JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 108, NO. D22, 8843, doi:10.1029/2002JD003246, 2003

Validation of the North American Land Data Assimilation System (NLDAS) retrospective forcing over the southern Great Plains

Lifeng Luo,[1,2] Alan Robock,[1] Kenneth E. Mitchell,[3] Paul R. Houser,[4] Eric F. Wood,[5] John C. Schaake,[6] Dag Lohmann,[3] Brian Cosgrove,[4] Fenghua Wen,[5] Justin Sheffield,[5] Qingyun Duan,[6] R. Wayne Higgins,[7] Rachel T. Pinker,[8] and J. Dan Tarpley[9]

Received 29 November 2002; revised 24 June 2003; accepted 15 July 2003; published 4 November 2003.

[1] Atmospheric forcing used by land surface models is a critical component of the North American Land Data Assimilation System (NLDAS) and its quality crucially affects the final product of NLDAS and our work on model improvement. A three-year (September 1996–September 1999) retrospective forcing data set was created from the Eta Data Assimilation System and observations and used to run the NLDAS land surface models for this period. We compared gridded NLDAS forcing with station observations obtained from networks including the Oklahoma Mesonet and Atmospheric Radiation Measurement/Cloud and Radiation Testbed at the southern Great Plains. Differences in all forcing variables except precipitation between the NLDAS forcing data set and station observations are small at all timescales. While precipitation data do not agree very well at an hourly timescale, they do agree better at longer timescales because of the way NLDAS precipitation forcing is generated. A small high bias in downward solar radiation and a low bias in downward longwave radiation exist in the retrospective forcing. To investigate the impact of these differences on land surface modeling we compared two sets of model simulations, one forced by the standard NLDAS product and one with station-observed meteorology. The differences in the resulting simulations of soil moisture and soil temperature for each model were small, much smaller than the differences between the models and between the models and observations. This indicates that NLDAS retrospective forcing provides an excellent state-of-the-art data set for land surface modeling, at least over the southern Great Plains region. *INDEX TERMS:* 3337 Meteorology and Atmospheric Dynamics: Numerical modeling and data assimilation; 3322 Meteorology and Atmospheric Dynamics: Land/atmosphere interactions; 3307 Meteorology and Atmospheric Dynamics: Boundary layer processes; 3354 Meteorology and Atmospheric Dynamics: Precipitation (1854); 1818 Hydrology: Evapotranspiration; *KEYWORDS:* surface fluxes, land surface modeling, LDAS

Citation: Luo, L., et al., 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, 2003.

1. Introduction

[2] The accuracy of numerical weather forecasting depends on both the forecast model and initial conditions. Current data assimilation systems essentially deal with atmospheric observations from meteorological stations, radiosondes, and satellites. There is not much information available about land surface states, such as soil moisture, soil temperature and snow cover; a land data assimilation system (LDAS) is needed to fill in the gap. In off-line mode, an LDAS can take the atmospheric information including precipitation, radiation and other major meteorological variables to drive a land surface model (LSM), and the LSM can calculate land surface states based on the meteorological information. North American LDAS (NLDAS) is such a system. Four LSMs (Mosaic, Noah, VIC, Sacramento or SAC) are implemented in NLDAS, running at 1/8° latitude/longitude resolution over a domain that covers the continental US, part of Canada, and part of Mexico (125°W–67°W, 25°N–53°N). Hourly fields of surface var-

[1]Department of Environmental Sciences, Rutgers University, New Brunswick, New Jersey, USA.
[2]Now at Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.
[3]Environmental Modeling Center, National Centers for Environmental Prediction, Camp Springs, Maryland, USA.
[4]Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[5]Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.
[6]Office of Hydrologic Development, National Weather Center, Silver Spring, Maryland, USA.
[7]Climate Prediction Center, National Centers for Environmental Prediction, 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, Silver Spring, Maryland, USA.

0148-0227/03/2002JD003246$09.00

iables including soil moisture, soil temperature and surface fluxes and more are output from the system [*Mitchell et al.*, 2003].

[3] The validity of these surface fields depends on two major components of the system, the LSMs and the atmospheric forcing used by the LSMs. The four models participating in this project are state-of-the-art LSMs, and have been calibrated and validated in many previous studies [e.g., *Koster and Suarez*, 1992, 1996; *Chen et al.*, 1997; *Chen and Mitchell*, 1999; *Liang et al.*, 1994, 1996a, 1996b; *Cherkauer and Lettenmaier*, 1999; *Anderson*, 1973; *Burnash et al.*, 1973; *Koren et al.*, 2000]. *Robock et al.* [2003] presents an evaluation of these models in the context of the NLDAS project. The focus of the present paper is to validate the atmospheric forcing. We use the word "validation" rather than "evaluation" here, because we are seeking to validate the forcing data set.

[4] The quality of atmospheric forcing is crucial for offline tests of LSMs, which require information such as precipitation, downward shortwave and longwave radiation, near surface air temperature, humidity, and wind speed. Precipitation and solar radiation are the most important forcings due to their significant impact on the water and energy budgets. Water enters the land surface system through precipitation and leaves through evapotranspiration, which is mostly driven by incoming solar radiation. Other variables such surface humidity, surface air temperature and wind speed also affect evapotranspiration, but their role is less important. The atmospheric forcing used by the NLDAS models is derived from Eta Data Assimilation System (EDAS) and observations [*Cosgrove et al.*, 2003]. The retrospective forcing data set was used to spin up the four models from a common state on September 30, 1996 through September 30, 1999 so that they could run in real-time from then on using the real-time forcing data set [*Mitchell et al.*, 2003; *Cosgrove et al.*, 2003]. This paper aims to validate this NLDAS retrospective forcing data set against station observations over a limited region.

[5] This paper is organized as follows. Section 2 briefly describes the NLDAS retrospective forcing data set and observations that are used to validate the forcing. Section 3 focuses on comparison of the NLDAS retrospective forcing and observations at the station scale. Then the forcing experiments and the results are described in section 4. Conclusions are drawn in section 5.

2. NLDAS Forcing and Validation Data

[6] The NLDAS retrospective forcing is a combination of EDAS model output and observations. It is based on 3-hourly NCEP EDAS data and 3-hourly and 6-hourly Eta forecasts when EDAS data are unavailable. Model fields are spatially and temporally interpolated to the NLDAS grid and an hourly time step. Surface pressure, incident longwave radiation, 2 m air temperature, and humidity are also adjusted to account for the significant difference in topographies between EDAS/Eta and NLDAS. Actual observations of precipitation and downward solar radiation are used in the retrospective forcing rather than EDAS output to avoid systematic biases in model-provided forcings of these two important variables. Stage II hourly Doppler Radar and River Forecast Center gauge data [*Baldwin and Mitchell*, 1997], Climate Prediction Center daily gauge data [*Higgins et al.*, 2000], and reprocessed daily gauge data are used to derive the hourly NLDAS observed precipitation forcing. The merged product makes use of the strength of each above data set [*Cosgrove et al.*, 2003]. The instantaneous downward solar radiation at each hour is interpolated to the NLDAS grid from the 1/2° product derived by the University of Maryland from NOAA's Geostationary Operational Environmental Satellites (GOES) [*Pinker et al.*, 2003]. These data are only available when sun angles are not too low, so modeled downward solar radiation has to be used as a supplement when GOES retrieval is not available near sunrise and sunset.

