The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins

Zong-Liang Yang,[1] Guo-Yue Niu,[1,2] Kenneth E. Mitchell,[3] Fei Chen,[4] Michael B. Ek,[3] Michael Barlage,[4] Laurent Longuevergne,[5] Kevin Manning,[4] Dev Niyogi,[6] Mukul Tewari,[4] and Youlong Xia[3]

[1] The augmented Noah land surface model described in the first part of the two-part series was evaluated here over global river basins. Across various climate zones, global-scale tests can reveal a model's weaknesses and strengths that a local-scale testing cannot. In addition, global-scale tests are more challenging than local- and catchment-scale tests. Given constant model parameters (e. g., runoff parameters) across global river basins, global-scale tests are more stringent. We assessed model performance against various satellite and ground-based observations over global river basins through six experiments that mimic a transition from the original Noah LSM to the fully augmented version. The model shows transitional improvements in modeling runoff, soil moisture, snow, and skin temperature, despite considerable increase in computational time by the fully augmented Noah-MP version compared to the original Noah LSM. The dynamic vegetation model favorably captures seasonal and spatial variability of leaf area index and green vegetation fraction. We also conducted 36 ensemble experiments with 36 combinations of optional schemes for runoff, leaf dynamics, stomatal resistance, and the β factor. Runoff schemes play a dominant and different role in controlling soil moisture and its relationship with evapotranspiration compared to ecological processes such as the β factor, vegetation dynamics, and stomatal resistance. The 36-member ensemble mean of runoff performs better than any single member over the world's 50 largest river basins, suggesting a great potential of land-based ensemble simulations for climate prediction.

Citation: Yang, Z.-L., G.-Y. Niu, K. E. Mitchell, F. Chen, M. B. Ek, M. Barlage, L. Longuevergne, K. Manning, D. Niyogi, M. Tewari, and Y. Xia (2011), The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins, *J. Geophys. Res.*, *116*, D12110, doi:10.1029/2010JD015140.

1. Introduction

[2] The Noah land surface model (LSM) has been used not only in short-term weather forecasting, but also in intra-seasonal to interannual climate predictions and downscaling of global climate model (GCM) projections. Its coupling with Global Forecast System (GFS) for operational weather and climate predictions by the National Centers for Environ-mental Prediction (NCEP) requires testing the Noah LSM at a global scale and a longer time scale. Most of previous Noah testing efforts have been focused at local scales [e.g., *Chen et al.*, 1997] (hereinafter Chen97) or continental U.S. at a shorter time scale [*Ek et al.*, 2003; *Mitchell et al.*, 2004]. While *Niu et al.* [2011] describe the augmented Noah LSM with multiple parameterization options (Noah-MP) and the testing results at local scales and a shorter time scale, this second paper focuses on testing Noah-MP and a recent version of the Noah LSM (Noah V3) at a global scale. Across various climate zones, global-scale testing can reveal models' weaknesses and strengths that a local-scale testing cannot. For instance, many previous local- and regional-scale studies revealed that Noah V3 produces a shallower snowpack [*Ek et al.*, 2003; *Pan et al.*, 2003; *Mitchell et al.*, 2004; *Livneh et al.*, 2010]. But as we will show in section 3.4, Noah V3 does not always produce shallower snow everywhere. It simulates too much snow in high-latitude river basins from forming excessive frost on the snow surface. In addition, global-scale tests are more challenging than local- and regional-scale tests. Given constant model parameters (e.g.,

[1]Department of Geological Sciences, John A. and Katherine G. Jackson School of Geosciences, University of Texas at Austin, Austin, Texas, USA.

[2]Now at Biosphere 2, University of Arizona, Tucson, Arizona, USA.

[3]Environmental Modeling Center, National Centers for Environmental Prediction, National Oceanic and Atmospheric Administration–National Weather Service, Camp Springs, Maryland, USA.

[4]Research Applications Laboratory, National Center for Atmospheric Research, Boulder, Colorado, USA.

[5]Bureau of Economic Geology, John A. and Katherine G. Jackson School of Geosciences, University of Texas at Austin, Austin, Texas, USA.

[6]Departments of Agronomy and Earth and Atmospheric Sciences, Purdue University, West Lafayette, Indiana, USA.

runoff parameters) across global river basins, global-scale tests are more stringent.

[3] Noah-MP increases the model complexity over Noah V3 due mainly to the model structural change, i.e., separating the vegetation canopy from the ground surface. But for global applications, *these changes are necessary* to solve the critical problems in Noah V3, such as, overestimation of runoff, too dry deep soil (due to free drainage), too impermeable frozen soil, and too rapid ablation of snow, etc. To solve these problems, one can either use simple ways, e.g., modifying a single formulation for parameterizing a physical process or calibrating model parameters, or pursuing more complex, physically based approaches. The former may improve a model's performance for a specific application, a specific variable, or a specific local site, but not universally for all variables and global applications. Of course, these more physically based schemes would involve new parameters and optimization of these new parameters. However, a more physically based scheme can alter the sensitivity behavior of a model to its parameters, facilitating optimization of model parameters. For instance, the snow scheme in Noah-MP is much more sensitive to model parameters that alter snow surface energy budgets than did Noah V3 [*Niu et al.*, 2011]. For instance, one of the major causes of the rapid snowmelt problem in Noah V3 is the inaccurate simulation of snow temperature, resulting in too long a duration of snow temperature being at melting point and lower cold content [*Niu et al.*, 2011]. Noah V3's surface layer is actually a bulk layer of topsoil, snow, and the vegetation canopy. For such a layer structure, it is difficult to accurately compute snow temperature and the ground heat flux, which is regarded as the forcing of the thermal state of the snowpack. Thus, Noah-MP first modified the model's structure for more accurately computing surface energy balance that controls the ground heat flux and snow temperature. The solution improves snow simulation globally both at midlatitudes and high latitudes as we will show in section 3.4.

[4] The multiple options of schemes for various processes in Noah-MP enable us to explore multimodel ensemble simulations. Model-based ensemble simulations are different from parameter-based ensemble simulations in two aspects: (1) parameter-based ensembles are model-dependent; ensembles executed by perturbing parameters in a simple model (e.g., a bucket model) may not embrace ensembles based on a more complex model (e.g., models including vegetation dynamics) and (2) perturbing parameters in a linear formulation (e.g., the Noah β factor for stomatal resistance shown in Figure 1 of *Niu et al.* [2011]) may be different from perturbing parameters in a nonlinear formulation (e.g., the SSiB β factor). Model-based ensembles may provide a wider range of ensembles, if combined with parameter-based ensembles. But the total numbers of ensembles would be super-large and not practical.

[5] The main purpose of this paper is to assess if Noah-MP that includes the proposed solutions (or augmentations) to the critical problems for hydrological simulations in Noah V3 can be efficient globally. At the same time, we preliminarily explore if ensembles of a limited number of selected models can improve runoff simulations over global 50 river basins. Section 2 describes the data sets used in this paper. Section 3 shows the model's improvements over the world's largest river basins. In section 4, we use a 36-member ensemble

simulation as an example to show the multiphysics model's potential for identification of critical processes and ensemble simulation of global runoff. Section 5 summarizes the paper.

2. Data Sets

[6] The data sets used in this study include global atmospheric forcing, vegetation and soil types, satellites retrievals of green vegetation fraction (GVF), leaf area index (LAI), and terrestrial water storage (TWS) change, ground-based measurements of soil moisture over Illinois, global runoff estimates, and North America snow water equivalent (SWE) and snow depth estimates.

2.1. Atmospheric Forcing

[7] We used a $1° \times 1°$, 3-hourly, near-surface meteorological data set processed by the Global Land Data Assimilation System (GLDAS) [*Rodell et al.*, 2004] to drive the model during the period 1980–2006. These include a spatially and temporally downscaled version of NOAA Climate Predictions Center's Merged Analysis of Precipitation (CMAP) [*Xie and Arkin*, 1997] and satellite based shortwave and longwave radiation. Other forcing fields are air temperature, air pressure, specific humidity, and wind speed based on NCEP reanalysis.

2.2. Vegetation and Soil Data

[8] We used the USGS 30 arc-second global vegetation type, the hybrid STATSGO/FAO soil texture data sets, and 0.144° monthly 5 year GVF climatology derived by *Gutman and Ignatov* [1998] of National Environmental Satellite, Data, and Information Service (NESDIS) from AVHRR-based NDVI as model inputs. We used 1° annual mean 2 m air temperature as the lower boundary condition for soil temperature. All the data sets are maintained by the Research Application Laboratory (RAL) of NCAR and are available at http://www.rap.ucar.edu/research/land/technology/lsm.php.

[9] We upscaled the high resolution vegetation and soil type data to 1°, as consistent with the atmospheric forcing data, using most predominant types. We aggregated the 0.144° monthly GVF data to 1° data using area-weighted average. Because the GVF data over boreal forest in wintertime are unreasonably low (K. Mitchell, personal communication, 2007), we replaced the wintertime GVF data over boreal forest grids with their annual mean values.

2.3. Illinois Soil Moisture

[10] We used soil moisture data from Illinois, USA [*Hollinger and Isard*, 1994] for its continuity and high quality. This data set consists of total soil moisture measured at 19 stations from 1981 to June, 2004, measured with the neutron probe technique and calibrated with gravimetric observations. The data are measured at 11 soil layers, the top 10 cm and then every 20 cm layer (e.g., 10–30 cm, 30–50 cm, etc.) down to a depth of 2 m. The data set is accessible through the Global Soil Moisture Databank (GSMDB) [*Robock et al.*, 2000] at http://www.ipf.tuwien.ac.at/insitu/. We aggregated the data into two layers, top 1 m and top 2 m, over the entire State of Illinois at a monthly interval using all the 19 stations data whenever the data were available at a station.

2.4. Grace Terrestrial, Water Storage Change

[11] We used the RL04 GRACE solutions released by the Center for Space Research (CSR), The University of Texas at Austin in this study to evaluate model performance in estimating terrestrial water storage change. GRACE terrestrial water storage change (ΔS) data were filtered following *Chen et al.* [2006] by first removing longitudinal stripes and then applying a 500 km Gaussian filter. To avoid misinterpretation due to leakage error, which is significant over coastal river basins and those whose adjacent river basins show opposite ΔS variations, we filtered model outputs in the same way. The accuracy of GRACE ΔS is mainly related to the atmospheric and oceanic mass change corrections produced by numerical models [*Seo et al.*, 2006]. A recent study by *Schmidt et al.* [2008] showed that RL04 improved its accuracy as compared with the first release of GRACE solutions, by a factor of 2, reducing the error to 24–30 mm.

2.5. MODIS LAI

[12] We used the Boston University MODIS (version 5) 0.25° monthly LAI products (2000 March to 2008 May) [*Myneni et al.*, 2002] to evaluate the model's ability to predict LAI. The LAI product is retrieved from surface reflectance measured by the moderate resolution imaging spectroradiometer (MODIS) aboard the TERRA platform, based on the physics of radiative transfer in vegetation canopies. The LAI products were evaluated at several field sites which are representative of six structural biomes [*Myneni et al.*, 2002]. The data set is available at ftp://primavera.bu.edu/pub/datasets/MODIS/MOD15_BU/C5/LAI/data/. We aggregated the 0.25° data to 1°.

