Uncertainties in North American Land Data Assimilation Systems over the Contiguous United States

KINGTSE C. MO
Climate Prediction Center/NCEP/NWS/NOAA, Camp Springs, Maryland

LI-CHUAN CHEN
Cooperative Institute for Climate and Satellites, Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland

SHRADDHANAND SHUKLA, THEODORE J. BOHN, AND DENNIS P. LETTENMAIER
Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington

(Manuscript received 13 October 2011, in final form 26 January 2012)

ABSTRACT

The Environmental Modeling Center (EMC) at the National Centers for Environmental Prediction (NCEP) and the University of Washington (UW) run parallel drought monitoring systems over the continental United States based on the North American Land Data Assimilation System (NLDAS). The NCEP system uses four land surface models (LSMs): Variable Infiltration Capacity (VIC), Noah, Mosaic, and Sacramento (SAC). The UW system uses VIC, SAC, Noah, and the Community Land Model (CLM). An assessment of differences in drought characteristics using both systems for the period 1979–2008 was performed. For soil moisture (SM) percentiles and runoff indices, differences are relatively small among different LSMs in the same system. However, the ensemble mean differences between the two systems are large over the western United States—in some cases exceeding 20% for SM and runoff percentile differences. These differences are most apparent after 2002 when the NCEP system transitioned to use the real-time North American Regional Reanalysis (NARR) and its precipitation gauge station data. (The UW system went into real-time operation in 2005.) Experiments were performed to address the sources of uncertainties. Comparison of simulations using the two systems with different model forcings indicates that the precipitation forcing differences are the primary source of the SM and runoff differences. While temperature, shortwave and longwave radiation, and wind speed forcing differences are also large after 2002, their contributions to SM and runoff differences are much smaller than precipitation.

1. Introduction

Droughts are among the costliest of natural hazards in the United States, drought costs over the last decade have averaged about $4 billion, and the inflation-adjusted (2007) cost of the 1988 Midwest drought exceeded $70 billion (NCDC 2011). To mitigate the impacts of droughts, the U.S. Drought Monitor (USDM) and Seasonal Drought Outlook provide operational monitoring and prediction of droughts over the continental United States (CONUS). These applications make use of drought indices, which are designed to provide objective measures of drought severity and to facilitate drought classification. Keyantash and Dracup (2002) suggested three types of droughts. “Meteorological drought” is defined by precipitation $P$ deficits and is often measured by the standardized precipitation index (SPI; McKee et al. 1993, 1995). “Agricultural drought” is caused by soil moisture (SM) deficits; total soil water storage or SM anomaly percentiles are now commonly used to monitor agricultural drought (Andreadis et al. 2005; Wang et al. 2009). “Hydrological drought” is reflective of runoff deficits, which can be monitored by using the standardized runoff index.

Corresponding author address: Kingtse Mo, Climate Prediction Center/NCEP/NWS/NOAA, 5200 Auth Rd., Camp Springs, MD 20746.

E-mail: kingtse.mo@noaa.gov

DOI: 10.1175/JHM-D-11-0132.1

© 2012 American Meteorological Society
(SRI) or streamflow anomaly percentiles (Shukla and Wood 2008; Mo 2008).

While precipitation station data are available over CONUS from the 1850s to present, there are few long-term, in situ measurements of SM. Furthermore, runoff measurements are not made on a fully (i.e., grid mesh) spatially distributed basis, although streamflow observations (essentially the convolution integral of runoff over a river network) are available at many locations over the CONUS—some with record lengths exceeding a century. For purposes of estimating drought characteristics on a spatial basis, model-derived drought indices for SM and runoff are being used increasingly, where the models are tested and evaluated using limited SM and more extensive streamflow observations. One such approach is based on variations of the North American Land Data Assimilation System (NLDAS; Xia et al. 2012a,b). One challenge in the use of NLDAS and similar products for drought classification is characterization of the biases and uncertainties that result from the use of different land surface models and model forcings. In this paper, we evaluate the relative contributions of these two sources of uncertainty to drought characterization over the CONUS.

