered (Fig. 3). Over central United States, v1 and v2 are grassland, crop-land, and cropland/natural vegetation mosaic with p1 typically from 25% to 75% and p2 typically from 5% to 25% (Figs. 2 and 3). Note that because v1 and v2 are decided based on land cover type rather than the vegetated area, p1 could be smaller than p2. For example, if barren had only slightly more area than grassland, its fractional vegetation coverage as p1 would be smaller than the grassland p2. Figure 4 shows the distribution of the area coverages of bare soil, wetland, and lakes. Bare soil covers up to 100% over North Africa, less than 10% over the Amazon region, and typically 5%–25% over central United States. Wetland is located primarily over Northern Hemisphere high latitudes, while lakes are distributed mainly over North America, followed by Eurasia and Africa. Wetland is treated as permanent in the model (i.e., with prescribed saturated soil) with vegetation properties given for each grid box.

### *c. Roughness length, vegetation root, and albedo data*

Land aerodynamic characteristics, including roughness length for momentum ($z_o$) and zero-plane displacement $d$, are proportional to the canopy height $h_c$. Both $z_o/h_c$ and $d/h_c$ are dependent on the frontal area index (i.e., the element silhouette area normal to the wind per unit surface area occupied by each element), which is, in general, a function of leaf-area index, fractional vegetation cover, and canopy height. Because of a lack of global $h_c$ data, $h_c$ is specified as a function of land cover type here following most other land models. Furthermore, since the variation of $h_c$ within a land cover type is significant but cannot be accurately determined at present, it is not unreasonable to neglect the secondary effect of LAI and $\sigma_v$ on $z_o$ and $d$. Therefore, $z_o$ and $d$ are prescribed as a function of land cover types based

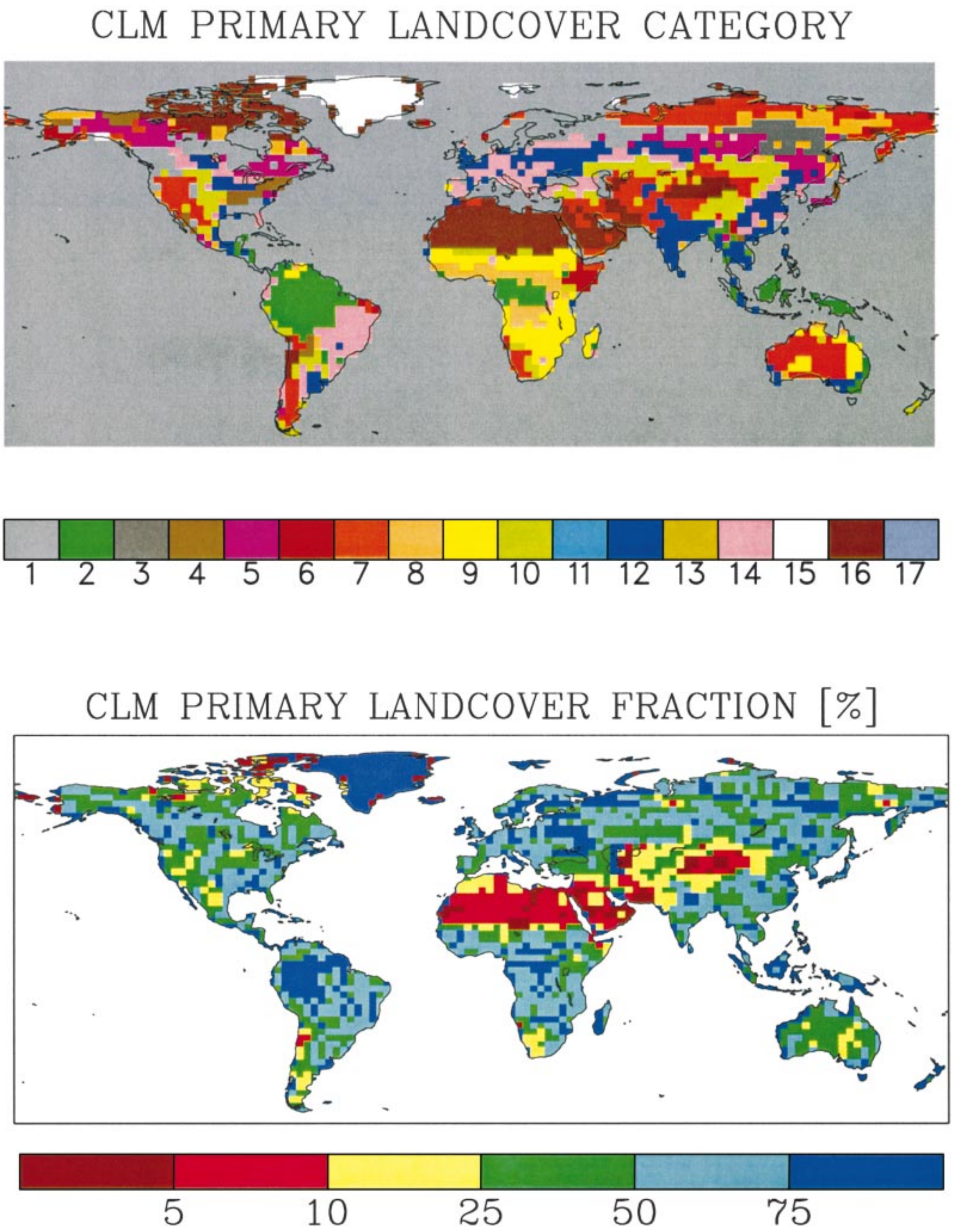


FIG. 2. The global distribution of the dominant vegetation type and its area coverage (v1 and p1, respectively) at T42 resolution.

on literature survey, as given in Table 2. Roughly speaking, $d$ is about 0.65 $h_c$, while $z_o$ varies between $0.04h_c$ and $0.2h_c$.

While aerodynamic roughness $z_o$ is used for wind, thermal roughness $z_{oh}$ is needed for heat and water vapor. In general, $z_o$ is different from $z_{oh}$, because the momentum transfer is effected by pressure fluctuations in the turbulent wakes behind the roughness elements, while no such dynamical mechanism exists for heat and water vapor transfer. Rather, heat and water vapor must ultimately be transferred by molecular diffusion across the interfacial sublayer. Instead of prescribing $z_{oh}$ for

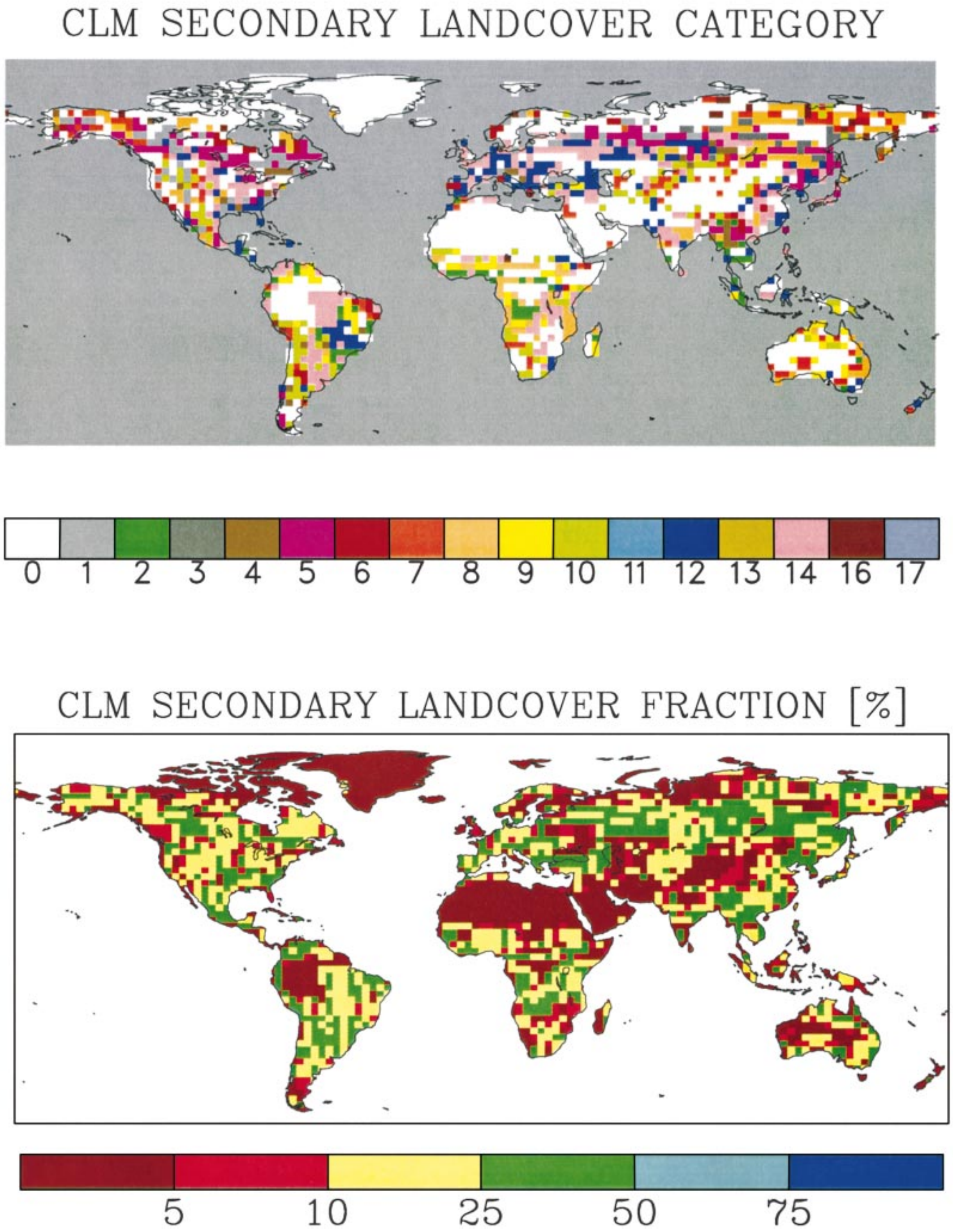
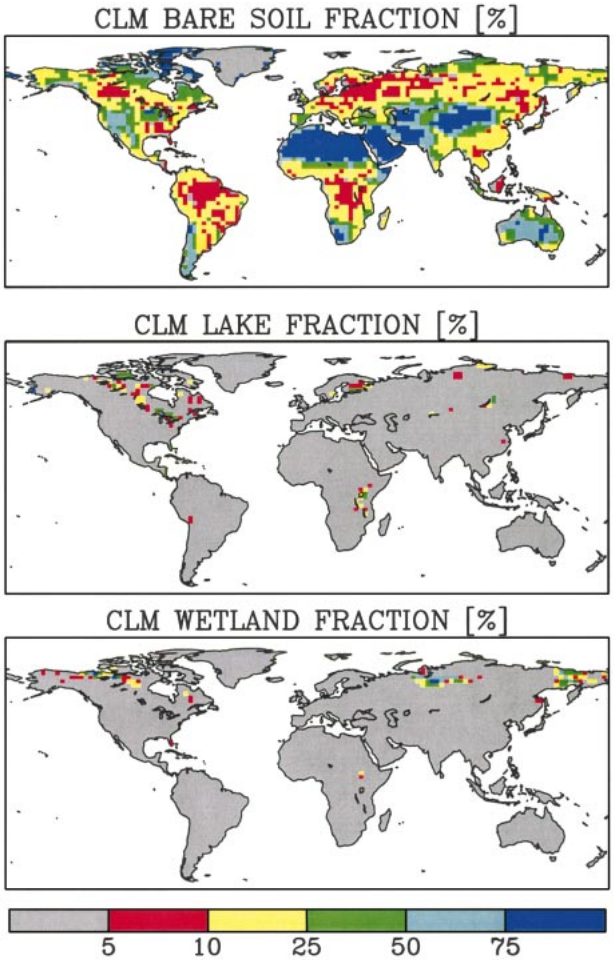


FIG. 3. The same as Fig. 3 except for v2 and p2, respectively. The vegetation type v2 is denoted by a numbered 0 color bar if its fractional cover p2 is 0.

each land cover, CLM uses the formulation from Zeng and Dickinson (1998) to compute $z_{oh}$ from $z_o$ and near-surface variables.