2.6. University of New Hampshire–Global Runoff Data Center Runoff

[13] We used the University of New Hampshire–Global Runoff Data Center (UNH-GRDC) data set. It provides monthly gridded climatological runoff composite fields, which are runoff outputs from a water balance model that is driven by observed meteorological data. The total amount of estimated runoff is then corrected with runoff fields that are disaggregated from the observed river discharges. Although a no-time delay assumption is applied when the gauge-observed discharge is distributed uniformly over a catchment, the resulting runoff fields over a large river basin approximate the real runoff, especially in river basins that contain a sufficiently dense network of streamflow gauges. The UNH-GRDC data set preserves the accuracy of the observed discharge measurements and maintains the spatial and temporal distributions of the simulated runoff, thereby providing the "best estimate" of terrestrial runoff over large domains [*Fekete et al.*, 2002; B. M. Fekete, C. J. Vörösmarty, and W. Grabs, Global composite runoff fields based on observed river discharge and simulated water balances (version 1.0), 2000, http://www.grdc.sr.unh.edu].

2.7. Canadian Meteorology Center Snow Depth and SWE

[14] We used the daily snow depth and SWE (1979–1996) at 0.25° resolution over North America (NA) for AMIP-2 (Second Atmospheric Model Intercomparison Project) [*Brown et al.*, 2003]. The data set is maintained at the Canadian Meteorology Center (CMC). The gridded snow depth combines in situ daily observations from about 8,000 U.S. cooperative stations and Canadian climate stations and 'first-guess' fields with an optimum interpolation scheme, which is employed operationally at CMC. The 'first-guess' fields were produced by a simple snow accumulation, aging and melt model driven by 6-hourly air temperature and precipitation from the European Centre for Medium-range Weather Forecasting (ECMWF) ERA-15 Reanalysis with extensions from the Tropical Ocean and Global Atmospheric Program (TOGA) operational data archive. The gridded snow depth and estimated SWE agreed well with available independent in situ and satellite data over midlatitude regions.

3. Improvements Over the Baseline Model

[15] We conducted one experiment with the baseline model, Noah V3, and six experiments with six different combinations of optional schemes described in Table 1. We designed these experiments to mimic transitional improvements from Noah V3 to the fully augmented version that includes all the modifications described by *Niu et al.* [2011]. For all the six experiments, we selected the same options for the following four processes: (1) the β factor using the Noah type, (2) radiation transfer using the modified two-stream approximation, (3) snow surface albedo using the CLASS type, and (4) partitioning precipitation into rainfall and snowfall using *Jordan*'s [1991] scheme. The experiments differ in options for other six processes. Table 1 lists the seven experiments and their corresponding selected schemes. All the experiments used the same soil type and vegetation type data sets (in section 3.2) to determine soil thermal and hydraulic parameters and vegetation optical and physiological parameters that are defined in look-up tables. EXP6 with dynamic vegetation scheme predicted LAI and GVF, while the rest used prescribed monthly LAI and NESDIS GVF data. All the seven experiments were run for 27 years (1980–2006) driven by the GLDAS forcing data at 1° resolution. All the experiments were initialized with the same arbitrary, globally constant thermal and hydrological conditions. Vegetation canopy temperature and ground surface temperature are initialized at 287 K. The 4 L soil temperatures, from the top to the bottom, are 274, 278, 282, and 286K respectively. The 4 L soil moistures are 60% of porosity. The depth to water table was initialized at 2.5 m. Snow depth and SWE are 0.0 m, although the integration was started in winter. For EXP6, leaf and stem carbon mass are both initialized at 50 g m^{-2}. Because we focused on short-term leaf dynamics, carbon mass stored in wood, root, and soil is given arbitrarily at a relatively low level. The leaf mass can reach an equilibrium state after a growing season. Thus, the spin-up time for leaf mass and LAI is not a major concern. Considering the spin-up time for soil moisture and groundwater is much longer, we excluded the first 3 year (1980–1982) results in the following analyses (note that the use of K_{bot} in *Niu et al.*'s [2007] equation (7) instead of K_a greatly reduces the time to reach an equilibrium state of soil moisture and groundwater). We averaged model outputs from 1983–2006 to form a basic climatology, while varying the time period to average model outputs to match the periods over which observational data are available (e.g., for snow, the averaging period is 1983–1996).

Table 1. Experiments With Different Combinations of Schemes[a]

	$\theta_{liq\ max,i}$	Frozen soil permeability	C_H	Runoff	r_s	Leaf Dynamics
Noah V3	Koren99	Koren99	Chen97	Schaake96	Jarvis	Off
EXP 1[b]	Koren99	Koren99	Chen97	Schaake96	Jarvis	Off
EXP 2	NY06	NY06	Chen97	Schaake96	Jarvis	Off
EXP 3	NY06	NY06	M-O	Schaake96	Jarvis	Off
EXP 4	NY06	NY06	M-O	SIMGM	Jarvis	Off
EXP 5	NY06	NY06	M-O	SIMGM	Ball-Berry	Off
EXP 6	NY06	NY06	M-O	SIMGM	Ball-Berry	On

[a]Koren99, *Koren et al.* [1999]; NY 06, *Niu and Yang* [2006]; Chen97, *Chen et al.* [1997]; Schaake96, *Schaake et al.* [1996].
[b]Although using the same selected processes, EXP1 differs from Noah V3 in many other aspects, such as shortwave and longwave radiation schemes, sensible and latent heat flux formulations, and the skin temperature solution.

3.1. Global Surface Energy and Water Budgets

[16] Table 2 presents global averages of land surface energy and water budgets. Experiments with the augmented version (EXP1– EXP6) produced greater net solar radiation, total net radiation, latent heat, ET, and less runoff than Noah V3 does (Table 2). EXP6 (with dynamic vegetation) produces 3 to 4 W m^{-2} lower net solar radiation than other five runs with prescribed phonology, mainly because EXP6 produces lower LAI and GVF over boreal forest regions in wintertime, exposing more snow surface to solar radiation. Table 2 also includes GSWP phase-2 (GSWP2; baseline run) 12-model mean values of surface energy and water budgets as a reference. The results from EXP1 to EXP6 are comparable to the surface fluxes from GWSP2, although the precipitation used in the GSWP2 baseline run (827 mm/year) was over-corrected for the wind-caused gauge undercatch of snowfall in cold regions [*Zhao and Dirmeyer*, 2003].

[17] EXP6 produces 16 W m^{-2} more net solar radiation than Noah V3 over arid and semiarid regions, mainly because surface albedo for barren surfaces in EXP6, which is a function of soil color, surface soil moisture, and wave band (visible and near-infrared) and around 0.30 on average, is lower than that in Noah V3, which is prescribed at 0.38 (Figure 1a). Over cold regions, EXP6 produces lower net solar radiation than Noah V3 does due mainly to its higher snow surface albedo and a longer snow season than Noah V3. The differences in net solar radiation between EXP6 and Noah V3 are partly balanced by the differences in net long-wave radiation (upward positive) (Figure 1b). As a net result, the global averaged net radiation from EXP6 (73 Wm^{-2}) is

5 Wm^{-2} more than that from Noah V3 (68 Wm^{-2}) (Table 2). Over arid and semiarid regions, EXP6 results in more sensible heat (Figure 1c) and latent heat fluxes (as indicated from ET in Figure 1d) than Noah V3 to further balance the residual differences in net radiation. EXP6 simulates more ET over most land areas by more than 100 mm/year over the majority of land areas. However, over tropical Amazon River and Congo River basins, EXP6 produces more sensible heat (Figure 1c) and less ET (Figure 1d) than Noah V3 does. This change is mainly attributed to the use of different C_H schemes as indicated from the difference between EXP2 and EXP3 (Table 2). Due mainly to the use of d_0 in the M-O scheme, EXP3 produces greater sensible heat than EXP2 does over tropical rain forests and boreal forests (figures not shown). The use of d_0, which is over 10 m in the forest regions, in the M-O scheme greatly enhances C_H and hence sensible heat fluxes. To balance the increase in sensible heat, latent heat (or ET) decreases (Figure 1d). The modeled runoff shows an opposite pattern to that of ET to balance the water budgets. EXP6 produces less soil water in the top 2 m soil due partly to the increased ET over most land areas except for the Amazon River basin, east Siberia, Northwest U.S., and central Canada, where EXP6 shows slightly higher soil moisture with less ET.

3.2. Runoff

[18] We compared the modeled monthly runoff climatology to the GRDC estimates. Because of the balance of precipitation with the sum of ET and runoff, evaluation of runoff can be regarded as an indirect way of evaluating ET. Noah V3 produces 388 mm/year runoff (Table 2), about 38.5% more

Table 2. Global (60°S–90°N) 10 Year (1986–1995) Area-Weighted Averages of Land Surface Energy and Water Budgets[a]

	S_a W/m^2	L_a W/m^2	R_{net} W/m^2	H W/m^2	LE W/m^2	P mm/a	ET mm/a	R mm/a	R_s mm/a	R_b mm/a
Noah-V3	133	65	68	37	30	769	376	388	84	305
EXP1[b]	141	65	76	38	37	769	460	308	98	211
EXP2	141	65	76	38	37	769	463	305	64	241
EXP3	140	64	77	43	33	769	416	352	69	283
EXP4	140	64	77	42	34	769	422	347	93	254
EXP5	140	64	77	42	34	769	422	347	93	254
EXP6	137	64	73	37	34	769	430	339	91	248
EN36[c]	139	64	75	41	34	769	421	347	121	226
GSWP2[d]	142	68	74	35	37	827	471	322	119	203

