The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system

Kenneth E. Mitchell,[1] Dag Lohmann,[1] Paul R. Houser,[2] Eric F. Wood,[3] John C. Schaake,[4] Alan Robock,[5] Brian A. Cosgrove,[2] Justin Sheffield,[3] Qingyun Duan,[4] Lifeng Luo,[5,6] R. Wayne Higgins,[7] Rachel T. Pinker,[8] J. Dan Tarpley,[9] Dennis P. Lettenmaier,[10] Curtis H. Marshall,[1,11] Jared K. Entin,[2] Ming Pan,[3] Wei Shi,[7] Victor Koren,[4] Jesse Meng,[1,2] Bruce H. Ramsay,[9] and Andrew A. Bailey[9]

Received 1 June 2003; revised 10 October 2003; accepted 28 October 2003; published 9 April 2004.

[1] Results are presented from the multi-institution partnership to develop a real-time and retrospective North American Land Data Assimilation System (NLDAS). NLDAS consists of (1) four land models executing in parallel in uncoupled mode, (2) common hourly surface forcing, and (3) common streamflow routing: all using a 1/8° grid over the continental United States. The initiative is largely sponsored by the Global Energy and Water Cycle Experiment (GEWEX) Continental-Scale International Project (GCIP). As the overview for nine NLDAS papers, this paper describes and evaluates the 3-year NLDAS execution of 1 October 1996 to 30 September 1999, a period rich in observations for validation. The validation emphasizes (1) the land states, fluxes, and input forcing of four land models, (2) the application of new GCIP-sponsored products, and (3) a multiscale approach. The validation includes (1) mesoscale observing networks of land surface forcing, fluxes, and states, (2) regional snowpack measurements, (3) daily streamflow measurements, and (4) satellite-based retrievals of snow cover, land surface skin temperature (LST), and surface insolation. The results show substantial intermodel differences in surface evaporation and runoff (especially over nonsparse vegetation), soil moisture storage, snowpack, and LST. Owing to surprisingly large intermodel differences in aerodynamic conductance, intermodel differences in midday summer LST were unlike those expected from the intermodel differences in Bowen ratio. Last, anticipating future assimilation of LST, an NLDAS effort unique to this overview paper assesses geostationary-satellite-derived LST, determines the latter to be of good quality, and applies the latter to validate modeled LST. *INDEX TERMS:* 1878 Hydrology: Water/energy interactions; 1836 Hydrology: Hydrologic budget (1655); 1860 Hydrology: Runoff and streamflow; 1818 Hydrology: Evapotranspiration; 3337 Meteorology and Atmospheric Dynamics: Numerical modeling and data assimilation; *KEYWORDS:* land modeling, land data assimilation, surface energy budget

Citation: Mitchell, K. E., et al. (2004), The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *J. Geophys. Res.*, *109*, D07S90, doi:10.1029/2003JD003823.

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

[2]Hydrological Sciences Branch and Data Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

[3]Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.

[4]Office of Hydrologic Development, National Oceanic and Atmospheric Administration–National Weather Service, Silver Spring, Maryland, USA.

[5]Department of Environmental Sciences, Rutgers University, New Brunswick, New Jersey, USA.

[6]Now at Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.

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

[8]Department of Meteorology, University of Maryland, College Park, Maryland, USA.

[9]Office of Research and Applications, National Environmental Satellite Data and Information Service, Camp Springs, Maryland, USA.

[10]Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington, USA.

[11]Now at Department of Atmospheric Sciences, Colorado State University, Fort Collins, Colorado, USA.

Copyright 2004 by the American Geophysical Union.
0148-0227/04/2003JD003823$09.00

1. Introduction

[2] Improving weather and seasonal climate prediction by dynamical models requires multidisciplinary advances in providing reliable initial states for the atmosphere, ocean and land components of the Earth system. For two decades, advances in providing atmospheric initial states via four-dimensional data assimilation (4DDA) have paved the way for emerging 4DDA systems for the ocean and land. The backbone of any 4DDA system is the geophysical model whose execution provides temporally and spatially continuous background states, into which generally discontinuous observations are assimilated from various observing platforms (in situ, satellite, radar). For example, present space-based microwave estimates of soil moisture sense only the top 1–5 cm of soil, far short of the root-zone depths needed for land-state initialization.

[3] Thus a land data assimilation system (LDAS) is needed to blend sparse land observations with the background fields of a land surface model (LSM). The accuracy of the LSM background field (and companion surface and subsurface water/energy fluxes) is crucial to LDAS viability. The chief objective of the North American Land Data Assimilation System (NLDAS) study here is to generate and validate, over a 3-year period over the continental U.S. (CONUS) domain, the background land states and surface fluxes of four LSMs: Noah, Mosaic, VIC, and Sacramento (denoted SAC) (hereinafter, all acronyms are defined in the Notation). Future NLDAS papers will address actual data assimilation experiments using such methods as adjoint models and Kalman filtering. As one step to assimilation of satellite land surface skin temperature (LST), this paper assesses geostationary-satellite-derived LST and uses it to validate NLDAS LST.

[4] It is instructive to consider the infancy of real-time large-scale land 4DDA. Global atmospheric 4DDA has been a mainstay of operational NWP centers since the late 1970s. Real-time ocean 4DDA on large-scale ocean basins followed in the middle to late 1980s [*Ji et al.*, 1994] on the heels of the TOGA program. Yet until the mid-1990s, initiatives in real-time continental or global land 4DDA were virtually nonexistent. The first viable examples of real-time land 4DDA on continental or global scales were the coupled land-atmosphere 4DDA systems at major NWP centers such as NCEP [*Kalnay et al.*, 1996] and the European Centre for Medium-Range Weather Forecasts [*Gibson et al.*, 1997]. Such coupled land-atmosphere 4DDA systems (including global reanalysis) often yield significant errors and drift in soil moisture/temperature and surface energy/water fluxes, owing to substantial biases in the surface forcing from the parent atmospheric models. To constrain such errors and drift, coupled land-atmosphere 4DDA systems temporally nudge the soil moisture by such means as (1) a climatology of soil moisture [*Kalnay et al.*, 1996], (2) differences between the observed and 4DDA background fields of precipitation [*Kanamitsu et al.*, 2002], or (3) screen-level air temperature and dew point [*Douville et al.*, 2000]. Such nudging methods, however, do not reduce the main error source, namely, large bias in the land surface forcing (especially precipitation and solar insolation) of the parent atmospheric model.

[5] Substantial biases in atmospheric model surface forcing also plague ocean 4DDA. To improve these surface fluxes, "flux corrections" are applied in ocean 4DDA [*Ji et al.*, 1994]. NLDAS here also applies surface flux corrections. As a pathfinder for this, the GEWEX Global Soil Wetness Project (GSWP) [*Dirmeyer et al.*, 1999] retrospectively demonstrated the viability of using nonmodel, observation-based precipitation analyses and nonmodel, satellite-based surface insolation fields (with all other surface forcing from atmospheric 4DDA) to drive uncoupled, land surface models over a global domain. However, the monthly satellite retrievals of precipitation and insolation used in GSWP are not conducive to the daily/weekly updates of land states needed to initialize operational prediction models. Hence the NLDAS project set and achieved the following key objectives: (1) develop and execute the first real-time operational prototype of a continental-scale uncoupled land 4DDA backbone (continuously cycled land-model states) executed daily at NCEP using real-time streams of hourly to daily data and (2) a companion retrospective mode for research. The NLDAS generates hourly surface forcing (using model-independent, observation-based precipitation and insolation fields) that drives four LSMs running in parallel to produce hourly output on a 1/8° grid over a CONUS domain.

[6] The retrospective NLDAS spans October 1996 to September 1999 and uses GCIP-supported archives of NOAA operational data streams. NLDAS thus provides a land 4DDA counterpart from the GEWEX community to complement the ocean 4DDA thrusts that followed TOGA. Moreover, a core objective of GCIP is the infusion of GCIP research into NOAA operational practice. The NLDAS partnership of operational and research investigators in both meteorology and hydrology is a flagship of GCIP success in such infusion. This paper gives an overview of the methodology and results of the initial development and evaluation of NLDAS, providing an overview of the nine papers by NLDAS partners given in Table 1 (hereinafter, each paper is cited with the label given in Table 1, denoting the last initials of first two authors and N for NLDAS).

[7] These papers and the sections that follow illustrate that a pillar feature of NLDAS is the integrated application of a multitude of GCIP-sponsored products, as listed in Table 2. Section 2 describes the NLDAS configuration, surface forcing, land models, and streamflow routing. Sections 3 and 4 evaluate the NLDAS surface water budget and surface energy budget, respectively. Section 5 presents conclusions and future plans.

2. NLDAS Configuration

2.1. General Configuration

[8] Pilot studies of ocean 4DDA began on relatively data-rich subglobal domains [*Ji et al.*, 1994], i.e., the tropical Pacific Ocean, which included the TOGA observing network. For the NLDAS domain, we also chose a relatively data-rich subglobal domain, thereby heeding a lesson from GSWP, namely that uncoupled land surface simulation is notably less viable over regions lacking moderately dense precipitation gages to anchor the precipitation forcing [*Oki et al.*, 1999]. Outside such regions, global precipitation analyses are dominated by satellite-based precipitation,

Table 1. List and Topics of the Nine Companion Papers by NLDAS Partners in the GCIP-3 Special Issue

	Label	Subject
Cosgrove et al. [2003a]	CL-N	generation of land surface forcing
Luo et al. [2003]	LR-N	validation of land surface forcing
Pinker et al. [2003]	PT-N	production/validation of GOES-based solar insolation
Lohmann et al. [2004]	LM-N	production/validation of streamflow and water budget
Robock et al. [2003]	RL-N	validation of energy budget, soil moisture/temperature
Schaake et al. [2004]	SD-N	evaluation of soil moisture storage and range
Sheffield et al. [2003]	SP-N	validation of simulated snow cover
Pan et al. [2003]	PS-N	validation of simulated snowpack content
Cosgrove et al. [2003b]	CM-N	evaluation and testing of spin-up

which may be only marginally better (or even worse) then model-based precipitation, especially in the extratropics and in winter. Hence we limit the NLDAS domain (shown in Figure 1) essentially to the CONUS, thereby benefiting from relatively dense precipitation gages and the CONUS-oriented GCIP-supported products in Table 2.

[9] On this domain, NLDAS applies the following in common across the four LSMs: a 1/8° regular latitude/longitude grid, land mask and terrain elevation, hourly input surface forcing, soil texture and vegetation classes, streamflow network and routing model, and content, frequency (hourly) and format (GRIB) of model input and output. The elevation was derived by averaging, in each 1/8° grid cell, the 30 arc-second (~1 km) digital elevation of the GTOPO30 database of *Verdin and Greenlee* [1996]. Of the four LSMs, VIC alone also employs subgrid elevation tiles (see section 3.4). The vegetation classification was derived from the global, 1-km, AVHRR-based, 13-class vegetation database of UMD [*Hansen et al.*, 2000]. For each 1/8° cell, the vegetation field includes the percent of each class based on its 1-km frequency. Mosaic and VIC use subgrid vegetation tiles, whose weights correspond to the percent of the classes. Noah uses the most predominant vegetation class. SAC omits explicit treatment of vegetation.

[10] The soil texture database over CONUS was derived from the 1-km STATSGO database of *Miller and White* [1998], which carries 16 texture classes by layer over 11 layers to 2-m depth. For each 1/8° grid cell, the NLDAS soil database carries the percent of each class by layer, based on the original 1-km frequency. Noah, Mosaic, and VIC assume a vertically uniform soil class based on the predominant soil texture of the top 5-cm layer. The exception is Mosaic's soil porosity, derived for each Mosaic soil layer based on weighted averages from the 11-layer soil textures. Though SAC uses conceptual soil-water storage reservoirs rather than explicit soil parameters of an explicit soil column, many a priori parameters for the NLDAS SAC were derived using all 11-layer soil types. Outside the CONUS, the NLDAS soil database applies the same 16 texture classes, but carries only a single, vertically uniform class at each cell, derived from the 5-min ARS FAO global data of *Reynolds et al.* [2000]. Spatial maps depicting the NLDAS vegetation and soils databases, as well as tables defining NLDAS vegetation and soil classes, may be viewed under the NLDAS tab at http://ldas.gsfc.nasa.gov, maintained by NASA.

[11] Although NLDAS control runs employ common fields of vegetation and soil class, the NLDAS partners chose NOT to impose additional commonality in the vegetation and soil properties, such as (1) parameter values, (2) configuration of a vegetation class (root depth and density) or the soil column (number and thickness of layers) and (3) seasonal cycle of vegetation. The desire was to avoid negating the legacy of calibration or tuning invested over the past decade in the LSMs. No extra calibration of LSMs was carried out for the control runs. (Additionally, the NLDAS project conducted various sensitivity tests, discussed throughout this overview.)

Table 2. GCIP-Supported Products Applied by the NLDAS Project

	Product
	A: For Producing NLDAS Surface Forcing
1	daily 1/8° gage-only CONUS precipitation analysis by NCEP
2	hourly 4-km radar-dominated (WSR-88D) CONUS precipitation analysis by NCEP and OHD
3	hourly 1/2° GOES-based CONUS surface insolation by NESDIS and UMD
4	3-hourly 40-km Eta-based 4DDA analyses of near-surface meteorology by NCEP
	B: For Validating NLDAS Surface Forcing
5	OU Mesonet surface meteorology observations
6	SURFRAD network of surface solar insolation observations (receives support from GCIP sister program in NOAA/OGP)
	C: For Validating NLDAS Land Model Output (States/Fluxes)
7	Oklahoma/Kansas ARM/CART surface flux stations (DOE)
8	Oklahoma Mesonet soil moisture/temperature observations (OU Climate Survey)
9	CONUS-wide GOES-based satellite LST (NESDIS and UMD)
10	Northern Hemisphere 23-km IMS daily snow cover analysis by NESDIS
11	Illinois Water Survey network of 18 soil moisture measuring stations
12	western U.S. network of SNOTEL observations (not GCIP supported)
13	USGS streamflow observations (not GCIP supported)
	D: For NLDAS Land Surface Characteristics
14	1-km CONUS soil texture database by Pennsylvania State University
15	NESDIS 0.144° global monthly NDVI-based vegetation greenness by NESDIS
	E: Improved Land/Hydrology Models (LSMs)
16	Noah LSM improvements (including in coupled EDAS) by NCEP, OHD and collaborators
17	VIC LSM improvements (Princeton University, University of Washington and collaborators)
18	SAC LSM improvements by OHD, NCEP and collaborators
19	Mosaic LSM improvements by NASA GSFC and collaborators
20	streamflow connectivity network and routing model by NCEP, University of Washington, Princeton University and OHD

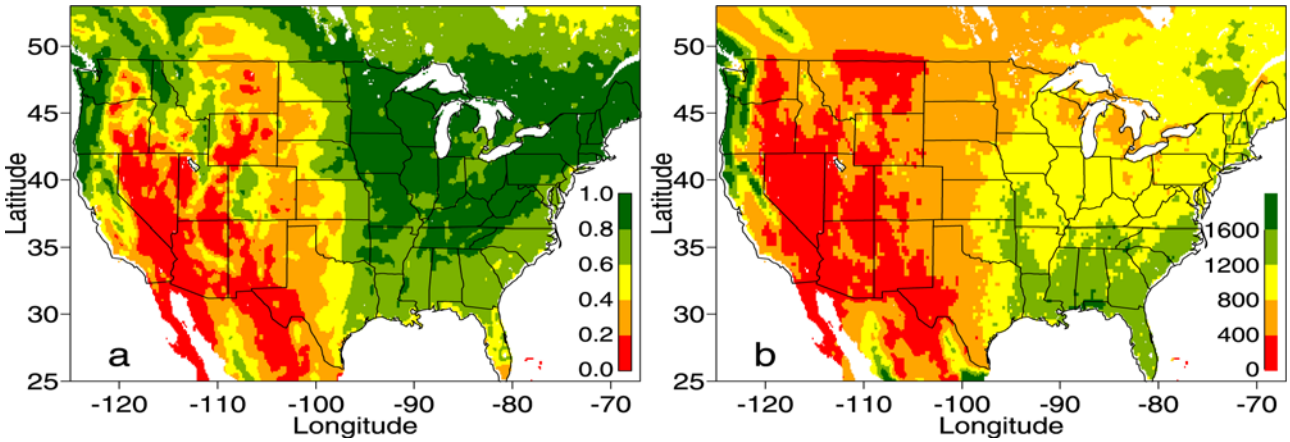

Figure 1. Depiction of NLDAS domain, showing (a) July mean green vegetation fraction from *Gutman and Ignatov* [1998] and (b) mean annual NLDAS precipitation (mm) for 1 October 1997 to 30 September 1999.

[12] For example, NLDAS does not impose a common treatment for the seasonality of vegetation, as the latter is central to a given model's canopy resistance formulation. In broad terms, Noah, VIC, and Mosaic runs here all use a satellite-derived, AVHRR-based, monthly seasonality of vegetation and all interpolate their respective monthly values to daily. Yet significant nuances exist between LSMs regarding whether the monthly values are for the given year (Mosaic) or from a multiyear climatology (Noah and VIC) and whether the seasonality is carried in LAI (VIC), or in vegetation fraction (Noah), or both (Mosaic). Noah uses the global, 0.144° (~15-km), monthly 5-year climatology of the green vegetation fraction (GVF) derived by *Gutman and Ignatov* [1998] of NESDIS from AVHRR-based NDVI. Figure 1a depicts this GVF climatology for the NLDAS domain for July, as a reference for later sections. For Mosaic, NASA obtained monthly 16-km AVHRR-based green LAI fields from Boston University for each month of each year (not climatology). NASA then first derived dead LAI (estimated from the difference in green LAI between consecutive months, along with vegetation-class-dependent values of minimum dead LAI) and then derived monthly total LAI (sum of green and dead LAI), from which GVF was computed (as green LAI divided by total LAI). VIC applies a global, AVHRR-based, multiyear monthly climatology of total LAI, used in conjunction with a vegetation-class-dependent look-up table of fixed annual-maximum vegetation fraction (dead and green).

[13] The LSMs of NLDAS provide common hourly output of about 50 required fields, including all terms of the surface energy and water budgets, all soil and snowpack moisture and temperature states, and ancillary fields. For a given LSM, some outputs are omitted if its physics omit the relevant process. NLDAS input/output fields are viewable at the NLDAS tab of http://ldas.gsfc.nasa.gov. For utility in NCEP operations, the input-output format of NLDAS is GRIB: a WMO standard at NWP centers. The NLDAS may be run in a "reduced-domain" mode by reducing the land mask, say to a single or handful of points (e.g., near flux stations), for purposes of lower output volume, faster execution, and locally focused sensitivity studies.

[14] The NLDAS requires initial values of all LSM state variables for the NLDAS start time of 00 UTC on 1 October 1996. Initial snowpack was set to zero (reasonable for 1 October over the NLDAS domain at 1/8° resolution), as was canopy interception storage. Initial states of soil moisture and temperature were derived from the soil states of the NCEP/DOE Global Reanalysis 2 [*Kanamitsu et al.*, 2002] valid at the start time. The soil moisture was provided to each LSM as a vertically uniform percent of saturation, which each LSM converted to its own absolute moisture state compatible with its parameters. The spin-up from this cold start was examined in all four LSMs by CM-N, who concluded that the practical drift in NLDAS land stores ceased within about one year. CM-N also conducted spin-up experiments in Mosaic, in which spin-up from the above initial states was found to be shorter than using saturated or dry initial states.

2.2. Surface Forcing

[15] The studies by *Cosgrove et al.* [2003a] (CL-N), *Pinker et al.* [2003] (PT-N), and *Luo et al.* [2003] (LR-N) summarized below describe the data sources, generation and validation of NLDAS forcing, produced in real-time and retrospectively on the NLDAS grid. Of the 16 fields in each forcing file (Table 3), nine fields required by Noah, Mosaic, and VIC are primary: U/V 10-m wind components, 2-m air temperature and specific humidity, surface pressure, downward longwave and shortwave radiation, and convective and total precipitation. SAC requires only total precipitation (P), air temperature and potential evaporation (PE). In NLDAS, SAC uses the PE computed in the Noah LSM. Mosaic alone requires convective precipitation.

[16] The chief source of NLDAS forcing is NCEP's Eta-model-based Data Assimilation System (EDAS) [*Rogers et al.*, 1995], a continuously cycled North American 4DDA system. It utilizes 3-hourly analysis-forecast cycles to derive atmospheric states by assimilating many types of observations, including station observations of surface pressure and screen-level atmospheric temperature, humidity and U and V wind components. EDAS 3-hourly fields of the latter five variables plus surface downward shortwave and longwave radiation and total and convective precipitation are provided

Table 3. Content and Data Sources of the Fields in the Hourly Surface Forcing Files of NLDAS[a]

Content	EDAS	GOES	Gage	Radar	Real-Time	Retrospective
Primary forcing						
2-m temperature, K	X				X	X
2-m specific humidity, kg/kg	X				X	X
10-m U-wind component, m/s	X				X	X
10-m V-wind component, m/s	X				X	X
Surface pressure, mb	X				X	X
Downward longwave radiation, W/m^2	X				X	X
Downward shortwave radiation, W/m^2		X			X	X
Convective precipitation, kg/m^2	X				X	X
Total precipitation, kg/m^2			X[b]	X[b]	X	X
Backup forcing						
Downward shortwave radiation, W/m^2	X				X	X
Total precipitation. kg/m^2	X				X	X
Auxiliary forcing						
WSR-88D precipitation, kg/m^2			X[c]	X[c]		
PAR, W/m^2		X			X	X
Downward diffuse radiation,[d] W/m^2		X			X	
CAPE	X				X	X
For validation (plus future assimilation)						
Land surface temperature (LST),[e] K		X			X	X

[a]To date, retrospective forcing is available from 1 October 1996 through 2002. Real-time forcing is available from 16 April 1999 to present.
[b]Daily total is gage-only. Radar estimate is used only to temporally partition gage-based daily into hourly.
[c]WSR-88D precipitation estimate is radar dominated, but some gage data are used.
[d]Diffuse radiation is present in forcing files since 15 November 1999.
[e]LST is present in forcing files since 1 May 1997 (LST derivation and application in section 4.3).

on a 40-km grid to NLDAS forcing software, which interpolates the fields spatially to the NLDAS grid and temporally to one hour. Last, to account for NLDAS versus EDAS surface-elevation differences, a terrain-height adjustment is applied to the air temperature and surface pressure using a standard lapse rate (6.5 K km^{-1}), then to specific humidity (keeping original relative humidity) and downward longwave radiation (for new air temperature, specific humidity). CL-N details the spatial/temporal interpolations and terrain-height adjustment.