[7] To evaluate this forcing data set, we need ground-based observations. Since these gridded forcings are valid at 1/8° resolution, which is roughly 11 × 14 km, it would be best if we had observations valid at the same scale to compare with the forcing. However, almost all ground-based observations are point measurements, and the representativeness and scale of these measurements must be determined by the scale of the variable that is measured.

[8] In the NLDAS domain, there are many observational networks and stations taking measurements of all the conventional meteorological variables, including temperature, pressure and, precipitation, but there are not many ground-based observations of solar radiation or longwave radiation available. The number of stations that take soil moisture, soil temperature, and surface flux observations is even smaller. These variables are important since they are the final NLDAS products that we want to evaluate. The southern Great Plains (SGP) of the United States is possibly the most intensively observed region in the world. There have been many field campaigns over this region (e.g., the SGP97 hydrology experiment [*Jackson et al.*, 1999], SGP99, and FIFE (First ISLSCP (International Satellite Land Surface Climatology Project) Field Experiment) [*Sellers et al.*, 1992]) and they helped spur the development of the two major networks that currently exist, the Oklahoma Mesonet and the Atmospheric Radiation Measurement/Cloud and Radiation Testbed (ARM/CART) (Figure 1).

[9] The Oklahoma Mesonet is a mesoscale meteorological monitoring network [*Brock et al.*, 1995], and consists of more than 100 automated meteorological stations (Figure 1). The network covers every county of Oklahoma, with an average station spacing of 32 km. Mesonet stations are predominantly located in rural areas, as free from anthropogenic influences as possible [*Shafer et al.*, 1993]. All the conventional meteorological variables and shortwave radiation are measured at these automated meteorological stations every 5 min. Wind speed and direction are measured at a height of 10 m and air temperature and relative humidity are measured at 1.5 m for consistency with existing NOAA cooperative observations and airport stations. The measurements are taken over natural vegetation. A tower stands near the center of a 10 m × 10 m plot of land, surrounded by a fence to secure the area from animals and nearby human activity. Quality assurance is performed by the Mesonet and quality flags are given in the data set. We only use data that are indicated as good by their quality assurance flag.

[10] The Atmospheric Radiation Measurement (ARM) program was founded by the U.S. Department of Energy

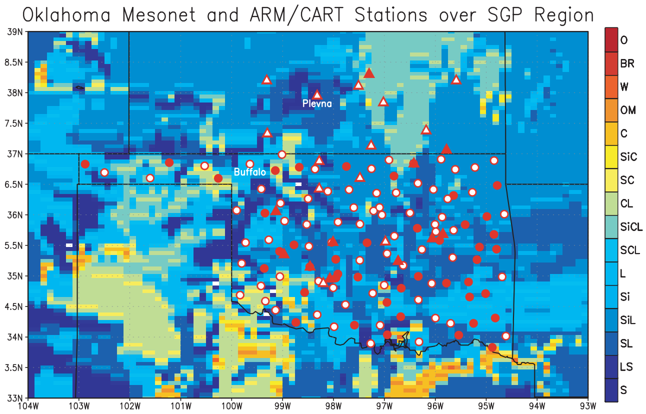

Figure 1. Surface stations of the Oklahoma Mesonet (circles) and ARM/CART stations (triangles) used in this paper. Open circles and open triangles are stations whose data was used in this study. Stations used as examples are indicated with their names. The background is the most predominant surface soil type, as specified by NLDAS (O, other; BR, bedrock; W, water; OM, organic materials; C, clay; SiC, silty clay; SC, sandy clay; CL, clay loam; SiCL, silty clay loam; SCL, sandy clay loam; L, loam; Si, silt; SiL, silty loam; SL, sandy loam; LS, loamy sand; S, sand).

to obtain field measurements and develop models to better understand the processes that control solar and thermal infrared radiative transfer in the atmosphere and at the Earth's surface. The SGP Cloud and Radiation Testbed (CART) site was the first field measurement site established by ARM, and consists of in situ and remote-sensing instrument clusters across north-central Oklahoma and south-central Kansas (Figure 1). A variety of instruments were installed at each ARM/CART site to obtain measurements of different variables. The Surface Meteorological Observation System (SMOS) uses conventional in situ sensors for surface wind speed, wind direction, air temperature, relative humidity, barometric pressure, and precipitation at the central facility and many extended facilities. The Solar Infrared Radiation Stations (SIRS) continuously measure broadband shortwave and longwave irradiances for both downwelling and upwelling components every minute and are implemented at many ARM/CART extended facilities. The estimated measurement uncertainties are less than 15 W m^{-2} for the shortwave and 6 W m^{-2} for the longwave [*Dutton et al.*, 2001]. In addition, an Energy Balance Bowen Ratio (EBBR) system and Soil Water And Temperature System (SWATS) are also available at many ARM/CART stations, which provide measurements of surface latent, sensible, and ground heat fluxes, as well as soil moisture and soil temperature. The data from EBBR and SWATS were not used in this study, but were used in the NLDAS evaluation [*Robock et al.*, 2003]. We combined the SMOS and SIRS data to produce a meteorological forcing data set at the extended facilities.

3. Comparison of NLDAS Forcing With Station Observations

[11] Direct comparison between station measurements and model output from nearby grid points from a weather forecast model or data assimilation system has produced very valuable information about the modeling systems [e.g., *Betts et al.*, 1993, 1998; *Robock et al.*, 1998; *Srinivasan et al.*, 2000]. Although the two quantities are from slightly different scales, systematic errors can still be easily identified from this simple direct comparison. As a first step of our work, we compare observations taken at one station with forcing from the NLDAS grid region where the station is located. The Mesonet meteorological data are available for 1998 and 1999 while the LDAS retrospective forcing spans 3 water years starting from October 1996. Because of this limitation, all the comparisons are shown only for the 21-month period of January 1998–September 1999.

[12] We conducted comparisons between station observations and LDAS forcing for 72 Mesonet stations and 13 ARM/CART stations. At each of these stations or NLDAS grids, the observed meteorological information was used to make the hourly forcing. The instantaneous values at the hours were used for all variables except precipitation, which was accumulated during the previous hour. We compared this local forcing with the NLDAS forcing. There are subtle differences between these two data sets besides their scales. Air temperature and humidity are observed at 1.5 m above ground while the NLDAS forcing is valid at 2 m [*Cosgrove et al.*, 2003]. The specific humidity is not available directly from the local observations and is calculated from temperature and dew point observations.

