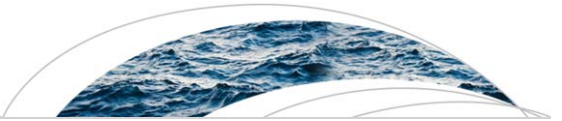
AGU PUBLICATIONS

Water Resources Research



RESEARCH ARTICLE
10.1002/2017WR021415

Mapping Surface Heat Fluxes by Assimilating SMAP Soil Moisture and GOES Land Surface Temperature Data

Yang Lu[1], Susan C. Steele-Dunne[1], Leila Farhadi[2], and Nick van de Giesen[1]

[1]Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, the Netherlands, [2]Department of Civil and Environmental Engineering, George Washington University, Washington, DC, USA

Key Points:
- Surface heat flux estimates can be improved by assimilating SMAP soil moisture and GOES land surface temperature (LST) data
- The resolution gap between SMAP soil moisture and GOES LST data can be bridged using a hybrid particle assimilation strategy
- Despite the coarse resolution, assimilating SMAP soil moisture data is crucial for flux estimation

Supporting Information:
- Supporting Information S1

Correspondence to:
Y. Lu,
y.lu-1@tudelft.nl

Citation:
Lu, Y., Steele-Dunne, S. C., Farhadi, L., & van de Giesen, N. (2017). Mapping surface heat fluxes by assimilating SMAP soil moisture and GOES land surface temperature data. *Water Resources Research*, *53*, 10,858–10,877. https://doi.org/10.1002/2017WR021415

Received 28 JUN 2017
Accepted 8 DEC 2017
Accepted article online 14 DEC 2017
Published online 29 DEC 2017

Abstract Surface heat fluxes play a crucial role in the surface energy and water balance. In situ measurements are costly and difficult, and large-scale flux mapping is hindered by surface heterogeneity. Previous studies have demonstrated that surface heat fluxes can be estimated by assimilating land surface temperature (LST) and soil moisture to determine two key parameters: a neutral bulk heat transfer coefficient (C_{HN}) and an evaporative fraction (EF). Here a methodology is proposed to estimate surface heat fluxes by assimilating Soil Moisture Active Passive (SMAP) soil moisture data and Geostationary Operational Environmental Satellite (GOES) LST data into a dual-source (DS) model using a hybrid particle assimilation strategy. SMAP soil moisture data are assimilated using a particle filter (PF), and GOES LST data are assimilated using an adaptive particle batch smoother (APBS) to account for the large gap in the spatial and temporal resolution. The methodology is implemented in an area in the U.S. Southern Great Plains. Assessment against in situ observations suggests that soil moisture and LST estimates are in better agreement with observations after assimilation. The RMSD for 30 min (daytime) flux estimates is reduced by 6.3% (8.7%) and 31.6% (37%) for *H* and *LE* on average. Comparison against a LST-only and a soil moisture-only assimilation case suggests that despite the coarse resolution, assimilating SMAP soil moisture data is not only beneficial but also crucial for successful and robust flux estimation, particularly when the uncertainties in the model estimates are large.

1. Introduction

Surface heat fluxes significantly influence the terrestrial water and energy cycle and are crucial to many aspects, such as ecology, meteorology, and climate change studies (Lu et al., 2016; McCabe & Wood, 2006). However, in situ measurements of surface heat fluxes are difficult and costly, and are only available from a handful of sparse flux networks (e.g., FluxNet, AmeriFlux, AsiaFlux) (Baldocchi et al., 2001; Kumar et al., 2017). Remote sensing techniques provide potential for land surface measurements. Unfortunately, surface heat fluxes cannot be observed directly from space since they do not have a unique signature that can be detected by remote sensing instruments (Sini et al., 2008). Large-scale flux mapping is further complicated by the spatial heterogeneity in the properties of the underlying land surface, such as surface roughness, soil moisture, and vegetation cover (Jung et al., 2009, 2011).

w?>Many studies have been conducted to estimate surface heat fluxes, either at point or large scale. Some studies build empirical relationships between surface heat fluxes and local indicators that are observable, such as land surface temperature (LST) and vegetation indices (Gillies et al., 1997; Nagler et al., 2005a, 2005b; Tang et al., 2010). These methods are often referred to as the "triangle method" (Carlson, 2007). Other studies estimate fluxes by solving the surface energy balance (SEB) using SEB models (Allen et al., 2007; Anderson et al., 1997, 2011; Bastiaanssen et al., 1998a, 1998b; Jiang & Islam, 2001; Kustas et al., 1996; Su, 2002; Timmermans et al., 2007) or land surface models (LSM) (Niu et al., 2011; Oleson et al., 2010; Zheng et al., 2015). These models require ancillary data input such as surface roughness length and leaf area index (LAI) to estimate fluxes, and a closure assumption is often imposed (Sini et al., 2008).

In a departure from the diagnostic methods, some studies estimate surface heat fluxes by assimilating LST time series into a heat transfer model (Bateni & Entekhabi, 2012; Bateni & Liang, 2012; Bateni et al., 2013a; Boni et al., 2001; Caparrini et al., 2003, 2004a, 2004b; Castelli et al., 1999; Xu et al., 2014, 2015). These studies aim to estimate two key parameters: a bulk heat transfer coefficient under neutral condition (C_{HN}) which

governs the sum of fluxes, and an evaporative fraction (EF) which represents the partitioning between sensible and latent heat fluxes. Sensible heat fluxes (H) are calculated using C_{HN}, and latent heat fluxes (LE) are derived from H and EF. Two assumptions are made: (1) C_{HN} stays constant during a certain time period (e.g., a month), and (2) EF stays constant during the daytime (09:00–16:00 LT) under clear-sky conditions. As a result of the simplified model structure, this group of methods greatly reduces the dimensions of parameter estimation, and only requires a limited amount of data input (e.g., solar radiation, precipitation, wind speed, and air temperature).

Despite the advantages of estimating fluxes from LST assimilation, some limitations have been demonstrated. First, the methods perform poorly on wet soil or dense vegetation conditions, as the constraint of LST on SEB is weakened when the energy partitioning becomes more energy limited (Bateni & Entekhabi, 2012; Caparrini et al., 2004a; Crow & Kustas, 2005; Lu et al., 2016; Xu et al., 2014). Second, an accurate initial guess of C_{HN} is crucial for flux estimation, which is often unavailable. Third, C_{HN} has been demonstrated to increase with LAI (Bateni et al., 2013b, 2014), hence assuming C_{HN} to be constant in a monthly period may degrade surface heat flux estimation, especially in the growing season when the surface characteristics change dramatically over a few weeks.

Several studies have been conducted to improve the methodology. Sini et al. (2008) first demonstrated that flux estimation can be improved by assimilating the antecedent precipitation index. Later, Farhadi et al. (2014, 2016) and Lu et al. (2016) introduced different forms of relationships between EF and soil wetness condition, and demonstrated that flux estimation is improved by putting stronger constraints on EF. An exponential relationship between C_{HN} and LAI was proposed by Farhadi et al. (2014), and proved effective in characterizing the temporal evolution of surface control on heat fluxes with vegetation phenology (Abdolghafoorian et al., 2017; Farhadi et al., 2016).