[a]S_a, net solar radiation; L_a, net longwave radiation (positive upward); R_{net}, net radiation; H, sensible heat; LE, latent heat; P, precipitation; ET, evapotranspiration; R, runoff; R_s, surface runoff; and R_b, base flow.
[b]Lake points are excluded in experiments from EXP1 to EXP6 (which compute lake surface temperature and ET) for comparison with Noah V3 (without lake).
[c]Ensemble mean of the 36 experiments (see section 5).
[d]GSWP2: Global Soil Wetness Project phase-2 12-model mean, which is available at http://hydro.iis.u-tokyo.ac.jp/GLASS/GSWP/ICC_Report01.html. The 12 model's results are averaged regardless of imbalance of water or energy of any model.
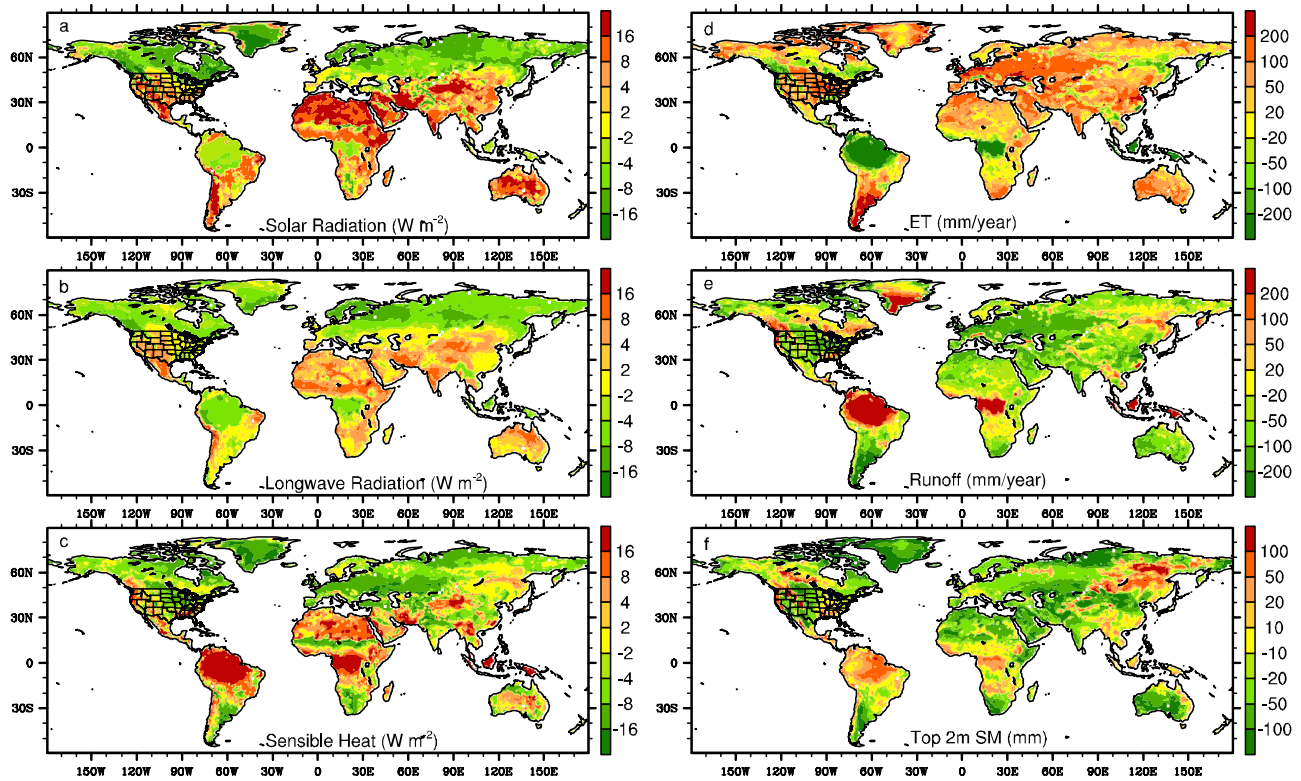
Figure 1. EXP6 minus Noah V3 (averaged over 1983–2006) (a) net solar radiation, (b) net longwave radiation (positive upward), (c) sensible heat, (d) ET, (e) runoff, and (f) top 2 m soil water.

than the GRDC estimate (280 mm/year). The six experiments (EXP1–EXP6) produce less runoff (305–352 mm/year) than Noah V3, but still 9–26% more than the GRDC estimates. Over the 50 global river basins, which cover most of the land area but excluding Greenland, EXP1–EXP6 produces 4.2–23.1% more runoff than GRDC estimates (see the legend of Figure 4c). Considering that industrial and agricultural water use and increased surface water evaporation from man-made reservoirs may reduce global runoff by 20–25% [*Vörösmarty et al.*, 2004], the total amount of runoff modeled by EXP1–EXP6 appears to be in a plausible range. EXP2 (using Chen97's C_H scheme) simulates 47 mm/year (16.8% relative to GRDC) less runoff than EXP3 (using the M-O C_H scheme). This significant difference suggests the important role of the surface exchange coefficient in controlling the total amount of runoff. Because of the limited accuracy of GRDC estimates, we cannot conclude that Chen97 is superior to M-O scheme.

[19] Noah V3 produces more runoff than GRDC estimates over most land areas including arid and semiarid regions, midlatitude wet regions (e.g., eastern U.S., southern China, and Europe), cold regions (e.g., west Siberia), and the tropical Congo River basin by up to 200 mm/year. Noah V3 simulates less runoff in the tropical Amazon River basin and slightly less along the boreal forest band in Canada, the corner of the Northwest U.S., and southern Tibet (Figure 2a). EXP6 reduces the biases produced by Noah V3 in most regions except for the tropical Congo River basin and southern Tibet (Figure 2b). However, the bias reduction appears insufficient in Northern Eurasia, Alaska, and southern South America (Figure 2c). The residual biases can be attributed to many

factors, such as limited accuracy of atmospheric forcing data and GRDC runoff estimates (e.g., over the Congo River basin and southern Tibet), inadequate representation of processes (e. g., $z_{0h} = z_{0m}$ in the M-O scheme), and/or globally constant runoff parameters and other model parameters (e.g., ground surface roughness length is globally set to 0.01 m). Some key parameters should be calibrated when applying the model to a specific river basin.

[20] We also evaluated the seasonal variations and annual magnitude of runoff in 24 selected river basins, representing four different climate zones of cold, midlatitude, tropical, and arid regions (Figure 3). We averaged the modeled runoff over these catchments using a global 1° river basin mask data set [*Graham et al.*, 1999]. For rivers in cold region, EXP6 consistently improves runoff simulation in terms of timing and amplitude mainly because of the more permeable frozen soil and the improvements in the simulations of snow cover processes (see section 4.4). However, in midlatitude and tropical river basins, the results are mixed. EXP6 improves the simulation of total amount of runoff over the Dabube, Yangtze, and Mississippi rivers but worsens the simulation in the Ganges River of North India. In most tropical rivers, the total amount of runoff increases (Figure 2). The increase favors the simulation in the Amazon River basin, but not others. In arid regions, both runoff seasonality and magnitude are improved (Figure 3).

[21] We used Taylor diagrams [*Taylor*, 2001] to compare seasonality of modeled runoff in terms of phase (measured by correlation coefficient between simulation and observation) and amplitude (measured by standard deviation of simulation normalized by that of the observation) for all the seven
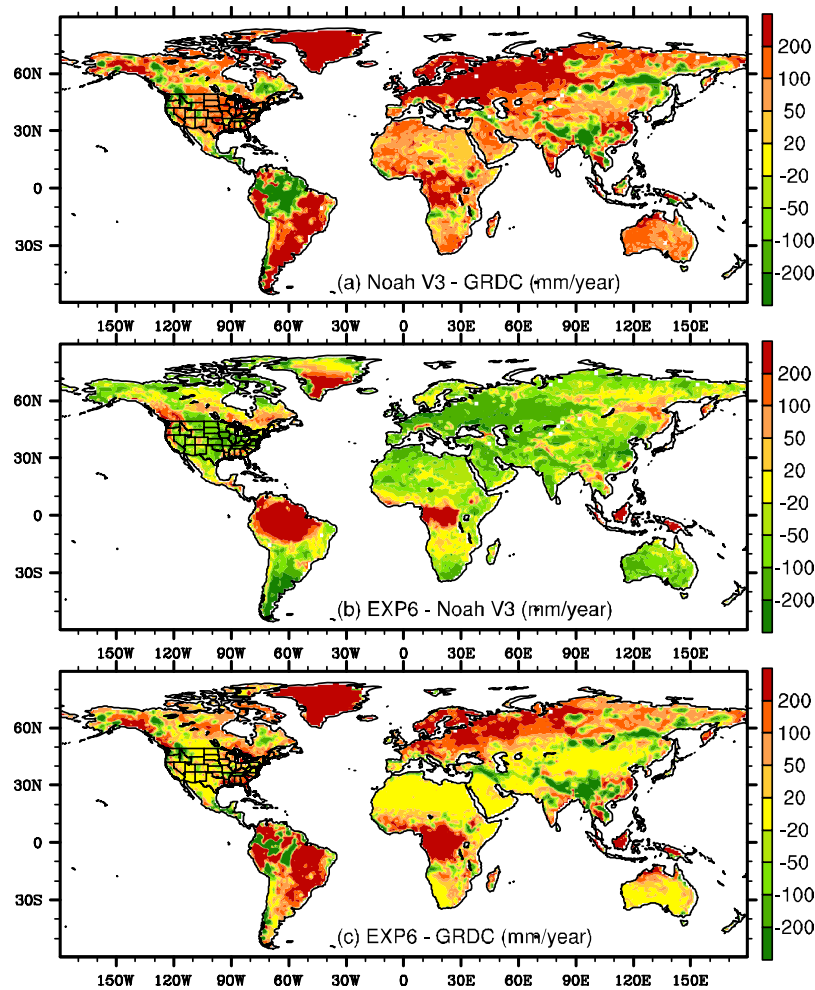
Figure 2. Differences of annual runoff climatology (averaged from 1983 to 2006) (a) Noah V3 minus GRDC, (b) EXP6 minus Noah V3, and (c) EXP6 minus GRDC.

experiments. Figure 4a shows, for each experiment, the area-weighted average of the correlation coefficients and normalized standard deviations obtained at six cold river basins. The area-weighted average correlations and normalized standard deviations for six arid river basins and the 50 largest river basins are presented in Figures 4b and 4c, respectively. In midlatitude and tropical river basins, the average metrics for each experiment are very close and overlapped (not shown), indicating EXP6 maintained Noah V3's performance over these regions. In cold-region river basins, Noah V3 shows on average a lower correlation and greater amplitude because of its earlier runoff peaks induced by more impermeable frozen soil as represented in the model. EXP1 shows the same level of correlation and even greater amplitude than Noah V3 because of the stronger snowmelt in springtime. EXP2 greatly reduced the amplitude by virtue of a more permeable frozen soil. EXP3 further improves the average correlation by changing the C_H scheme to one that produces a longer snow season due to decreases in sublimation from snow surfaces in melting seasons (figures not shown). An unconfined aquifer below the soil column plays a buffering effect on the freely drained soil water, and in such a way, EXP4 alleviates the higher amplitude produced by EXP3. EXP5 using the Ball-Berry scheme of stomatal

resistance does not show much improvement against EXP4 with the Jarvis stomatal resistance scheme. EXP6 with the dynamic vegetation shows a little improvement over EXP5 in the average correlation. In arid regions, the experiments with the augmented version (EXP1–EXP6) show significant improvements against the baseline model in both average standard deviation and correlation (magnitude and seasonality), but small differences between each other. Because almost a half of the world's largest 50 river basins are affected by frozen soil and snow cover, the model performance averaged over global 50 river basins shows a pattern similar to that for the six cold region river basins.