[13] Figure 2 presents the comparison at one ARM/CART station at Plevna, Kansas. This station has SMOS, SIRS, EBBR, and SWATS observations, so all the necessary variables are available to compare with the NLDAS retrospective forcing. There is generally a good agreement between the NLDAS retrospective forcing and the observations at this station. The bias and root mean square difference are generally small. Variables that have a larger temporal variability show a relatively larger root mean square difference, visually represented by larger scattering. For surface pressure, less variable in time and space, both bias and root mean square difference are small (Figure 2f). For shortwave radiation, more variable in time and space due to cloud cover, the bias is rather small but the root mean square difference is fairly large (Figure 2a). NLDAS tends to overestimate the downward shortwave radiation and underestimate the downward longwave radiation. For many hours NLDAS solar radiation (GOES solar radiation in most cases) is larger than the station observations. This might be due to local small scale cloud cover. Under those conditions, ground based radiation measurements are affected by local clouds but satellite measurements are not, since they would not consider the pixel as cloud-covered. The low bias in downward longwave radiation comes from EDAS and this finding agrees with the results from other EDAS validations from much smaller samples of stations [*Betts et al.*, 1997; *Hinkelman et al.*, 1999; *Berbery et al.*, 1999]. The high and low biases in the two different incoming energy terms cancel out somewhat, making the total amount of radiation received at the surface less biased. However, the solar radiation is also affected by the surface albedo.

[14] Specific humidity and air temperature generally agree well with observations at this station. There is more scatter in the wind speed comparison than in the

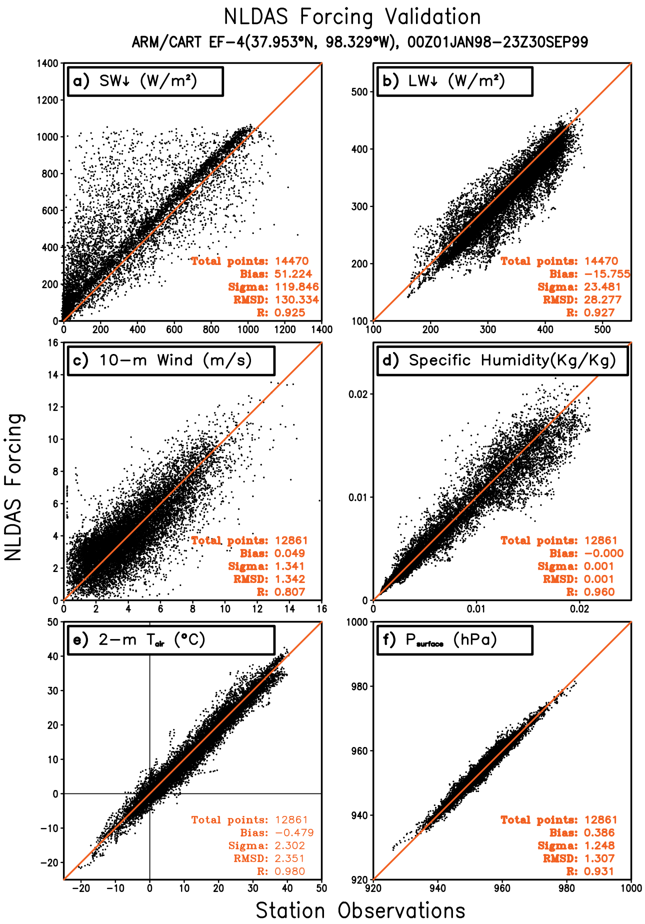

Figure 2. Comparison of NLDAS forcing with local forcing for standard meteorology variables for station EF-4 of ARM/CART at Plevna, Kansas, which is representative of other stations. Each point is for one hour during the period 0000 UT, 1 January 1998 through 2300 UT, 30 September 1999. The observed downward (a) shortwave and (b) longwave radiation are from the SIRS instruments. The other four observations are all from SMOS instruments. They are (c) 10-m wind speed (m/s), (d) specific humidity (kg/kg), (e) 2-m air temperature (°C), and (f) surface pressure (hPa). All state variables and fluxes are instantaneous values on the hour. The numbers at each panel are total number of points, bias, standard deviation (Sigma), root mean square different (RMSD) and correlation coefficient (R). The red solid line is the 1-to-1 line.

comparisons of specific humidity, air temperature, and surface pressure. This is related to the temporal and spatial variability of each variable. The comparison shown here includes all valid data during the 21-month period and there is no strong diurnal or seasonal dependence of the agreement found in the comparison.

[15] This one station is quite representative of all the stations. Table 1 summarizes the forcing differences for all Mesonet and ARM/CART stations. As expected, variables with larger spatial and temporal scales tend to agree more between NLDAS forcing and station observations. This finding is systematic at all stations across the SGP region, especially for downward solar radiation, which is spatially and temporally more variable, with a poorer agreement between the NLDAS forcing and local observations.

[16] In addition, there is a systematic difference between the NLDAS and observed downward shortwave radiation. Figure 3a shows the difference in monthly mean diurnal cycle of downward solar radiation between NLDAS forcing and observations averaged over all 72 Mesonet stations. The systematic overestimate of downward solar radiation in the local morning is obvious. This bias is fairly large given the absolute amount of downward solar radiation at low sun angles is not very large in the morning. Since NLDAS retrospective forcing is a merged product and downward solar radiation comes from both GOES retrieval and EDAS solar radiation, and EDAS downward solar radiation is only used when GOES retrievals are not available at very low sun angles, the high bias of solar radiation in the local morning exists both in EDAS and GOES fields. Surprisingly, the bias is not symmetric as we might have expected. There is not a high bias in the local afternoon although the sun angle is also low. To improve the forcing quality in the future, both GOES retrieval and EDAS modeled solar radiation have to be improved.

[17] This overestimate of downward solar radiation also seems to have a spatial pattern. Figure 3b is the difference in total amount of energy received by models for the 21-month period based on NLDAS forcing and station observations. The high bias in NLDAS forcing shows up in most parts of the region, but is highest in the East. We do not understand the reasons for this pattern at this stage.

[18] The precipitation in the NLDAS retrospective forcing comes from the observations as described above. Rain gage data from thousands of stations across the country are used in the unified precipitation analysis, and to improve its quality, almost all available observational networks are included in the data stream. It is impossible to find independent observations with good quality in this region to compare with NLDAS forcing. Therefore more precisely speaking, the following section is rather a comparison between the analyzed observed precipitation and the observed precipitation at individual stations. The goal of this comparison is to see how different they can be and what are the contributions to model simulations.

