



# On the utility of land surface models for agricultural drought monitoring

W. T. Crow<sup>1</sup>, S. V. Kumar<sup>2,3</sup>, and J. D. Bolten<sup>2</sup>

<sup>1</sup>USDA Hydrology and Remote Sensing Laboratory, Beltsville, MD, USA

<sup>2</sup>NASA Goddard Space Flight Center, Greenbelt, MD, USA

<sup>3</sup>Science Applications International Corporation, Beltsville, MD, USA

Correspondence to: W. T. Crow (wade.crow@ars.usda.gov)

Received: 16 March 2012 – Published in Hydrol. Earth Syst. Sci. Discuss.: 19 April 2012

Revised: 21 August 2012 – Accepted: 1 September 2012 – Published: 24 September 2012

**Abstract.** The lagged rank cross-correlation between model-derived root-zone soil moisture estimates and remotely sensed vegetation indices (VI) is examined between January 2000 and December 2010 to quantify the skill of various soil moisture models for agricultural drought monitoring. Examined modeling strategies range from a simple antecedent precipitation index to the application of modern land surface models (LSMs) based on complex water and energy balance formulations. A quasi-global evaluation of lagged VI/soil moisture cross-correlation suggests, when globally averaged across the entire annual cycle, soil moisture estimates obtained from complex LSMs provide little added skill (< 5 % in relative terms) in anticipating variations in vegetation condition relative to a simplified water accounting procedure based solely on observed precipitation. However, larger amounts of added skill (5–15 % in relative terms) can be identified when focusing exclusively on the extra-tropical growing season and/or utilizing soil moisture values acquired by averaging across a multi-model ensemble.

proposed in an effort to improve the early detection of agricultural drought and the efficiency of subsequent mitigation responses (Wardlow et al., 2012). One common approach has been the application of complex water balance formulations embedded within land surface models (LSMs) to track temporal anomalies in root-zone soil water availability (Mo et al., 2010; Sheffield et al., 2012). These models typically include water and energy balance formulations based on time-varying meteorological and radiative forcing as well as detailed vertical soil physics to describe sub-surface soil water flux and storage. As a result, these “modern” LSMs implicitly promise an enhanced representation of root-zone soil water dynamics relative to soil moisture proxy products based solely on the simple accounting of antecedent precipitation. Recent work has also focused on the potential for improving soil moisture predictions by averaging across a multi-model ensemble comprised of various LSMs (Guo et al., 2007). Despite this potential, quantifying the marginal value of modern LSMs for global drought monitoring is challenging due to a lack of adequate large-scale root-zone soil water datasets available for evaluation purposes (Bolten et al., 2010).

Recently, Peled et al. (2010) proposed a novel approach for evaluating LSM soil moisture predictions by examining the cross-correlation between model-estimated root-zone soil moisture anomalies and spatially concurrent anomalies in vegetation indices derived from visible/near-infrared (VIS/NIR) remote sensing. The use of VIS/NIR vegetation indices (VI) like the enhanced vegetation index (EVI) and the normalized difference vegetation index (NDVI) is

## 1 Introduction

Agricultural drought is commonly defined as the lack of sufficient soil water availability to maintain adequate crop growth and pasture productivity (Panu and Sharma, 2002). The development of large-scale drought agricultural monitoring systems has received considerable attention in the past decade, and a range of remote sensing, ground observation, and land surface modeling techniques has been

well-established for monitoring the extent and severity of agricultural drought (Kogan, 1995; Peters et al., 2002; Ji and Peters, 2003). The potential of root-zone soil moisture monitoring lies in its ability to provide a leading indicator of subsequent VI anomalies (Adegoke and Carleton, 2002; Ji and Peters, 2005; Musyimi, 2010). That is, under water-limited conditions, a negative soil moisture anomaly should temporally precede a detectable impact on vegetation health and biomass. The analysis in Peled et al. (2010) is based on the assumption that the strength of soil moisture/VI cross-correlation can be used as a large-scale proxy for the accuracy of a model-based, root-zone soil moisture product.

Here we expand the geographic scope of Peled et al. (2010) (from the European continent to all global land between 60° S and 60° N) and evaluate a wider range of potential land surface modeling strategies. In particular, this analysis will employ various global LSMs, ranging from complex, modern LSMs to a simple antecedent precipitation index to sample lagged rank-correlations between model-estimated soil moisture and remotely sensed VI products. These cross-correlations will then be examined for evidence that higher-order water and energy processes captured by modern LSMs, but neglected in simple accounting procedures based solely on antecedent precipitation, add significant marginal utility to agricultural drought monitoring. The intent here is not so much to promote the use of highly simplified models, but rather to critically evaluate the value of processes contained in more complex models. In addition to evaluating stand-alone LSM predictions via comparisons with NDVI, the advantages of acquiring soil moisture products from a multi-model ensemble will also be quantified, and an attempt will be made to verify results via comparisons between model-based, root-zone soil moisture products and an independent, satellite-derived surface soil moisture product.

## 2 Models and data

The analysis is based on root-zone soil moisture products extracted from four separate models: version 3.2 of the National Centers for Environmental Prediction, Oregon State University; Air Force Weather Office and National Weather Service Hydrologic Research Laboratory model (Noah) (Ek et al., 2003; Mitchell, 2005; Barlarge et al., 2010); version 2.0 of the Common Land Model (CLM2.0) (Dai et al., 2003); the Catchment Land Surface Model (CLSM) (Koster et al., 2000; Ducharme et al., 2000); and, as an obviously simplified baseline approach, the antecedent precipitation index (API). The three “modern” LSMs (i.e., Noah, CLM2.0 and CLSM) were selected due to their availability within the NASA Land Information System (LIS) data assimilation test-bed which provides a framework for the integrated use of several community LSMs (Kumar et al., 2006). All models are run on a 0.25° grid between 1 January 2000 and 31 December 2010 for all global land area between 60° S and 60° N. Noah,

CLM2.0, and CLSM are run on a half-hourly time step while API calculations are based on a daily time step. For all models, a 1 January 2000 initialization is derived by separately looping each model through three integrations of this time period.

### 2.1 Soil moisture models

All three modern LSMs dynamically predict vertically discretized profile soil moisture based on a complex vertical representation of water flow within the soil column and a surface energy balance approach for the estimation of evapotranspiration. In addition to precipitation, modern LSMs require air temperature, air pressure, relative humidity, wind speed, and radiation (both shortwave and longwave) forcing data as input. Vertical soil water processes (e.g., infiltration and drainage) vary as a function of soil hydraulic properties typically tied to soil textural classifications through pedo-transfer functions. Energy balance processes depend strongly on land surface parameters like albedo, surface roughness, and leaf area index parameters typically specified as a function of vegetation class or climatological VI information. While the focus here is on the growing season, Noah, CLM2.0, and CLSM all contain snow modules which account for the accumulation, retention and melting of snow water storage.