[17] EDAS precipitation and shortwave radiation serve only as backup (Table 3). Though *Roads et al.* [2003] found EDAS monthly precipitation to have rather smaller errors than other 4DDA systems, EDAS precipitation errors can be significant for daily events, as in summer convection (Figure 4 of CL-N). Thus NLDAS precipitation forcing over CONUS is anchored instead to NCEP's 1/4° gage-only daily precipitation analyses of *Higgins et al.* [2000] (available at http://www.cpc.ncep.noaa.gov/research_papers/ncep_cpc_atlas/7/index.html), which utilize about 6500 (real-time) or 13000 (retrospectively) gage observations of daily precipitation. In NLDAS, this daily analysis is interpolated to 1/8°, then temporally disaggregated to hourly (details given by CL-N) by applying hourly weights derived from hourly, 4-km, radar-based (WSR-88D) precipitation fields. The latter radar-based fields (saved as auxiliary field in Table 3) are used only to derive disaggregation weights and do not change the daily total precipitation. Last, convective precipitation is estimated by multiplying NLDAS total precipitation by the ratio of EDAS convective to EDAS total precipitation. Figure 1b shows the annual mean NLDAS precipitation for the two years of the water budget analysis done later in section 3.1. CL-N shows examples of hourly and daily NLDAS precipitation fields, as well as EDAS and radar-based counterparts.

[18] Downward shortwave radiation (solar insolation) in the EDAS and Eta model typically show high bias of 10–20% [*Betts et al.*, 1997], even higher in cloudy winter conditions. At two SURFRAD sites, Figure 7 of CL-N illustrates the high bias in EDAS insolation and the far less bias in GOES-based solar insolation, which provides the primary insolation forcing for NLDAS. PT-N describes the retrieval of this 1/2° surface insolation from GOES since January 1996, via collaboration of UMD, NESDIS/ORA and NCEP, and its hourly to monthly validation against fifty CONUS stations over a 1–2 year period. GOES insolation is not retrieved for zenith angles below 75° and so is supplemented with EDAS insolation near the day/night terminator (Figure 5 of CL-N). In validations by PT-N and LR-N, GOES insolation verifies well against flux stations, with some deterioration toward high bias at low sun angles and over snow cover (when cloud detection is difficult). CL-N shows that even in winter, the high bias of the GOES insolation is about half that of EDAS. Last from the GOES-based product suite, downward diffuse radiation, PAR and LST fields are included in the NLDAS forcing files (Table 3).

[19] NCEP originally selected the viable real-time NOAA data sources (Table 2, part A) on which to base NLDAS surface forcing fields and then developed the algorithms for their real-time production, which NCEP has sustained since April 1999. GCIP has supported the archiving back to 1996 of all NOAA data streams needed for this forcing. NASA GSFC acquired these archives and adapted NCEP software to produce retrospective forcing for October 1996 through 2002, the first 3 years of which force the NLDAS executions evaluated here and in the NLDAS papers. The retrospective forcing was created for purposes of (1) executing NLDAS over longer periods, especially those overlapping special validating observations (Table 2), such as

Table 4. Primary Attributes of the Four NLDAS Land Surface Models (LSMs)

	Mosaic	Noah	VIC	SAC
Full domain runs	yes	yes	yes	yes
Limited domain runs	yes	yes	yes	yes
Input surface forcing	seven forcing fields[a]	seven forcing fields[a]	seven forcing fields[a]	precipitation, Noah PE, 2-m air temperature
Energy balance	yes	yes	yes	n/a
Water balance	yes	yes	yes	yes
Model time step	15 min	15 min	1 hour	1 hour
Model soil layers	3	4	3	2 storages
Model soil layer depths	10, 30, 160 cm	10, 30, 60, 100 cm	10 cm, variable	variable
Tiling: Vegetation	yes	no	yes	no
Tiling: Elevation	no	no	yes	no
Snow model layers	1	1	2	1
Frozen soil: thermal	no	yes	disabled	n/a
Frozen soil: hydraulics	partial	yes	disabled	n/a
Soil thermodynamics	force-restore	heat conduction equation	heat conduction equation modified	no
Soil temperature profile	no	yes	yes	no
Soil water: drainage	yes	yes	yes	yes
Soil water: vertical diffusion	yes	yes	no	no
Snow-free albedo	vary wrt LAI, GVF, biome	monthly input background field	vary wrt LAI and biome	n/a
Diurnal albedo	yes	no	no	n/a
Explicit vegetation	yes	yes	yes	no
Canopy resistance	*Sellers et al.* [1986]	*Jarvis* [1976]	*Jarvis* [1976]	n/a
Rooting depth	0.4 m	variable (1 or 2 m)	variable (1.35–3 m)	n/a
Root density profile	constant	constant	exponential	n/a
Canopy capacity	0–1.6 mm	0.5 mm	0.1–1.0 mm	n/a

[a]Seven forcing fields: precipitation, downward solar and longwave radiation, 10-m wind speed, surface pressure, 2-m air temperature, 2-m air humidity.

the soil moisture used in RL-N, (2) leveraging the near doubling of gage observations (about 13000 versus 6500) of daily precipitation applied in the retrospective versus real-time CONUS precipitation analyses of *Higgins et al.* [2000], and (3) applying added quality control checks to and (when such checks warrant) reprocessing of the forcing that is not feasible in real time.

[20] The LR-N study assesses NLDAS retrospective forcing of January 1998 to September 1999 against hourly ARM/CART and OU Mesonet stations (yielding independent observations, not assimilated in EDAS). Except for precipitation, differences between NLDAS forcing and these observations were small at hourly to monthly timescales. For precipitation the agreement was marginal at hourly periods, but better at daily and rather good at 5-day and monthly periods. In net radiation, a small low bias in downward longwave partially offset the modest high bias in solar insolation. To investigate how these differences impacted NLDAS simulations, LR-N compared control simulations using standard NLDAS forcing with test simulations using site-specific, station-observed forcing. Simulation differences in soil moisture and temperature for each LSM were small: much smaller than differences between the LSMs and between LSMs and observations. Thus NLDAS provides quality forcing for land modeling, at least over the nonwestern CONUS. In section 3.4, the study by PS-N uncovers a 50% low bias in NLDAS precipitation at mountain SNOTEL sites at high elevations in western CONUS. In the final section, we discuss future remedies that will apply an adjustment to PRISM [*Daly et al.*, 1994].

2.3. Land Models

[21] Table 4 compares the attributes of the LSMs in NLDAS. Of the many LSMs, these four give a good cross-section of different early legacies, including small scale versus large scale, coupled versus uncoupled, distributed versus lumped, with and without explicit vegetation, tiled and nontiled, and significant versus minimal calibration. Mosaic and Noah emerged from the surface-vegetation-atmosphere transfer (SVAT) setting of coupled atmospheric modeling with little calibration. VIC and SAC grew from the hydrology community as uncoupled hydrology models with considerable calibration. Mosaic was developed for use in the NASA global climate model [*Koster and Suarez*, 1994, 1996; *Koster et al.*, 2000]. Noah was developed for use in the NCEP mesoscale Eta model [*F. Chen et al.*, 1997; *Betts et al.*, 1997; *Ek et al.*, 2003]. VIC was developed as a macroscale semi-distributed model [*Liang et al.*, 1994; *Wood et al.*, 1997]. SAC was developed as a lumped conceptual hydrology model [*Burnash et al.*, 1973], calibrated for small catchments and used operationally in NWS RFCs.

[22] Subsequent to their early heritage, Mosaic, Noah, and VIC have been widely executed coupled and uncoupled from small to large scales. Now all three models can be considered as both SVATs and semi-distributed hydrological models. All three have undergone testing on local and regional scales in the PILPS project [*T. H. Chen et al.*, 1997; *Wood et al.*, 1998; *Schlosser et al.*, 2000; *Bowling et al.*, 2003] and on the global scale in GSWP [*Dirmeyer et al.*, 1999]. We use "semi"-distributed to mean applied on a gridded basis with gridded state variables and gridded parameters, but no horizontal interaction between model grid cells, except for routing of gridded runoff into streamflow. Seeking SAC suitability over a broad range of scales, OHD of NWS recently developed a semi-distributed (nonlumped) version of SAC [*Koren et al.*, 2000] with a priori uncalibrated parameters. This SAC version is intended for testing from small basins to entire continents. NLDAS

provides the first tests of the semi-distributed SAC at continental scales. These are pilot tests, as SAC lacks the legacy of continental testing of the other LSMs.

[23] The three SVAT models simulate LST, the surface energy and water balance, snowpack, and soil moisture in several soil layers, though the number and thickness of the layers differ. Only Noah simulates soil freeze-thaw and its impact on soil heating or cooling and transpiration, after *Koren et al.* [1999]. The snowpack physics in the LSMs are described in section 3.4. In all three SVATs, the surface infiltration schemes account for subgrid variability in soil moisture and precipitation, but the treatments differ, as do the drainage approaches. All three SVATs include direct evaporation from soil, transpiration from vegetation, evaporation of interception, and snow sublimation; and all explicitly model canopy resistance, though their formulations (see Table 4) and parameters differ, as does their vegetation phenology (section 2.1) and root profiles. The aerodynamic conductance in the SVATS also differs, a focus of section 4.2.

[24] SAC is a conceptual rainfall-runoff, storage-type model [*Burnash et al.*, 1973]. It treats only the surface water budget, omitting the surface energy budget, and uses the snowpack model of *Anderson* [1973], called SNOW-17. Hereafter SAC means the SAC-SNOW-17 pair. SAC outputs evaporation E and runoff, with E being a fraction of input PE. SAC uses a "two-reservoir" soil water storage structure (a shallow upper reservoir and a deeper lower reservoir) and utilizes 28 parameters, 16 primary and 12 in SNOW-17. At the RFCs, the parameters are calibrated by catchment. Calibration was omitted in NLDAS SAC runs. Rather, the primary parameters are specified a priori, after *Koren et al.* [2000], as a function of the STATSGO-based soils of section 2.1 and other data sets. SNOW-17 parameters are prescribed uniformly over the domain.

[25] Mosaic was developed by *Koster and Suarez* [1994, 1996] to account for subgrid vegetation variability with a tile approach. Each vegetation tile carries its own energy and water balance and soil moisture and temperature. Each tile has three soil layers and the first two are the root zone. In NLDAS, Mosaic is configured to support a maximum of 10 tiles per grid cell with a 5% cutoff that ignores vegetation classes covering less than 5% of the cell. Additionally in NLDAS, all tiles of Mosaic in a grid cell have the predominant soil type of section 2.1 and three soil layers with fixed thickness values of 10, 30, and 160 cm (hence constant rooting depth of 40 cm and constant total column depth of 200 cm). This Mosaic configuration in NLDAS departs from the standard Mosaic configuration, for the purpose of easier comparison in NLDAS [*Robock et al.*, 2003] with the soil moisture observation levels of the OU Mesonet and the soil layers of the VIC and Noah models (e.g., their 10 cm top layer). Although never executed before with fixed layer thickness, Mosaic performed well in the PILPS experiments when configured in the standard way [*T. H. Chen et al.*, 1997; *Lohmann et al.*, 1998; *Wood et al.*, 1998]. The standard Mosaic configuration varies the soil type and layer thickness tile by tile by vegetation type and yields top-down layer thickness ranges of 1–2 cm, 1–150 cm, and 30–200 cm, total column depth ranges of 32–350 cm, and root depths of 2–49 cm for nonforest and 150 cm for forests.

[26] The Noah LSM [*Chen et al.*, 1996; *Koren et al.*, 1999; *Ek et al.*, 2003] is targeted for moderate complexity and computational efficiency for operational NWP and climate models. Thus it omits tiling and uses a single-layer snowpack, plus a linearized (noniterative) solution to the surface energy balance. Originating from the LSM of *Pan and Mahrt* [1987], Noah benefits from improvements arising from year-round assessment in the NCEP Eta model over North America by NCEP and collaborators [*Ek et al.*, 2003]. The Noah version here in NLDAS is that implemented in the NCEP Eta/EDAS suite on 19 June 2002 and includes four layers of fixed thickness (Table 4), of which the first three (nonforest) or four (forest) span the root zone. Virtually this same version of Noah was executed in NCEP's 24-year Regional Reanalysis. *Berbery et al.* [2003] examines the large-scale hydrology of the coupled Eta/Noah model over the Mississippi Basin for the period June 1995 to May 2002.

[27] The variable infiltration capacity (VIC) model was developed at the University of Washington and Princeton University [*Liang et al.*, 1994, 1996a, 1996b; *Cherkauer and Lettenmaier*, 1999]. In NLDAS, VIC executes with one-hour time step and uses three soil layers, with 10 cm top layer and varying depth for bottom two layers, partly determined from calibration. The root zone can span all three layers, depending on vegetation class. Like Mosaic, the VIC model carries subgrid vegetation tiles. Additionally, VIC is the only LSM of the four to apply subgrid elevation bands or tiles (see section 3.4). VIC has been tested over large river basins, such as the Columbia [*Nijssen et al.*, 1997] and Arkansas-Red [*Abdulla et al.*, 1996; *Wood et al.*, 1997], and over continental scales [*Maurer et al.*, 2002; *Roads et al.*, 2003] and global scales [*Nijssen et al.*, 2001]. Traditional executions of VIC [e.g., *Maurer et al.*, 2002] apply a uniform disaggregation of total daily precipitation to VIC time steps (typically 3-hourly). In NLDAS, VIC is executed with one-hour time steps and nonuniform, radar-based disaggregation of daily precipitation (see section 2.2). Implications of this departure from standard VIC are presented in section 3.2.

2.4. Streamflow Simulation and Assessment

[28] *Lohmann et al.* [2004] (LM-N) present the formulation of the streamflow modeling in NLDAS and the assessment of the control-run simulations of daily streamflow by the four LSMs. The latter study assesses model streamflow for 9 major and 1145 small to medium-sized CONUS basins (ranging from 23 km^2 to 10,000 km^2) using measured daily streamflow from the USGS. LM-N describes the criteria for choosing the 1145 basins. One criterion is the absence of obvious regulation signatures in the observed streamflow record. This yields few basins for assessment in the arid southwestern CONUS (see Figure 4).

[29] The streamflow routing requires both a river network (flow-direction mask) on the NLDAS grid and a routing model. LM-N derives and displays the river network. The chosen routing model is linear and identical to that in PILPS 2c and 2e [*Lohmann et al.*, 1998; *Bowling et al.*, 2003]. It calculates the timing of the runoff reaching the

grid-cell outlet, as well as the transport of water through the river network. It operates in two modes: (1) distributed, using a-priori grid-cell specific routing parameters common to all four models and (2) ''lumped'', in which constant routing parameters were separately calibrated for each of the 1145 basins for each model to minimize the least squares difference between modeled and measured daily streamflow. The lumped mode thus yielded a separate calibrated unit hydrograph for each basin for each model.

[30] LM-N shows the streamflow time series of the Nehalem River in Oregon and associates the derived lumped routing function for each LSM with the timing of the runoff produced by each LSM (Figures 10 and 11 of LM-N). In turn, a major cool season impact on the modeled runoff is the timing of snowmelt in each LSM. LM-N quantifies the delay between modeled and measured streamflow by means of the maximum of the temporal cross-correlation between measured and modeled streamflow. In general for all the LSMs, streamflow performance was degraded with increasing snowfall amounts (Figure 17 of LM-N). The worst case of peak streamflow timing occurs in the snowpack season of the mountain ranges of the northwest CONUS. There the LSMs vary by up to four months in the timing of peak streamflow (Figure 18 of LM-N). VIC showed the most realistic timing in such regions. Mosaic and SAC melted on the order of weeks too early, and Noah on the order of months too early. The SM-N and PS-N studies featured in section 3.4 further examine and elucidate the snowpack and snow cover simulations and forcing.

[31] In contrast, throughout the eastern half of CONUS, streamflow simulation skill as measured by the Nash-Sutcliffe efficiency (Figure 16 of LM-N) showed that Noah had the highest scores in general; VIC had the highest scores in the northeast, and Mosaic and SAC in the southeast Atlantic coastal states. The higher skill for Noah streamflow reflected Noah having the smallest bias in evaporation and runoff when assessed against the observed annual water budget (next in section 3.1). Finally, the validation of simulated monthly discharge for the 9 large river basins showed behavior and bias in each LSM consistent with that expected from spatial integration of the behavior and effects established in the assessment of the smaller basins.

3. Assessment of the NLDAS Water Budget

[32] Sections 3 and 4 assess the water and energy budgets of the LSMs, respectively, in the three-year NLDAS simulation with retrospective forcing for 1 October 1996 to 30 September 1999. The assessment focuses on the last two years, termed the ''control'' simulation, since the first year is a necessary and (mostly) sufficient spin-up year, as reported by *Cosgrove et al.* [2003b] (CM-N). Additionally, complimentary sensitivity tests of the LSMs in NLDAS are assessed. Altogether, the hallmark of the assessment is its breadth, addressing continental to local scales, all four seasons and multiple types of validating observations (Table 2, part C).

3.1. Annual Water Budget: Partitioning Between Evaporation and Runoff

[33] Over one or more annual cycles, the surface water budget is well approximated by mean annual precipitation being equal to the sum of mean annual evaporation and mean annual runoff, since mean annual storage change (in soil moisture, snowpack, etc.) is negligible by comparison. Thus analysis of observed mean annual precipitation minus observed mean annual runoff (from observed streamflow) yields crucial observation-based estimates of mean annual evaporation, as was carried out for NLDAS by *Lohmann et al.* [2004] (LM-N) and summarized here.

[34] Figure 1b depicts the observed mean annual precipitation forcing common to each LSM in the NLDAS control run. Figure 2 depicts the simulated mean annual evaporation of each control-run LSM. (See Figure 6 of LM-N for companion map of simulated mean annual runoff.) Last, Figure 3a shows the partitioning of mean annual NLDAS precipitation into mean annual evaporation and runoff by each LSM for the four CONUS quadrants of SW, NW, NE, and SE depicted in the figure inset. In Figure 3a, each diagonal denotes the mean-annual area-averaged precipitation of a given quadrant (given by the diagonal's x or y axis intercept). On each diagonal, each LSM's symbol projected onto the x axis (y axis) yields that LSM's quadrant-average mean annual runoff (evaporation). Since each LSM conserves water over the annual cycle, the tiny displacement in Figure 3a of a given LSM symbol from the diagonal represents the negligible change in that LSM's total water storage over the two years.

[35] The disparity in evaporation among LSMs in Figures 2 and 3a is striking, especially over the well-vegetated NE and SE (Figure 1a). Disparity is far less over the arid SW and NW. Over the NE and SE, Noah and VIC have notably lower evaporation and hence higher runoff than Mosaic and SAC, which give high evaporation and low runoff. Mosaic similarity to SAC in Figures 2 and 3 is surprising, as Mosaic (like Noah and VIC) includes explicit treatment of vegetation cover and canopy conductance, while SAC does not. Large disparity among LSMs in evaporation versus runoff partitioning was noted in PILPS and GSWP [*Wood et al.*, 1998; *Dirmeyer et al.*, 1999].

[36] The partitioning can be validated over subregions of the quadrants for which basin-observed streamflow is available. Figure 3b, which uses observed streamflow from the 1145 assessment basins of section 2.4, is the counterpart to Figure 3a obtained by area averaging the NLDAS observed precipitation and LSM simulated evaporation and runoff only over the quadrant subarea spanned by these basins (Figure 3b inset). For each basin, observed streamflow (m^3s^{-1}) is converted to mean-annual total discharge (m^3), in turn converted (using the basin area) to area-average mean-annual runoff (mm) for the basin. The plus symbol in Figure 3b depicts the area-mean of this observed runoff over the same quadrant subarea. Projecting the plus symbol onto the y axis yields the budget-based estimate of the area-average mean annual evaporation.

[37] The reliability of the observation-based evaporation estimate depends on the reliability of both the observed streamflow (high reliability) and the NLDAS precipitation forcing. Over NE and SE, which manifest relatively flat terrain and good density of precipitation gages, we trust the precipitation analysis and the estimates of evaporation there. For NE and SE, one sees in Figure 3b that evaporation and

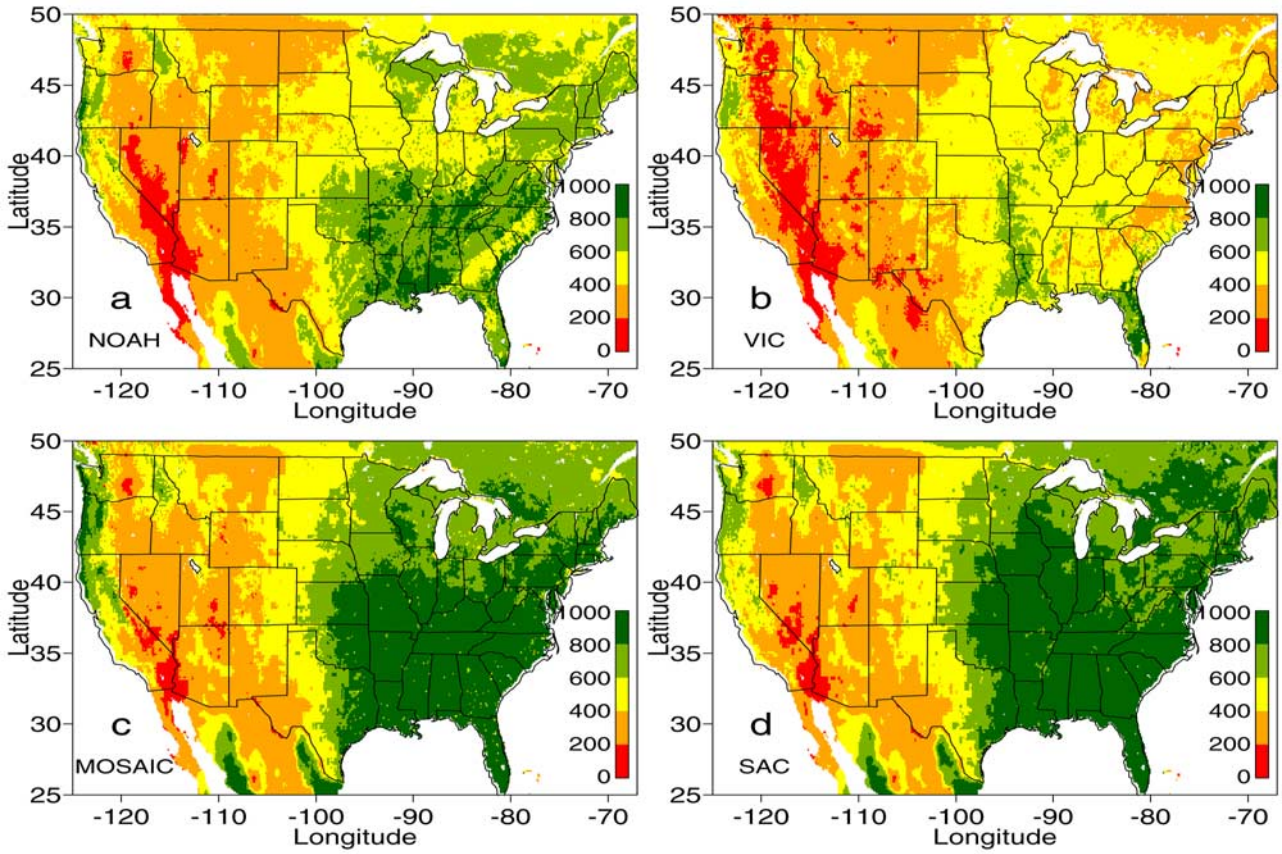

Figure 2. Mean annual evaporation (mm/year) in NLDAS from (a) Noah, (b) VIC, (c) Mosaic, and (d) SAC for 1 October 1997 to 30 September 1999.