Vegetation root distribution (including rooting depth) directly affects the water holding capacity of the land surface and the relative rates of water extraction from the different layers in the rooting zone. Using the comprehensive field survey data of vegetation root distribution (Jackson et al. 1996) and maximum rooting depth (Canadell et al. 1996), Zeng (2001) has generalized the approach in Zeng et al. (1998b) to provide vegetation root distribution (including rooting depth) that can be used by any land model with any number of soil layers in the vegetation rooting zone:



CLM BARE SOIL FRACTION [%]

CLM LAKE FRACTION [%]

CLM WETLAND FRACTION [%]

FIG. 4. The global distribution of the area coverages of bare soil, wetland, and lakes at T42 resolution. Note that bare soil fraction is given as zero for permanent snow and ice (IGBP type 15).

$$Y = 1 - \frac{1}{2}(e^{-az} + e^{-bz}), \tag{3}$$

where $Y$ is the cumulative root fraction from the surface to depth $z$ (positive), and $a$ and $b$ are coefficients depending upon vegetation types, as given in Table 2.

A simplified two-stream approach is used in CLM to compute the radiative transfer through canopy under the constraint that vegetation albedo approaches that of the underlying ground when leaf-area index goes to 0 and approaches a prescribed albedo value for each vegetation type when leaf-area index becomes very large (see section 2). The prescribed albedo values for thick canopy are given in Table 2. These values are initially prescribed based on values in BATS (Dickinson et al. 1993) and then adjusted (more specifically, slightly reduced) based on the analysis of AVHRR derived albedo data of Strugnell et al. (2001) in Xue et al. (2001).

As in BATS (Dickinson et al. 1993) and LSM (Bonan 1996), soil albedos in CLM vary with soil color class (as well as soil moisture). In addition to the eight color classes in BATS, LSM creates a ninth class with high albedos to better match the clear-sky top-of-the-atmosphere albedos derived from the Earth Radiation Budget Experiment (ERBE) for desert and semidesert surface types located in North Africa and the Arabian Peninsula. However, analysis of the AVHRR-derived albedo data of Strugnell et al. (2001) in Xue et al. (2001) suggests that soil color class 9 is unnecessary. Therefore, CLM uses LSM soil color data after replacing soil color class 9 by class 1 and increasing soil color class (or decreasing soil albedos) for some of the desert or semidesert grid boxes based on Xue et al. (2001). Since the quality of the AVHRR-derived albedo data is affected by the lack of aerosol distribution data (particularly over desert and semidesert), new satellite data and in situ measurements are still needed to accurately determine surface albedo over these regions. Finally, the soil texture data are the same as those in LSM.

## 4. Results

Results are based on a 15-yr simulation of the NCAR CCM3 coupled with CLM using observed sea surface temperature and sea ice for the period of 1979–93. Initial conditions for the land and atmosphere were obtained from a 10-yr spinup simulation using CCM3–CLM. Results from CCM3 coupled with LSM and BATS in previous research are used for comparison, although most of the discussions below will be focused on CLM versus LSM comparison. The same sea surface temperature and sea ice data for the same period and the same spinup procedure were used in the CCM3–LSM and CCM3–BATS simulations. For brevity, simulations of CCM3 coupled with CLM, BATS, and LSM will be referred to as CLM, BATS, and LSM, respectively, hereafter. Similar to Bonan (1998), only the surface air temperature and hydrology are emphasized here.

### *a. Surface air temperature*

Figure 5 compares the 2-m surface air temperature from LSM and CLM with the 0.5° × 0.5° temperature climatology in Willmott et al. (1998). The difference fields between LSM and observations are similar to those shown in Figs. 1 and 3 of Bonan (1998): in January, LSM has a cold bias over most of the regions except over part of the Northern Hemisphere high latitudes, while LSM has an overall cold bias in July. Figure 5 also shows that CLM significantly reduces the cold biases of LSM in July over most of the regions, particularly over Eurasia and North Africa. In January, CLM significantly reduces the warm bias of LSM over Eurasia in high latitudes and reduces the cold biases of LSM over North Africa and Southern Hemisphere midlatitudes.

To gain further insight to the model simulation of surface air temperature, Fig. 6 compares the annual cycle of surface air temperatures from CCM3 coupled with

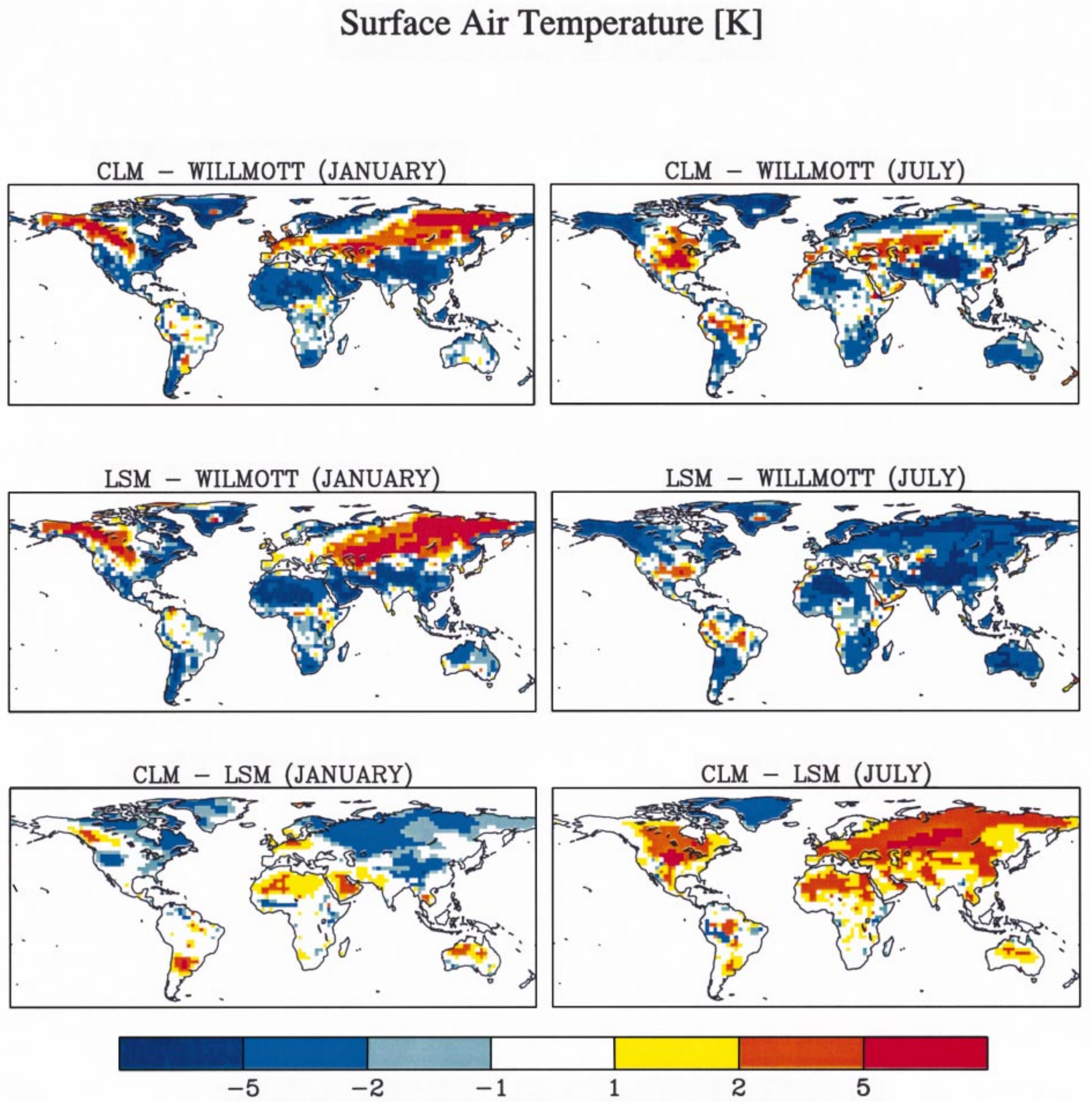


FIG. 5. Differences in surface air temperature in Jan and Jul between models (CCM3 − CLM and CCM3 − LSM) and observations (Willmott et al. 1998) and between the two models.

CLM, LSM, and BATS, respectively, over seven regions. The continental United States (29°–49°N and 85°–125°W), the Amazon region (0°–10°S and 50°–70°W), and western Siberia (50°–70°N and 60°–90°E) represent midlatitude land, tropical forest, and high-latitude forest, respectively, while the Sahara and the Arabian Peninsula (10°–30°N and 20°W–50°E) represent the subtropical desert. The shaded area in each panel indicates CLM's 95% confidence interval for the mean (i.e., the CLM vs observational temperature difference plus or minus twice the standard deviation based on CLM's 15-yr climatology). Note that there are well-documented shortcomings in the use of the *t* statistic for assessing differences between model results and observational data, partly because temperature or hydrological variables are not normally distributed. Nonetheless, if the temperature difference fields in the shaded area contain both positive and negative values, the CLM versus observed temperatures are more or less not significantly different (i.e., the observed temperatures fall