Root-zone soil moisture is nominally defined as LSM-predicted soil moisture for the top 1 m of the soil column ( $\theta$ ). For this particular implementation, Noah uses four soil layers with thicknesses of 10, 30, 60, and 100 cm (descending from the surface), and CLM2.0 uses ten soil layers with thicknesses of 1.75, 2.76, 4.55, 7.5, 12.36, 20.38, 33.60, 55.39, 91.33, and 113.7 cm. Consequently, the top three Noah layers and top eight CLM2.0 layers are averaged (using relative weights equal to the ratio of each layer thickness to the 1-m total root-zone depth) to obtain an integrated root-zone soil moisture product. The Catchment LSM, by contrast, is non-traditional in that the vertical soil moisture profile is diagnosed as a net variation in water storage relative to an equilibrium soil moisture profile calculated between the surface and the top of the water table. In this way, CLSM calculates soil moisture within both a 2-cm surface layer and a 1-m root-zone layer (Koster et al., 2000). All three modern LSMs are run on a half-hourly time step continuously throughout the year.

Finally, an API-based root-zone soil water proxy ( $\theta_{\text{API}}$ ) is calculated as a linear combination of the previous day's value ( $\theta_{\text{API},j-1}$ ) and accumulated precipitation (in mm) for the current day ( $P_j$ ):

$$\theta_{\text{API},j} = \gamma \theta_{\text{API},j-1} + P_j \quad (1)$$

where the constant parameter  $\gamma$  controls the effective memory of API levels to past rainfall accumulations. Unlike the modern LSMs described above, the API model neglects variations in root-zone soil moisture storage due to surface energy balance processes, the vertical and/or lateral movement

of water between multiple soil moisture states, and the impact of snow melt on soil water availability.

## 2.2 Model forcing data

The modern LSMs are driven by two separate forcing datasets which provide fine-scale (hourly to three-hourly) values of precipitation, insolation, air temperature, humidity, wind speed, and air pressure. The first dataset is derived from the Global Data Assimilation System (GDAS) obtained from the weather forecast model of the National Centers for Environmental Prediction (Derber et al., 1991). In order to mitigate known biases, GDAS precipitation fields are rescaled to match coarse-resolution ( $2.5^\circ$ , 5-day) precipitation accumulation totals from the NOAA Climate Prediction Center's (CPC) operational global Merged Analysis of Precipitation (CMAP) product (Xie and Arkin, 1997) which blends satellite and rain-gauge observations.

While this "GDAS + CMAP" product is representative of currently available global LSM forcing datasets, higher-quality forcing datasets are available in selected continental areas. To reflect this, the modern LSMs are also forced with the North American Land Data Assimilation System Version 2 (NLDAS-2) forcing dataset (Xia et al., 2012) within a regional domain centered on the contiguous United States (CONUS) ( $25.75^\circ$ – $52^\circ$  N,  $124^\circ$ – $68.75^\circ$  W). Relative to GDAS + CMAP, the NLDAS-2 dataset is based on regional (as opposed to global) reanalysis products and leverages a greater abundance of ground- and satellite-based observational resources. In particular, NLDAS-2 precipitation is based on the merger of daily CPC rain gauge accumulations, ground-based radar estimates, satellite-based retrievals, and North American Regional Reanalysis (NARR) precipitation fields (Cosgrove et al., 2003). Incoming longwave and shortwave radiation estimates are taken from the NASA/GEWEX Surface Radiation Budget (SRB) dataset and geostationary satellite observations. Remaining NLDAS-2 forcing variables (e.g., air temperature, wind speed, relative humidity, and air pressure) are based on NCEP North American Regional Reanalysis (NARR).

Modern LSMs generally use remotely sensed VI information to estimate either vegetation fractional coverage (Noah) or leaf area index (CLM2.0 and CLSM). Unless otherwise noted, all such parameters are derived from climatological VI information derived using long-term Advanced Very High Resolution Radiometer (AVHRR) surface reflectance products. Since they lack inter-annual variability, the use of climatological VI information as LSM input eliminates the risk of cross-correlated errors between LSM soil moisture estimates and inter-annual variations in VI products used for the purpose of evaluation. Additional land cover information is derived from the 1-km University of Maryland land cover classification (Hansen et al., 2000). Soil texture is obtained by merging the global Foreign Agricultural Office (FAO) soil classification product with the State Soil Geographic

(STATSGO) database within CONUS. In contrast, the API model is forced solely by daily (0:00–24:00 UTC) precipitation accumulations acquired by temporally aggregating sub-daily GDAS + CMAP and/or NLDAS-2 precipitation accumulations. The single parameter  $\gamma$  is assumed to be a fixed global constant (see below).

## 2.3 Evaluation data

The primary strategy for evaluating model-based soil moisture will be comparison against satellite-derived NDVI derived from the monthly Moderate Resolution Imaging Spectroradiometer (MODIS) MOD13C2 composite product (Collection 5) between February 2000 and December 2010. Only reliable MODIS VI retrievals categorized as "Good data – use with confidence" in the MOD13C2 pixel reliability field are included in the analysis and spatially aggregated to match the  $0.25^\circ$  LSM modeling grid.

In an attempt to verify NDVI-based results, model-based soil moisture estimates will also be compared to  $0.25^\circ$  Advanced Microwave-Scanning Radiometer-EOS (AMSR-E) surface soil moisture retrievals derived via the Land Parameter Retrieval Model (LPRM) (Owe et al., 2008) by Richard de Jeu and colleagues at VU University Amsterdam. Comparisons are made with LPRM AMSR-E retrievals obtained between July 2002 to December 2011. Note that retrievals are masked during periods of dense vegetation cover, heavy precipitation, and/or frozen surface condition, and assumed to reflect surface soil moisture conditions only in the top 1–3 cm of the soil column.

## 3 Analysis

The analysis is based on the assumption that higher-quality root-zone soil moisture datasets will exhibit stronger lagged correlations with future VI anomalies (Peled et al., 2010). However, secondary characteristics like climatological seasonality, distribution shape, and temporal auto-correlation can also impact soil moisture/VI cross-correlation. In order to minimize these effects, results are based on rank correlations sampled after the transformation of both raw VI and soil moisture data into a monthly rank time series and the standardization of soil moisture auto-correlation functions. See below for a description of this processing.

### 3.1 Rank time series calculation

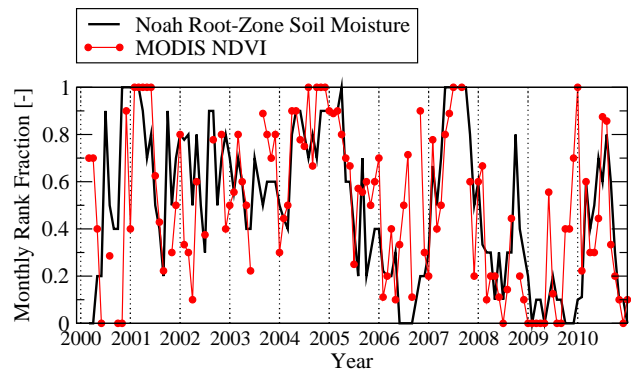
To begin, every model-based root-zone soil moisture product  $\theta$  is aggregated to create a monthly time series  $\bar{\theta}_i$  from January 2000 to December 2010 for each  $0.25^\circ$  land pixel between  $60^\circ$  S and  $60^\circ$  N. Next,  $\bar{\theta}_i$  for a single month  $i$  (and a single  $0.25^\circ$  land pixel) are ranked across all 11 occurrences of the same month of the year between 2000 and 2010 for the same pixel. As a result, the  $\bar{\theta}$  time series is transformed into a monthly time series of ranks – or  $\text{Rank}(\bar{\theta})_i$  – which

reflect the relative wetness of a particular month  $i$  relative to the same month of the year during all other years. The same ranking procedure is applied to monthly NDVI to create  $\text{Rank}(\overline{\text{NDVI}})_i$  and monthly-averaged LPRM AMSR-E surface soil moisture retrievals to create  $\text{Rank}(\overline{\text{LPRM}})_i$ . This rank transformation accomplishes two key objectives. First, it removes the seasonal cycle from each product so that the analysis focuses solely on inter-annual variations. Second, it ensures a consistent distribution for variables in the cross-correlation analysis and minimizes the potential impact of outliers. The use of a monthly time scale is intended as a compromise between minimizing the temporal resolution of the analysis while maximizing the spatial coverage and completeness of composited VI products.

