

DEVELOPMENT OF A HYDROMETEOROLOGICAL FORCING DATA SET FOR GLOBAL SOIL MOISTURE ESTIMATION

A. A. BERG,^{a,*} J. S. FAMIGLIETTI,^b M. RODELL,^c R. H. REICHLER,^{d,e} U. JAMBOR,^c S. L. HOLL^b and P. R. HOUSER^c

^a *Department of Geography, University of Guelph, Guelph, Ontario, Canada*

^b *Department of Earth System Science, University of California, Irvine, Irvine, California, USA*

^c *Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA*

^d *Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore, Maryland, USA*

^e *Global Modeling and Assimilation Office NASA/Goddard Space Flight Center, Greenbelt Maryland, USA*

Received 15 June 2004

Revised 24 December 2004

Accepted 29 March 2005

ABSTRACT

Off-line land surface modeling simulations require accurate meteorological forcing with consistent spatial and temporal resolutions. Although reanalysis products present an attractive data source for these types of applications, bias to many of the reanalysis fields limits their use for hydrological modeling. In this study, we develop a global 0.5° forcing data sets for the time period 1979–1993 on a 6-hourly time step through application of a bias correction scheme to reanalysis products. We then use this forcing data to drive a land surface model for global estimation of soil moisture and other hydrological states and fluxes. The simulated soil moisture estimates are compared to *in situ* measurements, satellite observations and to a modeled data set of root zone soil moisture produced within a separate land surface model, using a different data set of hydrometeorological forcing. In general, there is good agreement between anomalies in modeled and observed (*in situ*) root zone soil moisture. Similarly, for the surface soil wetness state, modeled estimates and satellite observations are in general statistical agreement; however, correlations decline with increasing vegetation amount. Comparisons to a modeled data set of soil moisture also demonstrates that both simulations present estimates that are well correlated for the soil moisture in the anomaly time series, despite being derived from different land surface models, using different data sources for meteorological forcing, and with different specifications of the land surfaces properties. Copyright © 2005 Royal Meteorological Society.

KEY WORDS: soil moisture; hydrometeorological forcing; land surface modeling; reanalysis

1. INTRODUCTION AND BACKGROUND

Water held in the soil is critical to the hydrologic cycle. It is an important source of moisture for evapotranspiration, it affects the physiology and distribution of vegetation, it recharges groundwater, controls runoff distributions, and it is essential for the partitioning of the land surface heat flux between sensible and latent heat components. In numerical climate simulations, numerous studies have demonstrated that soil wetness is an important atmospheric boundary condition affecting the prediction of precipitation (Fennessy and Shukla, 1999; Dirmeyer, 2000; Koster *et al.*, 2000) and temperature (Koster and Suarez, 2003).

A complicating issue for studies that examine the impact of the initial soil moisture state on climate prediction is obtaining realistic estimates from continental to global scales. To produce these estimates, comprehensive sets of land surface and atmospheric information are necessary at consistent temporal and spatial scales. Over many regions of the Earth, such information is only available from global reanalysis products such as the National Center for Environmental Prediction/National Center for Atmospheric Research

* Correspondence to: A. A. Berg, Department of Geography, University of Guelph, Guelph, Ontario, Canada N1G2W1; e-mail: aberg@uoguelph.ca

(NCEP/NCAR) 40-year reanalysis (Kalnay *et al.*, 1996) (NRA), and the European Center for Medium Range Weather Forecasts (ECMWF) (Gibson *et al.*, 1997) (ERA-15) and ERA-40 (Simmons and Gibson, 2000).

Soil moisture and other hydrologic stocks are integrated states; thus, systematic bias in forcing products is manifested in water storage estimates. Therefore, the use of meteorological forcing derived directly from reanalysis products for driving uncoupled land surface model simulations is not recommended because of bias in the reanalysis products. Bias in various reanalysis fields in the ERA and NRA has been the subject of much research (e.g. Betts *et al.*, 1996, 1998a, 1998b, 1998c; Maurer *et al.*, 2001a, 2001b; Roads and Betts, 2000, Roads *et al.*, 2003), and for this reason, studies that have incorporated reanalysis data sources have also incorporated some degree of bias correction to the reanalysis fields (Dirmeyer *et al.*, 1999; Arora *et al.*, 2000; Lenters *et al.*, 2000).

The reanalysis products also include global estimates of the soil moisture state. However, a number of studies have identified problems with methods implemented by the NRA and ERA-15 for soil water estimation (e.g. Roads and Betts, 2000; Maurer *et al.*, 2001a, 2001b). The NRA included an adjustment of the soil water cycle toward the soil moisture climatology of Mintz and Serafini (1992). While this procedure prevents the NRA from maintaining unrealistic soil moisture conditions, the values simulated are not always representative of actual soil moisture state or tendency (Roads *et al.*, 1999). Moreover, Robock *et al.* (1998) identify problems and inconsistencies with the Mintz and Serafini (1992) climatology. In the ERA-15, a tendency for the modeled soil moisture to become too dry was eliminated by systematically adding water to each of the soil layers (Betts *et al.*, 1999). Therefore, soil water properties derived from either of these sources, climatology or initial conditions, are not likely to be representative of the true soil water state (e.g. Srinivasan *et al.*, 2000).

Given the above problems with reanalysis meteorological forcing and soil moisture, the Global Soil Wetness Project (GSWP) (Dirmeyer *et al.*, 1999) and Nijssen *et al.* (2001) have produced soil moisture data sets in addition to those from the global reanalysis products. In the GSWP phase I, 11 different land surface models (LSMs) were used to produce retrospective estimates of hydrological fluxes and the soil water state for the years 1987 and 1988. The necessary meteorological forcing for this study was derived from a synthesis of observations and the ERA-15 (Sellers *et al.*, 1996). In the second phase of the GSWP, a hydrometeorological forcing data set was developed for over 10 years by hybridizing the NCEP/DOE (Department of Energy) reanalysis (Kanamitsu, *et al.*, 2002) with global observations (Zhao and Dirmeyer, 2003). Estimates of initial soil water field derived from the GSWP phase I have been used in a number of climate model studies (e.g. Dirmeyer, 2000; Douville and Chavin, 2000) and hydrologic analyses (e.g. Rodell and Famiglietti, 1999).

The study of Nijssen *et al.* (2001) expands on the work of GSWP phase I by providing estimates of the global root zone soil moisture storage for a 14-year period (1980–1993). To avoid previously identified errors with reanalysis forcing (e.g. Maurer *et al.*, 2000), an observation-based forcing product independent from the reanalysis products (except for wind speed) was developed. Grids of daily precipitation and temperature observations were produced using station observations together with monthly average temperature (Jones, 1994) and precipitation (Huffman *et al.*, 1997) products. Other necessary forcing variables (e.g. radiation, vapor pressure) were derived from relationships between the temperature and precipitation time series (e.g. Thornton and Running, 1999). Using this meteorological forcing data, the variable infiltration capacity (VIC) (Liang *et al.*, 1994) land surface model was run to produce a long-term daily data set of hydrological states and fluxes.

Recognizing the existence of errors in the reanalysis products and their potential to impact off-line land surface model simulations, Berg *et al.* (2003) created a forcing data set for North America through bias correction of the ERA and NRA reanalysis products. This bias-corrected reanalysis-based forcing was then used to simulate hydrologic fluxes over the continent and to demonstrate the sensitivity of a land surface model to the bias removal.

In the present study, we expand upon the techniques and methods presented in Berg *et al.* (2003) to develop a 0.5° hydrometeorological forcing data set for the period 1979–1993 for global land surface modeling applications. The techniques described in the Berg *et al.* (2003) study are adapted to enhance the spatial resolution of the precipitation forcing, and also to improve the radiation estimates for periods outside of the observed record. We then use this forcing data to drive a land surface parameterization scheme for global

estimation of soil moisture and the other hydrological states and fluxes. Finally, we compare our modeled soil moisture states to observations from the soil moisture data bank (Robock *et al.*, 2000) to a satellite-derived surface soil moisture product (Owe *et al.*, 2001), and to the modeled soil moisture product of Nijssen *et al.* (2001). Throughout this study, we limit our focus to descriptions of the soil water state. Analysis of other aspects of the hydrological cycle is the subject of ongoing research.

2. MODEL AND FORCING DATA

In this section, we discuss the application of a bias correction method similar to that of Berg *et al.* (2003) to the ERA-15 (Gibson *et al.*, 1997) and to the NRA-40 (Kalnay *et al.*, 1996) to produce a 0.5°, 6-hourly forcing data set for the global land surface over the period 1979–1993. Recently, the ERA-15 has been updated to the ERA-40 (Simmons and Gibson, 2000). The ERA-40 has numerous improvements to the ERA-15 including higher spatial and temporal resolutions, an updated land surface parameterization scheme and improvements to many of the errors and the bias identified in the ERA-15 (e.g. Betts *et al.*, 1998 a, 1998 b, 1998 c). Work by Seneviratne *et al.* (2004) demonstrates that the improvements to the ERA-40 allow for more accurate simulation of hydrological features such the seasonal soil moisture cycle and measures of interannual variability over the Mississippi basin. However, similar comparisons over the Mackenzie basin by CT Betts *et al.* (2003) demonstrate that precipitation bias remains an issue over some regions in the ERA-40. Bias has also been identified in the update to the NCEP/NCAR reanalysis, the NCEP-DOE reanalysis (e.g. Maurer *et al.*, 2001a; Zhao and Dirmeyer, 2003), and thus the hydrometeorological forcing data set created for the GSWP Phase II (Zhao and Dirmeyer, 2003) performs a hybridization strategy where global observation products are merged with NCEP-DOE reanalysis data. Therefore, bias reduction procedures such as those described in Berg *et al.* (2003) and in the discussion below will also be an appropriate strategy for the reduction of bias in the updated reanalysis products for producing forcing for off-line simulations of land surface models. Discussion in this section will focus on (1) updates to the bias correction described in the Berg *et al.* (2003) study, (2) the modeling strategy, and (3) data sets used for comparison with simulation results.

