Improvements to the Community Land Model and their impact on the hydrological cycle

K. W. Oleson,[1] G.-Y. Niu,[2] Z.-L. Yang,[2] D. M. Lawrence,[1] P. E. Thornton,[1] P. J. Lawrence,[3] R. Stöckli,[4,5,6] R. E. Dickinson,[7] G. B. Bonan,[1] S. Levis,[1] A. Dai,[1] and T. Qian[1]

[1] The Community Land Model version 3 (CLM3) is the land component of the Community Climate System Model (CCSM). CLM3 has energy and water biases resulting from deficiencies in some of its canopy and soil parameterizations related to hydrological processes. Recent research by the community that utilizes CLM3 and the family of CCSM models has indicated several promising approaches to alleviating these biases. This paper describes the implementation of a selected set of these parameterizations and their effects on the simulated hydrological cycle. The modifications consist of surface data sets based on Moderate Resolution Imaging Spectroradiometer products, new parameterizations for canopy integration, canopy interception, frozen soil, soil water availability, and soil evaporation, a TOPMODEL-based model for surface and subsurface runoff, a groundwater model for determining water table depth, and the introduction of a factor to simulate nitrogen limitation on plant productivity. The results from a set of offline simulations were compared with observed data for runoff, river discharge, soil moisture, and total water storage to assess the performance of the new model (referred to as CLM3.5). CLM3.5 exhibits significant improvements in its partitioning of global evapotranspiration (ET) which result in wetter soils, less plant water stress, increased transpiration and photosynthesis, and an improved annual cycle of total water storage. Phase and amplitude of the runoff annual cycle is generally improved. Dramatic improvements in vegetation biogeography result when CLM3.5 is coupled to a dynamic global vegetation model. Lower than observed soil moisture variability in the rooting zone is noted as a remaining deficiency.

Citation: Oleson, K. W., et al. (2008), Improvements to the Community Land Model and their impact on the hydrological cycle, *J. Geophys. Res.*, *113*, G01021, doi:10.1029/2007JG000563.

1. Introduction

[2] The circulation of water through the Earth climate system helps sustain life. The hydrological cycle is linked to the energy cycle and to biogeochemical processes including the carbon cycle. Simulating the various processes that interact to form the hydrological cycle is a daunting task for climate models. In particular, over land, interactions between precipitation and the vegetation/soil system determine the partitioning of water into various storage reservoirs and the subsequent release of water vapor to the atmosphere. Successful simulation of these interactions by the land surface component of a climate model requires detailed representation of processes such as interception, throughfall, canopy drip, snow accumulation and ablation, infiltration, surface and subsurface runoff, soil moisture, and the partitioning of evapotranspiration (ET) between canopy evaporation, transpiration, and soil evaporation. Depending on the capabilities of the model, the water cycle components may interact with and affect the simulation of biogeochemical processes such as the carbon and nitrogen cycles, dust and trace gas emissions, water and carbon isotopes, and vegetation dynamics.

[3] The Community Land Model version 3 (CLM3) represents land surface processes within the context of a global climate simulation [*Oleson et al.*, 2004]. *Dickinson et al.* [2006] described the climate statistics of CLM3 when coupled to the Community Climate System Model (CCSM3) [*Collins et al.*, 2006]. *Hack et al.* [2006] provided an analysis of selected features of the land hydrological

[1]Climate and Global Dynamics Division, National Center for Atmospheric Research, Boulder, Colorado, USA.
[2]Department of Geological Sciences, University of Texas at Austin, Austin, Texas, USA.
[3]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA.
[4]Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado, USA.
[5]Institute for Atmospheric and Climate Sciences, ETH Zurich, Zurich, Switzerland.
[6]NASA Earth Observatory, Goddard Space Flight Center, Greenbelt, Maryland, USA.
[7]Department of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia, USA.

cycle. *Bonan and Levis* [2006] evaluated global plant biogeography and net primary production from CLM3 when coupled to a dynamic global vegetation model (DGVM). *Lawrence et al.* [2007] examined the impact of changes in CLM3 hydrological parameterizations on partitioning of ET and its effect on the timescales of ET response to precipitation events, interseasonal soil moisture storage, soil moisture memory, and land-atmosphere coupling. *Qian et al.* [2006] evaluated CLM3's performance in simulating soil moisture content, runoff, and river discharge when forced by observed precipitation, temperature and other atmospheric data. Although the simulation of land surface climate by CLM3 is in many ways adequate [*Dickinson et al.*, 2006], most of the unsatisfactory aspects of the simulated climate described in these studies can be traced directly to a deficient simulation of the hydrological cycle.

[4] A poor simulation of the hydrological cycle in the Amazon basin is in part due to insufficient precipitation from the atmospheric model but is exacerbated by unrealistic partitioning of ET and deficiencies in runoff and soil water storage. This is also evident in off-line simulations forced with observed precipitation [*Qian et al.*, 2006]. The simulated present-day climate is biased warm and dry with lower runoff than observed [*Dickinson et al.*, 2006; *Lawrence et al.*, 2007; *Hack et al.*, 2006]. In particular, these studies indicate that the simulated ET is dominated by soil and canopy evaporation instead of by transpiration as observed. The deficiencies result in a poor simulation of vegetation biogeography with much less cover of broadleaf evergreen trees and more of deciduous trees than observed [*Bonan and Levis*, 2006]. On a global scale, forest cover is underestimated compared to observations in favor of grasses because of dry soils. *Lawrence and Chase* [2007] have noted that because of the unrealistic partitioning of ET, improved surface data sets of leaf and stem area index and plant functional type were unable to rectify temperature and precipitation biases in the coupled modeling system. Other hydrology-related problems in the model include low gross primary production (GPP) [*Bonan and Levis*, 2006] and poor simulation of the magnitude and seasonality of runoff and soil water storage in regions with frozen soil [*Niu and Yang*, 2006].

[5] As a community model, CLM has benefited from a number of scientists willing to scrutinize its scientific contents, offer constructive criticism, and improve its performance. Several new parameterizations designed to address the specific deficiencies in CLM3 have been proposed [*Niu et al.*, 2005; *Niu and Yang*, 2006; *Niu et al.*, 2007; *Thornton and Zimmermann*, 2007; *Lawrence and Chase*, 2007; *Lawrence et al.*, 2007]. Validation and sensitivity testing of the individual parameterizations have been addressed by the respective authors. While these individual parameterizations have clearly been shown to be beneficial in alleviating specific biases in the model, how they might interact with each other and the net effect on the simulation of the hydrological cycle have not previously been examined. This paper describes the implementation of these parameterizations and reports on the aggregated effects on the hydrology of CLM on a global scale. We show that in general the new parameterizations result in a more realistic depiction of the hydrologic cycle. We also demonstrate that the improved hydrology translates into better simulation of global GPP and present-day vegetation biogeography. On the other hand, the analysis of the new model presented here

is somewhat limited by the lack of observed data at global scales (e.g., surface energy partitioning). Thus, *Stöckli et al.* [2008] examine the performance of the new model at local scales by making use of a network of long-term ground-based ecosystem observations (FLUXNET [*Baldocchi et al.*, 2001]). These flux tower sites span a wide range of ecosystems and climate zones and are used to examine in more detail the model parameterizations of heat, water and carbon exchanges.

2. Materials and Methods

2.1. CLM3

[6] CLM3 is the land surface component of CCSM3, a community-developed global climate system model for application to studies of interannual and interdecadal variability, paleoclimate regimes, and projections of future climate change [*Collins et al.*, 2006]. Biophysical processes simulated by CLM3 include solar and longwave radiation interactions with vegetation canopy and soil, momentum and turbulent fluxes from canopy and soil, heat transfer in soil and snow, hydrology of canopy, soil, and snow, and stomatal physiology and photosynthesis. Vegetation cover in CLM3 is represented by several plant functional types (PFTs) which differ in their ecological and hydrological characteristics. Texture types determine thermal and hydrologic properties of soil. A detailed description of how processes of CLM3 are parameterized can be found in *Oleson et al.* [2004]. Specific detail on the parameterizations relevant to this paper is provided in the next section.

2.2. Summary of Model Improvements

[7] We implemented new surface data sets and parameterizations within CLM3.5. The modifications consist of surface data sets based on Moderate Resolution Imaging Spectroradiometer (MODIS) products [*Lawrence and Chase*, 2007], an improved canopy integration scheme [*Thornton and Zimmermann*, 2007], scaling of canopy interception [*Lawrence et al.*, 2007], a simple TOPMODEL-based model for surface and subsurface runoff [*Niu et al.*, 2005], a simple groundwater model for determining water table depth [*Niu et al.*, 2007], and a new frozen soil scheme [*Niu and Yang*, 2006]. In this paper, we also describe four additional modifications. Three of these, an improved description of soil water availability to plants, a resistance term to reduce excessive soil evaporation, and the introduction of a factor to simulate nitrogen limitation on plant productivity, can be categorized as new or improved parameterizations from the perspective of CLM3. The other may be categorized as fixing an algorithmically defective existing parameterization. In this section, we provide a brief overview of these modifications and summarize their individual effects on simulated hydrology as demonstrated by the respective authors. More detailed descriptions of the parameterizations and assessments of their performance can be found in the cited papers. However, we provide full details in Appendices A−G in order to fully document the new aspects of the model as compared to CLM3. The new model has been designated as CLM3.5.

2.2.1. Surface Data Sets

[8] Surface data sets of PFT and leaf and stem area index (LAI and SAI) in CLM3 are based on one year of data from

the Advanced Very High Resolution Radiometer (AVHRR) [*Bonan et al.*, 2002]. *Lawrence and Chase* [2007] developed new surface data sets for CLM3.5 that better reproduce the physical properties described in the multiyear MODIS land surface data products compared to the CLM3 representation, i.e. they created new PFT, glacier, and wetland maps, and LAI, SAI and soil color (which determines soil albedo) data sets and documented some improvements in simulated surface albedo, near-surface temperature, and precipitation. As noted above, however, the hydrologic deficiencies in the model limited the effectiveness of these improvements, the issue that we address in this paper. We have replaced the 0.5° resolution data sets used in CLM3 with these new data sets.

2.2.2. Canopy Integration

[9] Although the vegetation canopy in CLM3 is divided into shaded and sunlit fractions, all the direct and diffuse canopy intercepted radiation is assigned to the sunlit canopy fraction. *Thornton and Zimmermann* [2007] combined a logical framework relating the structural and functional characteristics of a vegetation canopy and a true two-leaf canopy model to produce a canopy integration scheme for land surface models. The framework posits a linear relationship between the ratio of leaf area to leaf mass (specific leaf area) and overlying leaf area index within the canopy. An inconsistency in the treatment of canopy radiation in CLM3 was also corrected. Incorporation of the new scheme in CLM3.5 resulted in significant increases in simulated global GPP. In separate simulations performed by us, we observed that the large increase in production was accompanied by unrealistic depletion in soil moisture in some regions because of increases in transpiration rates (not shown). In some regions the improvement in GPP was limited by the dry soils of CLM3, further motivating the development of a more realistic depiction of hydrology. Here, we implemented the canopy integration scheme in diagnostic canopy mode (using the remotely sensed LAI climatology from *Lawrence and Chase* [2007]) exactly as described by *Thornton and Zimmermann* [2007].

2.2.3. Canopy Interception

[10] The canopy in CLM3 intercepts too much water [*Hack et al.*, 2006] because of its lack of scaling of interception from point to grid cell [*Lawrence et al.*, 2007] (Appendix A). The excessively wet leaves lower transpiration rates. Only the dry fraction of the canopy can transpire and atmospheric evaporative demand is largely met by the evaporation of the intercepted water. A factor is implemented that scales the parameterization of interception from point to grid cell [*Lawrence et al.*, 2007] (Appendix A), and so results in lower canopy interception rates and increases in the amount of water reaching the soil surface and runoff. Consequently, the partitioning of ET is improved [*Lawrence et al.*, 2007].

2.2.4. Surface and Subsurface Runoff

[11] The runoff scheme in CLM3 combined concepts from the TOPMODEL [*Beven and Kirkby*, 1979] and BATS [*Dickinson et al.*, 1993] parameterizations. *Niu et al.* [2005] showed that it overestimates the runoff peaks and underestimates runoff in recession periods, resulting in a high ratio of surface to total runoff. They introduced a simple TOPMODEL-based runoff scheme (SIMTOP) that mitigated several problems associated with implementing

the TOPMODEL approach within a climate model. A key concept underlying their approach is that of fractional saturated area, which is determined by the topographic characteristics and soil moisture state of a grid cell. The topographic data is simplified to a single topographic parameter, the potential or maximum fractional saturated area, which is determined from 1 km resolution Digital Elevation Model (DEM) data. Surface runoff is parameterized in terms of the saturated fraction and an exponential function of water table depth. The scheme also includes the mechanism of infiltration excess which can generate additional surface runoff. Subsurface runoff is a product of an exponential function of the water table depth and a single coefficient for maximum subsurface runoff. *Niu et al.* [2005] demonstrated that modeling of runoff both for a small watershed and globally was much improved using SIMTOP compared to CLM3. We implemented SIMTOP in CLM3.5 as described in Appendix B.

2.2.5. Groundwater and Water Table Depth

[12] The assumptions in the original SIMTOP [*Niu et al.*, 2005] that were made to derive the water table depth restrict its applicability to regions where the water table is relatively shallow and to times when the water table is in approximate equilibrium with the model soil moisture. A simple lumped aquifer model was suggested by *Niu et al.* [2005] as a way to extend the SIMTOP approach to cases when the water table is deeper than the bottom of the model soil column. Furthermore, groundwater influences soil moisture and runoff generation and hence surface energy and water balances, making it desirable to include a groundwater component in land surface models. A simple groundwater model (SIMGM) was developed by *Niu et al.* [2007] to address these issues. The model represents groundwater recharge and discharge processes through a dynamic coupling between the bottom soil layer and an unconfined aquifer. The aquifer is added as a single integration element below the soil column (Figure C1). *Niu et al.* [2007] found that the modeled water storage anomaly compared favorably to the water storage anomaly estimated by the Gravity Recovery And Climate Experiment (GRACE) satellites for several river basins. *Gulden et al.* [2007] demonstrated that although a deeper model soil column and SIMGM were equally adept at simulating changes in water storage, explicit representation of groundwater reduced the sensitivity of simulated water storage changes to inappropriate parameter choices. SIMGM is implemented as described in Appendix C.

2.2.6. Frozen Soil

[13] The significant improvements in the magnitude and timing of runoff in tropical and arid regions from SIMTOP, as shown by *Niu et al.* [2005], were not apparent in arctic and boreal regions. *Niu and Yang* [2006] demonstrated that in these regions CLM3 soil has low permeability to water as a result of its treatment of frozen soil, which results in larger and earlier springtime runoff peaks than observed. They showed that introduction of the concepts of supercooled soil water and fractional impermeable area into CLM3 and a parameterization of soil hydraulic properties as a function of impermeable area would increase infiltration rates and improve the simulation of runoff in cold-region river basins of various spatial scales. In other similar experiments with CLM3, *Decker and Zeng* [2006] and *Yi et al.* [2006] showed

improvements in their simulations by accounting for supercooled soil water. The parameterizations described by *Niu and Yang* [2006] were implemented as described in Appendix D.

2.2.7. Soil Water Availability

[14] Plant water stress in CLM3 is linked to root distribution and soil matric potential which serves as a surrogate for negative leaf water potential. Root distribution is semi-unique for each PFT [*Oleson et al.*, 2004]. However, both the matric potential at which the initial reduction in stomatal conductance occurs (ψ_{open}) and the potential at which final reduction occurs (ψ_{close}) (leaf desiccation) are prescribed as constants for all PFTs ($\psi_{close} = -1.5 \times 10^5$ mm, $\psi_{open} = \psi_{sat}$ where ψ_{sat} is saturated matric potential, which varies by soil texture but not PFT). This is in contrast to numerous field studies that show that PFTs have unique values of ψ_{open} and ψ_{close} (e.g., as summarized by *White et al.* [2000]). Furthermore, since $\psi_{open} = \psi_{sat}$ in CLM3, plant water stress begins to occur immediately at soil moisture levels less than saturation. We implemented a parameterization for plant water stress that is functionally similar to that in CLM3 but allows for PFT variability in ψ_{open} and ψ_{close} using values from *White et al.* [2000] which lowers the soil moisture levels at which stress begins to occur (Appendix E). The new parameterization results in increased soil water availability for plants.

[15] CLM3 only supplies water to plants in soil layers with a temperature greater than the freezing temperature of fresh water (273.15K). However, significant amounts of liquid water may coexist with ice at freezing temperature and with the introduction of the concept of supercooled soil water, liquid water can exist at temperatures below freezing. The dependence of plant water stress on temperature has been removed in the new formulation (Appendix E), resulting in more realistic LAI when CLM is coupled to a prognostic treatment of carbon and nitrogen cycle dynamics (CLM-CN [*Thornton and Zimmermann*, 2007; *Thornton et al.*, 2007]).