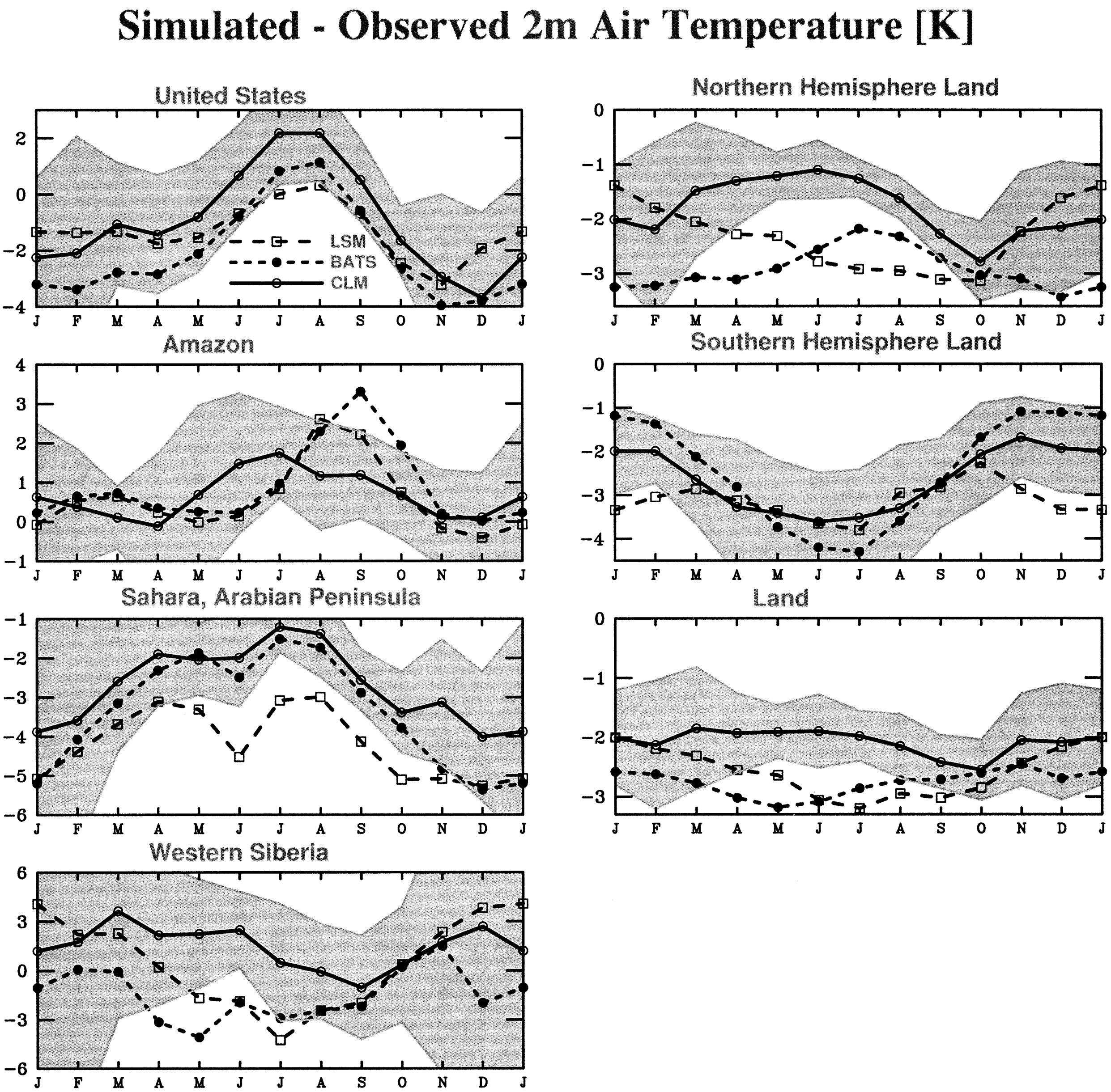


FIG. 6. Monthly differences in surface air temperature between CCM3 coupled with CLM, LSM, and BATS vs observations (Willmott et al. 1998) over seven regions. Four of the regions are the United States (29°–49°N and 85°–125°W), Amazon (0°–10°S and 50°–70°W), the Sahara and the Arabian Peninsula (10°–30°N and 20°W–50°E), and western Siberia (50°–70°N and 60°–90°E). The shaded area in each panel indicates CLM's 95% confidence interval for the mean based on its 15-yr climatology.

within CLM's interannual variability). Similarly, if the LSM results fall within the shaded area, LSM and CLM results are more or less not significantly different.

Figure 6 shows that over the continental United States, temperature simulations from all three models are better in warm seasons (May–September) than in other months. In the Amazon basin, monthly temperature biases of both CLM and LSM from observations are within 1.0°C for most months. Over western Siberia, LSM is too warm in winter but too cold in summer, which reduces the temperature annual range by up to 8°C compared with observational data. In contrast, the temperature bias from CLM is within 2°C for most months. In these three regions, the observed temperatures fall within CLM's interannual variability for most months, because the temperature difference fields in the shaded area contains both positive and negative values. Similarly, CLM and LSM results are not significantly different for most months. In the Sahara and the Arabian Desert, LSM significantly underestimates surface air temperature, which is partly caused by prescribed high soil albedos (Bonan 1998). The cold bias is reduced

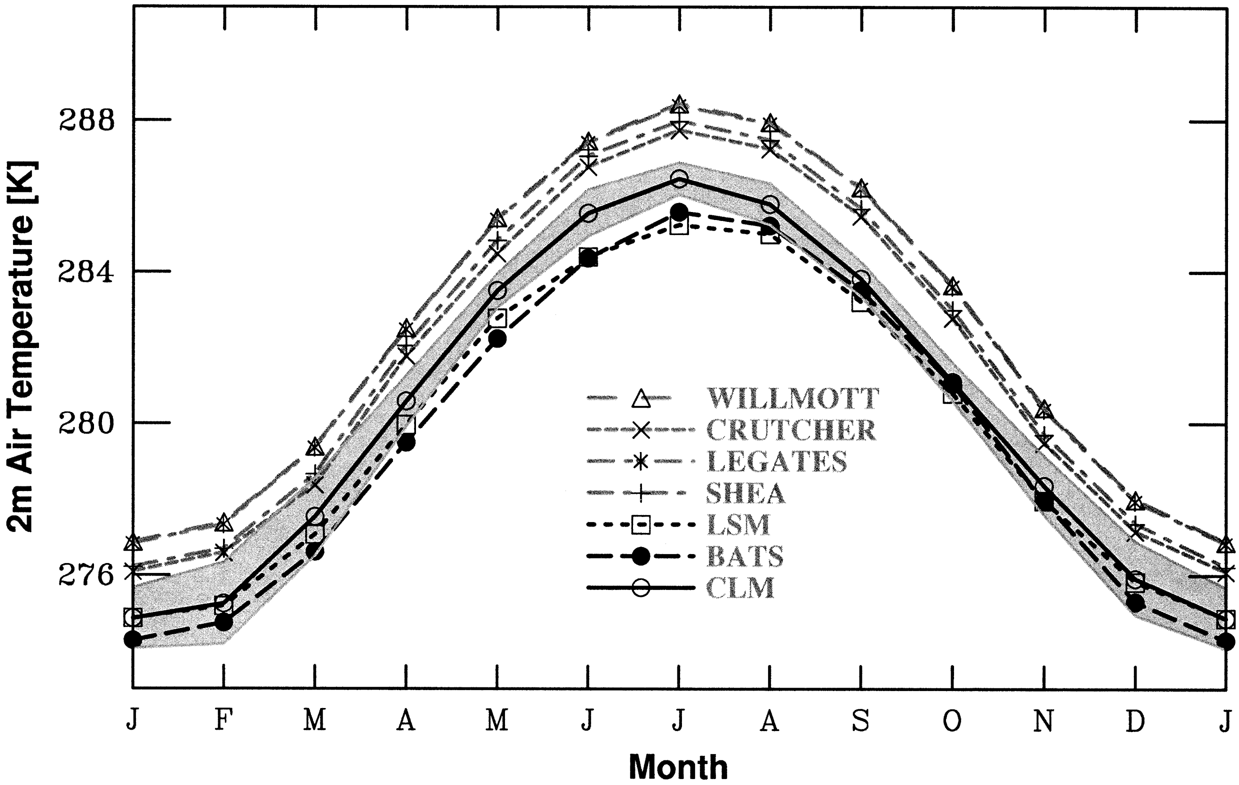


FIG. 7. Monthly surface air temperature over global land as simulated by CCM3 coupled with CLM, LSM, and BATS and from four observational datasets (Crutcher and Jenne 1969; Shea 1986; Legates and Willmott 1990a; Willmott et al. 1998). The shaded area indicates CLM's 95% confidence interval for the mean based on its 15-yr climatology.

throughout the year in CLM largely because lower soil albedos are prescribed based on data analysis in Xue et al. (2001). In particular, the reduction of the cold bias in LSM is statistically significant at the 95% level from June to October. Over both Northern Hemisphere land and global land, CLM significantly reduces the cold biases in LSM from May to September. For instance, the cold bias of 3.1°C in summer (June–August) over global land in LSM is reduced by 1.1°C in CLM. Over Southern Hemisphere land, CLM also significantly reduces the cold biases in LSM during austral summer (December–February).

To address the possible impact of uncertainties in observational data on the above results, Fig. 7 shows the annual cycle of surface air temperature over global land from three model simulations and four observational datasets. The Willmott et al. (1998) dataset represents the long-term average of surface observations at 0.5° × 0.5° resolution and is a revised version of the original Legates and Willmott (1990a) dataset, which is also used here for comparison. The Shea (1986) dataset represents the average temperature from 1950 to 1979 at 2.5° × 2.5° resolution, while the Crutcher and Jenne (1969) dataset represents the average temperature from 1950 to 1964 at 5° × 5° resolution. In contrast, the average temperatures of 15 yr (1979–93) are used from each of the three models. Among the four observed datasets, Crutcher and Jenne (1969) give the lowest surface air temperature while Willmott et al. (1998) and Legates and Willmott (1990a) give the highest. However, each of these datasets gives a higher temperature than any of the three models, and the difference among observed datasets is smaller than the difference between model results and observations. Furthermore, most of the observed values lie outside the 95% confidence interval of the CLM (LSM or BATS) results. Over a specific region, however, the temperature difference among the four observational datasets could be larger than the difference between these datasets and model results. For instance, over the Amazon region, the annual mean temperature difference from CLM is less than 0.5°C cooler than that from Legates and Willmott (1990a), while the difference between the Willmott et al. (1998) and Legates and Willmott (1990a) datasets is about 1°C (not shown).

The above results demonstrate that CLM improves the simulation of surface air temperature compared with LSM, and the possible physical reasons for these improvements will be explored later. However, it is difficult to judge to what extent the remaining biases are real or rather an artifact of incompatibility between the observations and model results. The major incompatibilities include the following: the model surface variables are determined as an average over the model surface, whereas surface observations, to the extent they are sited according to World Meteorological Organization standards, are for a well watered short grass surface; the model surface temperatures are for a 24-h average whereas the observations evaluate daily temperature as an average of daily minimum and maximum; the observational sites generally have a low elevation bias in mountainous terrain. These biases largely act to shift the observations to higher temperatures than those calculated in the models. Furthermore, as a diagnostic variable, surface air temperature can be computed by various methods in models. To directly compare CLM versus LSM results, CLM adopts the same method as LSM (Bonan 1996), which may introduce some biases over forests. Remaining discrepancies between CLM and observations cannot be assessed until these factors are adequately accounted for as well as the contributions of errors in the atmospheric simulation in terms of spatial patterns and contributions to surface radiation budgets. As an example, 30 months of 30-min average surface air temperature data obtained over midlatitude grassland (Betts and Ball 1998) are used here to illustrate the impact of the two averaging methods on the computed monthly temperatures. It is found that monthly mean temperature using hourly samples can be lower by 0.2°–1.2°C than that using daily maximum and minimum temperatures, and the temperature difference averaged for these 30 months using these two methods is 0.6°C. To illustrate the impact on surface air temperature of generally low elevation bias of observational sites, the Willmott et al. (1998) dataset (which does not use the station-height information directly) is compared with the Willmott and Matsuura (2000) dataset, which employs the digital-elevation-model (DEM)-assisted interpolation of air temperature with a constant lapse rate of 6 K km$^{-1}$ (independent of geographic location and season). Temperature differences of up to 20°C are found between the two datasets over part of the Tibet, Greenland, and Antarctic. The surface air temperature averaged over global land from Willmott and Matsuura (2000) is about 1.0°C cooler than that from Willmott et al. (1998).

### *b. Surface hydrology*

Figure 8 evaluates the model precipitation climatology in January and July using the Willmott et al. (1998) dataset. Overall, CCM3–LSM simulates the principal features of the observed precipitation distribution. Some of its notable deficiencies include the following: an overestimate of precipitation in January over western North America, Australia, South Africa, and part of South America; an underestimate of July precipitation over most tropical and subtropical land; and a slight overestimate of July precipitation over most of the Northern Hemisphere high latitudes. The precipitation differences between CLM and LSM are quite small over most of the regions, probably because both simulations were forced by the same sea surface temperature data. Evidently, the precipitation distribution in CCM3 is more controlled by the land–sea contrast, topography, and atmospheric processes than by the change of land models from LSM to CLM. Notable improvements in simulated precipitation in CLM include the reduction of January and July precipitation biases in LSM in Australia, and the reduction of precipitation overestimation in LSM over an east–west band in Eurasia in July.

Figure 9 compares the annual cycle of precipitation from CCM3 coupled with CLM, LSM, and BATS in seven regions. Over the United States, model precipitation bias is within 0.5 mm day$^{-1}$ for most months. In the Amazon region, CLM simulates precipitation better than LSM in October and November, which represent the transition period from the dry season to the wet season. The last month (during or after the dry season) when model precipitation is below observations is July, August, and September for CLM, LSM, and BATS, respectively, in Fig. 9, which is consistent with the month for the peak warm bias of models in Fig. 6. In the Sahara and Arabian Peninsula, model precipitation agrees with observations very well for most months. Over western Siberia, model precipitation bias is within 0.4 mm day$^{-1}$ for most months. Over Northern Hemisphere land, CLM and LSM overestimate precipitation in March to May but they underestimate it in July to September. Over Southern Hemisphere land, all models underestimate precipitation from March to August. Over global land, the model bias is within 0.2 mm day$^{-1}$ for most months. In each panel, the differences between CLM and LSM results are not significant at the 95% level for most (or all) months.