2.1. Bias correction and forcing data set development

The required climatic forcing for the majority of land surface schemes includes short- and longwave downwelling radiation, precipitation, air and dew point temperatures, wind speed, and surface pressure. The data products used in the correction procedure are summarized in Table I.

2.1.1. Downwelling short- and longwave radiation. Correction to the reanalysis (ERA and NRA) downwelling of long- and shortwave radiation was completed following a ratio approach (Berg *et al.*, 2003) to match data from the Surface Radiation Budget (SRB) (Gupta *et al.*, 1999) following:

$$cR_i = \frac{Obs}{R} \times R_i \quad (1)$$

where, cR_i is the corrected reanalysis forcing for time i , Obs are the monthly SRB observations, R is the raw monthly average radiation from the reanalysis, and R_i is the raw reanalysis also at time i .

Unfortunately, the SRB product was available only for the period 07/1983–06/1991. Therefore we expand upon the Berg *et al.* (2003) study to correct the reanalysis products outside of this 8-year period by correcting the ratio between the observations and reanalysis products (Obs/R in Equation 1) using a regression based approach. Assuming that differences between observations and the reanalysis are systematic, then removal of bias from the reanalysis products using regression equations between monthly observations and the reanalyses is valid. To this end, we calculate monthly anomalies for both the reanalysis products and the SRB monthly data for the period July 1983–June 1991. A regression analysis was then performed between the SRB and reanalysis anomalies for each grid cell; the statistical significance of each equation was evaluated using

Table I. Observational data sets used to perform bias correction to the ERA and NRA (the time frames given here reflect those used in the bias correction and are not necessarily those of the actual data set)

| Data sets used for bias correction of the reanalysis | | | |
|--|---|-------------------|-------------|
| Reanalysis field | Observed data set used | Resolution | |
| | | Temporal | Spatial |
| 2 m air temperature | Climate Research Unit (<i>New et al.</i> , 2000) | Monthly 1979–1993 | 0.5° × 0.5° |
| | Center for Climatic Research (<i>Willmott and Matsuura</i> , 2001) | Monthly 1979–1993 | 0.5° × 0.5° |
| 2 m dew point temperatures | Climate Research Unit (<i>New et al.</i> , 2000) | Monthly 1979–1993 | 0.5° × 0.5° |
| Precipitation | Global Precipitation Climatology Project Version 2 (<i>Adler et al.</i> , 2003) | Monthly 1979–1993 | 2.5° × 2.5° |
| | Climate Research Unit (<i>New et al.</i> , 2000) | Monthly 1979–1993 | 0.5° × 0.5° |
| | Center for Climatic Research (<i>Willmott and Matsuura</i> , 2001) | Monthly 1979–1993 | 0.5° × 0.5° |
| Long- and Shortwave radiation | Langley Eight Year Shortwave and Longwave Surface Radiation Budget CT (<i>Gupta et al.</i> , 2000) | Monthly 1983–1991 | 1.0° × 1.0° |

a Student's *t* test assuming a 0.05 confidence level ($p < 0.05$). Where the relationship between SRB and reanalysis anomalies was found to be statistically significant, the regression equation was then used to forecast the anomaly departure from the mean monthly SRB value. The anomaly forecast from the regression was added to the monthly mean SRB value to evaluate the value of *Obs* in Equation 1.

To ensure that any seasonal differences to the relationship between the reanalysis and SRB data are represented, each regression equation was evaluated separately over each of December–February (DJF), March–May (MAM), June–August (JJA), and September–November (SON). After the value of *Obs* was determined, the corrected value of the reanalysis radiation, both for long- and shortwave, was evaluated according to Equation 1. Overall, the regression equations were found to be statistically significant over much of the global land surface in both the reanalysis products. Table II documents the applicability of the regression equation approach for each period for correction of both the ERA and NRA.

For grid cells without a statistically significant regression relationship, a standardized anomaly calculated from the reanalysis relative to period July 1983–June 1991 was used to adjust the SRB mean (*Obs*) relative to the mean and standard deviation of the SRB observations. This procedure was repeated for both the shortwave and longwave radiation data for the time periods outside of July 1983–June 1991 on the monthly timescale.

The goal of the bias correction for the radiation data sets was to ensure that the monthly mean matches that of the SRB. However, by making this correction the diurnal cycle of radiation as forecast by the reanalysis has been impacted. At the time the radiation data sets were processed, data sets for the diurnal range were not available and thus were not included in the bias correction. Therefore, updates to this procedure must make use of these updated data sets and constrain the application of Equation 1 to within the observed diurnal range (e.g. Zhao and Dirmeyer, 2003). Despite this issue, sensitivity experiments that compare between off-line land surface model simulations using the corrected radiation fields and the raw ERA-15 reanalysis data demonstrate

Table II. Percentage of the global land area over which regression equations between anomalies to the SRB and the reanalysis products were found to be statistically significant

| Month | ERA | | NRA | |
|-------|-----------|----------|-----------|----------|
| | Shortwave | Longwave | Shortwave | Longwave |
| DJF | 0.71 | 0.94 | 0.68 | 0.78 |
| MAM | 0.99 | 0.99 | 0.99 | 0.99 |
| JJA | 0.99 | 0.99 | 0.71 | 0.98 |
| SON | 0.99 | 0.99 | 0.99 | 0.99 |

that the correction procedure results in the improved simulation of runoff and snow depth, particularly over midlatitudes drainage basins (Holl, 2004).

2.1.2. Air and dew point temperature. Production of the 2 m air and dew point temperature data closely follows Berg *et al.* (2003). For air temperature, the two reanalysis products (ERA-15 and NRA-40) were corrected to an average obtained from the products produced by CT New *et al.* (2000) and Willmott and Matsuura (2001). The globally averaged mean absolute monthly temperature difference between the two temperature observation data sets was found to be slightly over 1°, and because we cannot make an argument toward which data set is preferable we have decided to average the two data sets in this analysis. The key distinction between the methods used in this study and that of Berg *et al.* (2003) was interpolation to 0.5° grids (using inverse distance squares) rather than to a catchment-based resolution.

2.1.3. Precipitation. Precipitation was corrected in the Berg *et al.* (2003) study according to Equation 1, using the Global Precipitation Climatology Project GPCP version 2 observations (Adler *et al.*, 2003). For this analysis, we use a similar approach but only disaggregate the GPCP totals spatially. The monthly precipitation rate over each of the GPCP 2.5° × 2.5° grids was spatially distributed using the 0.5° degree data sets of Climate Research Unit (CRU) (CT New *et al.*, 2000) and the Center for Climatic Research (CCR) Monthly Precipitation time series (version 1.02) (Willmott and Matsuura, 2001). To spatially distribute the GPCP precipitation rates, first, the average of the CRU and CCR estimates were determined at the 0.5° resolution. An average between the two data sets was taken for reasons identical to those discussed above for temperature (the global mean absolute difference between the data sets was found to be approximately 91 mm (Fekete *et al.*, 2004)). From this combined 0.5° precipitation data set, the average monthly precipitation rate was determined for a grid size identical to that of the GPCP grid (2.5°). In the next step, the ratio of the GPCP estimate to the combined CRU and CCR estimate at 2.5° was multiplied by each of the combined 0.5° grid estimates to produce a precipitation data set where the total precipitation rate of each 2.5° grid cell was similar to the GPCP estimate; however, the spatial distribution of rainfall within that 2.5° grid is determined by the estimates of the CRU and CCR. In the final step, the reanalysis products were interpolated to a 0.5° grid using inverse distance squares and a ratio-based correction was applied to reanalysis following Equation 1 where *Obs* is the combined GPCP, CRU, and CCR precipitation estimate.