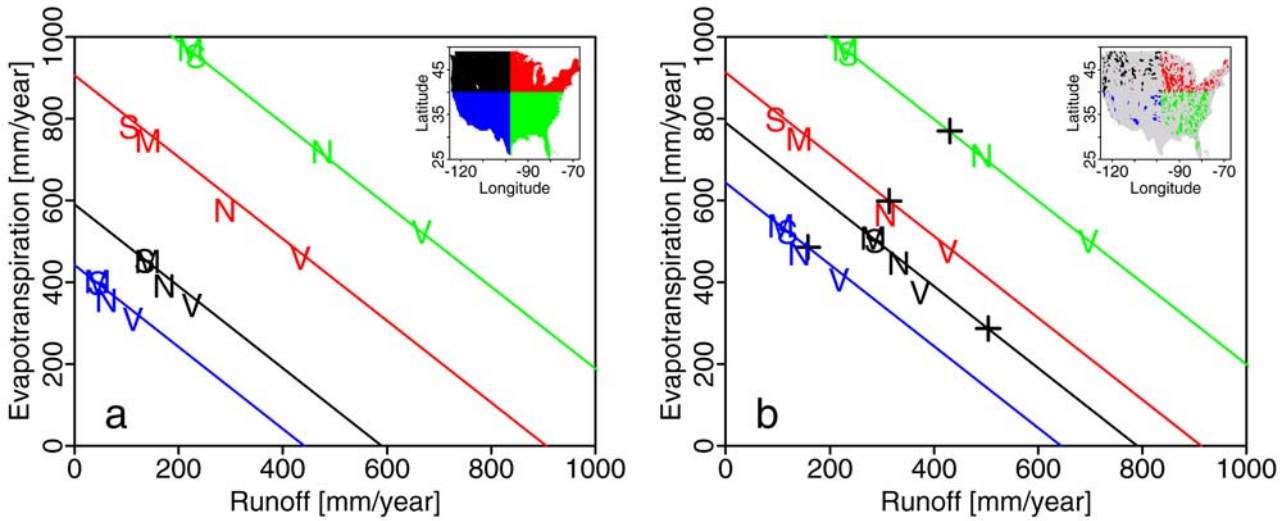

Figure 3. (a) Partitioning of mean annual area-mean precipitation (diagonal, mm/year) into mean annual area-mean runoff (*x* axis, mm/year) and evaporation (*y* axis, mm/year) for the CONUS quadrants (inset) of NW (black), NE (red), SW (blue), and SE (green) by Noah (N), VIC (V), Mosaic (M), and SAC (S) for 1 October 1997 to 30 September 1999. Model symbols below diagonal indicate (negligible) positive storage change. (b) As in Figure 3a, except area-mean is for subarea of basin set depicted in inset for each quadrant (same basins as in Figure 4) and the plus symbol depicts observed mean annual area-mean runoff for the same subarea.

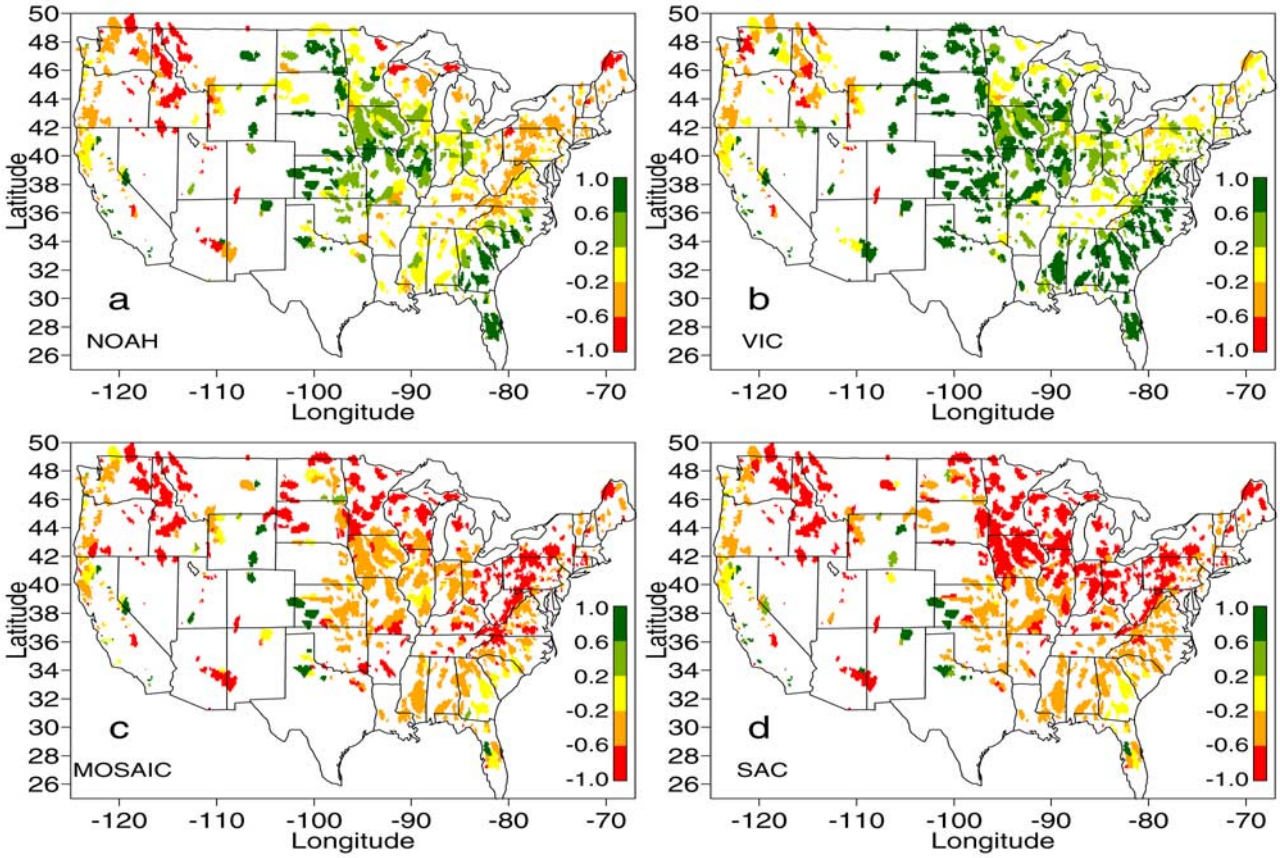

Figure 4. Relative bias [(model - observed)/observed] of mean annual runoff for selected basins for 1 October 1997 to 30 September 1999 from (a) Noah, (b) VIC, (c) Mosaic, and (d) SAC. Observed runoff for a given basin is calculated from basin area and observed basin-outlet stream discharge provided by USGS.

runoff of Noah are close to observed, while Mosaic and SAC show large biases of high evaporation and low runoff, with VIC yielding the reverse: large biases of low evaporation and high runoff. Figure 4 shows the variability of relative runoff bias [(model-observed)/observed] across the basins. Over NE and SE, Mosaic and SAC show similar patterns of pervasive underestimates of runoff, exceeding 60% (dark red) in the Appalachians and upper Midwest, while VIC runoff is highly overestimated (except for a corridor of near neutral bias west of the Appalachians), often by more than 60% (dark green) in the southeast and Midwest. Noah has the least runoff bias over NE and SE, with a more balanced likelihood of positive or negative bias, and fewer basins of (dark green) large bias.

[38] In NW, NLDAS precipitation has a large low bias (see section 3.4), owing to mountainous terrain, sparse precipitation gages, and lack of an adjustment to PRISM [*Daly et al.*, 1994] in the NLDAS precipitation analysis. This low bias thwarts the reliability of the NW budget-based evaporation estimate in Figure 3b and causes large low bias in runoff in all the LSMs over NW in Figures 3b and 4. For example, the LSMs substantially underestimate runoff in the Northern Rockies (dark red), though VIC yields much smaller bias there for reasons given in section 3.4. Last, over SW, in addition to precipitation gages being sparse, the number of assessment basins there in Figure 4 is small, casting uncertainty on observation-based SW evaporation estimates.

3.2. Monthly Water Budget and Soil Moisture Change

[39] In studying nonannual water budgets, storage changes of soil moisture and snowpack are important. Thus, on monthly scales, we apply the full surface water budget equation given by

$$\frac{dS_1}{dt} + \frac{dS_2}{dt} + \frac{dSn}{dt} = P - E - R_1 - R_2. \quad (1)$$

Each term is the area average, of storage change (left side) and accumulations (right side), of water mass per unit area (kg/m^2) per month, or depth of water (mm) per month. S_1 and S_2 are soil moisture stores of an upper and lower zone (defined later), respectively, and S_n is snowpack storage. P, E, R_1 and R_2 are precipitation, evaporation, surface runoff, and subsurface runoff, respectively. On monthly scales, canopy interception storage change is negligible and omitted in equation (1). Figure 5 gives the time series of area-average monthly evaporation E for each CONUS quadrant for the control runs.

[40] We focus here on the eastern quadrants, where evaporation in section 3.1 showed the most disagreement. During middle and late summer, Mosaic clearly has the

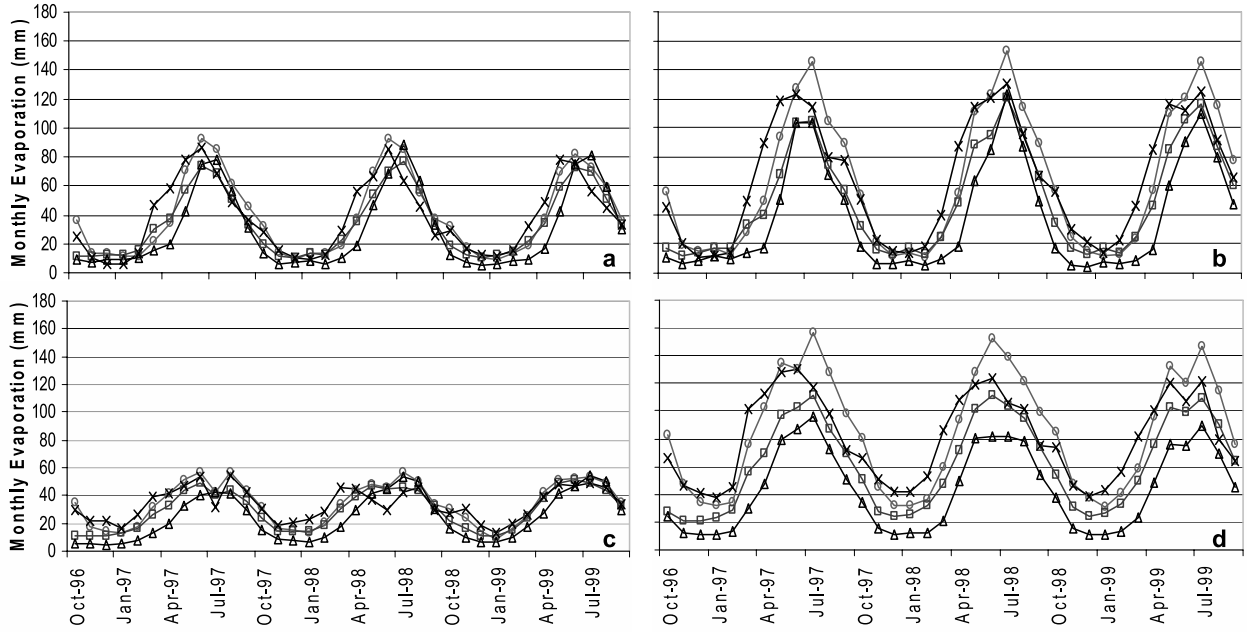

Figure 5. Time series of area-averaged monthly evaporation (mm/month) in NLDAS for Noah (squares), VIC (triangles) Mosaic (circles), and SAC (crosses) over the four CONUS quadrants of (a) NW, (b) NE, (c) SW, and (d) SE for October 1996 to September 1999.

highest evaporation, while SAC has the highest in winter and spring. Noah evaporation generally falls between that of Mosaic and VIC in the warm season. VIC has the lowest evaporation in virtually every month in the vegetated eastern quadrants, consistent with the earlier annual results. Figures 5b and 5d strongly suggest that Mosaic and VIC manifest rather different canopy conductance. This is most evident in the SE during 1998, where VIC monthly evaporation tops off at 80 mm during May–August while that of Mosaic sustains 120–150 mm. (Signatures of this middle and late summer stress over vegetation in VIC compared to Mosaic and Noah is evident again later in Figure 13b.)

[41] Such large differences in warm season evaporation imply large warm season differences in soil moisture storage change. Figure 6 shows the time series of area-average monthly mean total column soil moisture ($S_1 + S_2$) for all four quadrants. Figure 7 is the companion depiction for root zone soil moisture for the vegetated eastern quadrants. (Figure 7 excludes SAC, which has no explicit root zone.) The emphasis below is on years 2–3 of Figures 6 and 7, as nontrivial spin-up is evident in year one of Figure 6, though more so in Noah and VIC. After year one, spin-up is essentially complete (though small spin-up may still be occurring in NW in Noah).

[42] Focusing then on years 2–3, inspection of Figures 5–7 reveals the following: (1) very different levels of time-mean total soil moisture across the models, ranging in the SE from about 325 mm for VIC and SAC to 550 mm for Mosaic and 650 mm for Noah; (2) more similarity, yet important differences, in annual-cycle amplitude (seasonal change) of total soil moisture among the models, ranging in the SE from about 100–150 mm for Noah, VIC, and SAC to a notable high of 230 mm for Mosaic; (3) larger differences among the models over the wetter eastern quadrants than the drier western quadrants, in both total soil moisture and its seasonal change or range; (4) among the three models with a root zone, the contribution of the subroot zone to the change in total soil moisture varies widely, e.g. over the SE it is very large in Mosaic (about 180 of 230 mm), moderate in Noah (about 70 of 140 mm), and small in VIC (about 10 of 100 mm); (5) the model with the highest level of total soil moisture is not the model with the largest seasonal change in soil moisture, nor the model with the largest monthly/annual evaporation; (6) a model with high annual evaporation is not necessarily a model having a high annual range of soil moisture storage (as the counterpart of SAC illustrates).

[43] Intriguingly in Figure 6, while VIC and Noah have very different levels of total soil moisture in the SE and NE, they have more similar magnitudes of soil moisture in the root zone and seasonal change of soil moisture in the root zone in Figure 7. Thus the moisture source for the higher summer evaporation in Noah then VIC is Noah's subroot zone, consistent with Noah usually having a deeper total soil column (maximum storage capacity) than VIC, as shown in SD-N.

[44] Recalling that Mosaic executions apply a shallow root depth of 0.4 m, compared to 1–2 m in Noah and 1.35–3 m in VIC, it is a paradox in Figures 5 and 6 that Mosaic in SE and NE has the highest warm season evaporation and highest warm season change in total soil moisture. Comparing Mosaic's annual cycle amplitude in Figures 6b and 6d with Figure 7 provides the answer by revealing that Mosaic's subroot zone accounts for the bulk of Mosaic's annual storage range. Later, we show that Mosaic develops vigorous upward diffusion of water to its root zone from its subroot zone during the warm season. Mosaic's diffusion is not a dominant process in

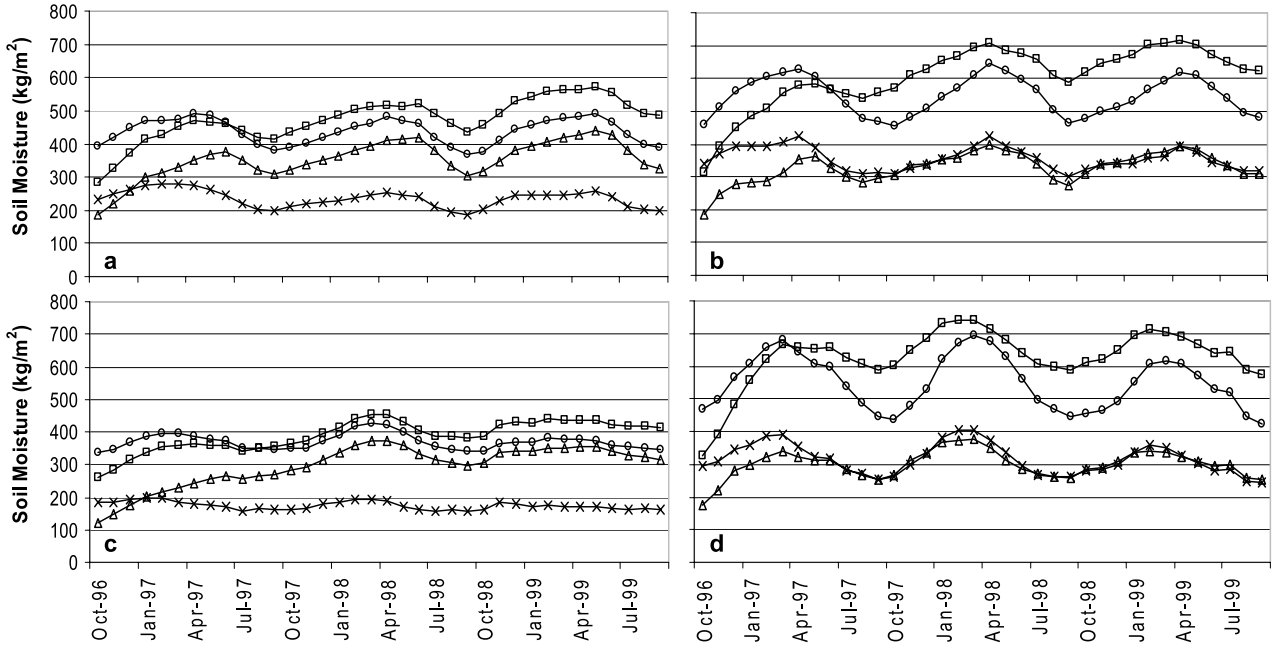

Figure 6. Time series of area-averaged monthly mean total column soil moisture (mm) in NLDAS for Noah (squares), VIC (triangles), Mosaic (circles), and SAC (crosses) for the CONUS quadrants of (a) NW, (b) NE, (c) SW, and (d) SE for October 1996 to September 1999.

the arid west, because root zone and subroot zone there are both typically dry.

[45] Figure 8 shows the May through September (nominal soil dry-down season) change in total soil moisture for 1999. Positive values denote soil drying. As an aside, the soil moistening (negative change: dark red) of Florida, south Texas, and Arizona/New Mexico is out of phase with the dry-down over the bulk of the CONUS, because their "wet season" is summer. Also, the Atlantic coastal states show moistening from two coastal hurricanes in September 1999. The hallmark of Figure 7 is vivid model differences in total soil moisture depletion over the southern Mississippi basin, which experienced significantly below normal precipitation during this period (not shown). Mosaic shows the largest depletion there and the largest evaporation (Figure 5d), followed by SAC and then Noah, with VIC showing the least depletion and lowest evaporation (Figure 5d). Together, Figures 5d, 6d, 7b, and 8 illustrate that over vegetated areas with a warm-season precipitation shortage, Mosaic taps notably more than the other LSMs from its deep (subroot) soil moisture to sustain evaporation, though too much so (SE in Figure 3b).

[46] Figure 8 shows large intermodel variation in the east-west gradient across CONUS of seasonal range in total soil moisture storage ($S_1 + S_2$). The study by *Schaake et al.* [2004] (SD-N) depicts this intermodel difference as a function of basin climatology, in terms of the *P*/PE ratio for the 12 CONUS RFCs. As one can infer from east-west inspection along latitude 37° N in Figure 8, SD-N finds that VIC's storage range is relatively invariant with respect to east-west gradient in *P*/PE between arid and moist regions, while SAC and Noah show more variability, and Mosaic the

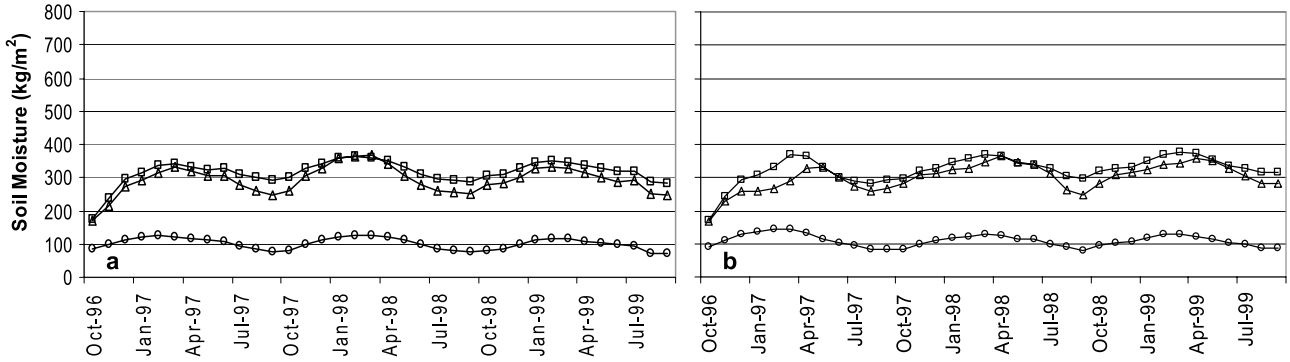

Figure 7. Time series of area-averaged monthly mean root zone soil moisture (mm) in NLDAS for Noah (squares), VIC (triangles), and Mosaic (circles) for the CONUS quadrants of (a) NE and (b) SE for October 1996 to September 1999.