Remote sensing techniques provide global measurements of LST (Li et al., 2013; Sobrino & Romaguera, 2004; Sun & Pinker, 2003) and soil moisture (Entekhabi et al., 2010; Kim et al., 2014; Mladenova et al., 2014), which greatly facilitates large-scale surface heat flux mapping. The commonly used soil moisture products are from the Advanced Microwave Scanning Radiometer for Earth observation science (AMSR-E) (Njoku et al., 2003), the Advanced Microwave Scanning Radiometer-2 (Imaoka et al., 2010), the Advanced Scatterometer (ASCAT) (Bartalis et al., 2007), the Soil Moisture and Ocean Salinity (SMOS) (Kerr et al., 2001), and the Soil Moisture Active Passive (SMAP) (Entekhabi et al., 2010). In particular, the SMAP mission launched in January 2015 provides global coverage of the top 5 cm soil moisture with a spatial resolution of about 36 km every 2–3 days. Various validation studies have suggested that SMAP characterizes the dynamics of soil moisture with a high accuracy (Cai et al., 2017; Colliander et al., 2017; Pan et al., 2016). Due to the coarse resolution of soil moisture estimates from remote sensing and the spatial heterogeneity of soil moisture, it is unclear if these soil moisture data can effectively constrain EF. In addition, the large spatiotemporal resolution gap between SMAP data and LST data poses a significant challenge for the joint assimilation. For example, LST data from Geostationary Operational Environmental Satellite (GOES) are provided every hour at 0.05° resolution.

Lu et al. (2016) proposed a methodology to estimate surface heat fluxes by joint assimilation of in situ LST and soil moisture data into a simple heat and moisture transfer model. Here this approach is developed further by: (1) employing a dual-source (DS) model to characterize the influence from the soil and canopy, (2) adopting the relationship proposed by Farhadi et al. (2014) to eliminate the need for initial C_{HN} values, (3) assimilating soil moisture data from SMAP and LST data from GOES, (4) bridging the resolution gap using a hybrid particle assimilation strategy, (5) using forcing data independent of ground measurements, and (6) making the strategy calibration free (surface heat flux observations are not required to calibrate the tuning factor in assimilation) using an adaptive approach. This study aims to answer the following questions: (1) Can surface heat fluxes be improved by assimilating SMAP soil moisture and GOES LST data? (2) How can the spatial and temporal resolution between the data sets be bridged? (3) What is the added value of the SMAP soil moisture data to flux estimation?

This paper is organized as follows: Section 2 describes the DS model, the data sets used, the hybrid particle assimilation strategy, and the study area. Section 3 provides results assessment and discussion of the estimated soil moisture, LST, 30 min and daytime heat fluxes as well as the spatial distribution of parameters. The added value of SMAP soil moisture data is also discussed. Finally, the conclusions are drawn in section 4.

2. Materials and Methods

2.1. Dual-Source Modeling

2.1.1. Surface Energy Balance

The methodology is fundamentally based on the SEB equation

$$R_n = H + LE + G \tag{1}$$

where R_n is net radiation, H is sensible heat flux, LE is latent heat flux, and G is ground heat flux. H can be calculated from the vertical gradient of temperature between the land surface and the near-surface air by

$$H = \rho C_p C_H U(\mathrm{LST} - T_a) \tag{2}$$

where ρ (kg/m^3) is air density, C_p (J/kg/K) is specific heat capacity of air, C_H (-) is the bulk coefficient for heat transfer, U (m/s) is wind speed, and LST (K) and T_a (K) are LST and near-surface air temperature.

C_H is mainly determined by influences from two factors: C_{HN} and atmospheric stability. Following Farhadi et al. (2014), C_{HN} is estimated by

$$C_{HN} = \exp(a + b \cdot \mathrm{LAI}) \tag{3}$$

Here the parameters a and b represent the surface control on energy fluxes from the soil and vegetation density, respectively.

The stability correction function proposed by Caparrini et al. (2003) is adopted to calculate C_H from C_{HN}. This approach has proved effective in several flux estimation studies (Bateni & Entekhabi, 2012; Bateni & Liang, 2012; Bateni et al., 2013a; Caparrini et al., 2003, 2004a, 2004b; Farhadi et al., 2014; Lu et al., 2016; Xu et al., 2014, 2015). The function is given by

$$C_H = C_{HN} \cdot f(R_i) = C_{HN} \cdot [1 + 2(1 - e^{10R_i})] \tag{4}$$

where R_i is the Richardson number which is an indicator of the atmospheric stability, and can be calculated by

$$R_i = \frac{g}{T_{pot}} \frac{\Delta T_{pot}}{\Delta z} \left(\frac{\Delta z}{\Delta U}\right)^2 \tag{5}$$

where g (m/s^2) is gravitational acceleration, T_{pot} (K) is potential temperature, z (m) is vertical height, and Δ represents the difference across height difference Δz.

H can be derived if C_{HN} and R_i are determined, and LE is calculated from H and EF. EF is defined as

$$\mathrm{EF} = \frac{LE}{H + LE} \tag{6}$$

which renders

$$LE = H \frac{\mathrm{EF}}{1 - \mathrm{EF}} \tag{7}$$

The primary assumption is that EF stays constant during daytime (09:00–16:00 LT) under clear-sky conditions (Crago, 1996; Crago & Brutsaert, 1996; Gentine et al., 2007). In this way, only one EF estimate is needed to calculate LE at every time step during daytime. This greatly reduces the number of parameters to be estimated and increases the application robustness (Caparrini et al., 2004a).

2.1.2. Dual-Source Scheme

The DS SEB scheme, as shown in Figure 1, was first introduced by Norman et al. (1995) and Kustas et al. (1996). In the DS scheme, the land surface is characterized as a mixed layer of soil and vegetation canopy, and the energy balance is constructed for soil and canopy, respectively. The surface heat fluxes are calculated as weighted combination of the contributions from both sources. The net radiation for soil (R_{ns}) is partitioned into sensible heat flux (H_s), latent heat flux (LE_s), and ground heat flux (G) by

$$R_{ns} = H_s + LE_s + G \tag{8}$$

The heat storage in the canopy is ignored, and the net radiation for canopy (R_{nc}) is partitioned into sensible (H_c) and latent (LE_c) heat fluxes

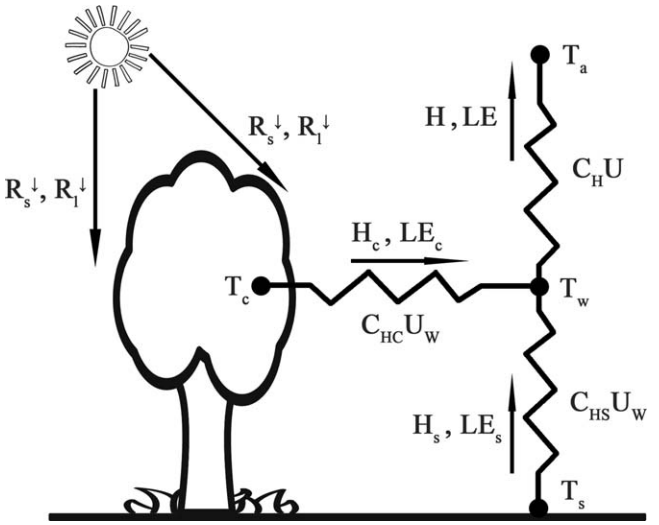


Figure 1. Schematic diagram of the DS SEB scheme.