2.2.8. Soil Evaporation

[16] *Lawrence et al.* [2007] found that even after implementing alterations to CLM3 to improve ET partitioning, soil evaporation was still an unreasonably large fraction of total ET. Similarly, preliminary simulations with the model changes discussed to this point yielded improved ET partitioning. However, the ratio of soil evaporation to total ET with these changes is still significantly larger than other model-based estimates of this fraction (e.g., as compared to the second Global Soil Wetness Project (GSWP-2) multi-model ensemble [*Dirmeyer et al.*, 2006] or to *Choudhury et al.* [1998]). *Lawrence et al.* [2007] showed that soil evaporation in their CLM3 experiments was reduced by altering two parameters in the formulation for the turbulent transfer coefficient between the soil and the canopy air, but that sensible heat flux was also reduced such that soil temperatures increased. Further testing of this approach by us in the context of land cover change experiments revealed that surface soil temperatures were unrealistically sensitive to changes in leaf and stem area (not shown) and that in certain regions, the air temperature response to changes in land cover types was largely controlled by this behavior. Hence, we have retained the turbulent transfer coefficient as formulated in CLM3 but added a soil resistance term that depends on soil moisture and thus affects only the soil latent heat flux. Justification and details of this parameterization are provided in Appendix F. *Stöckli et al.* [2008] demonstrates that this approach reduces evaporation from the soil and improves the simulation of turbulent surface fluxes.

2.2.9. Other Modifications

[17] Concurrent with development of the biophysical aspects of CLM3 discussed above have been extensive efforts to introduce the effects of biogeochemistry into the model. In particular, the option to include a prognostic treatment of carbon and nitrogen cycle dynamics has been implemented (CLM-CN). This inclusion of the carbon/nitrogen cycle in conjunction with most of the changes described above results in reasonable prognostic simulations of leaf area index and plant productivity [*Thornton et al.*, 2007]. However, many applications do not require a full carbon/nitrogen cycle. In these cases, the model assimilates too much carbon because of the lack of nitrogen limitation on plant productivity as is clearly illustrated by *Stöckli et al.* [2008] for several FLUXNET sites. Consequently, a simple approach is adopted that applies a PFT-dependent foliage nitrogen limitation factor to limit the maximum rate of carboxylation attainable by each PFT. More details can be found in Appendix G.

[18] A dimensionless factor is prognostically determined in CLM3 that provides for a fractional reduction in snow albedo due to snow aging (assumed to represent increasing grain size and dirt, soot content). The implementation of this algorithm in the code was found to be deficient (Y.-J. Dai, personal communication) and has been corrected. Consequently, snow age is increased more thereby lowering snow albedo and resulting in earlier snowmelt in certain regions (not shown).

2.3. Simulations

[19] The hydrology modifications were tested in uncoupled mode where the land model is forced by observations of atmospheric variables. Two uncoupled global simulations at T42 resolution (∼2.8° longitude by 2.8° latitude) were performed, a control simulation as described by *Oleson et al.* [2004] (U_CON; CLM3) and a simulation that incorporated all of the changes discussed in section 2.2 (U_HYD; CLM3.5). The simulations were driven by a 57-year (1948–2004) atmospheric forcing data set from *Qian et al.* [2006]. Land surface state variables (e.g., soil temperature and moisture) for the U_CON simulation were initialized from the land surface state at the end of a 425 year CCSM control run. The model was then run through the forcing data period four times for an additional 228 years of spin up to ensure that the deep soil water was at equilibrium before conducting the final 57-year simulation. The method of *Niu et al.* [2007] was used to ensure stable soil moisture and water table depth for the U_HYD simulation by cycling through the forcing data period a total of 12 times (684 years). The long integration times are not strictly necessary to ensure stability of soil moisture, surface fluxes, and temperatures, but were conducted mainly to demonstrate that the model could survive the long integration periods frequently required to spin up the coupled model (e.g., to ensure stability of the ocean simulation). In fact, under repeated yearly forcing, the year to year change in annual mean latent and sensible heat flux is less than 0.1 W m^{-2} (criterion defined by *Yang et al.* [1995]) for more than 99% of the grid cells within 30 years

Table 1. Annual Averages of Global Land (Excluding Grid Cells With Glaciers, Wetlands, and Lakes)[a]

	U_CON	U_HYD
Precipitation, mm day^{-1}	2.09	2.18
Infiltration, mm day^{-1}	0.47	1.21
Evapotranspiration, mm day^{-1}	1.37	1.39
Transpiration, mm day^{-1}	0.16 (12%)	0.57 (41%)
Canopy Evaporation, mm day^{-1}	0.61 (44%)	0.28 (20%)
Ground Evaporation, mm day^{-1}	0.60 (44%)	0.54 (39%)
Total Runoff, mm day^{-1}	0.72	0.78
Surface Runoff, mm day^{-1}	0.38	0.14
Subsurface Runoff, mm day^{-1}	0.34	0.64
Volumetric Soil Moisture, mm^3 mm^{-3}	0.18	0.32
Soil Moisture Stress Function	0.37	0.65
Gross Primary Production, Pg	59	155
Net Radiation, W m^{-2}	78	75
Sensible Heat, W m^{-2}	38	34
Latent Heat, W m^{-2}	40	40

[a]Partitioning of evapotranspiration is shown in parenthesis. Soil moisture stress function varies from zero (full stomatal closure) to one (stomata fully open) as described in Appendix E.

after starting the model from arbitrary initial conditions (Appendix C).

[20] In a separate simulation, CLM3.5 was coupled to the Dynamic Global Vegetation Model (DGVM) [*Levis et al.*, 2004]. A 180-year simulation was run, forced by the data from *Qian et al.* [2006], to assess the effects of the hydrological improvements on the simulation of vegetation biogeography. The simulation was initialized with bare soil and spun up for 160 years.

3. Results

3.1. Global and Zonal Averages

[21] In terms of global averages, the new model alters the partitioning of ET, moistens the soils, and shifts the dominant form of runoff from surface to subsurface (Table 1). The multimodel ensemble results from GSWP-2 gives that the ranking of partitioning of global annual mean ET should be transpiration (48%), soil evaporation (36%), and canopy evaporation (16%) [*Dirmeyer et al.*, 2006]. The new model (U_HYD) is in much better agreement with GSWP2 models than was the old model (U_CON). Although total ET remains basically the same (U_HYD ET is 64% of precipitation which is close to an estimate of 60–65% by *Brutsaert* [1984]), canopy and soil evaporation are reduced and transpiration is increased in the new model. Consequently, infiltration of water into the soil is more than doubled because of the reduction in canopy evaporation, and so soils are wetter (larger volumetric soil moisture) and have less plant water stress. As a result, GPP increases from 59 to 155 Pg. The simulated vegetation biogeography improves as we demonstrate in section 3.8. Total runoff increases slightly but there is a large shift from the fast component of runoff (surface) to the slow component (subsurface). The mean available surface energy and the mean partitioning of this energy into the turbulent fluxes are essentially unchanged in the new model.

[22] Changes in the meridional distribution of ET and its partitioning are shown in Figure 1. The partitioning of ET in the old model in the deep tropics (10°S–10°N) is clearly unrealistic with canopy evaporation dominating. The new model flattens the sharp peak in ET that is produced in the old model in the deep tropics (10°S–10°N) through a reduction in ground and canopy evaporation whose sum exceeds the increase in transpiration. Transpiration is increased at nearly all latitudes but most significantly at 20°S–20°N. At latitudes north of 20°N, reductions in ground and canopy evaporation are offset by increases in transpiration such that total ET remains nearly the same. South of about 40°S, ET is reduced in the new model because of reductions in ground and canopy evaporation. Because vegetation is sparse in this region, there is little gain in transpiration to offset these decreases in evaporation.

3.2. Runoff

[23] *Niu et al.* [2005], *Niu and Yang* [2006], and *Niu et al.* [2007] evaluated several of the hydrological improvements implemented in CLM3.5 using the University of New Hampshire-Global Runoff Data Center (UNH-GRDC) 0.5° monthly climatological composite runoff fields [*Fekete et al.*, 2000, 2002]. These fields were generated by combining observed river discharge information with output from a climate-driven water balance model. For consistency with these studies we make use of this data set to evaluate the annual mean and annual cycle of runoff. However, here we only use the UNH-GRDC composite runoff data for grid cells that have valid UNH-GRDC observed runoff (observed runoff fields are included in the UNH-GRDC data set). We limit our comparisons of modeled runoff to UNH-GRDC composite runoff to grid cells and regions that have UNH-GRDC observed runoff data for at least 50% of the considered area. This minimizes what would simply be a model-to-model comparison in regions without observed runoff data (B. Fekete, personal communication). Note that although the composite runoff fields are influenced by the accuracy of the water balance model, the runoff maps are at least calibrated to gauged streamflow. In the next section, we also evaluate the new model using observed river discharge directly.

[24] Annual mean runoff over most of the land surface in U_CON is within 0.5 mm day^{-1} of the UNH-GRDC runoff (Figure 2a). Larger biases occur in the southern Alaskan coast, the Amazon basin, India and Southeast Asia. The spatial distribution of the biases remains the same in the U_HYD simulation but the biases are somewhat alleviated in the Amazon basin and India (Figure 2b). In the Amazon basin, the biases are corrected by a small decrease in surface runoff countered by a larger increase in subsurface runoff (Figures 2c and 2d). In India, the positive bias in the old model is corrected by a decrease in surface runoff. The depth to the water table in the model (Figure 2f) controls the response of subsurface runoff with shallow water table depths producing large subsurface runoff (e.g., as in the Amazon basin and Indonesia). Deep water tables are simulated in regions with relatively small annual mean precipitation rates (Figure 2e) such as the southwest U.S., Sahara and Arabian Peninsula, Tibetan Plateau, south-central Australia, and Mongolia. In northern high latitudes, the water table is relatively shallow despite low precipitation rates because evaporation rates are also small and so the soil remains wet.

[25] We divided the land surface into 14 regions to examine the simulation of the annual cycle of runoff on regional scales. The Taylor plots [*Taylor*, 2001] shown in Figure 3 indicate the phase (correlation) and amplitude
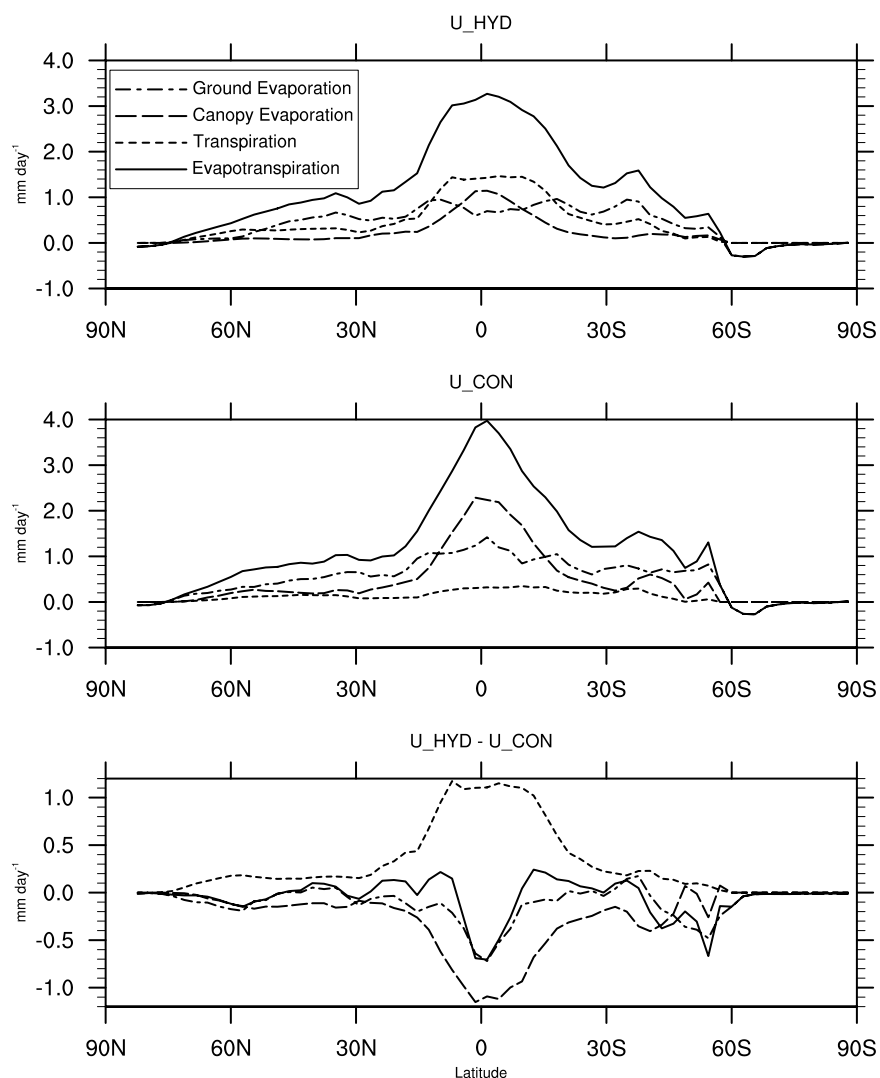
Figure 1. Zonally averaged annual mean evaporation (mm day^{-1}) for U_HYD (top), U_CON (middle), and U_HYD minus U_CON (bottom).

(normalized standardized deviations) of the simulated annual cycle of runoff compared to the UNH-GRDC composite runoff. Overall, the new parameterizations increase the agreement between model and UNH-GRDC runoff. The correlation between simulated and UNH-GRDC runoff increases by 0.2 in five regions in the new model compared to the old model. The normalized standardized deviation is improved by at least 20% in five regions in the new model. Applying these thresholds, the new model did not degrade the phase of runoff in any region and degraded the amplitude of runoff in only one region (region E in Figure 3).

[26] In polar and boreal regions, the old model generally overestimates the amplitude of the runoff annual cycle and the timing of runoff is poorly simulated in several regions (Figure 3a). The new model reduces the amplitude of runoff and improves the phase of runoff in all regions. Similar improvements in runoff were found by *Niu and Yang* [2006] for several cold region river basins and were attributed to the introduction of supercooled liquid water and fractional permeable area which combined to delay peak runoff. For example, in Western Siberia (region H), a large amount of snow melts in April and May (Figure 4). In the new model,

much of this water infiltrates into the soil and only a relatively small amount is transformed into surface runoff. The delay inherent in the infiltrated water raising the water table results in a later peak in subsurface runoff such that the annual cycle of total runoff agrees better with UNH-GRDC runoff.

[27] In general, although improved, the amplitude of the annual cycle of runoff is overcorrected in polar and boreal regions with the new model. For example, in Eastern Canada (region C), the timing of the spring runoff peak is quite good as indicated by the high correlation coefficient ($r = 0.93$), but the amplitude of the annual cycle is smaller than observed ($\sigma_m/\sigma_o = 0.72$). This discrepancy may indicate the global parameters applied to this region are inappropriate and require optimizing spatially explicit values for some of the global parameters.

[28] In the middle latitudes, smaller improvements in phase and/or amplitude are evident. In the Central U.S. (region G), the runoff phase is greatly improved and the amplitude remains well simulated. In tropical rainforest/savanna regions, the amplitude of the annual cycle is the primary aspect of runoff that is improved in the new model
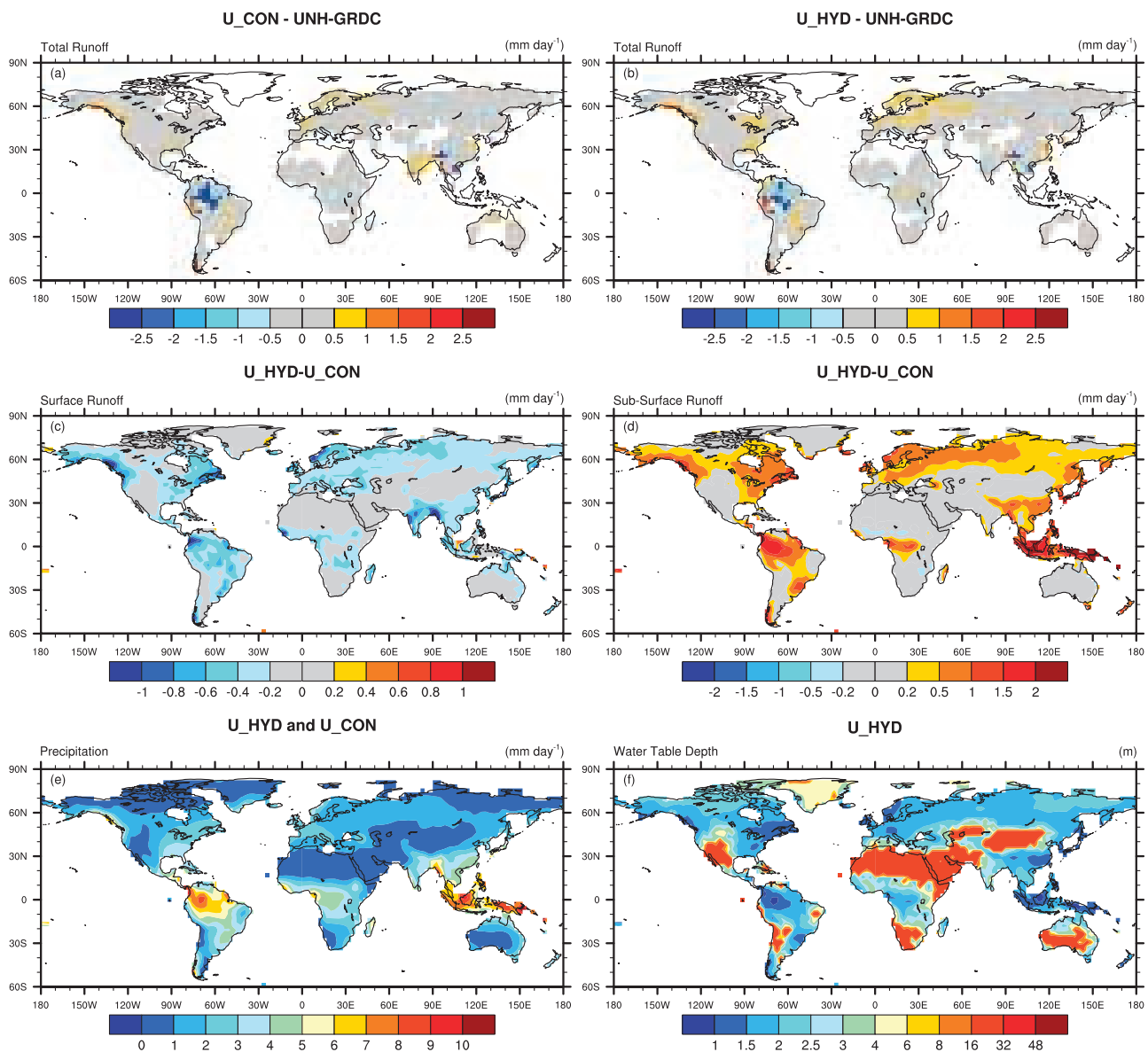
Figure 2. Annual mean of (a) U_CON minus UNH-GRDC total runoff, (b) U_HYD minus UNH-GRDC total runoff, (c) U_HYD-U_CON surface runoff, (d) U_HYD-U_CON subsurface runoff, (e) precipitation, and (f) U_HYD water table depth, at the model T42 resolution. All in mm day^{-1} except for water table depth in m. For Figures 2a and 2b, UNH-GRDC runoff is from 0.5° composite runoff fields [*Fekete et al.*, 2000, 2002] area-averaged to the model resolution but masked by UNH-GRDC observed runoff fields. Land areas in white indicate grid cells for which UNH-GRDC observed runoff is missing for more than 50% of the T42 grid cell.