To address the possible impact of uncertainties in observational data on the above results, Fig. 10 shows the annual cycle of surface precipitation over global land from three model simulations and five observational datasets. The Xie and Arkin (1997) dataset represents the climatology from 17 yr (1979–95) of data at 2.5° × 2.5° resolution as a combination of rain gauge data, satellite estimates, and National Center for Environmental Prediction–NCAR reanalysis data. In contrast, other observed datasets are all derived from rain gauges only: the Shea (1986) dataset represents the average precipitation from 1950 to 1979, while the Jaeger (1983) and Legates and Willmott (1990b) datasets represent the long-term averages of rain gauge data. The Willmott et al. (1998) dataset represents a revised version of the original Legates and Willmott (1990b) dataset. CLM and LSM results are not significantly different. They show low biases compared to the Willmott et al. (1998) and Legates and Willmott (1990b) datasets in June–August. In addition, they are biased high relative to the other datasets except for June–September. These model biases may be real but the variability among various datasets suggests caution in drawing such a conclusion.

Snow processes, particularly the timing of snowmelt and the subsequent fate of meltwater, play an important

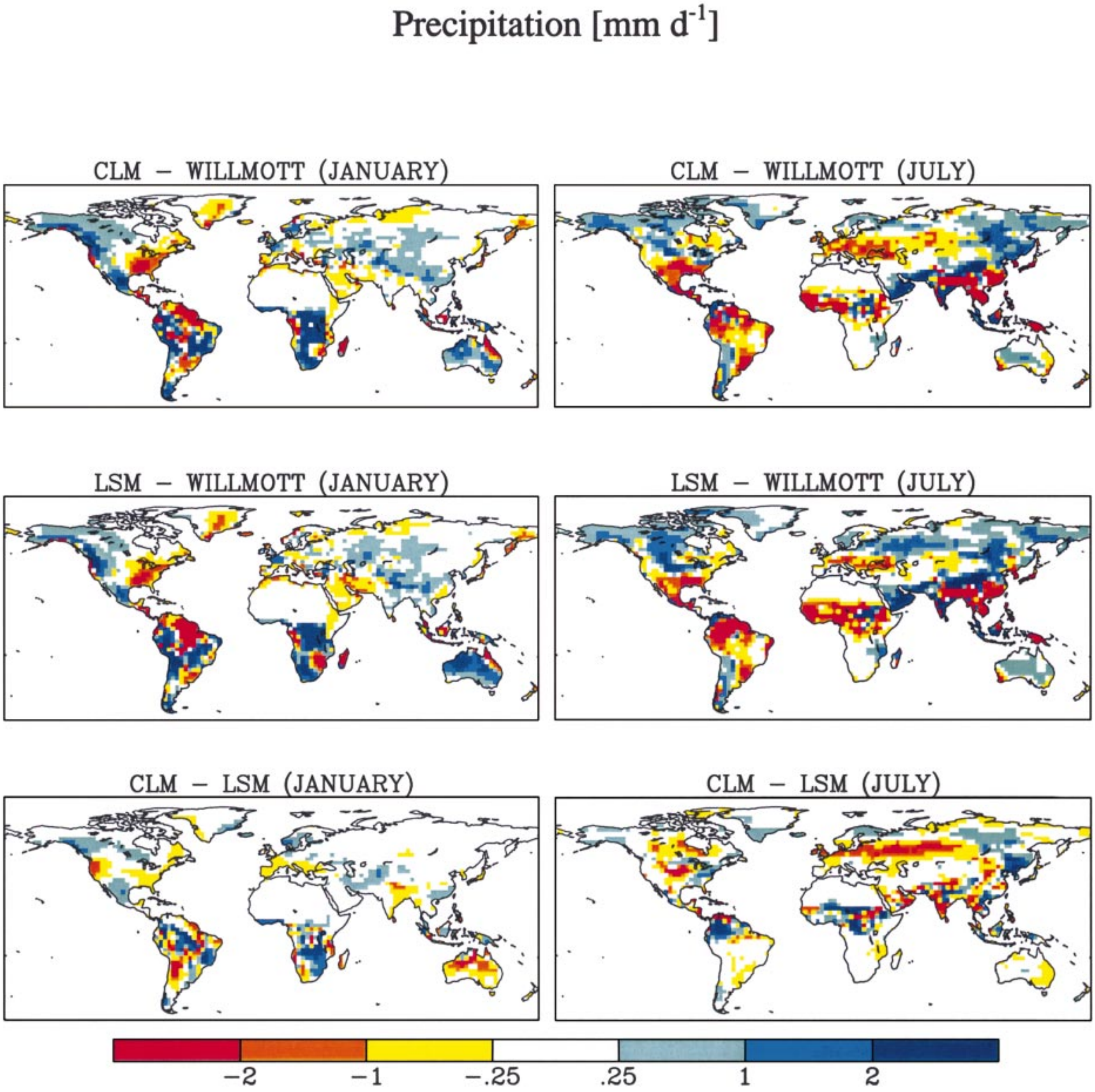


FIG. 8. Precipitation differences in Jan and Jul between models (CCM3 − CLM and CCM3 − LSM) and observations (Willmott et al. 1998) and between the two models.

role in the overall hydrological cycle. They also strongly affect the surface energy cycle primarily because of snow's high albedo and low thermal conductivity compared with soil. Both snow cover and depth data are already available from observations (e.g., Foster et al. 1996); however, comparison of model and observational snow cover data is strongly affected by at least two uncertainties; that is, the criterion used to define full snow cover in a grid box and the consideration of fractional snow cover. Neither snow cover nor depth data were saved in the integration of CCM3 coupled with CLM, BATS, or LSM. However, snow mass is saved in all global models that consider snow processes. Figure 11 evaluates snow mass over North America (here defined as the broad region between 0° and 90°N and 10° and 170°W but excluding Greenland) and Eurasia (here defined as the broad region between 0° and 90°N and 170° and 10°W through the date line) using data reported in Foster et al. (1996). The original snow depth data were compiled by the U.S. Air Force Environmental Technical Applications Center at Scott Air Force Base in Illinois based on climatological records, literature searches, surface weather synoptic reports, and data obtained at snow course sites (Foster and Davy

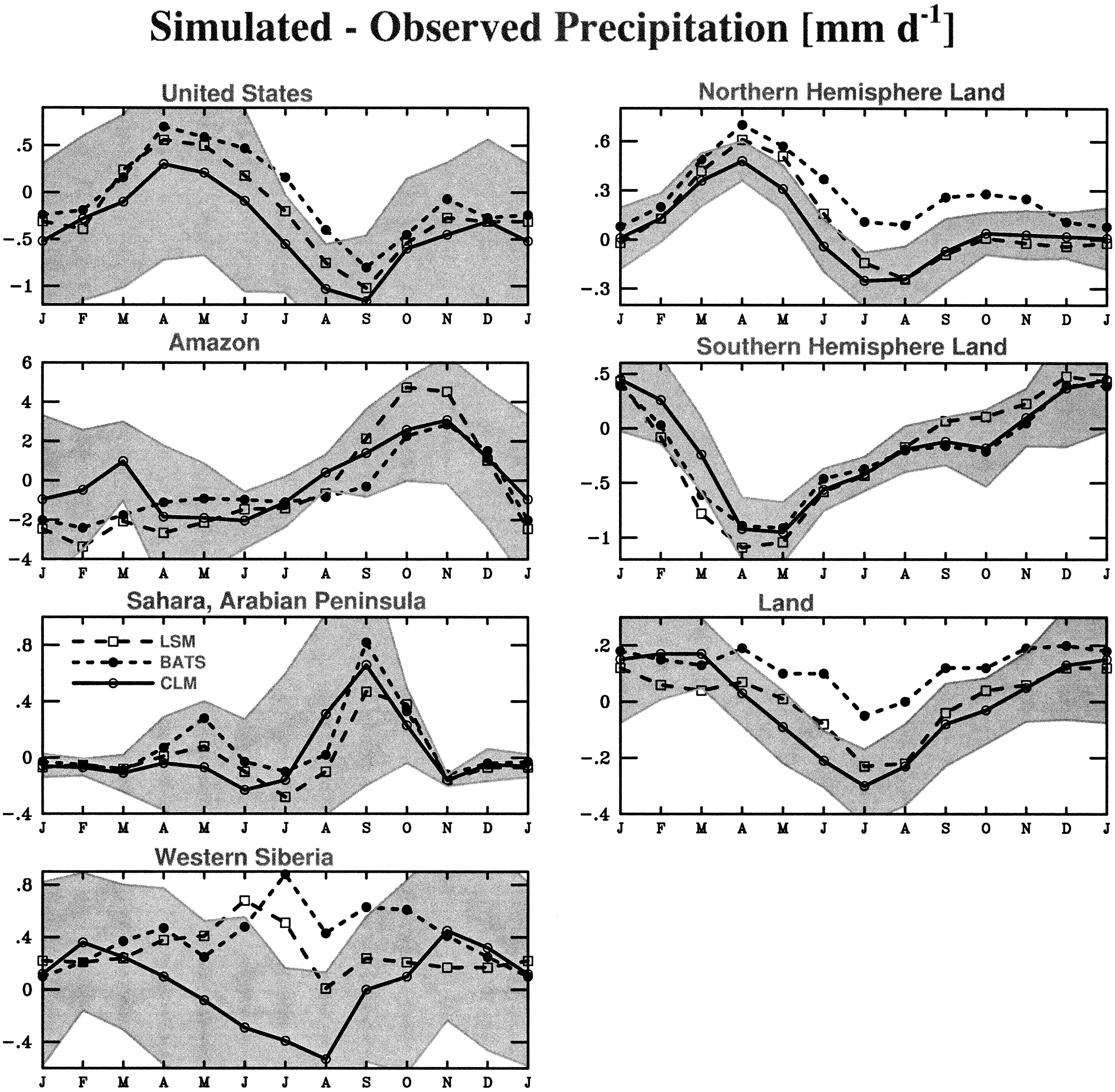


FIG. 9. Monthly precipitation differences between CCM3 coupled with CLM, LSM, and BATS vs observations (Willmott et al. 1998) over seven regions. Definitions of these regions are the same as in Fig. 6. The shaded area in each panel indicates CLM's 95% confidence interval for the mean based on its 15-yr climatology.

1988). A constant density of 300 kg m$^{-3}$ was then used to convert snow depth to mass in Foster et al. (1996). All three models reach the peak snow mass in March over North America, in agreement with observations, while they lag observations by 1 month to reach the peak snow mass over Eurasia, or equivalently, they are slow in their spring snowmelt. LSM underestimates snow mass for each month over North America and for most months over Eurasia. Overall, snow mass from CLM is closer to observations than that from LSM, provided the conversion from depth data has not seriously overestimated the mass.