$$R_{nc}=H_c+LE_c \tag{9}$$

The total fluxes are calculated by the combination of the fluxes from soil and canopy, weighted by the respective fractional cover

$$H=f_cH_c+(1-f_c)H_s \tag{10}$$

$$LE=f_cLE_c+(1-f_c)LE_s \tag{11}$$

where f_c (-) is the canopy cover fraction calculated from LAI following an semiemipirical relationship (Caparrini et al., 2004b; Choudhury, 1987; Norman et al., 1995) by

$$f_c=1-\exp(-0.5\cdot \mathrm{LAI}) \tag{12}$$

The net radiation for soil and canopy are calculated by building a balance between incident shortwave ($R_s^{\downarrow}$) and longwave ($R_l^{\downarrow}$) radiation and the corresponding outgoing radiation for each component

$$R_{ns}=(1-alb_s)R_s^{\downarrow}+R_l^{\downarrow}-\varepsilon_s\sigma T_s^4 \tag{13}$$

$$R_{nc}=(1-alb_c)R_s^{\downarrow}+R_l^{\downarrow}-\varepsilon_c\sigma T_c^4 \tag{14}$$

where alb_s (-) and alb_c (-) are the surface albedo of soil and canopy, ε_s (-) and ε_c (-) are the soil and canopy emissivity, and σ ($W/m^2/K^4$) is the Stefan-Boltzmann constant.

The conductances for heat transfer between soil and air within the canopy and between canopy and air within the canopy are represented by C_{HS} and C_{HC}, respectively, following Caparrini et al. (2004b). H_s and H_c are given by

$$H_s=\rho C_pC_{HS}U_w(T_s-T_w) \tag{15}$$

$$H_c=\rho C_pC_{HC}U_w(T_c-T_w) \tag{16}$$

where T_s (K) and T_c (K) are soil and canopy temperature, T_w (K) and U_w (m/s) are the air temperature and wind speed at a reference height within the canopy volume.

Assuming an exponential decay of conductance within the canopy, Caparrini et al. (2004b) proposed an approach to calculate $C_{HS}U_w$ and $C_{HC}U_w$ from C_{HU} by

$$\begin{cases} K_H=C_HU \\ K_{HS}=C_{HS}U_w=K_H\exp(-0.6\cdot \mathrm{LAI}) \\ K_{HC}=C_{HC}U_w=K_H\exp(-0.3\cdot \mathrm{LAI}) \end{cases} \tag{17}$$

H can also be calculated using C_H for the heat transfer between air within the canopy and air above the canopy by

$$H=\rho C_pC_HU(T_w-T_a) \tag{18}$$

The equation for T_w retrieval can be derived by solving equations (10) and (15–18):

$$T_w=\frac{1}{K_H+(1-f_c)K_{HS}}\left[\frac{f_c(1-\mathrm{EF}_c)}{\rho C_p}R_{nc}+K_HT_a+(1-f_c)K_{HS}T_s\right] \tag{19}$$

Likewise, T_c can be obtained as

$$T_c=\frac{\left[(1-alb_c)R_s^{\downarrow}+R_l^{\downarrow}+3\varepsilon_s\sigma T_a^4\right](1-\mathrm{EF}_c)+\rho C_pK_{HC}T_w}{4\varepsilon_s\sigma T_a^3(1-\mathrm{EF}_c)+\rho C_pK_{HC}} \tag{20}$$

Here EF_s and EF_c are the evaporative fraction for soil and canopy, respectively.

2.1.3. Forward Model

The force-restore model is used to propagate T_s in response to atmospheric forcing and the restoring effect of the deep soil

$$\frac{dT_s}{dt} = \frac{2\sqrt{\pi\omega}}{P_e}(R_{ns} - H_s - LE_s) - 2\pi\omega(T_s - T_d) + \epsilon \tag{21}$$

Here P_e (J/m^2/K/s$^{1/2}$) is the effective thermal inertia, ω (s^{-1}) is the diurnal frequency, T_d (K) is deep soil temperature, and ϵ represents model error. P_e is calculated from soil moisture, bulk density, and sand fraction using the method proposed by Lu et al. (2009), and T_d is estimated with a semidiurnal filter (Caparrini et al., 2003). An additive Gaussian error with a standard deviation of 0.1 K is added at each time step (Lu et al., 2016).

Soil moisture is modeled using the scheme from the simple biosphere model (SiB) (Sellers et al., 1986). For a soil column divided into n layers, the soil moisture variation is calculated by

$$\begin{cases} \frac{\partial W_1}{\partial t} = \frac{1}{\theta_s D_1}[I_1 - Q_{1,2} - \frac{1}{\rho_w}(E_s + E_{t,1})], & k=1 \\ \frac{\partial W_k}{\partial t} = \frac{1}{\theta_s D_k}[Q_{k-1,k} - Q_{k,k+1} - \frac{1}{\rho_w}E_{t,k}], & k=2\ldots n-1 \\ \frac{\partial W_n}{\partial t} = \frac{1}{\theta_s D_n}[Q_{n-1,n} - Q_n], & k=n \end{cases} \tag{22}$$

Here W_k (-) is the soil wetness of the kth layer, I_1 (cm/s) is the infiltration into the first layer from precipitation, D_k (cm) is the thickness of the kth layer, $Q_{k,k+1}$ (cm/s) is the flow between the kth and $k+1$th layer, ρ_w (g/cm^3) is the water density, E_s (g/cm^3/s) is the water loss from soil evaporation, $E_{t,k}$ (g/cm^3/s) is the water loss from vegetation transpiration in the kth layer, and Q_n (cm/s) is the gravitational drainage from the deepest layer. In this experiment, the soil column is divided into six layers, with layer thicknesses of 5, 10, 15, 15, 15, and 30 cm, respectively. The additive modeling error is assumed Gaussian with a standard deviation of 0.001 m^3/m^3. A more detailed description can be found in Lu et al. (2016).

To constrain EF_s and EF_c with modeled soil moisture, the relationship proposed by Dirmeyer et al. (2000) is adopted and improved to apply to the soil and canopy separately:

$$EF_s^{ref} = \frac{2EF_{max}}{\pi} arctan(\alpha_s \cdot SWI_s) \tag{23}$$

$$EF_c^{ref} = \frac{2EF_{max}}{\pi} arctan(\alpha_c \cdot SWI_c) \tag{24}$$

where EF_s^{ref} and EF_c^{ref} are the reference daytime EF values estimated from the corresponding soil wetness indices (SWI). EF_{max} is the maximum possible EF and is assumed to be unity (Lu et al., 2016). α_s and α_c are slope factors that control the shape of the curve and will be estimated during the assimilation. When the reference EF is determined, the prior guess of daytime EF is uniformly sampled within a ± 0.2 range following Lu et al. (2016).