(Figure 3c). In particular, the overly large amplitudes for the Sahel and India are corrected (regions M and N). In India (region N) for example, the precipitation peak in June creates large surface runoff in the old model (Figure 4). In the new model, surface runoff is greatly reduced which leads to much better agreement with UNH-GRDC runoff. The phase of runoff in tropical rainforest/savanna regions is well simulated by the old model ($r > 0.83$) with the exception of Central Africa (region K in Figure 3c). However, some small improvements are made by the new model. In India, the peak of subsurface runoff is delayed one month in the new model leading to improved phasing of the annual cycle (Figure 4).

3.3. River Discharge

[29] The River Transport Model (RTM) integrates grid cell runoff produced by CLM and transports it to the ocean via pathways that approximate the path of the real global river network [*Oleson et al.*, 2004]. Here, we evaluate the simulated annual mean and annual cycle of river discharge. The new model slightly improves the simulation of annual mean discharge at the farthest downstream station for world's largest 50 rivers (Figure 5). Both the linear and logarithmic correlation coefficients of modeled and observed river discharge are slightly higher in the new model. The mean bias (model-observations) improves from −45 km^3/yr to −8 km^3/yr. Following *Decharme and Douville* [2007], the
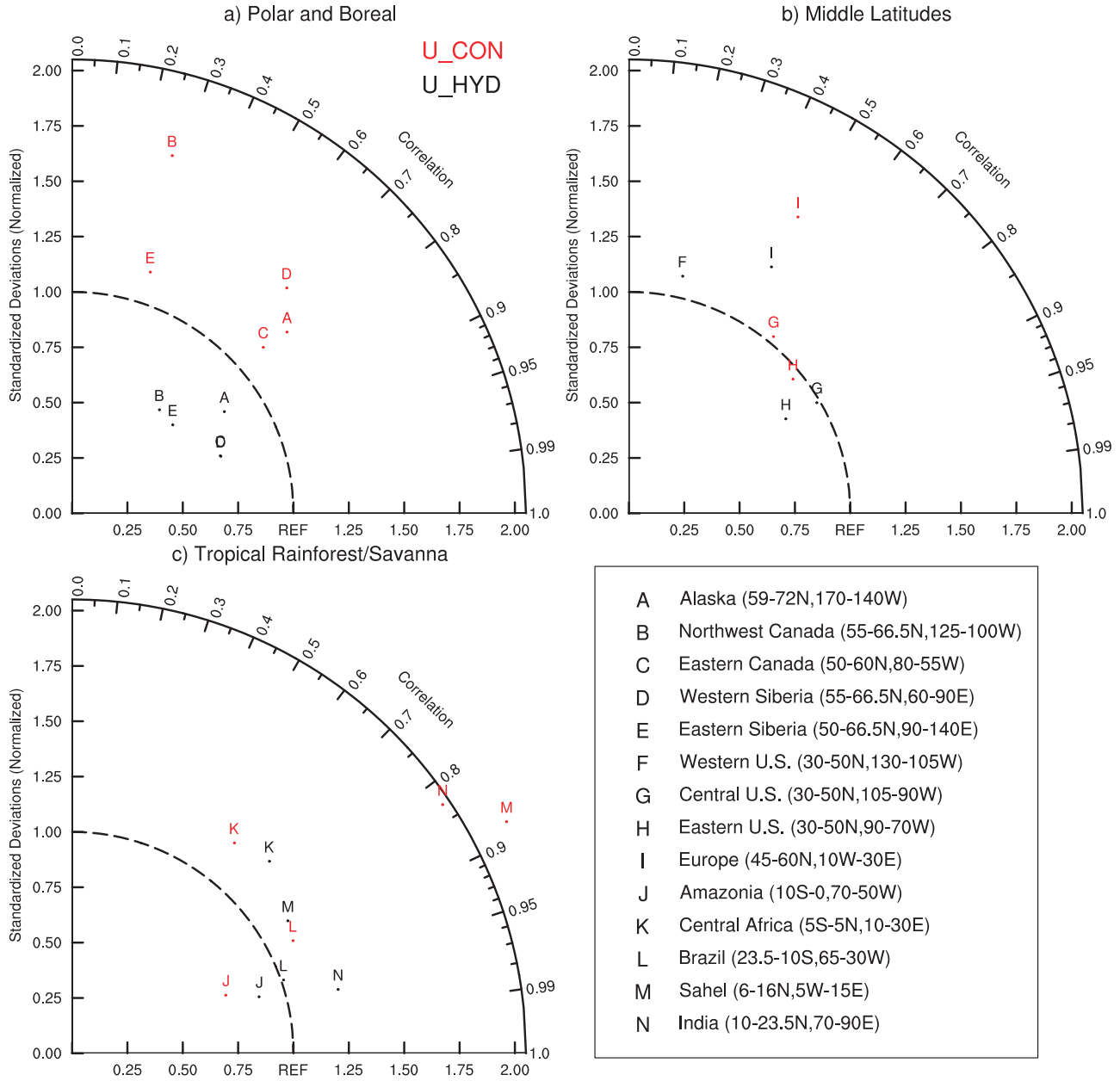
Runoff Annual Cycle

Figure 3. Statistics of the annual cycle of runoff from the U_CON (red letters) and U_HYD (black letters) simulations compared to UNH-GRDC composite runoff [*Fekete et al.*, 2000, 2002] for (a) polar and boreal, (b) middle latitudes, and (c) tropical rainforest/savanna regions. The radial distance from the origin to the letters is the standard deviation of the simulated monthly runoff normalized by the standard deviation of the observations (σ_m/σ_o). The azimuthal position of the letters is the linear correlation between simulated and observed monthly runoff (r). The r and σ_m/σ_o for the U_CON simulation in the Western U.S. (region F) do not fall within the plotted limits and are -0.15 and 1.83, respectively.

annual cycle of discharge from the largest 50 rivers is evaluated in Figure 6 using the efficiency criterion [*Nash and Sutcliffe*, 1970] defined as

$$Eff = 1.0 - \frac{\sum (Q_{sim}(t) - Q_{obs}(t))^2}{\sum (Q_{obs}(t) - \overline{Q_{obs}})^2} \qquad (1)$$

where $Q_{sim}(t)$ and $Q_{obs}(t)$ are the simulated and observed monthly discharge and $\overline{Q_{obs}}$ is the mean observed discharge.

According to *Boone et al.* [2004], *Eff* is 1 for a perfect simulation, above 0.7 for a very good one, between 0.5 and 0.7 for a reasonable one, and negative for a very poor one. The new model produces a reasonable to very good simulation of river discharge for about the same percentage of rivers as the old model (31% compared to 27%). An increase in the number of rivers with efficiency scores in the range of $0-0.5$ is simulated by the new model. These results are comparable to those generated by a range of land surface
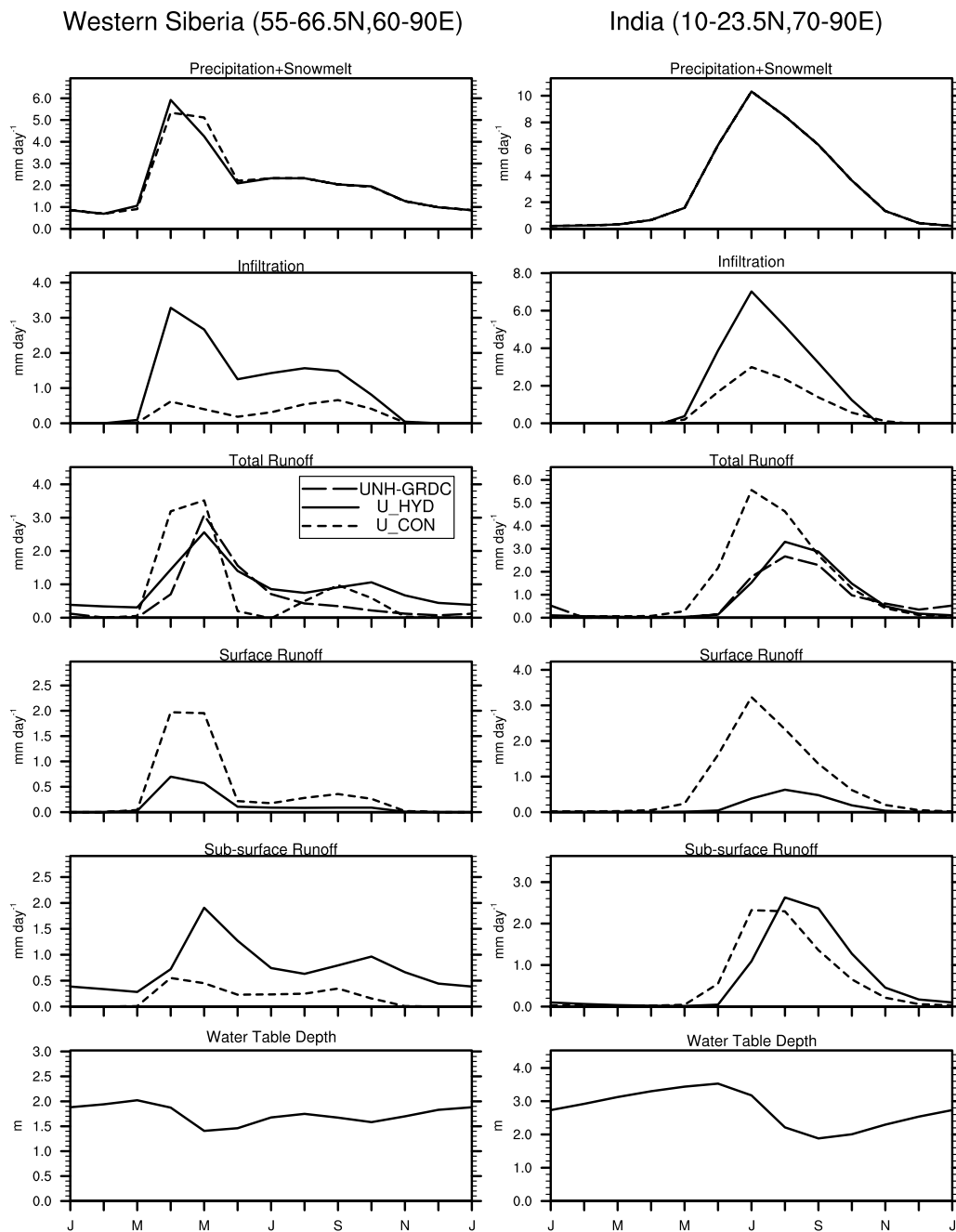
Western Siberia (55-66.5N,60-90E)

Precipitation+Snowmelt

Infiltration

Total Runoff

UNH-GRDC
U_HYD
U_CON

Surface Runoff

Sub-surface Runoff

Water Table Depth

India (10-23.5N,70-90E)

Precipitation+Snowmelt

Infiltration

Total Runoff

Surface Runoff

Sub-surface Runoff

Water Table Depth

Figure 4. Annual cycle of hydrologic quantities from the U_CON and U_HYD simulations for Western Siberia and India. UNH-GRDC runoff is from *Fekete et al.* [2000, 2002].

models evaluated by *Decharme and Douville* [2007]. In that study, the land surface models produced a reasonable to very good simulation for 20–40% of the gauging stations examined. This low percentage was attributed to various reasons including uncertainties in model parameterizations, possible anthropogenic influences (e.g., human withdrawal or impoundment of water), the quality of the atmospheric forcing, the use of a constant streamflow velocity in the river transport model, and the influence of seasonal floodplains, all of which may also apply to some extent to the current study.

[30] The new model has mixed effects on the annual cycle of discharge into the major oceans (Figure 7). The correla-tions between simulated and observed monthly discharge are quite good in the old model (>0.86 except for the Mediterranean/Black Seas). The high correlations are gen-erally maintained or improved in the new model, with the primary exception of the Indian Ocean. The new model delays peak discharge into the Indian Ocean by one month compared to the old model and the observations. Although the annual cycle of runoff over the India region in the new model appears to agree well with UNH-GRDC runoff (Figure 4), the simulated discharge for rivers with drainage areas outside this region (e.g., the Indus and Ganges rivers) have peaks that are 1–2 months later than the observed discharge (not shown).
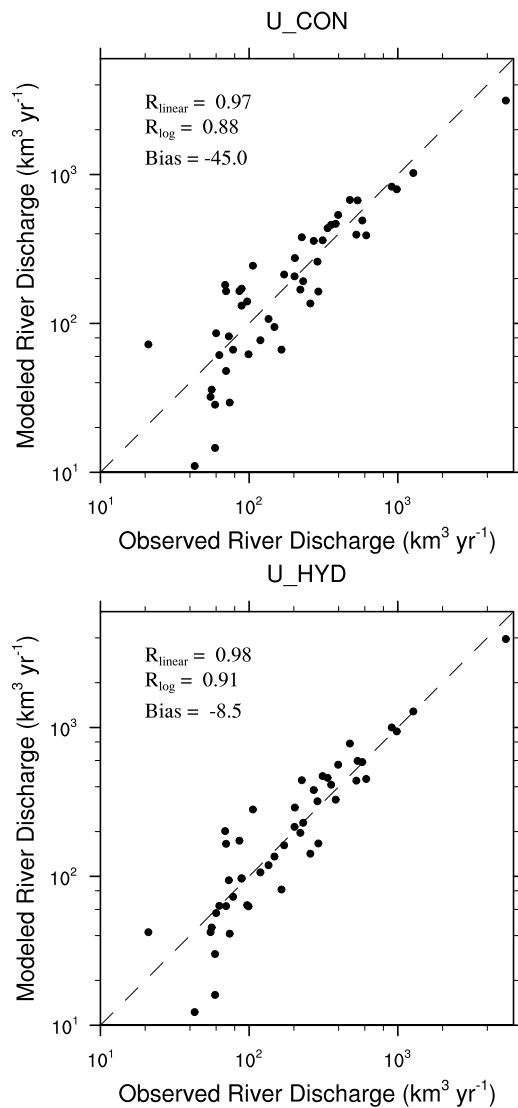
Figure 5. Comparison of observed and modeled (U_CON, U_HYD) annual discharge for the world's largest 50 rivers. R_{linear} and R_{log} are the linear and logarithmic correlation coefficients. The Niger and Zambeze rivers are excluded because the observed discharge is unrealistically low [*Qian et al.*, 2006]. Observed discharge is from *Dai and Trenberth* [2002].

[31] The sharp peak in June discharge to the Arctic Ocean due to snowmelt is somewhat more muted in the new model (Figure 7). Observed discharge data from several of the large Russian rivers (e.g., Ob, Yenisey) exhibit a sharp peak in June also that is not captured by the new model (not shown). Analysis of separate simulations conducted during CLM3.5 development indicate that a change made to the original formulation for water table depth (see equation (C10) and discussion in Appendix C) is primarily responsible for the dampened annual cycle in discharge to the Arctic Ocean. The fact that the new model's discharge to the Arctic Ocean agrees better with the discharge simulated by RTM-GRDC than with the *Dai and Trenberth* [2002] discharge may simply imply that the snowmelt runoff peak is not well resolved in the UNH-GRDC data.

[32] Figure 8 shows the interannual variability of streamflow for the world's top 10 rivers (except the Brahmaputra and Mekong which have records that are too short). As noted by *Qian et al.* [2006], the old model does a good job capturing the interannual variability in streamflow for these rivers with the exception of the two Russian rivers that flow north to the Arctic, the Yenisey and to a lesser extent the Lena. The new model performs generally as well or better than the old model for all rivers. Although the new model fails to improve upon the poor correlation in the old model for the Yenisey, interannual variability for other large Russian rivers with long records such as the Ob, Amur, and Severnaya Dvina remain well simulated (U_CON and U_HYD correlations for the Ob are 0.77 and 0.87, 0.89 and 0.92 for the Amur, and 0.89 and 0.84 for the Severnaya Dvina (not shown)).