[19] The agreement of precipitation between NLDAS forcing and station observations is not as good as the other variables at the hourly timescale (Figure 4a). As an example, the comparison (for illustration, we still use Plevna, Kansas) shows that there are not many precipitation events where the two data sets agree well. Station observations show higher precipitation rates at the hourly time interval on many occasions. As we average the precipitation to longer timescales the agreement becomes much better (Figures 4b–4d). Since the NLDAS forcing uses daily total precipitation as the base and is temporally interpolated to hours using Stage II hourly Doppler radar and River Forecast Center gauge data, the agreement at the hourly timescale is not as good as that at the daily timescale. When the averaging period is longer, for five days or one month, the agreement gets even better. This is not difficult to understand given the spatial and temporal scale of precipitation and the way the NLDAS precipitation field is generated. It is very likely that a small-scale thunderstorm can hit a meteorological station and dump a lot of rain there

Table 1. Comparison Between NLDAS Forcing and Local Forcing Observed at Mesonet Stations and ARM/CART Stations[a]

	Bias	Sigma	RMSD	R
Air temperature, K	−0.30 (0.34)	2.30 (0.26)	2.34 (0.27)	0.97 (0.02)
Surface pressure, hPa	0.53 (1.78)	1.38 (0.39)	2.11 (1.01)	0.97 (0.07)
Wind speed, m/s	−0.51 (0.77)	1.53 (0.20)	1.76 (0.37)	0.75 (0.04)
Specific humidity, g/kg	−0.012 (0.189)	1.13 (0.43)	1.12 (0.26)	0.96 (0.01)
Downward shortwave radiation, W m^{-2}	44.78 (6.81)	119.30 (5.84)	127.52 (7.40)	0.92 (0.01)
Downward longwave radiation, W m^{-2}	−14.89 (4.46)	23.60 (3.02)	28.26 (2.63)	0.93 (0.02)

[a]Shown are differences between station-observed atmospheric forcing and the NLDAS forcing at those grids where the stations are located. At each station, mean of the difference (NLDAS forcing—local forcing) is denoted as Bias and the standard deviation in time of the differences is denoted as Sigma. The root mean square difference (RMSD) and correlation (R) are also calculated. The quantity in parenthesis is the spatial standard deviation of each mean over all the stations. Downward longwave radiation calculations come only from the ARM/CART SIRS stations.

but that other rain gages do not get much or even get nothing. When we analyze all the available rain gages onto a grid, the precipitation field tends to be smoothed in space and time (see *Ebert et al.* [2003] for more discussion of this issue). Therefore the precipitation record from an individual station tends to be spikier than the analyzed field.

[20] To quantitatively evaluate the agreement between NLDAS and observed precipitation, we use skill scores traditionally used to evaluate weather forecasts: the hit rate (HR), probability of detection (POD), and bias (*B*) of detection [*Wilks*, 1995]. Considering precipitation at each hour as a yes/no event and two time series of precipitation (S_1 and S_2) spanning *n* hours, *a*, *b*, *c*, and *d* are the number of hours for the following conditions respectively: both S_1 and S_2 are yes, S_1 is yes and S_2 is no, S_2 is yes an S_1 is no, and both S_1 and S_2 are no. Then,

$$HR = \frac{(a+d)}{n}, \quad POD = \frac{a}{a+c}, \quad B = \frac{a+b}{a+c}. \qquad (1)$$

[21] Figure 5a presents the results at two different time-scales from 72 Mesonet stations. At the hourly scale, HR is 92% averaged over all 72 stations, which means that 92% of

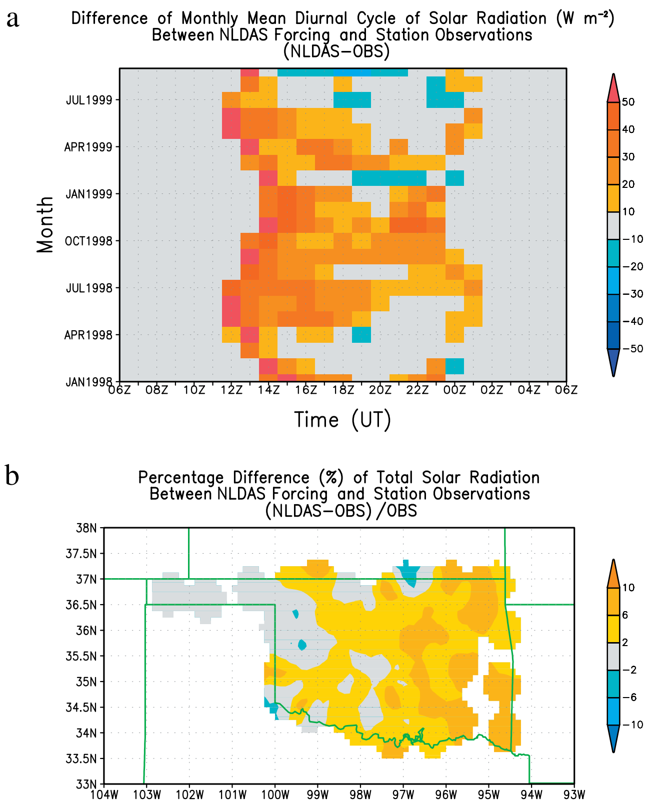

Figure 3. Comparison of NLDAS forcing with local forcing for downward solar radiation. (a) Difference in monthly mean diurnal cycle of downward solar radiation. (b) Percentage difference in total amount of energy received at the land surface from two data sets. The plot is made using objective analysis [*Cressman*, 1959] to show the spatial distribution of the differences.

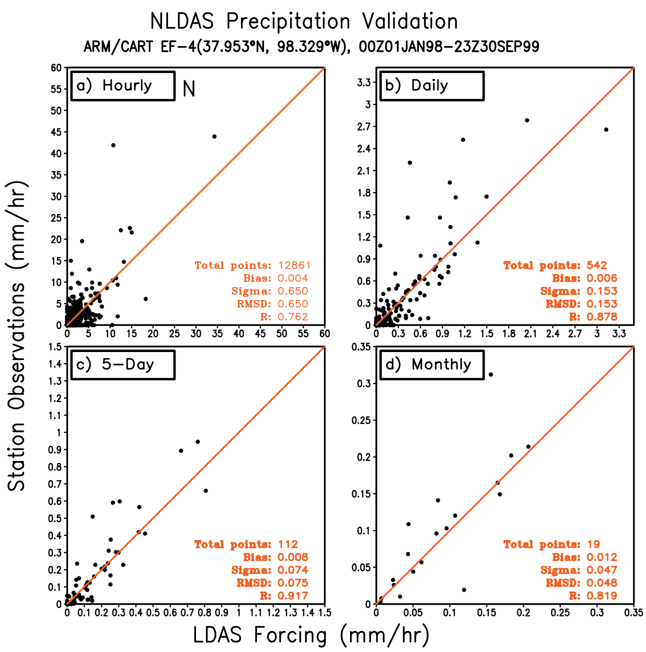

Figure 4. Comparison of NLDAS forcing with local forcing for precipitation for station EF-4 of ARM/CART at Plevna, Kansas, representative of other stations. For the hourly panel, each point is for one hour during the period 0000 UT, 1 January 1998 through 2300 UT, 30 September 1999. For the other panels, the averaging period is indicated.