3.3. Soil Moisture

[22] We evaluated modeled soil moisture against observations made in Illinois (Figure 5). Noah V3 simulates fairly well the top 2 m soil moisture, in terms of mean state and variability, but produces wetter top 1m soil with lower variability. EXP6 maintains Noah V3's capability to simulate the top 2 m soil moisture (a slightly greater variability though) and improves the simulation of the top 1 m soil moisture, especially the variability. EXP6 still produces drier soil during the 1988 summer drought. An experiment using CLM factor (e.g., EN5 in Table 3) further improved the simulation
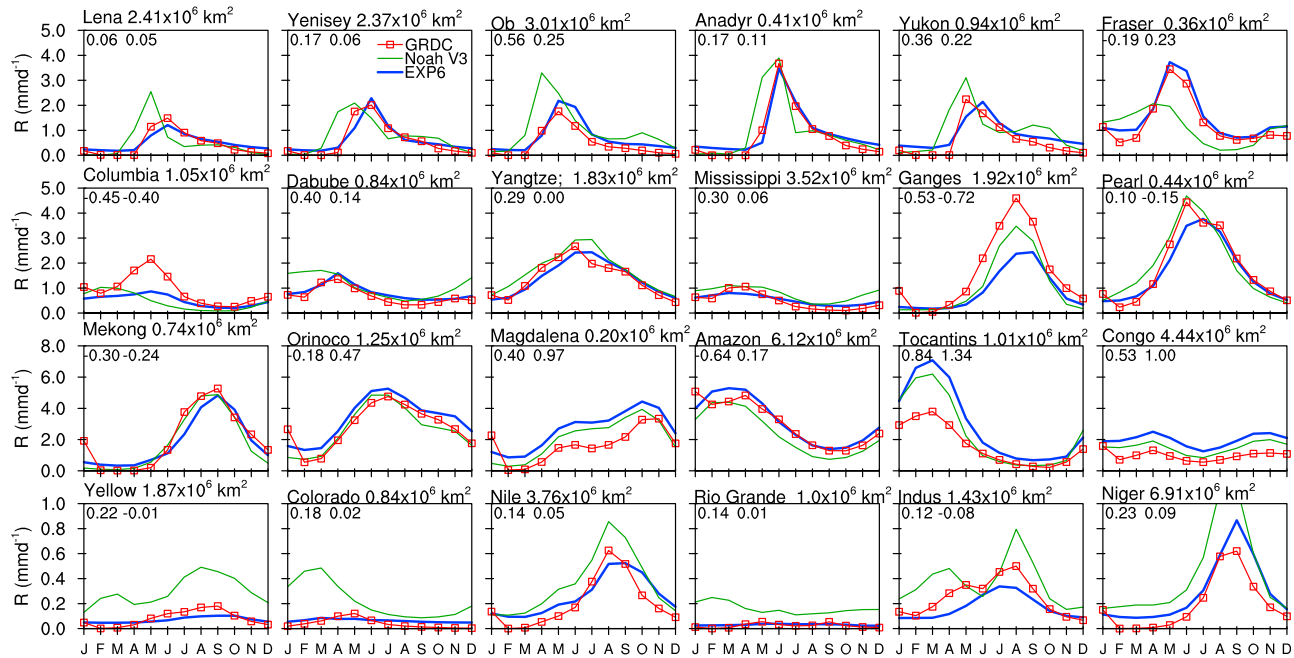
Figure 3. Monthly runoff climatology (averaged from 1983 to 2006) modeled by Noah V3 and EXP6 in comparison with GRDC runoff estimates in 24 selected river basins representing four different climate zones (top to bottom rows are for cold, midlatitude, tropical, and arid regions). Also shown at the upper left corner of each panel are the relative model biases [(model–observed)/observed] for Noah V3 and EXP6 in sequence.

of soil moisture during the 1988 summer drought without degrading other year's simulation (Figure 5). The CLM- or SSiB-β factor uses higher wilting point (converted from the wilting matrix potential, i.e., −150 m) for loam and clay [*Niu et al.*, 2011, Figure 3] and facilitates ceasing transpiration for dry soil conditions, and thus saving soil water under drought conditions.

[23] We compared modeled runoff against GRDC runoff monthly climatology, as an indirect way to evaluate ET over this region (Figure 6). Noah V3 produces less ET and thus greater runoff than EXP6 does. Compared to GRDC estimates,

Noah V3 produces too much runoff, indicating that the weaker ET is unfavorable. EXP6 and Noah V3 produce almost the same level of top 2 m soil moisture, despite the much greater ET produced by EXP6. This can be mainly attributed to groundwater buffering effect included in EXP6, which facilitates maintaining soil moisture and ET during dry seasons.

3.4. Snow

[24] Noah V3 and EXP6 produce very similar geographical distributions of SWE over Northern Hemisphere in February (Figures 7a and 7b). However, Noah V3 simulates more snow
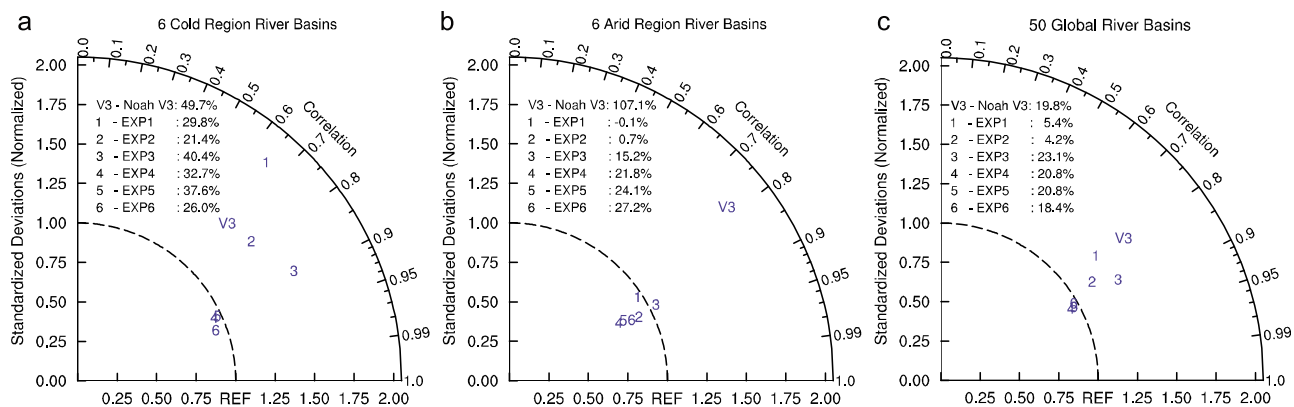
Figure 4. Statistics of the monthly runoff climatology (averaged from 1983 to 2006) for the seven experiments. The correlation coefficient and normalized standard deviation for each experiment are basin-area-weighted averages over (a) six cold region river basins, (b) six arid region river basins, and (c) 50 global largest river basins. Also shown in the legends are relative biases ([model–observation]/observation in percentage) of the modeled runoff.
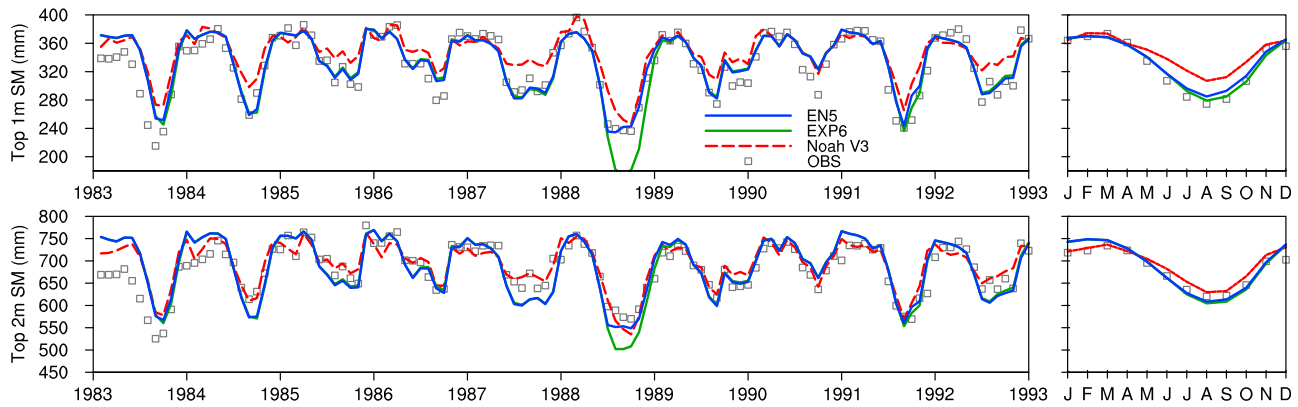
Figure 5. Modeled soil moisture averaged over Illinois by Noah V3, EXP6, and EN5 (see Table 3) in comparison with observations for top 1 m and top 2 m soil.

mass in most high-latitude regions but less over southern melting regions including Europe, the Middle East, and midlatitude North America than EXP6 does (Figure 7c).

[25] We compared the modeled SWE and snow depth simulated by Noah V3 and EXP6 against CMC estimates over the eight largest North American river basins (Figures 8a and 8b). Noah V3 simulates slightly more SWE over two high-latitude river basins (the Mackenzie and Yukon) but much less SWE over the other six river basins than CMC SWE (Figure 8a). EXP6 agrees well with CMC SWE over most of the river basins except for the mountainous Columbia and Colorado River basins, where EXP6 produces less snow. In both the catchments, precipitation (or snowfall) may be underestimated [*Pan et al.*, 2003] due to gauge undercatch errors [*Adam and Lettenmaier*, 2003; *Adam et al.*, 2006]. Overall, EXP6 with the multilayer structure, which facilitates representations of ground heat flux, retention, percolation, and refreezing of melt liquid water within the snowpack, and snow densification processes, improves the simulation of snowmelt and hence the simulations of snow mass over these river basins. Snowmelt occurs only during midday when snowpack temperature reaches the melting point. Thus, the diurnal cycle of snow layer temperatures, controlled by the net energy input into a snowpack, i.e., ground heat flux, is critical for snowmelt. Ground heat flux is further affected by surface energy budgets and their associated parameters (e.g., SCF, snow surface albedo, snow surface roughness length, and vegetation's cover fraction and interception capacity). Sensitivity experiments (figures not shown) indicate that all these factors may affect snow simulation in the augmented version. EXP1 and EXP2, which uses Chen97's C_H scheme, produces greater sublimation and thus slightly less snow mass in midlatitude river basins, while the other three runs (EXP3, EXP4, and EXP5) produces almost the same snow mass as EXP6 does.

[26] Noah V3 simulates smaller snow density than that of the CMC estimates as indicated from Figures 8a and 8b. For such a reason, Noah V3 produces much greater snow depth in the two high-latitude rivers (the Mackenzie and Yukon) and less snow depth than CMC snow depth (but in a closer agreement with the CMC snow depth than with the CMC SWE) in the other six rivers (Figure 8b). EXP6 modeled snow density agrees fairly well with that of the CMC estimates, as indicated from the same degree of agreement (between model

and the CMC) for SWE with that for snow depth (Figures 8a and 8b).

[27] Noah V3 produces more snow mass than EXP6 in most high-latitude river basins (Figure 9). The greater SWE simulated by Noah V3 is mainly because of greater frost (negative ET) being formed on the snow surface during deep winter. Noah V3 also produces greater sublimation (positive ET) from the snow surface in springtime than EXP6 does, conducing to faster depletion of the thicker snowpack during the melting season.

[28] Snow mass on the ground is a major contributor to seasonal variations of the total water storage over high-latitude river basins. To compare with GRACE ΔS estimates more strictly, we filtered the model outputs of snow mass, soil water, and groundwater using the same filtering algorithm as *Chen et al.* [2006]. Compared to GRACE estimates of TWS change, Noah V3 shows much greater amplitude, while EXP6 shows closer agreement in terms of amplitude and correlation. This indicates that the greater snow mass and excessive frost fall modeled by Noah V3 over high-latitude river basins are undesirable.