3.4. Soil Moisture

[33] There are acknowledged difficulties in comparing observed and simulated soil moisture. Among these are the mismatch of scales between in situ observations and climate model grid cells (both vertical and horizontal scales), measurement accuracy, and the very definition of soil moisture. Despite these challenges, several recent studies have successfully applied observations of soil moisture from the Global Soil Moisture Data Bank (GSMDB) [*Robock et al.*, 2000] to the evaluation of land surface models [*Entin et al.*, 1999; *Guo and Dirmeyer*, 2006; *Qian et al.*, 2006]. As in these studies we focus on the ability of the model to simulate the phasing of the annual cycle and interannual variability in terms of anomalies not absolute soil moisture. Data from the GSMDB for 19 stations in Illinois [*Hollinger and Isard*, 1994], 43 stations in China, and 50 stations in the former USSR (FSU) [*Vinnikov and Yeserkepova*, 1991] are used. The station data were averaged into regions by *Dai et al.* [2004].
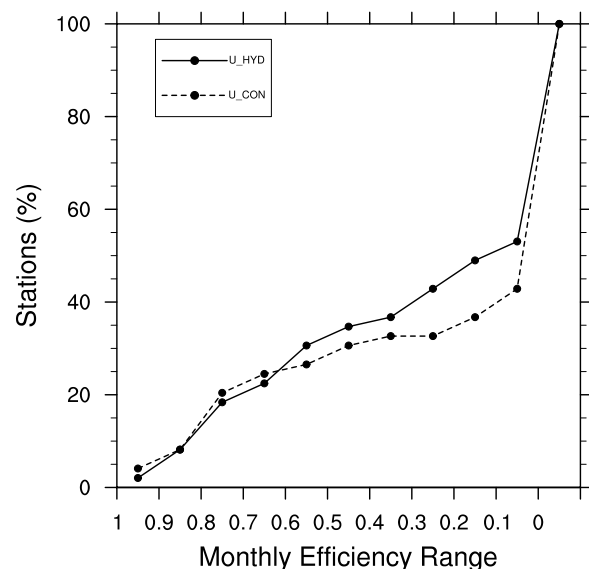

Figure 6. Cumulative efficiency distribution of monthly river discharge for the world's largest 50 rivers. The efficiency, *Eff*, is calculated from equation (1) in the text.
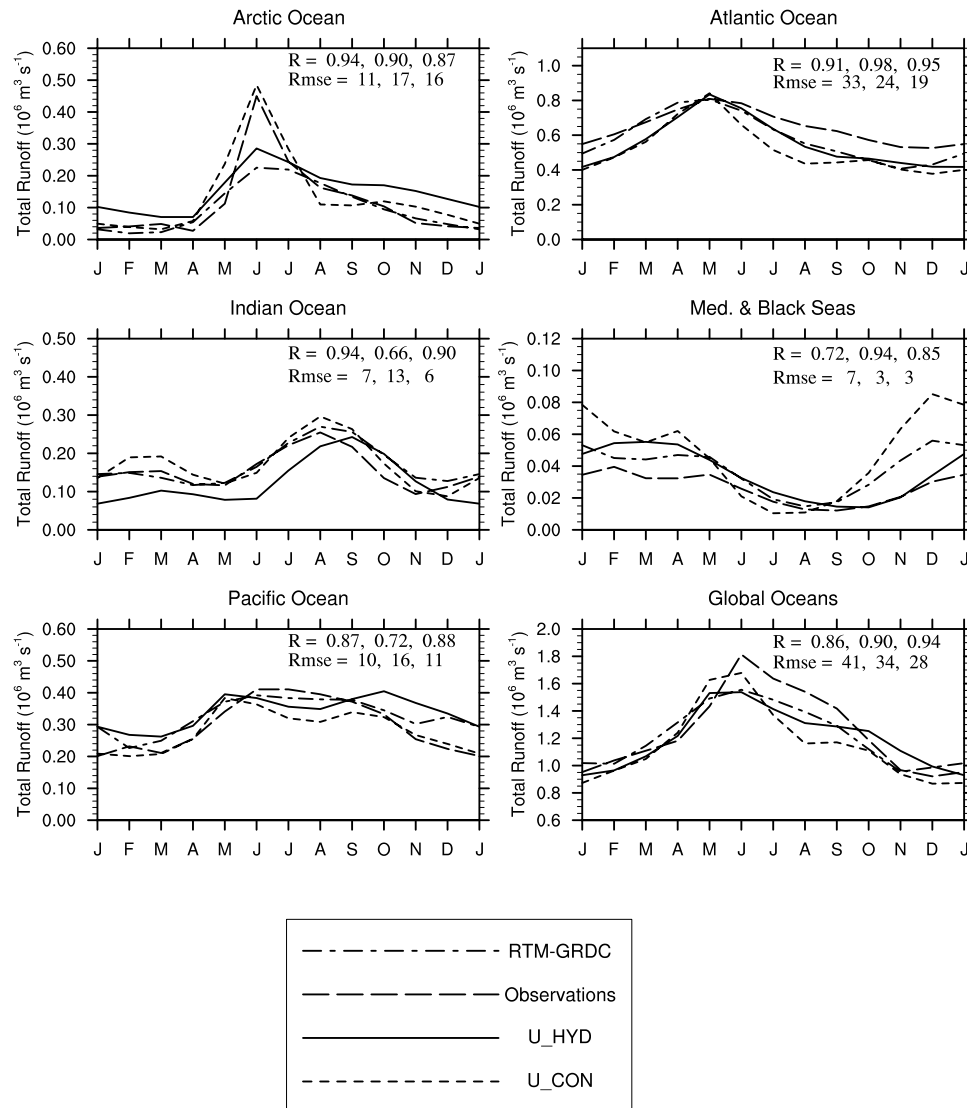
Mean Annual Cycle of Freshwater Discharge

Figure 7. Annual cycle of discharge into the Arctic, Atlantic, Indian, Mediterranean/Black Sea, Pacific, and Global oceans. R is the correlation coefficient between U_CON, U_HYD, and RTM_GRDC and observations from left to right and Rmse is the root mean square error. RTM-GRDC is discharge from the RTM driven by UNH-GRDC runoff. Observed and RTM-GRDC discharge is from *Dai and Trenberth* [2002].

[34] Figure 9 compares the simulated annual cycle of soil moisture anomalies in the rooting zone with the observed station data averaged for five regions. The new model does a reasonable job of simulating the soil moisture anomalies with correlation coefficients ranging from 0.71 to 0.96. The spring peak in soil moisture in the two FSU regions is better captured by the new model than the old model, resulting in a higher correlation coefficient for the FSU2 region. The new model has much lower surface runoff from snowmelt than the old model, allowing more water to infiltrate into the soil (not shown). For the FSU1 region, the correlation is somewhat poorer in the new model despite capturing the spring peak because soil moisture in other months is more constant than in the observations. The simulated annual cycle in the new model is improved in East China and about the same in South China.

[35] The ability of the models to reproduce multiyear monthly soil moisture anomalies in these regions is shown in Table 2. The first correlation coefficient r_1 is a measure of the model's ability to reproduce interannual variability (annual cycle removed). Results for the new model are mixed, with higher correlations in Illinois and FSU2, a lower correlation in FSU1, and about the same as the old model in the two China regions. Correlations are highest overall in Illinois, probably because the forcing data is of higher quality [*Guo and Dirmeyer*, 2006]. Inclusion of the annual cycle, which the models generally reproduce well (Figure 9), improves the correlation in all regions with the exception of South China which has a weak annual cycle. The new model also reproduces well total soil moisture for individual stations as shown in Table 3 for the 19 Illinois stations. Correlations for the top 1 m soil moisture range
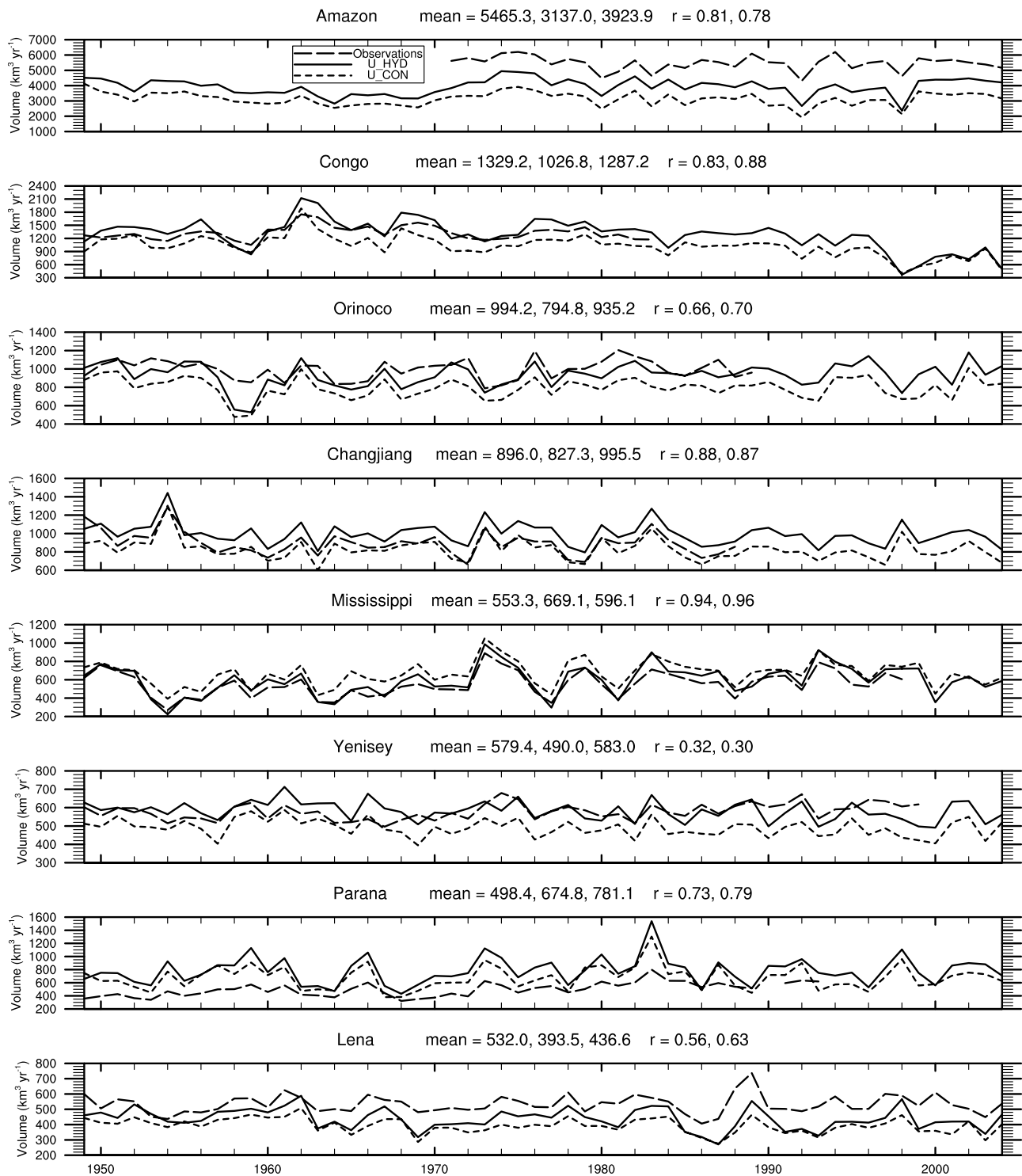
Figure 8. Water year (October–September) annual discharge for the world's top ten rivers (except for Brahmaputra and Mekong which have records that are too short). Also shown are the mean of the observations, U_CON, and U_HYD from left to right, and the correlation coefficient (r) between U_CON and observations, and U_HYD and observations from left to right. Observed discharge has been updated from *Dai and Trenberth* [2002] by A. Dai et al. (Changes in continental freshwater discharge from 1949 to 2004, submitted to *Journal of Climate*, 2007).

from 0.54 to 0.83 with the median correlation being 0.77. Correlations for the top 2 m are generally somewhat lower with the median correlation being 0.71. The temporal correlations presented in Tables 2 and 3 appear to be

reasonable when compared to the correlations produced by a range of land surface models participating in GSWP-2. For example, the median correlation for the top 1 m soil moisture for stations in Illinois for the period 1986–1995 is 0.83 in the
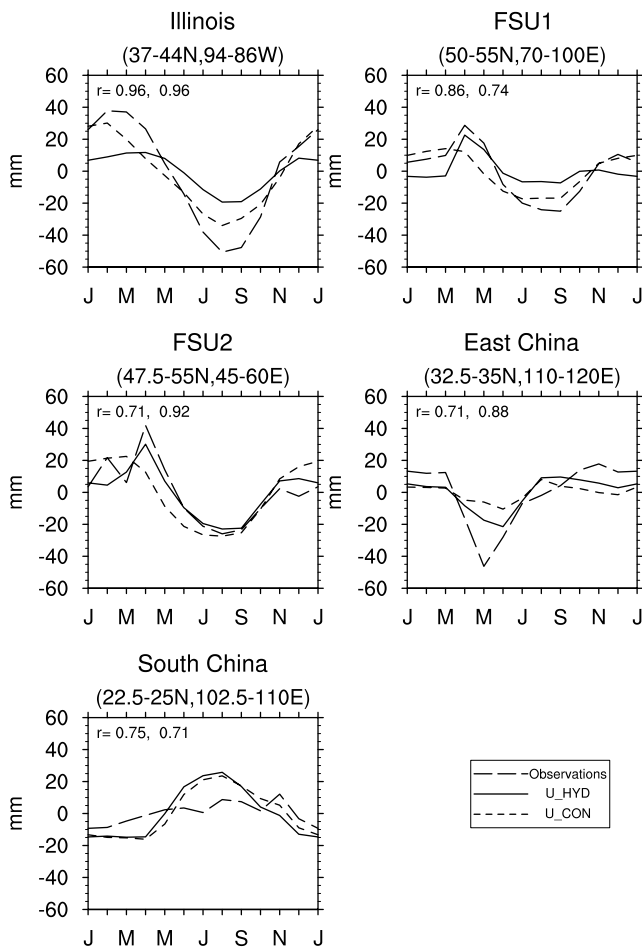
Figure 9. Mean annual cycle of observed and simulated soil moisture anomalies for five regions. Model results are calculated from same years as observed (Illinois: 1981–2001, FSU: 1972–1985, China: 1981–1991). Illinois soil moisture is total for the top 0.9 m. FSU and China soil moisture are plant available soil moisture for the top 1.0 m. Model wilting point soil moisture was subtracted from total soil moisture to obtain plant available soil moisture. Observed station data were averaged into regions by *Dai et al.* [2004]. r is the correlation coefficient between U_CON, U_HYD and observations from left to right.

new model which is at the higher end of the range of model results reported for GSWP-2 [*Guo and Dirmeyer*, 2006].

3.5. Terrestrial Water Storage

[36] Two GRACE data sets interpolated to the model resolution are used to evaluate the simulated terrestrial water storage anomalies of both models over several river basins [*Seo and Wilson*, 2005; *Chen et al.*, 2005] (GRACE1 and GRACE2 respectively in Figure 10). The observed data sets apply different filtering algorithms to the raw data and can be used to provide an estimate of the observational uncertainty. These are the same river basins selected by *Niu et al.* [2007] to evaluate the groundwater model. The spatial resolution of the model simulations is not optimal for any rigorous comparison with GRACE data because of the mismatch in spatial scales between the GRACE and model data and the

difficulty in delineating river basins at the coarse model resolution. The lack of resolution of the GRACE data probably results in some cross-contamination of the annual cycle between adjacent river basins (e.g., the Orinoco and Amazon as noted by *Niu et al.* [2007]). However, the comparison is sufficient to confirm that the new model has a much improved annual cycle of total water storage (Figure 10). The old model reasonably captures the phase of the annual cycle but the amplitude is generally much less than observed. The median correlation between GRACE1 and the old model is 0.79 for these river basins and improves to 0.90 in the new model (the correlation with GRACE2 improves from 0.76 to 0.87). The median root mean square error (rmse) between GRACE1 and the old model is 39 mm and improves to 22 mm in the new model (the median rmse with GRACE2 improves from 34 to 23 mm).

3.6. Soil Moisture Variability

[37] Global maps of the standard deviation of monthly soil moisture (annual cycle included) show that the new model has much larger soil moisture variability in the tropics than the old model in both the rooting zone (1 m) and the total column (Figure 11). As shown by a comparison with GRACE data described in the previous section (Figure 10), the annual cycle of total water storage is reproduced well by the new model in the tropics which suggests that this increase is reasonable. On the other hand, variability in the rooting zone in the mid to high latitudes remains similar to the old model or has even decreased in some regions (Alaskan Arctic, Europe, and eastern Siberia). The point-scale soil moisture observations suggest that this relatively low variability may be a deficiency in the new model. For example, the simulated decline in summer soil moisture in Illinois is less than observed (Figure 9). Similarly, the variability of the soil moisture anomalies in all regions but East China is about half that observed (as expressed by $\sigma_{MODEL}/\sigma_{OBS}$ in Table 2). The variability of soil moisture in the top 1 m for individual Illinois stations is also much less than observed ($\sigma_{U\_HYD}/\sigma_{OBS}$ in Table 3). Variability in the top 2 m is closer to observed but generally still too low (Table 3). Some of the higher values of $\sigma_{U\_HYD}/\sigma_{OBS}$ for the top 1 m of soil in Table 3 can likely be explained not by higher variability in the model but by lower variability in the observations [*Hollinger and Isard*, 1994]. For example, the Ina site ($\sigma_{U\_HYD}/\sigma_{OBS}$ = 0.90) is

Table 2. Temporal Correlations Between Simulated and Observed Monthly Soil Moisture Anomalies for Illinois (1981–2001), FSU (1972–1985), and China (1981–1991)[a]

	r_1		r_2		$\sigma_{MODEL}/\sigma_{OBS}$	
	U_CON	U_HYD	U_CON	U_HYD	U_CON	U_HYD
Illinois	0.69	0.72	0.85	0.86	0.41	0.38
FSU1	0.68	0.50	0.78	0.65	0.48	0.36
FSU2	0.54	0.66	0.65	0.82	0.53	0.49
East China	0.68	0.67	0.62	0.68	0.72	1.16
South China	0.57	0.54	0.47	0.46	0.44	0.48

[a]Here r_1 is correlation with annual cycle removed, and r_2 is with annual cycle included. Also shown is the modeled standard deviation of the anomalies (annual cycle removed) normalized by the observed standard deviation ($\sigma_{MODEL}/\sigma_{OBS}$). Illinois soil moisture is total for the top 0.9 m. FSU and China soil moisture are plant available soil moisture for the top 1.0 m. Model wilting point soil moisture was subtracted from total soil moisture to obtain plant available soil moisture.