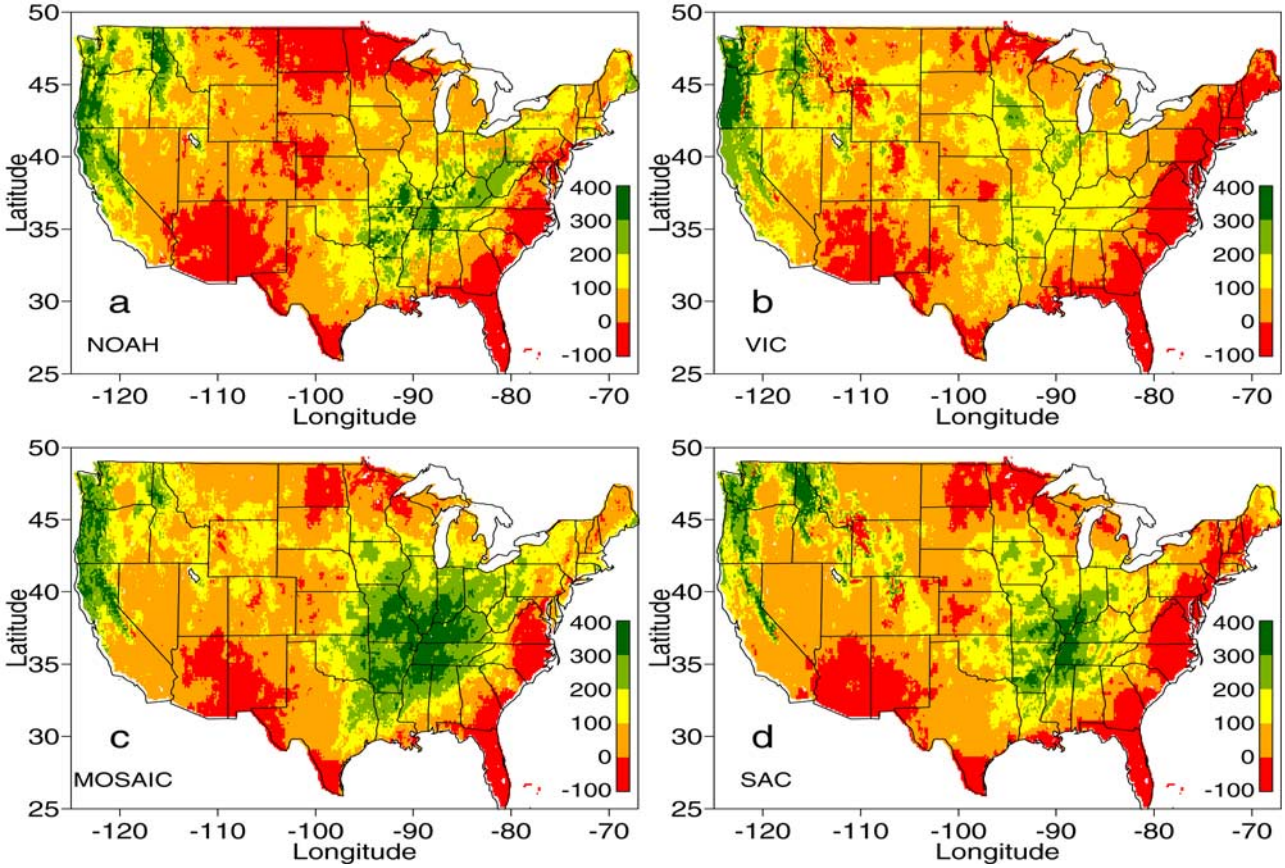

Figure 8. Warm season storage change (mm) of total column soil moisture, from difference of total column soil moisture (mm) of 30 April minus that of 30 September of 1999 at 23 UTC for (a) Noah, (b) VIC, (c) Mosaic, and (d) SAC. All colors (except dark red) are positive and denote a net drying during the period.

most variability. SD-N emphasizes that a given model's dynamic storage range in a given region is not reliably inferred from the model's maximum water holding capacity, but is rather the result of more complex interplay between a region's climatology and major facets of a model's physics. *Koster and Milly* [1997] show that a model's dynamic range of soil moisture is highly controlled by interaction between the model's runoff and evaporation formulations and the functional dependence of these formulations on the model's soil moisture state.

[47] We turn now to Figure 9, which depicts the annual cycle time series of every term in equation (1). Here storage S_1 is defined as the top two soil layers in Mosaic, Noah, and VIC, and the top storage reservoir in SAC and S_2 represents all remaining soil storage. The difference between the solid black line (total precipitation P) and dashed black line (liquid precipitation) is the snowfall. The red triangle is $P\text{-}dS_n/dt$ and the red triangle's departure below (above) the solid black line (P) equates to the monthly increase (decrease) in snowpack depth. In the snow season, the red triangle coincides with P if monthly snowfall is balanced by the sum of monthly snowmelt and sublimation, yielding zero net monthly change in snowpack.

[48] In Figure 9 we first examine region SE, which has negligible snow and rather small month-to-month changes in precipitation (in percent terms). Thus the annual cycle of the SE water budget is driven mainly by PE. The model differences in runoff response are vivid. Mosaic's dominance in soil moisture depletion in the warm season is mirrored by Mosaic having the largest soil recharge during November–February, leaving rather less precipitation available during cool months for runoff. SAC also has less cool season runoff than Noah or VIC, though not so much from high storage recharge as in Mosaic, but rather from SAC having the highest cool season evaporation. Noah and VIC have larger total runoff then Mosaic and SAC in most every month, with VIC having the notably largest runoff (mostly subsurface) throughout the fall, winter, and spring, as VIC requires less cool-season soil recharge to replenish its smaller summer depletion. In NE in Figure 9, the above tendencies in SE continue to hold in a broad sense, but other signatures arise from less precipitation in the cool season, greater monthly variability of precipitation in the warm season, and nonnegligible snowpack processes. Lower precipitation in the cool season compared to SE results in less cool season runoff in all four models, but in general VIC still produces the most monthly runoff, followed by Noah, then Mosaic and SAC. The lack of SAC runoff all year in NE is noteworthy, as is the high SAC evaporation in spring.

[49] Given that SAC in NLDAS takes its PE forcing from Noah PE output, the higher March–April evaporation in

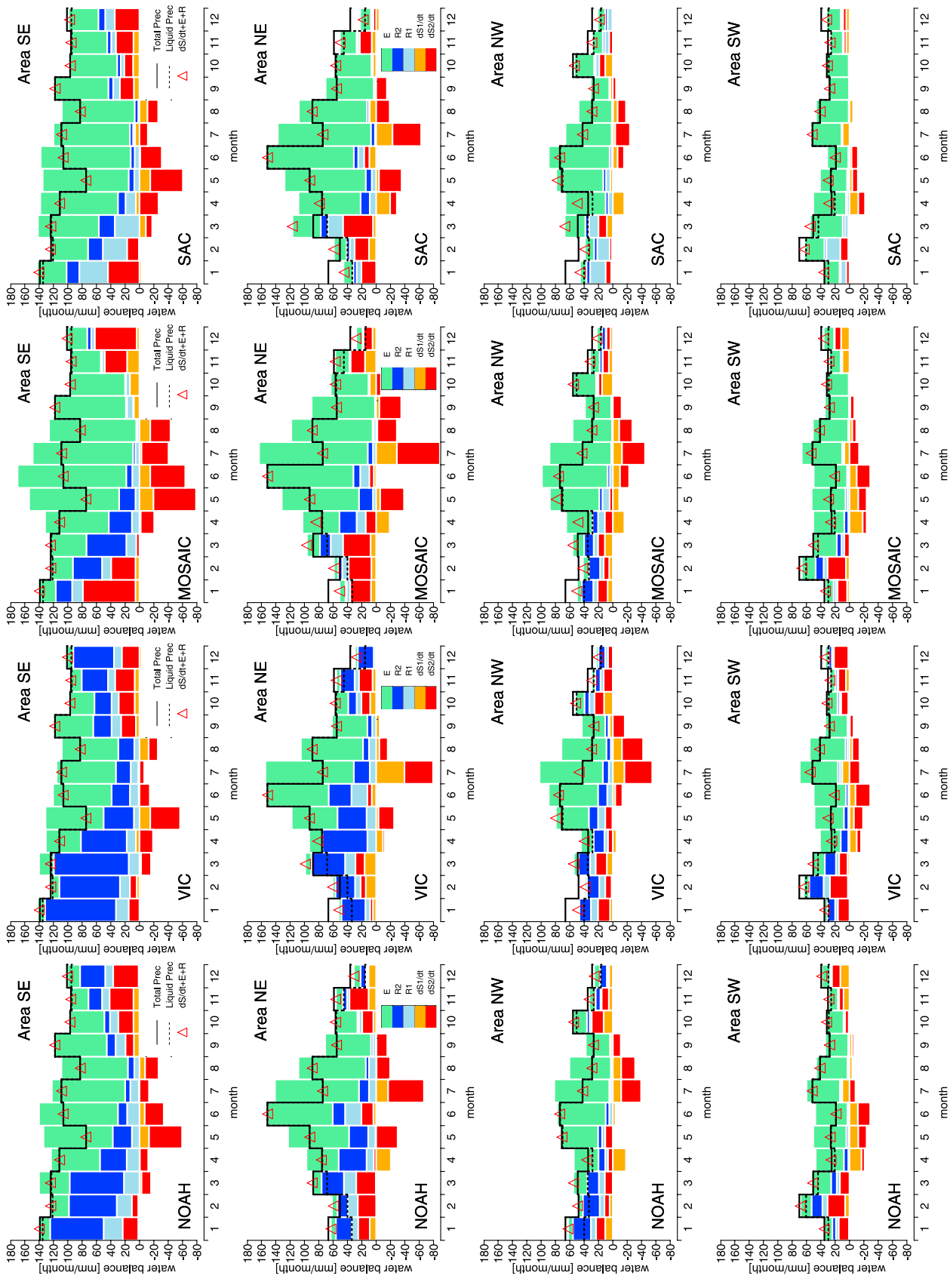
Area SE
Area NE
Area NW
Area SW
SAC
MOSAIC
VIC
NOAH
water balance [mm/month]
month
Total Prec
Liquid Prec
dS/dt+E+R
E
R2
R1
dS1/dt
dS2/dt

Figure 9.

SAC versus Noah in NE and NW is instructive, as vegetation greenness and hence transpiration are low then in both quadrants in Noah. Noah, Mosaic and VIC draw soil moisture for direct evaporation only from their first soil layer (0.1 m). This one layer is likely unable to sustain as high evaporation rates as SAC, which taps into both of its top two soil storages in response to PE demand. This structural difference in direct evaporation between SAC and the SVATS likely explains SAC's dominance in cool season evaporation (especially winter and spring), which contributes in turn to SAC's high mean annual evaporation. Thus, while Figures 2 and 3 show SAC and Mosaic having similar annual evaporation, the monthly water budget in Figures 5 and 9 (notably for NE and SE) shows SAC and Mosaic having rather different temporal character in evaporation, with SAC (Mosaic) having more in spring (summer).

[50] Moreover in Figure 9, the high precipitation in June in NE followed by a large drop of precipitation in July yields another vivid example of Mosaic's ability to draw upon its deep soil moisture. In the face of the steep drop in July precipitation, Mosaic is the only model in NE to yield July evaporation larger than June evaporation, and it does so via the largest July soil depletion. Following in August and September, Mosaic still sustains higher evaporation than the other models, despite its larger July storage depletion. Mosaic's layer 3 (subroot) storage change was dissected for July 1998 at one grid cell near 45°N latitude and 92°W longitude. For this month at this point, the Mosaic drainage (R_2) out the bottom of layer 3 ranged from 4 to 10 mm across its tiles there, while the upward diffusion of water to layer 2 ranged from 116 to 124 mm, confirming Mosaic's vigorous supply of subroot water to the root zone by diffusion. Though Noah also includes vertical diffusion of soil water, the magnitude is much larger in Mosaic.

[51] The contrast between models in Figure 9 is rather less in the western quadrants, wherein the warm-season water budget of the three SVAT models is quite similar, though SAC has less warm season evaporation and soil moisture depletion. The greater model similarity in SW and NW likely stems from the sparse vegetation (Figure 1a), whereby canopy conductance and root-zone processes are not dominant. In the cool season of NW and SW, VIC still tends to have the most runoff and is still dominated by subsurface runoff. Interestingly, SAC has virtually no subsurface runoff in NW and SW. One cool season contrast between models in NW (and NE) is the lower snowpack accumulation in Noah during winter. Comparison of the NW January water budget components of Noah with the other LSMs reveals model agreement in monthly snowfall amount (difference between solid and dashed lines), but Noah (unlike the accumulating snowpack in VIC, Mosaic and SAC) is melting and sublimating in the month about as much snow as it receives in snowfall, a topic revisited in section 3.4.

[52] We conclude this section with some discussion of the causes of the systematic biases noted in the models thus far. The high bias in Mosaic evaporation over CONUS-east is most likely a result of the vigorous diffusion of water from the subroot zone to the root zone. The high bias in SAC evaporation over CONUS-east is likely a consequence of SAC's use in NLDAS of (1) uncalibrated a-priori parameters (section 2.3) and (2) PE from Noah. As a counterexample, in NWS operations at the RFCs, SAC inputs a NOAA monthly climatology of PE, derived from evaporation-pan measurements. This climatological PE is then scaled during SAC runs by a monthly fractional coefficient. This coefficient is a key calibration parameter allowed to be moderately larger or smaller than 1.0, as determined from SAC calibration runs over a catchment. No such coefficient had been derived to date for use in SAC large-scale runs over the CONUS-wide domain, either with Noah PE or NOAA climatological PE, and hence a universal coefficient of 1.0 was used SAC control runs in NLDAS. Moreover, the Noah PE is known to be higher than NOAA PE climatology. These two factors contribute to high SAC evaporation (E) in NLDAS, but not always the highest E, and typically less than the E in Mosaic in the warm season over nonsparse vegetation. Hence the SAC results here are not outliers and they represent important pathfinder runs of SAC executed over a national domain in semi-distributed mode with uncalibrated parameters. Since the SAC control runs here in NLDAS, NWS/OHD has derived a CONUS-wide field of the PE coefficient from the vegetation greenness database cited in Figure 1a. SAC experiments of this field in NLDAS are imminent.

[53] We last address the unexpectedly low bias in VIC evaporation in the results here over CONUS-east. Two separate but related VIC modeling efforts have been conducted over the NLDAS grid and terrain heights; specifically, the 3-year retrospective runs executed here with 1-hour time steps, and the 50-year retrospective runs reported by *Maurer et al.* [2002], executed with 3-hour time steps and with different sources for the surface forcing. The 3-year VIC runs here use essentially the same parameters as the VIC runs of *Maurer et al.* [2002]. Yet two significant differences were hourly temporal disaggregation of the daily precipitation and subgrid spatial disaggregation within a grid box, both used in the VIC runs here but not in those of Maurer et al. Not having the advantage of hourly radar-anchored precipitation analyses, Maurer et al. used uniform distribution of the daily precipitation throughout the day and within each 3-hour forcing interval. *Maurer et al.* [2002, Figure 2] analyzed the impact of this uniform distribution versus nonuniform disaggregation to 3-hour time steps and the results showed that the differences for the subregion analyzed (Lower Mississippi basin) were modest. Nonetheless, subsequent comparisons between the retrospective runs of Maurer et al. and the 3-year retrospective runs of VIC here show that the combined and interactive impact of the three factors of temporal disaggregation, spatial disaggregation, and 1-hour versus 3-hour time steps can be significantly larger than suggested by the Lower Mississippi tests of the temporal

Figure 9. Monthly water budget in NLDAS for October 1997 to September 1998. In order, columns 1–4 are Noah, VIC, Mosaic, and SAC, and rows 1–4 are quadrants SE, NE, NW, and SW. Colors depict terms (mm/month) in equation (1): dS_1/dt (orange), dS_2/dt (red), R_1 (light blue), R_2 (dark blue), and E (green). Black solid line is total precipitation P (mm/month); black dashed line is liquid precipitation (mm/month). See text for definition of red triangles.

disaggregation alone. The differences (shown for a transect across the eastern and central United States at http://www.hydro.washington.edu/Lettenmaier/Models/VIC/VIChome.html) are evident in portions of the country with a high fraction of convective precipitation and full canopy cover (e.g., CONUS-east summer). More study of the differences, and development of parameter transformations to account for disaggregation and time step differences will be addressed in a future paper.

3.3. Regional Validation of Soil Moisture

[54] This section presents validation of NLDAS soil moisture over (1) Illinois from *Schaake et al.* [2004] (SD-N) and (2) Oklahoma from *Robock et al.* [2003] (RL-N). SD-N evaluated NLDAS soil moisture at 17 of 18 sites of the Illinois State Water Survey [*Hollinger and Isard*, 1994], which measures soil moisture at 11 levels down to 2 m. Figure 10 shows the resulting two-year scatterplot (and best-fit linear line) of model versus observed, state-wide average, total-column soil moisture over 2-m at bimonthly intervals. For VIC, two best-fit lines for northwest and southwest Illinois were required, because past VIC calibration yielded rather different soil moisture storage capacities in these regions. In Figure 10, a best-fit line having slope greater than one indicates a storage range greater than observed. Mosaic yields a storage range greater than the other models (as in sections 3.1–3.2) and about 50% more than observed. Noah and SAC agree with observations in both storage range and storage magnitude. VIC also shows good storage range over its two regions, but storage magnitude lower than observed. In addition to Illinois validation, SD-N validates NLDAS soil moisture storage range in the Arkansas-Red River basin (not shown). More broadly, SD-N intercompares the soil moisture storage capacity (explicit capacity) and storage range (''active'' capacity) of the four LSMs across the entire NLDAS domain, including mean statistical properties and spatial variation. The findings reveal significant LSM differences in soil moisture, as did section 3.2. Such differences challenge modelers using soil moisture from one LSM to initialize another.

[55] Similar contrasts in soil moisture between LSMs themselves and between observations and LSMs are found over Oklahoma by *Robock et al.* [2003] (RL-N), which included close scrutiny at individual stations. In-situ observations of soil moisture have been installed and calibrated at 72 Oklahoma Mesonet stations by the Oklahoma Climatological Survey. Figure 9 of RL-N (not shown) depicts a 21-month time series during 1998–1999 of observed and NLDAS-simulated daily mean, 0–40 cm total soil moisture averaged over all 72 Mesonet stations. The time series show substantial differences in soil moisture magnitude among the LSMs and between the LSMs and observations, with VIC showing the best agreement with the observations. Nonetheless, there is rather good agreement among the models and between models and observations in the soil moisture changes in time, a theme cited earlier here with Figure 6. RN-L also performs and evaluates important sensitivity tests in the three SVAT models at many of the OU Mesonet soil moisture measuring stations, wherein model soil type is set to match the station-reported soil type and the assigned soil parameters are unified across the models. This test improved a model's performance if the parameters were not incompatible with previous model calibration. Since Noah is not substantially calibrated, the matching with local soil type and use of unified soil parameters did improve Noah's soil moisture performance at the stations. In VIC, which has been regionally calibrated to streamflow over several large basins, including the Arkansas-Red river basin [*Abdulla et al.*, 1996], the use here over Oklahoma of local station-matching soil types and unified soil parameters had less consistent impact on improving model agreement with the soil moisture measurements, showing more variance of positive or negative impact depending on station. Thus soil parameter changes in an LSM to match local site characteristics may degrade LSM performance at those sites, especially if the changes disturb an LSM's calibration legacy.

3.4. Validation of Snow Cover and Snowpack Content

[56] This section presents results from *Pan et al.* [2003] (PS-N) and *Sheffield et al.* [2003] (SP-N), which perform large-scale assessment of NLDAS snowpack water equivalent (SWE) and snow cover extent (SCE), respectively. First, we summarize how the four LSMs treat snowpack. SAC simulates only the snowpack water balance, via the SNOW17 model [*Anderson*, 1973], which includes snowfall and snowmelt but not sublimation. Snowmelt is determined empirically via an index-method based on maximum-minimum daily air temperature. VIC, Noah and Mosaic also simulate the snowpack water balance, including sublimation as well as snowfall and snowmelt, plus the snowpack energy balance (net radiation, sensible, latent and subsurface heat fluxes, phase-change heat sources/sinks). The treatments for processes such as snow cover fraction, snow albedo and retention/refreezing of snowmelt differ among the models. SAC, VIC, and Noah explicitly account for retention of liquid water (snowmelt or rainfall) in the snowpack, but only VIC and Noah allow refreezing. Mosaic and VIC carry explicit subgrid vegetation tiles, but only VIC carries added subgrid tiles for elevation, known as ''elevation banding'', which is cited in PILPS studies as a key factor in VIC's good snowpack simulations [*Bowling et al.*, 2003; *Boone et al.*, 2004]. For each subgrid tile, Mosaic and VIC carry separate water-energy balances and separate soil, vegetation and snow states. More on the model snow physics is provided by SP-N, and by *Wigmosta et al.* [1994], *Koster and Suarez* [1996], *Koren et al.* [1999], and *Anderson* [1973] for VIC, Mosaic, Noah, and SAC, respectively.

[57] There is no distinction between rainfall and snowfall in NLDAS precipitation forcing. This requires criteria to infer snowfall. The input precipitation at each model time step was assumed to be all rainfall for surface air temperature >0°C and all snowfall otherwise. This criterion does not guarantee identical snowfall in the models owing to different model time steps and VIC's elevation banding. Noah and Mosaic interpolate hourly air temperature to their 15-min time steps, thus allowing rainfall and snowfall inside one hour, unlike the hourly steps of VIC and SAC. VIC adjusts the hourly air temperature to the elevations of its subgrid elevation bands, thus VIC

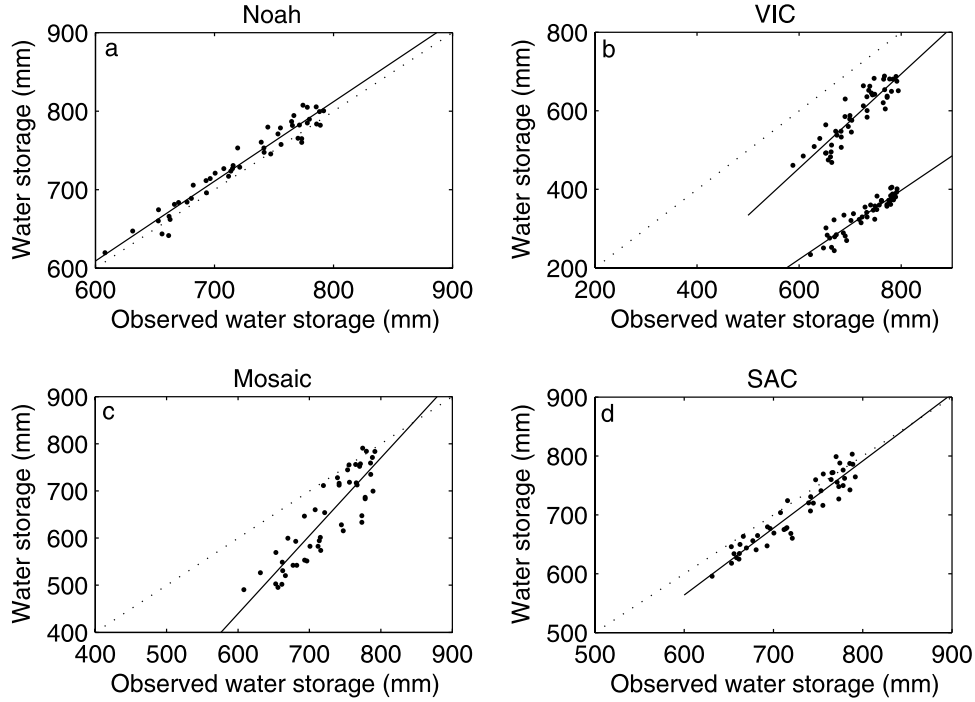

Figure 10. Comparison of NLDAS versus observed bimonthly total soil moisture (mm) in top 2 m, averaged over 17 sites throughout Illinois for October 1997 to September 1999 for (a) Noah, (b) VIC, (c) Mosaic, and (d) SAC. Note different x axis and y axis ranges. See text for discussion of two sets in Figure 10b.

allows both rainfall and snowfall inside a given grid cell for a given time step.

[58] PS-N validated NLDAS SWE simulations against NRCS SNOTEL stations west of 104°W. SNOTEL sites measure SWE, air temperature and precipitation every 15 min. The majority of SNOTEL elevations are above 1000 m, with mean elevation near 2500 m (see PS-N for references). Since NLDAS terrain resolution is 1/8°, comparing model SWE with point-wise SNOTEL is a challenge. Hence PS-N omitted use of SNOTEL sites whose elevation differs from that of the nearest NLDAS grid point by more than 50 m. This retains 110 SNOTEL sites.

[59] Figure 11 shows model versus observed mean-annual maximum SWE for the LSM control runs (and two VIC tests described later) at the SNOTEL sites by subset in four mountain ranges: 3 in Sierra Nevada (squares), 17 in Cascades (diamonds), 29 in southern Rockies (circles), 61 in northern Rockies (triangles). All the LSMs substantially underestimate maximum SWE in all four regions. Noah has the largest low bias and the lowest correlation. Mosaic also shows a rather low correlation. SAC and VIC have notably better bias and substantially higher correlation. The model with elevation tiling (VIC) yields the highest correlation, yet the simplest model (SAC) without elevation tiling or energy balance treatment is closely competitive.