To examine how well this method improves the representation of precipitation, in Figure 1 we compare correlations between corrected data and the raw reanalysis with the Climate Prediction Center 0.25 × 0.25 Daily US Unified Precipitation data set (Higgins *et al.*, 2000) averaged to the 0.5° resolution used in this study. Comparisons were completed for the daily, pentad (5 day) and monthly time periods. Unfortunately, few data sets are available for global comparisons on the daily time frame; therefore, we limit this comparison to the continental United States where reliable daily gauge-based precipitation products are available. The Higgins *et al.* (2000) data set is available daily from 1948 to 1998 at a spatial resolution of 0.25 × 0.25° for the continental United States. In Figure 1, we illustrate correlations between the gauge-based product and the raw reanalysis products (NRA, ERA; right column) and the bias-corrected (C) data (CNRA and CERA; left column) for each of the three time periods. In general, correlations between the raw reanalysis products

and Higgins *et al.* (2000) are best over the western coast of the continental United States, and improve with increasing amount of time considered (monthly *vs* daily). The well documented problem with precipitation in the NRA over the southeast of the United States (e.g. Roads and Betts, 2000) is evident in correlations between the gauge-based data set and the NRA for time periods greater than one day. As illustrated in Figure 1, the corrected products have higher correlations and lower root mean square errors (RMSE) and mean bias errors (MBE) when compared to the gauge-based product over all three time periods. A MBE calculated over the entire domain was found to be positive for all cases indicating that both the reanalysis and the corrected forcing are wetter than observations; however, in all cases the MBE falls very close to zero (not shown). Although the correction is applied only at the monthly time frame, improvements to the correlations between the gauge and corrected reanalysis product are observable particularly at the pentad time frame (improvement observed at the monthly scale represent the correlation between corrected data set and Higgins *et al.*, data (2000)). For both the corrected reanalysis products (CERA and CNRA), improvements to the correlations over the monthly time frame are best over the central and eastern United States. Over the Rocky Mountains correlations between Higgins *et al.* (2000) and the corrected reanalysis products are much lower than elsewhere in the United States. Over regions of high topographic complexity, approaches similar to that of Daly *et al.* (1994) maybe appropriate for further bias reduction.

2.1.4. Unmodified reanalysis fields. Identical to the Berg *et al.* (2003) study, no modifications were made to reanalysis wind speed. For surface pressure, the only correction made to the reanalysis product involves adjusting the surface pressure from the reanalysis surface height to the elevation of the 0.5° grid (Row *et al.*, 1995) following the hypsometric equation.

2.2. Modeling strategy

In this study, we utilize the Mosaic land surface modeling scheme (Koster and Suarez, 1992, 1996). Mosaic is a well-established land surface modeling scheme and a participant in a number of intercomparison studies including the Project for Intercomparison of Landsurface Parameterization Schemes (PILPS), the atmospheric model intercomparison project (for a recent discussion see Henderson–Sellers *et al.*, 2003) the GSWP (Dirmeyer *et al.*, 1999), the North American Land Data Assimilation System (NLDAS) (Mitchell *et al.*, 2004), and is one of the land surface models included in the GLDAS (Rodell *et al.*, 2004). Robock *et al.* (2003) and Schaake *et al.* (2004) present recent evaluations of energy and water balance of the Mosaic land surface model over the NLDAS study area.

An important attribute of the Mosaic model is its parameterization of subgrid scale variability by partitioning the land surface into a series of tiles based on vegetation type. The required land surface parameters, soil hydrologic properties and depths, vegetation type and leaf area index, and their specification within the Mosaic model are described in Rodell *et al.* (2004).

The modeling results discussed in this paper were performed as part of a retrospective simulation produced as part of the GLDAS. The modeling resolution for this simulation was fixed to 2° North/South × 2.5° East/West. Unfortunately, this required that the forcing fields also be scaled-up from the 0.5° × 0.5° to the modeling resolution through a simple arithmetic average. Therefore, we limit further discussion to features observed at the coarser modeled resolution. A separate simulation at the resolution of the forcing product (0.5°) is underway as part of the GLDAS (see Rodell *et al.*, 2004). The reader should also note that a recent study by Reichle *et al.* (2004) has produced similar results to those reported here using the described forcing product near the 0.5° resolution in a different land parameterization scheme.

Model spin-up was performed by driving the land surface model repetitively with the forcing produced for 1979 over a 45 year spin-up period; the extended spin-up period was completed as part of a separate analysis and is not critical for the present analysis. For the root zone and profile soil moisture analysis, simulations for 1979 were discarded for subsequent study to reduce the effects of model spin-up on the analysis. For comparisons with satellite-based observations, surface soil moisture observations, over the top 2 cm, from 1979 were included in the analysis because we expected little spin-up effects on surface fluctuations.

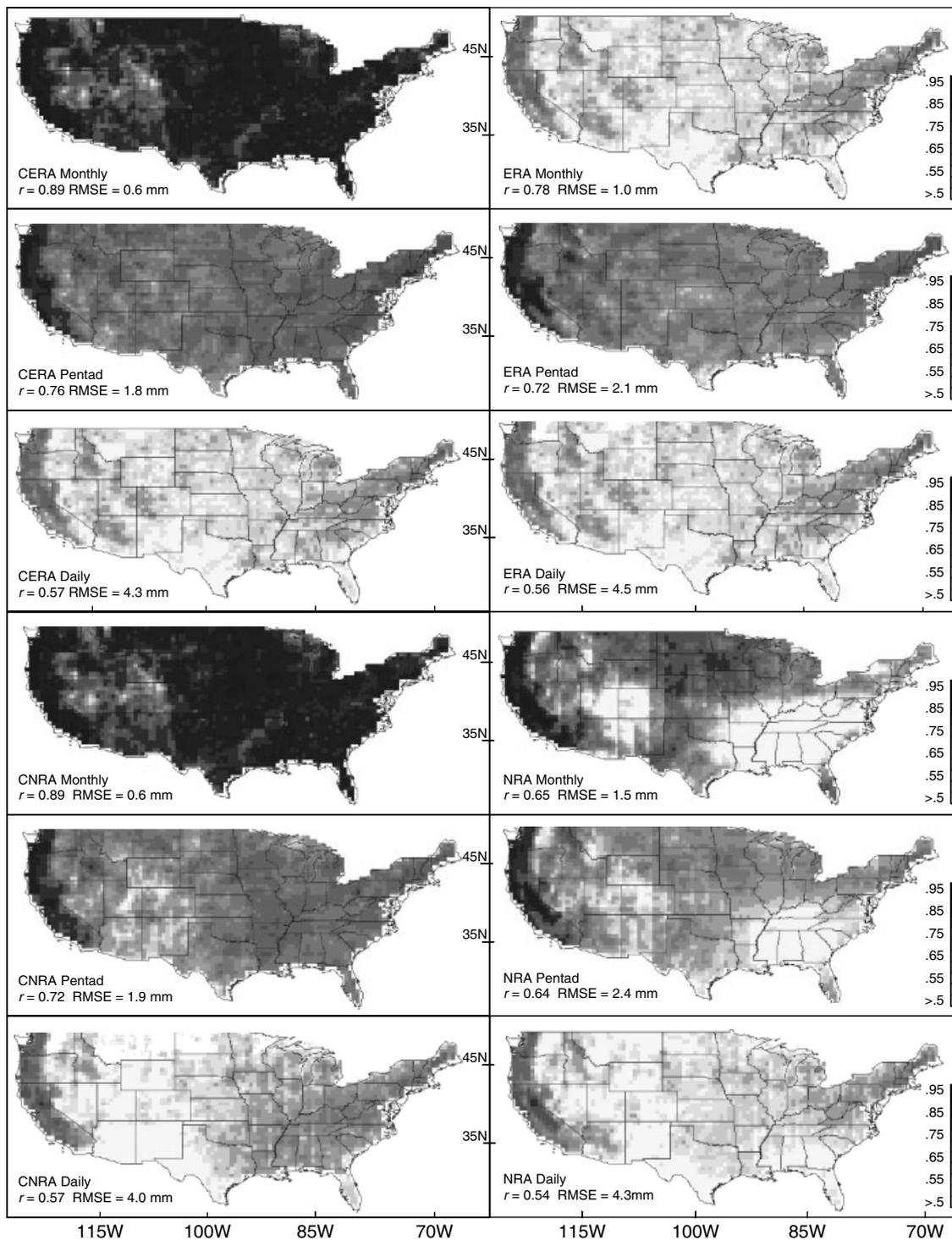


Figure 1. Correlations between the corrected ERA (CERA), the raw ERA, the corrected NRA (CNRA), and the raw NRA with the Higgins *et al.* (2000) daily gauge-based precipitation for the period 1985–1993

2.3. Data sets used for comparisons with simulation results

2.3.1. In situ soil moisture observations. For validation of modeled soil moisture, we use soil moisture observations obtained from the Soil Moisture Data Bank (SMDB) (Robock *et al.*, 2000). From the SMDB, we compare our simulated 1 m soil moisture estimates with observations (also from 1 m depth) from Iowa, Illinois, Russia, Mongolia, and China. To compare the SMDB observations with the Mosaic simulation, we interpolated the observed data to the identical grid resolution used in Mosaic. The search radius was set at 2° from each grid node, and the interpolation method was inverse distance squared. For each site, we record data only if observations are available from two or more soil moisture stations within the search window. While this approach does not consider the affect of covariance between the observations sites (e.g. Entin *et al.*, 1999; Vinnikov *et al.*, 1999), the size of the search radius window is conservative to spatial scales of soil moisture covariance dominated by the atmosphere (large-scale patterns of precipitation and evaporation) as identified in Entin *et al.* (2000).

The results of numerous PILPS studies and those of Entin *et al.* (1999) demonstrate that modeled mean soil moisture states will vary dramatically between observations and between different models because of different parameterizations schemes even if physical parameters such as soil texture and vegetation type are kept constant. Despite this issue, different models should still adequately capture the phase and amplitude of the observed soil moisture cycle. To address some of the discrepancies between observed and simulated soil moisture cycles, Maurer *et al.* (2000) proposed normalizing the soil moisture observations and modeling results by subtracting from each observation, or modeled value, the minimum data value during the selected period. This procedure, which was also described and implemented as part of the Nijssen *et al.* (2001) study, is adopted for the display of modeled and observed soil moisture in this study. We present the results of this analysis in Section 3.1.