3.5. Skin Temperature

[29] The modeled skin temperature is evaluated against GOES-EAST estimates of land surface skin temperature. GOES skin temperature was demonstrated reliable through comparison with Oklahoma ARM/CART (Atmospheric Radiation Measurement/Cloud and Radiation Test bed)

Table 3. The First Group of 12 Experiments and Their Corresponding Options of Schemes

| Experiment	Dynamic Vegetation	r_s	β	Runoff Schemes
EN1	On	Ball-Berry	Noah	SIMGM
EN2	On	Ball-Berry	Noah	SIMTOP
EN3	On	Ball-Berry	Noah	Schaake96
EN4	On	Ball-Berry	Noah	BATS
EN5	On	Ball-Berry	CLM	SIMGM
EN6	On	Ball-Berry	CLM	SIMTOP
EN7	On	Ball-Berry	CLM	Schaake96
EN8	On	Ball-Berry	CLM	BATS
EN9	On	Ball-Berry	SSiB	SIMGM
EN10	On	Ball-Berry	SSiB	SIMTOP
EN11	On	Ball-Berry	SSiB	Schaake96
EN12	On	Ball-Berry	SSiB	BATS
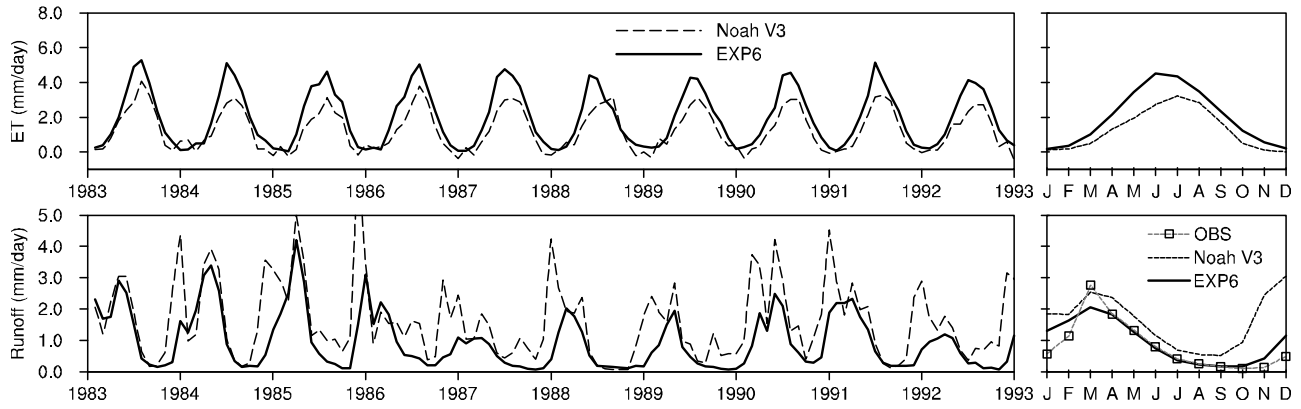
Figure 6. Modeled ET and runoff averaged over Illinois by Noah V3 and EXP6 in comparison with GRDC runoff monthly climatology (OBS).

ground-based measurements [*Mitchell et al.*, 2004]. To closely compare to the GOES estimates, we used the modeled surface radiative temperature (T_{rad}), and $T_{rad} = (L_{up}/\sigma)^{1/4}$, where L_{up} is the sum of upwelling longwave radiation fluxes from each of the surface components including the vegetation canopy, vegetated ground, and bare ground.

[30] Noah V3 shows a significant cold bias in arid Western U.S. (Figure 10). EXP2 (using Chen97's C_H scheme) shows the same level of cold bias. EXP3 (using the M-O C_H scheme) greatly improves the simulation, and EXP6 further improves

the simulation but the temperature is slightly lower than the GOES retrieval. EXP6a (EXP6 but with the M-O scheme being changed back to Chen97's scheme) shows the same cold bias as Noah V3 does, further confirming the important role of C_H in controlling land skin temperature. Under a neutral condition, the M-O scheme [*Niu et al.*, 2011] (equation (16)) should produce greater C_H than Chen97 [*Niu et al.*, 2011] (equation (15)) does by taking into account the zero-displacement height and a greater z_{0h}. However, Chen97 produced greater C_H than the M-O scheme because of the
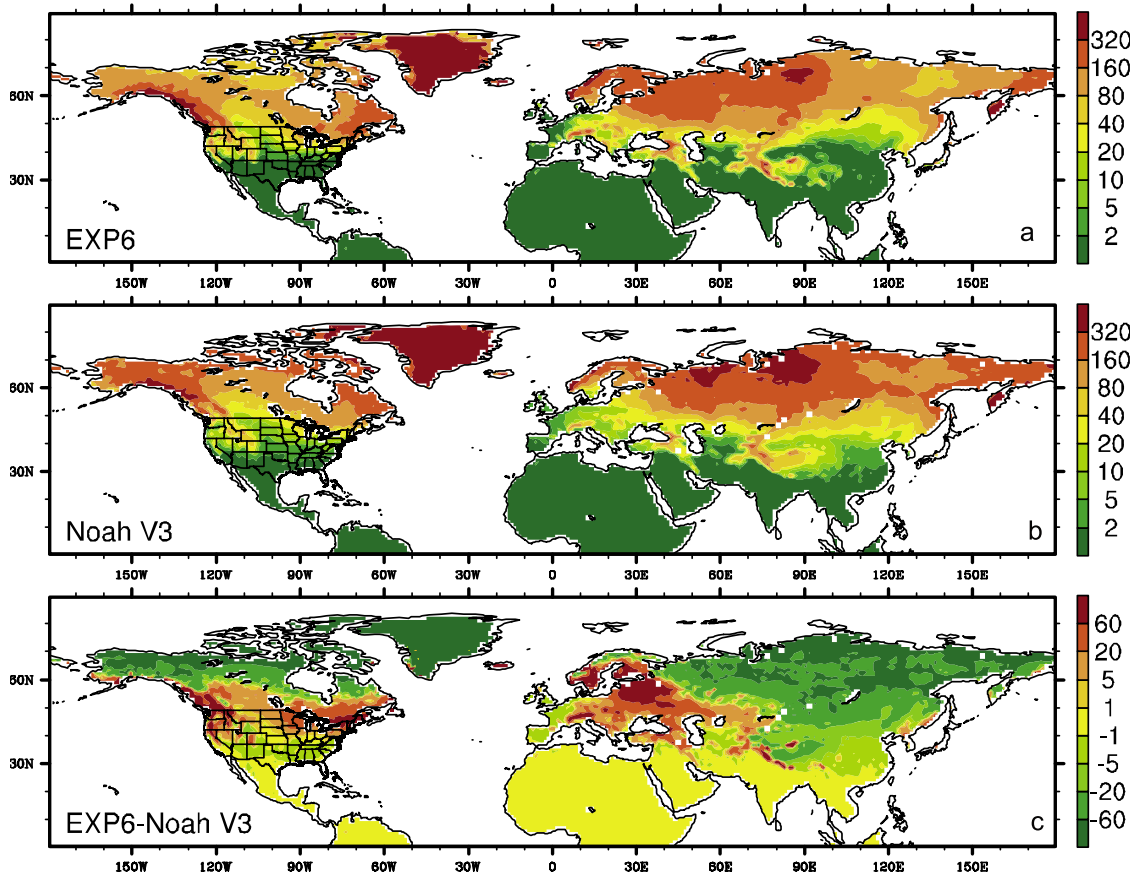
Figure 7. The 24 year averaged (1983–2006) February SWE (mm) from (a) EXP6, (b) Noah V3, and (c) the difference between EXP6 and Noah V3.
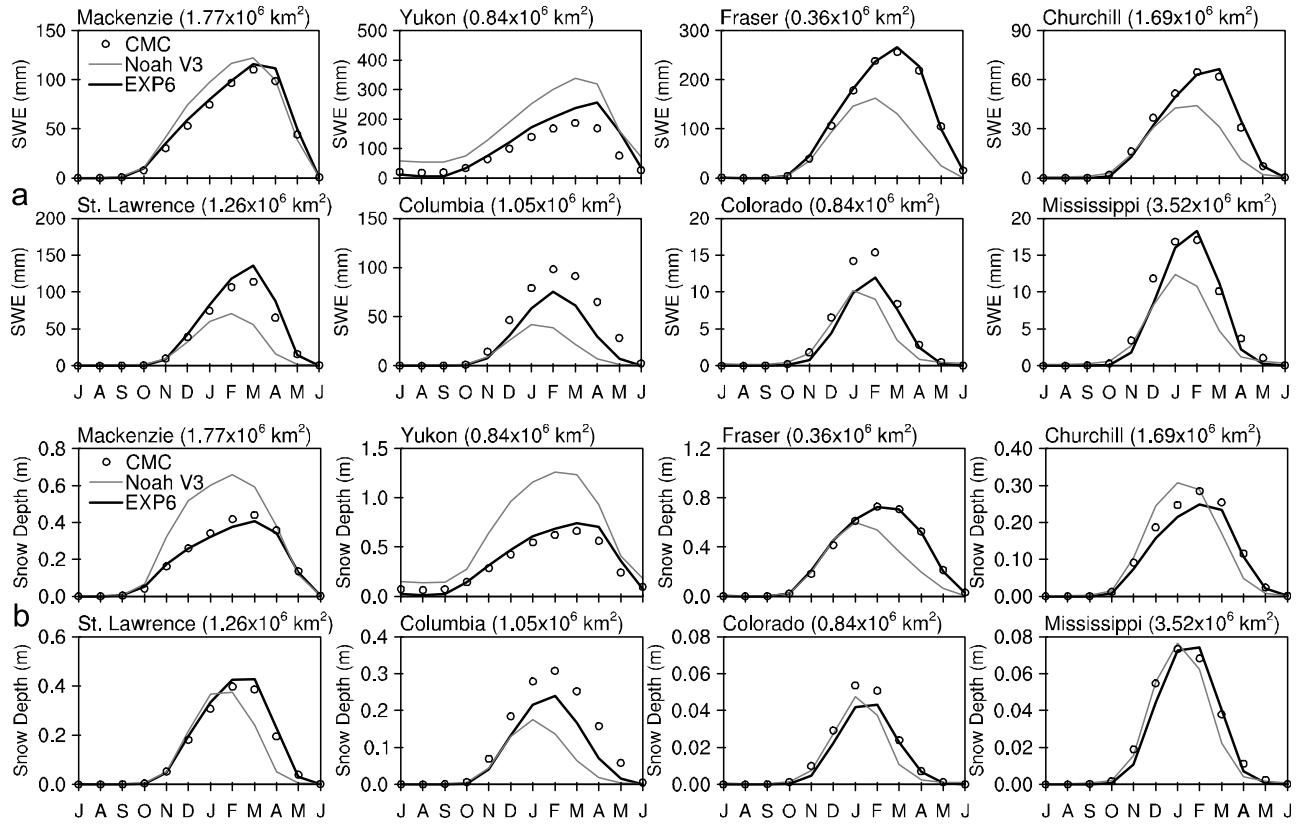
Figure 8. Monthly (a) SWE climatology and (b) snow depth (averaged from 1983 to 1996) modeled by Noah V3 and EXP6 in comparison with the CMC data.

different stability correction and the additional effects of planetary boundary layer (PBL) height on friction velocity. A greater C_H means more efficient ventilation of the land surface and greater cooling of the land surface during summer daytime.

