Geophysical Research Letters

RESEARCH LETTER
10.1029/2018GL080870

Key Points:
- A new representation for target soil moisture in irrigation modeling is presented by assimilating SMAP data into CLM4.5
- An a priori scaling approach is used to reduce the bias in SMAP data for simulations outside of the SMAP period
- 1-D Kalman filter assimilation approach provides promising improvements in irrigation and soil moisture simulations

Supporting Information:
- Supporting Information S1

Correspondence to:
Y. Pokhrel,
ypokhrel@egr.msu.edu

Citation:
Felfelani, F., Pokhrel, Y., Guan, K., & Lawrence, D. M. (2018). Utilizing SMAP soil moisture data to constrain irrigation in the Community Land Model. *Geophysical Research Letters*, *45*. https://doi.org/10.1029/2018GL080870

Received 12 OCT 2018
Accepted 16 NOV 2018
Accepted article online 21 NOV 2018

Utilizing SMAP Soil Moisture Data to Constrain Irrigation in the Community Land Model

Farshid Felfelani[1], Yadu Pokhrel[1], Kaiyu Guan[2], and David M. Lawrence[3]

[1]Department of Civil and Environmental Engineering, Michigan State University, East Lansing, MI, USA, [2]Department of Natural Resources and Environmental Sciences and National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, IL, USA, [3]National Center for Atmospheric Research, Boulder, CO, USA

Abstract Irrigation representation in land surface models has been advanced over the past decade, but the soil moisture (SM) data from SMAP satellite have not yet been utilized in large-scale irrigation modeling. Here we investigate the potential of improving irrigation representation in the Community Land Model version-4.5 (CLM4.5) by assimilating SMAP data. Simulations are conducted over the heavily irrigated central U.S. region. We find that constraining the target SM in CLM4.5 using SMAP data assimilation with 1-D Kalman filter reduces the root-mean-square error of simulated irrigation water requirement by 50% on average (for Nebraska, Kansas, and Texas) and significantly improves irrigation simulations by reducing the bias in irrigation water requirement by up to 60%. An a priori bias correction of SMAP data further improves these results in some regions but incrementally. Data assimilation also enhances SM simulations in CLM4.5. These results could provide a basis for improved modeling of irrigation and land-atmosphere interactions.

Plain Language Summary About 90% of consumptive water use is allocated for irrigation sector globally. Irrigation not only affects water resources and hydrology but also alters local to regional climate and weather systems. Therefore, it is important to represent irrigation in land and climate models for better assessment and prediction of water resources and to understand the potential feedback to climate system. Given the lack of global irrigation data, irrigation schemes currently included in large-scale land surface models often use simplified representation to simulate irrigation. The goal of this study is to improve the estimation of irrigation water in these global models by using the newly available soil moisture information from the recent National Aeronautics and Space Administration (NASA) satellite, the Soil Moisture Active Passive (SMAP). We use the widely used Community Land Model and apply it over the heavily irrigated areas in the central United States. Our results indicate that SMAP data assimilation yields a significant improvement in the simulation of irrigation water use. These results have important implications for better assessment and prediction of irrigation water use and for studying human-climate interactions.

1. Introduction

Irrigation water use accounts for ~90% of consumptive water use globally (Scanlon et al., 2012) and ~40% of total freshwater withdrawals in the United States (Dieter et al., 2018). Over the past decade, there has been increasing interest in better simulating irrigation processes in global land surface models (LSMs). Advances have consequently been made in irrigation schemes used in LSMs through improved representation of the amount, method, and timing of irrigation. However, challenges and opportunities to better simulate soil moisture (SM) in irrigated areas and hence the irrigation water requirement (IWR) still remain (Nazemi & Wheater, 2015; Pokhrel et al., 2016; Wada et al., 2017). In most irrigation schemes in LSMs (e.g., Haddeland et al., 2006; Ozdogan et al., 2010; Pokhrel et al., 2012), the timing and amount of irrigation are determined based on SM deficit in the root zone. Irrigation is triggered when root-zone SM drops below a specified threshold. IWR is then calculated as the amount required to bring the root-zone SM to the target level. Different studies have used a range of values for threshold and target SM, which are the key variables in irrigation schemes (Haddeland et al., 2006; Harding & Snyder, 2012; Lawston et al., 2015; Lawston, Santanello, Franz, & Rodell, 2017; Leng, Huang, Tang, Sacks, et al., 2013; Leng, Huang, Tang, Gao, & Leung, 2013; Ozdogan et al., 2010; Pei et al., 2016; Pokhrel et al., 2012; Sorooshian et al., 2011). Such differences in the irrigation representation can lead to discrepancies in the estimates of IWR and irrigation timing among models (Pokhrel et al., 2016), resulting in varying impacts on terrestrial water systems (Chaudhari et al., 2018; Felfelani et al., 2017) as well as surface energy balance and climate (Sacks et al., 2009). Further, the

threshold and target SM parameterizations in most irrigation schemes employ spatially constant bulk coefficients and parameters, causing a small temporal and spatial variability of threshold and target SM and underrepresenting the heterogeneity in irrigation attributes (e.g., irrigation practices, crop-specific water requirements, and irrigation timing).

A promising approach to address some of these limitations in large-scale LSMs is the integration of spatially explicit data from satellite measurements. The recent Soil Moisture Active Passive (SMAP) satellite provides global surface SM with generally low errors across different climate regions (Kumar et al., 2018). Numerous studies have evaluated SMAP data with ground-based observations (Chan et al., 2016; Pan et al., 2016) and used SMAP data to improve hydrological and carbon flux simulations (Alvarez-Garreton et al., 2016; He et al., 2017; Kumar et al., 2015; Lievens et al., 2017, 2015). Recently, Lawston, Santanello, and Kumar (2017) demonstrated that SMAP data can be used to detect seasonal timing and spatial signature of irrigation. In another study (Brocca et al., 2018), SMAP data were incorporated into the soil water balance equation to quantify IWR. These recent findings imply that SMAP data could potentially be used to better constrain and improve irrigation simulations in large-scale LSMs; however, to the authors' best knowledge, such potential has not yet been investigated.

We propose a novel approach for assimilating SMAP data to overcome the above limitations and enhance irrigation simulations by (1) presenting a parsimonious parameterization for extrapolating the 5-cm SMAP SM to the entire root zone, (2) using a 1-D Kalman filter (KF) for SMAP data assimilation into the irrigation scheme of the Community Land Model version-4.5 (CLM4.5), and (3) accounting for bias correction of the SMAP data using an a priori scaling approach. Our hypothesis is that the assimilation of SMAP data to set the target SM can significantly improve the simulation of IWR and SM, thus enabling advancements in the representation of irrigation in global LSMs.

2. Study Domain, Data, and Methods

2.1. Study Domain and Data

The model is set up for a region overlying the High Plains Aquifer (HPA) in central United States, ranked first for groundwater withdrawal among all U.S. aquifers (Pokhrel et al., 2015) and a heavily irrigated and data-rich region for ground-based SM observations (30–50°N, 116–92°W; Figure S1 in the supporting information). The annual total freshwater withdrawals in the HPA region are estimated at ~22–27 km^3/year during 2005–2015 (Dieter et al., 2018), of which >90% is used for irrigation.