[60] PS-N found the NLDAS precipitation forcing to be substantially low when compared to observed precipitation at all 110 SNOTEL sites. Sites with the highest observed precipitation are where NLDAS precipitation has the largest low bias. PS-N determined the linear regression between SNOTEL and NLDAS precipitation at the 110 sites to be $P_{\mathrm{SNOTEL}} = 2.1693\ P_{\mathrm{NLDAS}}$, with an R^2 value of 0.64, revealing a factor-of-2 underestimation by NLDAS precipitation at the sites on average. The low precipitation bias is consistent with that anticipated from the low bias in annual streamflow in the LSMs in the Northwest in section 3.1 (Figure 4).

[61] PS-N executed two tests in VIC with two methods of bias-adjusted precipitation. Both tests executed VIC in the reduced-grid mode (section 2.1) at grid cells nearest the SNOTEL sites. Test 1 used NLDAS precipitation forcing scaled upward by a separate regional factor for the four mountain ranges, based on regional regression fit of the NLDAS and SNOTEL mean annual precipitation. In test 1 (Figure 11e), the model SWE bias is dramatically less, with the scatter rather evenly balanced about the 1:1 line. Yet the scatter remains substantial and the correlation is only modestly improved, as the regional scaling does not eliminate site-specific bias. VIC in test 2 (Figure 11f) was forced with site-specific adjusted NLDAS precipitation, scaled to match the observed annual total precipitation at each site. Test 2 yields a much smaller bias, substantially reduced scatter and increased correlation of $R^2 = 0.82$. PS-N also evaluated NLDAS air temperature bias at SNOTEL sites, finding it small in the cool season (exceptions at some stations), contributing much less to model SWE bias than precipitation. The high bias in NLDAS solar insolation over snow (section 2.2) also contributes to a low bias in SWE in the three SVAT models. The impact of this insolation bias on SWE has not been quantified yet.

[62] SP-N validated NLDAS simulations of areal fraction of snow cover extent (SCE) against the NESDIS operational, daily, 23-km, Northern Hemisphere snow cover product known as the Interactive Multisensor Snow (IMS) [*Ramsay*, 1998], viewable at http://www.ssd.noaa.gov/PS/SNOW. NESDIS analysts produce the IMS using an interactive

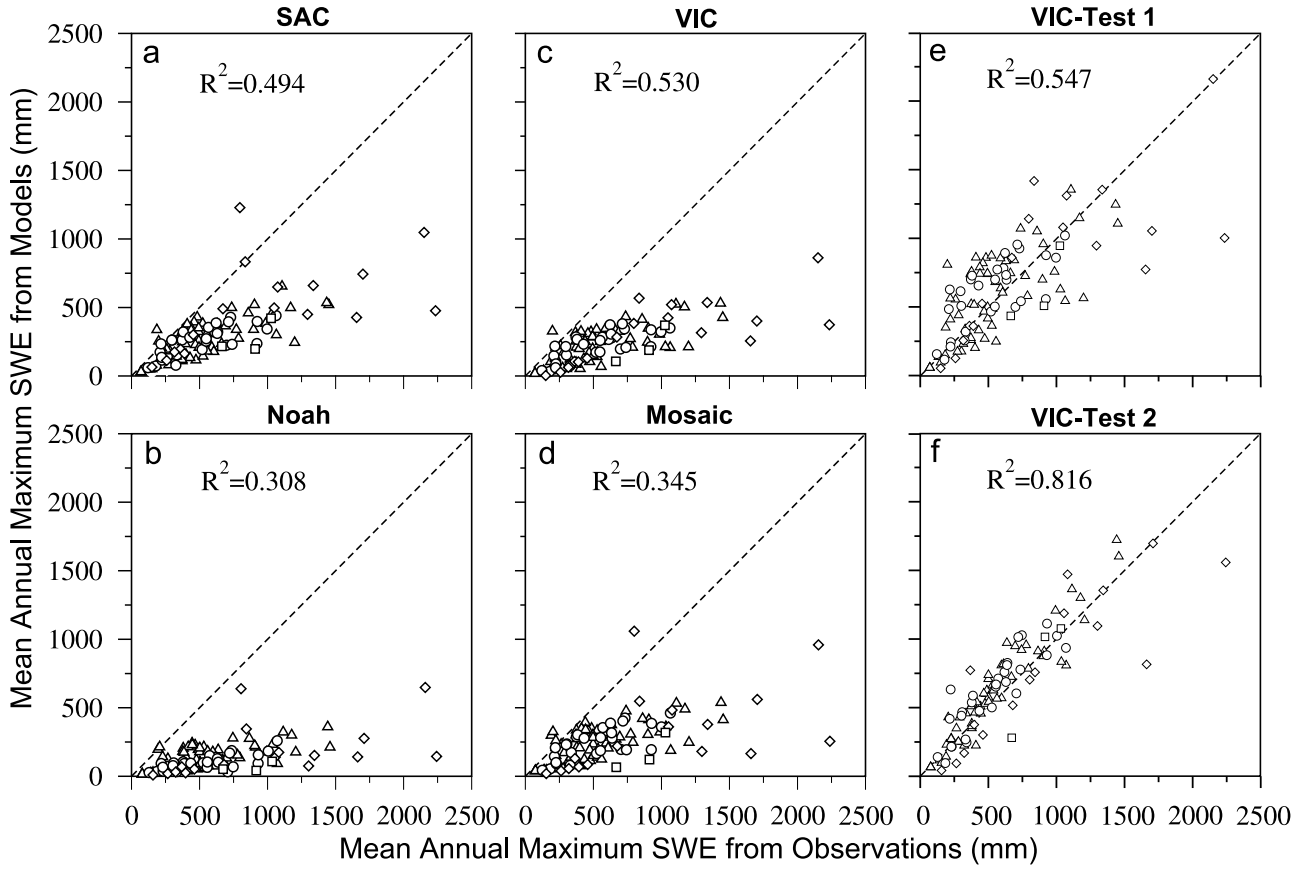

Figure 11. Comparison of mean annual maximum snow water-equivalent (SWE) during October 1996 to September 1999 between observations (*x* axis) and model simulations (*y* axis) at 110 SNOTEL sites for the control runs of (a) SAC, (b) Noah, (c) VIC, and (d) Mosaic, and two VIC tests runs forced with (e) regionally adjusted precipitation and (f) locally adjusted precipitation.

workstation to assess snow cover related visible, infrared, and microwave satellite products, as well as in situ snow depth observations. Details of mapping the 23-km IMS field to the NLDAS 1/8° grid and mapping NLDAS snow cover fraction (0–1 range) to the IMS binary snow cover state (yes/no) are given in SP-N.

[63] All four LSMs diagnose SCE (0–1 fraction) as an empirical function of the model simulated SWE, but the treatments differ significantly, namely, the critical SWE needed to reach SCE = 1 and the form of the function relating SCE to SWE. SP-N provides details, but in broad terms SAC and Noah require relatively deep snowpack for high SCE values, while Mosaic requires notably less, and VIC very little. For a nondeep SWE value, VIC will yield the highest snow cover, followed by Mosaic, then Noah, and finally SAC. For illustration, assuming a nonforest vegetation type, a SWE value of 4 mm yields SCE = 1.0, 0.67, and 0.24 in VIC, Mosaic, and Noah, respectively. Correspondingly (shown later), VIC, Mosaic, and Noah in order yield the highest, intermediate, and lowest albedo over snow. (SAC requires no albedo.) The high SCE fractions in VIC arise in part because VIC assigns SCE = 1 over any tile with any nonzero SWE.

[64] SP-N carried out validation of SCE separately over the 8 of 12 CONUS RFCs that exhibit substantial winter snow cover. In general, all models simulate reasonably well the regional-scale spatial and seasonal dynamics of snow cover. Yet systematic biases exist, with (on average over 8 RFCs) underestimation of SCE by Noah (−22.5%) and MOSAIC (−19.8%) and overestimation of SCE by VIC (22.3%), with SAC being essentially unbiased. The level of bias over individual RFC regions varies (see Figure 4 of SP-N). The more mountainous RFC regions (Northwest, Colorado, California-Nevada) show the largest model differences with IMS observations and between models. Here VIC further overestimates SCE, while Noah further underestimates SCE and manifests an early bias in spring snowmelt. VIC's high SCE bias in the west is surprising at first, since all four LSMs had notably low bias in SWE at SNOTEL sites in Figure 11, owing to the low bias in NLDAS precipitation. However, the number of pixels at SNOTEL sites in any RFC domain is relatively small, so the effect on the regional mean is difficult to judge. Yet we surmise that VIC's high SCE bias stems from VIC assigning SCE = 1 at any subgrid tile with any nonzero SWE, however small.

[65] The low bias in Noah SCE appears to result from not only the high SWE threshold required in Noah for high SCE, but also the low snow albedo in Noah and its positive feedback effect on the energetics of snowmelt. For one RFC domain, Figure 12 shows the time series of monthly domain-mean (1) snowmelt, (2) snow sublimation and (3) albedo from the four models (only snowmelt for SAC, as SAC excludes sublimation and albedo). There are large differences in snow albedo among the models, with Noah

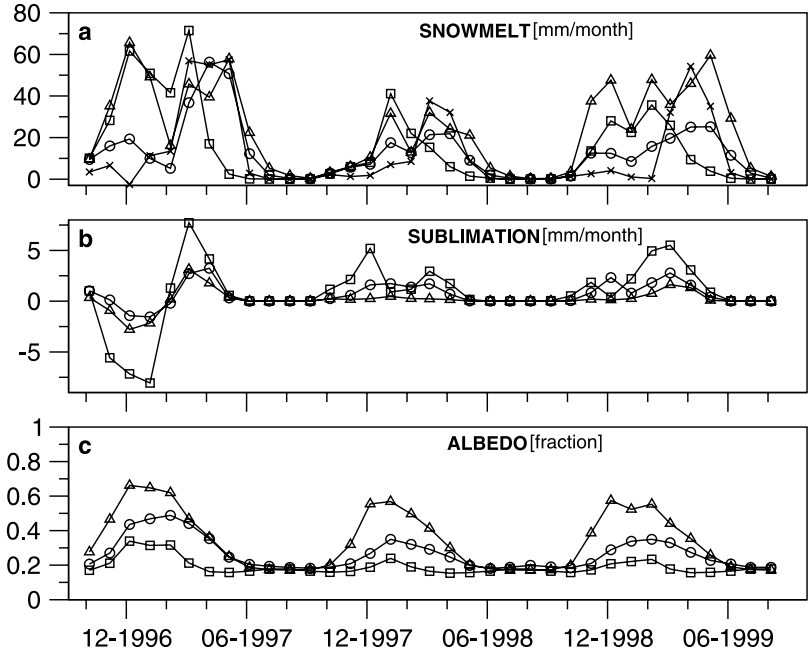

Figure 12. For the Northwest RFC domain, time series of monthly domain-mean (a) snowmelt (mm/month), (b) snow sublimation (mm/month), and (c) albedo in NLDAS for Noah (squares), VIC (triangles), Mosaic (circles), and SAC (crosses) for the time period October 1996 to September 1999. (Note different *y* axis range in top two panels.)

yielding the lowest (0.2–0.3), Mosaic being intermediate (0.3–0.5), and VIC yielding the highest (0.5–0.65). Not surprisingly then in Figure 12, Noah yields notably higher snowmelt and sublimation in early and midwinter, while VIC yields very little sublimation, with Mosaic in between. The larger midwinter snowpack sinks in Noah are consistent with Noah having the greatest low bias in annual maximum SWE in Figure 11. By spring, Noah's snowmelt and sublimation reduce to almost zero, as much of Noah's snowpack has already melted or sublimated earlier in the winter, while VIC and SAC have the largest spring snowmelt volumes, in part because they had the smallest sublimation sinks during early and midwinter. SAC and Mosaic tend to have higher melt in the spring than the winter months while VIC melts at a more quasi-steady rate throughout the winter and spring.

[66] A low bias in snow albedo is vulnerable to positive feedback problems in the surface energy balance. A low bias in albedo contributes to a high bias in net solar insolation, which melts more snow and reduces the snow cover, yielding still lower albedo and so on. The feedback is amplified by the high bias here in the incoming solar insolation over snow. High albedo and high snow cover, such as in VIC, is also vulnerable to positive feedback in the opposite direction, but the high albedo in VIC likely acted to offset the high bias in NLDAS incoming insolation over snow. A counterpart to such feedback risk is the simplicity of snow physics in the SAC/SNOW17 model. Its snowpack predictions perform rather well when assessed at the large regional scales of NLDAS. One reason is SAC's simple temperature-index approach to snowmelt, which avoids the feedback loops that can plague energy balance models over snow. The recent PILPS high latitude modeling experiments [*Bowling et al.*, 2003] found large differences in snow ablation and snowmelt among 21 LSMs and also concluded that differences in model parameterizations of albedo and SCE have large effects on energy available to the snowpack.

4. Validation of NLDAS Surface Energy Fluxes and LST

4.1. In Situ Validation of Surface Energy Fluxes Over the Southern Great Plains (SGP)

[67] This section presents key results from the surface energy-validation portion of the *Robock et al.* [2003] study (RL-N), which validates energy fluxes during January 1998 to September 1999 using the 24 extended facility (EF) flux stations of the ARM/CART network in Oklahoma and Kansas. The results include the three SVATS (Noah, Mosaic, VIC) but not SAC, as SAC omits the physics of surface energy balance. Multistation spatial averaging and hourly temporal averaging are used to reduce the influence of scale differences between NLDAS grid cells ($\sim$12 km) and pointwise flux stations. Radiation fluxes were averaged over the 22 of 24 EF stations using Solar and Infrared Radiation Station instruments (SIRS). Heat fluxes were averaged over the 14 of 24 stations using Energy Balance Bowen Ratio systems (EBBR). Though the energy budget is not exactly closed in the averaging, discrepancy is less then 20 W m^{-2} in most months.

[68] The ARM network spans central and northern Oklahoma and southern Kansas (see map in RL-N). It is instructive to consider, a priori, what one would expect to uncover in the model surface energy budgets in this region based on the annual water budget validation in section 3.1. This region is a transition zone between quadrants SE and SW in Figure 3. Yet over the bulk of this region, Figures 1a and 1b show that warm-season green vegetation fraction is not sparse, ranging

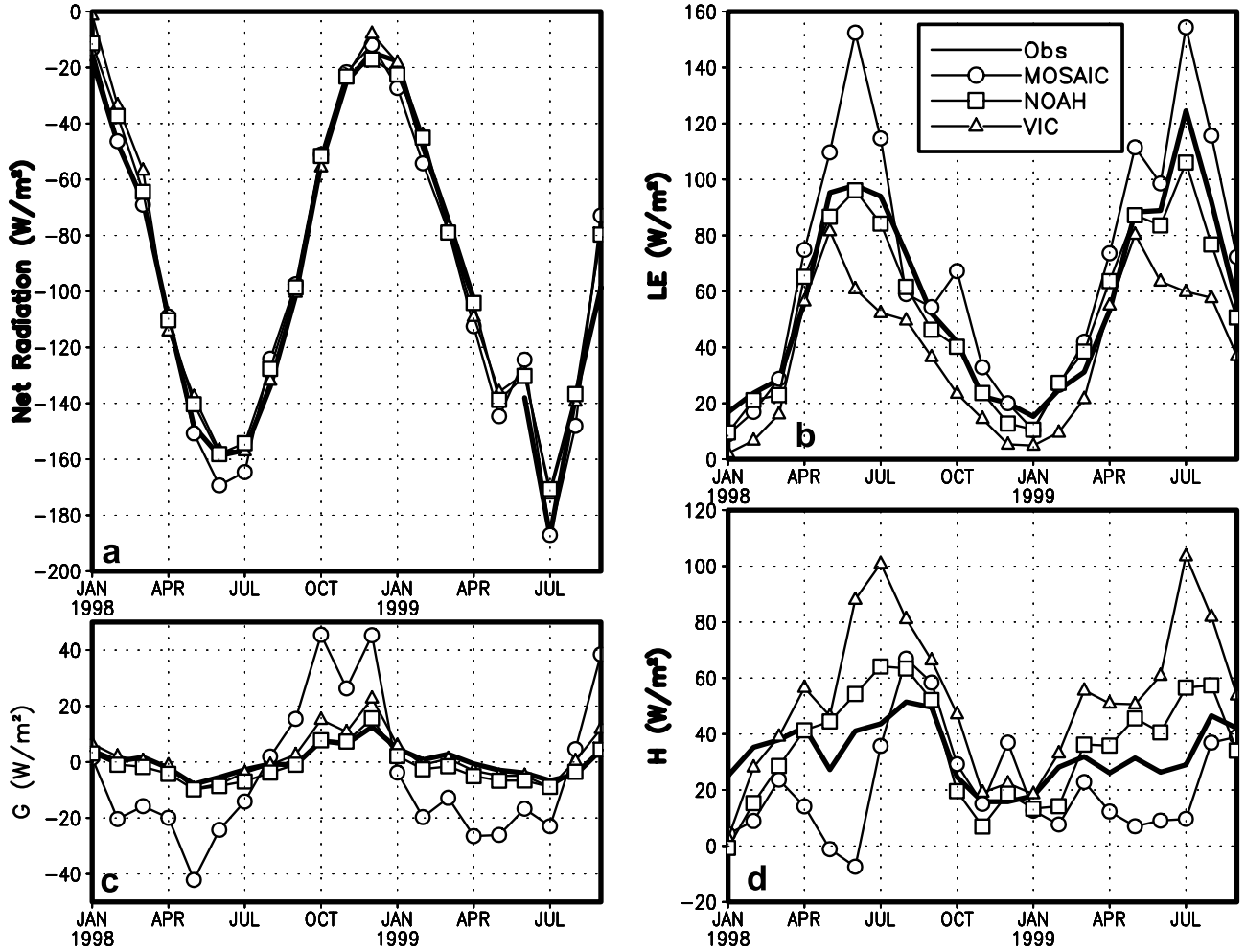

Figure 13. Time series of monthly mean surface energy fluxes (W m^{-2}) of (a) net radiation *R*, (b) latent heat LE, (c) ground heat *G*, and (d) sensible heat *H* averaged over the ARM/CART sites during January 1998 to September 1999 from observations (bold line, no symbols) and control runs for Noah (squares), VIC (triangles), and Mosaic (circles). The *y* axis range varies among panels. Positive flux is heat sink to surface, except for *G*.

from 0.4–0.8, and the mean annual precipitation exceeds 800 mm during the study period. Hence inspection of Figure 1 leads one to expect the mean surface water and energy budget of the region to be more similar in nature to that of SE then SW. Therefore, given that warm-season evaporation dominates annual evaporation, the following expectation is warranted: latent heat fluxes of the three SVATS over the ARM network are likely to reflect the model evaporation biases established over the SE quadrant in Figure 3b, namely, that Mosaic, Noah, and VIC will manifest latent heat flux averages that are substantially higher than observed, close to observed but somewhat low, and substantially less than observed, respectively. The validations below in Figures 13 and 14 confirm this expectation.

[69] Figure 13 gives a 21-month time series of monthly mean observed versus modeled surface energy fluxes of the NLDAS control runs, including net radiation (*R*), latent heat flux (LE), sensible heat flux (*H*), and ground heat flux (*G*). Figure 14 shows corresponding monthly mean diurnal cycles for the two months of July and April 1999. Columns 1 and 2 of Figure 14 depict only control-run results, while columns 3 and 4 additionally show experiments described later. There is rather good agreement between observed and simulated R in all the models. Figure 14 shows some small model phase errors in R in Noah and VIC, which are further diagnosed in RL-N.

[70] Of more interest are the several situations of substantial bias in simulated LE, *H*, and *G*, examined first in Figure 13. As correctly anticipated above, in spring and summer, Mosaic has a substantial high bias in LE and, correspondingly, a substantial low bias in *H*. VIC has a substantial low bias in LE and high bias in *H* throughout most of the year (except spring), while Noah shows much smaller bias in LE (slightly low in warm season) and *H* (modestly high in warm season). The LE results here agree with the sign and relative magnitude of the model evaporation biases inferred in section 3.1. The counterpart to evaporation bias in section 3.1 was runoff bias of opposite sign. Here the analog counterpart to LE bias is *H* bias of opposite sign. The warm season LE and *H* biases of Mosaic, VIC, and Noah are highlighted further in the July 1999 midday biases in the diurnal cycles of Figure 14, again showing Mosaic with a significant positive bias in daytime LE and substantial negative bias in daytime *H*, thus very low Bowen ratio (BR =

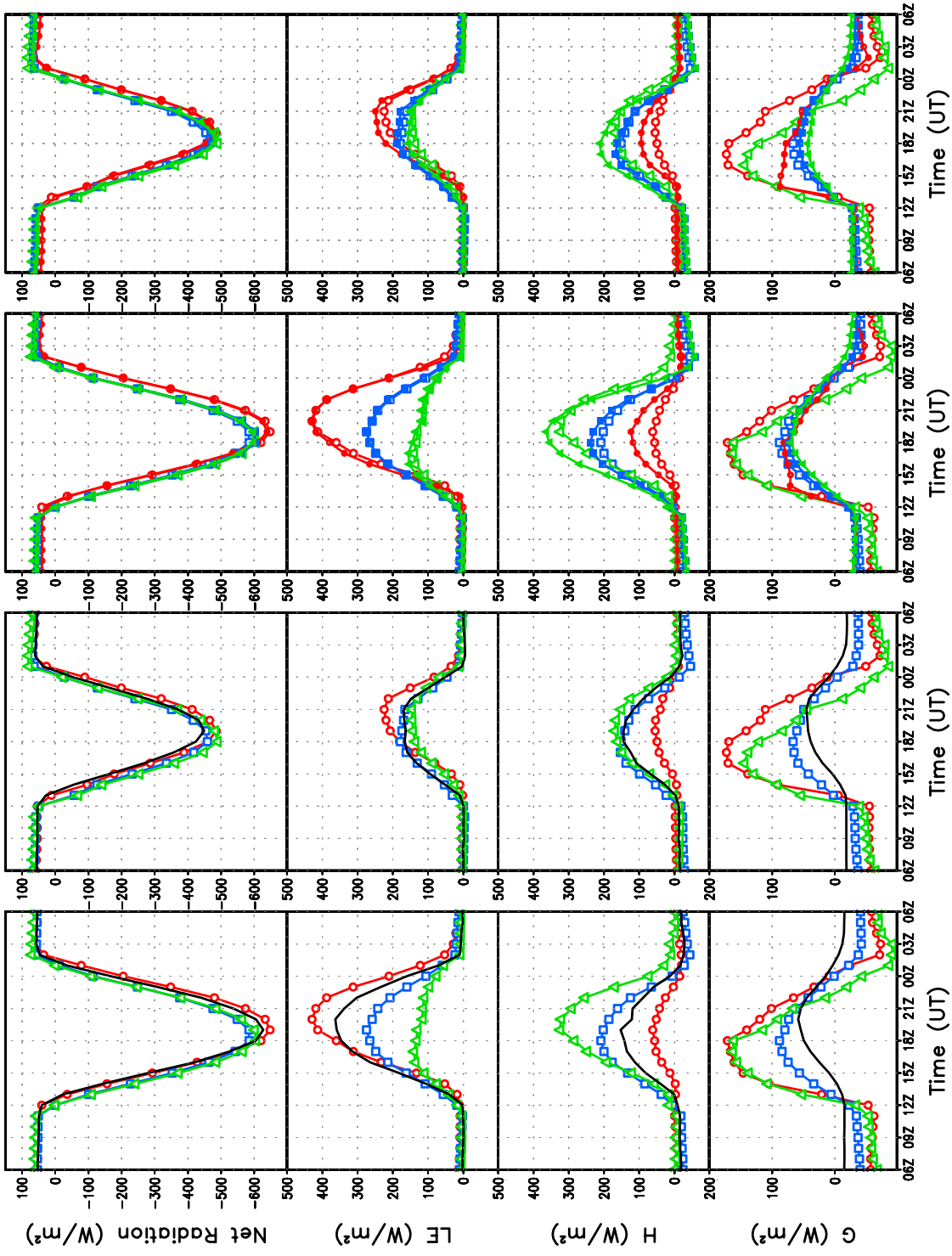
JUL99
APR99
JUL99
APR99
Net Radiation (W/m²)
LE (W/m²)
H (W/m²)
G (W/m²)
Time (UT)
062 092 122 152 182 212 002 032 062

Figure 14.