Figure 1 shows example times series of monthly  $\text{Rank}(\bar{\theta}_{\text{Noah}})$  and  $\text{Rank}(\overline{\text{NDVI}})$  for a single  $0.25^\circ$  pixel in the South Central United States. Formally, the y-axis describes the fractional rank of month  $i$  relative to the same month of the year found in other years of the 2000 to 2010 time period (i.e., the fraction of the same month of the year in different years with lower  $\bar{\theta}$  or lower  $\overline{\text{NDVI}}$ ). Periodic gaps in the NDVI time series reflect months where MODIS-based NDVI products are deemed unreliable.

### 3.2 Rank auto-correlation analysis and standardization

Despite the fact that Noah, CLM2.0, and CLSM root-zone products are all defined to provide top-1-m soil moisture products, differences in evapotranspiration and soil water parameterizations between models can induce variations in the effective persistence of soil moisture anomalies. Such differences can, in turn, impact sampled soil moisture/VI cross-correlation. Since we hope to interpret such correlations as being reflective solely of skill in model-based soil moisture estimates, large variations in soil moisture auto-correlation can potentially confound our analysis. For example, otherwise skillful variations in profile soil moisture estimates from a given model could be unfairly evaluated if sampled at a depth which did not adequately express their true cross-correlation with NDVI. To address this, the auto-correlation function of  $\text{Rank}(\bar{\theta})_i$  – or  $\rho(L)$  – is standardized across all models prior to further cross-correlation analysis. With this goal in mind, Fig. 2a plots quasi-global averages (i.e., land areas between  $60^\circ\text{S}$  and  $60^\circ\text{N}$ ) of  $\rho(L)$  for root-zone soil moisture estimates from Noah, CLM2.0, and CLSM. For top-1-m LSM results, sampled CLM2.0 and CLSM  $\rho(L)$  match relatively closely and will therefore be left unmodified. However, 1-m Noah results show significantly more temporal auto-correlation. Consequently, all subsequent Noah results are instead based on a shallower vertical integration of soil moisture (i.e., top 40 cm versus top 1 m). Unlike the original 1-m results, 40-cm Noah soil moisture  $\rho(L)$  results provide a close match to 1-m CLM2.0 and CLSM  $\rho(L)$  results. Also note that all sampled  $\rho(L)$  functions show considerable temporal auto-correlation at lags of  $\pm 1$ -month – suggesting that



**Fig. 1.** Example monthly  $\text{Rank}(\overline{\text{NDVI}})$  and  $\text{Rank}(\bar{\theta}_{\text{Noah}})$  time series for a  $0.25^\circ$  pixel in the South Central United States. A rank of zero (one) means that a given month has a lower (higher) soil moisture/NDVI than all other occurrences of its month of the year within the 2000–2010 time period.

a monthly time scale represents a reasonable temporal support for capturing root-zone soil moisture dynamics.

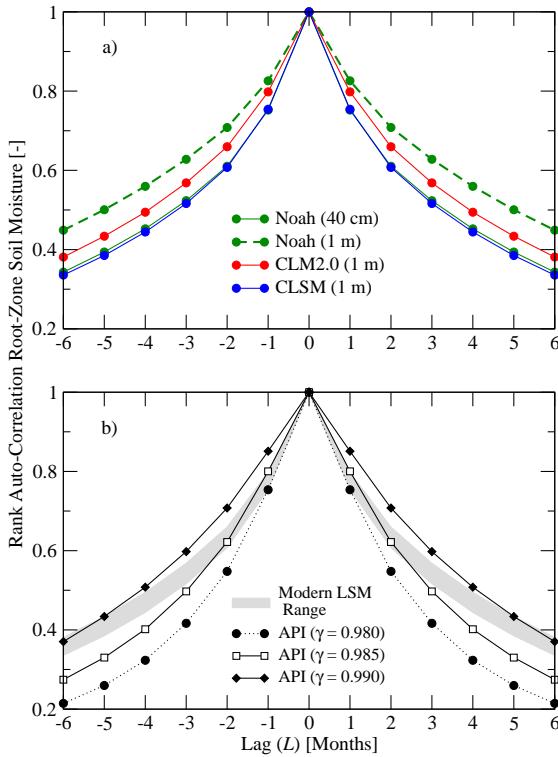
API results are based on calibrating  $\gamma$  in Eq. (1) to produce quasi-globally averaged API-based  $\rho(L)$  results which approximate that of the modern LSMs. However, due to differences in the shape of API's  $\rho(L)$  function relative to the modern LSMs' functions, there is some ambiguity in this calibration. Figure 2b illustrates this effect by comparing quasi-globally averaged API  $\rho(L)$  for  $\gamma = 0.98, 0.985,$  and  $0.99$  to the absolute range of  $\rho(L)$  results for Noah, CLM2.0, and CLSM. Note that  $\gamma = 0.98$  represents a plausible fit to the modern LSM range for  $|L| = 1$  but drifts badly for larger  $|L|$ . Conversely,  $\gamma = 0.99$  is adequate at large  $|L|$  but poor for small  $|L|$ . While the middle choice of  $\gamma = 0.985$  minimizes misfits over the entire range of  $L$ , it still performs badly at large  $|L|$ . Unless otherwise noted, all future API results will be for the middle case  $\gamma = 0.985$ . However, given the ambiguity noted in Fig. 2b, the sensitivity of key API results to  $\gamma$  will also be noted.

### 3.3 Ensemble-mean product

As described above, a final soil moisture product is created by averaging across soil moisture results within a multi-model ensemble. This product is based on transforming each of the three monthly, modern LSM root-zone soil moisture products ( $\bar{\theta}_{\text{Noah}}$ ,  $\bar{\theta}_{\text{CLM2.0}}$ , and  $\bar{\theta}_{\text{CLSM}}$ ) into standard normal deviates:

$$\bar{\theta}'_i = \frac{\bar{\theta}_i - \mu_{\bar{\theta}}}{\sigma_{\bar{\theta}}} \quad (2)$$

where  $\mu_{\bar{\theta}}$  and  $\sigma_{\bar{\theta}}$  are the sampled mean and standard-deviation, respectively, for each  $\bar{\theta}$  product during all occurrences of the month of the year associated with month  $i$ . Next, all three anomaly products are averaged to create a monthly ensemble-averaged product:



**Fig. 2.** Quasi-global land averages of  $\rho(L)$  for (a) modern LSMs (i.e., Noah, CLM2.0, and CLSM) and (b) various API cases. Noah results in (a) are shown for both a 40-cm and 1-m root-zone depth case. “Modern LSM Range” shading in (b) is defined as the absolute range of 40-cm Noah, 1-m CLM2.0, and 1-m CLSM  $\rho(L)$  results in part (a).