We use version-4 of the SMAP level-3 radiometer SM data for 2015–2017 period to generate the daily climatology, which is then regridded to 3-arc-min model grids. The SMAP data record is short, which means that extreme events such as wet/drought cycles that might have occurred in the period could have an outsize effect on our results. However, this concern is less important than at first glance because of the rationale that irrigation maintains optimal SM level for crop growth, preventing SM anomaly caused by wet/drought cycles. Nevertheless, since SMAP data might have been affected by other biases (Dong et al., 2018), we apply bias correction using an a priori scaling approach, in which the cumulative distribution function of SMAP data (CDF_{SMAP}) is matched with the CDF of ground observations (CDF_{grnd}) located within a sampling window of 0.5° radius and assumed as the true data (Kumar et al., 2012; Reichle & Koster, 2004). Solving the equation $CDF_{grnd}(x') = CDF_{SMAP}(x)$ for x', SMAP data (i.e., x) are scaled to the bias-corrected SM (i.e., x').

The ground-based SM data are taken from three monitoring networks, namely, Soil Climate Analysis Network (SCAN), U.S. Climate Reference Network (USCRN), and SNOwpack TELemetry (SNOTEL), which are used for both bias correcting SMAP data and validating simulated SM (Bell et al., 2013; Bitar et al., 2012; Schaefer et al., 2007). To validate simulated IWR, we use the U.S. Geological Survey (USGS) census data of irrigation withdrawals (Dieter et al., 2018; Maupin et al., 2014), available every 5 years since 1985.

2.2. Existing Irrigation Scheme in CLM4.5

The CLM (Lawrence et al., 2011; Oleson et al., 2013) is the land component of the Community Earth System Model (CESM). The CLM4.5 includes an irrigation scheme based on Ozdogan et al. (2010), which is used in conjunction with a prognostic crop module (Levis et al., 2012; Peng et al., 2018). Irrigated areas are prescribed based on the high-resolution global irrigated and rainfed crop areas from Portmann et al. (2010). The irrigation scheme uses the SM deficit approach, in which irrigation is activated when the crop leaf area

index is greater than zero and water is limiting for photosynthesis depending on crop type, wilting factor, and root fraction. IWR is then estimated based on the difference between the prescribed target SM and the simulated SM in all soil layers within the root zone (i = 1,n), as

$$SW_{def} = \Sigma_{i=1}^{n} \max\left(SW_{tar,i} - SW_{sim,i}, 0\right) \quad (1)$$

where SW_{def} is the soil water deficit and $SW_{sim,i}$ and $SW_{tar,i}$ are the simulated and target soil water amount (all in kg/m^2), corresponding to the simulated and target SM, respectively, at ith layer from the surface. If irrigation is required and $SW_{def} > 0$, irrigation water (SW_{def}) is withdrawn from runoff and applied directly to the surface at a constant rate during 6–10 a.m. each day. Runoff is allowed to become negative when IWR is larger than runoff (Leng, Huang, Tang, Sacks, et al., 2013) so that irrigation water is not artificially suppressed. The target soil water is simply the weighted arithmetic mean of minimum soil water ($SW_{min,i}$)—that prevents crop water stress—and soil water at saturation ($SW_{sat,i}$):

$$SW_{tar,i} = \left(1 - F_{irrig}\right) \times SW_{min,i} + F_{irrig} \times SW_{sat,i} \quad (2)$$

where F_{irrig} is an empirical factor that is set globally at 0.7 to roughly replicate the global annual irrigation amount observed circa year 2000. Leng, Huang, Tang, Sacks, et al. (2013); Leng, Huang, Tang, Sacks, et al. (2015) suggest that the globally calibrated F_{irrig} parameter may not be suitable for regional studies and that calibration at the scale of administrative units can result in improved regional irrigation simulations. Minimum SM (mm^3/mm^3) corresponding to $SW_{min,i}$ is calculated based on the Clapp and Hornberger (1978) relation:

$$SM_{min,i} = \phi_{e,i} \times \left(\frac{\psi_{o,i}}{\psi_{sat,i}}\right)^{-\frac{1}{B_i}} \quad (3)$$

where $\phi_{e,i}$ is the effective porosity, B_i is the Clapp and Hornberger parameter representing the soil type (i.e., organic matter, sand, and clay fractions), and $\psi_{o,i}$ is the matric potential when stomata are fully open, set to −74,000 mm, the average of matric potential at wilting point (−150,000 mm) and field capacity (−3,400 mm).

The existing irrigation scheme in CLM4.5 uses constant $\psi_{o,i}$ for all crops, and other effective terms in equations (2) and (3) such as $\phi_{e,i}$, $\psi_{sat,i}$, B_i, and $SW_{sat,i}$ only represent the soil type. Thus, other parameters that play crucial roles in irrigation estimation (e.g., crop type and irrigation practices) are ignored altogether due to the lack of global data. Further, as in other LSMs (e.g., Lawston et al., 2015; Pei et al., 2016), CLM4.5 irrigation scheme uses a constant bulk coefficient (here F_{irrig}; equation (2)) globally, which is a major limitation of the existing scheme as described above. Finally, the use of fixed irrigated areas representing approximately 2000 is another structural limitation in CLM4.5 because irrigation location and extent can have significant interannual variability, especially during wet-dry transitions (Deines et al., 2017).

2.3. Improved Representation for Target SM in Irrigation Modeling

Our rationale is that since SMAP can detect seasonal timing and spatial signature of irrigation (Lawston, Santanello, & Kumar, 2017), SMAP SM retrievals can be used to constrain the target SM in irrigation parameterizations, enabling us to capture the effects of some of the missing irrigation attributes (e.g., IWR for different crops) and practices (i.e., drip, sprinkler, and flood systems). In using SMAP data, we set the target SM of each irrigated grid cell on any particular day of the year as a function of daily climatology of SMAP data for the given grid cell and day of the year. The daily climatology of SMAP data is assumed to represent the daily average level of SM maintained by local farmers based on the crop type, atmospheric conditions, and irrigation practice during the SMAP period.

A realistic use of SMAP data for SM-based irrigation modeling requires vertical extrapolation of SM from 5 cm to the entire root zone. Various filtering approaches have been suggested to relate the root-zone SM to surface SM (observed by satellite sensors), by applying recursive equations on time series of surface SM to update certain parameters (e.g., characteristic time length and gain factor) and then to estimate the root-zone SM (Albergel et al., 2008; Sabater et al., 2007). Here for simplicity, we propose a parsimonious (i.e., requires minimal parameters) formulation based on the concept of Clapp and Hornberger (1978) to

vertically extrapolate the SMAP SM as a function of the model soil type (represented in B_i) and the degree of saturation ($\theta^* = \frac{SM_{sim,i}}{\phi_{e,i}}$). The target SM for soil layers deeper than 5 cm from the surface is then computed as

$$SM_{tar,i} = SM_{SMAP,5cm} + SM_{SMAP,5cm} \times \left(1 - \left(1 - \frac{SM_{sim,i}}{\phi_{e,i}}\right)^{B_i}\right) \quad (4)$$

where $SM_{SMAP,5cm}$ is the SMAP SM and $SM_{tar,i}$ is the target SM in layer i, both in mm^3/mm^3. Overall, the multiplying term $\left(1 - (1 - \theta^*)^{B_i}\right)$ starts from 0 and approaches unity as the degree of saturation ranges from 0 to 1. Therefore, $SM_{tar,i}$ ranges from $SM_{SMAP,5cm}$ in the top two layers (where depth is less than 5 cm) to twice the $SM_{SMAP,5cm}$ close to the water table (where $\theta^* \approx 1$). The vertical profile of extrapolated SMAP SM thus derived is presented in Figure S2, which is comparable with the temporally averaged vertical SM profile discussed in previous studies (e.g., Zeng & Decker, 2009).