Table 3. Temporal Correlations (r) Between U_HYD and Observed Monthly Total Soil Moisture (Top 1 m and 2 m) for 19 Stations in Illinois for 1981–2004[a]

Station	Location	Top 1 m		Top 2 m	
		r	$\sigma_{\mathrm{U\_HYD}}/\sigma_{\mathrm{OBS}}$	r	$\sigma_{\mathrm{U\_HYD}}/\sigma_{\mathrm{OBS}}$
Bondville	40.05N, 88.22W	0.80	0.41	0.77	0.67
Dixon Springs-Bare	37.45N, 88.67W	0.64	0.77	0.40	1.09
Brownstown	38.95N, 88.95W	0.77	0.40	0.68	0.67
Orr Center (Perry)	39.80N, 90.83W	0.71	0.36	0.64	0.59
De Kalb	41.85N, 88.85W	0.71	0.47	0.63	0.77
Monmouth	40.92N, 90.73W	0.75	0.46	0.69	0.86
Peoria	40.70N, 89.52W	0.71	0.39	0.62	0.60
Springfield	39.52N, 89.62W	0.70	0.64	0.71	1.03
Belleville	38.52W, 89.88W	0.82	0.31	0.80	0.50
Carbondale	37.72N, 89.23W	0.83	0.37	0.80	0.64
Olney	38.73N, 88.10W	0.80	0.78	0.73	1.36
Freeport	42.28N, 89.67W	0.54	0.34	0.50	0.52
Ina	38.13N, 88.92W	0.77	0.90	0.67	1.49
Stelle	40.95N, 88.17W	0.83	0.40	0.84	0.74
Topeka	40.30N, 89.90W	0.71	0.65	0.71	1.01
Oak Run	40.97N, 90.15W	0.79	0.42	0.74	0.70
Fairfield	38.38N, 88.38W	0.83	0.57	0.83	1.05
Champaign	40.08N, 88.23W	0.82	0.50	0.80	1.03
Dixon Springs-Grass	37.45N, 88.67W	0.81	0.45	0.68	0.77

[a]Also shown is the modeled standard deviation of soil moisture normalized by the observed standard deviation ($\sigma_{\mathrm{U\_HYD}}/\sigma_{\mathrm{OBS}}$). Observed soil moisture is from *Hollinger and Isard* [1994].

close to the shore of a lake which evidently keeps the soil fairly moist. The Dixon Springs-Bare site ($\sigma_{\mathrm{U\_HYD}}/\sigma_{\mathrm{OBS}} = 0.77$) has no vegetation and reduced soil moisture variability compared to the adjacent Dixon Springs-Grass site ($\sigma_{\mathrm{U\_HYD}}/\sigma_{\mathrm{OBS}} = 0.45$). The Topeka site ($\sigma_{\mathrm{U\_HYD}}/\sigma_{\mathrm{OBS}} = 0.65$) has anomalously sandy soil compared to the silty loam of the majority of sites. The sandy soils are characterized by very little soil structure, low porosity, and low water holding capacity which reduces variability. The apparently low variability in the new model is discussed further in section 4.

3.7. Water Cycle and Energy Balance in the Amazon

[38] As noted in the introduction, the simulation of water and energy balance in Amazonia in the old model is particularly poor. The new model improves the partitioning of ET, with more transpiration and lower canopy and soil evaporation (Figure 12). Its annual cycle of sensible and latent heat flux is much less strong than that of the old model. Soil moisture is now only slightly limiting in the dry season, i.e. plants are able to photosynthesize at a fairly constant rate throughout the year. The annual cycle of runoff is improved though still low compared to observations. Although there are no detailed observations of these variables at this spatial scale, the behavior of the new model is more consistent with what is known about energy and water balance in the Amazon in general [*Matsuyama*, 1992; *Fu and Li*, 2004; *Shuttleworth*, 1988; *Malhi et al.*, 2002; *Nepstad et al.*, 2002; *Ubarana*, 1996; *Marin et al.*, 2000]. *Stöckli et al.* [2008] found that CLM3.5 had larger soil water storage which improved the simulation of latent heat flux in the dry season at a tropical forest site in Brazil. This result is consistent with observations that suggest dry season ET is maintained by a deep soil moisture reservoir

recharged by rainfall during the previous wet season [*Juárez et al.*, 2007].

3.8. Vegetation Biogeography

[39] Shortcomings in the CLM3 hydrology became evident early on when the optional Dynamic Global Vegetation Model (DGVM) was introduced [*Levis et al.*, 2004]. *Bonan and Levis* [2006] documented a 480-year CLM3-DGVM simulation in which the model's dry soil moisture bias resulted in overestimated bare ground at the expense of grasses (e.g., in the western U.S.), overestimated grass cover at the expense of trees (e.g., eastern U.S.), and overestimated raingreen trees at the expense of evergreens (e.g., tropical rainforest regions). Global plant productivity was underestimated.

[40] Simulated vegetation by CLM3.5-DGVM illustrates the beneficial effects of wetter soils in this version of the CLM (Figure 13). The model now simulates a continuous circumpolar evergreen boreal forest, a distinct deciduous boreal forest in eastern Siberia, and a more realistic distribution of trees and grasses in the midlatitudes and tropics. Some discrepancies between the simulated and observed vegetation cover are due to the absence of shrubs and crops in the DGVM. Plant productivity is now overestimated mainly due to changes made in the autotrophic respiration and canopy integration methods when the optional prognostic treatment of carbon and nitrogen cycle dynamics was introduced (CLM-CN [*Thornton and Zimmermann*, 2007; *Thornton et al.*, 2007]). This overestimation would likely be corrected by coupling the DGVM and CN components. This coupling is not available in CLM3.5.

4. Discussion

[41] The new hydrological parameterizations that are implemented here to alleviate biases in CLM3 are evaluated using observed data sets of runoff, river discharge, soil moisture, and total water storage. Although the degree of performance of the new model compared to CLM3 differs according to the aspect of simulation being evaluated, the overall impression given by an analysis is that it has improved in most aspects. One of the most significant improvements is a better partitioning of the ET components; transpiration, canopy evaporation, and soil evaporation. Partitioning is reordered in the new model with transpiration now the largest component followed by ground and canopy evaporation. This change in relative magnitudes is a consequence of several new parameterizations that act together to produce more realistic behavior in the model. Canopy interception is reduced, which increases the amount of water that can reach the ground. A reduction in surface runoff as a consequence of the introduction of a prognostic groundwater model and a more realistic frozen soil scheme allows more of this water to infiltrate into the soil and increase soil moisture. The addition of a resistance term to the soil evaporation scheme prevents excessive loss of soil moisture to evaporation. The net effect is to increase soil water storage so that more soil moisture is available during the dry season. A new canopy integration scheme coupled with an improved description of soil water availability to plants is able to make use of this additional soil moisture, thereby increasing transpiration. Predictably, photosynthesis, which
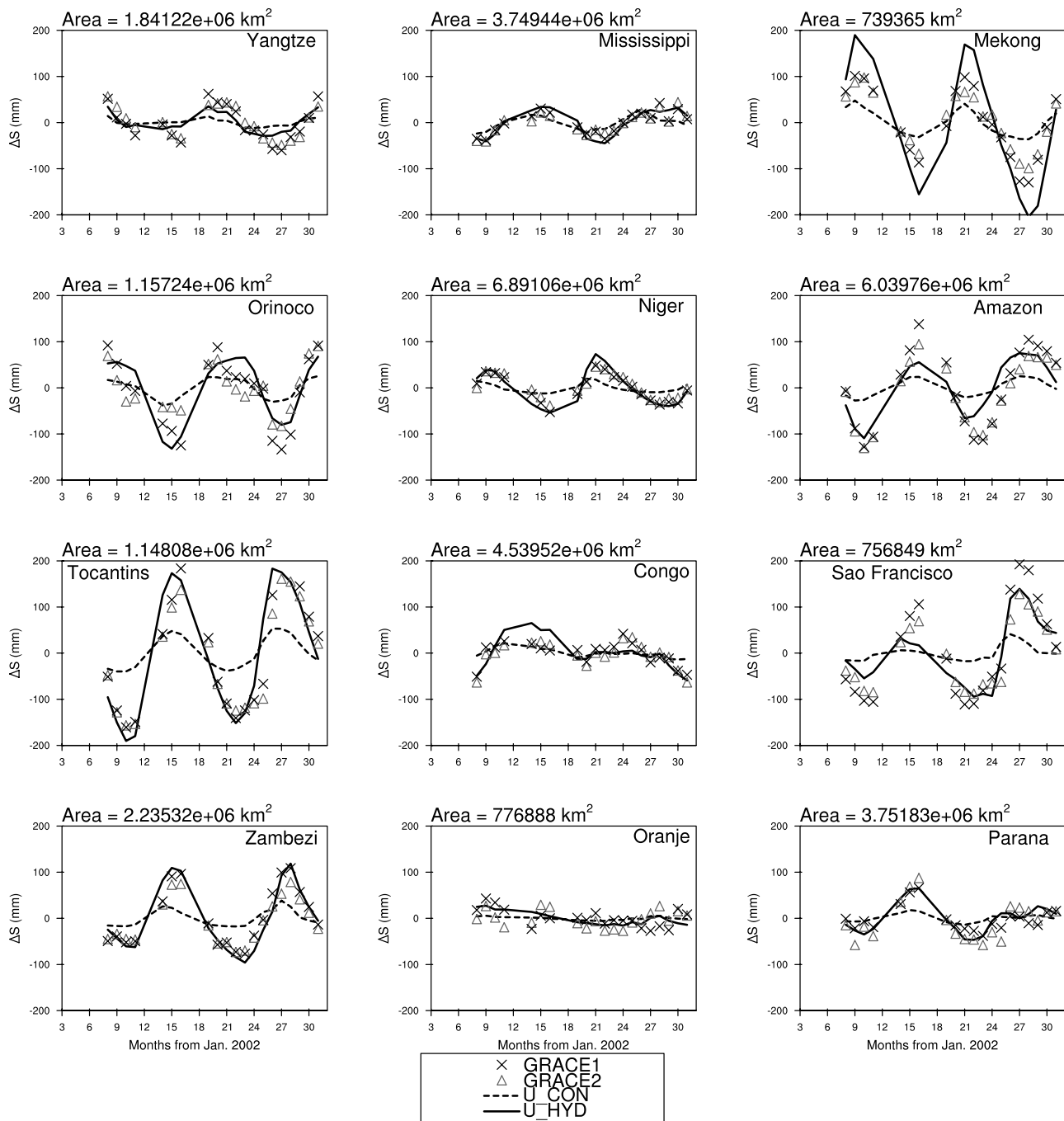
Figure 10. Total water storage anomalies (mm) for U_HYD and U_CON compared to two sources of GRACE data (GRACE1 [*Seo and Wilson*, 2005] and GRACE2 [*Chen et al.*, 2005]). Model total water storage anomalies are calculated from the sum of snow water, canopy water, total column soil water, and aquifer water. GRACE data were interpolated to the model resolution.

is linked to hydrology through soil moisture, stomatal conductance, and transpiration, is substantially increased in the new model. When the new model is coupled to a DGVM, the simulation of vegetation biogeography under modern-day climate is closer to that observed. The influence of the new parameterizations is nicely summarized by a qualitatively better simulation of water and energy fluxes and vegetation biogeography in the Amazon. In addition, although the turbulent surface fluxes are not evaluated here due to lack of global data, *Stöckli et al.* [2008] clearly show

that the improvements in the hydrological cycle translate into substantial quantitative improvements in the simulation of latent heat and smaller biases in sensible heat.

[42] Despite the poor partitioning of ET, the old model is able to simulate certain aspects of hydrology fairly well in an annual mean sense. For example, the annual mean runoff over most of the land surface and annual mean discharge for the world's top 50 rivers are reasonable in the old model. Consequently, CLM3.5 does not further improve these aspects very much, other than its notable increase in
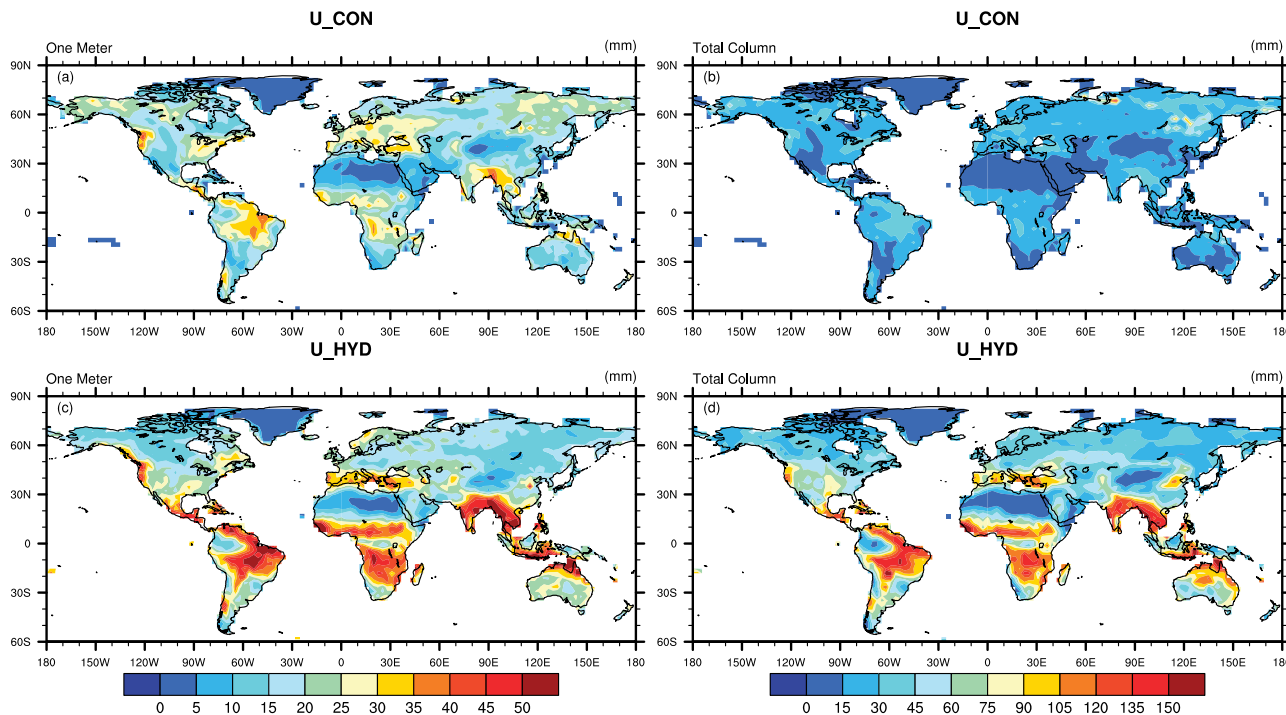
Standard Deviation of Soil Moisture

Figure 11. Standard deviation of 1948–2004 monthly soil moisture for U_CON (a) top 1 m, (b) total soil column, and U_HYD, (c) top 1 m, and (d) total soil column.

discharge from the Amazon. In general, the phase and amplitude of the annual cycle of runoff regionally averaged are improved in the new model (Figure 3). However, this does not necessarily translate into similar improvements in the simulated annual cycle of river discharge (e.g., Figure 6). The inconsistency between grid cell runoff and river discharge makes it difficult to use the observations to motivate improvements in the model.

[43] The process of determining a set of solutions for the hydrological biases in CLM3 has been one of trial and error and discovery. As noted by *Stöckli et al.* [2008] missing processes or deficiencies in existing formulations sometimes may only be revealed through correction of the representation of other processes (e.g., the need for nitrogen limitation on photosynthesis). The low soil moisture variability in the rooting zone in certain regions noted in a comparison of simulated soil moisture with observed for Illinois is a deficiency in the new model that may require further attention. It may be due to a missing process, remaining deficiencies in the parameterization of existing processes, or perhaps incorrectly tuned free parameters in the model. The causes and potential solutions for this discrepancy are being explored. In preliminary simulations, we have found that because of the presence of shallow water table depths in these regions, there is a strong upward flux of water from the saturated lower soil layers to the rooting zone. This upward flux is sufficient to prevent substantial drying of the rooting zone even during times of drought (e.g., the 1988 drought in the Midwest U.S.). Sensitivity simulations currently point to reductions in hydraulic conductivities and/or more roots in the upper soil layers as

possible solutions. The poor simulation of snowmelt induced June peak in discharge to the Arctic Ocean suggests that further improvements to high-latitude hydrology are also desirable.