For calculating correlations between the modeled and simulated soil moisture, we attempt to remove the seasonal cycle by converting both modeled and observed data to anomaly departures from their seasonal mean state (DJF, MAM, JJA, and SON).

2.3.2. Satellite soil moisture observations. We compared modeled estimates of the surface soil water content (top 2 cm in Mosaic) with observations from Owe *et al.* (2001) and de Jeu (2003). They describe the production of a surface soil moisture data set derived from Scanning Multichannel Microwave Radiometer (SMMR) satellite observations (1978–1987) (similar data sets can be derived using the current AMSR-E observations (Njoku and Li, 1999)). The soil moisture product is available for the global land surface with a resolution of 140 km although the final product was interpolated to a 1/4° grid (de Jeu, 2003). The depth of the SMMR-derived observations are dependent upon the thermal sampling depth which at 6.6 GHz is typically less than 2 cm (Schmugge, 1983). No data is observed over regions where the retrieval of soil moisture is impacted by high vegetation water contents, or by the presence of ice, snow, and frozen soils. We have completed further screening of the Owe *et al.* (2001) data set by removing observations over regions with high vegetation optical depth, and over regions where the Mosaic land surface model simulation predicts snow cover.

To compare the Mosaic simulations completed with the corrected forcing data set, the SMMR-based soil moisture observations were scaled-up to a resolution that corresponds to the Mosaic grid (2.5°E–W × 2°N–S). The scaling-up procedure is necessary for two reasons; firstly, it was completed to match resolutions between the satellite observations and modeled soil moisture and secondly for reducing the noise to the instantaneous SMMR-based soil moisture retrievals (see Reichle *et al.*, 2004). The scaling-up procedure uses a quadrant approach for determining the average moisture content of the larger Mosaic grid cell. Briefly, this approach derives average soil moisture content for the Mosaic grid on the basis of information from each of four quadrants. The average for the larger grid cell is then determined on the basis of the four subquadrants. Where one quadrant is not represented, because of no available data, the average for the entire cell was not considered. We adopted the quadrant approach to prevent bias due to unequal satellite coverage, where a satellite overpass retrieves information for half of the larger grid, and because of a sampling bias, where part of the larger grid cell is sampled at a higher rate because of more optimal soil moisture retrieval conditions. Results of this comparison are presented in Section 3.2.

In a related paper, Reichle *et al.* (2004) compare modeled and the SMMR-derived soil moisture to the SMDB observations. An important result of this study demonstrates that the satellite-derived soil moisture does not agree well with absolute soil moisture values obtained from the SMDB. However, with respect to the time series and anomaly correlation, agreements between the data sets (SMMR-based and SMDB) were found to be statistically significant. Therefore, in the present study we limit the majority of our analysis to the congruent features between the SMDB and SMMR data sets, namely, the time series correlation.

2.3.3. Intermodel comparison. We compare our modeled soil moisture results with those of Nijssen *et al.* (2001) (hereafter also referred to as the VIC simulation). Through comparison with the VIC simulation results we intend to highlight key similarities and differences between these soil moisture data sets. Numerous studies completed as part of GSWP (e.g. Entin *et al.*, 1999) and in the NLDAS (e.g. Schaake *et al.*, 2004) clearly demonstrate that despite tight control over the forcing and other parameters (e.g. soil type and vegetation cover) the mean soil moisture produced in varying LSMs are dramatically different. Therefore, we should expect that any comparison between two different models where the forcing and other modeling parameters are not identical will vary significantly for the mean soil moisture states. Despite these concerns, we hope that the different simulations should compare well for other soil water features, namely, the correlation between standardized anomaly departures from the modeled mean monthly soil moisture state, and the range or amplitude of the annual soil moisture cycle. Therefore, in this study we will limit our comparison between the VIC and Mosaic simulations to these features.

We examined differences between the amplitude of seasonal soil moisture cycle in both models by subtracting the maximum root zone soil water content (top 1 m in both models) by the minimum value for each year over the 14-year period (1980–1993). We evaluate the statistical significance of the differences observed using the nonparametric Wilcoxon signed-ranks test. Nonparametric statistical tests were used because of the limited size of the data set (14 years).

To simplify comparisons between the two data sets, we interpolated the VIC simulations from the 2° grids that were used in Nijssen *et al.* (2001) study to match the 2° × 2.5° resolution used in this study using a simple areal weighting approach.

3. VALIDATION OF THE MOSAIC SOIL MOISTURE SIMULATIONS

For the evaluation of the Mosaic soil moisture simulations, we limit our comparisons to the soil moisture simulations produced from the bias-corrected ERA (CERA) simulations (1980–1993). Although there are some important differences between the CERA and CNRA simulations (e.g. Berg and Famiglietti, 2003) for the general features of the soil water cycle described above (e.g. anomaly correlation and amplitude of the yearly cycle), the CERA and CNRA simulations are similar. Therefore, we chose to limit the focus of this section to comparisons only between the CERA simulation and the observed data. The sensitivity of the bias correction for modeling components of the water budget was included as part of the Berg *et al.* (2003) study, where it was shown to dramatically improve hydrologic simulation, and is not repeated here.

3.1. Comparisons to soil moisture data bank observations

Correlations between simulated and observed soil moisture at 1-m depth are displayed in Figure 2(a) (also see, Entin *et al.* (1999) and Nijssen *et al.* (2001)). As discussed above, for both the modeled and simulated data, the seasonal cycle was removed by converting each estimate and data value to anomaly departures from their seasonal mean state (DJF, MAM, JJA, SON) prior to calculation of the correlation statistic (seasons are used instead of months because of a lack of data at a number of points over some months). The black dots in Figure 2(a) represent the soil moisture stations of the SMDB, and the shaded circles represent correlations between the modeled and observed anomaly departures from the mean seasonal state (observations interpolated to the Mosaic grid center). Crosses through a shaded data point identify locations where the correlations shown are not statistically significant ($p < 0.05$), and open circles identify

locations where observations are available for 8 months or more of a given year. The mean level of correlation between simulated and observed root zone soil moisture (top 1 m) shown in Figure 2(a) was 0.38, with statistically significant correlations observed at more than 60% of stations shown. In general, correlations between modeled soil moisture and observations are marginally higher over regions with wetter mean soil moisture conditions ($r = 0.40$, $p < 0.01$), including the regions of North America, western Russia, and eastern China. Weak relationships between modeled and observed soil moisture data occur over the dry regions of western Mongolia and Russia north of Mongolian–Russian boundary where yearly precipitation totals less than 40 cm, and over regions where few observations are available. Over all the sites shown in Figure 2(a), the Mosaic simulation is slightly drier than observations (MBE is -0.5 mm) with a root mean square error (RMSE) of 5 cm.

In Figure 2(b), line graphs of the observed and modeled mean soil moisture seasonal cycle are plotted for a number of the sites numbered in Figure 2(a). The outer tick marks surrounding the observation data points depict one standard deviation estimate from the mean. We have screened the sites selected for further analysis to contain available soil moisture data for eight or more months of the year (open circles as identified in Figure 2(a)) with low frequencies of missing monthly data (which allowed for a more reliable estimate of the mean and standard deviation). As previously addressed, the data presented in Figure 2(b) was ‘normalized’ by subtracting the minimum water content to show the active soil water range for both modeled and observed data.

In Figure 2(b.1–3) observations from Iowa and Illinois are shown together with simulated soil moisture estimates. In Figure 2(b.1) (Iowa) and 2(b.2) (northern Illinois), the modeled soil moisture cycle corresponds very closely to the observed cycle. In Figure 2(b.3) (over southern Illinois), the amplitude of the observed cycle is well simulated despite the presence of low bias to the normalized soil moisture.

For a number of the Russian sites, especially sites north of 53° , and specifically at sites 4 and 10–14 the vegetation type specified in the Mosaic simulation does not match the vegetation type from which observations were obtained. This discrepancy may explain some of the inconsistencies between observed and simulated soil moisture for various locations in Russia and China in Figures 2(b).4–16. Over sites 5, 8–10, 15, and 16, the range of soil moisture cycle (top 1 m) is well simulated, although location 8 exhibits a wet bias to the normalized soil moisture volume. At sites 4, 6, 7, and 11–14, there are some inconsistencies between the modeled and observed cycle especially apparent during the late winter and spring (sites 4, 7, 11, 12–14). After this period, the observed cycles are better approximated in the simulation. At the very dry sites over central Russia (sites 13 and 14) there is little range in the observed soil moisture cycle, while in the simulation there is a wet bias to the soil moisture estimates. The simulated soil moisture over eastern China (sites 15 and 16) illustrates better correspondence to the observed mean annual cycle.