We test the original CLM4.5 irrigation scheme and two new representations for target SM using SMAP data assimilation by (1) directly integrating raw SMAP data and (2) assimilating SMAP data using KF with and without bias correction. In the direct integration approach, the target SM at each time step is set by evaluating equation (4), given the daily climatology of SMAP SM for the day of the year and the grid cell considered. If there is no SMAP observations for the day of the year, the scheme relies on the CLM4.5 estimation of irrigation water (section 2.4). In the second approach, original and bias-corrected SMAP data are assimilated into the irrigation scheme using 1-D KF to set the target SM based on the SMAP data for the day of the year and the grid cell considered as well as adjacent spatial and temporal grid cells. That is, some degree of ergodicity is assumed in the assimilation framework following previous studies (e.g., Reichle & Koster, 2004). We note that SMAP data are assimilated into CLM to modify the target SM representation in irrigation parameterization and not to directly adjust SM, meaning that SM simulation is not constrained by SMAP data and hence SMAP can be used for an independent validation of simulated SM. See Text S1 in the supporting information for 1-D KF assimilation framework.

2.4. Experimental Design

We conduct five sets of off-line simulations (i.e., CLM decoupled from CESM and forced by meteorological data) using CLM4.5 with (1) no crop and irrigation schemes (NOirrig simulation), (2) the default irrigation scheme (control simulation; CTRL), (3) the improved representation for target SM by directly integrating raw SMAP data (SMAP_raw), (4) the improved representation for target SM enhanced by 1-D KF (SMAP_KF), and (5) the improved representation for target SM enhanced by using a priori bias reduced SMAP data and 1-D KF (SMAP_KF_BC). The crop model is activated for all simulations except for NOirrig. The model is set up at high resolution of 3 arc-min (0.05°) to capture fine-scale details of irrigation processes and reduce scale mismatch with field observations. The model is first spun up for 100 years; simulations are then conducted for 1985–2016 period using the North America Land Data Assimilation System phase II (NLDAS2) forcing data (Xia et al., 2012).

3. Results and Discussion

Figure 1 shows the spatial variability of top-5 cm SM from SMAP and CTRL, along with differences between the SM from SMAP and NOirrig, and SMAP_KF and CTRL, all averaged for June–August (JJA) 2015–2016. It is evident from Figures 1a and 1b that broad wet/dry patterns of SM are reasonably reproduced in CTRL. However, a significant wet bias (up to ~0.18 mm^3/mm^3) compared to SMAP observations can be discerned especially in regions over the HPA, Snake River Plain (SRP), and western portions of the domain (Figures 1a and 1b). This SM overestimation in CTRL is due in part to the overestimation of irrigation water (discussed in Figure 2). However, multiple other factors such as evapotranspiration and soil resistance in CLM (Lawrence et al., 2011; Sakaguchi & Zeng, 2009; Swenson & Lawrence, 2014), biases in precipitation, and uncertainties in runoff parameterizations could have also contributed to a certain extent. To isolate the potential SM bias caused by these factors, surface SM from NOirrig is deducted from SMAP data (Figure 1c); a large wet bias (up to ~0.15 mm^3/mm^3) is found even in the absence of irrigation. Further, the large wet bias in surface SM seen in CTRL is reduced by up to 15% and 40% over the HPA and SRP, respectively, in SMAP_KF (Figure 1d). This bias reduction results from the lower target SM in the improved irrigation representation in SMAP_KF as

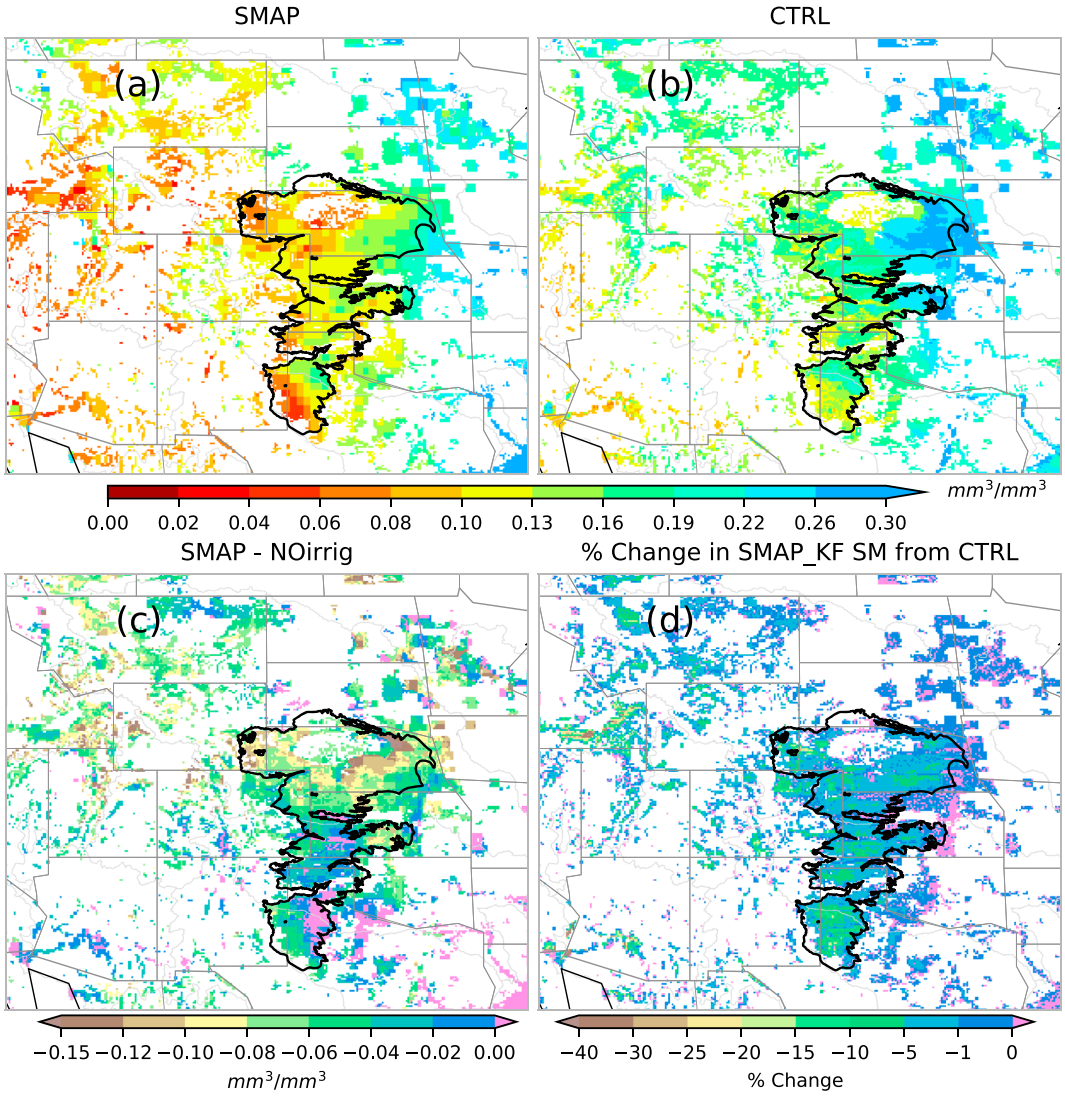

Figure 1. Spatial distribution of top-5 cm soil moisture (averaged for June–August of 2015–2016) from (a) SMAP satellite observations, (b) CTRL, (c) the difference between NOirrig simulation and SMAP observations, and (d) the change (percentage) in surface soil moisture from SMAP_KF relative to CTRL.