5. Conclusions

[44] Previous evaluations and the data presented here establish that the Community Land Model version 3 (CLM3) has several deficiencies in its simulation of the hydrological cycle. Chief among these are a poor partitioning of evapotranspiration with adverse implications for runoff, soil water storage dynamics, plant productivity, and vegetation biogeography. A set of new data sets and parameterizations developed by the CLM user community was implemented into the new model CLM3.5 with generally beneficial effects on the reduction of these biases. These individual improvements act in combination to improve the partitioning of evapotranspiration into transpiration, and ground and canopy evaporation. The result is wetter soils, less plant water stress, increased transpiration and photosynthesis, and improved annual cycle of runoff and total water storage. The new hydrology interacts with a DGVM to produce more realistic distributions of global vegetation. These effects are supported by observations. The benefits of the new formulations for other aspects of the simulated hydrology (mainly annual mean quantities) are not as apparent, primarily because the old model already produced a reasonable simulation in this regard. Lower soil moisture variability in the rooting zone than observed is identified as a deficiency.
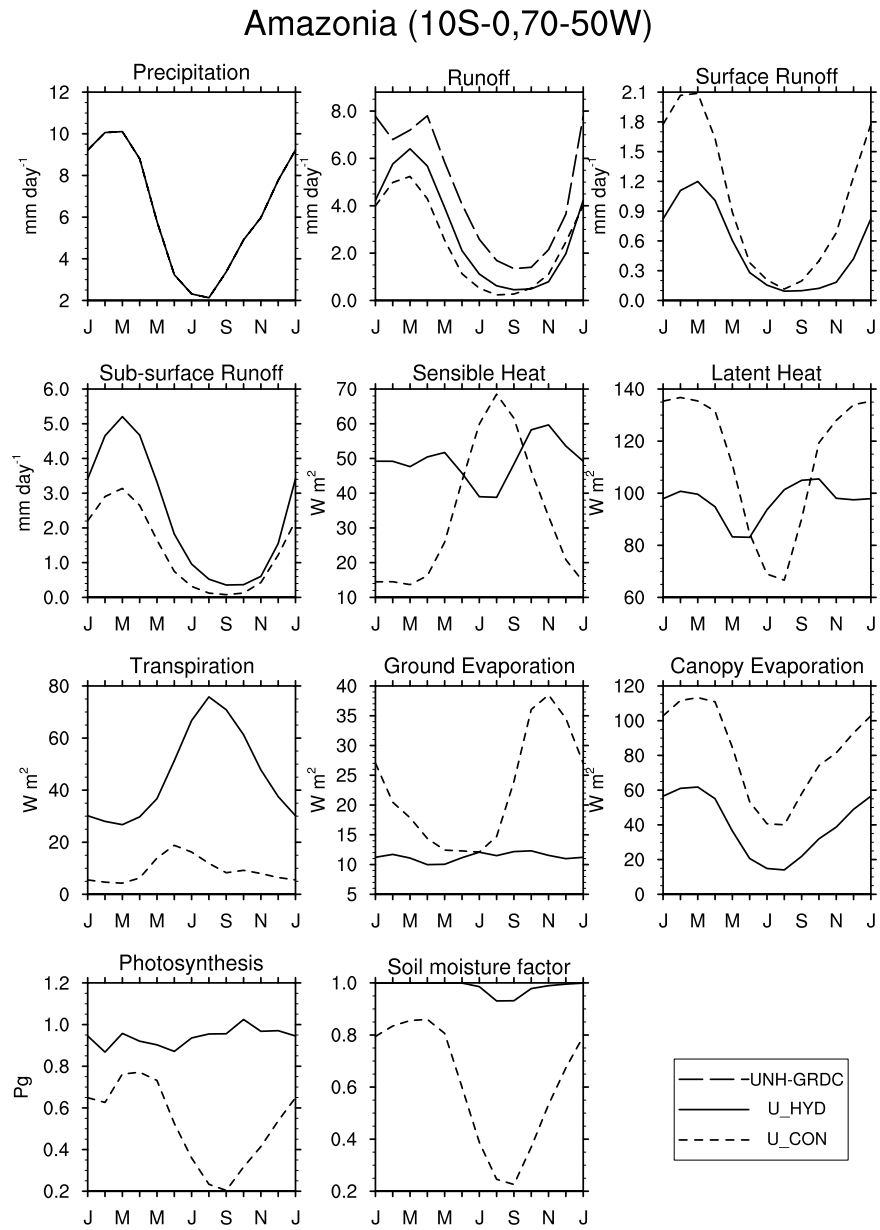
Amazonia (10S-0,70-50W)

Figure 12. U_CON and U_HYD hydrology in Amazonia. UNH-GRDC runoff is from *Fekete et al.* [2000, 2002].
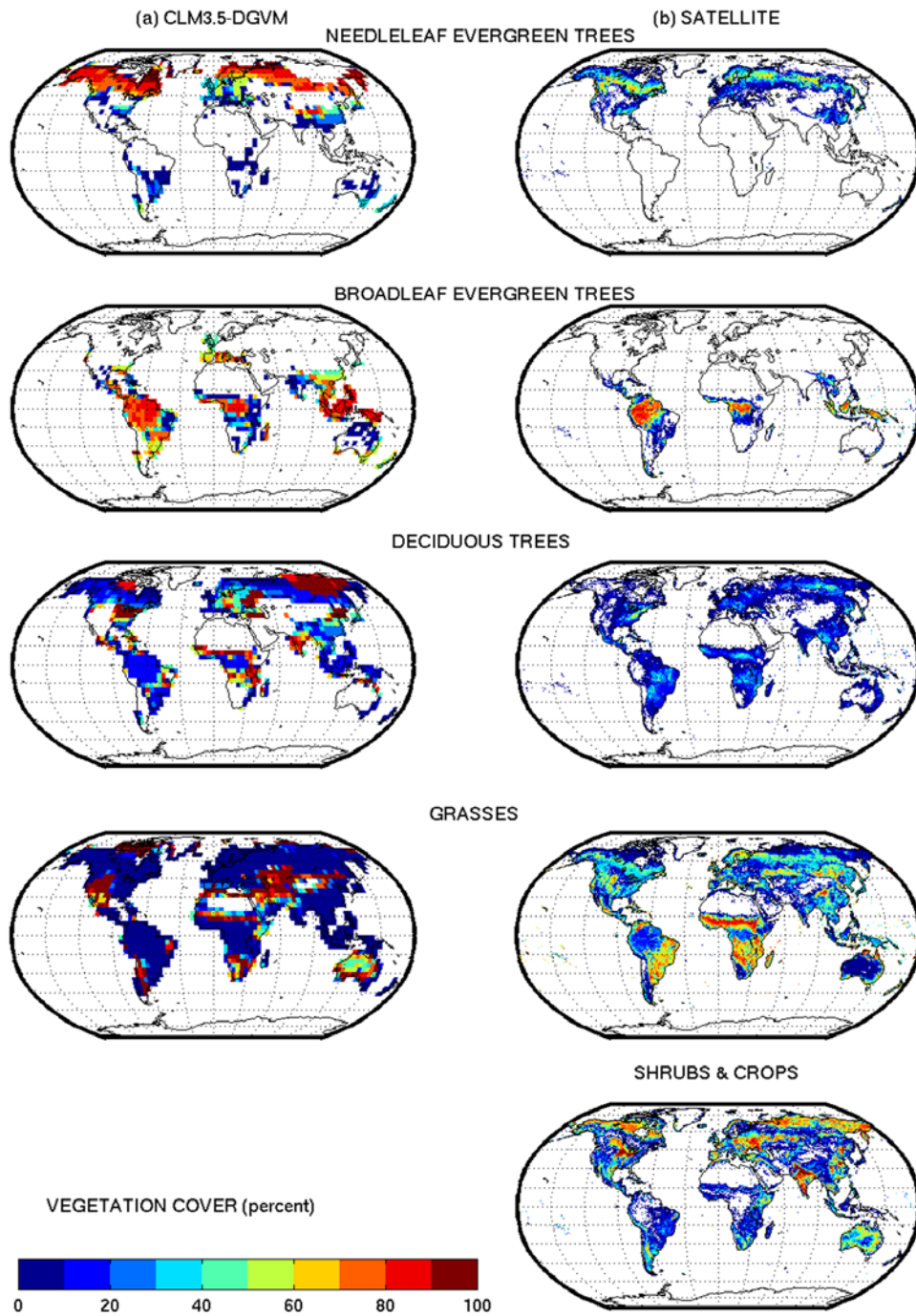
Figure 13. Global vegetation cover as a percentage of the soil-covered portion of the grid cell simulated by (a) CLM3.5-DGVM and (b) observed. CLM3.5-DGVM data are averages from the last 20 years of the simulation. Observed data are from *Lawrence and Chase* [2007]. The simulations were initialized with CLM's arbitrary initial conditions [*Oleson et al.*, 2004] (i.e., not spun up and 0% vegetated).

Appendix A: Canopy Interception

[45] The rate of water intercepted by the canopy (kg m^{-2} s^{-1}) is

$$q_{intr} = \alpha(q_{rain} + q_{sno})\{1 - \exp[-0.5(L + S)]\} \qquad \text{(A1)}$$

where q_{rain} and q_{sno} are the liquid and solid precipitation rates (kg m^{-2} s^{-1}) and L and S are the exposed leaf and stem area index. The factor α has been changed from 1.0 to 0.25 to scale the interception from point to grid cell [*Lawrence et al.*, 2007].

Appendix B: Surface and Subsurface Runoff

[46] The simple TOPMODEL-based runoff model (SIM-TOP) described by *Niu et al.* [2005] is implemented. SIM-TOP parameterizes surface runoff as consisting of overland flow from Dunne (runoff over saturated ground) and Horton (infiltration excess) mechanisms as

$$q_{over} = f_{sat}q_{liq,0} + (1 - f_{sat})\max(0, q_{liq,0} - q_{infl,\max}) \qquad \text{(B1)}$$

where $q_{liq,0}$ is liquid precipitation reaching the ground plus any melt water from snow (kg m^{-2} s^{-1}) and $q_{infl,\max}$ is a maximum soil infiltration capacity (kg m^{-2} s^{-1}).

[47] The variable f_{sat} is the saturated fraction of a grid cell. In the work of *Niu et al.* [2005], f_{sat} was solely determined by the water table depth and the subgrid topographic characteristics of the grid cell and represented the potential or maximum saturated fraction ($f_{\max}$). *Niu and Yang* [2006] modified the expression for f_{sat} to include a dependence on impermeable area fraction $f_{frz,1}$ of the top $i = 1$ soil layer (defined in Appendix D) as

$$f_{sat} = (1 - f_{frz,1})f_{\max}\exp(-0.5fz_{\nabla}) + f_{frz,1} \qquad \text{(B2)}$$

where $f_{\max}$ is the maximum saturated fraction, f is a decay factor (m^{-1}), and z_{∇} is the water table depth (m). The decay factor f for global simulations was determined through sensitivity analysis and comparison with observed runoff to be 2.5 m^{-1}.

[48] The maximum saturated fraction $f_{\max}$ is defined as the discrete cumulative distribution function (CDF) of the topographic index when the grid cell mean water table depth is zero. Thus, $f_{\max}$ is the percent of pixels in a grid cell whose topographic index is larger than or equal to the grid cell mean topographic index. It should be computed explicitly from the CDF at each grid cell at the resolution that the model is run. However, because this is a computationally intensive task for global applications, $f_{\max}$ is calculated once from the CDF at 0.5 degree resolution following *Niu et al.* [2005] and then area-averaged to the desired resolution. The 0.5 degree resolution is compatible with resolution of the other CLM input surface data sets (e.g., plant functional types, leaf area index).

[49] The maximum infiltration capacity in equation (B1) $q_{infl,\max}$ is determined from soil texture and soil moisture [*Entekhabi and Eagleson*, 1989] as

$$q_{infl,\max} = k_{sat,1}[1 + v(s - 1)]. \qquad \text{(B3)}$$

The liquid water content of the top soil layer relative to effective porosity and adjusted for saturated fraction is determined from

$$s = \frac{\dfrac{\theta_{liq,1}}{\max(\theta_{imp}, \theta_{sat,1} - \theta_{ice,1})} - f_{sat}}{1 - f_{sat}} \qquad \text{(B4)}$$

where $\theta_{liq,1}$ and $\theta_{ice,1}$ are the volumetric liquid water and ice contents of the top soil layer, and $\theta_{imp} = 0.05$ is a minimum effective porosity. The variable v

$$v = -\left(\frac{d\psi}{ds}\right)_{s=1}\frac{1}{0.5\Delta z_1} \qquad \text{(B5)}$$

where Δz_1 is the thickness of the top soil layer (mm) and

$$\left(\frac{d\psi}{ds}\right)_{s=1} = -B_1\psi_{sat,1}. \qquad \text{(B6)}$$

The saturated hydraulic conductivity $k_{sat,1}$ (kg m^{-2} s^{-1}), volumetric water content at saturation (i.e., porosity) $\theta_{sat,1}$, exponent B_1, and saturated soil matric potential $\psi_{sat,1}$ (mm) are determined from soil texture (%*sand*, %*clay*) [*Oleson et al.*, 2004].

[50] In the work of *Niu et al.* [2005], the subsurface runoff or drainage q_{drai} (kg m^{-2} s^{-1}) was formulated as

$$q_{drai} = q_{drai,\max}\exp(-fz_{\nabla}) \qquad \text{(B7)}$$

where $q_{drai,\max} = 4.5 \times 10^{-4}$ kg m^{-2} s^{-1} is the maximum subsurface runoff when the grid-averaged water table depth is zero. To restrict drainage in frozen soils, *Niu et al.* [2005] added the following condition

$$q_{drai} = 0 \qquad \text{for } w_{ice,10} > w_{liq,10} \qquad \text{(B8)}$$

where $w_{ice,10}$ and $w_{liq,10}$ is the ice and liquid water content of the 10th soil layer (kg m^{-2}). In preliminary testing we found that a more gradual restriction of drainage was required so that the water table depth remained dynamic under partially frozen conditions. We implemented the following

$$q_{drai} = (1 - f_{imp})q_{drai,\max}\exp(-fz_{\nabla}) \qquad \text{(B9)}$$

where f_{imp} is the fraction of impermeable area determined from the ice content of the soil layers interacting with the water table

$$f_{imp,i} = \left\{\exp\left[-\alpha\left(1 - \frac{\sum\limits_{i}^{10}\frac{w_{ice,i}}{w_{ice,i}+w_{liq,i}}\Delta z_i}{\sum\limits_{i}^{10}\Delta z_i}\right)\right] - \exp(-\alpha)\right\} \geq 0. \qquad \text{(B10)}$$

where $\alpha = 3$ is an adjustable scale-dependent parameter, i is the index of the layer directly above the water table, and $w_{ice,i}$ and $w_{liq,i}$ are the ice and liquid water contents of soil layer i (kg m^{-2}). This expression is functionally the same as

that used to determine the permeability of frozen soil (Appendix D).

[51] If the water table depth z_∇ is below the soil column, then the drainage q_{drai} is removed from the aquifer (Appendix C). If z_∇ is within the soil column, q_{drai} is extracted from the soil liquid water in soil layers within the water table (Appendix C). The value of $q_{drai,max}$ was determined from a calibration against the averaged observed water table depth for sixteen wells in Illinois [*Niu et al.*, 2007]. Future work will focus on optimizing spatially explicit values for parameters $q_{drai,max}$ and the decay factor f.

[52] Two numerical adjustments are implemented to keep the liquid water content of each soil layer ($w_{liq,i}$) within physical constraints $0.01 \leq w_{liq,i} \leq (\theta_{sat,i} - \theta_{ice,i})\Delta z_i$. These adjustments, $w_{liq}^{deficit}$ and w_{liq}^{excess}, may decrease or increase subsurface runoff, respectively. First, to help prevent negative $w_{liq,i}$, each layer is successively brought up to $w_{liq,i} = 0.01$ by taking the required amount of water from the layer below. If the total amount of water in the soil column is insufficient to accomplish this, the water is taken from the unconfined aquifer and the subsurface runoff ($w_{liq}^{deficit}$). Second, beginning with the bottom soil layer, any excess liquid water in each soil layer is successively added to the layer above. Any excess liquid water that remains after saturating the entire soil column (plus a maximum ponding depth $w_{liq}^{pond} = 10$ mm [*Oleson et al.*, 2004]), w_{liq}^{excess}, is added directly to subsurface runoff. These two adjustments are rarely necessary.

[53] Two other changes were made following *Niu et al.* [2005]. First, the exponentially decaying saturated hydraulic conductivity was removed and replaced with a conductivity that depends on soil texture alone [*Cosby et al.*, 1984]. The saturated hydraulic conductivity is

$$k_{sat} = 0.0070556 \times 10^{-0.884 + 0.0153(\%sand)_i} \qquad (B11)$$

where $(\%sand)_i$ is the sand content of the ith soil layer.

[54] Second, a no-drainage bottom boundary condition is imposed on the solution for soil moisture. Groundwater recharge (including gravitational drainage and upward flow driven by capillary forces) is now employed to update the bottom layer soil moisture (Appendix C). The coefficients of the tridiagonal set of equations for soil layer $i = 10$ [*Oleson et al.*, 2004, equations (7.105)–(7.108)] are now

$$a_i = -\frac{\partial q_{i-1}}{\partial \theta_{liq,i-1}} \qquad (B12)$$

$$b_i = -\left[\frac{\partial q_{i-1}}{\partial \theta_{liq,i}} + \frac{\Delta z_i}{\Delta t}\right] \qquad (B13)$$

$$c_i = 0 \qquad (B14)$$

$$r_i = e_i + q_{i-1}^n. \qquad (B15)$$

Appendix C: Groundwater and Water Table Depth

[55] The determination of the water table depth z_∇ is based on work by *Niu et al.* [2007]. In this approach, a groundwater component was added to CLM in the form of an unconfined aquifer lying below the model soil column (Figure C1). The groundwater solution is dependent on whether the water table is within or below the soil column. Two water stores are used to account for these solutions. The first, W_a, is the water stored in the unconfined aquifer (mm) and is proportional to the change in water table depth when the water table z_∇ is below the lower boundary of the model soil column (3.43 m). The second, W_t, is the actual groundwater which can be within the soil column. When the water table is below the soil column $W_t = W_a$. When the water table is within the soil column, W_a is constant because there is no water exchange between the soil column and the underlying aquifer, while W_t varies with soil moisture conditions. W_a, W_t, and z_∇ are prognostic variables within the model.