EF_s is assumed to be dependent only on the wetness condition of the surface soil (θ_1), and EF_c is assumed to be affected by the total transpirable water of the soil column (θ_{col}):

$$SWI_s = \frac{\theta_1 - \theta_r}{\theta_s - \theta_r} \tag{25}$$

$$SWI_c = \frac{\theta_{col} - \theta_{wp}}{\theta_{fc} - \theta_{wp}} \tag{26}$$

where θ_s and θ_r are saturated and residual soil moisture, while θ_{fc} and θ_{wp} are field capacity and wilting point. θ_{col} is the mean soil moisture of all layers weighted by root fractions proposed by Zeng (2001).

2.2. Study Area and Data

The developed methodology is tested over an area in the U.S. Southern Great Plains as illustrated in Figure 2. The major part of the study area (35.75°N to 37.24°N, 96.72°W to 98.21°W) is in Oklahoma and a small portion is in Kansas. The area is mostly flat with elevation ranging from about 230 to 450 m above sea level. The area is covered by 4 × 4 SMAP soil moisture grid cells, or 30 × 30 GOES LST grid cells. The major land cover types are grassland and cropland according to ESA Climate Change Initiative (CCI) land cover data set

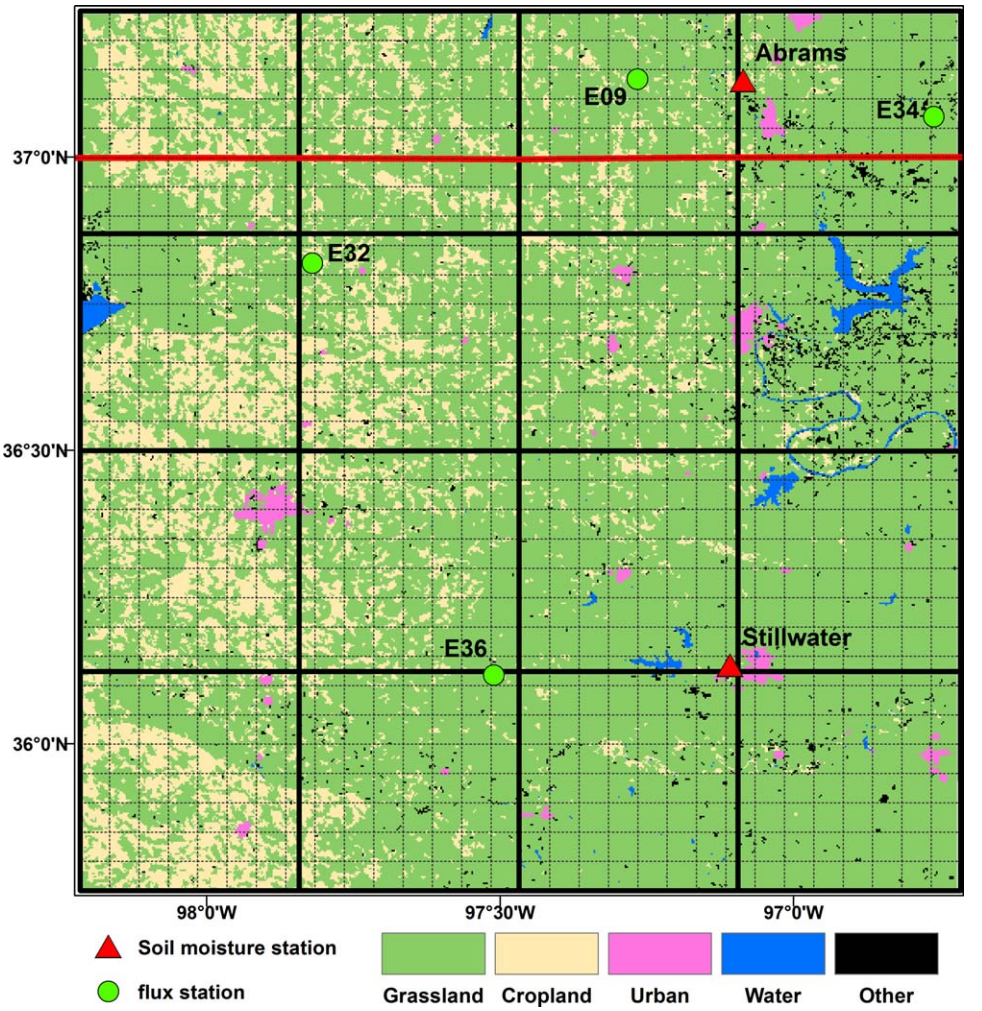


Figure 2. Land cover and stations in the study area. The thick black grids represent SMAP soil moisture grids, while the thin dashed grids represent GOES LST grids. The thick red line demarcates the border between Kansas and Oklahoma.

(v1.6.1). The main soil texture classes are silt loam and sandy loam, and the soil hydraulic properties are estimated from soil texture data using ROSETTA software (Schaap et al., 2001). Thirty-minute in situ flux measurements are available from four Energy Balance Bowen Ratio (EBBR) flux stations provided by the Atmospheric Radiation Measurement (ARM) network. Hourly soil moisture measurements at the depths of 5, 10, 20, 50, and 100 cm are obtained at two stations: Abrams from the Soil Climate Analysis Network (SCAN; Schaefer et al., 2007) and Stillwater from the U.S. Climate Reference Network (CRN; Bell et al., 2013).

Precipitation forcing data are obtained from Global Precipitation Measurement (GPM) (Hou et al., 2014) 3IMERGHH product. The data set provides multisatellite precipitation estimates with gauge calibration, which are available every 30 min at a spatial resolution of 0.1° (Huffman et al., 2015). Other atmospheric forcing data ($R_s^{\downarrow}$, $R_l^{\downarrow}$, T_a, U and air pressure [P_a]) are provided by the North American Land Data Assimilation System project phase 2 (NLDAS-2) Forcing File A (Xia et al., 2012). The majority of forcing data are interpolated from the North American Regional Reanalysis (NARR) and are provided hourly at 0.125° resolution.

Soil texture data, including sand fraction, clay fraction, and bulk density data are the same data sets used for SMAP soil moisture retrieval (Das, 2013) and are obtained from the National Snow and Ice Data Center (NSIDC). These data sets are provided at 3 km resolution. LAI data are extracted from the Moderate Resolution Imaging Spectroradiometer (MODIS) MCD15A2 product (Knyazikhin et al., 1999). The data are available at 1 km resolution every 8 days.

Table 1
Summary of Data Sets

			Original resolution	
Category	Source	Data set	Spatial	Temporal
Forcing	NLDAS-2	$R_s^{\downarrow}, R_l^{\downarrow}, U, T_a, P_a$	0.125°	1 h
	GPM	P	0.1°	30 min
Ancillary	NSIDC	soil texture	3 km	-
	MODIS	LAI	1 km	8 days
Assimilation	SMAP	soil moisture	36 km	2–3 days
	GOES	LST	0.05°	1 h

The SMAP soil moisture data used for assimilation are the Level-3 soil moisture data (L3_*SM_P*) retrieved from the L-band radiometer. The data are posted on 36 km Equal-Area Scalable Earth-2 (EASE-2) grids, with a typical revisit interval of 2–3 days (Entekhabi et al., 2014). The data are acquired and mapped to geographic coordinates from https://reverb.echo.nasa.gov/. Previous studies have suggested that SMAP surface soil moisture data may be biased (Colliander et al., 2017; Zeng et al., 2016). Before assimilation, bias correction is performed using cumulative distribution function (CDF) matching between the observations and simulations for each SMAP grid cell separately to reconcile the differences in long-term mean, variance, and higher moments (De Lannoy & Reichle, 2016).