dictated by SMAP data; consequently, IWR is essentially reduced in SMAP_KF. We note that the 3-month average SMAP SM is generally drier than in NOirrig (Figure 1c); regardless, irrigation is triggered when daily SMAP SM becomes wetter than that in NOirrig even when the monthly/seasonal SMAP SM is drier. This is likely caused by a short surface SM memory during spring-summer time and in dry regions (Rahman et al., 2015; Wu & Dickinson, 2004).

Figure 2 presents the county-level comparison of simulated annual total IWR with USGS data for census years during 2005–2015; the data for years other than 2015 are used as out-of-sample test data to evaluate irrigation simulations for non-SMAP period. Results for earlier years during 1985–2005 are shown in Figure S4. We note that CLM4.5 simulates IWR without considering field losses (e.g., conveyance and application losses), but USGS data represent total withdrawals. Thus, for consistency, we convert USGS withdrawals to equivalent water requirements by multiplying withdrawals by traditional irrigation efficiency obtained from Jägermeyr et al. (2015; Figure S5c).

As it is evident in Figure 2, significant improvements in the simulated IWR are achieved in other simulations compared to CTRL. The CTRL overestimates IWR for regions over the HPA, Southwest Alluvial Basins in Arizona (ABA), the SRP (green and red rectangles in Figure 2), Montana, New Mexico, and eastern Colorado

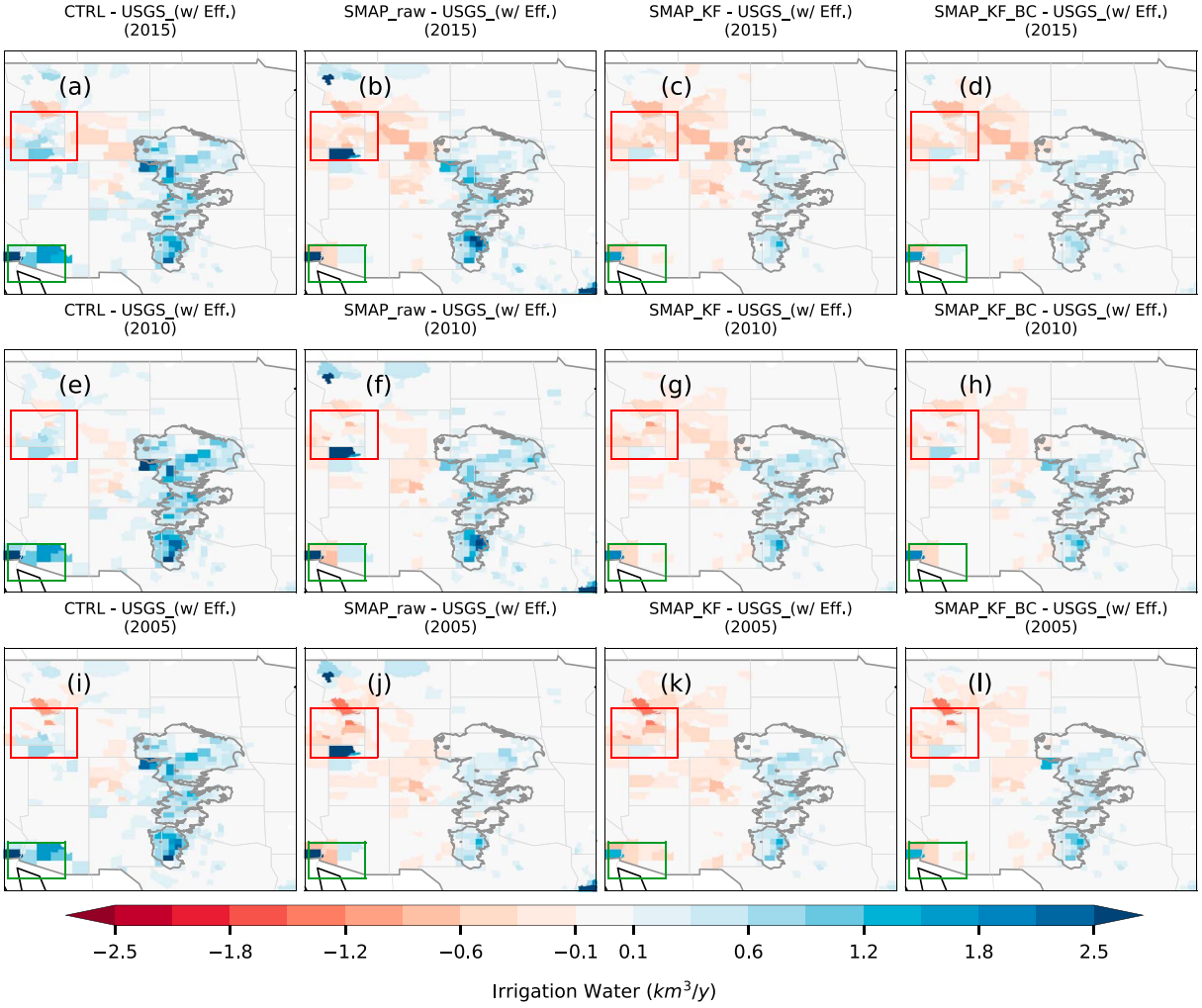

Figure 2. County-level difference between annual total IWR from different simulation settings and U.S. Geological Survey (USGS) data during 2005–2015. USGS withdrawals are converted to equivalent water requirements, and 3-arc-min model results are aggregated for each county. The dark black outline indicates the High Plains Aquifer, and the red and green rectangles show the Snake River Plain and Southwest Alluvial Basins in Arizona regions, respectively.

for all years (Figures 2a, 2e, and 2i). Conversely, CTRL underestimates irrigation in western Colorado, Wyoming, and southwest of Montana. These comparisons clearly demonstrate the deficiency in the default CLM4.5 irrigation scheme in accurately capturing the amount of irrigation water applied, especially over highly irrigated areas (e.g., HPA and ABA; Figure S5a). We make the following key observations.

First, notable improvements are found in SMAP_raw (Figures 2b, 2f, and 2j) compared to CTRL (Figures 2a, 2e, and 2i), most noticeably over the HPA. We find that SMAP detects a likely accurate SM in these highly irrigated areas over the HPA, thus improving the target SM and hence the IWR compared to that in CTRL, which uses a static (i.e., temporally invariant; Figure S3) soil-type-based irrigation trigger threshold (section 2.2). In some counties in the west of the domain, irrigation is underestimated because of relatively low SM in SMAP data over large SMAP grids (36 km) that are sparsely irrigated (Figure S1); that is, a decreased target SM is set in the model. The assessment of SMAP data quality against ground networks indicates that the SMAP error estimate is higher in the western half of the domain, covered mainly by SNOTEL stations (Table S1 in the supporting information). This implies that the potentially lower SMAP data quality in the western half of the domain could have contributed to the underestimation of irrigation. On the contrary, there are areas of overestimated IWR in northwestern Utah, western Montana, and western ABA, which coincide with high SM in SMAP data (Figure 1a).