H/LE, not shown), with VIC showing the opposite, and Noah showing bias similar to VIC in sign but notably smaller in magnitude. In April, Mosaic retains high LE bias and low *H* bias, while VIC and Noah show little bias in LE or *H*.

[71] Figures 13 and 14 exhibit serious errors in ground heat flux *G* in VIC (diurnal cycle), and especially in Mosaic (both diurnal and annual cycles). Noah shows comparatively little error in G, with virtually no bias in monthly mean (Figure 13) and rather modest high bias in diurnal daytime G from phase error. Mosaic has a large bias in monthly mean *G* during most months (Figure 13) and most hours (Figure 14), though the sign changes depending on the time in the diurnal/annual cycle. VIC has large hourly biases in *G* for most hours (Figure 14), but small monthly mean biases (Figure 13). VIC's daytime and nighttime biases in *G* are rather symmetric and opposite in sign, so they nearly cancel on a daily or monthly mean basis. Both Mosaic and VIC show large daytime diurnal high bias and phase error in G in both July and April of Figure 14, with daytime maximums 2–3 times larger than observed. Finally, the simultaneous and very high daytime biases in *G* and LE in Mosaic during April–July conspire to yield dramatically low sensible heat flux (*H*) during these months. Mosaic monthly mean *H* values during May–June are less than the annual winter minimums in observed *H*.

[72] With NLDAS infrastructure now in place (section 2), increasing attention is being given to model experiments. The large biases in *G* in Mosaic and VIC prompted such experiments, described in RL-N and presented here briefly. The VIC and Noah models use a surface energy balance approach for their surface radiative temperature, calculated in VIC for a thin but nonvanishing skin layer that has nontrivial heat storage, while Noah assumes an infinitesimally thin skin layer with negligible (zero) heat storage. As a test, VIC was re-executed by imposing zero heat storage in its surface energy balance treatment. The formulation G in Mosaic [*Koster and Suarez*, 1996] is based on the standard force-restore or "slab" treatment. It assumes that the "aggregate" surface/canopy medium of vegetation and near-surface soil has nonnegligible heat-storage capacity, specified by a heat capacity parameter (C_H) that strongly impacts G. In NLDAS, the C_H value in Mosaic's control run (175,000 J m^{-2} K^{-1}) was one calibrated in an earlier, independent temperature data assimilation system [*Radakovich et al.*, 2001], and not the lower traditional C_H value (70,000 J m^{-2} K^{-1}) specified by *Koster and Suarez* [1996] and used in several Mosaic PILPS experiments. To gage the impact of C_H, a Mosaic test was executed in NLDAS using the traditional lower value. Columns 3–4 of Figure 14 give the results of these Mosaic and VIC tests. (Aside: the Noah test in Figure 14 is presented later in section 4.2 and does not involve a change to surface heat capacity.) In both Mosaic and VIC, model ground heat flux was dramatically improved in the tests, both in July and April, becoming competitive with that in Noah (though Mosaic manifests an unusual anomaly in *G* during the early morning). This improvement in simulated G holds throughout the year in both models, as shown in RL-N.

[73] However, the improvement in *G* in the Mosaic and VIC tests provided no improvement in the large LE biases in Mosaic or VIC in Figure 14. Rather, the increase in daytime available energy (R-G) gained by reducing daytime high bias in G acted only to increase the sensible heat flux *H*, in both models, which helped the low *H* bias in Mosaic and worsened the high *H* bias in VIC. The nonresponsiveness of LE and the high response in H strongly suggests that the canopy resistance is substantially higher than the aerodynamic resistance in both models in this vegetated region in the warm season. Sensitivity tests of canopy resistance will be a focus in all three SVAT models in NLDAS follow-on studies. As a start, Mosaic tests (not shown) have been run in which the fixed thicknesses of Mosaic's soil layers (with fixed 40-cm root zone) in Mosaic's control run was replaced with the traditional Mosaic approach of letting soil layer thickness and root depth vary tile by tile according to the tile's vegetation type. The impact on the warm season latent heat flux of Mosaic was significant, but the improvement with respect to ARM flux observations was mixed, being either negative or positive, depending on which warm season month was examined. Thus further evaluation is underway.

4.2. In Situ Validation of Land Surface Skin Temperature Over the SGP

[74] A chief goal of NLDAS is assimilation of satellite data to improve soil moisture, and in turn, surface fluxes. One keen interest is the assimilation of satellite-derived LST. Positive impact from LST assimilation will be greatly enhanced if errors in modeled LST stem primarily from errors in the background model's Bowen ratio that arise from errors in model soil moisture states. Prospects for success are much lower if LST errors arise from Bowen ratio errors caused not by soil moisture, but by errors in the model's treatment of (1) vegetation cover and nonsoil moisture attributes of its canopy conductance, (2) the surface air layer and its aerodynamic conductance, (3) albedo and net solar insolation, or (4) ground heat flux, *G*, and its impact on the available energy ($R - G$). Section 4.1 uncovered substantial errors in *G*. This section uncovers significant impact on model LST from intermodel differences in aerodynamic conductance.

[75] At the top of the hour in NLDAS, all three LSMs output an instantaneous, grid-cell mean, radiometric surface temperature, referred to here as the land surface skin temperature, LST. To obtain LST, each model applies the Stephan-Boltzmann Law given by $L = \varepsilon\sigma(\mathrm{LST})^4$, in which ε is the surface emissivity (=1 in all three LSMs), σ is the Stephan-Boltzmann constant, LST is the skin temperature (K), and *L* is the upwelling longwave radiation (W m^{-2}). In Noah, which is nontiled, the surface energy budget is solved once for each grid cell to obtain LST and then *L*. In Mosaic

Figure 14. Monthly mean diurnal cycle of surface energy fluxes (W m^{-2}) of net radiation *R* (row 1), latent heat LE (row 2), sensible heat *H* (row 3) and ground heat *G* (row 4) for July (columns 1 and 3) and April (columns 2 and 4) of 1999. Columns 1 and 2: observed (bold black line, no symbols) and control runs for Noah (open squares, blue), VIC (open triangles, green), and Mosaic (open circles, red). Columns 3 and 4: same control runs as columns 1 and 2, plus one test each for Noah (solid squares, blue), VIC (solid triangles, green), and Mosaic (solid circles, red). See text for test features. The *y* axis range varies between rows. Positive flux is surface heat sink.

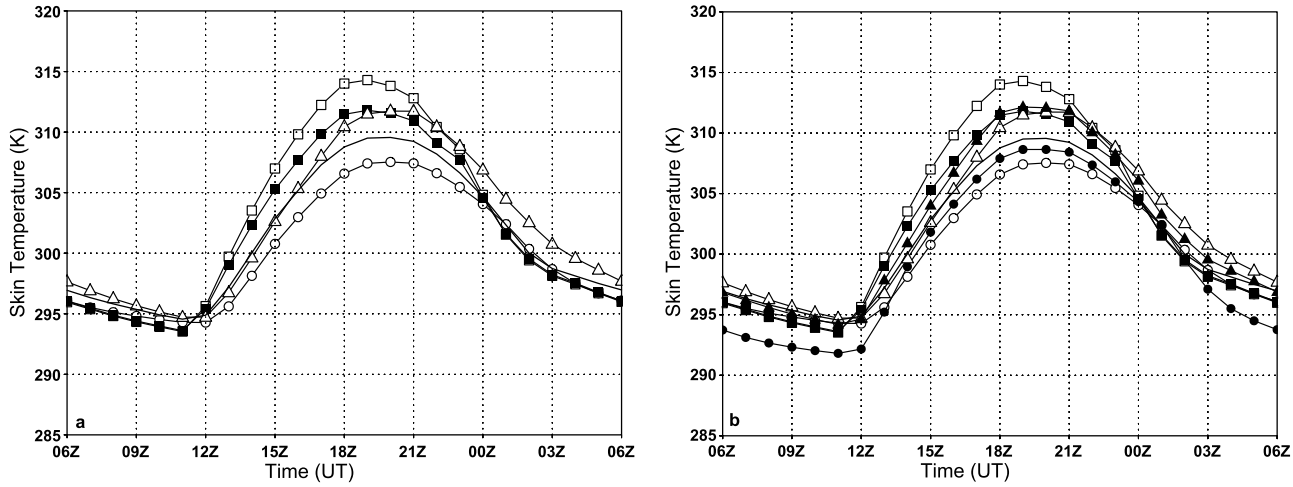

Figure 15. Monthly mean diurnal cycle of LST (K) averaged over all ARM/CART SIRS sites for July 1999 from observations (solid line, no symbols), control runs (open symbols) and test runs (solid symbols) for Noah (squares), VIC (triangles), and Mosaic (circles). (a) All three control runs plus Noah test of aerodynamic conductance. (b) As in Figure 15a, plus Mosaic and VIC tests of lower surface heat capacity.

and VIC, LST and L are obtained for each tile from a tile-specific energy budget, then the tile-weighted mean L over the grid cell is obtained, from which the grid-cell mean LST is derived from $L = \varepsilon\sigma(\mathrm{LST})^4$. Finally, at each SIRS station, the observed L is time averaged to the top of the hour, and LST is obtained also from the latter relation using $\varepsilon = 1$.

[76] Figure 15a shows the multistation average, monthly mean diurnal cycle of SIRS-observed and LSM co-located LST of the control runs, averaged over the SIRS sites, for July 1999. (The Noah test in Figure 15a is described later.) Mosaic has a midday cool bias in both months, as expected, given its high bias in LE and G and low bias in H in Figure 14 (columns 1 and 2). VIC and Noah have midday warm biases in July, also as expected, given their low LE and high H bias in July. While the sign of the models' midday LST bias in Figure 15a is as expected, the comparative magnitude of the bias between the models is perplexing at first, given the LE and H fluxes in Figure 14. Specifically, the VIC midday (19–20 UTC) warm bias in July (about +2 K) is about half as large as Noah (about +4 K), despite VIC's Bowen ratio (BR = 2.91) at this time being much higher than Noah's (BR = 0.70) and the observed (BR = 0.38). Thus VIC does not yield the largest midday warm bias, despite having by far the largest high bias in Bowen ratio.

[77] The daytime high bias in G in VIC's control run does not answer the paradox. Figure 15b shows the LST of the Mosaic and VIC tests of lower heat capacity (and the later Noah test) alongside control runs. Figure 15b reveals that the lower heat capacity (and its lower midday ground heat flux) in the VIC test does raise VIC's midday LST as physically expected, but only slightly, leaving it still well below the warmer LST of Noah. Figure 15b shows that the lower heat capacity test in Mosaic does yield a nontrivial increase in midday LST, reducing Mosaic's July midday cool bias by roughly half (though introducing a nighttime cool bias, because of the reduced nighttime release of stored ground heat diminished by the reduced heat capacity).

[78] The chief explanation of the paradox of VIC versus Noah midday summer LST lies in significant intermodel differences in aerodynamic conductance. In the three models, the sensible heat flux H (W m^{-2}) is computed from the typical bulk transfer formulation given by

$$H = -\rho\, c_p\, C_h |V| (T_a - \mathrm{LST}) \tag{2}$$

where ρ is the air density (kg m^{-3}), c_p the heat capacity for air (1004.5 J kg^{-1} K^{-1}), $|V|$ the wind speed (m s^{-1}), T_a the air temperature (K), and C_h the surface turbulent exchange coefficient for heat. The product $C_h|V|$ is the aerodynamic conductance (m s^{-1}), and its reciprocal is the aerodynamic resistance. C_h manifests a strong diurnal cycle with larger values during daytime heating. In (2), positive H means a heat source to the atmosphere and heat sink to the land surface for daytime LST exceeding T_a. The models get common surface forcing values of ρ, $|V|$, and T_a. Only C_h and LST in (2) are computed uniquely in each model. Therefore Noah can have higher midday values of LST than VIC simultaneously with lower midday values of H than VIC if and only if Noah has lower values of C_h. Figure 16 depicts the July 1998 monthly mean diurnal cycle of C_h for each model, averaged across the 14 EBBR stations. The line with solid squares in Figure 16 is a Noah test, discussed later, using a modified roughness length for heat. Indeed, the Noah control run has substantially smaller daytime values of aerodynamic conductance, and hence C_h, than Mosaic, and far smaller values than VIC. Follow-on research will seek to derive C_h explicitly from the EBBR observations.

[79] The smaller C_h values for Noah inferred from Figure 16 motivated a sensitivity run. The treatment of C_h in Noah was the subject of the NCEP study by *F. Chen et al.* [1997], which examined the impact on C_h of the chosen formulation for the roughness length for heat, z_{0t}. From a suite of tests, Chen et al. recommended the z_{0t} formulation of *Zilitinkevich* [1995], which is based on the dynamic roughness Reynolds number and includes an adjustable parameter, denoted here C_z, in the range 0–1. The Noah control run in NLDAS uses $C_z = 0.2$. Decreasing C_z increases z_{0t}, which increases C_h (thus increasing aerodynamic conductance) and the land/atmosphere coupling, thereby decreasing daytime

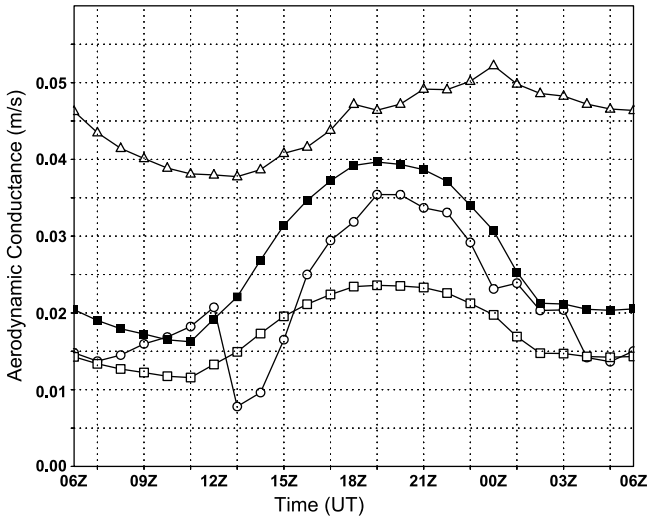

Figure 16. Monthly mean diurnal cycle of aerodynamic conductance (ms^{-1}) averaged over all ARM/CART SIRS sites for July 1998 for the control runs of Noah (open squares), VIC (open triangles), and Mosaic (open circles) and the Noah test (solid squares) using a modified formulation for roughness length for heat.

LST. This was investigated by a Noah sensitivity test, depicted in Figure 16, which used $C_z = 0.05$ in place of $C_z = 0.2$.

[80] The July monthly mean, midday aerodynamic conductance values in the Noah test in Figure 16 exceed the control values by nearly 70%. The surface fluxes and LST of this Noah test are depicted in Figure 14 (right columns) and Figure 15, respectively. Figure 15 shows a pleasing 2–3 K decrease in Noah's July peak daytime LST, cutting the midday warm bias by about half. Inspection of Noah test versus control surface fluxes in Figure 14 reveals small midday changes in R (increase) and G (decrease) of about 10–20 W m^{-2} each, as expected from the lower midday LST. This slightly improved the already small biases in Noah R and G and yielded a roughly 20–40 W m^{-2} increase in midday "available energy" ($R - G$) for sensible and latent heat fluxes. Once again however, very similar to the earlier response in Mosaic and VIC to increases in (R – G), the LE change in the Noah test in both July and April was negligible, especially so in July, with the $R - G$ increase in July once again being realized almost entirely by an increase in H.

[81] The lack of change in LE in July strongly indicates that the evaporative resistances to canopy transpiration (canopy resistance) and bare soil evaporation are much larger than, and thus dominant over, the aerodynamic resistance in influencing LE in this situation. Again, Figure 1a shows that the vegetation cover over the bulk of the ARM-SGP region in July is of order 40–80%. The analysis of *Vogel et al.* [1995] shows that even over an irrigated midlatitude wheat field in June, the LE change from a 20% change in aerodynamic resistance yielded only a 2% change in LE; that is, canopy resistances over nonsparse vegetation in midlatitude summer are typically much larger than, and dominant over, aerodynamic resistances, even when the soil is wet and contributing to a rise in canopy conductance. These results strongly suggest that the remaining July bias of +2 K in the Noah test in Figure 15 is caused by an overly high canopy resistance. Moreover, the aerodynamic conductance of VIC in Figure 16 may well be too high, thus acting to preclude a much higher midday warm bias in LST that would more properly reflect VIC's high bias in Bowen ratio (H/LE) in July in Figure 14. Sensitivity tests of aerodynamic resistance and canopy resistance will be one focus in follow-on NLDAS studies.

4.3. Satellite-Based Validation of Land Surface Skin Temperature

[82] The GOES LST fields are produced by the GCIP partnership of NESDIS and UMD in GOES land surface products. In this section, after validating GOES LST against ARM LST as a benchmark, we use GOES LST to assess NLDAS LST over north central CONUS. The study is limited to nonmountainous regions, to avoid shadowing effects on the GOES LST retrievals. The retrievals are obtained from GOES-East (GOES-8) and provide fields of hourly LST at 0.5° spatial resolution in cloud-free conditions during daytime. The LST retrieval provides a single aggregate LST for each 0.5° target scene. We bilinearly interpolate the LST fields to the 1/8° NLDAS grid. The GOES LST is retrieved only at 0.5° targets deemed 100% cloud-free. Cloud detection is based on that of earlier GOES insolation-retrieval studies such as *Tarpley* [1979], as refined in later studies such as PT-N. Despite the 100% cloud-free criteria, clouds may still be present in the scene owing to (1) optically thin cirrus, (2) subresolution or "subpixel" cloud (fair weather cumulus), and (3) difficulty of cloud detection over snow cover.

[83] GOES LST is retrieved by the so-called "split-window" technique of *Wu et al.* [1999], in which LST is obtained from a linear regression of the GOES brightness temperatures in the 11 μm and 12 μm bands. The regression coefficients were derived assuming a surface emissivity of $\varepsilon = 1$. This assumption is valid over land surfaces of nonsparse vegetation or snowpack, but less valid over rather bare soils (wherein $\varepsilon = 0.91–0.97$). Uncertainty from emissivity issues is avoided in this study by staying over nonsparse vegetation and by our universal application of $\varepsilon = 1$ in (1) the NLDAS models, (2) the in situ ARM/SIRS sites, and (3) the GOES retrievals.

[84] We assess GOES LST here against the in situ LST observations of the 22 SIRS sites. We limit the assessment to nonwinter, as our future assimilation of GOES LST will generally be confined to the warm season of stronger coupling between LST and soil moisture. Figure 17 presents the monthly and multistation mean of the daytime hourly diurnal cycle of GOES LST and ARM LST for April and July in 1998 and 1999. The data samples for Figure 17 (and Figure 18) represent only locations and times when the GOES cloud screening detected zero cloud. In Figure 17, the GOES LST demonstrates a remarkable ability to match the station-observed mean diurnal cycle, though it shows a small cool bias (likely from undetected clouds) of order 0–1.5 K before 18 UTC (local noon) and 1–2.5 K thereafter. The smaller cool bias in the morning is likely from less prevalent cloud cover then (e.g., subpixel cumulus). In future data assimilation, one may mitigate this cool bias by assimilating the 3-hour rise in GOES LST before noon (about 15–18 UTC here), rather than LST itself. *Tarpley* [1994] applied the morning rise of GOES LST to infer

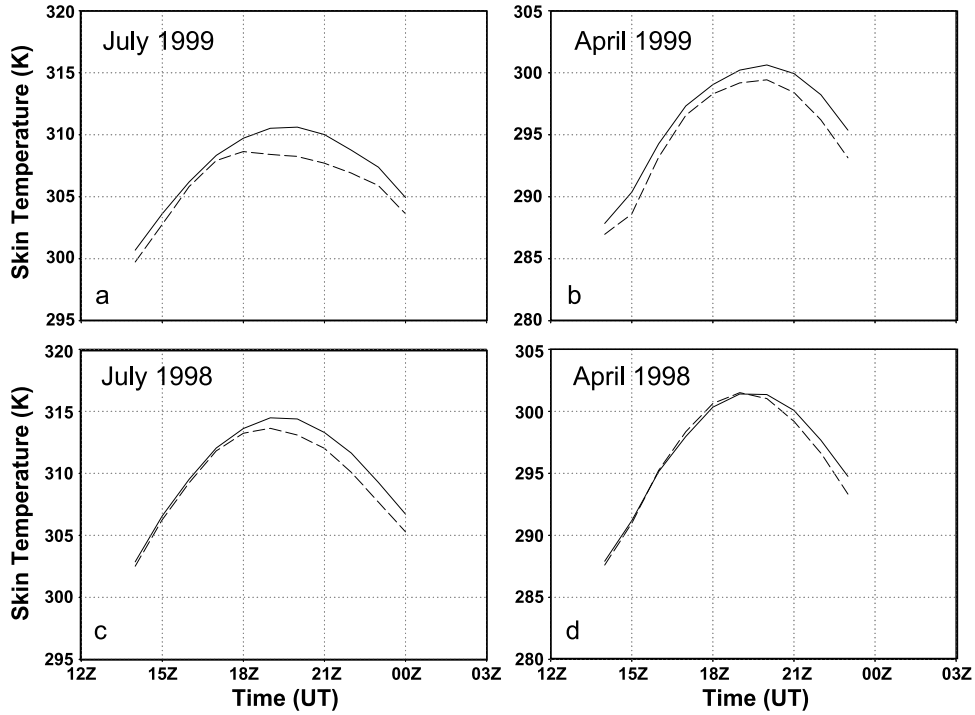

Figure 17. Monthly mean diurnal cycle of LST (K) averaged over all ARM/CART SIRS sites for (left) July and (right) April during (a and b) 1999 and (c and d) 1998 from SIRS observations (solid lines) and GOES-East retrieval (dashed lines).

monthly mean surface-moisture availability. Like the ARM LST, the 18 UTC GOES LST in Figure 17 is warmer (3–4 K) and its preceding 3-hour morning rise is larger (by about 1K) in July 1998 than July 1999. This interannual LST variability reflects the drought episode over the ARM region in July 1998, thus conditions were warmer and drier than in 1999 (and likely less cloudy, hence the smaller GOES LST afternoon cool bias in July 1998 versus 1999).