GOES LST data are provided by Copernicus Global Land Service (http://land.copernicus.eu/global/). The LST data are generated from infrared data from a collection of geostationary satellites covering different areas to provide global coverage (Freitas et al., 2013). The data are provided hourly at 0.05° resolution. Following previous studies (Bateni et al., 2013b), the LST data are directly assimilated into the forward model without bias correction. One concern is that the bias between modeled and observed LST data can be caused by not only the climatological differences between the forward model and the observation system but also by the uncertainties of initial parameter ranges at local scale. Hence, prior scaling may disturb the diurnal variations of LST evolution, and degrade the algorithm performance in updating parameters. A summary of the data sets is provided in Table 1, and all data are converted to geographic coordinates.

2.3. Hybrid Particle Assimilation Strategy

Since SMAP soil moisture data and GOES LST data are provided at different spatial and temporal resolution, a hybrid particle assimilation strategy is adopted. In this strategy, soil moisture states are updated by assimilating SMAP soil moisture data using the particle filter (PF), while LST states and all parameters (i.e., a and b for calculating C_{HN}, and α_s and α_c for calculating EF) are updated by assimilating GOES LST data using an adaptive particle batch smoother (APBS).

2.3.1. Particle Filter for Soil Moisture Assimilation

Particle filters have their origin in Bayesian estimation. Instead of updating model states directly, the PF updates the particle weights based on a likelihood function, and model states are estimated as the weighted sum of all particle estimates (Dong et al., 2015; Moradkhani et al., 2012; Yan et al., 2015). The rationale for using the PF is that the "snapshot" of soil moisture is available every 2–3 days, hence it is more appropriate to update the instantaneous soil moisture using a filter than a smoother. In addition, the distribution of modeled soil moisture realizations is often non-Gaussian, therefore the performance of ensemble methods is often suboptimal (Lu et al., 2016). Using the PF, the need for the assumption of probability densities is relaxed, and the non-Gaussian distributions can be effectively handled (DeChant & Moradkhani, 2012; Lu et al., 2016; Moradkhani et al., 2005; Yan & Moradkhani, 2016).

The evolution of model states in time can be described by

$$\mathbf{x}_t^i = f(\mathbf{x}_{t-1}^i, \mathbf{u}_t^i, \mathbf{b}_t^i) + \mathbf{w}_t^i \tag{27}$$

where f is the forward model, $\mathbf{x}_t^i$ is the model state vector of the ith particle at time step t, $\mathbf{u}_t^i$ is the perturbed forcing data, $\mathbf{b}_t^i$ is the model parameter vector, and $\mathbf{w}_t^i$ represents model error. Here $\mathbf{w}_t^i$ is assumed to be normally distributed.

The model states are related to the observations by

$$\hat{\mathbf{y}}_t^i = h(\mathbf{x}_t^i) + \mathbf{v}_t^i \tag{28}$$

where $\hat{\mathbf{y}}_t^i$ is the simulated observation, h is the observation operator that translates modeled states to the observations, and $\mathbf{v}_t^i$ is the observation error.

Initially, the particles are given uniform weights. At time step t when assimilation is conducted, the weights are updated by

$$w_t^{i*} \propto w_{t-1}^i p(\mathbf{y}_t|\mathbf{x}_t^i) \tag{29}$$

$$w_t^i = \frac{w_t^{i*}}{\sum_{i=1}^{N} w_t^{i*}} \tag{30}$$

where w_t^i is the weight of the ith particle, w_t^{i*} is the unnormalized weight from importance sampling, N is particle size, and $p(\mathbf{y}_t|\mathbf{x}_t^i)$ is the likelihood function, which is expressed as

$$p(\mathbf{y}_t|\mathbf{x}_t^i) \propto e^{\left[-0.5(\mathbf{y}_t-\hat{\mathbf{y}}_t^i)^T\mathbf{R}^{-1}(\mathbf{y}_t-\hat{\mathbf{y}}_t^i)\right]} \tag{31}$$

Here $\mathbf{y}_t$ is the observation, and $\mathbf{R}$ is observation variance.

At the SMAP overpass time (06:00 LT), SMAP soil moisture data are assimilated if available. Since SMAP soil moisture data are much coarser than the model grids, an averaging operator is used, which essentially averages the predicted surface soil moisture from the model so that they can be compared to the coarser resolution observations. First, all the model grid cells within a SMAP grid cell are found. Second, for each model grid cell, the particles are sorted based on the soil moisture of the first layer (i.e., the first particle is the wettest and the last particle is the driest). Finally, the particles from all the model grid cells within a SMAP grid cell are grouped based on their ranking (i.e., the first group contains all the particles which rank first), and the simulated "observation" of this group is generated using an averaging operator

$$\hat{\mathbf{y}}_t^i = \frac{\sum_{j=1}^{M} \theta_{j,1,t}^i}{M} \tag{32}$$

where M is the number of model grid cells within a SMAP grid cell, and $\theta_{j,1,t}^i$ is the soil moisture of the first layer of the ith particle in jth model grid at time step t. The SMAP observation error is assumed 0.04 m^3/m^3 (Cai et al., 2017; Chan et al., 2016; Colliander et al., 2017), and soil moisture is updated after assimilation. It is worth noting that the SMAP observation error used is the nominal value at the original 36 km scale, and is prone to influence from soil texture, vegetation water content and surface heterogeneity. Furthermore, the actual observation error should be greater when the coarse soil moisture data are used to update simulations at finer resolution. Soil moisture estimates may be improved if locally realistic observation error is available.

2.3.2. APBS for LST Assimilation

The APBS was proposed by Dong et al. (2016a). The particle batch smoother (PBS) (Dong et al., 2015; Margulis et al., 2015) can be seen as an extension of the PF. The difference is that the PBS updates model states within an assimilation window in a batch using all available observations in that window, while the PF assimilates observations sequentially (Lu et al., 2016). In the PBS, a tuning factor is estimated using in situ observations to prevent particle degeneracy. The APBS is improved from the PBS by determining the tuning factor automatically in an adaptive manner. The rationale for using the APBS for LST assimilation is that the information of surface control on energy fluxes is contained in the LST time series rather than each individual observation, hence smoothing is better suited than filtering. In addition, estimating the tuning factor adaptively in the APBS eliminates the dependence on in situ observations and is more appropriate for large-scale applications using remote sensing data.