Second, the positive bias in IWR over the HPA is further reduced for most of the years when the 1-D KF is applied (Figures 2c, 2g, and 2k, Figure S4, and Table S2). For example, the bias in total IWR in HPA reduces by up to 60% from CTRL to SMAP_KF when compared with USGS data for years 2005, 2010, and 2015. Moreover, application of the filter also enables a significant improvement in the highly overestimated IWR in regions of ABA, Utah, and Montana in SMAP_raw (Figures 2b, 2f, and 2j). These improvements are achieved through dampening of the potential noise in SMAP data by considering the underestimated signals from the temporally and spatially neighboring grid cells (see section 2.3 and Text S1). However, KF does not reduce the underestimation of IWR for some counties in SRP and ABA because the neighboring SMAP observations used in KF are also dry, meaning that the surrounding regions are also underirrigated (Figures 2c, 2g, and 2k compared to Figures 2b, 2f, and 2j).

Finally, improvements are found in results from the application of bias correction of SMAP data using ground observations (Figures 2d, 2h, and 2l) compared to those obtained from the application of KF alone (Figures 2c, 2g, and 2k), especially over the HPA and for years 2015, 2010, 1990, and 1985 (Table S2). However, these improvements are relatively small and results from SMAP_KF_BC resemble SMAP_KF results, due primarily to similarity in target SM between the two simulations (Figure S3). Nevertheless, results from bias correction confirm that the bias in simulated IWR is not primarily due to the use of SMAP data for the 2015–2016 period for all simulation years.

Comparisons for other census years (Figure S4) show similar performance as for 2005–2015 discussed above, which is primarily because of (1) the target SM that is constant for all years despite the year-round variability in SMAP simulations and (2) small temporal variability of USGS data throughout 1985–2015 (Figure S5b). In general, Figures 1c and 2 suggest that CLM4.5 tends to overestimate irrigation and the improvements in irrigation simulation are achieved likely due to the reduced wet bias resulting from the constrained target SM through assimilation of SMAP data. Utilizing higher quality SMAP products with longer record could further improve irrigation estimation.

A summary of statistical measures obtained from the state-level comparison of simulated IWR with USGS data for all census years is provided in Text S2 and Table S2 for the three states that cover the majority of the HPA (i.e., Nebraska, Kansas, and Texas). These statistics corroborate our findings that promising improvements are achieved in all modified model settings compared to CTRL. For example, the root-mean-square error (against USGS data) of simulated irrigation water is reduced on average by 50% for above states in SMAP_KF compared with CTRL.

Figure 3 depicts the comparison of simulated SM using different settings with ground observations from SCAN and USCRN networks (red squares in Figure S1) for JJA of 2005–2006, a period chosen as a non-SMAP period. Top panels (Figures 3a–3d) show time series of top-5 cm SM at four of these stations located over the HPA. For clarity of view, only CTRL, SMAP_raw, and SMAP_KF_BC are shown; results from SMAP_KF are highly similar to that from SMAP_KF_BC due to the similitude of target SM in most of the areas across the HPA (Figure S3).

Overall, CTRL fails to capture the episodes of low SM and often simulates false peaks, likely due to false irrigation timing during the growing season. The overall temporal dynamics is improved in SMAP_raw and SMAP_KF_BC, especially in terms of better capturing the episodes of low SM. Occasionally, SMAP_raw outperforms, even the SMAP_KF_BC in reproducing low SM (e.g., in year 2006); however, it fails to capture the overall temporal variability. For example, while observations show a descending trend in SM, which is closely followed by SMAP_KF_BC, SMAP_raw exhibits false peaks (e.g., SCAN_2105 and SCAN_2106 during 2005 June–July, SCAN_2107 during 2005 June–July, and SCAN_2111 in 2006 June).

Interestingly, SMAP_KF_BC realistically simulates the periods of low SM (e.g., SCAN_2105, SCAN_2106, and SCAN_2107 during 2006 June–July and SCAN_2111 during 2006 July–August) as in SMAP_raw, while also capturing the observed SM dynamics (e.g., SCAN_2105, SCAN_2106, and SCAN_2107 during 2005 June–August and SCAN_2111 in 2006 August), which is more accurately simulated in CTRL than in SMAP_raw. This demonstrates the promising performance of KF that predicts the target SM considering the uncertainties associated with SMAP and the model estimation of target SM (see Text S1). Additionally, the bar plot in Figure 3 shows daily precipitation stacked above IWR from SMAP_KF_BC, in which the coincidence of most of observed dry SM events with irrigation application suggests that irrigation is triggered expectedly during relatively dry periods.