[85] Our goal is to use GOES LST retrievals to assess NLDAS LST over large regions that lack in situ observations of LST or surface fluxes. As a benchmark for that goal, Figure 18 illustrates, at the SIRS sites during July and April 1999 for 18 UTC, a pleasing similarity between GOES-based and ARM-based site-by-site match-ups with model LST. Moreover, all three models show good skill in either the GOES or ARM validation setting by yielding rather tight clusters close to the diagonal (and hence high correlations, shown later). In each month, the separate GOES and ARM match-ups use the same sample of instances where the GOES deemed the site to be cloud free. In Figure 18, the sample size of 198 in April (out of a possible 660 = 30 days x 22 stations) is notably smaller than that of 334 (out of 682) in July, as the GOES cloud screening detects cloud more often in the spring. One would expect this from the natural trend of decreasing cloud cover from spring to summer and the greater likelihood in July of shallow, subresolvable cumulus. Indeed in July, the GOES LST in Figure 18 manifests a small (but nonnegligible) leftward-pointing "cold tail" of outlier values that are not present in either the ARM observations or the models and thus likely represent GOES cloud detection failures. Similarly, the ARM observations in April 1999 show several warm outlier values (near 315 K), not present in either the GOES or model LST, likely representing bad ARM station observations.

[86] Most importantly, as desired, the GOES versus model match-ups yield the same sense of model midday LST bias as we derived from ARM data alone in the prior section. Table 5 compares the GOES-based versus ARM-based model bias, error standard deviation and correlation obtained from the Figure 18 match-up and listed top-down from warmest to coldest model bias. The table shows good agreement between the sign and magnitude of the GOES-based and ARM-based model bias. The GOES-based model bias is order 1 K warmer than the ARM-based model bias, owing to the aforementioned GOES LST cool bias of order 1 K versus ARM LST. The GOES-based model LST bias essentially reproduced the ARM-based model bias, both in (1) the absolute sense of correct sign and reasonably good magnitude and (2) the relative sense between models and between spring and summer season. Specifically, in agreement with the ARM-based LSM signatures of control-run midday LST bias presented for the entire annual cycle in RL-N, the GOES-based model bias results in Figure 18 show that in summer (1) Noah has the largest warm bias, which becomes much smaller in spring, (2) VIC has a smaller and modest warm bias, which becomes virtually zero in spring, and (3) Mosaic has a modest cool bias, which becomes larger in spring. Similarly, the GOES-based and ARM-based standard deviations in Table 5 from Figure 18 are in reasonable agreement, both in the range of 3–4 K. Finally, the GOES-based correlations with model LST in Table 5 are very encouraging, ranging between 0.66 and 0.78, with

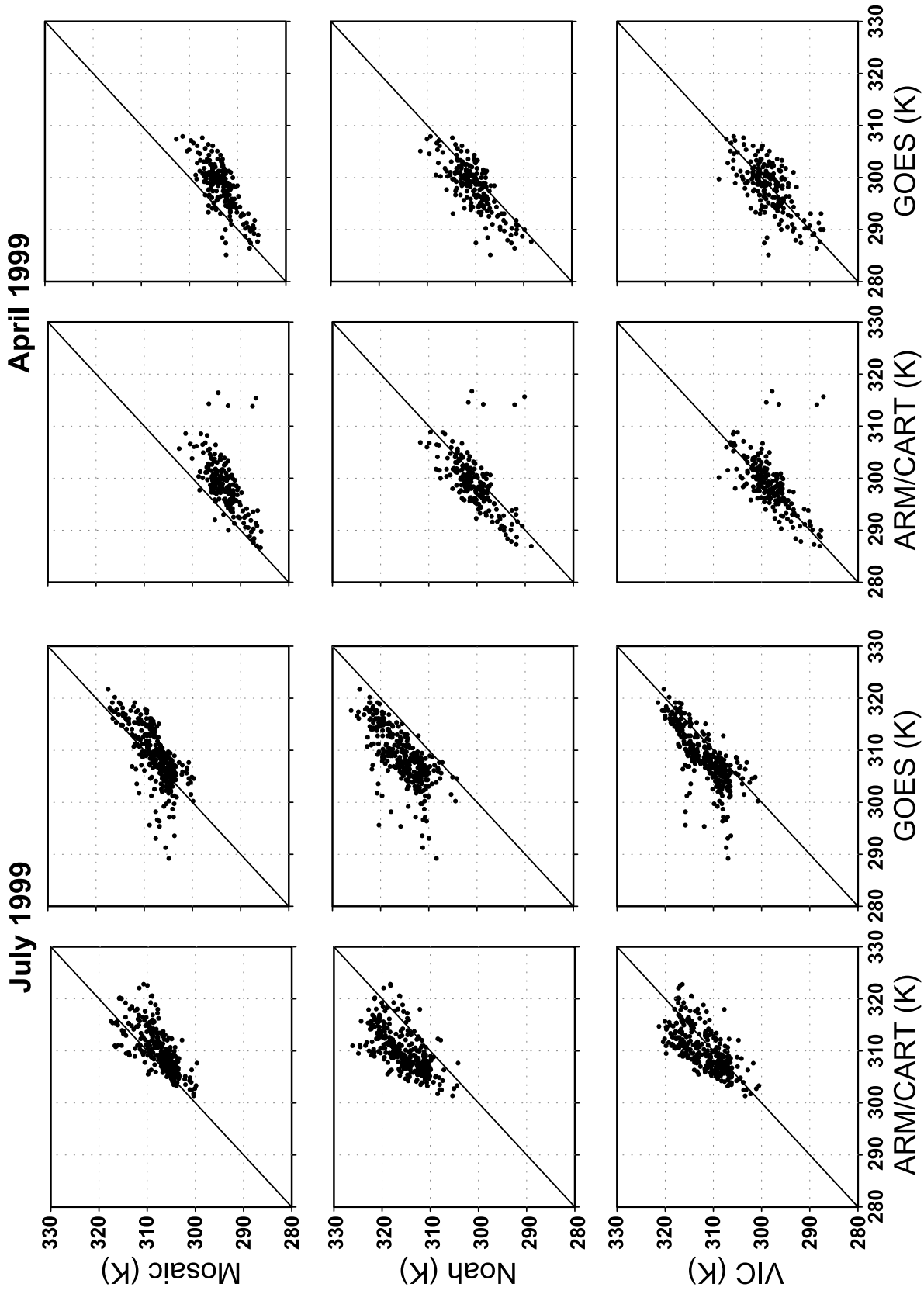
July 1999
April 1999
Mosaic (K)
Noah (K)
VIC (K)
ARM/CART (K)
GOES (K)
280
290
300
310
320
330

Figure 18.

Table 5. Bias, Error Standard Deviation, and Correlation of Model LST Versus Both ARM LST and Unscreened GOES LST Over the SGP From Figure 18[a]

	July 1999 (GOES/ARM)			April 1999 (GOES/ARM)		
Model	Bias, K	STDE, K	Correlation	Bias, K	STDE, K	Correlation
Noah	+6.5/+5.4	3.8/3.5	0.70/0.64	+2.4/+1.7	2.9/4.1	0.78/0.61
VIC	+2.8/+1.7	3.4/3.3	0.76/0.68	+0.3/−0.5	3.9/4.3	0.66/0.57
Mosaic	−1.2/−2.3	3.7/3.0	0.70/0.72	−4.7/−5.4	3.0/4.0	0.77/0.62

[a]STDE, error standard deviation.

five of six values of 0.70 or more, all without any screening of the GOES cold LST outliers. The correlations of model LST with GOES LST are consistently higher than the ARM-based correlations, likely from better match of the spatial scale of the GOES footprint and the NLDAS grid-cell size, versus the point scale of ARM values.

[87] Encouraged by Figures 17 and 18, we evaluated model LST against GOES LST in Figure 19 across a large region of the northern Midwest, bounded by latitudes 39°N and 53°N and longitudes 82°W and 98°W (straddling 90°W longitude, representing strict local noon at 18 UTC). This region is chosen for its (1) spatial separation from SGP, (2) nonsparse green vegetation in summer (Figure 1a), and (3) vivid model differences in evaporation in Figure 2. GOES versus model LST over this region at 18 UTC for July and April 1999 are presented in Figure 19. The ''screened'' results therein are described later. Sample counts in Figure 19 are 70,000–100,000 (2–3 orders larger than Figure 18, owing to the larger region). In Figure 19, we binned the data into 1 K intervals (for display only, kept full precision in statistics) and used colors to denote 4 orders of data counts: 0–10 (red), 10–99 (yellow), 100–999 (light green), and 1000–9999 (dark green). The green shades depict the vast majority of the sample and they manifest well-behaved, elongated clusters lying near and parallel to the diagonal, as desired.

[88] Yet the red and yellow ''tails'' of cold GOES LST in the unscreened panels of Figure 19 depict a nontrivial number of points that likely represent cloud detection failures, similar to the cold tails in the July GOES panels of Figure 18. Hence we used model simulated LST to screen the cold GOES LST tails in columns 2 and 4 in Figure 19. Specifically, we rejected a GOES LST if model-minus-GOES LST exceeded the unscreened, sample-wide model-minus-GOES LST bias by more than two times the model-minus-GOES LST standard deviation of the unscreened sample. Figure 19 shows that this screening preserves the high-density core region of the original data cluster, while eliminating the cold tails. The data counts of the three panels in any column of Figure 19 are identical in the unscreened case, but differ slightly (less than 1.2%) in the screened case, as the screening for each panel uses the given model's LST simulation. Last, in Figure 19, we point out the ''lower lobe'' of cold model LST in the April results of Mosaic and VIC. The lower edge of the lobe ends at freezing, suggesting that Mosaic and VIC have sustained remnants of melting snowpack too late into the spring in this region.

[89] Table 6 shows the GOES-based model LST bias, error standard deviation and correlation obtained from Figure 19. The screened model-bias results are warmer, as expected, but by a modest 0.5 to 1.0 K. More notably, the standard deviation of the model errors are significantly less by around 1.0–1.5 K and the already high correlations increase by around 0.1 to 0.76 or higher for almost all entries. It is revealing to compare the unscreened results from the north central CONUS in Table 6 with the SGP results in Table 5 (all unscreened). Table 6 preserves the relative nature and order noted in the biases in Table 5; namely, Noah is warmest and Mosaic is coolest in July, with Noah notably less warm and Mosaic notably more cool in April, while VIC falls in-between in both months. Yet the unscreened biases in Table 6 are consistently 1–3 K warmer than those in Table 5, reflecting that either the models are warmer in this region relative to GOES LST than over the ARM SGP region, or the GOES LST has a larger cool bias (of order 1–3 K) than the GOES cool bias over the ARM region. We suspect the latter owing to more cloud contamination in this more humid region.

[90] We look forward to use of GOES LST in future assimilation studies and LSM assessments. The validation and utility of GOES LST in mountainous regions awaits future study. Efforts continue at UMD and NESDIS to improve cloud detection and spatial resolution in GOES LST.

5. Summary and Conclusions

[91] A multi-institution partnership under the GCIP program has developed and evaluated the backbone for a North American Land Data Assimilation System (NLDAS). This paper is the overview of nine NLDAS papers (Table 2), which appear together in the online HTML version of the GCIP3 special section of *Journal of Geophysical Research*. These partners assembled a wide set of GCIP-sponsored products and other data sources into robust forcing data sets and multiscale validation databases. Validation applied surface stations measuring energy fluxes, surface meteorology, soil moisture and temperature, and mountain snowpack, plus daily streamflow observations and satellite-derived land surface temperature (LST) and snow cover.

[92] NLDAS features nonmodel sources of precipitation and solar insolation and the four LSMs of Noah, VIC,

Figure 18. Comparison of model (*y* axis) versus observed (*x* axis) LST (K) at 18 UTC over all ARM/CART SIRS sites for July (columns 1 and 2) and April (columns 3 and 4) 1999 for (top) Mosaic, (middle) Noah, and (bottom) VIC versus SIRS observations (columns 1 and 3) and GOES-East observations (columns 2 and 4). Match-up point included only if GOES LST is available (cloud free), yielding sample sizes of 334 in columns 1 and 2 and 198 in columns 3 and 4.

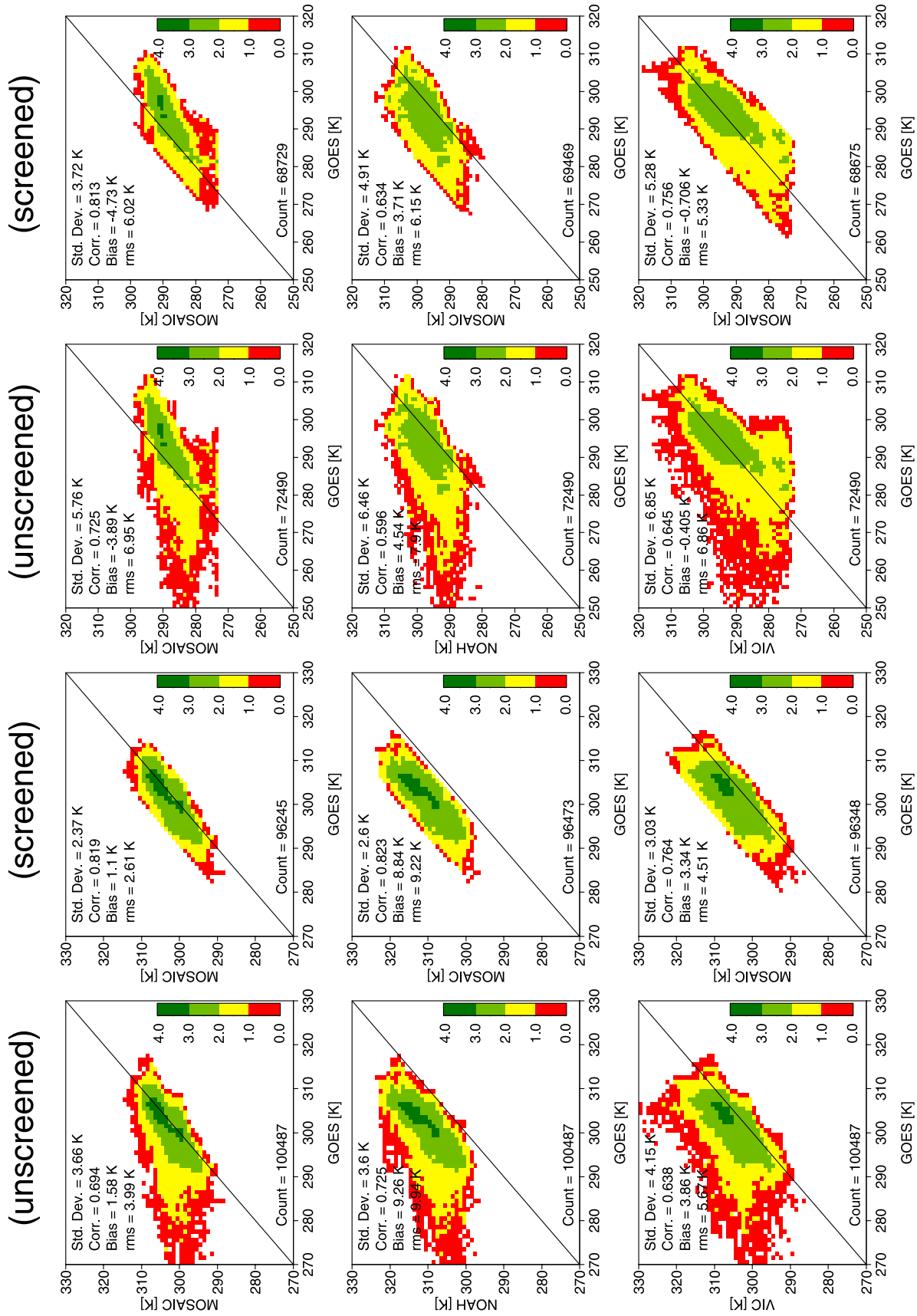
July 1999
(unscreened)
July 1999
(screened)
April 1999
(unscreened)
April 1999
(screened)
MOSAIC [K]
NOAH [K]
VIC [K]
GOES [K]
Std. Dev. = 3.66 K
Corr. = 0.694
Bias = 1.58 K
rms = 3.99 K
Count = 100487
Std. Dev. = 2.37 K
Corr. = 0.819
Bias = 1.1 K
rms = 2.61 K
Count = 96245
Std. Dev. = 5.76 K
Corr. = 0.725
Bias = -3.89 K
rms = 6.95 K
Count = 72490
Std. Dev. = 3.72 K
Corr. = 0.813
Bias = -4.73 K
rms = 6.02 K
Count = 68729
Std. Dev. = 3.6 K
Corr. = 0.725
Bias = 9.26 K
rms = 9.94 K
Count = 100487
Std. Dev. = 2.6 K
Corr. = 0.823
Bias = 8.84 K
rms = 9.22 K
Count = 96473
Std. Dev. = 6.46 K
Corr. = 0.596
Bias = 4.54 K
rms = 7.91 K
Count = 72490
Std. Dev. = 4.91 K
Corr. = 0.634
Bias = 3.71 K
rms = 6.15 K
Count = 69469
Std. Dev. = 4.15 K
Corr. = 0.638
Bias = 3.86 K
rms = 5.67 K
Count = 100487
Std. Dev. = 3.03 K
Corr. = 0.764
Bias = 3.34 K
rms = 4.51 K
Count = 96348
Std. Dev. = 6.85 K
Corr. = 0.645
Bias = -0.405 K
rms = 6.86 K
Count = 72490
Std. Dev. = 5.28 K
Corr. = 0.756
Bias = -0.706 K
rms = 5.33 K
Count = 68675

Figure 19.

Table 6. Bias, Error Standard Deviation, and Correlation of Model LST Versus Screened and Unscreened GOES LST for North Central CONUS From Figure 19[a]

	July 1999 (GOES: Unscreened/Screened)			April 1999 (GOES: Unscreened/Screened)		
Model	Bias, K	STDE, K	Correlation	Bias, K	STDE, K	Correlation
Noah	+9.3/+8.8	3.6/2.6	0.73/0.82	+4.5/+3.7	6.5/4.9	0.60/0.63
VIC	+3.9/+3.3	4.2/3.0	0.64/0.76	−0.4/−0.7	6.9/5.3	0.65/0.76
Mosaic	+1.6/+1.1	3.7/2.4	0.69/0.82	−3.9/−4.7	5.8/3.7	0.73/0.81

[a]STDE, error standard deviation.

Mosaic, and Sacramento (SAC) executing in parallel on a 1/8° CONUS domain to provide land-state background fields for data assimilation experiments. The infrastructure of NLDAS includes streamflow routing and provides both real-time and retrospective execution to support both operations and research. The hourly NLDAS forcing, now spanning seven years from October 1996 to present (at time of writing), represents an important NLDAS by-product.

[93] The NLDAS thrust here was the forerunner to recent companion initiatives in 50 + year retrospective executions of VIC on the NLDAS grid by *Maurer et al.* [2002] and Noah on the NLDAS grid by H. van den Dool of NCEP (private communication, 2003). We encourage researchers to compare our NLDAS water and energy budgets here with (1) these 50 + year retrospectives, (2) operational global and regional coupled 4DDA, and (3) global and regional reanalysis.

[94] A central distinction between the above suites is the source and bias in the surface forcing. In NLDAS here, the forcing is anchored by gage-based daily precipitation analyses (with hourly disaggregation using radar-derived precipitation) and hourly surface insolation derived from GOES satellites. All remaining forcing is from NCEP's mesoscale 4DDA system, known as EDAS. NLDAS surface forcing compared well against Mesonet observations over the SGP. In tests that replaced NLDAS forcing with local-station forcing, the test versus control differences in states and fluxes were pleasingly small. Yet we continue thrusts to further improve the forcing. The GOES-based solar insolation shows some high bias at low sun angles and over snowpack, though less bias then the fallback insolation from EDAS. At mountain SNOTEL sites in western CONUS, NLDAS precipitation has a nearly 50% low bias. Thus NLDAS partners have implemented a PRISM-based technique [*Daly et al.*, 1994] into the real-time forcing as of 1 February 2002 (with plans to reproduce the retrospective forcing using PRISM).

[95] Observed precipitation and streamflow applied to the annual water budget provided observation-based estimates of evaporation and runoff over large regions. This revealed substantial biases and intermodel differences in evaporation. The ARM-observed surface energy budget of the SGP confirmed the same evaporation bias anticipated from the annual water budget analysis. The three SVAT-type models, though they treat vegetation cover explicitly, nevertheless yield strikingly different warm season evaporation over vegetation. This canopy conductance disparity among the models is a foremost issue. Moreover, evapotranspiration bias can run counter to intuition. Though Mosaic has the shallowest root zone in NLDAS of the three SVATs, it has the highest warm season evaporation rates and hence highest warm season storage change in soil moisture, as it allows vigorous upward diffusion of water from the subroot zone.

[96] Aerodynamic conductance ($C_h|V|$) was a second area of large disparity. Overly large or small midday C_h values were found to substantially distort the expected correlation between daytime LST bias and Bowen ratio bias. Such distortion has crucial implications for the prospects of successful assimilation of satellite LST. In summer of the SGP, though Noah had a small high bias in Bowen ratio while VIC had a large high bias, Noah had the largest midday LST warm bias and VIC the smallest. The cause was the substantially lower C_h values in Noah versus VIC.

[97] Soil moisture storage emerged as a third area of large disparity, similar to previous PILPS studies. The forward radiative transfer models that are crucial to modern-era assimilation of satellite data are sensitive to absolute moisture states. Thus, while one can simulate evaporation and runoff well from good simulation of temporal change in soil moisture, land assimilation of satellite data brings a more stringent need for good absolute states of soil moisture.

[98] NLDAS simulations of snowpack water equivalent (SWE) at mountain SNOTEL sites showed a substantial low bias in all four LSMs, with an attendant low bias in runoff, owing to the cited high insolation bias over snow and low precipitation bias in mountains in NLDAS. Yet there was still notable disparity across the models in snow cover fraction, snow albedo and timing of spring snowmelt. Noah exhibited particularly low snow albedo, which conspired with the high insolation bias in the forcing to yield a very early bias in Noah seasonal snowmelt. The VIC and SAC models yielded the smallest biases in simulated SWE and regional snow cover, with VIC yielding the best snowmelt timing. The elevation tiling unique to VIC in NLDAS provides an advantage in snow state modeling. SAC snowpack simulations also performed well, as SAC's simple temperature index-based snow model bypasses surface energy balance and snow albedo, and thus avoids the

Figure 19. Comparison of model (y axis) versus GOES-East (x axis) LST (K) at 18 UTC over the northern Midwest during July (columns 1 and 2) and April (columns 3 and 4) 1999 for (top) Mosaic, (middle) Noah, and (bottom) VIC versus unscreened (columns 1 and 3) and screened (columns 2 and 4) GOES LST. See text for color scale definition and latitude/longitude range of region.

positive feedback loops that can plague snowpack simulations in surface energy balance models [*Slater et al.*, 2001; *Bowling et al.*, 2003].