In the APBS, the likelihood function is given by

$$p(\mathbf{y}_{t-L+1:t}|\mathbf{x}_{t-L+1:t}^i) \propto \prod_{j=t-L+1}^{t} e^{\left[-0.5\beta^2(\mathbf{y}_j-\hat{\mathbf{y}}_j^i)^T\mathbf{R}^{-1}(\mathbf{y}_j-\hat{\mathbf{y}}_j^i)\right]} \tag{33}$$

Here L is the length of the assimilation window, $\mathbf{R}$ is the covariance matrix of observations, and β is a tuning factor to avoid particle degeneracy (Dong et al., 2016b). Although resampling of the posterior is performed after each update (Moradkhani et al., 2005), degeneracy cannot be prevented for cases when the observations are nearly perfect (i.e., very small $\mathbf{R}$) or when the model estimates are very inaccurate (e.g., bad initialization or parameterization) by resampling alone (Stordal et al., 2011). In the APBS, β for each assimilation window is determined by maximizing the reliability of model states (0 for zero reliability and 1 for perfect reliability). The probability metric of reliability is estimating using the Quantile-Quantile (Q-Q) plot, which indicates whether the estimated uncertainty is appropriate. See Dong et al. (2016a) for details of the APBS.

In the daytime window (09:00–16:00 LT), a maximum number of eight GOES LST observations are available. At the end of the daytime window (16:00), the assimilation is performed for each model grid cell separately, if at least half of the observations (i.e., four observations) are available for that grid cell. GOES LST observations are related to modeled T_s and T_c following Kustas et al. (1996)

$$\mathrm{LST}=\left[f_c T_c^4+(1-f_c)T_s^4\right]^{1/4} \tag{34}$$

The GOES LST observation error is assumed to be 3 K (Lu et al., 2016). In the APBS, β is varied from 0.05 to 1 with an increment of 0.05, and the β value that yields the largest reliability is selected. After assimilation, particles of modeled T_s and T_c as well as parameters are updated, and fluxes are estimated as the weighted sum of all particle estimates.

2.4. Experiment Setup

The model is run for 120 days from DOY (day of year) 155 to 274, 2015 at 30 min time step for each 0.05° grid cell using 600 particles. Off-line sensitivity analysis confirms that this number of particles is large enough to yield stable results. Grid cells with dominant land cover of water or urban area are masked out in the modeling. Since the model grids are of finer resolution than the forcing data, the forcing data are extracted from the forcing data grid cell in which the center of the model grid cell falls. The perturbations used for the forcing data are listed in Table 2. The perturbations are used to characterize the data uncertainties caused by the uncertainty of the forcing data as well as the heterogeneity within the forcing data grid cell. Sensitivity analysis reveals that flux estimation is robust as long as the perturbations are reasonable. The model and observation errors used are adopted from Lu et al. (2016), and prove to have limited influence on the flux estimates. The ancillary data are spatially aggregated to the modeling scale. The LAI data are also temporally interpolated to get daily LAI data.

The unknown parameters are a and b in estimating C_{HN}, and α_s and α_c in estimating reference EF. Based on the acceptable ranges of soil roughness and vegetation density, the valid ranges are determined for a ($-7 < a < -5$) and b ($0<b<1$) following Abdolghafoorian et al. (2017). The initial range of (1,10) is used for α_s and α_c based on Dirmeyer et al. (2000). This range allows sufficient EF dynamics based on soil moisture condition. Sensitivity analysis demonstrated that the model would generate reliable results as long as the parameters were generated within physically reasonable ranges.

The main input data are solar radiation ($R_s^{\downarrow}$ and $R_l^{\downarrow}$) and meteorological data (P, T_a, U and P_a). At the start of the experiment, parameters are uniformly sampled within the given ranges, and the energy balance is constructed at each time step for soil and canopy separately. To calculate net radiation, soil albedo alb_s is calculated from surface soil moisture based on Idso et al. (1975), and canopy albedo alb_c is based on the values given in Houldcroft et al. (2009) and assumed constant.

During daytime, H is calculated using propagated T_s and T_c, and LE is calculated using H and estimated EF. G is derived as the residual of the energy balance equation and is used to propagate T_s of the next time step. From 16:00 to 09:00, the next day, G is assumed a random fraction of R_{ns} as EF is no longer conservative, and T_s is propagated. T_c is then assumed equal to T_s. SMAP soil moisture data and GOES LST data are assimilated as described in section 2.3. The reader is referred to Lu et al. (2016) for more details.

Table 2
Perturbations used for forcing data[a]

Forcing	Perturbation	Standard deviation
$R_s^{\downarrow}$	Gaussian, ×	×0.1
$R_l^{\downarrow}$	Gaussian, ×	×0.1
U	Gaussian, +	1 m/s
T_a	Gaussian, +	5 K
P	Lognormal, ×	×0.2

[a]"×" and "+" represent multiplicative and additive perturbations, respectively.

3. Results and Discussion

3.1. Soil Moisture and LST Estimation

Modeled surface soil moisture after assimilation (hereafter DA) and from open loop (hereafter OL, i.e., no assimilation case) is assessed against in situ measurements in Abrams and Stillwater. It should be noted that the modeled results are at 0.05° resolution, while in situ data are point measurements. The same is true for flux data; therefore, the term "assessment" is used instead of "validation." The implicit assumption is that the point measurements are representative of the model grid cell condition, which may be influenced by surface heterogeneity. The assessment results are shown in Figure 3. In general, the modeled soil moisture from OL and DA follows the dynamics of in situ

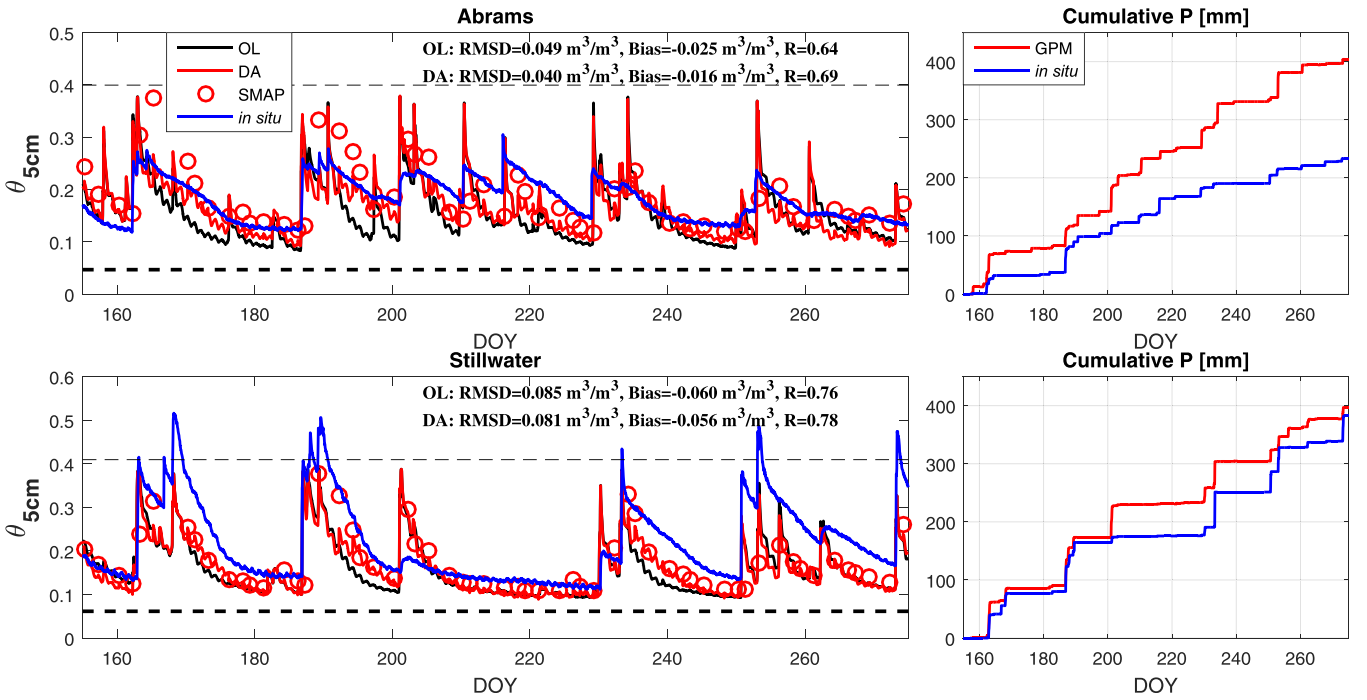


Figure 3. Surface soil moisture assessment against in situ measurements. The thick and thin dashed lines represent residual and saturated soil moisture used in the modeling, respectively.