$$\bar{\theta}'_{ENS,i} = \frac{1}{3} \left( \bar{\theta}'_{Noah,i} + \bar{\theta}'_{CLM2.0,i} + \bar{\theta}'_{CLSM,i} \right). \quad (3)$$

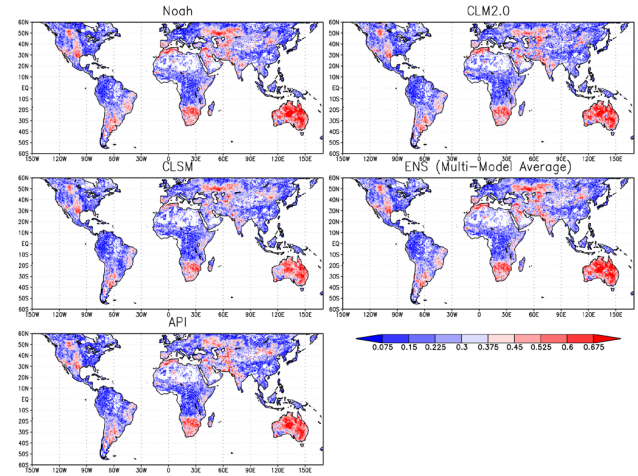
The resulting time series of  $\bar{\theta}'_{ENS}$  are then ranked to create  $\text{Rank}(\bar{\theta}_{ENS})$ . Note that the anomaly notation is dropped when referring to this rank product since  $\text{Rank}(\bar{\theta}_{ENS}) = \text{Rank}(\bar{\theta}'_{ENS})$ .

### 3.4 Rank cross-correlation calculation

The rank cross-correlation  $R(L)$  at lag  $L$  between NDVI and all five root-zone soil moisture rank products (i.e.,  $\text{Rank}(\bar{\theta}_{Noah})$ ,  $\text{Rank}(\bar{\theta}_{CLM2.0})$ ,  $\text{Rank}(\bar{\theta}_{CLSM})$ ,  $\text{Rank}(\bar{\theta}_{ENS})$ , and  $\text{Rank}(\bar{\theta}_{API})$ ) is calculated as the sampled correlation coefficient between  $\text{Rank}(\bar{\theta})_{i+L}$  and  $\text{Rank}(\overline{NDVI})_i$  over all possible  $i$ . Based on this definition,  $R(L)$  for  $L < 0$  relates the ability of *current* soil moisture conditions to forecast *future* NDVI. Since the Fischer transformation,

$$F(R) = \frac{1}{2} \ln \left( \frac{1+R}{1-R} \right), \quad (4)$$

of sampled  $R$  yields a normal distribution with variance  $1.06/(n-3)$  (Fieller et al., 1957),  $Z$ -scores for  $R(-1)$  differences between an LSM and API can be calculated as:



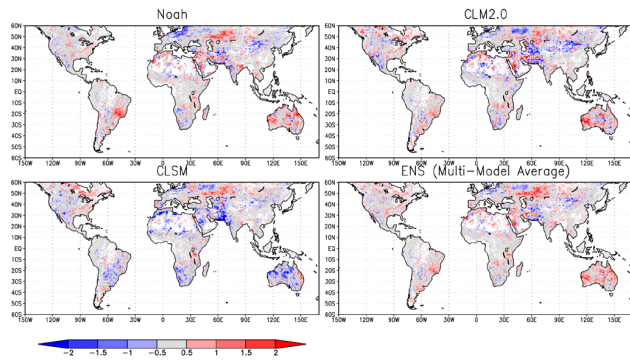
**Fig. 3.** Quasi-global map of  $0.25^\circ R(-1)$  for the Noah, CLM2.0, CLSM, ENS, and API cases.

$$Z = \sqrt{\frac{n-3}{2 \cdot 1.06}} \left( F[R_{LSM}(-1)] - F[R_{API}(-1)] \right) \quad (5)$$

where  $n$  is taken to be the number of months sampled to obtain  $R(-1)$ . Note that, since Eq. (5) neglects the impact of temporal auto-correlation in both  $\text{Rank}(\overline{NDVI})$  and  $\text{Rank}(\bar{\theta})$ , these  $Z$ -scores are likely not appropriate for formal hypothesis testing. Nevertheless, they represent a useful tool for standardizing observed model-to-model differences.

It is important to note that positive  $R(L)$  is not expected for all biomes or land cover types. For example, in energy-limited areas, relatively dry periods may be associated with enhanced VI due to reduced cloudiness (Huete et al., 2006). In these areas, an increase in  $R(L)$  (i.e., making it less negative) cannot be reliably linked to improved soil moisture skill. Therefore, all  $0.25^\circ$  pixels in which the null hypothesis  $R(-1) >= 0$  can be rejected (at 80% significance) for *any* model product are subsequently masked from the entire analysis. In order to minimize the impact of cold-season conditions, months with an average daily high air temperature below  $5^\circ\text{C}$  are also removed (on a month-by-month and pixel-by-pixel basis).

This cross-correlation analysis will be repeated for the case of replacing the target variable  $\text{Rank}(\overline{NDVI})$  with  $\text{Rank}(\overline{LPRM})$  and sampling the equivalent rank cross-correlation  $R^{LPRM}(L)$ . All spatial and temporal rank procedures described above (for the NDVI analysis) are adhered to. In addition, since temporal gaps exist in the LPRM surface soil moisture retrieval product (e.g., due to snow cover, dense vegetation and/or frozen soil), all months lacking at least five separate LPRM retrievals are masked from the calculation of  $R^{LPRM}(L)$ . Note that  $R^{LPRM}(L)$  results play a secondary role in the analysis and are used only to verify NDVI-based  $R(L)$  results.



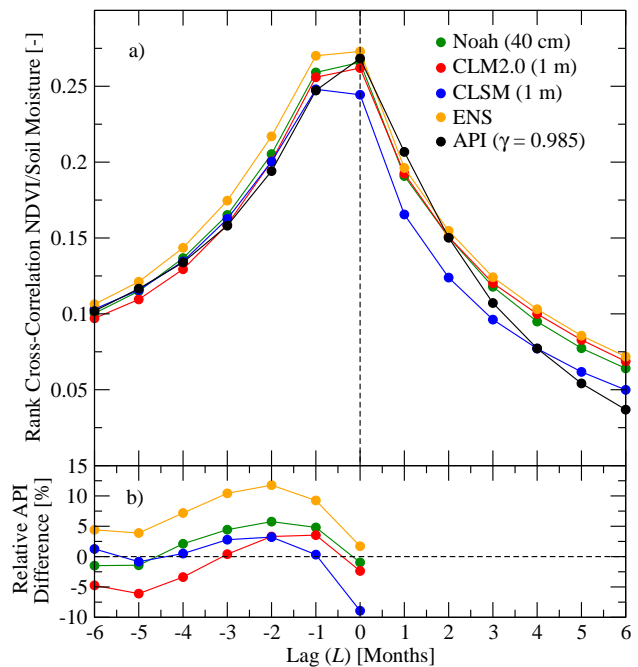
**Fig. 4.** Quasi-global map of  $0.25^\circ$  Z-scores for the difference  $R(-1)_{\text{LSM}}$  minus  $R(-1)_{\text{API}}$  where LSM estimates are obtained from Noah, CLM2.0, CLSM, and ENS.