3.2. Comparisons to satellite observations

In Figure 3, we present correlations between the Mosaic surface soil moistures and Owe *et al.* (2001) SMMR-based soil moisture observations taken during the evening overpass (near midnight in local time) 1979–1987. Correlations over the regions in grey scale in Figure 3 are statically significant ($p < 0.05$), while the regions blacked out are not. Over the regions in the lightest shade of grey the correlations were not shown because the standard deviation in soil moisture values is less than the approximate noise level in the SMMR observations (Reichle *et al.*, 2004). For observable locations on the Earth’s surface, the mean level of correlation between the SMMR observations and the Mosaic simulation was 0.43 and 0.56 for daytime (not shown) and nighttime observations respectively. Overall, the results of Figure 3 are encouraging because they represent correlations between daily observations and modeled output for the same 6-h period. Therefore, over many regions of the globe, aspects of daily precipitation and soil moisture cycle are being captured correctly by the Mosaic simulation. In a related paper, Reichle *et al.* (2004) present similar correlation values, despite using a separate land surface model (catchment, Koster *et al.*, (2000)) and averaging both data sets to monthly means. It is important to note that both studies use the forcing data described in this study.

Similar to the Reichle *et al.*, (2004) study, modeled soil moisture is drier than the satellite-derived observations. RMSE (MBE) between the SMMR-based observations and the Mosaic simulation are 0.11

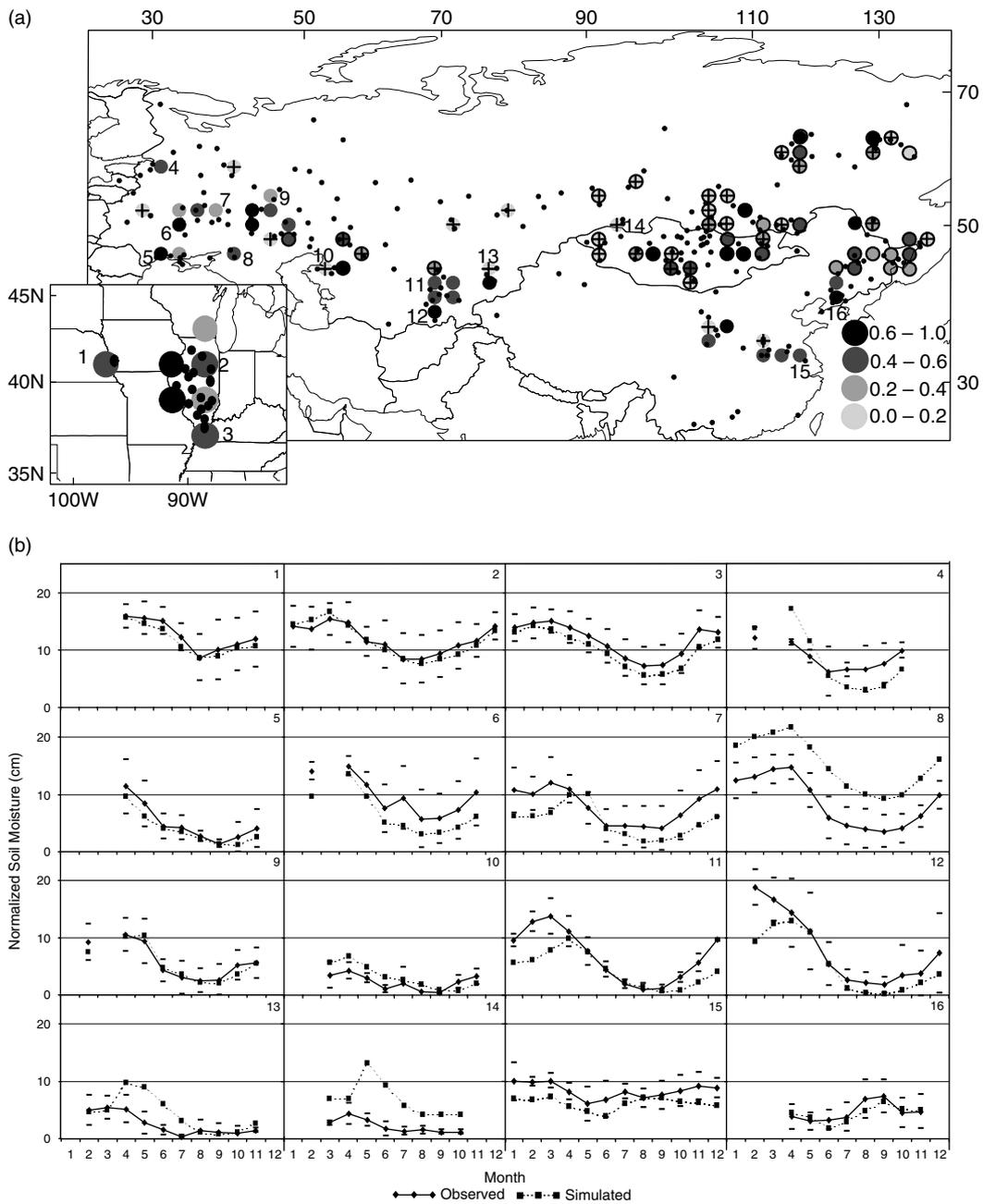


Figure 2. (a) Correlations between simulated and observed root zone soil moisture (observation locations are black dots). Open circles identify locations where observations are available for 8 months or more, and crosses identify locations where the correlations shown are not statistically significant. (b) Mean seasonal-simulated (dashed line) and observed soil moisture conditions for a number of locations identified in (a); outer tick marks identify one standard deviation from the mean observations

(0.072) and 0.10 (0.079) m^3/m^3 for noon and midnight observations respectively. RMSE (correlations) are highest (lowest) for in the day and lowest (highest) in the evening, and increase (decrease) with increasing soil moisture. This is largely affected by accuracy of the simultaneously estimated ground temperature, which were found to be more stable at night (Owe *et al.*, 2001); however, the length of time over which the modeled

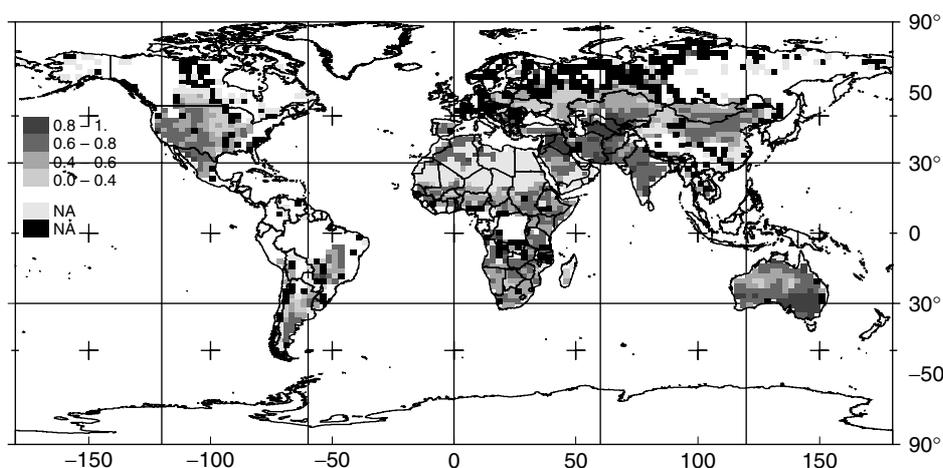


Figure 3. Correlations between the Mosaic simulation and midnight SMMR-based observations; grey scaled regions are statistically significant ($p < 0.05$) while over the regions in black are not. Over the regions in the lightest shade of grey the standard deviation in soil moisture or values are less than the approximate noise level in SMMR observations (Reichle *et al.*, 2004). No data is observed over regions where the retrieval of soil moisture is impacted by high vegetation water contents, or by the presence of ice, snow, and frozen soils

data is averaged may also have some effect. The SMMR observations, although interpolated to the larger Mosaic grid, can be regarded as instantaneous measurements, whereas the Mosaic output is averaged over the 6-h period that corresponds to the SMMR overpass time either at noon or midnight local time at the equator. Therefore, the evening observations may have lower RMSE and higher correlations due to lower evaporation rates from the surface, which allow wet surface soils to persist for longer periods.

In general, correlations between Mosaic simulation and the SMMR observations are highest over the semiarid regions of Australia, the Middle East, southwestern North America, Mongolia/northern China, and North Africa. High correlations are also observed over regions such as the Indian subcontinent. Negative, weak and statically nonsignificant correlations between the Mosaic output and the SMMR observations are observed over much of northern Russia, and Canada (black regions of Figure 3).

Regions that show weak, negative, or statistically insignificant correlations between modeled and observed surface soil moisture generally have high vegetation cover. Although we attempted to remove areas covered with high vegetation (this was completed as part of the Owe *et al.* (2001) product, and additional processing carried out as part of this study), clear boundaries of low to nonsignificant correlations are evident along the boundaries between the highly vegetated regions, where observations are not possible, and the observable semiarid regions. Over these areas, the failure of the model to accurately reproduce features of the daily surface soil type is likely related to a combination of factors including modeling and sampling scale issues, especially considering the large grid cells used in this analysis. However, the increased RMSE and lower correlations between observations and the Mosaic simulation are statistically related (as evaluated by Student's t tests) to increased vegetation optical depth ($r = 0.39$, $p < 0.01$).