[31] Noah V3 uses its surface energy balance equation to compute the overall skin temperature of the combined soil and vegetation without iteration, while EXP6 separately computes skin temperatures of the bare ground, vegetation canopy, and vegetated ground using energy balance equations
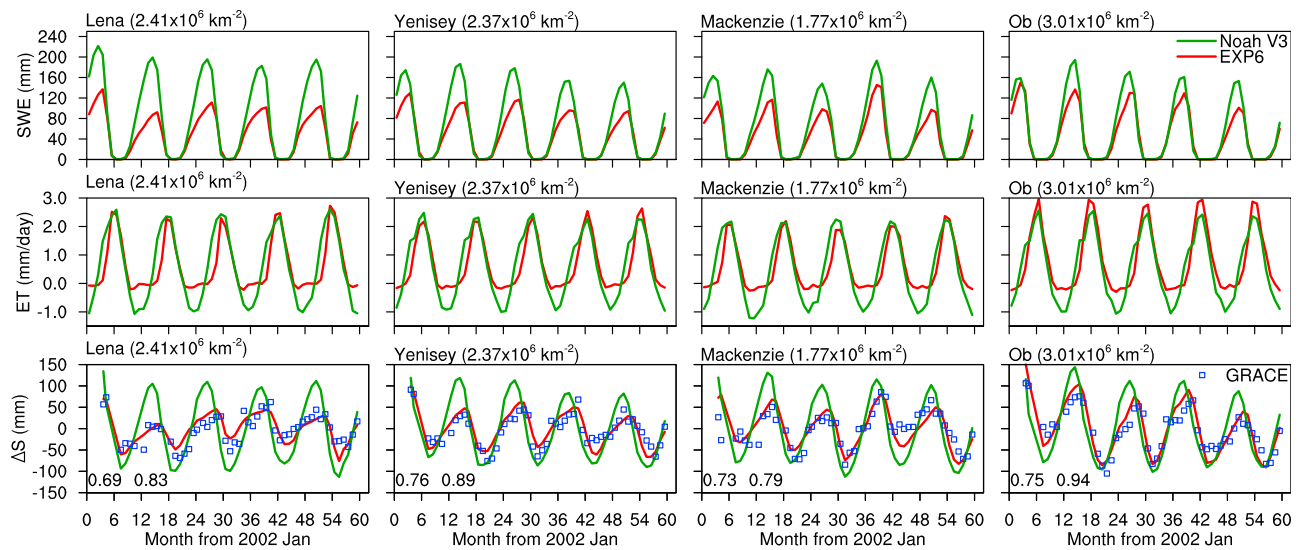
Figure 9. Modeled monthly (top) SWE, (middle) ET, and (bottom) anomaly of total water storage (filtered) in comparison with GRACE estimates over four large high-latitude river basins (from 2002 to 2006). Also shown in the lower left corner are of the bottom panels are correlation coefficients for Noah V3 and EXP6, in order, with GRACE.
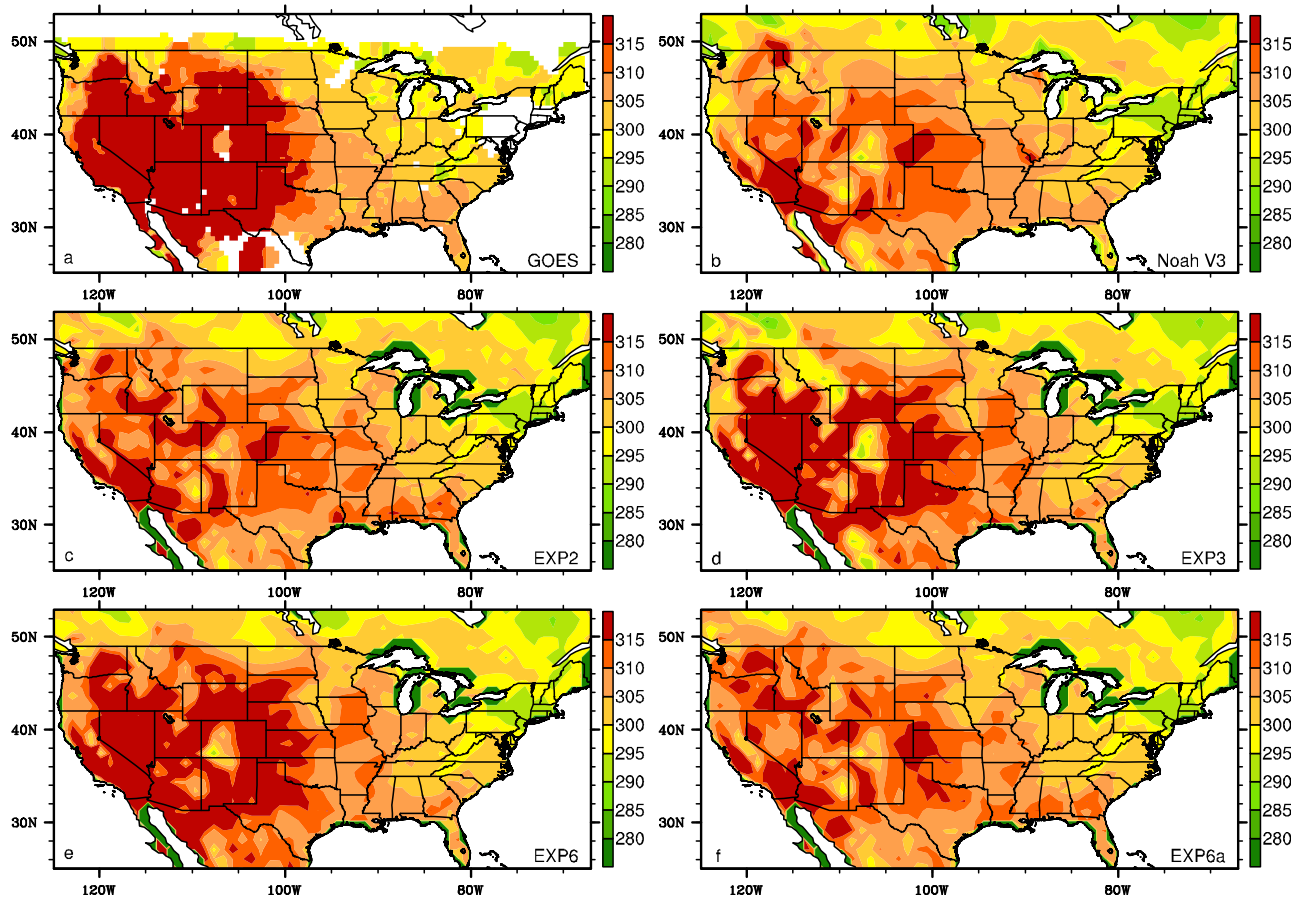
Figure 10. Land skin temperature at 2100 UST of 12 July 2004 (a) GOES-EAST retrieval, (b) modeled by NoahV3, (c) by EXP2, (d) by EXP3, (e) by EXP6, and (f) modeled by EXP6 but with Chen97's surface exchange coefficient scheme.

for each of the surface components with iterations. Additionally, EXP6 differs from Noah V3 in many other aspects, such as formulations of surface energy fluxes and representations of hydrological processes. The experiments indicate that C_H is the most important factor for modeling land skin temperature. Chen97 should be able to improve the simulation in Western U.S. by further reducing z_{0h} (by calibrating the "C" parameter in equation (15) [*Niu et al.*, 2011]; see also the discussion by Chen97). However, to maintain the simulation in Eastern U.S., a more elaborate parameterization of distributed z_{0h} should be applied (X. Zeng, personal communication, 2010). The M-O scheme has a potential to further improve the simulation by also accounting for the different roughness length for heat from that for momentum as considered by Chen97.

3.6. LAI and GVF

[32] Prediction of LAI and GVF is a unique attribute of EXP6. We compared the model predicted LAI and GVF against the MODIS LAI and the NESDIS GVF product.

[33] We calibrated some major parameters related to leaf phenology including maximum carboxylation rate at 298 K (equation (B7) [*Niu et al.*, 2001]), specific leaf area, growth respiration fraction, foliage maintenance respiration rate, leaf turnover rate, leaf death rate due to drought stress and cold stress for various vegetation types. The resulted global annual

gross primary productivity (GPP) and net primary production (NPP) of carbon are 143.2 Pg (10^{15} g) and 62.0 Pg, respectively, slightly larger than other estimates (120 Pg for GPP and 60 Pg for NPP [*Janzen*, 2004]) due mainly to the assumption of saturated nitrogen in the model. The spatial distributions of LAI and GVF simulations show a favorable agreement with those of satellite retrievals (Figure 11). Major ecosystems such as boreal forests and tropical forests are well simulated. The transition from arid to wet regions is also well captured by the model, indicating the drought stress and soil moisture state are also favorably represented in the model. The modeled GVF showed a wider band of boreal forests, where GVF is greater than 0.2, due mainly to the too low NESDIS GVF in wintertime.

[34] Over 24 selected river basins in different climate zones (cold, midlatitude, tropical, and arid regions), the model well simulates both LAI (Figure 12) and GVF (Figure 13) in terms of seasonal variability and interbasin variability. The model shows a fairly good agreement with satellite retrievals in midlatitude and arid river basins with a much higher correlation with satellite retrievals than in tropical regions. In tropical river basins, the simulated LAI has a quite low correlation with MODIS LAI, with the lowest correlation coefficient in the Amazon River basin (−0.88). However, the modeled GVF shows a much higher correlation with the NESDIS data over tropical river basins, especially the
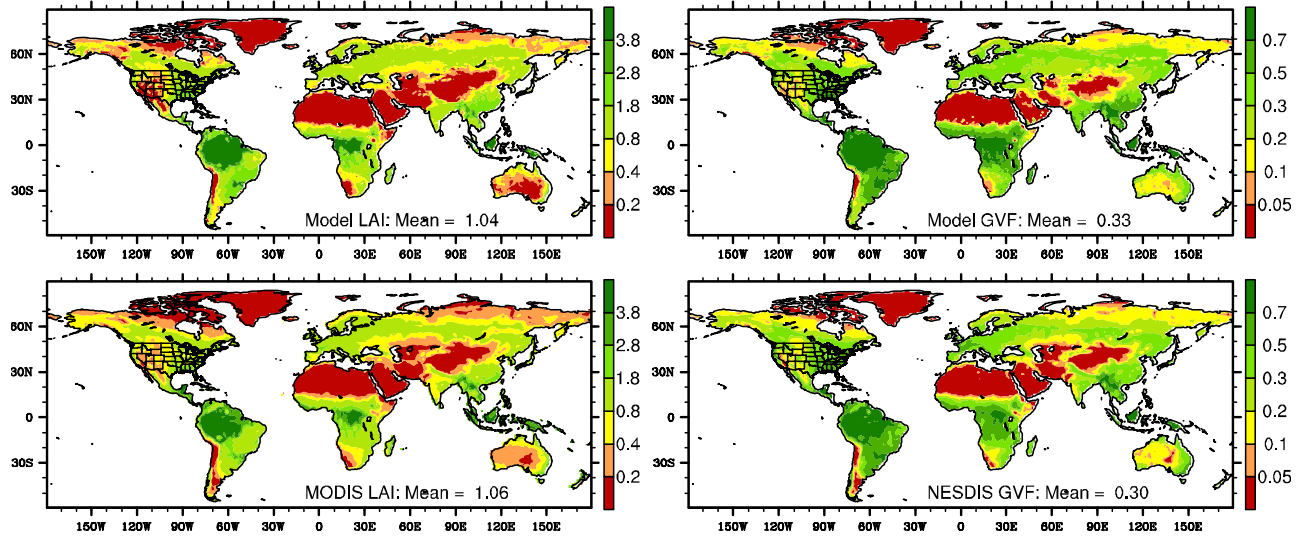
Figure 11. Model predicted LAI and GVF climatology (averaged from 2001 to 2006) in comparison with the MODIS LAI and NESDIS GVF.

basins of the Mekong, Orinoco, and Tocantins rivers. Modeling tropical ecosystems, especially modeling the Amazon forests, is a challenging task because of the complex response of vegetation to availabilities of sunlight and water [*Saleska et al.*, 2007]. Over most of the cold regions, both modeled LAI and GVF show a delayed spring outburst of leaves. In the model, the dying process due to cold stress functions all the year around including springtime. Thus, the newly grown leaves in spring may die whenever a cold front comes. This spurious phenomenon as represented by the model can be partly corrected by decreasing the death rate. But decreasing

the death rate would delay the dying of leaves in fall season. We suggest that in the future, a timing system should be set up to the dynamic leaf model, so that the model leaves can distinguish spring season from fall season.