Runoff (including surface runoff and base flow) is another important component of the surface hydrological cycle, and is driven by precipitation and snowmelt. Recently, Fekete et al. (2000) have developed global composite runoff fields between 83°N and 55.5°S [denoted as the University of New Hampshire (UNH)–Global Runoff Data Center (GRDC) dataset] by combining observed river discharge information with a climate-driven water balance model. First, a water balance model was used to simulate global runoff as driven by the long-term averaged monthly precipitation and surface air temperature data of Willmott et al. (1998). The

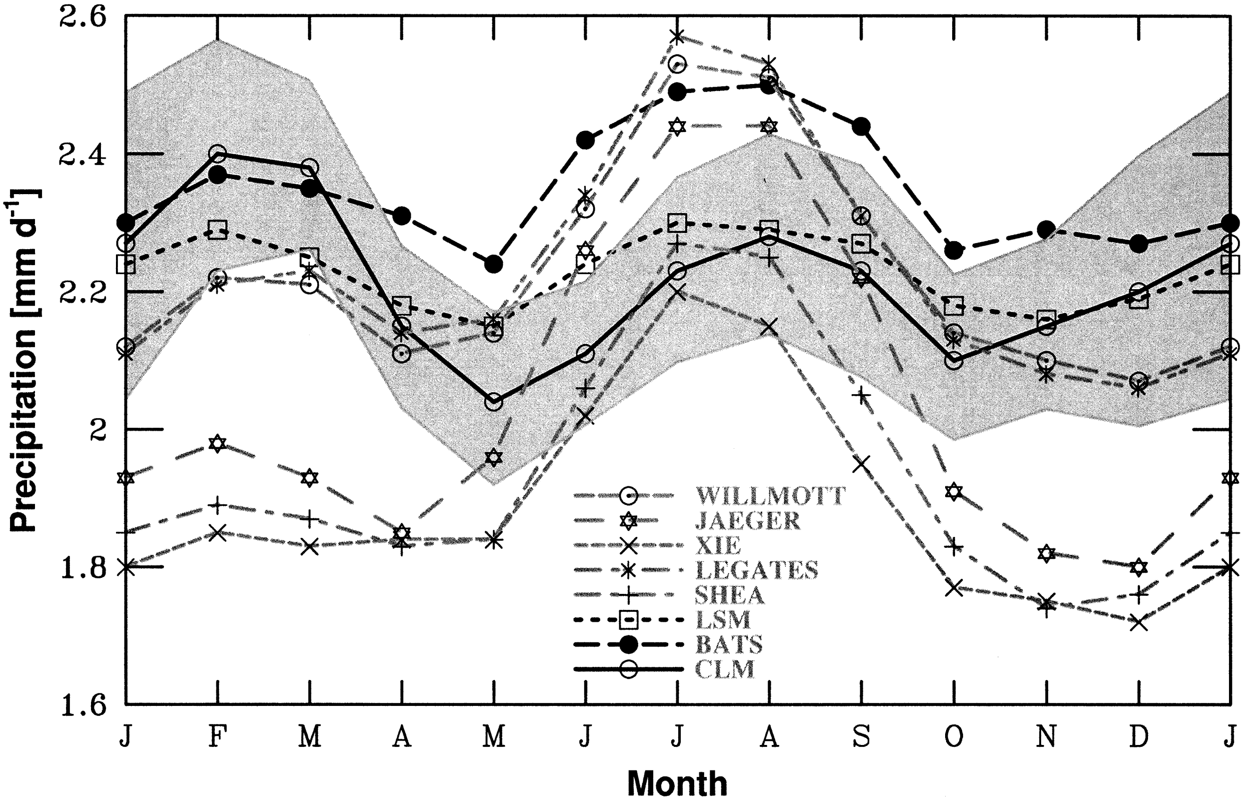


FIG. 10. Monthly precipitation over global land as simulated by CCM3 coupled with CLM, LSM, and BATS and from five observational datasets (Jaeger 1983; Shea 1986; Legates and Willmott 1990b; Xie and Arkin 1997; Willmott et al. 1998). The shaded area indicates CLM's 95% confidence interval for the mean based on its 15-yr climatology.

simulated annual mean runoff was then adjusted to fit the observed river discharge data. The seasonal cycle of the simulated runoff in the UNH–GRDC dataset is affected by several assumptions of the water balance model: precipitation is considered snow when the monthly surface air temperature is below −1°C; snowmelt is a prescribed function of temperature and elevation; and runoff is formed either as snowmelt or when the surplus from the difference between precipitation and evaporation exceeds the soil moisture deficit from the soil's water holding capacity (Fekete et al. 2000). In contrast, precipitation, snowmelt, and runoff are computed at each time step (i.e., 20 min) in the climate model. Precipitation is assumed to be snow when the surface air temperature is below 2.2°C, and snowmelt is computed as part of the snow parameterization. Runoff in CLM is generated by a saturated area runoff component that responds rapidly to water added to the surface and a base-flow term that also will generally be largest near the time of maximum soil moisture but will decay on a longer timescale. Overall CLM appears to generate a larger fraction of its runoff rapidly near the time of maximum soil moisture.

Figure 12 evaluates the annual cycle of model runoff over seven regions using the UNH–GRDC dataset. Over the United States, which represents one of the regions with the highest density of discharge and precipitation gauge stations, CLM simulates runoff very well throughout the year, including the correct timing of peak runoff (in April) and minimum runoff (in September). In contrast, LSM incorrectly gives the peak runoff in May and minimum runoff in December, and the amplitude of the seasonal cycle in the LSM runoff is just 0.4 mm day$^{-1}$, or about 40% of that from the UNH–GRDC dataset (or CLM), apparently because of a stronger base-flow component in LSM. In the Amazon basin, peak precipitation occurs in February or March from the five observational datasets used in Fig. 10 (not shown). CLM gives the correct month (March) of peak precipitation (not shown) and hence gives the peak runoff in March (Fig. 12), which is 1 month earlier than indicated by the UNH–GRDC dataset. In contrast, LSM gives the peak precipitation in November (with a weaker peak in March) (not shown) and hence gives the peak runoff incorrectly in December. Over the Sahara and Arabian Desert, both CLM and LSM give the correct seasonal variation of precipitation (not shown). Figure 12 shows that CLM-simulated runoff is nearly zero except between July and October, and reaches its peak in August, both in agreement with the data, while LSM gives an incorrect month (October) for the peak runoff and maintains runoff in February and March when the

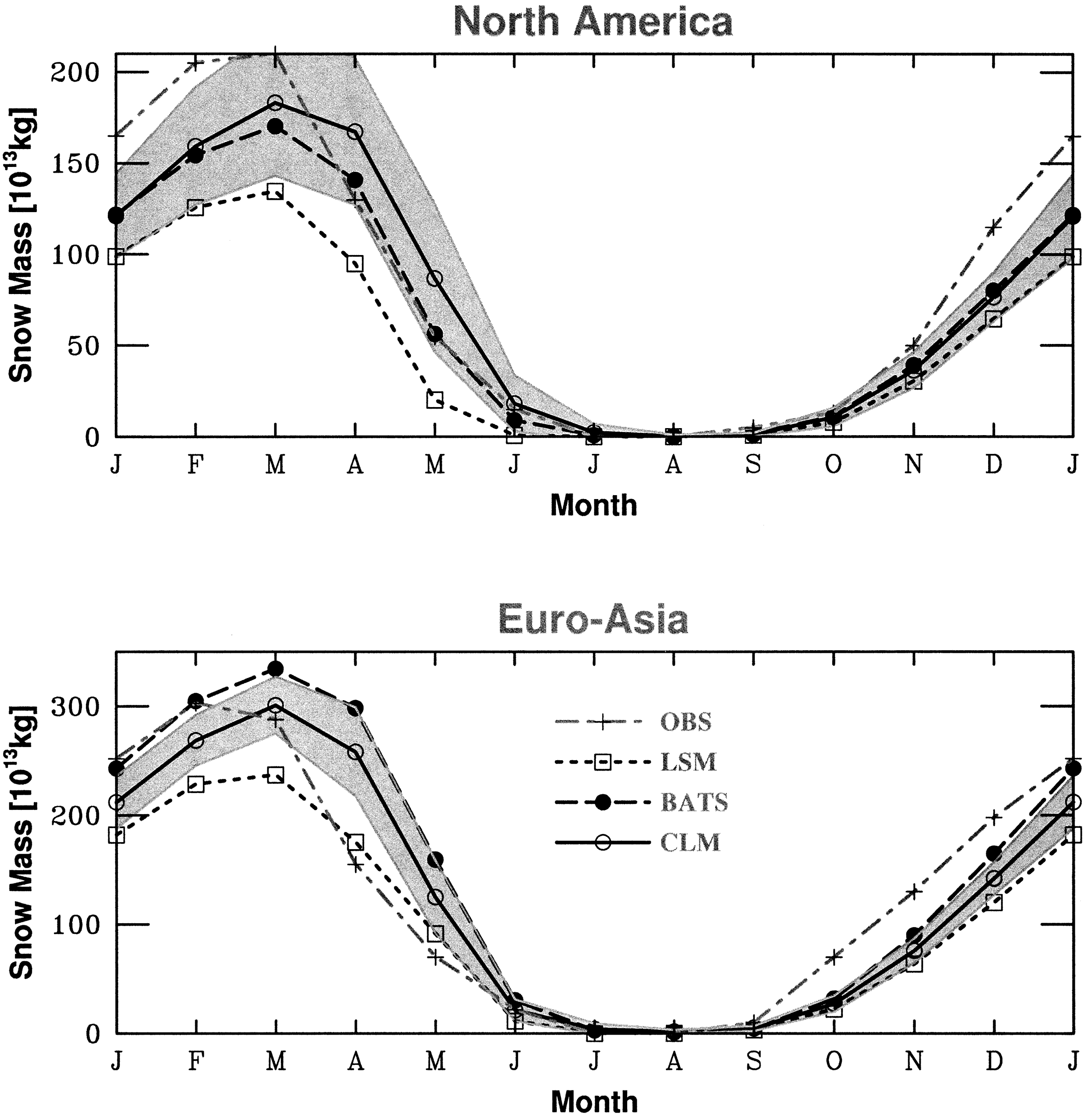


FIG. 11. Monthly snow mass as simulated by CCM3 coupled with CLM, LSM, and BATS and from observations (Foster and Davy 1988) over North America (0°–90°N and 10°–170°W but excluding Greenland) and Eurasia (0°–90°N and 170°–10°W through the date line). The shaded area in each panel indicates CLM's 95% confidence interval for the mean based on its 15-yr climatology.

runoff is zero from the data. Over the western Siberia, the peak runoff is too high in CLM but too low in LSM. CLM gives the peak runoff in April probably because of the early snowmelt associated with the warm bias in winter (e.g., Fig. 5). In contrast, LSM does not show any peak runoff associated with snowmelt; instead it incorrectly shows the peak in October (after the peak precipitation in summer).