3.3. Comparisons with the results of Nijssen *et al.* (2001)

We present correlations between monthly standardized anomalies in root zone soil moisture in the Mosaic and VIC simulations in Figure 4(a). The regions not colored were omitted from the analysis because of the presence of significant snow cover throughout much of the year (10 months or more). Over the stippled regions, apparent over the Himalayas and eastern China, the anomaly correlation between the two simulations was not statistically significant ($p < 0.05$). In general, standardized anomalies between the two simulations are well correlated over much of the global land surface. Locations where correlations are not statistically

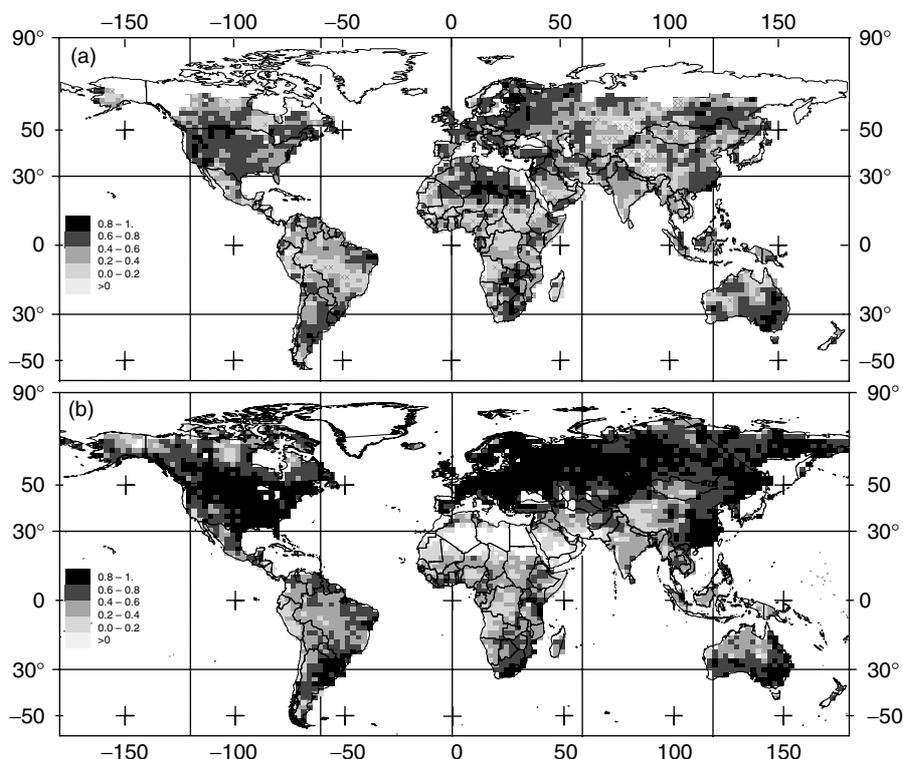


Figure 4. (a) Correlations between monthly standardized anomalies in root zone soil moisture for the Mosaic and VIC (Nijssen *et al.*, 2001) simulations. Over the stippled regions, the anomaly correlation was not statistically significant. (b) Identical to (a) except for precipitation

significant are observed over tropical Africa (extending to the Sahel), central South America, and eastern China.

Given the number of factors that can contribute to the differences observed between these simulations (e.g. differences to forcing, model parameterization, specification of land surface characteristics), identifying and controlling all this variability becomes exceedingly difficult. However, when a map identical to Figure 4(a) is developed for precipitation (Figure 4(b)) some important similarities between Figures 4(a) and 4(b) are apparent (correlation between the two maps is 0.39, $p < 0.01$). Although this result maybe anticipated, it is of interest because both of the precipitation data sets were scaled to GPCP observations. At the time when the Nijssen *et al.* (2001) data set was developed, GPCP observations were available for the period 1987–1993 (GPCP version 1a). Therefore, to supplement this data source for the period 1979–1986 they used the Hulme (1995) precipitation data set. To prevent a change in the precipitation statistics for the earlier (1979–1986) period, this data was then scaled to be consistent with the 1987–1993 GPCP climatology. The globally averaged correlation coefficient between this and Nijssen *et al.* (2001) precipitation forcing (between standard anomalies) is 0.73 for the 1987–1993 period and 0.54 for the period 1980–1986 due to the change in data sets. Correlations between the two data sets 1987–1993 are not closer to 1 because of multiple interpolations, merging the GPCP product with other data sets, and because of the use of different versions of the GPCP. Specifically, GPCP version 1a (Huffman *et al.*, 1997) was later superceded by GPCP versions 1c and most recently by version 2 (Adler *et al.*, 2003) because of changes in methods. For root zone soil moisture, similar differences to the relationship between the two data occur when we compare the correlations to standardized anomalies over these two time frames (the globally averaged correlation coefficient is 0.55 for the 1980–1986 period and 0.62 for the 1987–1993 period). Although this issue was unavoidable in the case of Nijssen *et al.* (2001) study, it underscores the necessity of a long, consistent time series for deriving the precipitation forcing.

Because of the differences identified between the precipitation data sets used in this study and in Nijssen *et al.* (2001), comparison to other soil water characteristics such as the yearly range of soil moisture is not as informative for understanding parameterization differences between the two models as studies where some of the sources of variability are controlled. However, for researchers using these soil moisture data sets for estimation of initial soil moisture states, or for large-scale water flux estimates, some discussion of differences between these simulations is necessary. In Figure 5, we compare the absolute bias between the VIC and Mosaic simulation for the annual range of soil water stored in the top 1 m. Globally averaged RMSE (MBE) between the two simulations is 12 (5.7) mm. In general, the VIC simulation estimates larger annual soil moisture ranges. This feature maybe related to the relatively shallow soil column used in the version of the VIC simulations analyzed here (1 m) (Nijssen *et al.*, 2001) *versus* the much deeper soil column specified in the Mosaic simulation (2.5 m). The differences shown are statically significant ($p < 0.05$), as evaluated with the nonparametric Wilcoxon signed-ranks tests, over approximately 50% of the Earth's surface shown in the stippled regions of Figure 5. Differences to the annual range of soil moisture are greatest over Europe, and South America, and are more closely associated with yearly total precipitation ($r = 0.25$, $p < 0.01$) than the differences between precipitation data sets ($r = 0.12$, $p < 0.01$). Although this relationship suggests the importance of parameterization differences or specification of soil moisture depth between the models, we had to limit our investigation into this matter due the amount of variability from other sources including the other forcing fields and specification of land surface hydrologic characteristics and vegetation types and amount.

4. SUMMARY AND CONCLUSIONS

Off-line land surface modeling simulations require accurate meteorological forcing with consistent spatial and temporal resolutions. For this reason, reanalysis products have been an attractive data source for numerous researchers. However, a number of studies have identified bias to many of the reanalysis fields, and therefore the use of raw reanalysis forcing is not recommended for water and energy balance simulations. Recognizing the existence of these errors, previous works (Berg *et al.*, 2003) have created hydrometeorological forcing data sets for the North America continent through implementation of a relatively simple bias correction procedure applied to the ECMWF and NCEP/NCAR reanalysis products. In this study, we expand upon this work to develop a global 0.5° forcing data sets for the time period 1979–1993 on a 6-hourly time step). We then use our forcing data to drive the Mosaic (Koster and Suarez, 1992) land surface process model for global estimation of soil moisture and other hydrological states and fluxes. The simulated soil moisture results are compared to observations obtained from the soil moisture data bank (Robock *et al.*, 2000) to satellite-derived surface soil moisture (Owe *et al.*, 2001), and to the modeled soil moisture product of Nijssen *et al.* (2001).

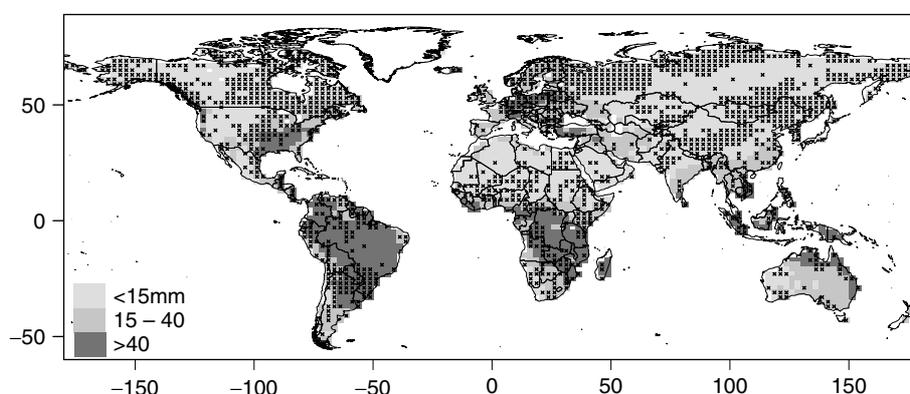


Figure 5. Absolute bias (in millimeters) between the VIC and Mosaic simulation for the annual range of soil water stored in the top 1 m. Over the stippled regions, differences shown are statistically significant ($p < 0.05$)

Observations from the global soil moisture data bank were compared to simulated results for a number of locations in the United States, Russia, China and Mongolia. Overall, statistically significant correlations are observed at a majority of these sites, and the seasonal cycle of water storage in the root zone is generally well represented. Weak relationships between modeled and observed soil moisture are generally limited to drier regions and are especially evident over western Mongolia.