4. A 36-Member Physically Based Ensemble Simulation

[35] The total number of combinations of alternative schemes for the 10 processes described in section 4 of *Niu et al.*'s [2011] work can be up to 4608, representing 4608
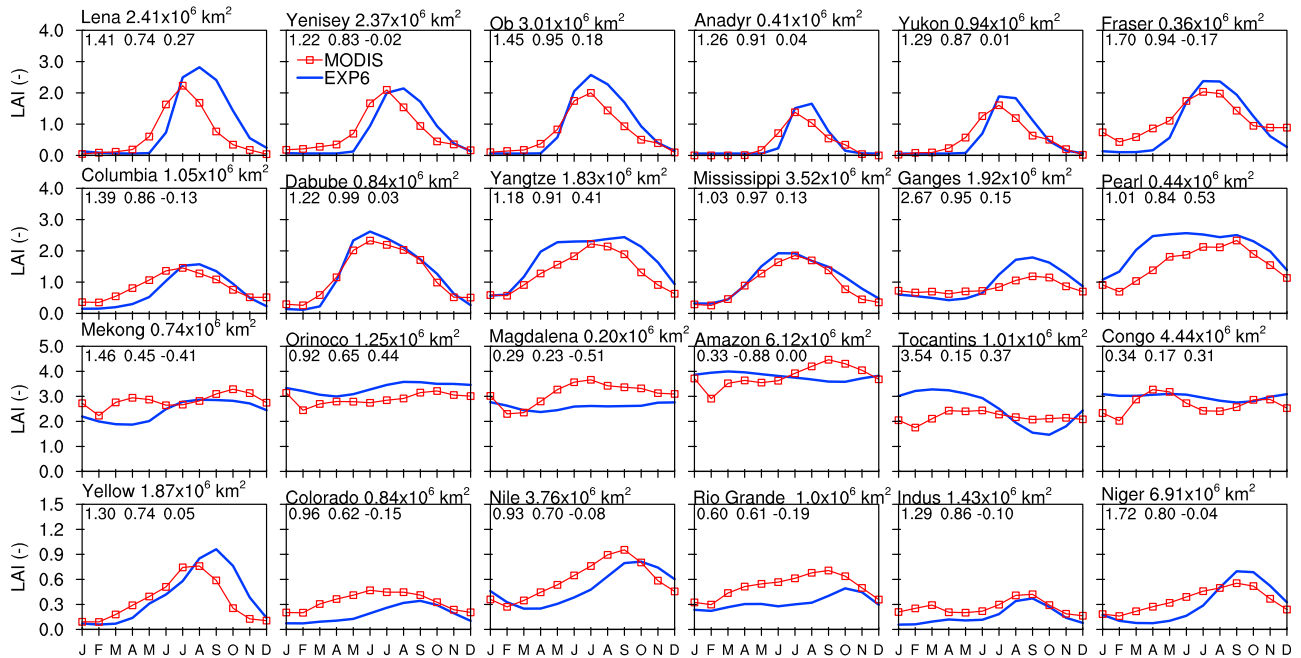
Figure 12. Modeled monthly climatology (averaged over 2001–2006) of LAI in comparison with MODIS LAI (averaged over 2000–2007) over 24 river basins representing (top to bottom) cold region, midlatitude, tropical region, and arid region. Also shown in the upper left corner of each panel are, in order, normalized standard deviation, correlation coefficient, and bias [model – observation].
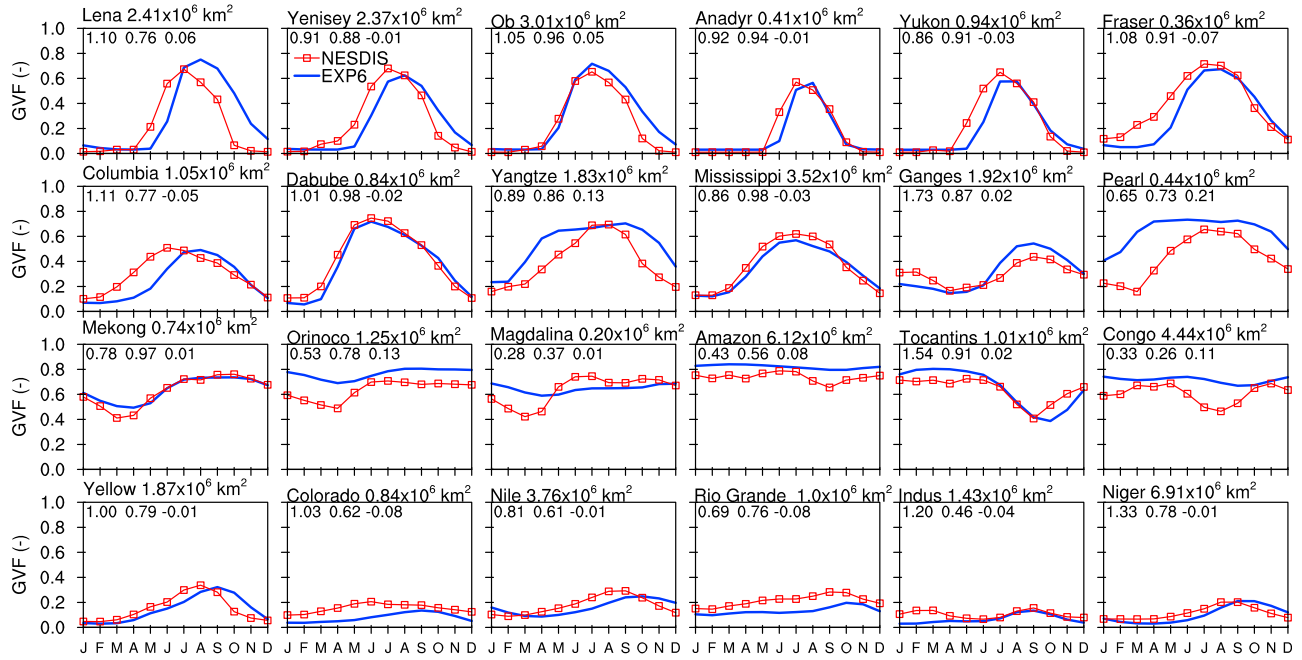
Figure 13. The same as Figure 12 but for GVF in comparison with NESDIS GVF.

models. We selected some critical processes that may more significantly affect land surface fluxes, especially ET flux. The interplay of the β and runoff schemes determines a model's overall performance [*Koster and Milly*, 1997]. We were inspired to select these two processes and two additional processes: leaf dynamics and stomatal resistance. The total number of combinations of optional schemes for the four processes is 36 (note that the dynamic leaf model can only work with the Ball-Berry stomatal resistance scheme). The first group of 12 experiments uses dynamic vegetation and Ball-Berry stomatal resistance but vary in the β schemes and runoff schemes as listed in Table 3. Based on the first group of experiments, the second group of 12 experiments turns off the dynamic vegetation model but uses prescribed vegeta-

tion (LAI look-up table and the monthly NESDIS GVF). The third group of 12 experiments changes the stomatal resistance scheme from the Ball-Berry to the Jarvis type while still keeping the dynamic vegetation turned off as in the second group of 12 experiments.

[36] The modeled ET and runoff, averaged over global landmass for the 24 year period (1983–2006), are well balanced with precipitation (Figure 14a). We also show ET varying with the top 1m soil moisture averaged over global land (Figure 14b) and over Illinois (Figure 14c).

[37] Runoff schemes play the most important role in differentiating the 36 models in the ET–soil moisture plot. The four runoff schemes are designated with different colors in Figure 14. SIMTOP produces the wettest soil and greatest ET;
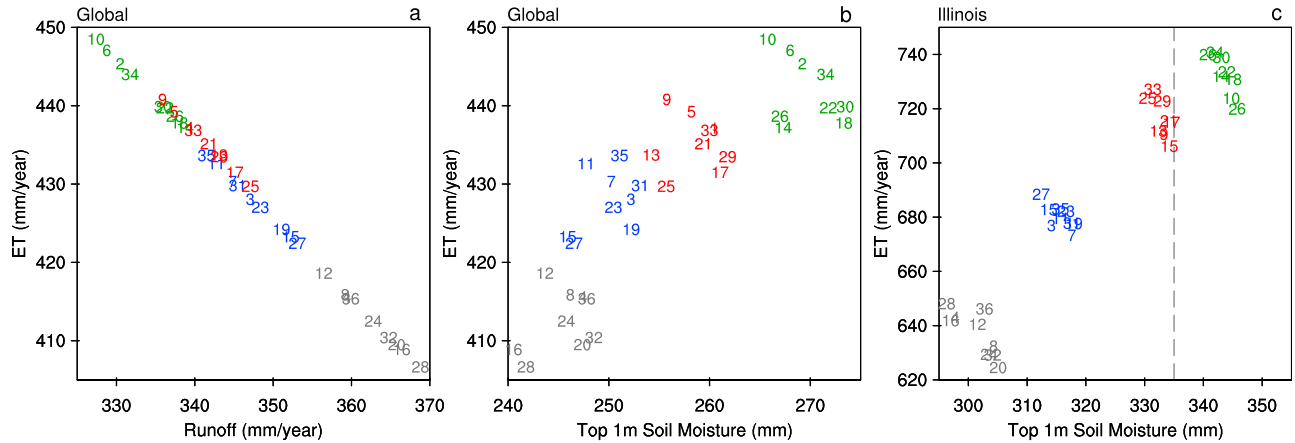
Figure 14. Relationships between (a) ET and runoff climatology (1983–2006) averaged over global 50 largest river basins, (b) ET and top 1 m soil moisture (1983–2006) averaged over global 50 largest river basins, and (c) ET and top 1 m soil moisture (1983–2006) over Illinois. Runs with SIMTOP runoff scheme are green, SIMGM is red, Schaake96 scheme is blue, and BATS runoff scheme is gray. The dashed vertical line in Figure 14c represents the observed top 1 m soil moisture (averaged over 1983–2003).
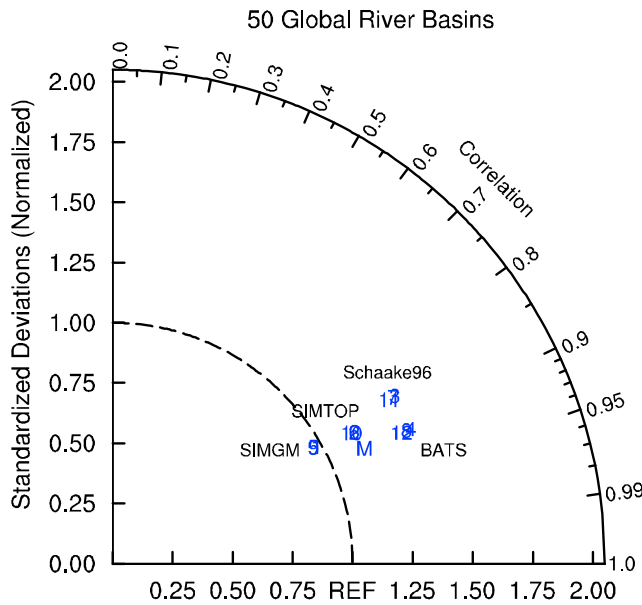
Figure 15. Statistics of the monthly runoff climatology (1983–2006) of the first group of 12 experiments. The correlation coefficient and normalized standard deviation for each experiment are basin-area-weighted averages over global 50 largest river basins. "M" stands for the statistics of the ensemble mean of the 36 experiments.

because it seals the soil bottom (i.e., zero-flux lower boundary condition), it retains more water in the soil column, which is first available for ET and then for lateral subsurface runoff. BATS produces the driest soil and smallest ET; the largest surface runoff results in the least water infiltrating into the soil. Schaake96 and SIMGM produce modest soil moisture and ET between those equipped with BATS and SIMTOP (Figure 14b); their surface to total runoff ratios are quite similar and much smaller than those of BATS. SIMGM produces a slightly wetter soil and more ET than the Schaake96 scheme because of the implementation of the buffering effects of aquifers (upward water flow from aquifers to soil in dry season) over the global land (Figure 14b). However, these effects are more obvious for a smaller region, e.g., Illinois (Figure 14c). Compared to the long-term (1983–2003) average of the observed soil moisture in the State of Illinois, the SIMGM group shows the best performance.