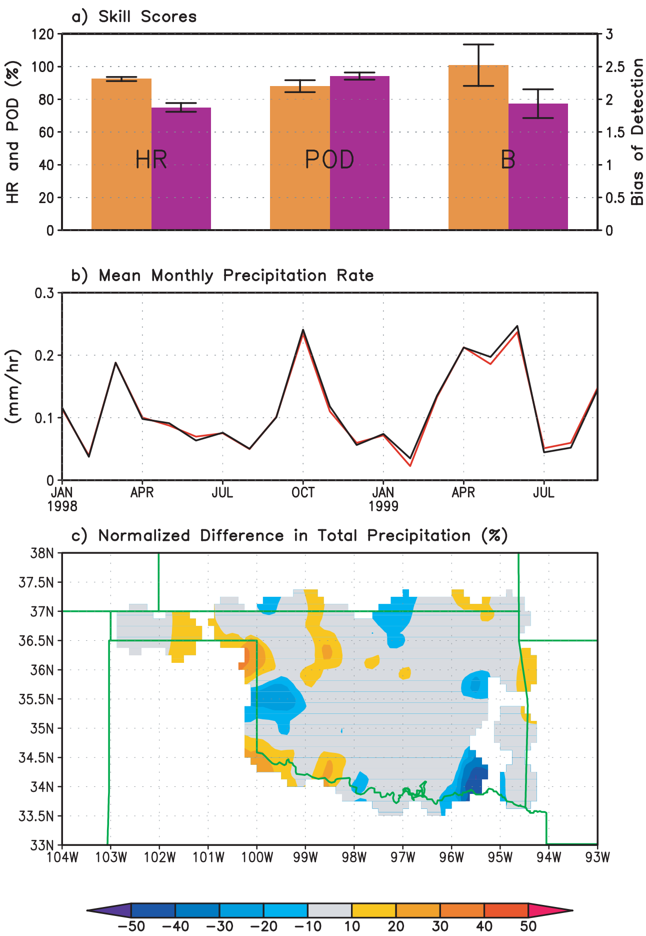

Figure 5. Comparison of NLDAS forcing with local forcing for precipitation for the 72 Oklahoma Mesonet stations indicated as open circles in Figure 1, for the period 0000 UT, 1 January 1998, through 2300 UT, 30 September 1999. (a) Mean Hit rate (HR), probability of detection (POD) and bias (B) of detection at two different timescales (hourly and daily, indicated by brown (left) and purple (right) bars, respectively) for all 72 stations. The error bars indicate the spatial standard deviation. (b) Time series of monthly mean precipitation rate averaged over all 72 Mesonet stations. The black line is the local observations and red line is NLDAS forcing. (c) Map of the percentage difference in total precipitation between two data sets for the period, normalized by the total amount of precipitation observed at stations. The plot uses objective analysis [*Cressman*, 1959] to show the spatial distribution of the differences.

the hours during the 21-month period the NLDAS forcing and the station observations show the same conditions, i.e., both have some precipitation or neither has precipitation. However, most of the time there is no precipitation and the agreement is not that meaningful. POD, a more meaningful ratio, is 88% averaged over all 72 stations. This means that 88% of the time when the stations observed some precipitation, NLDAS forcing also has some precipitation, regardless of the precipitation rate or amount of precipitation. The bias of detection *B* is about 2.5 at the hourly scale averaged over all 72 stations. Whenever *B* is larger than 1, it indicates an overestimate or overprediction (if it is for forecast verification). At the daily scale the average HR drops to 75%, POD goes up to 94%, and *B* drops to 1.9. The drop in HR from hourly to daily scale is simply because the number of days without rain in both time series gets smaller much more quickly than the size of the time series does. The change of POD expresses that NLDAS forcing is able to catch 94% of the occurrence of precipitation events during the 21-month over the SGP region. The bias is still larger than 1 although it gets smaller. The analyzed precipitation field tends to have more rain events than observed at one individual station. Once any of the rain gages in an analysis grid shows precipitation, the grid will show precipitation. Indeed, when different precipitation thresholds are chosen, the skill scores change accordingly. Lower POD is found for heavier precipitation events. This agrees well with other precipitation verification studies [*Ebert et al.*, 2003].

[22] Besides the timing of precipitation, land surface models also respond to the amount of precipitation they receive. We calculated the difference in total precipitation between the NLDAS forcing and local precipitation (Figure 5c). At one individual station, the monthly mean precipitation rate might differ a little between two data sets, but the average of all stations gives a very good agreement. In Figure 5c, at most stations the normalized difference in total precipitation is less than 10%, while some stations have fairly large values. A close check of the data shows that there were cases of very heavy precipitation reported by the Mesonet stations but only light rainfall in the NLDAS forcing. This is related to the aforementioned analysis technique. The quality control processes in the precipitation analysis or NLDAS forcing generation might also contribute to these differences. Some very large rainfall events are likely to have been eliminated by maximum allowable precipitation limits in the analysis by the lack of heavy precipitation at nearby stations. Fortunately, there are only a few stations showing such big differences. The rest of the region still agrees fairly well. On the basis of the above analysis, we conclude that the NLDAS precipitation field agrees reasonably well with station observations from Mesonet and ARM/CART over SGP region.

4. Forcing Experiments and Results

[23] To measure how important the differences between the NLDAS retrospective forcing and the local forcing are we now examine their impact on land surface model simulations by conducting a forcing experiment. *Robock et al.* [2003] show that there are significant differences between the modeled fields and observations including soil moisture, soil temperature and surface fluxes. Because these are all point observations, we really want to know whether the differences in models simulations are related to the differences in atmospheric forcing. If so, what are the relative contributions of these forcing differences?

[24] This experiment consists of two simulations with each model, one using the NLDAS retrospective forcing and the other using forcing constructed from station observations from 72 Mesonet stations and 13 ARM/CART stations over the SGP region. Hereafter, for convenience

we name the two simulations as "control" and "local" experiments, respectively. The observed downward solar radiation, precipitation, and other meteorological data are used in the local forcing. Downward longwave radiation, however, is only available at ARM/CART stations, and therefore is not included in the local forcing, but comes from the NLDAS forcing. There are some occasions or periods when one or more variables from station observations are missing. In those cases the values from the NLDAS retrospective forcing were used to fill in the gap. The local forcing only covers the 21 months. To avoid possible spin-up problems, the four models (Mosaic, Noah, VIC, and SAC) were run from September 1996 to January 1998 using the NLDAS forcing initialized at a common soil moisture and soil temperature state [*Mitchell et al.*, 2003]. The control run continues with NLDAS forcing throughout the rest of the period, while the local experiment switches to local forcing at 0000 UT 1 January 1998. The difference in the model simulated fields from the two simulations will be an indicator of the impact of the forcing differences.