Simulated soil moisture in the top 2 cm was compared to satellite observations derived from the SMMR (Owe *et al.*, 2001). Correlations were found to be strongest for data derived from the midnight overpass of the SMMR satellite ($r = 0.56$) than observations obtained during midday ($r = 0.43$). Statistical relationships between modeled and observed surface soil moisture weaken with increasing vegetation (as evaluated by vegetation optical depths). Agreement between simulated and observed surface soil moisture demonstrates that aspects of the daily soil water cycle (precipitation and soil drying) are captured in the reanalysis-derived precipitation and are correctly simulated within the Mosaic land surface model; however, similar to the analysis by Reichle *et al.* (2004), we observe high RMSE and MBE between the SMMR-based observations and the modeled results with the modeled soil moisture values drier than observations.

In addition to comparisons with observations, we also compare the soil moisture estimates of this study to those of Nijssen *et al.* (2001) for the period 1980–1993. In general, both data sets are well correlated in their prediction of wet and dry anomalies to root zone soil moisture, despite being derived from different land surface models, using different data sources for meteorological forcing, and with different specifications of the land surface properties. However, when we compare features such as the annual range of root zone soil moisture we find that the two simulations are not in general agreement, a result that complicates the transferability of soil moisture characteristics between separate models beyond the anomaly time series (also see Schaake *et al.*, 2004 and Reichle *et al.*, 2004). Further investigation into the causes of the differences between model simulations must follow a strategy where parameters such as hydrometeorological forcing and land surface characterization are more tightly constrained.

The soil moisture estimates produced as part of this analysis are available to other researchers for providing initial soil water estimates, or for understanding soil moisture variability over this 14-year time frame. We anticipate that this data can be used independently or as a companion product to similar data sets (e.g. Nijssen *et al.*, 2001). When used in combination with other data sets, we expect that this data set could provide some further information regarding the range in a given soil moisture estimate. Users combining estimates should be aware of the difficulties of applying soil moisture estimates across different models (e.g. Pittman *et al.*, 1999; Reichle *et al.*, 2004; Schaake *et al.*, 2004), and the uncertainty of given soil moisture value due to differences in parameterization (e.g. Entin *et al.*, 1999), in global forcing estimates (e.g. Fekete *et al.*, 2004), and in the specification of the land surface states (e.g. soil and vegetation types). In addition to the forcing data sets described above and in Nijssen *et al.* (2001), Zhao and Dirmeyer (2003) have recently developed a forcing data set for forcing off-line land surface models as part of the second phase of the global soil wetness project. There are several distinctions between these data sets, most notably due to differences, in the observations and reanalysis products used to create the forcing data sets, methods employed in their development, and to the temporal and spatial resolutions in all three data sets. Therefore, each of these forcing data sets could be used to further evaluate uncertainty of modeled soil moisture estimates due to differences in global forcing estimates (e.g. Berg and Famiglietti, 2003; Fekete *et al.*, 2004). At the time of writing, the results of these soil moisture simulations, including the forcing data and other hydrological flux data were available by request through the Global Land Data Assimilation System (GLDAS) (Rodell *et al.*, 2004) website (<http://www.ldas.gsfc.nasa.gov>).

ACKNOWLEDGEMENTS

This work was supported by NASA grants NAG5-11645, 12344, and the NASA Earth System Science Fellowship Program.

REFERENCES

Adler RF. *et al.* 2003. The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–Present). *Journal of Hydrometeorology* **4**: 1147–1167.

- Arora VK, Chiew FHS, Grayson RB. 2000. The use of river runoff to test CSIRO9 land surface scheme in the Amazon and Mississippi river basins. *International Journal of Climatology* **20**: 1077–1096.
- Berg AA, Famiglietti JS. 2003. Characterizing regional uncertainty in the initial soil moisture status. *Geophysical Research Letters* **30**: 10.1029/2003GL017075.
- Berg AA, Famiglietti JS, Walker JP, Houser PR. 2003. Impact of bias correction to reanalysis products on simulations of North American soil moisture and hydrological fluxes. *Journal of Geophysical Research* **108**(D16): 4490. 10.1029/2002JD003334.
- Betts AK, Ball JH, Viterbo P. 1999. Basin-scale surface water and energy budgets for the Mississippi from the ECMWF reanalysis. *Journal of Geophysical Research* **104**: 19 293–19 306.
- Betts AK, Ball JH, Viterbo P. 2003. Evaluation of the ERA-40 surface water budget and surface temperature for the Mackenzie River basin. *Journal of Hydrometeorology* **4**(6): 1194–1211.
- Betts AK, Hong S-Y, Pan H-L. 1996. Comparison of NCEP/NCAR reanalysis with 1987 FIFE data. *Monthly Weather Review* **124**: 1480–1498.
- Betts AK, Viterbo P, Beljaars ACM. 1998a. Comparison of the ECMWF reanalysis with the 1987 FIFE data. *Monthly Weather Review* **126**: 186–198.
- Betts AK, Viterbo P, Beljaars A, Pan H-L, Hong S-L, Goulden M, Wofsy S. 1998b. Evaluation of land-surface interaction in ECMWF and NCEP/NCAR reanalysis models over grassland (FIFE) and boreal forest (BOREAS). *Journal of Geophysical Research* **103**: 23 079–23 085.
- Betts AK, Viterbo P, Wood E. 1998c. Surface energy and water balance for the Arkansas-Red River basin from the ECMWF reanalysis. *Journal of Climate* **11**: 2881–2897.
- Daly C, Neilson RP, Phillips DL. 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology* **33**: 140–158.
- de Jeu RAM. 2003. *Retrieval of Land Surface Parameters Using Passive Microwave Remote Sensing*. PhD Thesis, Vrije Universiteit, Amsterdam.
- Dirmeyer PA. 2000. Using a global soil wetness dataset to improve seasonal climate simulation. *Journal of Climate* **13**: 2900–2922.
- Dirmeyer PA, Dolman AJ, Sato N. 1999. The pilot phase of the global soil wetness project. *Bulletin of the American Meteorological Society* **80**: 851–878.
- Douville H, Chavin F. 2000. Relevance of soil moisture for seasonal climate predictions: a preliminary study. *Climate Dynamics* **16**: 719–736.
- Entin JK, Robock A, Vinnikov KY, Hollinger SE, Liu S, Namkhai A. 2000. Temporal and spatial scales of observed soil moisture variations in the extratropics. *Journal of Geophysical Research* **105**: 11 865–11 877.
- Entin JK, Robock A, Vinnikov KY, Zabelin V, Liu S, Namkhai A, Adyasuren TS. 1999. Evaluation of global soil wetness project soil moisture simulations. *Journal of Meteorological Society of Japan* **77**: 183–198.
- Fekete BM, Vörösmarty CJ, Roads JO, Willmott CJ. 2004. Uncertainties in precipitation and their impacts on runoff estimates. *Journal of Climate* **17**: 294–304.
- Fennessy MJ, Shukla J. 1999. Impact of initial soil wetness on seasonal atmospheric prediction. *Journal of Climate* **12**: 3167–3180.
- Gibson JK, Källberg P, Uppala S, Nomura A, Hernandez A, Serrano E. 1997. ERA Description, *ECMWF Reanalysis Project Report Serial 1*, European Center For Medium-Range Weather Forecasts; Reading, England: 72.
- Gupta SK, Ritchey NA, Wilber AC, Whitlock CH, Gibson GG, Stackhouse PW Jr. 1999. A climatology of surface radiation budget derived from satellite data. *Journal of Climate* **12**: 2691–2710.
- Henderson-Sellers A, Pitman AJ, Love PK, Irannejad P, Chen TH. 1995. The project for Intercomparison of Land Surface Parameterisation Schemes (PILPS): Phases 2 and 3. *Bulletin of the American Meteorological Society* **76**(14): 489–503.
- Henderson-Sellers A, Irannejad P, McGuffie K, Pitman AJ. 2003. Predicting land-surface climates-better skill or moving targets? *Geophysical Research Letters* **30**(14): 1777. doi:10.1029/2003GL017387.
- Higgins RW, Shi W, Yarosh E, Joyce R. 2000. *Improved US Precipitation Quality Control System and Analysis*, NCEP/Climate Prediction Center ATLAS No. 7, U. S. Department of Commerce, National Oceanic and Atmospheric Administration.
- Holl SL. 2004. *The Sensitivity of Land Surface Model Simulations to Bias Reduction of ERA-15 Downward Radiation Forcing*. M.S. Thesis, Department of Geological Sciences, The University of Texas at Austin.
- Huffmann EJ, Adler RF, Arkin PA, Chang A, Ferrano R, Gruber A, Janowiak J, McNab A, Rudolf B, Schneider U. 1997. The Global Precipitation Climatology Project (GPCP) Combined Precipitation Dataset. *Bulletin of the American Meteorological Society* **78**(1): 5–20.
- Hulme M. 1995. Estimating global changes in precipitation. *Weather* **50**: 34–42.
- International GEWEX Project Office. 2002. *GSWP-2: The Second Global Soil Wetness Project Science and Implementation Plan*, Publication Series No. 37. IGPO: 65.
- Jones PD. 1994. Hemispheric surface air temperature variations: A reanalysis and update to 1993. *Journal of Climate* **7**: 1794–1802.
- Kanamitsu M, Ebisuzaki W, Woolen J, Yang S-K, Hnilo JJ, Fiorino M, Potter GL. 2002. NCEP-DOE AMIP-II reanalysis (R-2). *Bulletin of the American Meteorological Society* **83**: 1631–1643.
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Leetmaa A, Reynolds B, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Jenne R, Joseph D. 1996. The NCEP/NCAR 40-year Reanalysis Project. *Bulletin of the American Meteorological Society* **77**(3): 437–471.
- Koster RD, Suarez MJ. 1992. Modeling the land surface boundary in climate models as a composite of independent vegetation stands. *Journal of Geophysical Research* **97**: 2697–2715.
- Koster RD, Suarez MJ. 1996. *Energy and Water Balance Calculations in the Mosaic LSM*, Vol. 9, NASA Technical Memorandum: 104 606.
- Koster RD, Suarez MJ. 2003. Impact of land surface initialization on seasonal precipitation and temperature prediction. *Journal of Hydrometeorology* **4**: 408–423.
- Koster RD, Suarez MJ, Heiser M. 2000. Variance and predictability of precipitation at seasonal to interannual timescales. *Journal of Hydrometeorology* **1**: 26–46.