Over Northern Hemisphere land, Fig. 12 also shows that the peak runoff occurs in May in CLM, which is consistent with the maximum decrease of snow mass from April to May during the snowmelt season (Fig. 11). The 1-month difference in the timing of the peak runoff between CLM and the UNH–GRDC dataset could be caused in part by the different methods used for computing snowmelt, as mentioned earlier. LSM shows the incorrect seasonal cycle with the peak runoff in October even though the seasonal cycle of precipitation is quite similar between CLM and LSM (Fig. 9). The snowmelt in late spring (Fig. 11) only slightly in-

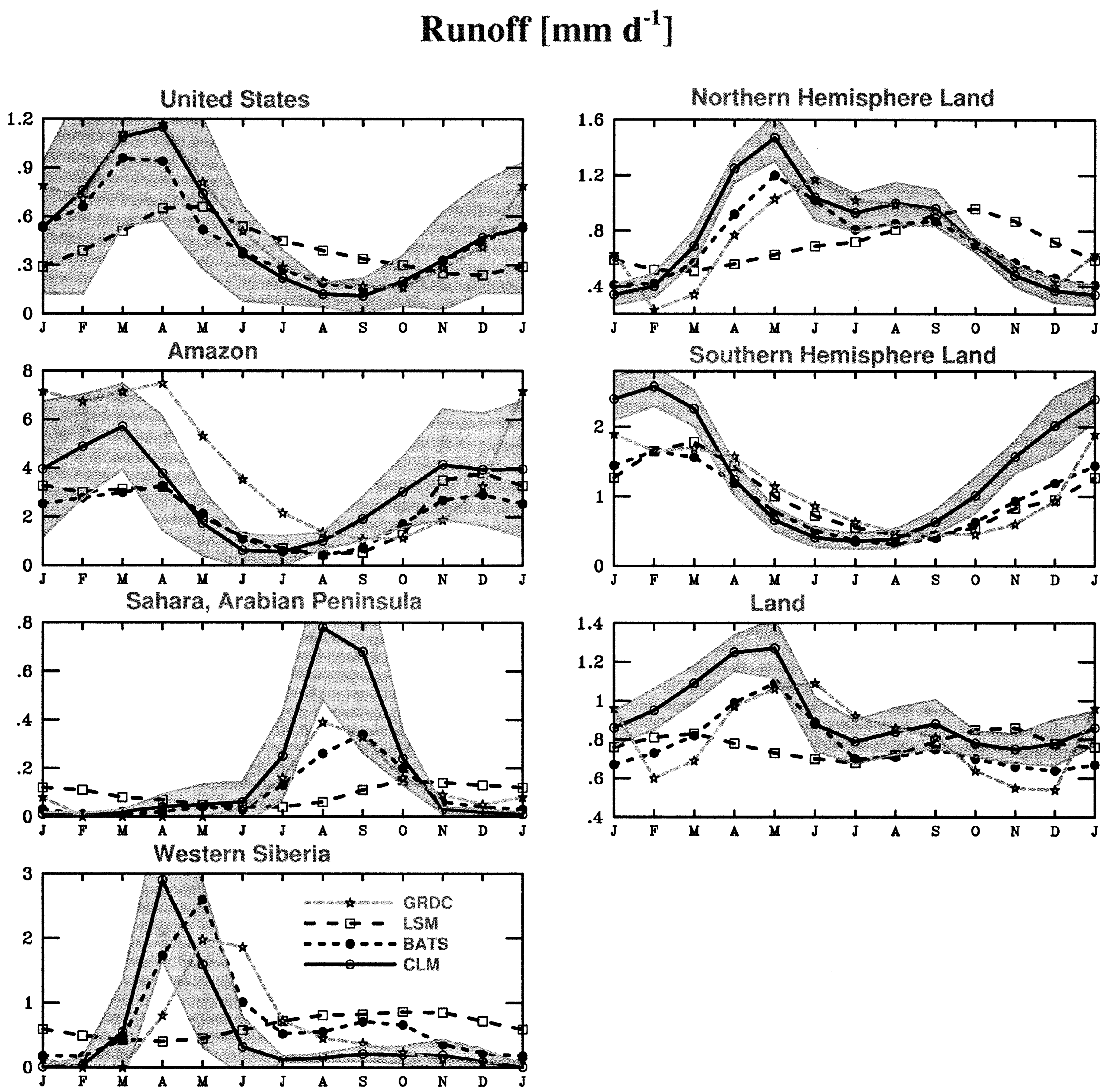


FIG. 12. Monthly surface runoff over seven regions as simulated by CCM3 coupled with CLM, LSM, and BATS and from the UNH–GRDC dataset (Fekete et al. 2000). Definitions of these regions are the same as in Fig. 6 except that Antarctica is excluded. The shaded area in each panel indicates CLM's 95% confidence interval for the mean based on its 15-yr climatology.

creases runoff in LSM. Evidently, LSM and CLM have different partitionings of meltwater among soil water, evaporation, and runoff, with the partitioning from CLM apparently more consistent with observations. Over Southern Hemisphere land (excluding Antarctic), the month for the peak runoff is February in CLM (or BATS) and March in LSM, while it is January in the UNH–GRDC dataset. CLM significantly overestimates runoff during austral summer (December–February), partly because of the overestimate of precipitation (Fig. 9). Over global land (excluding Antarctic), the timing for the peak (or minimum) runoff is within 1 month between CLM and the data, while the timing differs by 5 or 6 months between LSM and the data. CLM overestimates the peak runoff by 0.18 mm day$^{-1}$, while LSM underestimates it by 0.23 mm day$^{-1}$. The annual range of monthly runoff is very small (0.18 mm day$^{-1}$) in LSM, 0.52 mm day$^{-1}$ in CLM, and 0.55 mm day$^{-1}$ in the UNH–GRDC dataset. The annual average runoff over global land (excluding Antarctic) is 0.81, 0.93, 0.78, and 0.77 mm day$^{-1}$ from the UNH–GRDC dataset, CLM, BATS, and LSM, respectively. Note that, since the observed river discharge is affected by human activities (e.g., dams and agricultural, industrial, and

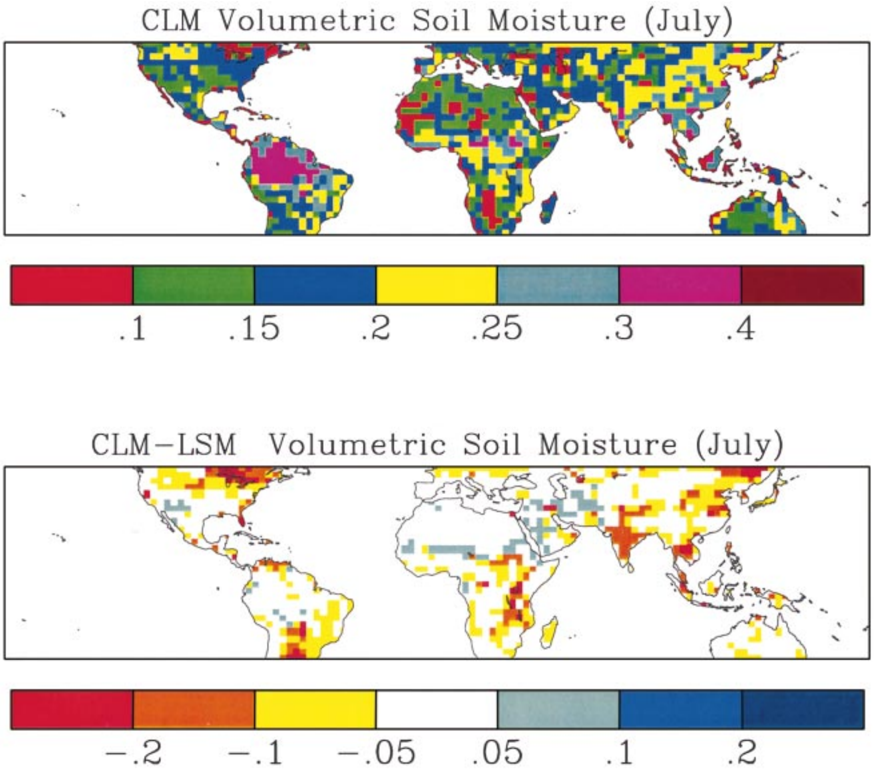


FIG. 13. (top) Volumetric soil water in the top 1 m of soil in CCM3–CLM and (bottom) volumetric soil moisture difference in the top 1 m of soil between CCM3–CLM and CCM3–LSM for the month of July.

household consumptions of water), the UNH–GRDC dataset may underestimate the natural annual mean runoff.

Another important component of the hydrological cycle is soil moisture. LSM determines only soil water, whereas CLM considers both soil water and ice. The ice content, however, was not saved. Figure 13 compares the volumetric soil water [mm(water)$^3$ mm(soil)$^{-3}$] for the top 1 m of soil in July from CLM and LSM in the regions between 30°S and 50°N where soil ice is expected to be insignificant. Figure 13 shows that CLM reproduces the expected spatial patterns of soil water: wet soil over tropical forests, very dry soil over arid regions (such as the Sahara and the Arabian Desert), and relatively dry soil over Northern Hemisphere midlatitude land due to depletion of soil water during the growing season. There are few large-scale observations for the quantitative evaluation of model soil water. Soil moisture data from various reanalyses depend primarily on precipitation, surface energy, and water balance, and the land model, none of which is likely comparable to that from CCM3 coupled with CLM, LSM, or BATS. Therefore, it would be difficult to obtain anything useful from the comparison of soil moisture from reanalyses and models. Quantitative comparison is also inherently difficult among models that treat soil water, root distribution, vegetation transpiration, and bare soil evaporation differently, because the same volumetric soil water would result in different vegetation transpiration and bare soil evaporation in different models. Soil texture data in CLM are from LSM and vegetation biophysics in CLM is primarily based on LSM. However, CLM and LSM have different root distribution data and use different formulations for the bare soil evaporation. Therefore, quantitative comparison of volumetric soil water from CLM and LSM is not entirely appropriate. Overall, CLM and LSM give a similar spatial distribution of soil water but CLM has slightly drier soils. Over some areas with large soil water differences (e.g., India), the negative (or positive) differences in soil water are also consistent with the positive (or negative) differences in temperature (in Fig. 5) and negative (or positive) differences in precipitation (in Fig. 8).

### c. Interpretation of the results

Land surface processes are primarily driven by precipitation and solar radiation; they also feedback to the atmosphere through surface energy, water, and momentum fluxes. Because of this interactive nature, it is a challenging (if not impossible) task to provide exact reasons for the improvement of CLM in the simulation of surface air temperature, runoff, and snow mass. Here we attempt to at least give some possible explanations.

Runoff in CLM is generated by a saturated area runoff component that responds rapidly to water added to the surface (from precipitation or snowmelt) and a base-flow term that also will generally be largest near the time of maximum soil moisture but will decay on a longer timescale. In contrast, LSM appears to contain too strong a base-flow term and its surface runoff term does not respond rapidly to water added to the surface from snowmelt. This explains the improved simulation of the annual cycle of runoff in CLM (Fig. 12).

During the snow accumulation stage (October–March) over Northern Hemisphere land, the CLM- and LSM-simulated precipitation is nearly the same (Fig. 9). Precipitation is also taken as snow in both CLM and LSM when the surface air temperature is below 2.2°C. Compared with the one-layer snow parameterization in LSM, the more realistic multilayer snow parameterization in CLM reduces the thermal coupling of snow to the underlying soil, hence allowing a colder snow surface with less sublimation. The increased stability in the near-surface atmosphere reduces the turbulent mixing, which also contributes to reduced sublimation. Indeed, offline tests show that CLM can realistically simulate the snow accumulation and snowmelt (Dai et al. 2002, manuscript submitted to *Bull. Amer. Meteor. Soc.*). These are probably the main reasons for the more realistic (i.e., faster) accumulation of snow mass in CLM (Fig. 11).

In the Sahara and the Arabian Desert, CLM significantly reduces the cold bias in LSM (Figs. 5 and 6) largely because lower soil albedos are prescribed based on data analysis in Xue et al. (2001). For instance, the net solar radiation at surface in July in CLM is 17 W $m^{-2}$ higher than that in LSM. Over Australia, which is dominated by open shrublands (Fig. 2), CLM also reduces the cold bias in LSM (Fig. 5). Here the net radiative flux is unchanged, but its partitioning into the sensible and latent heat fluxes is different, with a higher sensible and lower latent heat fluxes in CLM, and hence the higher surface air temperature in CLM. In July, CLM's reduction of the cold bias over Eurasia in LSM (Fig. 5) results from the increase of net solar radiation (and net solar and longwave radiation) as well as the higher sensible and lower latent heat fluxes associated with the lower soil moisture in CLM (Fig. 13). In winter, CLM's reduction of the warm bias over Eurasia in LSM (Fig. 5) is accompanied by a decrease of the (negative) sensible heat flux. Evidently, changes in the micrometeorology reduce the coupling of the surface to the overlying warmer atmosphere. The multilayer snow parameterization reduces the coupling of snow to the underlying soil, which also contributes to the colder surface. Although LSM has a higher snow fraction and hence higher albedo (Xue et al. 2001), the net solar radiation is nearly the same in CLM versus LSM, because the downward solar radiation is small in winter over Eurasia. In addition to the overall improvement in model physics in CLM, the use of the satellite data may also contribute to the reduction of temperature bias over Eurasia. For instance, Buermann et al. (2001) showed that replacing the default leaf-area index data in LSM by the satellite LAI (Myneni et al. 1997) would slightly reduce the cold (or warm) bias over part of the Eurasia in July (or January). At the same time, however, this would further increase the cold bias over northern Asia in July. Similarly, when the fractional vegetation cover of each plant functional type in LSM was replaced by the satellite data (Zeng et al. 2000), we also found that the cold (or warm) bias in July (or January) in LSM was slightly reduced over part of the Eurasia.