measurements very well. This reveals the good timing of precipitation events from GPM data, which can be seen more clearly in the cumulative precipitation plots on the right. When SMAP observations are available, the surface soil moisture is updated toward SMAP observations and gets closer to in situ soil moisture, which is most evident from DOY 190–200 at both stations. This demonstrates that assimilating SMAP soil moisture data improves surface soil moisture estimates, despite the coarse resolution.

The improvement is also reflected in the statistical metrics, especially in Abrams. In Abrams, the RMSD decreases from 0.049 m^3/m^3 for OL to 0.040 m^3/m^3 for DA, and the estimation bias decreases from -0.025 m^3/m^3 for OL to -0.016 m^3/m^3 for DA. The R also increases from 0.64 for OL to 0.69 for DA. The improvement is mostly due to a better characterization of the dry-down events after assimilating SMAP soil moisture data. The estimation is less accurate in Stillwater, reflected by larger RMSD and bias. Only small improvement can be seen from the RMSD (0.085 to 0.081 m^3/m^3), bias (–0.060 to -0.056 m^3/m^3) and R (0.76 to 0.78) metrics after assimilation. This is caused by the erroneous soil hydraulic properties and disagreement in precipitation magnitude between GPM and in situ precipitation measurements. For example, the saturated soil moisture (thin dashed line) estimated using ROSETTA is about 0.4 m^3/m^3, but the maximum in situ soil moisture measurement exceeds 0.5 m^3/m^3 during many rain events. This would lead to a general underestimation of soil moisture, which is reflected in the large negative bias in both OL and DA results. On DOY 201, a precipitation event is detected in both GPM and in situ data, but the magnitudes differ significantly ($>$50 mm from GPM and $<$10 mm from in situ data). As a result, modeled soil moisture increases sharply to near saturation, while only a small increase is seen from in situ soil moisture measurements. A large increase in surface soil moisture is also seen in SMAP observations between DOY 200 and 202, which is in line with GPM measurements. This may be caused by the spatial heterogeneity of precipitation, which appears to be lighter at the station location than in the surrounding areas.

Figure 4 shows that after assimilation, the magnitude of surface soil moisture agrees better with SMAP observations, and that the spatial patterns are preserved at 0.05°. A dry-wet gradient from the northwest to the southeast is revealed by SMAP observations. The pattern is also seen in OL estimates but less clearly, and large gap in the magnitude exists between SMAP observations and OL estimates. After assimilation, the gradient is much clearer, and the spatial heterogeneity of soil moisture within each SMAP grid cell is preserved in the DA results.

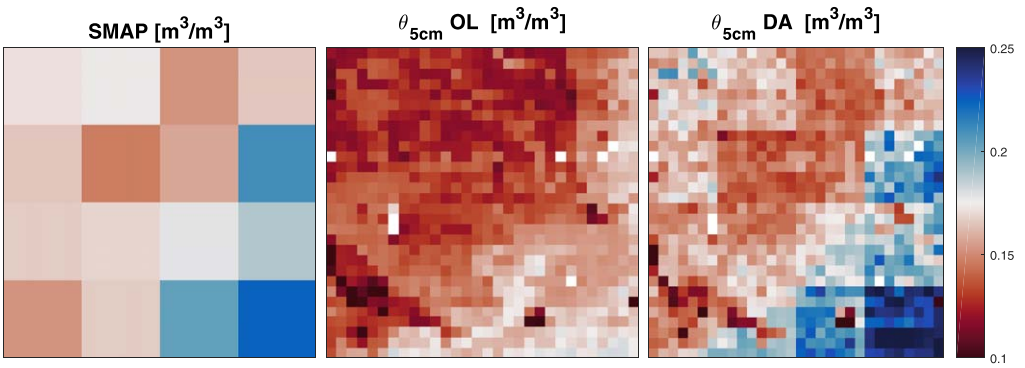


Figure 4. Surface soil moisture from SMAP, OL, and DA at 06:00, DOY 173.

Thirty-minute LST estimates from OL and DA within the daytime window are compared to GOES LST observations for each model grid cell in Figure 5. To perform the assessment, all GOES LST observations available within the daytime window each day are found. The corresponding modeled LST from OL and DA are extracted to calculate statistical metrics for each model grid cell separately. After assimilation, 30 min LST estimates improve significantly based on all the metrics. The range of R^2 improves from 0.6 to 0.7 for OL to