Robock et al. (2003) compared four land surface models from an early version of the National Centers for Environmental Prediction (NCEP) NLDAS. They found that although the differences of SM are large among models, anomalies—defined as the departures from the monthly mean climatology—are much more similar. Dirmeyer et al. (2006) analyzed model outputs from the Global Soil Wetness Project (GSWP) and reached similar conclusions. These studies suggested that SM anomaly percentiles are suitable for drought classification because they are only weakly dependent on the specific model used. For both studies, all LSMs were driven by the same input forcings. Mo (2008) compared SM percentiles and SRI from the Variable Infiltration Capacity (VIC) model output in the University of Washington (UW) NLDAS and the Noah model output in the NCEP NLDAS. She found that differences in SM percentiles over the western United States between the two models were so large as to be problematic for drought classification. In that study, the NCEP and UW systems used both different models and forcings, which precluded isolating the sources of the differences.

The Environmental Modeling Center (EMC) at NCEP (http://www.emc.ncep.noaa.gov/mmb/nldas) and UW (Surface Water Monitor; http://www.hydro.washington.edu/forecast/monitor/) run parallel drought monitoring systems based on variations of NLDAS in near–real time. Both systems are used by USDM authors and the National Integrated Drought Information System (NIDIS) community to classify droughts. For SM percentiles and runoff indices, systematic differences have been observed between the two systems. The objective of this study is to assess the sources of these differences (uncertainties) in drought indices derived from the two systems. To do so, we first compared drought indices derived from the NCEP and UW NLDAS systems to quantify differences in the structure of the two systems. We then conducted numerical experiments to address the sources of uncertainties. We describe the two systems briefly in section 2. In section 3, we compare drought indices derived from the two systems. Section 4 presents experiments that examine the effects of differences in forcings used by the two systems on SM and runoff indices. A summary and conclusions are provided in section 5.

2. The NLDAS systems

a. The UW system

The UW multimodel system (Wang et al. 2009) as used in this study includes four land surface models: VIC.4.0.6 (Liang et al. 1994), Noah 2.8 (Koren et al. 1999; Ek et al. 2003), Sacramento/Snow-17 (SAC) (Burnash et al. 1973; Anderson 1973), and the Community Land Model 3.5 (CLM3.5) (Oleson et al. 2007). Model descriptions and properties can be found in Wang et al. (2009). All models are driven by gridded forcings that cover the period from 1915 to the present over CONUS.

The UW forcings consist of daily total precipitation, maximum temperature $T_{\text{max}}$ and minimum temperature $T_{\text{min}}$, and surface wind speed. The forcings are derived from a set of index stations (Wood and Lettenmaier 2006) that are a subset of the National Climatic Data Center’s (NCDC) Cooperative Observer stations (Andreadis et al. 2005). Specifically, the index stations—totaling about 231 over CONUS—were selected based on data quality and stability of the stations. These are stations that are currently operating in real time, and have data records that cover most of the period from 1915 to present. The advantage of using index stations is the continuity of data records, because the same stations are used for the entire simulation period. Figure 1 shows the spatial distribution of the index stations. The coverage is denser over the eastern United States than over the western United States. The station data are interpolated to 0.5° grid cells using methods outlined in Maurer et al. (2002) and Niijssen et al. (2001). To include an adjustment for topographic effects, the Precipitation Regression on Independent Slopes Method (PRISM) data of Daly et al. (1994) is applied on a seasonally varying but climatologically fixed basis. Within each day, the precipitation data are equally distributed to hourly time steps, while temperature and
radiative forcings have an imposed diurnal cycle, as described in Maurer et al. (2002). Where data are missing, they are interpolated from nearby stations, which has the effect of creating a somewhat smoother gridded field, especially in the mountainous parts of the western United States, which have large topographic variations. The surface wind speed data are taken from the NCEP–National Center for Atmospheric Research (NCAR) reanalysis (Kalnay et al. 1996).