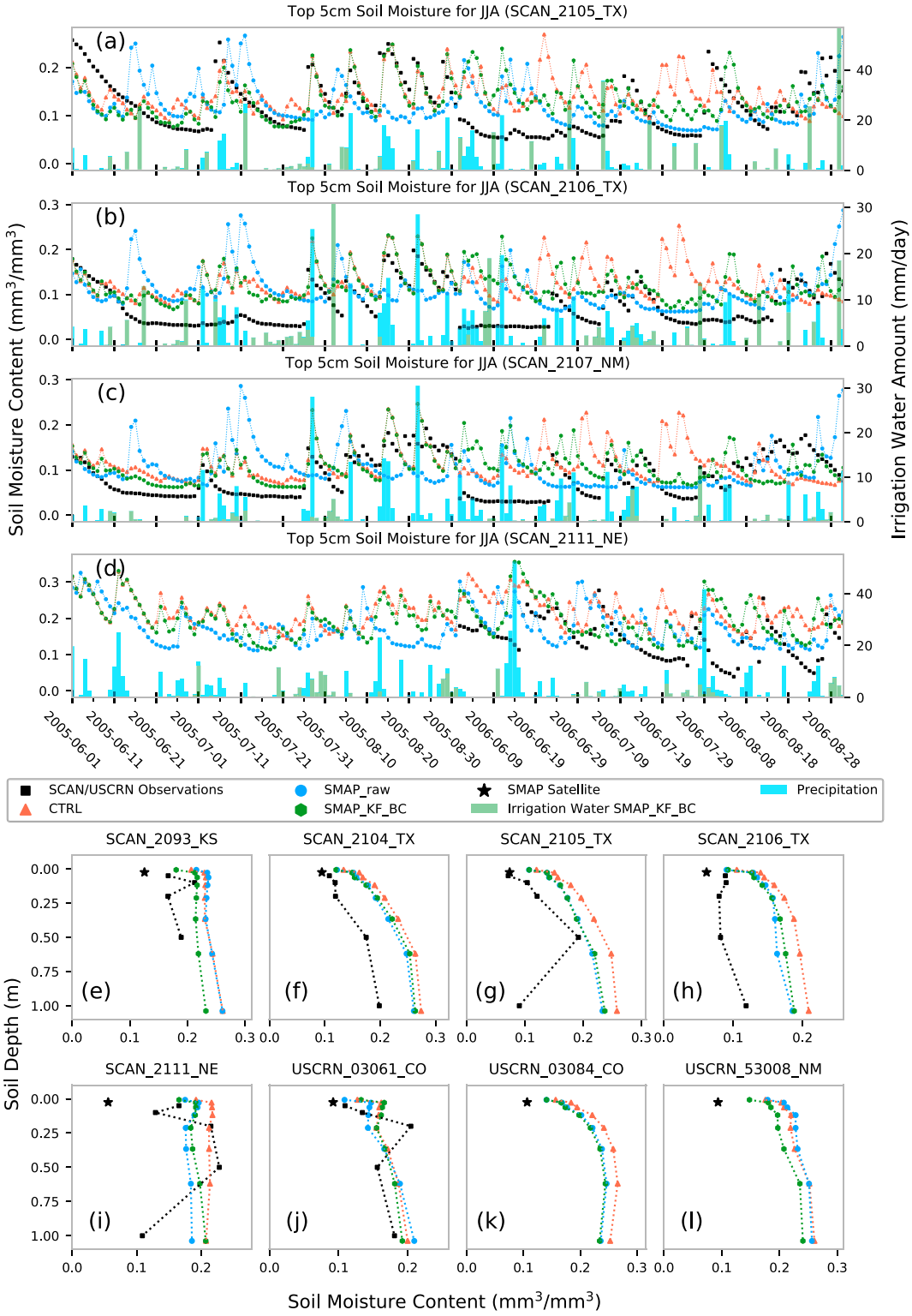

Figure 3. (a-d) Temporal variability of top-5 cm soil moisture from CTRL, SMAP_raw, and SMAP_KF_BC simulations and Soil Climate Analysis Network (SCAN) observations for June–August of 2005–2006 at stations not located in irrigated areas. (e–l) Vertical profiles of averaged SM over June–August (JJA) during 2015–2016 from SMAP, ground observations, and CTRL, SMAP_raw, and SMAP_KF_BC simulations.

The bottom two rows in Figures 3e–3l show vertical profiles of SM in the top meter averaged over JJA of 2015–2016 from different simulations compared against SMAP and ground observations for eight stations over the HPA (additional points are shown in Figure S6). In general, a shift in the SM profile toward the observed profile and SMAP data can be observed in the improved irrigation schemes, suggesting an improvement also in SM simulations due to SMAP data assimilation. Note that SMAP data are used only to improve irrigation representation by revising the threshold-based irrigation application and not to directly alter SM in the model. Therefore, a perfect match between the simulated and observed SM is not expected. Further, since most of the ground stations are located in nonirrigated areas, wetter SM profile in the model is expected. The comparison of SMAP data and ground observations further suggests a fair degree of dry bias in SMAP data across the domain (Figures 3 and S6), more pronounced over the western states; the underestimation of irrigation water in SMAP_raw, SMAP_KF, and SMAP_KF_BC (Figure 2) discussed above is caused by this dry bias in SMAP. While these comparisons demonstrate that SMAP data assimilation in the irrigation scheme also improves SM simulations, the results should be interpreted with caution because this comparison is done between grid-based results and point observations from sparse networks. Despite certain caveats (Chan et al., 2016), such a comparison is common (Pan et al., 2016; Zhang et al., 2017) owing to the lack of dense observation networks and the difficulty in quantifying and comparing SM across observations and models (Dirmeyer et al., 2017, 2016).

4. Conclusions

This study presents a new approach to assimilate SM from SMAP satellite into global LSMs to improve irrigation representation. Results suggest that the simulation of irrigation can be improved by directly integrating SMAP data to constrain the target SM. However, we find that further improvements in simulation of irrigation can be achieved if the 1-D KF assimilation framework is applied. The use of bias-corrected SMAP further improves results in some regions, but the improvements are relatively small compared to those achieved from the KF application. We conclude that despite certain limitations, the use of SMAP data with 1-D KF better represents the target SM, thus providing robust improvements in simulation of IWR and SM, and generally reducing wet bias in IWR in the control simulation (e.g., by up to 60% over the HPA). These results demonstrate the potential of the new parameterizations for constraining target SM using SMAP data and KF, which can be incorporated into any LSM, and applied and validated globally, even for the regions where ground-based data are not available for bias correction. Future research directions include (1) the use of higher quality SMAP data from level-4 products with a longer record, (2) incorporation of other available information from SMAP (e.g., surface flag and land cover class) in the analysis, (3) refinement in model spatial resolution (e.g., 1 arc-min) for better comparison of results with point measurements, and (4) consideration of field-scale details and uncertainties (e.g., in irrigation extent and practices) in irrigation representation with increased spatial resolution.

Acknowledgments
This study was partly supported by the National Science Foundation (award 1752729). We acknowledge computing support from Cheyenne (doi:10.5065/D6RX99HX). D. M. L. is supported by U.S. Department of Agriculture grant 2015-67003-23489. We thank Jonas Jägermeyr for providing irrigation efficiency data sets. SMAP data are available at https://nsidc.org/data/smap/smap-data.html, ground-based SM is available at https://www.wcc.nrcs.usda.gov, and USGS irrigation data are available at https://water.usgs.gov/watuse/data/index.html.

References

Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J.-C., Fritz, N., Froissard, F., et al. (2008). From near-surface to root-zone soil moisture using an exponential filter: An assessment of the method based on in-situ observations and model simulations. *Hydrology and Earth System Sciences*, *12*(6), 1323–1337. https://doi.org/10.5194/hess-12-1323-2008

Alvarez-Garreton, C., Ryu, D., Western, A. W., Crow, W. T., Su, C.-H., & Robertson, D. R. (2016). Dual assimilation of satellite soil moisture to improve streamflow prediction in data-scarce catchments. *Water Resources Research*, *52*, 5357–5375. https://doi.org/10.1002/2015WR018429

Bell, J. E., Palecki, M. A., Baker, C. B., Collins, W. G., Lawrimore, J. H., Leeper, R. D., et al. (2013). U.S. climate reference network soil moisture and temperature observations. *Journal of Hydrometeorology*, *14*(3), 977–988. https://doi.org/10.1175/JHM-D-12-0146.1

Bitar, A. A., Leroux, D., Kerr, Y. H., Merlin, O., Richaume, P., Sahoo, A., & Wood, E. F. (2012). Evaluation of SMOS soil moisture products over continental U.S. using the SCAN/SNO_x2121; network. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(5), 1572–1586. https://doi.org/10.1109/TGRS.2012.2186581

Brocca, L., Tarpanelli, A., Filippucci, P., Dorigo, W., Zaussinger, F., Gruber, A., & Fernández-Prieto, D. (2018). How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products. *International Journal of Applied Earth Observation and Geoinformation*, *73*, 752–766. https://doi.org/10.1016/j.jag.2018.08.023

Chan, S. K., Bindlish, R., O'Neill, P. E., Njoku, E., Jackson, T., Colliander, A., et al. (2016). Assessment of the SMAP passive soil moisture product. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(8), 4994–5007. https://doi.org/10.1109/TGRS.2016.2561938

Chaudhari, S., Felfelani, F., Shin, S., & Pokhrel, Y. (2018). Climate and anthropogenic contributions to the desiccation of the second largest saline lake in the twentieth century. *Journal of Hydrology*, *560*, 342–353. https://doi.org/10.1016/j.jhydrol.2018.03.034

Clapp, R. B., & Hornberger, G. M. (1978). Empirical equations for some soil hydraulic properties. *Water Resources Research*, *14*(4), 601–604. https://doi.org/10.1029/WR014i004p00601

Deines, J. M., Kendall, A. D., & Hyndman, D. W. (2017). Annual irrigation dynamics in the U.S. northern High Plains derived from Landsat satellite data. *Geophysical Research Letters*, *44*, 9350–9360. https://doi.org/10.1002/2017GL074071

Dieter, C. A., Maupin, M. A., Caldwell, R. R., Harris, M. A., Ivahnenko, T. I., Lovelace, J. K., et al. (2018). Estimated use of water in the United States in 2015. *U.S. Geological Survey Circular*, *1441*(65). https://doi.org/10.3133/cir1441. [Supersedes USGS Open-File Report 2017–1131.]