[25] Time series of simulated top 40 cm soil moisture show that the differences between the control and local experiments are much smaller than the differences between the control simulations of the different models, and between each control simulation and the observations (Figure 6a). (The reasons for these inter-model differences are investigated by *Robock et al.* [2003].) With the local forcing, all the models tend to have a slightly drier soil. This is a bit surprising given that NLDAS downward solar radiation has a high bias. This has to be explained by the precipitation difference. As we mentioned in the previous comparison, local precipitation tends to be spikier because NLDAS uses an analyzed precipitation field. The change of temporal distribution of precipitation has some impact on model simulations [*Wood et al.*, 1998]. Although each model's soil moisture has slightly different responses to the forcing change, their relative changes are all less than 10% in most cases.

[26] Figure 7 summarizes the differences in some other variables between the control and local forcing experiment at the hourly timescale over all 72 Mesonet stations. All four models ran the two simulations, but only three of them are shown here. The SAC model is physically different from the others and does not calculate the variables shown. We use median, inter-quartile range (IQR) to describe the distribution of the differences. The differences of these variables between two simulations are all closely distributed around 0. Although there are occasions when the differences are quite large, the frequency for these occasions is very low, indicated by the small IQR. If we use half of the IQR to represent the average changes, then for radiative skin temperature, it is only 1°C, for top 10 cm soil moisture it is less than 10% and for 10–40 cm soil moisture is less than 10%. The surface fluxes are more temporally variable and so their changes are slightly larger, but still less than 20 W m^{-2}. Compared with the average summer midday fluxes, the changes are still small. The amplitude of change in the fluxes in each model is somewhat different and is related to the amplitude of the fluxes simulated by the models, i.e., Mosaic has the largest latent heat flux in the control run and so its changes in latent heat are also

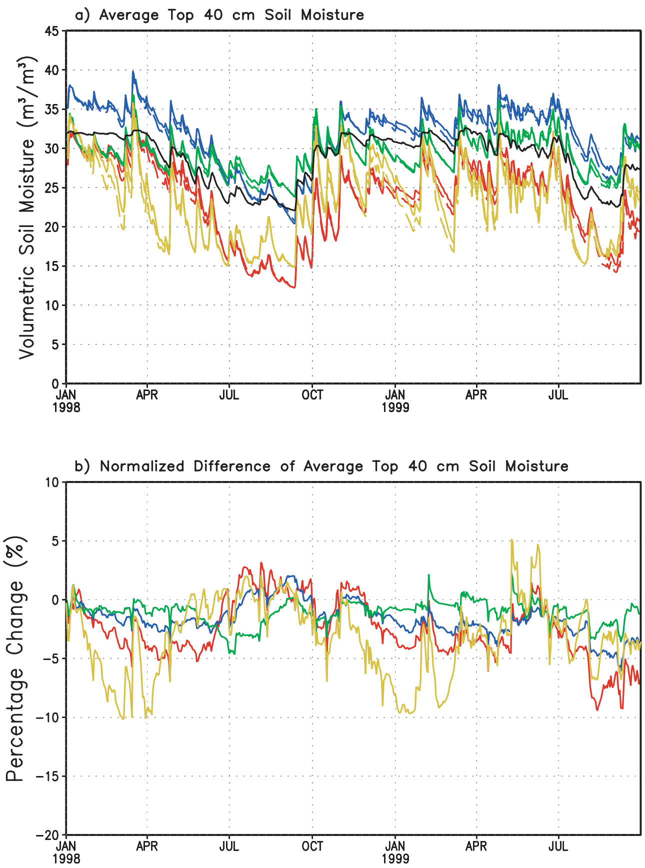

Figure 6. (a) Time series of the top 40 cm soil moisture from model simulations driven by NLDAS forcing (solid lines) and local forcing (dashed lines), averaged over 72 stations shown in Figure 1. The black line is the observations. Mosaic, Noah, VIC, and SAC are plotted in red, blue, green and gold respectively. (b) Differences between the models' two simulations (local forcing run – NLDAS control run) divided by NLDAS control run.

the largest. If we were to normalize the changes with respect to each model's climatology, the differences between the models would be smaller.

[27] In general, the differences in atmospheric forcing do not produce significant differences in modeled land surface conditions. However, this does not mean that forcing difference has no impact on model simulations. When a difference in forcing creates differences in some of the modeled land surface states at one time step, these differences will then be further modified by other model internal processes and propagate to other part of the system at the subsequent time steps. We now examine how a difference in precipitation forcing propagates in each model using an example from the 21-month simulation. Figure 8 shows the soil moisture differences induced by precipitation differences and how they change over time. During the 35-day period (1 June–6 July 1998), the Mesonet station at Buffalo, Oklahoma, received 5 precipitation events. However, the NLDAS forcing and station observed precipitation are slightly different both in timing and amount of precipitation (Figures 8a–8b). The models respond to the two forcing data sets accordingly, which