For temperature, \( T_{\text{min}} \) and \( T_{\text{max}} \) are obtained from the index stations, and gridded to 0.5\(^\circ\) following the method described in Maurer et al. (2002) and Wood and Lettenmaier (2006) in the UW system. The daily temperature range (\( T_{\text{max}} - T_{\text{min}} \)) is then used to estimate daily downward solar radiation, vapor pressure, and downward longwave radiation, following the methods of Thornton and Running (1999), Kimball et al. (1997), and the Tennessee Valley Authority (TVA 1972; Idso 1981), respectively. Hourly versions of these variables as well as \( T_{\text{air}} \) are computed by fitting a spline to the daily \( T_{\text{min}} \) and \( T_{\text{max}} \) data.

b. The NCEP system

The NCEP system uses four land surface models: VIC, Noah, SAC, and Mosaic. While three of the models used in the NCEP system have the same names as in the UW system, the model versions differ in some cases. Detailed descriptions and configurations of the NCEP system can be found in Ek et al. (2011) and Xia et al. (2012b) with updated information available from http://ldas.gsfc.nasa.gov/nldas/NLDAS2forcing.php.

In contrast to the UW system, which uses 0.5\(^\circ\) spatial resolution, the spatial resolution of the NCEP system is 0.125\(^\circ\). NCEP simulations start from 1979 and continue to present and cover the same CONUS domain as the UW system, although the UW system also covers Mexico. Forcings for the NCEP system other than precipitation come from the North American Regional Reanalysis (NARR). For surface variables, 2-m temperature and specific humidity and 10-m winds are used. The data have been adjusted to the surface for terrain height (Cosgrove et al. 2003), and biases in the NARR downward solar radiation climatology are corrected using satellite observations from Pinker et al. (2003).

The precipitation data used in the NCEP system are the Climate Prediction Center (CPC)/Office of Hydrologic Development (OHD) 0.125\(^\circ\) daily precipitation gauge dataset. These data are derived from three sources: NCDC daily cooperative observer stations (1948–present), the River Forecast Center first-order stations (1992–present), and daily accumulations from the hourly precipitation dataset (1948–present) (Higgins et al. 2000). PRISM (Daly et al. 1994) is used to adjust precipitation for topographical influences in a similar manner to that used in the UW system. The NCEP system uses the 3-hourly mean 2-m temperature from NARR with elevation adjustment rather than using temperature indexing methods. The downward solar radiation and longwave radiation and specific humidity are taken from NARR and are also elevation adjusted.

The NCEP system used gridded precipitation gauge data from the sources noted above through 2001. From 2002 on, NARR was transitioned to real-time operation and since then the NARR and the \( P \) analysis in the NCEP system uses the same real-time rain gauge station input data. To distinguish the two periods, we denote the period 1979–2001 as the historical period and the period 2002–08 as the real-time period. For the real-time period, station data counts dropped by more than 1/3 in comparison to the historical period. Figure 2 shows the total number of precipitation gauge reports per 0.5\(^\circ\) grid box per month for \( P \) from the NCEP system averaged for selected years. If the reports come in once daily, then the counts for each station each month should be between 28 and 31. Some grid boxes contain multiple stations and some stations report two to four times a day during some time periods; hence the numbers are higher. Similar to the UW system (Fig. 1), the station coverage is denser over the eastern United States than over the west. In general, there are sufficient reports to obtain realistic representation of precipitation even for the real-time period over the eastern region. For the western region, however, the number of station reports drops significantly in real time. When data are sparse, gridded precipitation estimates are strongly dependent on the method used to interpolate the data. By comparing Fig. 2 to Fig. 1, it can be seen that some areas in the western United States—such as Nevada, New Mexico, eastern Texas, western Washington, and Montana—have no stations reporting in real time. The sparse coverage
introduces uncertainties in the forcings and outputs from the model simulations.