[56] For the case when the water table is below the soil column, the temporal variation of the water stored in the unconfined aquifer W_a (mm) is

$$\frac{dW_a}{dt} = q_{recharge} - q_{drai} \qquad (C1)$$

where $q_{recharge}$ is the recharge to the aquifer (kg m^{-2} s^{-1}) and the subsurface runoff q_{drai} is equivalent to the groundwater discharge. The recharge rate is derived from Darcy's law and is defined as positive when water enters the aquifer

$$q_{recharge} = -k_a \frac{-10^3 z_\nabla - (\psi_{10} - 10^3 z_{10})}{10^3(z_\nabla - z_{10})} \qquad (C2)$$

where k_a is the hydraulic conductivity of the aquifer (kg m^{-2} s^{-1}), z_∇ is the water table depth (m), and ψ_{10} is the matric potential of the bottom (10th) soil layer (mm) at the node depth $z_{10} = 2.865$ m. The matric potential of the bottom soil layer is determined from

$$\psi_{10} = \psi_{sat,10}(s_{10})^{-B_{10}} \qquad (C3)$$

where $\psi_{sat,10}$ and B_{10} are the saturated matric potential (mm) and soil texture-dependent *Clapp and Hornberger* [1978] exponent for the bottom soil layer. The wetness of the bottom soil layer $0.01 < s_{10} < 1$ is determined from the volumetric liquid water content $\theta_{liq,10}$ and effective porosity $\theta_{sat,10} - \theta_{ice,10} \geq 0$

$$s_{10} = \frac{\theta_{liq,10}}{\theta_{sat,10} - \theta_{ice,10}}. \qquad (C4)$$

[57] The hydraulic conductivity below the model soil column is assumed to decay with depth from the hydraulic
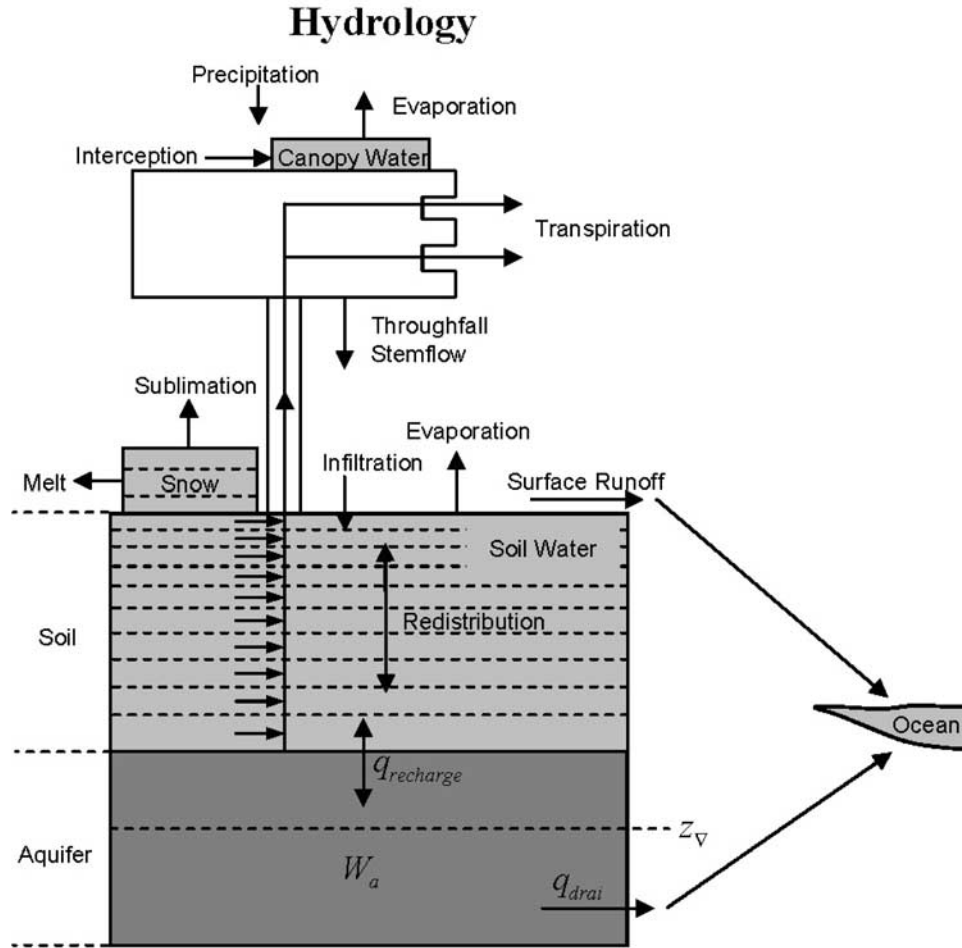
Hydrology

Figure C1. Hydrologic processes simulated by CLM3.5. An unconfined aquifer is added to the bottom of the CLM3 soil column [*Niu et al.*, 2007]. The depth to the water table is z_∇ (m). Changes in aquifer water content W_a (mm) are controlled by the balance between drainage from the aquifer water q_{drai} and the aquifer recharge rate $q_{recharge}$ (kg m^{-2} s^{-1}) (defined as positive from soil to aquifer).

conductivity of the bottom layer ($k_{10} \exp[-f(z - z_{10})]$). Thus, the hydraulic conductivity of the aquifer is

$$k_a = \frac{k_{10}\{1 - \exp[-f(z_\nabla - z_{10})]\}}{f(z_\nabla - z_{10})}. \quad \text{(C5)}$$

where k_{10} is the hydraulic conductivity of the bottom layer (kg m^{-2} s^{-1}).

[58] The water table depth is calculated from the aquifer water storage scaled by the average specific yield $S_y = 0.2$ (the fraction of water volume that can be drained by gravity in an unconfined aquifer)

$$z_\nabla = z_{h,10} + 25 - \frac{W_a}{10^3 S_y} \quad \text{(C6)}$$

where $z_{h,10}$ is the depth of the bottom of the soil column (3.43 m). The form of equation (C6) originates from the assumption that the initial amount of water in the aquifer is 4800 mm and the corresponding initial water table depth is 1 m below the bottom of the soil column. The water table depth is at the bottom of the soil column ($z_\nabla = z_{h,10}$)

when the aquifer water is at its prescribed maximum value (5000 mm). The change in soil water in the bottom layer is

$$\Delta w_{liq,10} = -q_{recharge}\Delta t + \max(0, W_a - 5000) \quad \text{(C7)}$$

where Δt is the model time step (s).

[59] For the case when the water table is within the model soil column, there is no water exchange between the soil column and the underlying aquifer. However, variations of the water table depth are still computed from equations (C1) and (C2), but the variables of the bottom layer are replaced with those of the layer directly above the water table. Hence,

$$\frac{dW_t}{dt} = q_{recharge} - q_{drai}. \quad \text{(C8)}$$

The recharge rate is

$$q_{recharge} = -k_i \frac{(\psi_{sat,i+1} - 10^3 z_\nabla) - (\psi_i - 10^3 z_i)}{10^3(z_\nabla - z_i)} \quad \text{(C9)}$$

where $\psi_{sat,i+1} - 10^3 z_\nabla$ is the water head at the water table depth and i is the index of the layer directly above the water table.

[60] In the work of *Niu et al.* [2007], the water table depth is computed from equation (C6) but with the specific yield determined by the volume of air pores (the pore space not filled with water) within the soil to convert W_t to water table depth. In preliminary global simulations we found that this approach resulted in unstable water table calculations for a significant proportion of grid cells. More specifically, when repeatedly forcing the model with a single year of atmospheric data, the temporal evolution of water table depth was significantly different from year to year for some grid cells, with occasional rapid (within a few days) movement of the water table depth to the soil surface in some cases. This occurred in grid cells with soil water contents near saturation because of the small amount of available pore space. This had deleterious implications for stability of surface fluxes and temperature. For example, we found that 4% of the grid cells would not satisfy the imposed spin up criterion (year to year change in annual mean surface fluxes less than 0.1 W m^{-2} [*Yang et al.*, 1995]). Here, we implement a calculation based on effective porosity only. Although less defensible from a physical viewpoint, the new approach stabilizes the water table calculation for these grid cells and eliminates unrealistic oscillations in surface fluxes and temperature. The spin up criterion is now satisfied for more than 99% of the grid cells in a global simulation within 30 years after starting the model from arbitrary initial conditions. The water table depth calculation is then

where Δt is the time step (s) and i is the index of the layer directly above the water table.

Appendix D: Frozen Soil

[61] The heat conduction equation is solved numerically to calculate the soil and snow temperatures for a ten-layer soil column with up to five overlying layers of snow [*Oleson et al.*, 2004]. The temperature profile is calculated first without phase change and then readjusted for phase change. Melting is still treated as in CLM3 [*Oleson et al.*, 2004]. Melting occurs if

$$T_i > T_f \quad \text{and} \quad w_{ice,i} > 0 \qquad (D1)$$

where T_i and $w_{ice,i}$ are the soil temperature (K) and ice content (kg m^{-2}) of layer i, and T_f is freezing temperature (K). The amount of ice that is melted is assessed from the energy needed to change T_i to T_f.

[62] For the freezing process, *Niu and Yang* [2006] incorporated the concept of supercooled soil water in CLM. The supercooled soil water is the liquid water that coexists with ice over a wide range of temperatures below freezing and is implemented through a freezing point depression equation

$$w_{liq,\max,i} = \Delta z_i \theta_{sat,i} \left[\frac{10^3 L_f (T_f - T_i)}{g T_i \psi_{sat,i}} \right]^{-1/B_i} \quad T_i < T_f \qquad (D2)$$

$$z_\nabla = \begin{cases} z_{h,i+1} - \left[\dfrac{W_t - 10^3 \times 25 S_y - \sum\limits_{j=i+2}^{10} \Delta z_j \left(\theta_{sat,j} - \theta_{ice,j} \right)}{10^3 \left(\theta_{sat,i+1} - \theta_{ice,i+1} \right)} \right] & 1 \le i \le 8 \\[4ex] z_{h,i+1} - \left[\dfrac{W_t - 10^3 \times 25 S_y}{10^3 \left(\theta_{sat,i+1} - \theta_{ice,i+1} \right)} \right] & i = 9 \end{cases} \qquad (C10)$$

where θ_{ice} is the volumetric ice content of a layer and i is the index of the layer directly above the water table. In this case, the subsurface runoff q_{drai} is extracted from the soil liquid water in layers within the water table instead of from the aquifer. The partitioning of discharge to these layers is proportional to the layer depth-weighted hydraulic conductivity as

$$\Delta \theta_j = -\frac{q_{drai} k_j \Delta t \Delta z_j}{\sum\limits_{j=i+1}^{10} k_j \Delta z_j} \quad j = i+1, 10 \qquad (C11)$$

where $w_{liq,\max,i}$ is the maximum liquid water in layer i (kg m^{-2}) when the soil temperature T_i is below the freezing temperature T_f, L_f is the latent heat of fusion (J kg^{-1}), and g is the gravitational acceleration (m s^{-2}). Freezing occurs if

$$T_i < T_f \quad \text{and} \quad w_{liq,i} > w_{liq,\max,i}. \qquad (D3)$$

The ice content at model step $n + 1$ is calculated from

$$w_{ice,i}^{n+1} = \begin{cases} \min\left(w_{liq,i}^n + w_{ice,i}^n - w_{liq,\max,i}^n, \ w_{ice,i}^n - \dfrac{H_i \Delta t}{L_f} \right) & w_{liq,i}^n + w_{ice,i}^n \ge w_{liq,\max,i}^n \\[3ex] 0 & w_{liq,i}^n + w_{ice,i}^n < w_{liq,\max,i}^n \end{cases} \qquad (D4)$$

Table E1. Soil Water Potential at Stomata Fully Open (ψ_{open}) and at Full Stomatal Closure (ψ_{close}) for Plant Functional Types

Plant Functional Type	ψ_{open}, $\times 10^5$ mm	ψ_{close}, $\times 10^5$ mm
Needleleaf evergreen tree–temperate[a]	−0.66	−2.55
Needleleaf evergreen tree–boreal[a]	−0.66	−2.55
Needleleaf deciduous tree–boreal[b]	−0.66	−2.55
Broadleaf evergreen tree–tropical[b]	−0.66	−2.55
Broadleaf evergreen tree–temperate[b]	−0.66	−2.55
Broadleaf deciduous tree–tropical[a]	−0.35	−2.24
Broadleaf deciduous tree–temperate[a]	−0.35	−2.24
Broadleaf deciduous tree–boreal[a]	−0.35	−2.24
Broadleaf evergreen shrub–temperate[a]	−0.83	−4.28
Broadleaf deciduous shrub–temperate[a]	−0.83	−4.28
Broadleaf deciduous shrub–boreal[a]	−0.83	−4.28
C_3 arctic grass[a]	−0.74	−2.75
C_3 grass[a]	−0.74	−2.75
C_4 grass[a]	−0.74	−2.75
Crop1[c]	−0.74	−2.75
Crop2[c,d]	−0.74	−2.75

[a]White et al. [2000].
[b]Assigned values of needleleaf evergreen tree.
[c]Assigned values of grass.
[d]Two types of crops are specified to account for the different physiology of crops, but currently only the first crop type is specified in the surface data set.

where H_i is the amount of energy needed to change T_i to T_f ($H_i < 0$) [Oleson et al., 2004]. Because part of the energy H_i may not be released in freezing, the energy is recalculated as

$$H_i* = H_i - \frac{L_f \left(w_{ice,i}^n - w_{ice,i}^{n+1} \right)}{\Delta t} \quad (D5)$$

and the energy H_{i*} is used to cool the soil layer.

[63] The impermeable fraction $f_{frz,i}$ (used in equation (B2) to determine the saturated fraction of the grid cell) is parameterized as a function of soil ice content [Niu and Yang, 2006]

$$f_{frz,i} = \left\{ \exp\left[-\alpha \left(1 - \frac{w_{ice,i}}{w_{ice,i} + w_{liq,i}} \right) \right] - \exp(-\alpha) \right\} \geq 0 \quad (D6)$$

where $\alpha = 3$ is an adjustable scale-dependent parameter, and $w_{ice,i}$ and $w_{liq,i}$ are the ice and liquid water contents of soil layer i (kg m^{-2}). The hydraulic properties of the soil are also modified. The hydraulic conductivity is defined at the depth of the interface of two adjacent layers $z_{h,i}$ (m) and is a function of the saturated hydraulic conductivity $k_{sat} [z_{h,i}]$, the volumetric soil moisture of the two layers, and the impermeable fraction

Appendix E: Soil Moisture Availability

[64] The effect of soil moisture stress on plant transpiration and photosynthesis is parameterized through a soil moisture limitation function acting on the leaf-scale maximum carboxlyation capacity of Rubisco [Thornton and Zimmermann, 2007]. The stress function is

$$\beta_t = \sum_i w_i r_i \quad (E1)$$

where w_i is a soil dryness or plant wilting factor for soil layer i and r_i is the fraction of roots in layer i [Oleson et al., 2004]. The plant wilting factor w_i is

$$w_i = \left\{ \begin{array}{ll} \left(\dfrac{\theta_{sat,i} - \theta_{ice,i}}{\theta_{sat,i}} \right) \left(\dfrac{\psi_i - \psi_{close}}{\psi_{open} - \psi_{close}} \right) \leq 1 & \theta_{liq,i} > 0 \\ 0 & \theta_{liq,i} = 0 \end{array} \right\} \quad (E2)$$

The soil water matric potential ψ_i (mm) is

$$\psi_i = \psi_{sat,i} s_i^{-B_i} \geq \psi_{close} \quad (E3)$$

Soil wetness s_i is defined as

$$s_i = \frac{\theta_{liq,i}}{\theta_{sat,i} - \theta_{ice,i}} \geq 0.01. \quad (E4)$$

where $\theta_{liq,i} \leq \theta_{sat,i} - \theta_{ice,i}$. The wilting point potential (full stomatal closure) ψ_{close} and the potential at which the stomata are fully open ψ_{open} (both in mm) are PFT-dependent parameters defined in Table E1.

Appendix F: Soil Evaporation

[65] For vegetated surfaces, the water vapor flux from the soil beneath the canopy E_g (kg m^{-2} s^{-1}) is

$$E_g = -\rho_{atm} \frac{(q_s - q_g)}{r'_{aw}} \quad (F1)$$

where q_s is the specific humidity of air at height $z_{0w} + d$ (the

$$k[z_{h,i}] = \left\{ \begin{array}{ll} \left[1 - 0.5\left(f_{frz,i} + f_{frz,i+1} \right) \right] k_{sat}[z_{h,i}] \left[\dfrac{0.5(\theta_i + \theta_{i+1})}{0.5(\theta_{sat,i} + \theta_{sat,i+1})} \right]^{2B_i+3} & 1 \leq i \leq 9 \\ \left(1 - f_{frz,i} \right) k_{sat}[z_{h,i}] \left[\dfrac{\theta_i}{\theta_{sat,i}} \right]^{2B_i+3} & i = 10 \end{array} \right\} \quad (D7)$$

where θ is the total (ice plus liquid) volumetric soil moisture. The soil matric potential is determined from the liquid water content as

$$\psi_i = \psi_{sat,i} \left(\frac{\theta_i}{\theta_{sat,i}} \right)^{-B_i} \geq -1 \times 10^8 \quad 0.01 \leq \frac{\theta_i}{\theta_{sat,i}} \leq 1 \quad (D8)$$

canopy air specific humidity), and r'_{aw} is the aerodynamic resistance (s m^{-1}) to water vapor transfer between the ground at height z'_{0w} and the canopy air at height $z_{0w} + d$ (water vapor roughness length plus displacement height (m)).