[99] We emphasize that the Mosaic, VIC, and SAC LSM configurations in NLDAS differed in important aspects from their traditional configurations. For example, Mosaic executed with fixed soil-layer thicknesses and root depth, rather then vegetation-dependent spatial variability. VIC executed with one-hour rather than its typical three-hour time steps and with hourly rather than uniform daily disaggregation of precipitation. Moreover, the SAC runs, by design, were the first executions over a continental scale of the newly developed, semi-distributed version of SAC, with a priori noncalibrated parameters. Thus NLDAS execution of SAC provided an essential benchmark for future CONUS-wide SAC execution in semi-distributed mode.

[100] Indeed, all the model intercomparisons thus far in NLDAS must be viewed as incomplete and providing only a benchmark. The models were not calibrated to NLDAS configuration (e.g., spatial resolution, model time step, given fields of surface characteristics, temporal character of precipitation disaggregation). As in PILPS, our purpose is not to rank the models, but rather to build an enduring test bed via the NLDAS infrastructure for development of objective calibration approaches on very large continental scales that far exceed and complement the regional scales applied recently in PILPS [*Wood et al.*, 1998; *Bowling et al.*, 2003; *Boone et al.*, 2004].

[101] Last, this paper concluded with assessments and application of GOES-based LST at 1/2° hourly resolution. The diurnal cycle of GOES LST validated well against SGP flux-stations. Validation of model LST by means of GOES LST over the SGP gave validation scores similar to those from ground-based ARM observations. By using GOES LST to validate model LST over the northern Midwest, we found model LST biases consistent with those over the SGP. Hence GOES LST offers a powerful large-scale LSM validation tool.

[102] We are now assembling the tools to perform actual land data assimilation experiments. For this purpose, we will be adding one or two forward radiative transfer models into our common NLDAS infrastructure. Additionally, we are pursuing development of adjoint models and ensemble Kalman filter approaches. Finally, the NLDAS initiative here represented a pathfinder for a companion extension to a Global Land Data Assimilation System (GLDAS) by the NASA and NCEP partners of NLDAS. The GLDAS is described by *Rodell et al.* [2003] and is presently being ported from NASA to NCEP. With NLDAS and GLDAS together, NCEP and NASA and partners are striving to provide land state initial conditions for (1) land-memory predictability studies and (2) operational weather and climate model forecasts on daily to seasonal timescales.

Notation

4DDA	four-dimensional data assimilation
ARM/CART	Atmospheric Radiation Measurement/Cloud and Radiation Testbed (DOE)
ARS FAO	Agricultural Research Service Food and Agriculture Organization
AVHRR	Advanced Very High Resolution Radiometer on NOAA polar satellites
CAPE	convective available potential energy
CONUS	continental United States
CPC	Climate Prediction Center
CL-N	NLDAS paper by *Cosgrove et al.* [2003a]
CM-N	NLDAS paper by *Cosgrove et al.* [2003b]
DMIP	Distributed Model Intercomparison Project
DOE	Department of Energy
EDAS	NCEP Eta-model-based 4-D Data Assimilation System
EMC	Environmental Modeling Center of NCEP
EBBR	energy balance Bowen ratio flux stations in ARM/CART network
EF	extended facility flux stations in ARM/CART network
GAPP	GEWEX America Prediction Project
GCIP	GEWEX Continental-Scale International Project
GEWEX	Global Energy and Water Cycle Experiment
GOES	Geosynchronous Operational Environmental Satellite (USA)
GRIB	Gridded Binary data file format (WMO standard)
GSFC	Goddard Space Flight Center (NASA)
GSWP	Global Soil Wetness Project
GTOPO30	Global (30 arc seconds) digital elevation database
GVF	green vegetation-cover fraction
IMS	Interactive MultiSensor Snow (NESDIS)
LAI	leaf area index
LDAS	land data assimilation system
LM-N	NLDAS paper by *Lohmann et al.* [2004]
LR-N	NLDAS paper by *Luo et al.* [2003]
LSM	land surface model
LST	land surface skin temperature
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCDC	National Climatic Data Center
NCEP	National Centers for Environmental Prediction
NDVI	normalized difference vegetation index
NESDIS	National Environmental Satellite, Data, and Information Service
NLDAS	North American LDAS
NOAA	National Oceanic and Atmospheric Administration
NRCS	National Resources Conservation Service (USDA)
NWIS	National Water Information System
NWP	numerical weather prediction
NWS	National Weather Service (NOAA)
OGP	Office of Global Programs (NOAA)
OHD	Office of Hydrologic Development (NWS, formerly Office of Hydrology)
ORA	Office of Research and Applications of NESDIS
PAR	photosynthetically active radiation
PILPS	Project for Intercomparison of Land-Surface Parameterization Schemes

PRISM	Parameter-Elevation Regressions on Independent Slopes Model
PS-N	NLDAS paper by *Pan et al.* [2003]
PT-N	NLDAS paper by *Pinker et al.* [2003]
RFC	River Forecast Center (NWS)
RL-N	NLDAS paper by *Robock et al.* [2003]
SAC	Sacramento model (Sacramento Soil Water Accounting Model)
SCE	snow cover extent
SD-N	NLDAS paper by *Schaake et al.* [2004]
SGP	Southern Great Plains (field program)
SIRS	Solar and Infrared Radiation Station
SNOTEL	Snowpack Telemetry network of the NRSC
SNOW-17	Snow accumulation and ablation model (NWS/OHD)
SP-N	NLDAS paper by *Sheffield et al.* [2003]
SST	sea surface temperature
STATSGO	State Soil Geographic Database
SURFRAD	Surface Radiation Budget Network (NOAA/OAR-ARL, OGP)
SVAT	Surface-Vegetation-Atmosphere Transfer (model)
SWE	snowpack water equivalent
TOGA	Tropical Ocean Global Atmosphere
UMD	University of Maryland
USGS	U.S. Geological Survey
VIC	Variable Infiltration Capacity LSM
WMO	World Meteorological Organization
WSR-88D	Weather Service Radar-Doppler

[103] **Acknowledgments.** The work by NCEP/EMC, NWS/OHD, and NESDIS/ORA was supported by the NOAA OGP grant for the NOAA Core Project for GCIP/GAPP (K. Mitchell, J. Schaake, J. Tarpley, Co-PIs). The work by NASA/GSFC was supported by NASA's Terrestrial Hydrology Program (P. Houser, PI). The work by Rutgers University was supported by NOAA OGP GAPP grant GC99-443b (A. Robock, PI), the Cook College Center for Environmental Prediction, and the New Jersey Agricultural Experiment Station. The work by Princeton was supported by NOAA OGP GAPP grant NA86GPO258 (E. Wood, PI). The work by NCEP/CPC was supported by NOAA/NASA GAPP Project 8R1DA114 (R. Higgins, PI). The work by University of Maryland was supported by grants NA56GPO233, NA86GPO202, and NA06GPO404 from NOAA OGP and by NOAA grant NA57WC0340 to University of Maryland's Cooperative Institute for Climate Studies (R. Pinker, PI). Some figures were drawn with GrADS, created by Brian Doty. We thank DOE for the ARM/CART meteorological and heat flux data that were provided to the project at no cost and the NOAA Office of Global Programs and NASA Land Surface Hydrology Program for their purchase of the Oklahoma Mesonet meteorological and soil moisture and temperature data for their funded investigators. Dag Lohmann worked on the GCIP Core Project at NCEP under the University Corporation for Atmospheric Research Visiting Scientist Program.

References

Abdulla, F. A., D. P. Lettenmaier, E. F. Wood, and J. A. Smith (1996), Application of a macroscale hydrologic model to estimate the water balance of the Arkansas-Red river basin, *J. Geophys. Res.*, *101*, 7449–7459.

Anderson, E. A. (1973), National Weather Service River Forecast System: Snow Accumulation and Ablation Model, *NOAA Tech. Memo. NWS Hydro-17*, Natl. Weather Serv., Silver Spring, Md.

Berbery, E. H., Y. Luo, K. E. Mitchell, and A. K. Betts (2003), Eta model estimated land surface processes and the hydrological cycle of the Mississippi basin, *J. Geophys. Res.*, *108*(D22), 8852, doi:10.1029/2002JD003192.

Betts, A., F. Chen, K. Mitchell, and Z. Janjic (1997), Assessment of the land surface and boundary layer models in two operational versions of the NCEP Eta model using FIFE data, *Mon. Weather Rev.*, *125*, 2896–2916.

Boone, A., et al. (2004), The Rhone-Aggregation Land Surface Scheme Intercomparison Project: An overview of results, *J. Clim.*, *17*, 187–298.

Bowling, L. C., et al. (2003), Simulation of high latitude hydrological processes in the Torne-Kalix basin: PILPS Phase 2(e): 1. Experiment description and summary intercomparisons, *Global Planet. Change*, *38*, 1–30.

Burnash, R. J. C., R. L. Ferral, R. A. McGuire (1973), A generalized streamflow simulation system: Conceptual models for digital computers, technical report, Joint Fed.-State River Forecast Cent., U.S. Natl. Weather Serv. and Calif. Dep. of Water Resour., Sacramento, Calif.

Chen, F., K. Mitchell, J. Schaake, Y. Xue, H.-L. Pan, V. Koren, Q. Y. Duan, M. Ek, and A. Betts (1996), Modeling of land-surface evaporation by four schemes and comparison with FIFE observations, *J. Geophys. Res.*, *101*, 7251–7268.

Chen, F., Z. Janjic, and K. 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.

Chen, T. H., et al. (1997), Cabauw experimental results from the Project for Intercomparison of Landsurface Parameterization Schemes (PILPS), *J. Clim.*, *10*, 1194–1215.

Cherkauer, K. A., and D. P. Lettenmaier (1999), Hydrologic effects of frozen soils in the upper Mississippi River basin, *J. Geophys. Res.*, *104*, 19,599–19,610.

Cosgrove, B. A., et al. (2003a), Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project, *J. Geophys. Res.*, *108*(D22), 8842, doi:10.1029/2002JD003118.

Cosgrove, B. A., et al. (2003b), Land surface model spin-up behavior in the North American Land Data Assimilation System (NLDAS), *J. Geophys. Res.*, *108*(D22), 8845, doi:10.1029/2002JD003316.

Daly, C., R. P. Neilson, and D. L. Phillips (1994), A statistical-topographic model for mapping climatological precipitation over mountainous terrain, *J. Appl. Meteorol.*, *33*, 140–158.

Dirmeyer, P., A. Dolman, and N. Sato (1999), The pilot phase of the Global Soil Wetness Project, *Bull. Am. Meteorol. Soc.*, *80*, 851–878.

Douville, H., P. Viterbo, J.-F. Mahfouf, and A. C. M. Beljaars (2000), Evaluation of the optimum interpolation and nudging techniques for soil moisture analysis using FIFE data, *Mon. Weather Rev.*, *128*, 1733–1756.

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 advances in the National Centers for Environmental Prediction operational mesoscale Eta model, *J. Geophys. Res.*, *108*(D22), 8851, doi:10.1029/2002JD003296.

Gibson, J. K., P. Kallberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano (1997), ERA description, *ERA-15 Proj. Rep. Ser. 1*, 72 pp., Eur. Cent. for Medium-Range Weather Forecasts, Reading, UK.

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.

Hansen, M. C., R. S. DeFries, J. R. G. Townshend, and R. Sohlberg (2000), Global land cover classification at 1 km spatial resolution using a classification tree approach, *Int. J. Remote Sens.*, *21*, 1331–1364.

Higgins, R. W., W. Shi, E. Yarosh, and R. Joyce (2000), Improved United States precipitation quality control system and analysis, *NCEP/Clim. Predict. Cent. Atlas 7*, Clim. Predict. Cent., Camp Springs, Md.

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

Jarvis, P. G. (1976), The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field, *Philos. Trans. R. Soc. London, Ser. B.*, *273*, 593–610.

Ji, M., A. Kumar, and A. Leetmaa (1994), A multiseason climate forecast system at the National Meteorological Center, *Bull. Am. Meteorol. Soc.*, *75*, 569–577.

Kalnay, E., et al. (1996), NCEP/NCAR 40-year Reanalysis Project, *Bull. Am. Meteorol. Soc.*, *77*, 437–472.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. Hnilo, M. Fiorino, and G. Potter (2002), NCEP-DOE AMIP-II Reanalysis (R-2), *Bull. Am. Meteorol. Soc.*, *83*, 1631–1643.

Koren, V., J. Schaake, K. Mitchell, Q. Duan, F. Chen, and J. Baker (1999), A parameterization of snowpack and frozen ground intended for NCEP weather and climate models, *J. Geophys. Res.*, *104*, 19,569–19,585.

Koren, V. I., M. Smith, D. Wang, and Z. Zhang (2000), Use of soil property data in the derivation of conceptual rainfall-runoff model parameters, paper presented at 15th Conference on Hydrology, Am. Meteorol. Soc., Long Beach, Calif., 10–14 Jan.

Koster, R., and P. Milly (1997), The interplay between transpiration and run-off formulations in land-surface schemes used with atmospheric models, *J. Clim.*, *10*, 1578–1591.

Koster, R., and M. Suarez (1994), The components of a SVAT scheme and their effects on a GCM's hydrological cycle, *Adv. Water Resour.*, *17*, 61–78.

Koster, R., and M. Suarez (1996), Energy and water balance calculations in the Mosaic LSM, *NASA Tech. Memo., 104606*, vol. 9, 60 pp.

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

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land surface water and energy fluxes for GCMs, *J. Geophys Res.*, *99*, 14,415–14,428.

Liang, X., D. P. Lettenmaier, and E. F. Wood (1996a), One-dimensional statistical dynamic representation of subgrid spatial variability of precipitation in the two-layer variable infiltration capacity model, *J. Geophys. Res.*, *101*, 21,403–21,422.

Liang, X., E. F. Wood, and D. P. Lettenmaier (1996b), Surface soil moisture parameterization of the VIC-2L model: Evaluation and modifications, *Global Planet. Change*, *13*, 195–206.

Lohmann, D., et al. (1998), The project for intercomparison of land-surface parameterization schemes (PILPS) phase 2(c) Red-Arkansas river basin experiment: 3. Spatial and temporal analysis of water fluxes, *Global Planet. Change*, *19*, 161–179.

Lohmann, D., et al. (2004), Streamflow and water balance intercomparison of four landsurface models in the North American Land Data Assimilation System (NLDAS), *J. Geophys. Res.*, *109*, D07S91, doi:10.1029/2003JD003517.

Luo, L., et al. (2003), Validation of the North American Land Data Assimilation System (NLDAS) retrospective forcing over the southern Great Plains, *J. Geophys. Res.*, *108*(D22), 8843, doi:10.1029/2002JD003246.

Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen (2002), A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States, *J. Clim.*, *15*, 3237–3251.

Miller, D. A., and R. A. White (1998), A conterminous United States multilayer soil characteristics data set for regional climate and hydrology modeling, *Earth Inter.*, 2, Paper No. 2.

Nijssen, B., D. P. Lettenmaier, X. Liang, S. W. Wetzel, and E. F. Wood (1997), Streamflow simulation for continental-scale river basins, *Water Resour. Res.*, *33*, 711–724.

Nijssen, B., R. Schnur, and D. P. Lettenmaier (2001), Global retrospective estimation of soil moisture using the variable infiltration capacity land surface model, 1980–93, *J. Clim.*, *14*, 1790–1808.

Oki, T., T. Nishimura, and P. Dirmeyer (1999), Assessment of annual runoff from land surface models using total runoff integrating pathways (TRIP), *J. Meteorol. Soc. Jpn.*, *77*, 69–89.

Pan, H.-L., and L. Mahrt (1987), Interaction between soil hydrology and boundary layer development, *Boundary Layer Meteorol.*, *38*, 185–202.

Pan, M., et al. (2003), Snow process modeling in the North American Land Data Assimilation System (NLDAS): 2. Evaluation of model-simulated snow water equivalent, *J. Geophys. Res.*, *108*(D22), 8850, doi:10.1029/2003JD003994.

Pinker, R. T., et al. (2003), Surface radiation budgets in support of the GEWEX Continental-Scale International Project (GCIP) and the GEWEX Americas Prediction Project (GAPP), including the North American Land Data Assimilation System (NLDAS) project, *J. Geophys. Res.*, *108*(D22), 8844, doi:10.1029/2002JD003301.

Radakovich, J. D., P. R. Houser, A. da Silva, and M. G. Bosilovich (2001), Results from global land-surface data assimilation methods, paper presented at 5th Symposium on Integrated Observing Systems, Am. Meteorol. Soc., Albuquerque, N. M., 14–19 Jan.

Ramsay, B. H. (1998), The interactive multisensor snow and ice mapping system, *Hydrol. Processes*, *12*, 1537–1546.

Reynolds, C. A., T. J. Jackson, and W. J. Rawls (2000), Estimating soil water-holding capacities by linking the Food and Agriculture Organization soil map of the world with global pedon databases and continuous pedotransfer functions, *Water Resour. Res.*, *36*, 3653–3662.

Roads, J. O., et al. (2003), GCIP water and energy budget synthesis (WEBS), *J. Geophys. Res.*, *108*(D16), 8609, doi:10.1029/2002JD002583.

Robock, A., et al. (2003), Evaluation of the North American Land Data Assimilation System over the southern Great Plains during the warm season, *J. Geophys. Res.*, *108*(D22), 8846, doi:10.1029/2002JD003245.

Rodell, M., et al. (2003), global land data assimilation System, *Bull. Am. Meteorol. Soc.*, in press.

Rogers, E., D. Deaven, and G. Dimego (1995), The Regional Analysis System for the operational "early" Eta model: Original 80-km configuration and recent changes, *Weather Forecast.*, *10*, 810–825.

Schaake, J. C., et al. (2004), An intercomparison of soil moisture fields in the North American Land Data Assimilation System (NLDAS), *J. Geophys. Res.*, *109*, D01S90, doi:10.1029/2002JD003309.

Schlosser, C. A., A. G. Slater, A. Robock, A. J. Pitman, J. K. Y. Vinnikov, A. Henderson-Sellers, N. A. Speranskaya, K. Mitchell, and PILPS 2(d) Contributors (2000), Simulations of a boreal grassland hydrology at Valdai, Russia: PILPS Phase 2(d), *Mon. Weather Rev.*, *128*, 301–321.

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

Sheffield, J., et al. (2003), Snow process modeling in the North American Land Data Assimilation System (NLDAS): 1. Evaluation of model-simulated snow cover extent, *J. Geophys. Res.*, *108*(D22), 8849, doi:10.1029/2002JD003274.

Slater, A. G., et al. (2001), The representation of snow in land surface schemes: Results from PILPS 2(d), *J. Hydrometeorol.*, *2*, 7–25.

Tarpley, J. (1979), Estimating incident solar radiation at the surface from geostationary satellite data, *J. Appl. Meteorol.*, *18*, 1172–1181.

Tarpley, J. D. (1994), Monthly evapotranspiration from satellite and conventional meteorological observations, *J. Clim.*, *7*, 704–713.

Verdin, K. L., and S. K. Greenlee (1996), Development of continental scale digital elevation models and extraction of hydrographic features, paper presented at Third International Conference/Workshop on Integrating GIS and Environmental Modeling, Natl. Cent. for Geogr. Inf. and Anal. Santa Fe, N. M., 21–26 Jan.

Vogel, C. A., D. D. Baldocchi, A. K. Luhar, and R. K. Shankar (1995), A comparison of a hierarchy of models for determining energy balance components over vegetation canopies, *J. Appl. Meteorol.*, *34*(10), 2182–2196.

Wigmosta, M. S., L. W. Vail, and D. P. Lettenmaier (1994), A distributed hydrology-vegetation model for complex terrain, *Water Resour. Res.*, *30*, 1665–1679.

Wood, E. F., D. P. Lettenmaier, X. Liang, B. Nijssen, and S. W. Wetzel (1997), Hydrological modeling of continental-scale basins, *Annu. Rev. Earth Planet. Sci.*, *25*, 279–300.

Wood, E., et al. (1998), The Project for Intercomparison of Land-Surface Parameterization Schemes (PILPS) Phase 2(c) Red-Arkansas River basin experiment: 1. Experimental description and summary intercomparisons, *Global Planet. Change*, *19*, 115–136.

Wu, X., W. P. Menzel, and G. S. Wade (1999), Estimation of sea surface temperatures using GOES-8/9 radiance measurements, *Bull. Am. Meteorol. Soc.*, *80*, 1127–1138.

Zilitinkevich, S. S. (1995), Non-local turbulent transport: Pollution dispersion aspects of coherent structure of convective flows, in *Air Pollution III*, vol. 1, *Air Pollution Theory and Simulation*, edited by H. Power, N. Moussiopoulos, and C. A. Brebbia, pp. 52–60, Comput. Mech., Billerica, Mass.

A. A. Bailey, B. H. Ramsay, and J. D. Tarpley, Office of Research and Applications, NESDIS, E/RA1 WWBG Room 712, 5200 Auth Road, Camp Springs, MD 20746, USA. (baileya@comcast.net; bruce.r.ramsay@noaa.gov; dan.tarpley@noaa.gov)

B. A. Cosgrove, J. K. Entin, and P. R. Houser, Hydrological Sciences Branch and Data Assimilation Office, NASA Goddard Space Flight Center, Mail Code 974.1, Greenbelt, MD 20771, USA. (brian.cosgrove@gsfc.nasa.gov; jared.entin@gsfc.nasa.gov; paul.r.houser@nasa.gov)

Q. Duan, V. Koren, and J. C. Schaake, Office of Hydrologic Development, NOAA/NWS, 1325 East-West Highway, SSMC2, Room 8356, Silver Spring, MD 20910, USA. (qingyun.duan@noaa.gov; victor.koren@noaa.gov; john.schaake@noaa.gov)

R. W. Higgins and W. Shi, Climate Prediction Center, National Centers for Environmental Prediction, NOAA/NWS, 5200 Auth Road, Room 605, Camp Springs, MD 20746-4304, USA. (wayne.higgins@noaa.gov; wshi@ncep.noaa.gov)

D. P. Lettenmaier, Department of Civil and Environmental Engineering, University of Washington, Roberts Hall, FX-10, Box 352700, Seattle, WA 98195-2700, USA. (dennisl@u.washington.edu)

D. Lohmann, J. Meng, and K. E. Mitchell, Environmental Modeling Center, National Centers for Environmental Prediction, NOAA/NWS, 5200 Auth Road, Camp Springs, MD 20746-4304, USA. (dag.lohmann@noaa.gov; jesse.meng@noaa.gov; kenneth.mitchell@noaa.gov)

L. Luo, M. Pan, J. Sheffield, and E. F. Wood, Department of Civil and Environmental Engineering, Princeton University, Room E208, E-Quad, Olden Street, Princeton, NJ 08544, USA. (lluo@princeton.edu; mpan@princeton.edu; justin@princeton.edu; efwood@princeton.edu)

C. H. Marshall, Department of Atmospheric Sciences, Colorado State University, Fort Collins, CO 80523, USA. (curtis@atmos.colostate.edu)

R. T. Pinker, Department of Meteorology, University of Maryland, College Park, 2213 Computer and Space Sciences Building, College Park, MD 20742-2425, USA. (pinker@atmos.umd.edu)

A. Robock, Department of Environmental Sciences, Rutgers University, 14 College Farm Road, New Brunswick, NJ 08901-8551, USA. (robock@envsci.rutgers.edu)