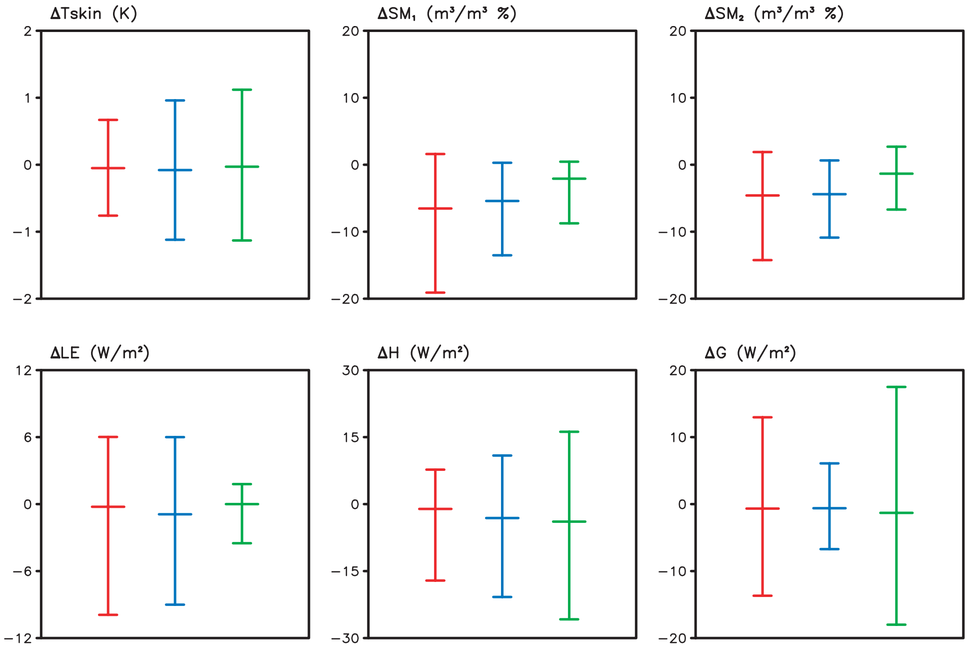

Figure 7. Differences between the control and local forcing runs (local forcing run – NLDAS control run) for the 72 Mesonet stations (Figure 1) for skin temperature (T_{skin}), top 10 cm volumetric soil moisture (SM_1), 10–40 cm volumetric soil moisture (SM_2), surface latent heat flux (LE), surface sensible heat flux (H), and ground heat flux (G). The samples include differences in hourly data at all 72 Mesonet stations. The sample size is 1,102,464. The median, upper quartile, and lower quartile of are plotted.

produces differences in their soil moisture simulations between two runs (Figures 8c–8f). Originally, the difference in top 40 cm soil moisture is virtually zero for all three models. When the precipitation comes on June 6 in the NLDAS forcing but not in the local forcing, the soil moisture in the control run increases, causing a negative difference in soil moisture between these two runs (local run minus control run). This difference then decreases in an e-folding fashion in the first layer (solid lines). On June 8, the station received much heavier rain in the local forcing than in the NLDAS forcing, which brought the soil moisture up significantly in the local run resulting in a positive difference in soil moisture between two simulations. After several days, however, the difference is rather small. This kind of behavior repeats whenever there is a difference in precipitation. Because the differences decrease over time, the soil moisture states are always very close in two simulations. Because NLDAS forcing and local forcing provide a very similar climate to the land surface, the subtle differences in precipitation and radiation on short timescales only generate differences over short timescales. Over a long period, they are not significantly different.

5. Discussion and Conclusions

[28] We have compared the NLDAS retrospective forcing with observations from Mesonet and ARM/CART stations during the period January 1998 to September 1999. We find a good agreement between the two data sets for all the meteorological variables except precipitation at the hourly timescale. Because the NLDAS precipitation is derived from observations, the comparison of station observations with the analyzed field gives a good agreement at longer timescales. Although differences exist between the two data sets, the differences are primarily due to the nature of spatial variation of the variables and possible sampling errors. However, a warm bias in GOES solar radiation and cool bias in EDAS downward longwave radiation consistently appear at all the grids studied. EDAS downward solar radiation also has a warm bias when it is used by the models in the morning when the sun is at a very low angle and GOES solar radiation is not available.

[29] Differences in precipitation exist between NLDAS forcing and local observations mostly at the hourly timescale. Most of the differences can be attributed to the natural characteristics of precipitation and the precipitation analysis scheme. The NLDAS forcing is able to capture almost 90% of the precipitation events. The amount of precipitation reported in the two data sets agrees within 10% for the 21-month period, although the difference can be very large for individual precipitation events, especially for small-scale convective precipitations. On the basis of these comparisons, we conclude that the two forcing data sets agree fairly well and NLDAS forcing is a good resource for land surface modeling applications in the SGP region.

[30] We used these two forcing data sets to drive four LSMs to quantify the differences in forcing based on their impacts on land surface model simulations. We found that three models all show relatively small differences in their land surface states and surface fluxes on day to month timescales. The differences in surface fluxes are more variable than those of state variables on short timescales, as expected. Although these changes are significant at

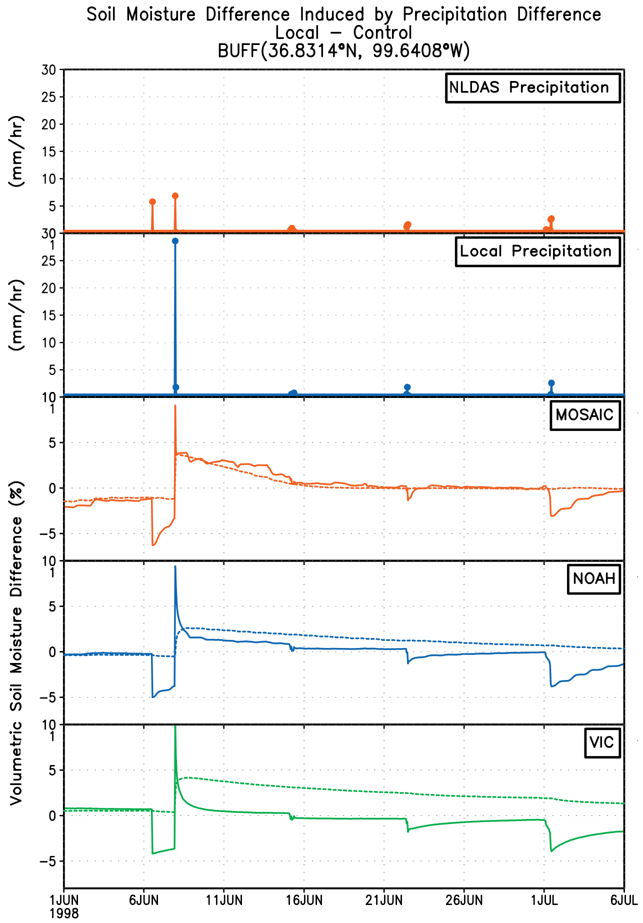

Figure 8. Soil moisture difference induced by precipitation difference between control and local simulations: NLDAS precipitation and local precipitation are shown. Differences in modeled first layer soil moisture (0–10 cm, solid lines) and second layer soil moisture (10–40 cm, dashed lines) are plotted for Mosaic, Noah and VIC in different panels.

certain periods at individual stations, the overall differences are not large at longer timescales and the differences are much smaller than the differences between the models and the observations. This supports our hypothesis that the differences between model output and station observations are not due primarily to differences in atmospheric forcing, but due to model physics.

[31] **Acknowledgments.** 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 NCEP/EMC, NWS/OHD, and NESDIS/ORA was supported by the NOAA OGP grant for the NOAA Core Project for GCIP/GAPP (co-PIs K. Mitchell, J. Schaake, J. Tarpley). The work by NASA/GSFC/HSB was supported by NASA's Terrestrial Hydrology Program (P. Houser, PI). 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). 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.

References

Anderson, E. A., *National Weather Service River Forecast System: Snow Accumulation and Ablation Model*, *Tech. Memo. NWS Hydro-17*, National Oceanic and Atmos. Admin., Silver Spring, MD, 1973.