To understand the reasons for the CLM- and LSM-simulated surface air temperature differences over global land, and to illustrate the overall impact of land processes on the surface and top-of-the-atmosphere budgets, Figs. 14 and 15 show the various quantities averaged over global land that are related to the energy and water budgets at surface and top of the atmosphere (TOA). Figure 14 shows that the partitioning of the net surface radiative flux is quite different between CLM and LSM, with a higher sensible heat flux, lower latent heat flux, and hence a much higher Bowen ratio (i.e., the ratio of sensible over latent heat fluxes) from CLM. Note that nearly the same photosynthesis-conductance parameterization is used in CLM and LSM. However, the partitioning of the net radiation into the sensible and latent heat fluxes can still be significantly different, because the latent heat flux along with runoff is also constrained by the total precipitation, which differs little between the two models. The more realistic runoff parameterization and typically higher runoff in CLM need to be balanced by a lower latent heat flux, which, in turn, needs to be balanced by a higher sensible heat flux. Note that 1 mm $day^{-1}$ of precipitation or runoff is equivalent to about 29 W $m^{-2}$ of heat fluxes. Other differences in model physics (e.g., the turbulent parameterization) can also affect the partitioning of the net radiation, but whether they increase or decrease the Bowen ratio is difficult to assess. The higher surface air temperature in CLM (particularly in boreal summer) primarily results from this increase of sensible heat flux (e.g., by about 10 W $m^{-2}$ in July) and the corresponding decrease of latent heat flux. The higher surface air temperature in CLM leads to a higher net (upward) longwave radiation, which largely compensates the higher solar flux in CLM (e.g., by 7 W $m^{-2}$ in July), so the net surface radiative flux differs little (e.g., the difference is only 0.6 W $m^{-2}$ between CLM and LSM in July). The higher surface air temperature in July in CLM is also consistent with the lower precipitation and drier soil (e.g., Fig. 13). In February and March, CLM and LSM still give nearly the same surface air temperature, even though precipitation is higher in CLM. The average snow water equivalent depth over global land is primarily determined by snow depth over Antarctic and Greenland, and the large

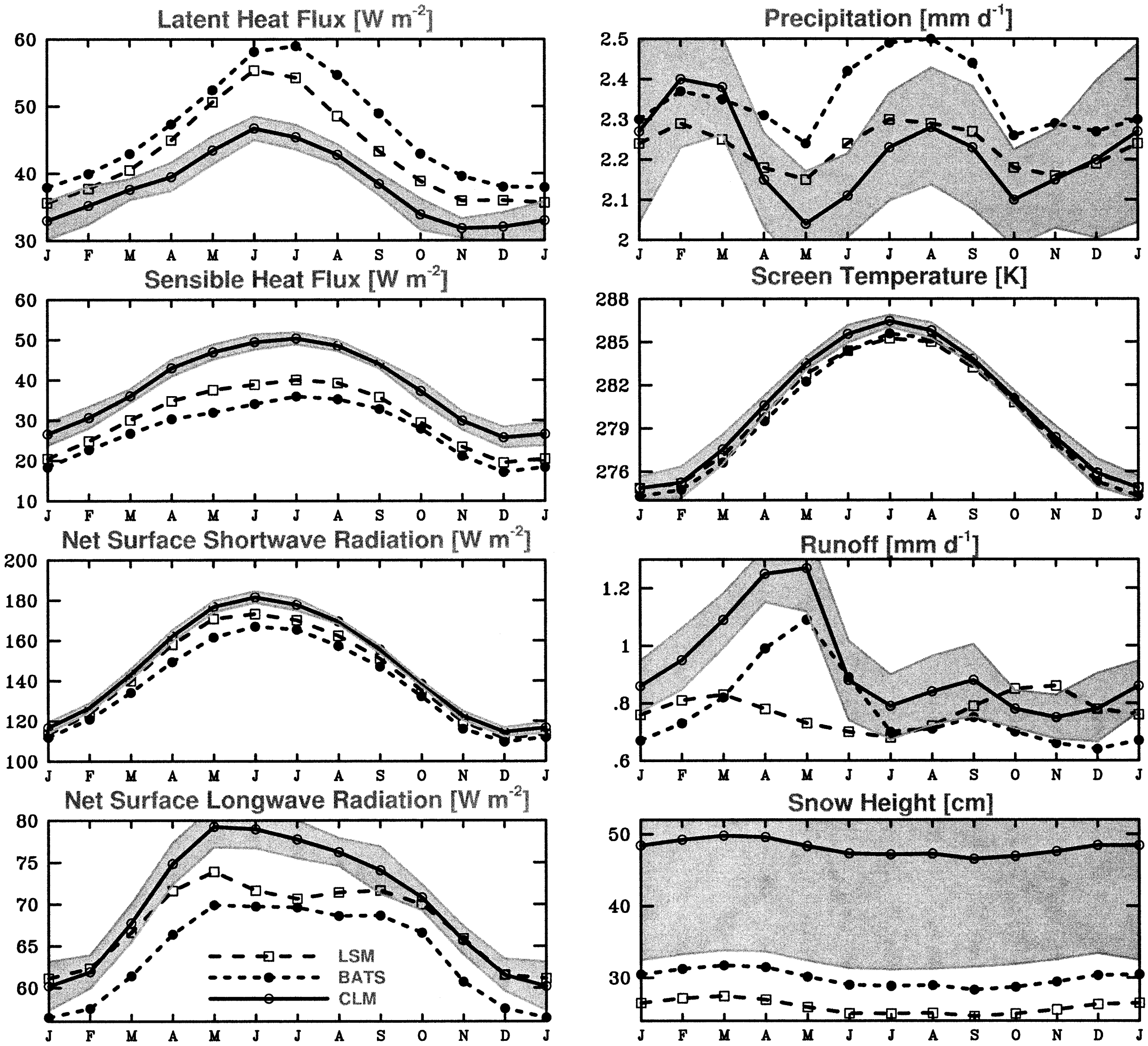


FIG. 14. Various components related to surface energy and water cycles over global land as simulated by CCM3 coupled with CLM, LSM, and BATS. The shaded area in each panel indicates CLM's 95% confidence interval for the mean based on its 15-yr climatology.

difference over these two regions among models does not affect surface energy or water cycle.

Figure 15 shows that CLM yields a higher planetary boundary layer depth in agreement with its higher Bowen ratio as discussed earlier. The difference in sea level pressure or precipitable water between CLM and LSM is not significant at the 95% level. CLM also has a lower total cloud cover partly because of a drier atmosphere (particularly in boreal summer). The average clear-sky planetary albedo over land at the top of the atmosphere (TOA) in boreal summer (June–August) differ by less than 0.5% between CLM and LSM, partly because the same atmospheric model (CCM3) is used. However, due to the smaller total cloud cover in CLM, the total planetary albedo at TOA over land averaged in June–August is lower by 2% in CLM, which is statistically significant at the 95% level. Because of this smaller cloud cover and a higher surface air temperature in CLM, the TOA clear-sky and total outgoing longwave radiation fluxes are also higher over land (particularly in boreal summer).

## 5. Conclusions

The land boundary data, the coupling strategy, and the results of the Common Land Model (CLM) coupled with the NCAR CCM3 have been presented in this paper. The high-resolution vegetation data include the global 1-km fractional vegetation cover and International Geosphere–Biosphere Program (IGBP) land cover classification, both of which are pixel-dependent but season-independent (Zeng et al. 2000), and the global

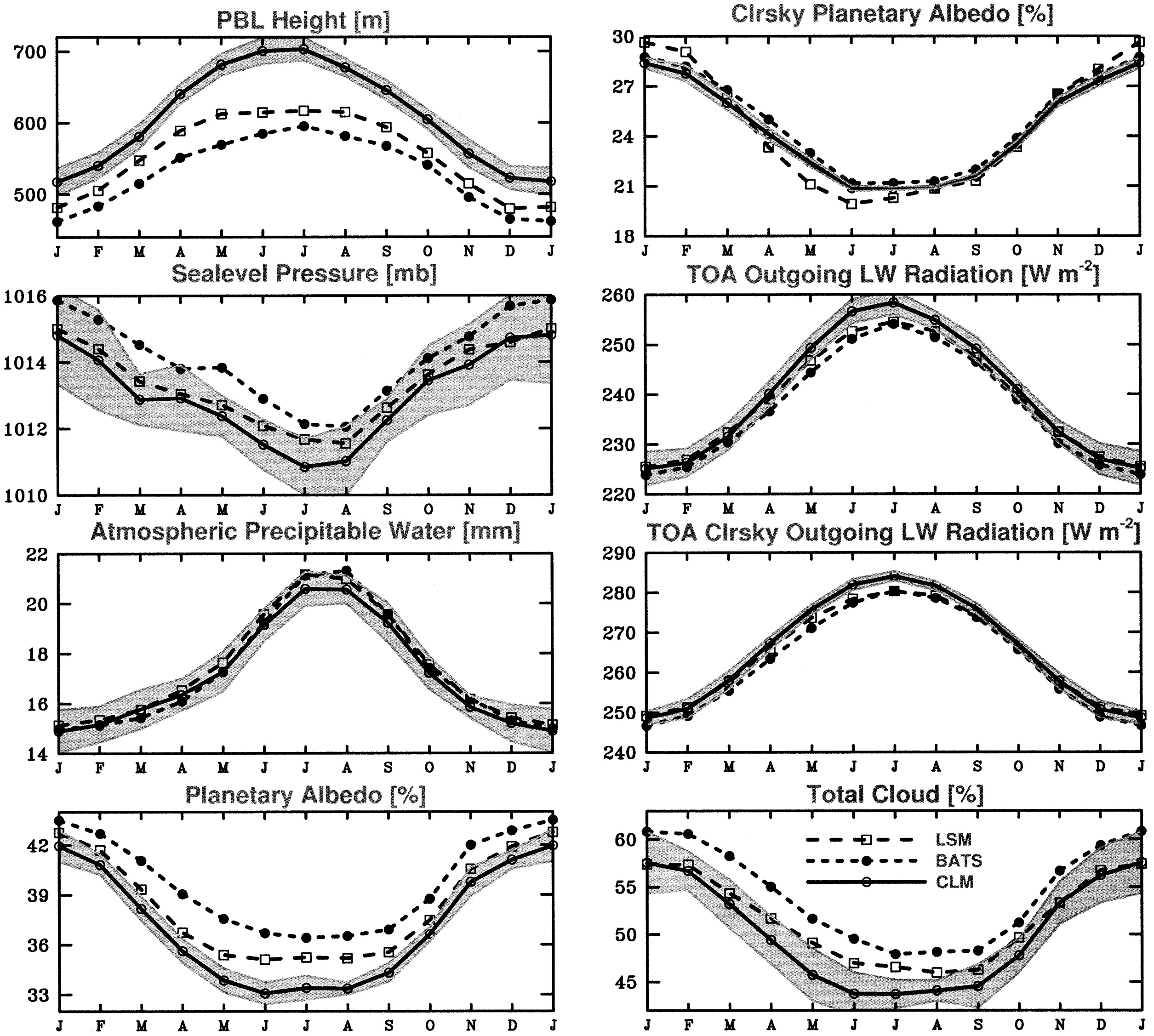


FIG. 15. Various components related to surface and TOA energy cycles over global land as simulated by CCM3 coupled with CLM, LSM, and BATS. The shaded area in each panel indicates CLM's 95% confidence interval for the mean based on its 15-yr climatology.