## 4 Results

For the case  $L = -1$  (i.e.,  $\text{Rank}(\bar{\theta})$  temporally precedes  $\text{Rank}(\overline{\text{NDVI}})$  by 1 month), Fig. 3 plots global  $0.25^\circ$  Noah, CLM2.0, CLSM, ENS and API  $R(-1)$  results. White masked areas represent a combination of open water surfaces, areas with non-significant positive  $R(-1)$  (see above), and barren areas where no temporal NDVI variability is observed. Substantial coupling ( $R(-1) > 0.50$ ) is found in semi-arid areas of the world prone to water-limited plant growth (e.g., Australia, Southern Africa, and the Western United States). Conversely, humid areas of the Eastern United States, Europe, and Southeastern Asia demonstrate weak soil moisture/VI cross-correlation ( $R(-1) < 0.20$ ). A secondary cause of low  $R(-1)$  is poor accuracy in model-based soil moisture predictions. For example, low sampled  $R(-1)$  in arid regions of Sub-Saharan Africa are likely caused by inadequate rain gauge coverage which prevents LSMs from accurately capturing relative soil moisture variations in data-poor regions.

Figure 4 examines model-to-model differences in performance between models by plotting spatially distributed Z-scores for sampled  $R(-1)$  differences between the four approaches based on modern LSM simulations (i.e., Noah, CLSM, CLM2.0, and ENS) and the API baseline. While regions of significantly improved NDVI forecasting (relative to API) exist in Noah, CLM2.0, and CLSM predictions (i.e., positive Z-scores indicated by red shading in Fig. 4), they are balanced by areas where API-based soil moisture products are superior (i.e., negative Z-scores indicated by blue shading in Fig. 4). Only the multi-model ENS case appears to consistently improve upon the API baseline.

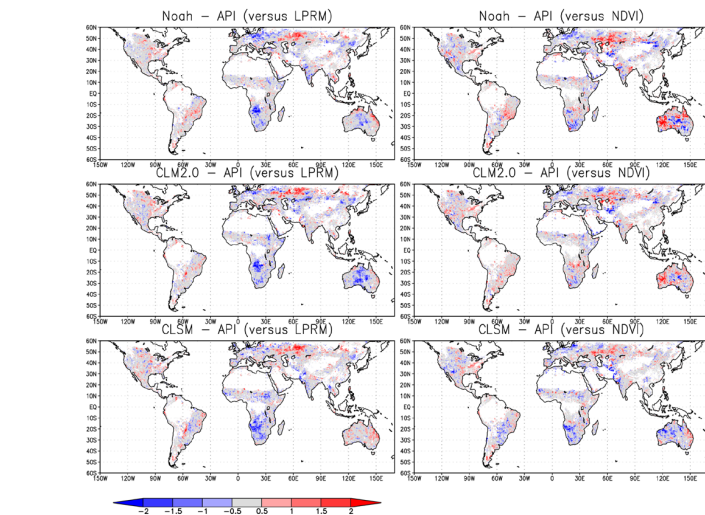
Figure 5a compares modeling results on a quasi-global scale by plotting average  $R(L)$  across all unmasked land areas in Fig. 3 for a range of  $L$ . Sampled  $R(L)$  functions are not symmetric with respect to  $L = 0$ , and instead are larger for  $L < 0$ . This lack of symmetry underscores the predictive role for soil moisture where the largest  $R(L)$  is sampled for



**Fig. 5.** Quasi-global land averages of (a)  $R(L)$  for Noah, CLM2.0, CLSM, ENS, and API root-zone soil moisture predictions, and (b) percentage relative  $R(L)$  difference for the modern LSMs (Noah, CLM2.0, CLSM, and ENS) versus an API baseline (i.e.,  $100 \cdot [R(L)_{\text{LSM}} - R(L)_{\text{API}}]/R(L)_{\text{API}}$ ).

the  $L < 0$  case in which  $\text{Rank}(\bar{\theta})_{i+L}$  precedes  $\text{Rank}(\overline{\text{NDVI}})_i$ . Using Eqs. (4) and (5), error bars can be constructed for individual points in Fig. 5. However, even if conservative reductions in effective degrees of freedom are made to account for potential spatial and temporal autocorrelation in  $\text{Rank}(\overline{\text{NDVI}})$  and  $\text{Rank}(\bar{\theta})$ ,  $1\sigma$  sampling uncertainty associated with these quasi-global averages of  $R(-1)$  remains on the order of 0.005 [-] to 0.001 [-] and therefore smaller than the size of plotted symbols in Fig. 5. Consequently, it is safe to assume that all visible differences in plotted  $R(L)$  are significant at a  $1\sigma$  certainty level.

Nevertheless, among the stand-alone models, the relative magnitude of model-to-model variations is small. For  $L < 0$ , Noah, CLM2.0, and CLSM results are associated with  $R(L)$  values that fall within about  $\pm 5\%$  of baseline API results (Fig. 5b). That is, none of the stand-alone modern LSMs demonstrates any substantial advantage over API in anticipating the near-term impact of agricultural drought on NDVI anomalies. However, using a multi-model ensemble average acquired from Eq. (3) leads to a larger (and more consistent) amount of improvement relative to the API baseline (see Fig. 5b). As a result, the only viable method for consistently increasing  $R(L)$  via modified model physics appears to lie in the use of multi-model ensembles. It should be noted that ENS results here are based on a very small ensemble size and



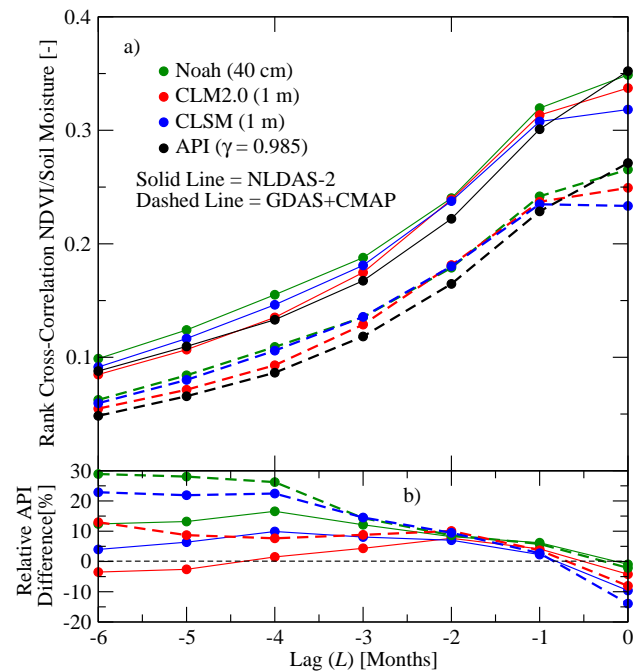
**Fig. 6.** Quasi-global map of  $0.25^\circ$  Z-scores for the difference  $R^{\text{LPRM}}(-1)_{\text{LSM}}$  minus  $R^{\text{LPRM}}(-1)_{\text{API}}$  (use of LPRM surface soil moisture retrievals as target variable; left column) and  $R(-1)_{\text{LSM}}$  minus  $R(-1)_{\text{API}}$  (use of NDVI as target variable; right column) where LSM estimates are obtained from Noah, CLM2.0, and CLSM.

would almost certainly be enhanced by considering a larger number of LSMs.