3. Comparisons between the NCEP and UW systems

a. Methodology and procedures

To compare model results, we used the monthly mean fields from the two systems and processed each land surface model separately. The comparison covers the common base period from 1979 to 2008 over the contiguous United States. Because the two systems have different spatial resolutions, the NCEP output was aggregated from 0.125° to 0.5° to be consistent with the UW output. For each month and each variable, the monthly mean climatology and standard deviation were computed for the base period (1979–2008). The monthly mean anomaly is defined as the departure from the monthly mean climatology. The standardized monthly mean anomaly was obtained by normalizing the monthly mean anomaly by its standard deviation for that month. For each model and each month, the 6-month standardized precipitation index (SPI6) (McKee et al. 1993, 1995), 6-month standardized runoff index (SRI6), and SM anomaly percentile were computed. The SM anomaly percentile was calculated using a fitted Gaussian distribution for each month as in Mo (2008). Each model is treated as an ensemble member.

Drought is defined as occurring when the SM percentile is below 20% or the SRI index is less than −0.8 for a season or longer (Svoboda et al. 2002). If the differences between the two systems exceed this threshold, then, as in Mo (2008), we categorize them as too large for drought classification.

The ensemble mean of an index is the equally weighted average over the ensemble members. The spread among the members of the ensemble is estimated using a method similar to that of Dirmeyer et al. (2006), but with weightings by variance rather than standard deviation. The formula is described as follows.

Let $X(t)$ be the equally weighted ensemble mean of a drought index at time $t$. The ensemble can be four members from the UW or NCEP system or all eight members from the two systems combined. Here $S_m(t)$ is the departure of the drought index for model $m$ at time $t$ from $X(t)$. The variance of the intermodel spread $\Omega_{sp}$ is defined as

$$\Omega_{sp} = \frac{1}{M \times T} \sum_{t=1}^{T} \sum_{m=1}^{M} S_m^2(t),$$

where $T$ is the length of the base period and $M$ is the number of models in the ensemble. The variance of interannual variability of the ensemble means $\Omega_{int}$ is defined as the variance of the ensemble means for the base period. The measure of spread $R$ is defined as the ratio of
the variance of the intermodel spread and the interannual variance of the ensemble means—that is,

\[ R = \frac{\Omega_{sp}}{\Omega_{int}}. \]  

(2)

b. SM and runoff differences between NCEP and UW

The \( R \) values for SM percentiles (SMP) are computed for the four UW members, the four NCEP members, and all eight members together. When \( R \) is close to 1, the spread is close to the interannual variability and the SMP differences between the two systems are so large as to provide little useful information for drought classification. The \( R \) values are about 0.2 among the members in the same system (Figs. 3a,b). The \( R \) values for the two systems combined (all eight members) are larger (see Fig. 3c) in comparison to \( R \) for either system alone. Over the eastern United States (east of 95°W), differences of the \( R \) values for the combined and either system alone are generally less than 0.4. Over the interior of the western United States,
the combined $R$ is almost double the value for either system alone. Figure 3d shows the root-mean-square (RMS) difference of the ensemble mean of SMP between the NCEP and UW systems. Over the eastern United States, the RMS values are about 10%–15%. Over the western United States (west of 95°W), the RMS values are greater than 20%, which as a practical matter is too large for drought classification. The RMS difference of SR16 shows a similar pattern (Fig. 3c), with large differences over the western United States. The largest differences occur after 2002 for both variables.

The SMP and SR16 differences between the ensemble means from the NCEP and UW systems for four 7-yr periods are presented in Fig. 4. The differences for the historical period are less than 20% (Figs. 4a–c), whereas after 2002, the differences over the western United States can be as large as 20%–30% (Fig. 4d). The UW system is usually wetter than the NCEP. The differences of SR16 between two systems are also larger in the real-time period in comparison to the historical period.

The differences over the western United States after 2002 are systematic and they are not caused by one particular model. For example, Fig. 5 shows the SMP averaged over the area (38.25°–42.25°N, 105.25°–110.25°W) over the western region (Fig. 3f) for the ensemble means (Fig. 5a) and each individual member (Fig. 5b) for both systems. The corresponding plot for SR16 is given in Fig. 5c. Differences are comparatively small within the members of the same system in comparison with the differences of the ensemble means between the NCEP and UW systems. The differences of SMP and SR16 between the two systems are comparably smaller in the historical period than in the real-time period. After 2002, SMP and SR16 from the members in the NCEP system (red) are consistently higher than those in the UW system (green; Figs. 5b,c). The differences between the two systems are consistent and are not caused by one or two models as seen in Figs. 5b,c. Similar differences can also be found in SPI6 (Fig. 5d).

c. Differences in forcings

We have shown that the differences of SMP and runoff among models of the same system driven by the same forcing are relatively small, but the differences between the two systems are large. This finding suggests that the differences are mainly caused by the input forcings. In addition to precipitation, air temperature ($T_{\text{air}}$) is also an important forcing.