Dirmeyer, P. A., Chen, L., Wu, J., Shin, C.-S., Huang, B., Cash, B. A., et al. (2017). Verification of land–atmosphere coupling in forecast models, reanalyses, and land surface models using flux site observations. *Journal of Hydrometeorology*, *19*(2), 375–392. https://doi.org/10.1175/JHM-D-17-0152.1

Dirmeyer, P. A., Wu, J., Norton, H. E., Dorigo, W. A., Quiring, S. M., Ford, T. W., et al. (2016). Confronting weather and climate models with observational data from soil moisture networks over the United States. *Journal of Hydrometeorology*, *17*(4), 1049–1067. https://doi.org/10.1175/JHM-D-15-0196.1

Dong, J., Crow, W. T., & Bindlish, R. (2018). The error structure of the SMAP single and dual channel soil moisture retrievals. *Geophysical Research Letters*, *45*, 758–765. https://doi.org/10.1002/2017GL075656

Fan, Y., Miguez-Macho, G., Weaver, C. P., Walko, R., & Robock, A. (2007). Incorporating water table dynamics in climate modeling: 1. Water table observations and equilibrium water table simulations. *Journal of Geophysical Research*, *112*, D10125. https://doi.org/10.1029/2006JD008111

Felfelani, F., Wada, Y., Longuevergne, L., & Pokhrel, Y. N. (2017). Natural and human-induced terrestrial water storage change: A global analysis using hydrological models and GRACE. *Journal of Hydrology*, *553*, 105–118. https://doi.org/10.1016/j.jhydrol.2017.07.048

Haddeland, I., Lettenmaier, D. P., & Skaugen, T. (2006). Effects of irrigation on the water and energy balances of the Colorado and Mekong river basins. *Journal of Hydrology*, *324*(1-4), 210–223. https://doi.org/10.1016/j.jhydrol.2005.09.028

Harding, K. J., & Snyder, P. K. (2012). Modeling the atmospheric response to irrigation in the Great Plains. Part I: General impacts on precipitation and the energy budget. *Journal of Hydrometeorology*, *13*(6), 1667–1686. https://doi.org/10.1175/JHM-D-11-098.1

He, L., Chen, J. M., Liu, J., Bélair, S., & Luo, X. (2017). Assessment of SMAP soil moisture for global simulation of gross primary production. *Journal of Geophysical Research: Biogeosciences*, *122*, 1549–1563. https://doi.org/10.1002/2016JG003603

Jägermeyr, J., Gerten, D., Heinke, J., Schaphoff, S., Kummu, M., & Lucht, W. (2015). Water savings potentials of irrigation systems: Global simulation of processes and linkages. *Hydrology and Earth System Sciences*, *19*(7), 3073–3091. https://doi.org/10.5194/hess-19-3073-2015

Kumar, S. V., Dirmeyer, P. A., Peters-Lidard, C. D., Bindlish, R., & Bolten, J. (2018). Information theoretic evaluation of satellite soil moisture retrievals. *Remote Sensing of Environment*, *204*, 392–400. https://doi.org/10.1016/j.rse.2017.10.016

Kumar, S. V., Peters-Lidard, C. D., Santanello, J. A., Reichle, R. H., Draper, C. S., Koster, R. D., et al. (2015). Evaluating the utility of satellite soil moisture retrievals over irrigated areas and the ability of land data assimilation methods to correct for unmodeled processes. *Hydrology and Earth System Sciences*, *19*(11), 4463–4478. https://doi.org/10.5194/hess-19-4463-2015

Kumar, S. V., Reichle, R. H., Harrison, K. W., Peters-Lidard, C. D., Yatheendradas, S., & Santanello, J. A. (2012). A comparison of methods for a priori bias correction in soil moisture data assimilation. *Water Resources Research*, *48*, W03515. https://doi.org/10.1029/2010WR010261

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., et al. (2011). Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model. *Journal of Advances in Modeling Earth Systems*, *3*, M03001. https://doi.org/10.1029/2011MS00045

Lawston, P. M., Santanello, J. A. Jr., Franz, T. E., & Rodell, M. (2017). Assessment of irrigation physics in a land surface modeling framework using non-traditional and human-practice datasets. *Hydrology and Earth System Sciences*, *21*(6), 2953–2966. https://doi.org/10.5194/hess-21-2953-2017

Lawston, P. M., Santanello, J. A. Jr., & Kumar, S. V. (2017). Irrigation signals detected from SMAP soil moisture retrievals. *Geophysical Research Letters*, *44*, 11,860–11,867. https://doi.org/10.1002/2017GL075733

Lawston, P. M., Santanello, J. A. Jr., Zaitchik, B. F., & Rodell, M. (2015). Impact of irrigation methods on land surface model spinup and initialization of WRF forecasts. *Journal of Hydrometeorology*, *16*(3), 1135–1154. https://doi.org/10.1175/JHM-D-14-0203.1

Leng, G., Huang, M., Tang, Q., Gao, H., & Leung, L. R. (2013). Modeling the effects of groundwater-fed irrigation on terrestrial hydrology over the conterminous United States. *Journal of Hydrometeorology*, *15*(3), 957–972. https://doi.org/10.1175/JHM-D-13-049.1

Leng, G., Huang, M., Tang, Q., & Leung, L. R. (2015). A modeling study of irrigation effects on global surface water and groundwater resources under a changing climate. *Journal of Advances in Modeling Earth Systems*, *7*, 1285–1304. https://doi.org/10.1002/2015MS000437

Leng, G., Huang, M., Tang, Q., Sacks, W. J., Lei, H., & Leung, L. R. (2013). Modeling the effects of irrigation on land surface fluxes and states over the conterminous United States: Sensitivity to input data and model parameters. *Journal of Geophysical Research: Atmospheres*, *118*, 9789–9803. https://doi.org/10.1002/jgrd.50792

Levis, S., Bonan, G. B., Kluzek, E., Thornton, P. E., Jones, A., Sacks, W. J., & Kucharik, C. J. (2012). Interactive crop management in the Community Earth System Model (CESM1): Seasonal influences on land–atmosphere fluxes. *Journal of Climate*, *25*(14), 4839–4859. https://doi.org/10.1175/JCLI-D-11-00446.1

Lievens, H., Reichle, R. H., Liu, Q., De Lannoy, G. J. M., Dunbar, R. S., Kim, S. B., et al. (2017). Joint Sentinel-1 and SMAP data assimilation to improve soil moisture estimates. *Geophysical Research Letters*, *44*, 6145–6153. https://doi.org/10.1002/2017GL073904