Utilizing EVI as the target VI (not shown) produces a qualitatively similar plot except sampled  $R(L)$  values are somewhat lower than those found using NDVI for all modeling cases. Likewise, API  $R(L)$  results in Fig. 5 are slightly improved when using values of  $\gamma$  along the upper edge of the feasible range identified in Fig. 2b; however, the overall effect is very small (on the order of a  $\sim 2\%$  relative increase in  $R(L)$  for  $-2 \leq L \leq 0$ ).

#### 4.1 Comparisons with satellite-based surface soil moisture

Figure 6 replicates  $R(-1)$  results previously shown in Fig. 4 (right column), but also adds  $R^{\text{LPRM}}(-1)$  results based on comparisons between LPRM surface soil moisture retrievals derived from AMSR-E observations (left column). As in Fig. 4, Fig. 6 shows Z-scores for differences in rank correlations achieved using modern LSMs versus an API baseline. Results in Fig. 6 are based on the time period July 2002 to December 2010 and areas in which adequate LPRM retrievals are available after masking of densely vegetated and frozen surfaces. As before, red shading indicates regions where modern LSMs add skill to the API baseline (by providing a better predictor of either satellite-based surface soil moisture retrievals or NDVI), and blue shading indicates areas where modern LSMs perform worse than the API baseline. Results in Fig. 6 indicate relatively strong consistency between results based on comparisons against LPRM and NDVI. In particular,  $R^{\text{LPRM}}(-1)$  differences (relative to API) mirror earlier  $R(-1)$  results in suggesting that none of



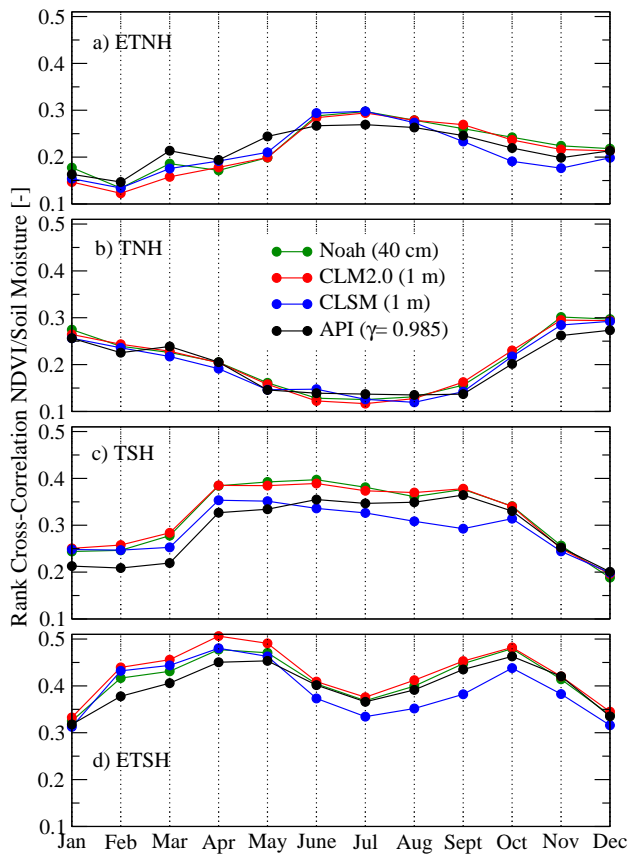
**Fig. 7.** CONUS land averages of (a)  $R(L)$  for Noah, CLM2.0, CLSM, and API root-zone soil moisture predictions based on either the GDAS + CMAP or NLDAS-2 forcing datasets, and (b) percentage relative  $R(L)$  percentage difference for the modern LSMs (Noah, CLM2.0, and CLSM) versus an API baseline (i.e.,  $100 \cdot [R(L)_{\text{LSM}} - R(L)_{\text{API}}]/R(L)_{\text{API}}$ ).

the modern LSMs examined provides a clear global advantage over API. Patterns of relative degradation/improvement versus an API baseline are also roughly comparable for  $R^{\text{LPRM}}(-1)$  and  $R(-1)$  difference results (see e.g., Noah results in the first row of Fig. 6).

This general consistency between results obtained from independent verification strategies lends credibility to our earlier interpretation of  $R(-1)$  as a reliable performance metric for LSM soil moisture estimates. Nevertheless, two caveats should be noted for  $R^{\text{LPRM}}(-1)$  results. First, they are based on the (vertically inconsistent) comparison of surface soil moisture retrievals with root-zone soil moisture model estimates. Second, the use of  $L = -1$  [month] as the lag scale lacks any clear physical justification and is done simply for consistency with earlier NDVI-based  $R(-1)$  results.

#### 4.2 Impact of LSM forcing data

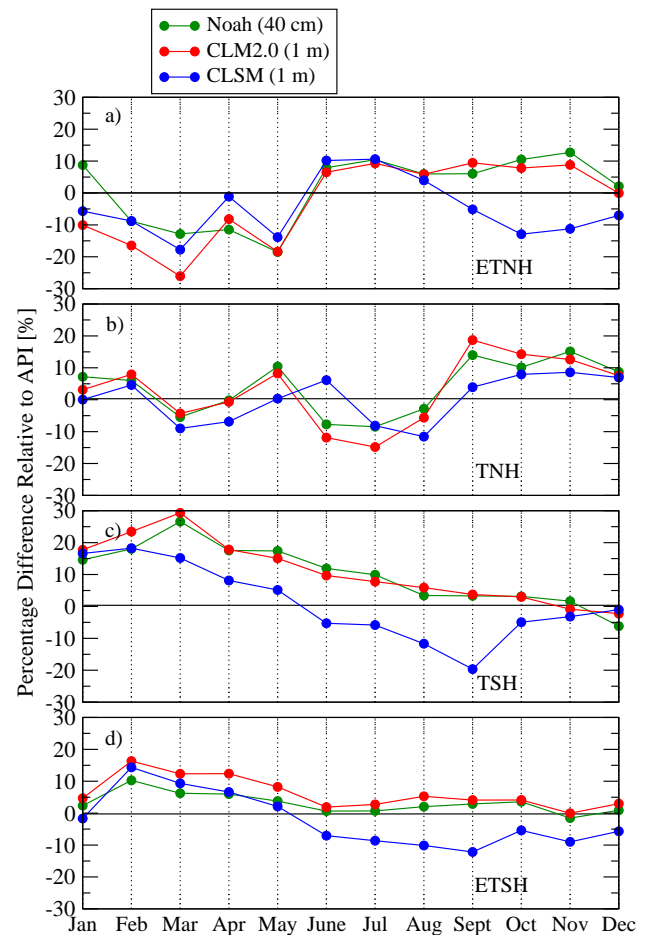
Since modern LSMs attempt to exploit temporal variations in non-rainfall-based forcing (e.g., air temperature and insolation) to better predict soil moisture anomalies, one factor impacting the performance spread between modern LSMs and an API baseline may be the quality of non-rainfall forcing data. Figure 7 shows the impact of replacing the global GDAS + CMAP forcing dataset with the higher-quality (but non-global) NLDAS-2 dataset. Dashed lines in Fig. 7a show



**Fig. 8.** Spatial averages of  $R(-1)$  for Noah, CLM2.0, CLSM, and API broken down by month of the year for Rank( $\bar{\theta}$ ) within the (a) extra-tropical Northern Hemisphere (ETNH), (b) tropical Northern Hemisphere (TNH), (c) tropical Southern Hemisphere (TSH), and (d) extra-tropical Southern Hemisphere (ETSH).