Figures 6a–d show the differences of normalized $P$ and $T_{\text{suf}}$ anomalies between the NCEP and UW systems in four 7-yr periods. For the historical period, the UW $P$ anomalies are wetter than the NCEP, but differences are small. The differences are large in the real-time period over the western United States where the UW $P$ is drier than the NCEP. In addition to the mean difference, the UW $P$ anomalies are smoother and have less variability than NCEP (not shown). While the differences between two systems are large for the real-time period, the differences of the annual mean precipitation for the same system between the historical period and the real-time period are small (not shown).

To illustrate the uncertainties in air temperature, we computed the monthly mean $T_{\text{air}}$ anomalies for both systems, thus removing model biases. The differences of temperature overall are larger. The NARR assimilates precipitation but not surface temperature. For the NCEP system, $T_{2m}$ is taken from the model outputs and that may cause differences. The differences of $T_{\text{air}}$ anomalies between the two systems averaged over four 7-yr periods are plotted in Figs. 6e–h. For the western United States, the UW system was cooler than the NCEP before 1995. After 1996, the UW system was warmer. Over the eastern United States, the UW system was warmer than the NCEP from 1981 to 1987, but cooler than the NCEP for the real-time period. Overall differences in the historical period (1979–2001) are smaller than the differences for the real-time period.

The differences in longwave radiation between two systems are systematic and do not have large seasonal variations. Differences averaged over the historical and the real-time periods are also similar. Figure 7a shows the differences in annual mean longwave radiation. The UW system has higher longwave radiation over the Southwest with a maximum close to 30 W m$^{-2}$ and lower longwave radiation along the East Coast and the Northeast in comparison to the NCEP system. This pattern corresponds to the spatial distribution of the differences between the two systems’ mean $T_{\text{air}}$ over the real-time period (Fig. 6h), and may have resulted from the UW system’s use of the TVA algorithm (TVA 1972) to estimate downward longwave radiation as a function of $T_{\text{air}}$.

The differences in shortwave radiation are seasonally dependent, but the spatial patterns and mean differences for the historical period and the real-time period are similar (Figs. 7d,e). The mean differences for the cold season (November–March) averaged for the study period are small (Fig. 7b), but the differences are larger for the warm season (April–October) (Fig. 7c). The UW system has less shortwave radiation over the Southern Plains and the Gulf states and more shortwave radiation over the Northeast and the western United States, with a maximum above 30 W m$^{-2}$ located in Colorado and Utah. No single factor appears to be clearly responsible
for the differences. The UW system’s shortwave and vapor pressure algorithms depend on each other (via iteration) as well as on daily temperature range and on the timing of precipitation events; thus, biases should be expected to differ between different climate regimes.

4. VIC experiments to identify the sources of uncertainties

The previous analyses show that there are large forcing differences between two systems. Large $P$ differences
between the two systems occurred after 2002 when the NCEP system began to use input forcings from the real-time NARR and the real-time precipitation gauge reports used by NARR. The UW system went to real-time operation in 2005. Thereafter the number of precipitation gauge reports are somewhat low over the western U.S. region in real time. Therefore, we hypothesize that precipitation forcing is the main source of uncertainties in SM and runoff indices. To verify our hypothesis, we conducted numerical experiments to examine the effects of input forcings on drought indices. Since anomalies are relatively insensitive to the model used, we chose the VIC model version 4.0.6 from the UW system to perform the experiments. The model simulation covers the CONUS with grid resolution of 0.5°. We did not include Mexico, which uses a different data source in the UW system, and is not included in NCEP. All simulations were made for the period 1 January 1979–31 December 2008 with initial conditions on 31 December 1978 taken from the VIC model simulation from the UW system.