- Liang X, Lettenmaier DP, Wood EF, Burges SJ. 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research* **99**: 14415–14428.
- Lenters JD, Coe MT, Foley JA. 2000. Surface water balance of the continental United States, 1963–1995: Regional evaluation of a terrestrial biosphere model and the NCEP/NCAR reanalysis. *Journal of Geophysical Research* **105**: 22393–22425.
- Maurer EP, Nijssen B, Lettenmaier DP. 2000. Use of reanalysis land surface water budget variables in hydrologic studies. *Gewex News* **10**(4): 6–8.
- Maurer EP, O'Donnell GM, Lettenmaier DP, Roads JO. 2001a. Evaluation of the land surface water budget in NCEP/NCAR and NCEP/DOE reanalyses using an off-line hydrologic model. *Journal of Geophysical Research* **106**: 17841–17862.
- Maurer EP, O'Donnell GM, Lettenmaier DP, Roads JO. 2001b. Evaluation of NCEP/NCAR Reanalysis water and energy budgets using macroscale hydrologic model simulations. *Land Surface Hydrology, Meteorology and Climate: Observations and Modeling: Water Science and Applications 3*, Venkataraman L, Albertson J, Schaake J (eds). AGU: Washington DC.
- Mintz Y, Serafini YV. 1992. A global monthly climatology of soil moisture and water balance. *Climate Dynamics* **8**: 13–27.
- Mitchell KE, Lohmann D, Houser PR, Wood EF, Schaake JC, Robock A, Cosgrove BA, Sheffield J, Duan Q, Luo L, Higgins RW, Pinker RJ, Tarpley JR, Lettenmaier DP, Marshall CH, Entin JK, Pan M, Shi W, Koren V, Meng J, Ramsay BH, Bailey AA. 2004. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrologic modeling system. *Journal of Geophysical Research* **109**: DOI: 10.1029/2003JD003823.
- New M, Hulme M, Jones P. 2000. Representing twentieth-century space – time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate. *Journal of Climate* **13**: 2217–2238.
- Nijssen B, Schnur R, Lettenmaier DP. 2001. Global retrospective estimation of soil moisture using the variable infiltration capacity land surface model, 1980–93. *Journal of Climate* **14**: 1790–1808.
- Njoku EG, Li L. 1999. Retrieval of land surface parameters using passive microwave measurements at 6–18 GHz. *IEEE Transactions on Geoscience and Remote Sensing* **37**: 79–93.
- Owe M, De Jeu R, Walker JP. 2001. A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index. *IEEE Transactions on Geoscience and Remote Sensing* **39**: 1643–1654.
- Pittman AJ, Henderson-Sellers A, Desborough CE, Yang ZL, Abramopoulou F, Boone A, Dickinson RE, Gedney N, Koster R, Kowalczyk KE, Lettenmaier D, Liang X, Mahfouf JF, Noilhan J, Polcher J, Qu W, Robock A, Rosenzweig C, Schlosser CA, Shmakin AB, Smith J, Suarez M, Verseghy D, Wetzel P, Wood E, Xue Y. 1999. Key results and implications from phase 1(c) of the project for intercomparison of land-surface parameterization schemes. *Climate Dynamics* **15**: 673–684.
- Reichle RH, Koster RD, Dong J, Berg AA. 2004. On global soil moisture from satellite observations, land surface models, and ground data: Implications for data assimilation. *Journal of Hydrometeorology* **5**: 430–442.
- Roads JO, Chen S-C, Kanmitsu M, Juang H. 1999. Surface water characteristics in NCEP global spectral model and reanalysis. *Journal of Geophysical Research* **104**: 19307–19327.
- Roads J, Betts A. 2000. NCEP-NCAR and ECMWF reanalysis surface water and energy budgets for the Mississippi basin. *Journal of Hydrometeorology* **1**: 88–94.
- Roads JR, Lawford R, Bainto E, Berbery E, Chen S, Fekete B, Gallo K, Grundstein A, Higgins W, Kanamitsu M, Krasewski W, Lakshmi V, Leathers D, Lettenmaier D, Luo L, Maurer E, Meyers T, Miller D, Mitchell K, Mote T, Pinker R, Reichler T, Robinson D, Robock A, Smith J, Srinivasan G, Verdin K, Vinnikov K, Vonderttaurt A, Vörösmarty C, Williams S, Yarosh E. 2003. GCIP Water and Energy Budget Synthesis (WEBS). *Journal of Geophysical Research* **108**(D16): doi: 10.1029/2002JD002583.
- Robock A, Schlosser CA, Vinnikov KY, Speranskaya NA, Entin JK, Qiu S. 1998. Evaluation of the AMIP soil moisture simulations. *Global and Planetary Change* **19**: 181–208.
- Robock A, Vinnikov KY, Srinivasan G, Entin JK, Hollinger SE, Speranskaya NA, Liu S, Namkhai A. 2000. The global soil moisture data bank. *Bulletin of the American Meteorological Society* **81**: 1281–1299.
- Robock A, Luo L, Wood EF, Went F, Mitchell KE, Houser PR, Schaake JC, Lohmann D, Cosgrove B, Sheffield J, Duan Q, Higgins RW, Pinker RT, Tarpley JD, Basara JB, Crawford KC. 2003. Evaluation of the North American Land Data Assimilation System over the southern Great Plains during the warm season. *Journal of Geophysical Research* **108**(D22): 8846. doi:10.1029/2002JD003245.
- Rodell M, Famiglietti JS. 1999. Detectability of variations in continental water storage from satellite observations of the time dependent gravity field. *Water Resources Research* **35**: 2705–2723.
- Rodell M, Houser PR, Jambor JG, Mitchell K, Meng J, Arsenault K, Cosgrove B, Radakovich J, Bosilovich M, Entin JK, Walker JP, Lohmann D, Toll D. 2004. The Global Land Data Assimilation system. *Bulletin of the American Meteorological Society* **85**: 381–394.
- Row LW, Hastings DA, Dunbar PK. 1995. *TerrainBase Worldwide Digital Terrain Data Documentation Manual*. National Geophysical Data Center: Boulder, CO.
- Schaake JC, Duan Q, Koren V, Mitchell KE, Houser PR, Wood EF, Robock A, Lettenmaier DP, Lohmann D, Cosgrove B, Sheffield J, Luo L, Higgins RW, Pinker RT, Tarpley JD. 2004. An intercomparison of soil moisture fields in the North American Land Data Assimilation System (NLDAS). *Journal of Geophysical Research* **109**: D01S90. doi: 1029/2002JD003309.
- Schmugge TJ. 1983. Remote sensing of soil moisture: Recent advances. *IEEE Transactions on Geoscience and Remote Sensing* **21**: 336–344.
- Sellers PJ, Collatz J, Hall FG, Meeson BW, Closs J, Corprew F, McManus J, Myers D, Sun KJ, Dazlich D, Kerr Y, Koster R, Los S, Mitchell K, Try P. 1996. The ISLSCP initiative I global datasets: surface boundary conditions and atmospheric forcings for land-atmosphere studies. *Bulletin of the American Meteorological Society* **77**: 1987–2006.
- Simmons AJ, Gibson JK (eds). 2000. *ERA-40 Project Plan*. ERA-20 Project Report 1, ECMWF: 62.
- Srinivasan G, Robock A, Entin JK, Luo L, Vinnikov KY, Viterbo P. Participating AMIP Modeling Groups. 2000. Soil moisture simulations in revised AMIP models. *Journal of Geophysical Research* **105**: 26635–26644.
- Seneviratne SI, Viterbo P, Luthi D, Schar C. 2004. Inferring changes in terrestrial water storage using ERA-40 reanalysis data: The Mississippi River basin. *Journal of Climate* **17**: 2039–2057.
- Thornton PE, Running SW. 1999. An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. *Agricultural and Forest Meteorology* **93**: 211–228.

- Vinnikov KY, Robock A, Qiu S, Entin JK. 1999. Optimal design of surface networks for observation of soil moisture. *Journal of Geophysical Research* **104**: 19 743–19 749.
- Willmott CJ, Matsuura K. 2001. *Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1950–1999) Version 1.02*. Data set archived at <http://www.siesip.gmu.edu/home.html>.
- Zhao M, Dirmeyer PA. 2003. Production and analysis of GSWP-2 near-surface meteorology data sets. *COLA Technical Report 159*, [Available from <http://www.iges.org/pubs/tech.html>], 38.