[38] Although the various β factors appear to show very different slopes against soil moisture, simulated global averages of soil moisture and ET by different β groups (e.g., EN1, EN5, EN9 using Noah, CLM, and SSiB β factors, respectively, with dynamic vegetation and SIMGM) show much smaller differences than those resulting from different runoff schemes. The degree of impact of the β factor on the soil moisture–ET relationship should depend on the soil moisture regime. For instance, over a specific region where the soil moisture ranges from 0.15 to 0.25 m^3/m^3 for a clay soil [*Niu et al.*, 2011, Figure 3c] (the Noah $\beta = 0.0$–0.5, while CLM and SSiB $\beta = 0.0$), the β factor should play a more important role than for global average.

[39] Within a runoff scheme group (e.g., SIMGM), predicted LAI (by the dynamic vegetation) (e.g., EN5 versus

EN17 under the CLM β factor and EN9 versus EN21 under the SSiB β factor) produces nonnegligibly different soil moisture and ET compared to models with prescribed LAI (Figure 14b). Runs with the Ball-Berry type and Jarvis type stomatal resistance schemes behave very similarly under both CLM-β (e.g., EN17 versus EN29) and SSiB-β factors (e.g., EN21 versus EN33) but more significantly different under the Noah β factor (e.g., EN13 and EN25).

[40] We conclude that the runoff schemes are predominant over the β factor, dynamic vegetation, and stomatal resistance schemes in controlling globally averaged soil moisture and ET and their relationship. However, the effects on the soil moisture and ET relationship may vary with different regions according to variations in soil moisture regimes and vegetation conditions. It should be pointed out that the present implementation does not include the β factor for ground surface evaporation, although possible options can be made available based on works by *Mahfouf and Noilhan* [1991] and *Sakaguchi and Zeng* [2009].

[41] Figure 15 shows the statistics of the 36-member ensemble mean runoff (denoted by M) in comparison with those of the first group of 12 runs. Runs with the BATS runoff scheme (EN4, EN8, and EN12) and Schaake96 scheme (EN3, EN7, and EN11) show the largest seasonal variations (greater monthly runoff peaks), while runs with SIMGM (EN1, EN5, and EN9) show the smallest seasonality. The BATS runoff scheme shows the largest correlation, while Schaake96 shows the least correlation. It is promising that the ensemble mean runoff (averaged monthly runoff of the 36 models) performs the best (the closest to "REF" in the Taylor plot) with the largest correlation (0.92) and less seasonal variability than the BATS and Schaake96 schemes. Runs with the Schaake96 runoff scheme (EN3, EN7, and EN11) appear to degrade the correlation of the ensemble mean runoff to GRDC estimates, but removing the 9 runs with the Schaake96 runoff scheme does not show significant improvement for the new 27-member ensemble. Due to the limited accuracy of the monthly GRDC runoff estimates, further demonstrations should be better conducted over small river basins where river discharges are recorded more frequently (e.g., daily). To further improve hydrological simulations, more elaborate methods to generate the ensemble average should be pursued [*Duan et al.*, 2007]. It is promising that the 36-member ensemble mean of runoff outperforms any single member over the world's 50 largest river basins. These results show the potential not only for physically based ensemble hydrological predictions, but also for land-based ensemble climate predictions.

5. Summary

[42] We evaluated the augmented Noah LSM performance against global and regional observational data sets, including MODIS LAI, NESDIS GVF, GRDC runoff, GRACE TWS change, CMC SWE and snow depth, GOES skin temperature, and Illinois soil moisture measurements. With different combinations of schemes, we conducted six experiments to mimic a transition from the original Noah model to the fully augmented version with dynamic vegetation using GLDAS atmospheric forcing data. The results show transitional improvements in modeling runoff, snow, soil moisture, and land skin temperature.

[43] The fully augmented version (EXP6) reduces the runoff bias produced by Noah V3 over most land areas but degrades runoff simulation in the tropical Congo River basin and southern Tibet where observational data, either the atmospheric forcing or runoff, may be scarce. Over 50 global river basins, which cover most of the land area but excluding Greenland, EXP6 produces 18.4% more than GRDC estimates. Considering that industrial and agricultural water use and increased surface water evaporation from man-made reservoirs may reduce global runoff by 20–25% [*Vörösmarty et al.*, 2004], the total amount of runoff modeled by EXP6 should be in a plausible range. The total amount of runoff is sensitive to different choices of C_H schemes, most significantly in forest regions due to the effects of d_0 on C_H. The M-O scheme that includes the effect of d_0 on C_H should further take into consideration z_{0h} that is different from z_{0m} as considered by Chen97. The fully augmented version greatly improves runoff seasonality over cold-region river basins due to the improved representation of frozen soil permeability and snow cover processes, such as snow albedo, snow sublimation and frost fall, and snowmelt.

[44] The fully augmented version maintains Noah V3′s capability of simulating the top 2 m soil moisture and improves the top 1 m soil moisture variability over Illinois. EXP6 produces greater summertime ET than Noah V3 does with almost the same level of soil moisture due to the groundwater buffering effects. The mean state and variability of soil moisture are mainly controlled by ET but modulated by the buffering effects due to groundwater.

[45] Noah V3 produces less SWE in the melting season due to greater snowmelt and sublimation but too much SWE during the accumulation season due to the excessive frost formation on the snow surface. The fully augmented version improves snow simulations in both melting and accumulation seasons as evaluated against CMC SWE and snow depth data sets and GRACE TWS change data.

[46] The fully augmented version improves the simulation of midday land skin temperature over arid Western U.S. by reducing the cold bias produced by Noah V3. The warmer midday land surface skin temperature produced by the fully augmented version is mainly attributed to the use of M-O scheme for the surface exchange coefficient, which produces a smaller C_H than Chen97's scheme.

[47] Prediction of LAI and GVF is a unique attribute of the fully augmented version. The model well simulates the spatial distributions of LAI and GVF compared to satellite retrievals. The model simulates LAI and GVF seasonality better in midlatitude and arid river basins than in tropical regions. In tropical river basins, the modeled LAI has a low correlation with MODIS LAI, while the modeled GVF shows a much higher correlation with the NESDIS GVF data.

[48] We also conducted a 36-member physically based ensemble simulation with 36 different combinations of optional schemes for leaf dynamics, stomatal resistance, the β factor, and runoff. The runoff schemes largely determine the modeled soil moisture–ET relationship. A runoff scheme with a sealed bottom (zero-flux lower boundary condition, like SIMTOP) produces the wettest soil and greatest ET. A model with a greater surface runoff (like BATS) produces the driest soil and smallest ET. A runoff scheme with gravitational free drainage like Noah V3 and BATS tends to underestimate soil moisture and produces less ET than that

with a groundwater model. The β factor, dynamic vegetation, and stomatal resistance schemes play a less important role than the runoff schemes for globally averaged soil moisture and ET. However, the effects on the soil moisture and ET relationship vary with different regions according to variations in soil moisture regimes and vegetation conditions. Although the Ball-Berry type stomatal resistance scheme behaves very similarly to the Javis type in modeling transpiration, the former scheme allows a direct coupling of terrestrial water and carbon fluxes and improves the simulation of vegetation–atmosphere interactions [*Niyogi et al.*, 2009].

References

Adam, J. C., and D. P. Lettenmaier (2003), Adjustment of global gridded precipitation for systematic bias, *J. Geophys. Res.*, *108*(D9), 4257, doi:10.1029/2002JD002499.

Adam, J. C., E. A. Clark, D. P. Lettenmaier, and E. F. Wood (2006), Correction of global precipitation products for orographic effects, *J. Clim.*, *19*, 15–38, doi:10.1175/JCLI3604.1.

Brown, R. D., B. Brasnett, and D. Robinson (2003), Gridded North American monthly snow depth and snow water equivalent for GCM evaluation, *Atmos. Ocean*, *41*(1), 1–14, doi:10.3137/ao.410101.

Chen, F., Z. Janjic, and K. E. Mitchell (1997), Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model, *Boundary Layer Meteorol.*, *85*, 391–421, doi:10.1023/A:1000531001463.

Chen, J. L., C. R. Wilson, and K. W. Seo (2006), Optimized smoothing of Gravity Recovery and Climate Experiment (GRACE) time-variable gravity observations, *J. Geophys. Res.*, *111*, B06408, doi:10.1029/2005JB004064.

Duan, Q., N. K. Ajami, X. Gao, and S. Sorooshian (2007), Multi-model ensemble hydrologic prediction using Bayesian model averaging, *Adv. Water Resour.*, *30*(5), 1371–1386, doi:10.1016/j.advwatres.2006.11.014.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley (2003), Implementation of Noah land surface model advancements in the National Centers for Environmental Prediction operational mesoscale Eta model, *J. Geophys. Res.*, *108*(D22), 8851, doi:10.1029/2002JD003296.

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.

Graham, S. T., J. S. Famiglietti, and D. R. Maidment (1999), Five minute, 1/2°, and 1° data sets of continental watersheds and river networks for use in regional and global hydrologic and climate system modeling studies, *Water Resour. Res.*, *35*(2), 583–587, doi:10.1029/1998WR900068.

Gutman, G., and A. Ignatov (1998), The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models, *Int. J. Remote Sens.*, *19*, 1533–1543, doi:10.1080/014311698215333.

Hollinger, S. E., and S. A. Isard (1994), A soil moisture climatology of Illinois, *J. Clim.*, *7*, 822–833, doi:10.1175/1520-0442(1994)007<0822:ASMCOI>2.0.CO;2.

Janzen, H. H. (2004), Carbon cycling in Earth systems—A soil science perspective, *Agric. Ecosyst. Environ.*, *104*(3), 399–417, doi:10.1016/j.agee.2004.01.040.

Jordan, R. (1991), A one-dimensional temperature model for a snow cover, *Spec. Rep. 91–16*, Cold Reg. Res. and Eng. Lab., U.S. Army Corps of Eng., Hanover, N. H.

Koren, V., J. C. Schaake, K. E. Mitchell, Q.-Y. Duan, F. Chen, and J. M. Baker (1999), A parameterization of snowpack and frozen ground intended for NCEP weather and climate models, *J. Geophys. Res.*, *104*, 19,569–19,585, doi:10.1029/1999JD900232.

Koster, R. D., and P. C. D. Milly (1997), The interplay between transpiration and runoff formulations in land surface schemes used with atmospheric models, *J. Clim.*, *20*(7), 1578–1591.