a. Impacts of $P$, $T_{\text{max}}$, and $T_{\text{min}}$

To assess the impacts of $P$, we separated the forcing into two parts: precipitation alone ($P$) and other forcings $F$ including $T_{\text{max}}$, $T_{\text{min}}$, and wind speed. The first simulation, labeled as Exp ($F_{\text{uw}}, P_{\text{uw}}$), is the control run for the UW system. Both $F$ and $P$ were taken from the UW forcings. The outputs from this simulation and those from the VIC model in the UW system are the same. The second simulation, labeled as Exp ($F_{\text{ncep}}, P_{\text{ncep}}$), represents the control run for the NCEP system. Both $F$ and $P$ in this simulation were obtained from the NCEP forcing. The third simulation, labeled as Exp ($F_{\text{ncep}}, P_{\text{uw}}$), uses $F$ forcings taken from the NCEP system and $P$ from the UW system. The last simulation, labeled as Exp ($F_{\text{uw}}, P_{\text{ncep}}$), is opposite to the previous run and uses $F$ forcings taken from the UW system and $P$ from the NCEP system. SMP and SRI6 were computed from the outputs of all four simulations.

The control experiments capture the major differences between the two systems. Figures 8a and 8d show the RMS differences of SMP and SRI6 for the period from 1979 to 2008, respectively. There are large differences over the western United States in comparison to the east, which is consistent with the RMS differences between the two systems from the ensemble means (Figs. 3d,e). The differences of SMP and SRI6 between the two control experiments for the historical period (1979–2001) are small (Figs. 8b,e). Large differences between the two control experiments occur in the real-time period 2002–08 (Figs. 8c,f). The largest differences are over the interior of the western United States and the Southwest where precipitation gauge densities are sparse during the real-time period (Fig. 2). This result is consistent with the findings from the ensemble means of the two systems (Fig. 3).

The RMS differences of SMP and SRI6 for the period 1979–2008 from the four experiments are given in Fig. 9. The threshold for drought is 20% for SMP and 0.8 for SRI. If the differences exceed the threshold, then we argue that they are too large for drought classification. The influence of the $F$ forcings including temperature and wind is small. The differences between the simulations with different $P$ forcings and the same $F$ forcings are large (Figs. 9a,b,e,f). They are comparable to the differences between the two control runs (Figs. 8a and 8d). The differences of SMP over the eastern United States are about 10% or less while the differences over the west are mostly larger than 20%. The differences between the simulations with different $F$ forcings and the same $P$
Forcings are small (Figs. 9c,d,g,h)—less than 10%, which is within the threshold for drought classification. Similar differences are also found for SRI6. Therefore, we conclude that precipitation is the major contributor to the uncertainties.

b. Impact of radiation terms and $q_{\text{surf}}$ specific humidity

For the VIC model, the changes in $T_{\text{max}}$ and $T_{\text{min}}$ will lead to changes in radiation terms and $q_{\text{surf}}$. This is not
the case for the NCEP system, which takes radiation terms and specific humidity from the NARR, but adjusted for elevation. Therefore, additional experiments were performed to assess the impact of radiation terms. The two experiments were performed using a newer version of VIC (VIC 4.1.2) that allows us to prescribe the radiation terms and surface humidity. (While some of the model physics in VIC 4.1.2 differ slightly from those of VIC 4.0.6, VIC 4.1.2 has an option that essentially replicates VIC 4.0.6, given the same input forcings and parameters, through the appropriate settings of run-time options.) The different versions of VIC are described at the website http://www.hydro.washington.edu/Lettenmaier/Models/VIC/Development/CurrentVersion.shtml. The experiment period is 1979–2008. SMP and SRI6 were computed from the SM and runoff outputs as in the earlier experiments.