Lievens, H., Tomer, S. K., Al Bitar, A., De Lannoy, G. J. M., Drusch, M., Dumedah, G., et al. (2015). SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia. *Remote Sensing of Environment*, *168*, 146–162. https://doi.org/10.1016/j.rse.2015.06.025

Maupin, M. A., Kenny, J. F., Hutson, S. S., Lovelace, J. K., Barber, N. L., & Linsey, K. S. (2014). Estimated use of water in the United States in 2010. *U.S. Geological Survey Circular*, *1405*, 56. https://dx.doi.org/10.3133/cir1405

Nazemi, A., & Wheater, H. S. (2015). On inclusion of water resource management in Earth system models—Part 1: Problem definition and representation of water demand. *Hydrology and Earth System Sciences*, *19*(1), 33–61. https://doi.org/10.5194/hess-19-33-2015

Oleson, K., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., et al., 2013. Technical description of version 4.5 of the Community Land Model (CLM). NCAR Tech. Note NCARTN-503STR 420 pp. https://doi.org/10.5065/D6RR1W7M

Ozdogan, M., Rodell, M., Beaudoing, H. K., & Toll, D. L. (2010). Simulating the effects of irrigation over the United States in a land surface model based on satellite-derived agricultural data. *Journal of Hydrometeorology*, *11*(1), 171–184. https://doi.org/10.1175/2009JHM1116.1

Pan, M., Cai, X., Chaney, N. W., Entekhabi, D., & Wood, E. F. (2016). An initial assessment of SMAP soil moisture retrievals using high-resolution model simulations and in situ observations. *Geophysical Research Letters*, *43*, 9662–9668. https://doi.org/10.1002/2016GL069964

Pei, L., Moore, N., Zhong, S., Kendall, A. D., Gao, Z., & Hyndman, D. W. (2016). Effects of irrigation on summer precipitation over the United States. *Journal of Climate*, *29*(10), 3541–3558. https://doi.org/10.1175/JCLI-D-15-0337.1

Peng, B., Guan, K., Chen, M., Lawrence, D. M., Pokhrel, Y., Suyker, A., et al. (2018). Improving maize growth processes in the community land model: Implementation and evaluation. *Agricultural and Forest Meteorology*, *250-251*, 64–89. https://doi.org/10.1016/j.agrformet.2017.11.012

Pokhrel, Y., Hanasaki, N., Koirala, S., Cho, J., Yeh, P. J.-F., Kim, H., et al. (2012). Incorporating anthropogenic water regulation modules into a land surface model. *Journal of Hydrometeorology*, *13*(1), 255–269. https://doi.org/10.1175/JHM-D-11-013.1

Pokhrel, Y., Hanasaki, N., Wada, Y., & Kim, H. (2016). Recent progresses in incorporating human land–water management into global land surface models toward their integration into Earth system models. *Wiley Interdisciplinary Reviews Water*, *3*(4), 548–574. https://doi.org/10.1002/wat2.1150

Pokhrel, Y., Koirala, S., Yeh, P. J.-F., Hanasaki, N., Longuevergne, L., Kanae, S., & Oki, T. (2015). Incorporation of groundwater pumping in a global Land Surface Model with the representation of human impacts. *Water Resources Research*, *51*, 78–96. https://doi.org/10.1002/2014WR015602

Portmann, F. T., Siebert, S., & Döll, P. (2010). MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles*, *24*, GB1011. https://doi.org/10.1029/2008GB003435

Rahman, M. M., Lu, M., & Kyi, K. H. (2015). Variability of soil moisture memory for wet and dry basins. *Journal of Hydrology*, *523*, 107–118. https://doi.org/10.1016/j.jhydrol.2015.01.033

Reichle, R. H., & Koster, R. D. (2004). Bias reduction in short records of satellite soil moisture. *Geophysical Research Letters*, *31*, L19501. https://doi.org/10.1029/2004GL020938

Sabater, J. M., Jarlan, L., Calvet, J.-C., Bouyssel, F., & De Rosnay, P. (2007). From near-surface to root-zone soil moisture using different assimilation techniques. *Journal of Hydrometeorology*, *8*(2), 194–206. https://doi.org/10.1175/JHM571.1

Sacks, W. J., Cook, B. I., Buenning, N., Levis, S., & Helkowski, J. H. (2009). Effects of global irrigation on the near-surface climate. *Climate Dynamics*, *33*(2-3), 159–175. https://doi.org/10.1007/s00382-008-0445-z

Sakaguchi, K., & Zeng, X. (2009). Effects of soil wetness, plant litter, and under-canopy atmospheric stability on ground evaporation in the Community Land Model (CLM3.5). *Journal of Geophysical Research*, *114*, D01107. https://doi.org/10.1029/2008JD010834

Scanlon, B. R., Faunt, C. C., Longuevergne, L., Reedy, R. C., Alley, W. M., McGuire, V. L., & McMahon, P. B. (2012). Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley. *Proceedings of the National Academy of Sciences*, *109*(24), 9320–9325. https://doi.org/10.1073/pnas.1200311109

Schaefer, G. L., Cosh, M. H., & Jackson, T. J. (2007). The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN). *Journal of Atmospheric and Oceanic Technology*, *24*(12), 2073–2077. https://doi.org/10.1175/2007JTECHA930.1

Sorooshian, S., Li, J., Hsu, K., & Gao, X. (2011). How significant is the impact of irrigation on the local hydroclimate in California's Central Valley? Comparison of model results with ground and remote-sensing data. *Journal of Geophysical Research*, *116*, D06102. https://doi.org/10.1029/2010JD014775

Swenson, S. C., & Lawrence, D. M. (2014). Assessing a dry surface layer-based soil resistance parameterization for the Community Land Model using GRACE and FLUXNET-MTE data. *Journal of Geophysical Research: Atmospheres*, *119*, 10,299–10,312. https://doi.org/10.1002/2014JD022314

Wada, Y., Bierkens, M. F. P., de Roo, A., Dirmeyer, P. A., Famiglietti, J. S., Hanasaki, N., et al. (2017). Human–water interface in hydrological modelling: Current status and future directions. *Hydrology and Earth System Sciences*, *21*(8), 4169–4193. https://doi.org/10.5194/hess-21-4169-2017

Wu, W., & Dickinson, R. E. (2004). Time scales of layered soil moisture memory in the context of land–atmosphere interaction. *Journal of Climate*, *17*(14), 2752–2764. https://doi.org/10.1175/1520-0442(2004)017<2752:TSOLSM>2.0.CO;2

Xia, Y., Kenneth, M., Michael, E., Justin, S., Brian, C., Eric, W., et al. (2012). 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. *Journal of Geophysical Research*, *117*, D03109. https://doi.org/10.1029/2011JD016048

Zeng, X., & Decker, M. (2009). Improving the numerical solution of soil moisture–based Richards equation for land models with a deep or shallow water table. *Journal of Hydrometeorology*, *10*(1), 308–319. https://doi.org/10.1175/2008JHM1011.1

Zhang, X., Zhang, T., Zhou, P., Shao, Y., & Gao, S. (2017). Validation analysis of SMAP and AMSR2 soil moisture products over the United States using ground-based measurements. *Remote Sensing*, *9*(2), 104. https://doi.org/10.3390/rs9020104