Baldwin, M., and K. Mitchell, The NCEP hourly multi-sensor U. S. precipitation analysis for operations and GCIP research, paper presented at the 13th AMS Conference on Hydrology, Am. Meteorol. Soc., Long Beach, Calif., 1997.

Berbery, E., K. Mitchell, S. Benjamin, T. Smirnova, H. Ritchie, R. Hogue, and E. Radeva, Assessment of land-surface energy budgets from regional and global models, *J. Geophys. Res.*, *104*, 19,329–19,348, 1999.

Betts, A. K., J. H. Ball, and A. C. M. Beljaars, Comparison between the land surface response of the European Centre model and the FIFE-1987 data, *Q. J. R. Meteorol. Soc.*, *119*, 975–1001, 1993.

Betts, A., F. Chen, K. Mitchell, and Z. Janjic, 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, 1997.

Betts, A. K., P. Viterbo, and A. C. M. Beljaars, Comparison of the land surface interaction in the ECMWF reanalysis model with the 1987 FIFE data, *Mon. Weather Rev.*, *126*, 186–198, 1998.

Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. Johnson, and M. D. Eillts, The Oklahoma Mesonet: A technical overview, *J. Atmos. Oceanic Technol.*, *12*, 5–19, 1995.

Burnash, R. J. C., R. L. Ferral, and R. A. McGuire, *A Generalized Stream Flow Simulation System: Conceptual Models for Digital Computers*, Joint Fed. State River Forecast Cent., Sacramento, Calif., 1973.

Chen, F., and K. Mitchell, Using the GEWEX/ISLSCP forcing data to simulate global soil moisture fields and hydrological cycle for 1987–1988, *J. Meteorol. Soc. Japan*, *77*, 167–182, 1999.

Chen, F., Z. Janjic, and K. Mitchell, Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model, *Boundary Layer Meteorol.*, *85*, 391–421, 1997.

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

Cosgrove, B. A., et al., 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, in press, 2003.

Cressman, G. P., An operational objective analysis system, *Mon. Weather Rev.*, *87*, 367–374, 1959.

Dutton, E. G., J. J. Michalsky, T. Stoffel, B. W. Forgan, J. Hickey, D. W. Nelson, T. L. Alberta, and I. Reda, Measurement of broadband diffuse solar irradiance using current commercial instrumentation with a correction for thermal offset errors, *J. Atmos. Oceanic Technol.*, *18*, 297–306, 2001.

Ebert, E. E., U. Damrath, W. Wergen, and M. E. Baldwin, The WGNE assessment of short-term quantitative precipitation forecasts, *Bull. Am. Meteorol. Soc.*, *84*, 481–492, 2003.

Higgins, R. W., W. Shi, and E. Yarosh, Improved United States precipitation quality control system and analysis, *Atlas 7*, 40 pp., Climate Prediction Center, Nat. Cent. for Environ. Predict., Camp Springs, Md., 2000.

Hinkelman, L. M., T. P. Ackerman, and R. T. Marchaud, An evaluation of NCEP Eta model predictions of surface energy and cloud properties by comparison with measured ARM data, *J. Geophys. Res.*, *104*, 19,535–19,550, 1999.

Jackson, T. J., D. M. Le Vine, A. Y. Hsu, A. Oldak, P. J. Starks, C. T. Swift, J. D. Isham, and M. Haken, Soil moisture mapping at regional scales using microwave radiometry: The Southern Great Plains Hydrology Experiment, *IEEE Trans. Geosci. Remote Sens.*, *37*, 2136–2151, 1999.

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

Koster, R. D., and M. J. Suarez, Modeling the land surface boundary in climate models as a composite of independent vegetation stands, *J. Geophys. Res.*, *97*, 2697, 1992.

Koster, R. D., and M. J. Suarez, Energy and water balance calculations in the Mosaic LSM, *NASA Tech. Memo. 104606*, vol. 9, 1996.

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

Liang, X., D. P. Lettenmaier, and E. F. Wood, 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, 1996a.

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

Mitchell, K. E., The multi-institution North Americann Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *J. Geophys. Res.*, *108*(D22), 8841, doi:10.1029/2003JD003823, in press, 2003.

Pinker, R. T., et al., 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, in press, 2003.

Robock, A., C. A. Schlosser, K. Y. Vinnikov, N. A. Speranskaya, J. K. Entin, and S. Qiu, Evaluation of AMIP soil moisture simulations, *Global Planet. Change*, *19*, 181–208, 1998.

Robock, A., et al., 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, in press, 2003.

Sellers, P. J., F. G. Hall, G. Asrar, D. E. Strebel, and R. E. Murphy, An overview of the First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment, *J. Geophys. Res.*, *97*, 18,345–18,371, 1992.

Shafer, M. A., T. Hughes, and J. D. Carlson, The Oklahoma Mesonet: Site selection and layout, paper presented at Eighth Symposium on Meteorological Observations and Instrumentation, Am. Meteorol. Soc., Anaheim, Calif., 1993.

Srinivasan, G., A. Robock, J. K. Entin, L. Luo, K. Y. Vinnikov, and participating AMIP modeling groups, Soil moisture simulation in revised AMIP models, *J. Geophys. Res.*, *105*, 26,635–26,644, 2000.

Wilks, D. S., *Statistical Methods in the Atmospheric Sciences*, Academic, San Diego, Calif., 1995.

Wood, E. F., et al., The Project for Intercomparison of Land-surface Parameterization Schemes PILPS Phase 2c Red–Arkansas River basin experiment: 1. Experiment description and summary intercomparisons, *Global Planet. Change*, *19*, 115–135, 1998.

B. Cosgrove and P. R. Houser, Hydrological Sciences Branch, NASA Goddard Space Flight Center, Mail Code 974, Greenbelt, MD 20771, USA.

Q. Duan and J. C. Schaake, Office of Hydrologic Development, NOAA, National Weather Center, SSMC2, W/OHD12, 1325 East-West Highway, Silver Spring, MD 20910, USA.

R. W. Higgins, Climate Prediction Center, National Centers for Environmental Prediction, NOAA Science Center, 5200 Auth Road, Camp Springs, MD 20746, USA.

D. Lohmann and K. E. Mitchell, Environmental Modeling Center, National Centers for Environmental Prediction, NOAA Science Center, 5200 Auth Road, Camp Springs, MD 20746, USA.

L. Luo, J. Sheffield, F. Wen, and E. F. Wood, Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ 08544, USA.

R. T. Pinker, Department of Meteorology, University of Maryland, Space Sciences Building, College Park, MD 20742, USA.

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

J. D. Tarpley, Office of Research and Applications, National Environmental Satellite Data and Information Service, World Weather Building, Room 711, 5200 Auth Road, Camp Springs, MD 20746, USA.