We repeated the UW control run with VIC 4.1.2 set to VIC 4.0.6 options. This experiment is referred to as Exp (UW rad). Both $P$ and $F$ forcing terms were the same as Exp ($F_{uw}$, $P_{uw}$); only the VIC model version was (slightly) different. The radiation terms were computed from $T_{min}$ and $T_{max}$ as in Exp ($F_{uw}$, $P_{uw}$). The RMS differences of SMP and SRI6 between Exp (UW rad) and Exp ($F_{uw}$, $P_{uw}$) were small (Figs. 10a,c). This gives us confidence to perform the next experiment. The radiation experiment Exp (NCEP rad) uses the same $P$ and $F$ forcings from the UW system, but with radiation terms and $q_{surf}$ from the NCEP system. The differences of SMP and SRI6 between Exp (NCEP rad) and Exp (UW rad) were less than 10% (Figs. 10b,d). While the differences of radiation terms between two systems were large, their impact on SM and runoff was small.

5. Summary and conclusions

The NCEP/EMC and UW run parallel drought monitoring systems based on variations of NLDAS gridded
forcing data over CONUS. Both systems use a suite of land surface models. We have noticed consistent differences in soil moisture percentile and runoff indices after 2002, when the NCEP system began to use real-time forcings. We assessed uncertainties in the NLDAS by comparing drought indices derived from the two systems and performing model experiments to determine the sources of uncertainties.

For SMP and SRI, the differences are small among land surface models in both systems, which is consistent with the findings of Robock et al. (2003) and Dirmeyer et al. (2006). However, differences of the ensemble means between the two systems are large, especially over the western United States and especially post-2002, where and when the differences of SM percentiles can be as large as 20%—too large for drought classification. The differences are not due to any particular model and they are apparent after the NCEP system transitioned to use the real-time NARR and its precipitation gauge station data source in 2002. The UW system went to real-time operation in 2005. Thus, caution must be used when using model-derived drought indices from either of the systems for drought monitoring purposes. While both systems are able to capture the same drought events, the uncertainties

![Figure 8](image-url)
between both systems are too large to determine the drought category, especially over the interior of the western United States, unless care is taken to assure that the source of precipitation data is climatically consistent and that the fraction of stations reporting in real time is consistent with historic data availability.

Comparison of simulations using the two systems with different model forcings indicates that the primary cause

Fig. 9. (a) RMS difference of SMP for the experimental period (1979–2008) between Exp (F_{ncep}, P_{uw}) and Exp (F_{ncep}, P_{ncep}) (contour interval is given by the color bar); (b) as in (a), but for the difference between Exp (F_{uw}, P_{ncep}) and Exp (F_{uw}, P_{uw}); (c) as in (a), but for the difference between Exp (F_{uw}, P_{ncep}) and Exp (F_{ncep}, P_{ncep}); (d) as in (a), but for the difference between Exp (F_{ncep}, P_{uw}) and Exp (F_{uw}, P_{uw}); and (e)–(h) as in (a)–(d), but for SRI6.
of the differences in drought indicators is differences in the precipitation forcings. Post-2002, the availability of real-time $P$ data over the western United States became especially sparse for the NCEP system. In general, the effects of changes in $P$ station availability appeared to be lessened in the UW system because it relies on a set of index stations that have been relatively stable over the entire period of record for the system, which begins in 1915. While the NCEP precipitation forcings use all available station reports each day (potentially as many as 6000–8000 in the post-2002 real-time period), the UW system relies on a smaller number of about 2131 stations, which have more consistent climatologies, and are more likely to report on most days.

The systems also use different sources of surface air temperature, wind, humidity, and downward solar and longwave radiation (the UW system uses algorithms indexed to surface air temperature or its diurnal range, whereas the NCEP system takes these variables from NARR with the elevation adjustment). Differences in shortwave radiation between the two systems can be as large as 30 W m$^{-2}$ in summer. Nonetheless, experiments designed to isolate the effects of differences in variables other than $P$ showed that they had much smaller effects on SM and SRI than did $P$.

The differences among two systems with different $P$ forcing are not limited to the drought indices. The differences also appear in evapotranspiration and other variables. Therefore, users of the NLDAS need to be aware of the uncertainties in the NLDAS.

Acknowledgments. This project is supported by NOAA’s Climate Program Office Grant J8R1RP7-P01 to the University of Washington through its Climate Testbed project.

REFERENCES