CONUS GDAS+CMAP  $R(L)$  results for each model and solid lines NLDAS-2 results for the same CONUS domain. For clarity, ENS results are omitted. The transition between GDAS+CMAP and NLDAS-2 forcing clearly improves the performance of the models. However, nearly all of this improvement is attributable to improved rainfall since there is no discernible improvement in modern LSM results relative to API (Fig. 7b). In fact, for  $L < -4$ , utilizing NLDAS-2 forcing actually degrades the quality of the modern LSM predictions relative to API (Fig. 7b). Consequently, there is no evidence that enhancing the quality of non-rainfall forcing data improves the performance of modern LSMs relative to an API baseline. Note that the focus here is on the impact of errors in dynamic model forcings, as opposed to static inputs like soil texture, since inter-annual model variability, and thus Rank( $\bar{\theta}$ ) time series, is more strongly impacted by time-varying model inputs.

A related issue is the use of a long-term NDVI climatology to estimate LSM vegetation parameters. The neglect of inter-annual variability in this climatology could conceivably impair the ability of LSMs to accurately characterize relative



**Fig. 9.** Spatial averages of relative change in Noah, CLM2.0, and CLSM  $R(-1)$  results versus the API baseline (i.e.,  $100 \cdot [R(-1)_{\text{LSM}} - R(-1)_{\text{API}}] / R(-1)_{\text{API}}$ ) broken down by month of the year for Rank( $\bar{\theta}$ ) within the (a) extra-tropical Northern Hemisphere (ETNH), (b) tropical Northern Hemisphere (TNH), (c) tropical Southern Hemisphere (TSH), and (d) extra-tropical Southern Hemisphere (ETSH).

variations in root-zone soil moisture availability. To test for this possibility, Noah results were duplicated for the case of obtaining fractional vegetation cover from monthly NDVI composites (as opposed to an NDVI climatology). Surprisingly, this inclusion of inter-annual variability in LSM vegetation parameters leads to a slight reduction in globally averaged Noah  $R(-1)$  and  $R^{\text{LPRM}}(-1)$  results. While the reason(s) for this reduction are unclear, it suggests that our use of an NDVI climatology does not significantly degrade the performance of modern LSMs.

### 4.3 Impact of seasonality

Large seasonal variability in soil water availability, and thus  $R(L)$ , is also expected in certain climate zones. To examine such seasonal variability, Fig. 8 plots spatially averaged



$R(-1)$  within various latitude bands according to the month of the year for Rank( $\theta$ ) obtained using GDAS + CMAP forcings. To maximize the spatial consistency of sampled  $R(-1)$  for different months, the monthly air temperature mask (see Sect. 3) is not applied here. Observed monthly trends in Fig. 8 conform well to expected seasonal patterns. For instance, in the extra-tropical Northern Hemisphere (ETNH; Fig. 8a), the highest soil moisture/vegetation coupling, and thus sampled  $R(-1)$ , occurs during the boreal summer when root-zone soil moisture is generally minimized. Likewise, seasonal  $R(-1)$  trends in tropical regions (Fig. 8b and c) reflect the expected progression of the tropical rain belt with relatively lower  $R(-1)$  found during the rainy seasons for both the tropical Northern Hemisphere (TNH; May to October) and tropical Southern Hemisphere (TSH; November to April).

Figure 9 mirrors Fig. 5b by plotting results in Fig. 8 in terms of percentage variation versus an API baseline. Despite relatively modest model-to-model variability in Fig. 8, several potential trends can be noted. For example,  $R(-1)$  for Noah and CLM2.0 consistently improves upon the API baseline during mid-to-late portions of the ETNH (see June to November results in Fig. 9a) and ETSH (see February to April in Fig. 9d) growing seasons. The enhanced importance of evapotranspiration to the soil water balance in these periods may increase the value of energy balance calculations made by modern LSMs. Likewise, during the end of both the TNH rainy season (September to November in Fig. 9b) and the TSH rainy season (January to March in Fig. 9c), all three modern LSMs appear to maintain some advantage relative to API.

## 5 Conclusions

Given the wide variety of remote sensing, ground observation, and modeling strategies currently being proposed for global agricultural drought monitoring (Wardlow et al., 2012), it is important to define benchmarking strategies capable of objectively evaluating the relative merits of each. Here, we quantify the added benefit of modern LSMs for anticipating future vegetation health and biomass anomalies relative to a baseline case of utilizing a much simpler antecedent precipitation index (API). Unlike API, modern LSMs offer a complex parameterization of the surface energy balance and detailed vertical water balance physics in an attempt to more accurately characterize temporal variations in root-zone soil moisture availability. However, when objectively evaluated at global scales over the entire seasonal cycle, modern LSMs offer little relative advantage versus an API baseline in terms of anticipating the impact of agricultural drought on vegetation condition (Figs. 3, 4, and 5) and/or matching independent surface soil moisture retrievals obtained from satellite remote sensing (Fig. 6). The relative utility of modern LSMs versus API is not enhanced by improving the quality

of LSM forcing data (Fig. 7). Taken as a whole, results suggest that non-rainfall forcing data and modern LSM energy balance calculations currently contribute relatively little towards the accuracy of agricultural drought monitoring systems. As such, results are broadly consistent with past work in Abramowitz et al. (2008) questioning the general utility of existing LSM energy balance calculations. However, clear additive value does emerge when root-zone soil moisture estimates obtained from various models (including multiple modern LSMs) are merged into a single ensemble-mean prediction (Figs. 3, 4, and 5). In addition, more added value (around 5 % to 15 % in relative terms) for modern LSMs is found during specific points along the seasonal cycle – particularly during middle to late portions of the extra-tropical growing season (Figs. 8 and 9).

In summary, it should be stressed that modern LSMs possess a range of functionality (e.g., data assimilation capabilities, compatibility with crop growth models, and compatibility with atmospheric models) that cannot be duplicated by simpler models like API. In addition, even if their contribution to off-line drought prediction is marginal, modern LSMs possess substantial utility by describing the subsequent impact of drought conditions on water balance processes like ET and runoff. Therefore, results here should not be interpreted as advocating the use of API in operational drought monitoring systems. However, for the specific objectives examined here, complex LSM treatment of surface water and energy balance processes contributes relatively little marginal utility above and beyond that obtainable from highly simplified modeling approaches. This, in turn, implies additional model development and benchmarking work are required before modern LSMs can meet their full potential as agricultural drought monitoring tools.

*Acknowledgements.* Research was supported by NASA Applied Sciences Grant entitled “Enhancing the USDA Global Crop Production Decision Support System with the NASA Land Information System and Water Cycle Satellite Observations” (W. T. Crow – Principal Investigator). Computing was partially supported by the resources at the NASA Center for Climate Simulation.