8-km green leaf-area index (LAI) data, which are both season- and pixel-dependent (Myneni et al. 1997). All these data were derived using the satellite AVHRR data. The stem and dead leaf-area index data were derived from the seasonal variation of LAI. The vegetation root distribution depends on vegetation type and was derived using data from global field surveys (Zeng et al. 1998b; Zeng 2001). The vegetation albedo (under dense canopy conditions) depends on vegetation type and was derived based on Dickinson et al. (1993) with modifications using remote sensing data in Strugnell et al. (2001) and Xue et al. (2001). The soil texture data are the same as those in the NCAR Land Surface Model (LSM) (Bonan 1996), while soil color data (related to soil albedo) are based on Dickinson et al. (1993) with modifications using remote sensing data in Strugnell et al. (2001) and Xue et al. (2001). Together, the above data represent a significant improvement over those currently used by regional or global models.

Using the above data, a 15-yr simulation of CCM3–CLM has been done from 1979 to 1993 with observed sea surface temperature and sea ice. The subgrid variability of vegetation is treated using the mosaic approach with each atmospheric model grid box containing up to five tiles (i.e., the first two dominant vegetation types, bare soil, lakes, and wetland). The CLM code results from a 3-yr joint effort among seven land modeling groups in the United States, and has been developed primarily based on the Biosphere–Atmosphere Transfer Scheme (Dickinson et al. 1993), LSM (Bonan 1996), and the multilayer snow model of Dai and Zeng (1996). The 15-yr averaged land surface climatology from CCM3–CLM has been compared with those from CCM3–LSM and CCM3–BATS, although most of the discussions were focused on CLM versus LSM comparison.

The simulation of land surface climate by global models has shown much improvement over the last decade as a result of both improved land surface treatments and atmospheric simulations. Consequently, errors in observational data and incompatibilities between model and observed variables now can introduce differences between model and observed variables that are not much smaller than the departures of the model fields from reality. Even given these caveats, the results reported here appear to indicate that the simulation of land climate in CCM3 has been substantially improved through use of the CLM. It is found that CLM significantly reduces the cold bias of surface air temperature in LSM (particularly in summer) and reduces the warm bias in LSM over some regions in winter. For instance, the cold bias of −3.2°C in LSM in July over global land is reduced to −2.0°C in CLM, while the warm bias of 4.2°C in LSM in January over western Siberia is reduced to 1.2°C in CLM. Another significant improvement from CLM is the simulation of the annual cycle of runoff. While LSM incorrectly reaches the peak runoff in October (rather than in June as given by observed data) over Northern Hemisphere land, CLM realistically simulates the overall annual cycle of runoff. CLM also simulates the snow mass better than LSM during the snow accumulation period, while their performance is similar during the snowmelt period. These improvements are primarily caused by the improved parameterizations in runoff, snow, and other processes (e.g., turbulence) in CLM. The new land boundary data (e.g., leaf-area index, fractional vegetation cover, albedo) also contribute to the improvement in surface air temperature simulation over some regions.

In agreement with the above results, CLM increases both the net surface solar and long-wave fluxes over global land. These increases, however, largely cancel each other, resulting in little change in net surface radiative flux between LSM and CLM. The partitioning of the net radiative flux is quite different in CLM and LSM with higher sensible heat fluxes and lower latent heat fluxes in CLM. Precipitation differs little between LSM and CLM, but its partitioning is different with lower latent heat fluxes and higher runoff in CLM.

The higher surface air temperature and lower latent heat fluxes in CLM are consistent with a drier atmosphere, higher planetary boundary layer depth, lower sea level pressure, lower total cloud cover, and higher outgoing longwave radiation at the top of the atmosphere (TOA) over global land. While the TOA clear-sky planetary albedo differs little between CLM and LSM in boreal summer, the TOA total planetary albedo over global land is reduced by about 2% in CLM in June–August, partly because of the reduced total cloud cover.

*Acknowledgments.* This work was supported by the NASA through its EOS IDS Program (429-81-22;428-81-22), Department of Energy under Grant DE-FG02-91ER61216, and NOAA under Grant NA06GP0569. P. Rao and A. Walker are thanked for helping process some of the data and preparing Fig. 1 (PR). R. DeFries, J. Qi, and R. B. Jackson are thanked for helpful discussions on various data issues. We also thank other CLM members (I. Baker, G. Bonan, M. Bosilovich, S. Denning, P. Dirmeyer, P. Houser, K. Oleson, A. Schlosser, and Z.-L. Yang) for stimulating discussions, and three anonymous reviewers for helpful comments.

## REFERENCES

Betts, A. K., and J. H. Ball, 1998: FIFE surface climate and site-average dataset: 1987–1989. *J. Atmos. Sci.,* **55,** 1091–1108.

——, J. H. Ball, A. C. M. Beljaars, M. J. Miller, and P. Viterbo, 1996: The land surface–atmosphere interaction: A review based on observational and global modeling perspectives. *J. Geophys. Res.,* **101,** 7209–7225.

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

——, 1998: The land surface climatology of the NCAR Land Surface Model coupled to the NCAR Community Climate Model. *J. Climate,* **11,** 1307–1326.

Buermann, W., J. Dong, X. Zeng, R. B. Myneni, and R. E. Dickinson, 2001: Evaluation of the utility of satellite-based vegetation leaf area index data for climate simulations. *J. Climate,* **14,** 3536–3550.

Canadell, J., R. B. Jackson, J. R. Ehleringer, H. A. Mooney, D. E. Sala, and E. D. Schulze, 1996: Maximum rooting depth of vegetation type at the global scale. *Oecologia,* **108,** 583–595.

Clapp, R. B., and G. M. Hornberger, 1978: Empirical equations for some soil hydraulic properties. *Water Resour. Res.,* **14,** 601–604.

Cosby, B. J., G. M. Hornberger, and T. R. Ginn, 1984: A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resour. Res.,* **20,** 682–690.

Crutcher, H., and R. Jenne, 1969: Monthly Northern and Southern Hemisphere climatologies for 1950 to 1964. [Available online at http://dss.ucar.edu/datasets/ds205.0/.]

Dai, Y., and Q. C. Zeng, 1996: A land surface model (IAP94) for climate studies. Part I: Formulation and validation in off-line experiments. *Adv. Atmos. Sci.,* **14,** 433–460.

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

——, M. Shaikh, R. Bryant, and L. Graumlich, 1998: Interactive canopies for a climate model. *J. Climate,* **11,** 2823–2836.

Fekete, B. M., C. J. Vorosmarty, and W. Grabs, 2000: UNH/GRDC composite runoff fields version 1.0. [Available online from http://www.watsys.unh.edu.]

Foster, D. J., Jr., and R. D. Davy, 1988: Global snow depth climatology. USAF Publ. USAFETAC/TN-88/006, 48 pp.

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

Gutman, G., and A. Ignatov, 1998: The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. *Int. J. Remote Sens.,* **19,** 1533–1543.

Jackson, R. B., J. Canadell, J. R. Ehleringer, H. A. Mooney, O. E. Sala, and E. D. Schulze, 1996: A global analysis of root distributions for terrestrial biomes. *Oecologia,* **108,** 389–411.

Jaeger, L., 1983: Monthly and areal patterns of mean global precipitation. *Variations in the Global Water Budget,* A. Street-Perrot, M. Beran, and R. Rateliffe, Eds., D. Reidel, 129–140.

James, M. E., and S. N. V. Kalluri, 1994: The Pathfinder AVHRR land data set—An improved coarse resolution data set for terrestrial monitoring. *Int. J. Remote Sens.,* **15,** 3347–3363.

Jordan, R., 1991: A one-dimensional temperature model for a snow cover. U.S. Army Corps of Engineering, Cold Regions Research and Engineering Laboratory Special Report 91-16, 49 pp. [Available from CRREL, 72 Lyme Rd., Hanover, NH 03755-1290.]

Kiehl, J. T., J. J. Hack, G. B. Bonan, B. A. Boville, D. L. Williamson, and P. J. Rasch, 1998: The National Center for Atmospheric Research Community Climate Model: CCM3. *J. Climate,* **11,** 1131–1149.

Koster, R. D., M. J. Suarez, and M. Heiser, 2000: Variance and predictability of precipitation at seasonal-to-interannual timescales. *J. Hydrometeor.,* **1,** 26–46.

Legates, D. R., and C. J. Willmott, 1990a: Mean seasonal and spatial variability in global surface air temperature. *Theor. Appl. Climatol.,* **41,** 11–21.

——, and ——, 1990b: Mean seasonal and spatial variability in gauge-corrected, global precipitation. *Int. J. Climatol.,* **10,** 111–127.

Loveland, T. R., Z. L. Zhu, D. O. Ohlen, J. F. Brown, B. C. Reed, and L. Yang, 1999: The IGBP-DIS global 1-km land cover data set, DISCover: A project overview. *Photogramm. Eng. Remote Sens.,* **9,** 1013–1020.

Manabe, S., 1969: Climate and ocean circulation. I. The atmospheric circulation and the hydrology of the earth's surface. *Mon. Wea. Rev.,* **97,** 739–744.

Myneni, R. B., R. R. Nemani, and S. W. Running, 1997: Estimation of global leaf area index and absorbed par using radiative transfer models. *IEEE Trans. Geosci. Remote Sens.,* **35,** 1380–1393.

Sellers, P. J., S. O. Los, C. J. Tucker, C. O. Justice, D. A. Dazlich, G. J. Collatz, and D. A. Randall, 1996: A revised land surface parameterization (SiB2) for atmospheric GCMs. Part II: The generation of global fields of terrestrial biophysical parameters from satellite data. *J. Climate,* **9,** 706–737.

——, and Coauthors, 1997: Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science,* **275,** 502–509.

Shea, D. J., 1986: Climatological atlas: 1950–1979 surface air temperature, precipitation, sea-level pressure, and sea-surface temperature (45S-90N). NCAR Tech. Note NCAR/TN-269+STR. [Available online at http://dss.ucar.edu/datasets/ds290.0/.]

Strugnell, N. C., W. Lucht, and C. Schaaf, 2001: A global albedo data set derived from AVHRR data for use in climate simulations. *Geophys. Res. Lett.,* **28,** 191–194.

Willmott, C. J., and K. Matsuura, 2000: Terrestrial air temperature and precipitation: Monthly and annual climatologies (version 3.01). [Available online at http://climate.geog.udel.edu/~climate/html_pages/download.html.]

——, ——, and D. L. Legates, 1998: Global air temperature and precipitation: Regridded monthly and annual climatologies (Version 2.01). [Available online at http://climate.geog.udel.edu/~climate/html_pages/download.html.]

Xie, P. P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.,* **78,** 2539–2558.

Xue, W., R. E. Dickinson, A. Hahmann, Z.-L. Yang, X. Zeng, K. J. Schaudt, C. B. Schaaf, and N. Strugnell, 2001: Comparison of the albedo computed by land surface models and their evaluation against remotely sensed data. *J. Geophys. Res.,* **106,** 20 687–20 698.

Zeng, X., 2001: Global vegetation root distribution for land modeling. *J. Hydrometeor.,* **2,** 525–530.

——, and R. E. Dickinson, 1998: Effect of surface sublayer on surface skin temperature and fluxes. *J. Climate,* **11,** 537–550.

——, M. Zhao, and R. E. Dickinson, 1998a: Intercomparison of bulk aerodynamic algorithms for the computation of sea surface fluxes using the TOGA COARE and TAO data. *J. Climate,* **11,** 2628–2644.

——, Y. Dai, R. E. Dickinson, and M. Shaikh, 1998b: The role of root distribution for land climate simulation. *Geophys. Res. Lett.,* **25,** 4533–4536.

——, R. E. Dickinson, A. Walker, M. Shaikh, R. S. DeFries, and J. Qi, 2000: Derivation and evaluation of global 1-km fractional vegetation cover data for land modeling. *J. Appl. Meteor.,* **39,** 826–839.