Table G1. Nitrogen Limitation Factor for Plant Functional Types

Plant Functional Type	f(N)
Needleleaf evergreen tree–temperate	0.72
Needleleaf evergreen tree–boreal	0.78
Needleleaf deciduous tree–boreal	0.79
Broadleaf evergreen tree–tropical	0.83
Broadleaf evergreen tree–temperate	0.71
Broadleaf deciduous tree–tropical	0.66
Broadleaf deciduous tree–temperate	0.64
Broadleaf deciduous tree–boreal	0.70
Broadleaf evergreen shrub–temperate	0.62
Broadleaf deciduous shrub–temperate	0.60
Broadleaf deciduous shrub–boreal	0.76
C_3 arctic grass	0.68
C_3 grass	0.61
C_4 grass	0.64
Crop1	0.61
Crop2	0.61

[66] The specific humidity of the soil surface q_g is assumed to be proportional to the saturation specific humidity

$$q_g = \alpha q_{sat}^{T_g} \qquad (F2)$$

where $q_{sat}^{T_g}$ is the saturated specific humidity at the ground surface temperature T_g. The factor α is a weighted combination of values for soil and snow

$$\alpha = \alpha_{soi,1}(1 - f_{sno}) + \alpha_{sno}f_{sno} \qquad (F3)$$

where f_{sno} is the fraction of ground covered by snow, and $\alpha_{sno} = 1.0$. $\alpha_{soi,1}$ refers to the surface soil layer and is a function of the surface soil water matric potential ψ as in *Philip* [1957]

$$\alpha_{soi,1} = \exp\left(\frac{\psi_1 g}{10^3 R_{wv} T_g}\right) \qquad (F4)$$

where R_{wv} is the gas constant for water vapor (J kg^{-1} K^{-1}), g is the gravitational acceleration (m s^{-2}), and ψ_1 is the soil water matric potential of the top soil layer (kg m^{-2}).

[67] The term α is supposed to be the air relative humidity at the humidity roughness height z'_{0w}. As pointed out by *Kondo et al.* [1990] and *Wetzel and Chang* [1987], however, the term frequently used is the relative humidity of the air adjacent to the water in the soil pore (i.e., the relationship from *Philip* [1957] is used in CLM), which is not the same as α. Some studies have found that the resistance to water vapor transport by molecular diffusion from the water surface in the soil pores to the soil surface needs to be accounted for, even for thin soil layers [*Kondo et al.*, 1990; *Lee and Pielke*, 1992; *Wu et al.*, 2000]. Indeed, in our own global and point simulations, we found excessive soil evaporation (not shown). To account for this, we added an additional soil resistance term R_{soil} based on work by *Sellers et al.* [1992]

$$R_{soil} = (1 - f_{sno})\exp(8.206 - 4.255 s_1) \qquad (F5)$$

where f_{sno} is the fractional soil covered by snow and s_1 is the soil moisture of the top layer relative to saturation determined from

$$s_1 = \frac{\theta_{ice,1} + \theta_{liq,1}}{\theta_{sat,1}} \leq 1 \qquad (F6)$$

where $\theta_{ice,1}$, $\theta_{liq,1}$, and $\theta_{sat,1}$ are the volumetric ice, liquid water, and saturation water contents. R_{soil} is set to zero in the case of dewfall.

Appendix G: Nitrogen Limitation

[68] PFT-dependent scale factors to represent nitrogen limitations on plant productivity were derived from a simulation with CLM coupled to a carbon/nitrogen cycle (CLM-CN [*Thornton and Zimmermann*, 2007; *Thornton et al.*, 2007]). The factor, f(N), represents the proportion of potential photosynthesis (gross primary production, or GPP) that is realized in the face of nitrogen limitation, as predicted by CLM-CN, for each PFT (Table G1). The simulation from which these factors are derived is a fully spun-up preindustrial state, driven by 25-year cyclic NCEP drivers (1949–1972). The f(N) is imposed on the maximum rate of carboxylation V_{max} in a manner similar to plant water stress [*Oleson et al.*, 2004] as

$$V_{max} = V_{max\,25}(a_{v\,max})^{\frac{T_v - 25}{10}}f(T_v)\beta_t f(N) \qquad (G1)$$

where $V_{max\,25}$ is the value at 25°C (μ mol CO_2 m^{-2} s^{-1}), $a_{v\,max}$ is the Q_{10} parameter, T_v is leaf temperature (C), $f(T_v)$ is a function that mimics thermal breakdown of metabolic processes, and β_t is a soil moisture stress function.

References

Baldocchi, D., et al. (2001), FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, *Bull. Am. Meteorol. Soc.*, *82*, 2415–2433.

Beven, K. J., and M. J. Kirkby (1979), A physically based, variable contributing model of basin hydrology, *Hydrol. Sci. Bull.*, *24*, 43–69.

Bonan, G. B., and S. Levis (2006), Evaluating aspects of the Community Land and Atmosphere Models (CLM3 and CAM3) using a dynamic global vegetation model, *J. Clim.*, *19*, 2290–2301.

Bonan, G. B., S. Levis, L. Kergoat, and K. W. Oleson (2002), Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models, *Global Biogeochem. Cycles*, *16*(2), 1021, doi:10.1029/2000GB001360.

Boone, A., et al. (2004), The Rhone-Aggregation land surface scheme intercomparison project: An overview, *J. Clim.*, *17*, 187–208.

Brutsaert, W. (1984), *Evaporation Into the Atmosphere*, 299 pp., Kluwer Acad., Dordecht, Netherlands.

Chen, J. L., M. Rodell, C. R. Wilson, and J. S. Famiglietti (2005), Low degree spherical harmonic influences on Gravity Recovery and Climate Experiment (GRACE) water storage estimates, *Geophys. Res. Lett.*, *32*, L14405, doi:10.1029/2005GL022964.

Choudhury, B. J., N. E. DiGirolamo, J. Susskind, W. L. Darnell, S. K. Gupta, and G. Asrar (1998), A biophysical process-based estimate of

global land surface evaporation using satellite and ancillary data II. Regional and global patterns of seasonal and annual variations, *J. Hydrol.*, *205*, 186–204.

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

Collins, W. D., et al. (2006), The Community Climate System Model version 3 (CCSM3), *J. Clim.*, *19*, 2122–2143.

Cosby, B. J., G. M. Hornberger, R. B. Clapp, 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.

Dai, A., and K. E. Trenberth (2002), Estimates of freshwater discharge from continents: Latitudinal and seasonal variations, *J. Hydrometeorol.*, *3*, 660–687.

Dai, A., K. E. Trenberth, and T. Qian (2004), A global data set of Palmer Drought Severity Index for 1870–2002: Relationship with soil moisture and effects of surface warming, *J. Hydrometeorol.*, *5*, 1117–1130.

Decharme, B., and H. Douville (2007), Global validation of the ISBA subgrid hydrology, *Clim. Dyn.*, *29*, 21–37, doi:10.1007/s00382-006-0216-7.

Decker, M., and X. Zeng (2006), An empirical formulation of soil ice fraction based on in situ observations, *Geophys. Res. Lett.*, *33*, L05402, doi:10.1029/2005GL024914.

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*, Natl. Cent. for Atmos. Res., Boulder, Colo.

Dickinson, R. E., K. W. Oleson, G. Bonan, F. Hoffman, P. Thornton, M. Vertenstein, Z.-L. Yang, and X. Zeng (2006), The Community Land Model and its climate statistics as a component of the Community Climate System Model, *J. Clim.*, *19*, 2302–2324.

Dirmeyer, P. A., X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki (2006), GSWP2: Multimodel analysis and implications for our perception of the land surface, *Bull. Am. Meteorol. Soc.*, *87*, 1381–1397.

Entekhabi, D., and P. S. Eagleson (1989), Land surface hydrology parameterization for atmospheric general circulation models including subgrid scale spatial variability, *J. Clim.*, *2*, 816–831.

Entin, J. K., A. Robock, K. Y. Vinnikov, V. Zabelin, S. Liu, A. Namkhai, and T. Adyasuren (1999), Evaluation of global soil wetness project soil moisture simulations, *J. Meteorol. Soc. Jpn.*, *77*, 183–198.

Fekete, B. M., C. J. Vörösmarty, and W. Grabs (2000), Global composite runoff fields based on observed river discharge and simulated water balances (version 1.0). (Available at http://www.grdc.sr.unh.edu)

Fekete, B. M., C. J. Vörösmarty, and W. Grabs (2002), High-resolution fields of global runoff combining observed river discharge and simulated water balances, *Global Biogeochem. Cycles*, *16*(3), 1042, doi:10.1029/1999GB001254.

Fu, R., and W. Li (2004), The influence of the land surface on the transition from dry to wet season in Amazonia, *Theor. Appl. Climatol.*, *78*, 97–110.

Gulden, L. E., E. Rosero, Z.-L. Yang, M. Rodell, C. S. Jackson, G.-Y. Niu, J.-F. P. Yeh, and J. Famiglietti (2007), Improving land-surface model hydrology: Is an explicit aquifer model better than a deeper soil profile?, *Geophys. Res. Lett.*, *34*, L09402, doi:10.1029/2007GL029804.

Guo, Z., and P. A. Dirmeyer (2006), Evaluation of the second global soil wetness project soil moisture simulations: 1. intermodel comparison, *J. Geophys. Res.*, *111*, D22S02, doi:10.1029/2006JD007233.

Hack, J. J., J. M. Caron, S. G. Yeager, K. W. Oleson, M. M. Holland, J. E. Truesdale, and P. J. Rasch (2006), Simulation of the global hydrological cycle in the CCSM Community Atmosphere Model version 3 (CAM3): mean features, *J. Clim.*, *19*, 2199–2221.

Hollinger, S. E., and S. A. Isard (1994), A soil moisture climatology of Illinois, *J. Clim.*, *7*, 822–833.

Juárez, R. I. N., M. G. Hodnett, R. Fu, M. L. Goulden, and C. von Randow (2007), Control of dry season evapotranspiration over the amazonian forest as inferred from observations at a southern amazon forest site, *J. Clim.*, *20*, 2827–2839.

Kondo, J., N. Saigusa, and T. Sato (1990), A parameterization of evaporation from bare soil surfaces, *J. Appl. Meteorol.*, *29*, 385–389.

Lawrence, D. M., P. E. Thornton, K. W. Oleson, and G. B. Bonan (2007), The partitioning of evapotranspiration into transpiration, soil evaporation, and canopy evaporation in a GCM: Impacts on land-atmosphere interaction, *J. Hydrometeorol.*, *8*, 862–880.

Lawrence, P. J., and T. N. Chase (2007), Representing a new MODIS consistent land surface in the Community Land Model (CLM3.0), *J. Geophys. Res.*, *112*, G01023, doi:10.1029/2006JG000168.

Lee, T. J., and R. A. Pielke (1992), Estimating the soil surface specific humidity, *J. Appl. Meteorol.*, *31*, 480–484.

Levis, S., G. B. Bonan, M. Vertenstein, and K. W. Oleson (2004), The Community Land Model's Dynamic Global Vegetation Model (CLM-DGVM): Technical description and user's guide, 50 pp., *NCAR Tech. Note NCAR/TN-459+IA*, Natl. Cent. for Atmos. Res., Boulder, Colo.

Malhi, Y., E. Pegoraro, A. D. Nobre, G. P. Pereira, J. Grace, A. D. Culf, and R. Clement (2002), Energy and water dynamics of a central Amazonian rain forest, *J. Geophys. Res.*, *107*(D20), 8061, doi:10.1029/2001JD000623.

Marin, C. T., W. Bouten, and J. Sevink (2000), Gross rainfall and its partitioning into throughfall, stemflow and evaporation of intercepted water in four forest ecosystems in western Amazonia, *J. Hydrol.*, *237*, 40–57.

Matsuyama, H. (1992), The water budget of the Amazon river basin during the FGGE period, *J. Meteorol. Soc. Jpn.*, *70*, 1071–1083.

Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models. 1. A discussion of principles, *J. Hydrol.*, *10*, 282–290.

Nepstad, D. C., et al. (2002), The effects of partial throughfall exclusion on canopy processes, aboveground production, and biogeochemistry of an Amazon forest, *J. Geophys. Res.*, *107*(D20), 8085, doi:10.1029/2001JD000360.

Niu, G.-Y., and Z.-L. Yang (2006), Effects of frozen soil on snowmelt runoff and soil water storage at a continental scale, *J. Hydrometeorol.*, *7*, 937–952.

Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, and L. E. Gulden (2005), A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, *J. Geophys. Res.*, *110*, D21106, doi:10.1029/2005JD006111.

Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, L. E. Gulden, and H. Su (2007), Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, *J. Geophys. Res.*, *112*, D07103, doi:10.1029/2006JD007522.

Oleson, K. W., et al. (2004), Technical description of the Community Land Model (CLM), *NCAR Tech. Note NCAR/TN-461+STR*, 173 pp., Natl. Cent. for Atmos. Res., Boulder, Colo.

Philip, J. R. (1957), Evaporation, and moisture and heat fields in the soil, *J. Meteorol.*, *14*, 354–366.

Qian, T., A. Dai, K. E. Trenberth, and K. W. Oleson (2006), Simulation of global land surface conditions from 1948 to 2004: Part I: Forcing data and evaluations, *J. Hydrometeorol.*, *7*, 953–975.

Robock, A., K. Y. Vinnikov, G. Srinivasan, J. K. Entin, S. E. Hollinger, N. A. Speranskaya, S. Liu, and A. Namkhai (2000), The Global Soil Moisture Data Bank, *Bull. Am. Meteorol. Soc.*, *81*, 1281–1299.

Sellers, P. J., M. D. Heiser, and F. G. Hall (1992), Relations between surface conductance and spectral vegetation indices at intermediate (100 m^2 to 15 km^2) length scales, *J. Geophys. Res.*, *97*(D17), 19,033–19,059.

Seo, K. W., and C. R. Wilson (2005), Simulated estimation of hydrological loads from GRACE, *J. Geod.*, *78*, 442–456.

Shuttleworth, W. J. (1988), Evaporation from Amazonian rainforest, *Proc. R. Soc. London, Ser. B*, *233*, 321–346.

Stöckli, R., D. M. Lawrence, G.-Y. Niu, K. W. Oleson, P. E. Thornton, Z.-L. Yang, G. B. Bonan, A. S. Denning, and S. W. Running (2008), Use of Fluxnet in the Community Land Model development, *J. Geophys. Res.*, doi:10.1029/2007JG000562, in press.

Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, *106*(D7), 7183–7192.

Thornton, P. E., and N. E. Zimmermann (2007), An improved canopy integration scheme for a land surface model with prognostic canopy structure, *J. Clim.*, *20*, 3902–3923.

Thornton, P. E., J.-F. Lamarque, N. A. Rosenbloom, and N. Mahowald (2007), Influence of carbon-nitrogen cycle coupling on land model response to CO_2 fertilization and climate variability, *Global Biogeochem. Cycles*, *21*, GB4018, doi:10.1029/2006GB002868.

Ubarana, V. N. (1996), Observations and modelling of rainfall interception at two experimental sites in Amazonia, in *Amazonian Deforestation and Climate*, edited by J. H. C. Gash et al., Wiley, Chickester, UK.

Vinnikov, K., and I. B. Yeserkepova (1991), Soil moisture: empirical data and model results, *J. Clim.*, *4*, 66–79.

Wetzel, P. J., and J. T. Chang (1987), Concerning the relationship between evaporation and soil moisture, *J. Clim. Appl. Meteorol.*, *26*, 18–27.

White, M. A., P. E. Thornton, S. W. Running, and R. R. Nemani (2000), Parameterization and sensitivity analysis of the BIOME-BGC terrestrial ecosystem model: net primary production controls, *Earth Interactions*, *4*, 1–85.

Wu, A., T. A. Black, D. L. Verseghy, M. D. Novak, and W. G. Bailey (2000), Testing the α and β methods of estimating evaporation from bare and vegetated surfaces in CLASS, *Atmos. Ocean*, *38*, 15–35.

Yang, Z.-L., R. E. Dickinson, A. Henderson-Sellers, and A. J. Pitman (1995), Preliminary study of spin-up processes in land surface models with the first stage data of Project for Intercomparison of Land Surface Parameterization Schemes Phase 1(a), *J. Geophys. Res.*, *100*, 16,553–16,578.

Yi, S., M. A. Arain, and M.-K. Woo (2006), Modifications of a land surface scheme for improved simulation of ground freeze-thaw in northern en-

vironments, *Geophys. Res. Lett.*, *33*, L13501, doi:10.1029/2006GL026340.

—————————————

G. B. Bonan, A. Dai, D. M. Lawrence, S. Levis, K. W. Oleson, T. Qian, and P. E. Thornton, Climate and Global Dynamics Division, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000, USA. (oleson@ucar.edu)

R. E. Dickinson, Department of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA.

P. J. Lawrence, Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309, USA.

G.-Y. Niu and Z.-L. Yang, Department of Geological Sciences, University of Texas at Austin, Austin, TX 78712, USA.

R. Stöckli, Department of Atmospheric Science, Colorado State University, Fort Collins, CO 80523, USA.