Edited by: H. -J. Hendricks Franssen

## References

- Abramowitz, G., Leuning, R., Clark, M., and Pitman, A.: Evaluating the performance of land surface models, *J. Climate*, 21, 5468–5481, 2008.
- Adegoke, J. O. and Carleton, A. M.: Relations between soil moisture and satellite vegetation indices in the U.S. Corn Belt, *J. Hydrometeorol.*, 3, 395–405, 2002.
- Barlage, M., Chen, F., Tewari, M., Ikeda, K., Gochis, D., Dudhia, J., Rasmussen, R., Livneh, B., Ek, M., and Mitchell, K.: Noah land

- surface model modifications to improve snowpack prediction in the Colorado Rocky Mountains, *J. Geophys. Res.*, 115, D22101, doi:10.1029/2009JD013470, 2010.
- Bolten, J. D., Crow, W. T., Jackson, T. J., Zhan, X., and Reynolds, C. A.: Evaluating the utility of remotely-sensed soil moisture retrievals for operational agricultural drought monitoring, *IEEE J. Sel. Top. Appl.*, 3, 57–77, doi:10.1109/JSTARS.2009.2037163, 2010.
- Cosgrove, B. A., Lohmann, D., Mitchell, K. E., Houser, P. R., Wood, E. F., Schaake, J., Robock, A., Marshall, C., Sheffield, J., Luo, L., Duan, Q., Pinker, R. T., Tarpley, J. D., Higgins, R. W., and Meng, J.: Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project, *J. Geophys. Res.*, 108, 8842, doi:10.1029/2002JD003118, 2003.
- Dai, Y., Zeng, X., Dickinson, R. E., Baker, I., Bonan, G., Bosilovich, M., Denning, S., Dirmeyer, P., Houser, P., Niu, G., Oleson, K., Schlosser, A., and Yang, Z.-L.: The common land model (CLM), *B. Am. Meteorol. Soc.*, 84, 1013–1023, doi:10.1175/BAMS-84-8-1013, 2003.
- Derber, J., Parrish, D., and Lord, S.: The new global operational analysis system at the National Meteorological Center, *Weather Forecast.*, 6, 538–547, 1991.
- Ducharne, A., Koster, R., Suarez, M., Stieglitz, M., and Kumar, P.: A catchment-based approach to modeling land surface processes in a general circulation model. 2. Parameter estimation and model demonstration, *J. Geophys. Res.*, 105, 24823–24838, 2000.
- Ek, M. B., Mitchell, K. E., Yin, L., Rogers, P., Grunmann, P., Koren, V., Gayno, G., and Tarpley, J. D.: Implementation of Noah land-surface model advances in the NCEP operational mesoscale Eta model, *J. Geophys. Res.*, 108, 8851, doi:10.1029/2002JD003296, 2003.
- Fieller, E. C., Hartley, H. O., and Pearson, E. S.: Tests for rank correlation coefficients, *Biometrika*, 44, 470–481, 1957.
- Guo, Z., Dirmeyer, P. A., Gao, X., and Zhao, M.: Improving the quality of simulated soil moisture with a multi-model ensemble approach, *Q. J. Roy. Meteor. Soc.*, 133, 731–747, 2007.
- Hansen, M., DeFries, R., Townshend, J., and Sohlberg, R.: Global land cover classification at 1 km spatial resolution using a classification tree approach, *Int. J. Remote Sens.*, 21, 1331–1364, 2000.
- Huete, A. R., Didan, K., Shimabukuro, Y. E., Ratana, P., Saleska, S. R., Hutya, L. R., Yang, W., Nemani, R. R., and Myneni, R.: Amazon rainforests green-up with sunlight in dry season, *Geophys. Res. Lett.*, 33, L06405, doi:10.1029/2005GL025583, 2006.
- Ji, L. and Peters, A. J.: Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices, *Remote Sens. Environ.*, 87, 85–98, doi:10.1016/S0034-4257(03)00174-3, 2003.
- Ji, L. and Peters, A. J.: Lag and seasonality considerations in evaluating AVHRR NDVI response to precipitation, *Photogramm. Eng. Rem. S.*, 71, 1053–1061, 2005.
- Kogan, F. N.: Droughts of the late 1980s in the United States as derived from NOAA polar orbiting satellite data, *B. Am. Meteorol. Soc.*, 76, 655–668, 1995.
- Koster, R. D., Suarez, M. J., Ducharme, A., Stieglitz, M., and Kumar, P.: A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure, *J. Geophys. Res.*, 105, 24809–24822, doi:10.1029/2000JD900327, 2000.
- Kumar, S. V., Peters-Lidard, C. D., Tian, Y., Houser, P. R., Geiger, J., Olden, S., Lighty, L., Eastman, J. L., Doty, B., Dirmeyer, P., Adams, J., Mitchell, K., Wood, E. F., and Sheffield, J.: Land Information System – An interoperable framework for high resolution land surface modeling, *Environ. Modell. Softw.*, 21, 1402–1415, 2006.
- Mitchell, K.: The community Noah land-surface model: User Guide Public Release Version 2.7.1, available at: <http://www.emc.ncep.noaa.gov/mmb/gcp/noahlsn> (last access: 17 September 2012), 2005.
- Mo, K. C., Long, L. N., Xia, Y., Yang, S. K., Schemm, J. E., and Ek, M. B.: Drought indices based on the Climate Forecast System Reanalysis and ensemble NLDAS, *J. Hydrometeorol.*, 12, 185–210, 2010.
- Musyimi, Z.: Temporal Relationships Between Remotely-Sensed Soil Moisture and NDVI over Africa: Potential for Drought Early Warning, Master Thesis, University of Twente, Netherlands, 2010.
- Owe, M., de Jeu, R., and Holmes, T.: Multisensor historical climatology of satellite-derived global land surface moisture, *J. Geophys. Res.*, 113, F01002, doi:10.1029/2007JF000769, 2008.
- Panu, U. S. and Sharma, T. C.: Challenge in drought research: Some perspectives and future directions, *Hydrolog. Sci. J.*, 47, S19–S30, 2002.
- Peled, E., Dutra, E., Viterbo, P., and Angert, A.: Technical Note: Comparing and ranking soil drought indices performance over Europe, through remote-sensing of vegetation, *Hydrol. Earth Syst. Sci.*, 14, 271–277, doi:10.5194/hess-14-271-2010, 2010.
- Peters, A. J., Walter-Shea, E. A., Ji, L., Vina, A., Hayes, M., and Svoboda, M. D.: Drought monitoring with NDVI-based Standardized Vegetation Index, *Photogramm. Eng. Rem. S.*, 68, 71–75, 2002.
- Sheffield, J., Xia, Y., Luo, L., Wood, E. F., Ek, M., Mitchell, K. E., and the NLDAS Team: Drought monitoring with the North American Land Data Assimilation System (NLDAS): Current capabilities and future challenges, in: *Remote Sensing and Drought*, edited by: Wardlow, B., Anderson, M., and Verdin, J., CRC Press, 480 pp., 2012.
- Wardlow, B., Anderson, M., and Verdin, J.: *Remote Sensing and Drought*, CRC Press, 480 pp., 2012.
- Xia, Y., Mitchell, K. E., Ek, M. B., Cosgrove, B., Sheffield, J., Luo, L., Alonge, C. J., Wei, H., Meng, J., Livneh, B., Duan, Q., and Lohmann, D.: Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products, *J. Geophys. Res.*, 117, D03109, doi:10.1029/2011JD016048, 2012.
- Xie, P. and Arkin, P.: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model output, *B. Am. Meteorol. Soc.*, 78, 2539–2558, 1997.