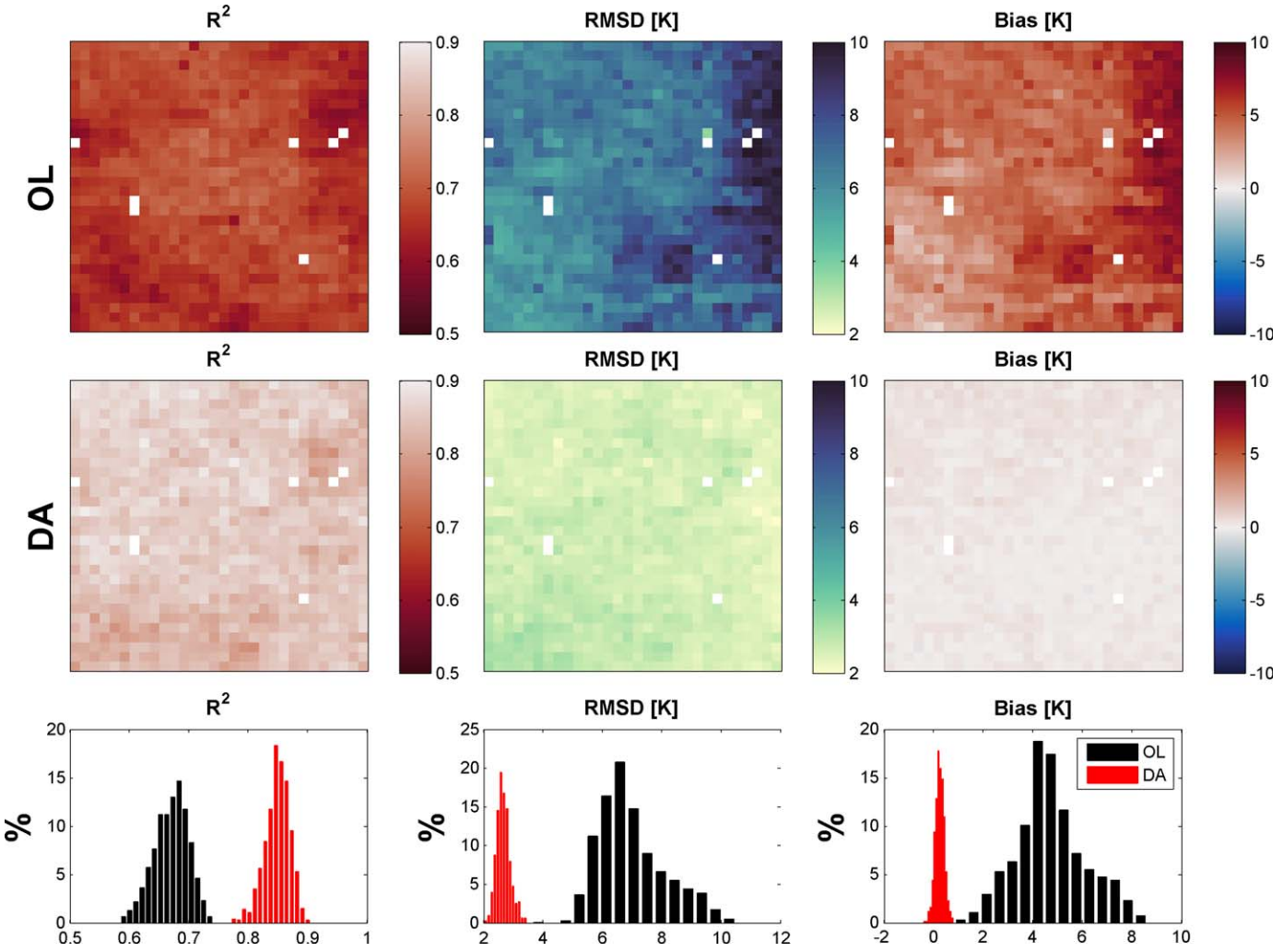


Figure 5. Thirty-minute LST assessment against GOES observations. The histograms of the metrics are plotted in the third row.

0.8 to 0.9 for DA, and the RMSD improves from >5 K for OL to <4 K for DA. The estimation bias for OL is generally between 2 and 8 K, and is reduced to almost zero after assimilation.

3.2. Flux Estimation

The 30 min *H* estimates are compared to in situ observations at the four flux stations on days when assimilation is performed, as shown in Figure 6. Generally, the improvement in *H* is moderate, with the largest improvement in R and small improvement in the RMSD. On average, R is increased by 0.14 and the RMSD is decreased by 6.3%. This is caused by the compensation effect between C_{HN} magnitude and LST dynamics on *H* estimates. For example, if C_{HN} is very large for a particle, *G* will be small based on the energy balance, which will lead to a small LST at the next time step. As C_{HN} and LST compensate each other on the influence on *H*, the variation of *H* with C_{HN} will be small, and vice versa for small C_{HN} values.

Figure 7 shows the skill of 30 min *LE* estimates. In contrast to the *H* estimates, the improvement after assimilation is more evident in all the stations. Since *LE* is determined by *H* and EF calculated from soil moisture, high *R* values are achieved in both OL and DA results, which proves the validity of the EF∼*SWI* relationship. The assimilation improves *LE* estimates in two ways: (1) smaller and more reasonable ranges of α_s and α_c through data assimilation to characterize EF dynamics, and (2) better daytime LST evolution to get more accurate R_n estimates. As shown in Figure 7, RMSD and bias decrease dramatically after assimilation. Averaged among the four stations, RMSD decreases from 135.8 to 92.8 W/m^2, and bias decreases from −97.4 to −28.0 W/m^2.

Similar conclusions can be drawn from daytime *H* and *LE* estimates (an assessment of daytime fluxes is provided in the supporting information). For *H* estimates, R values increase dramatically after assimilation in all four stations, and RMSD is reduced by 8.7% on average. The improvement in *LE* estimates is much more

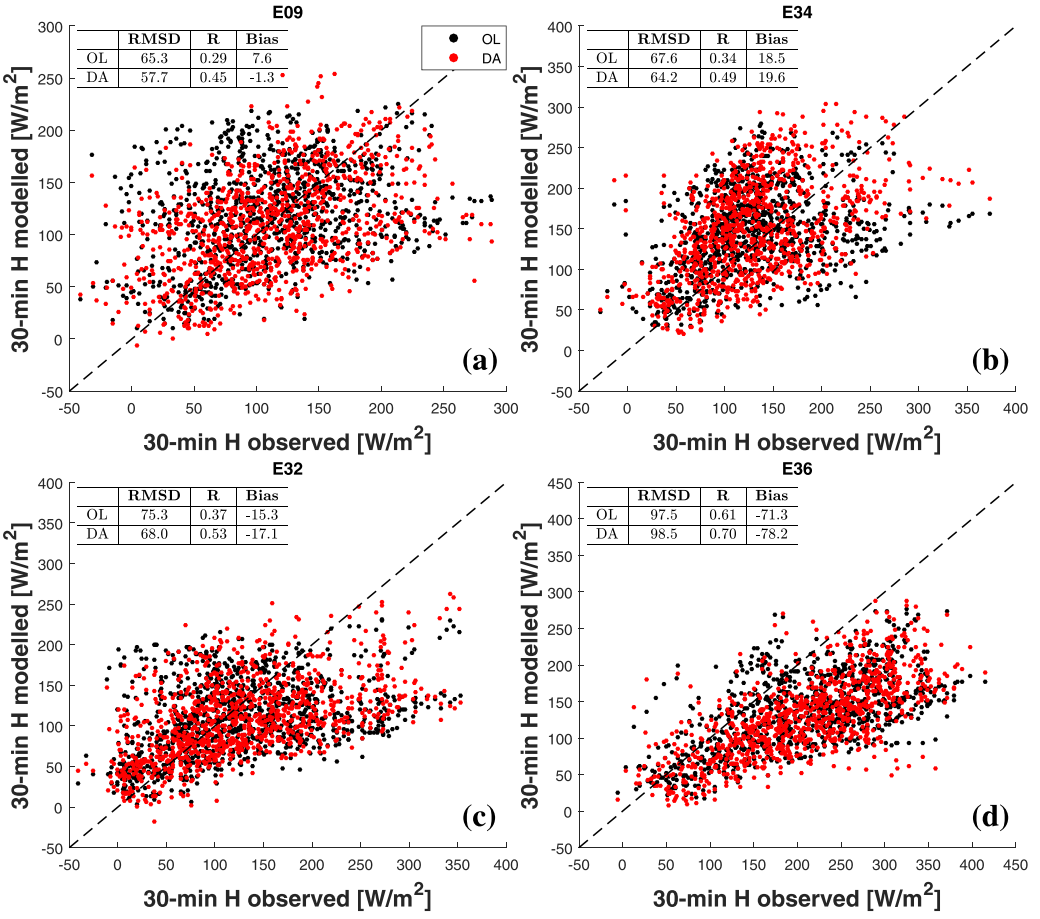


Figure 6. Thirty-minute *H* assessment against in situ observations. The unit of RMSD and Bias is W/m^2.

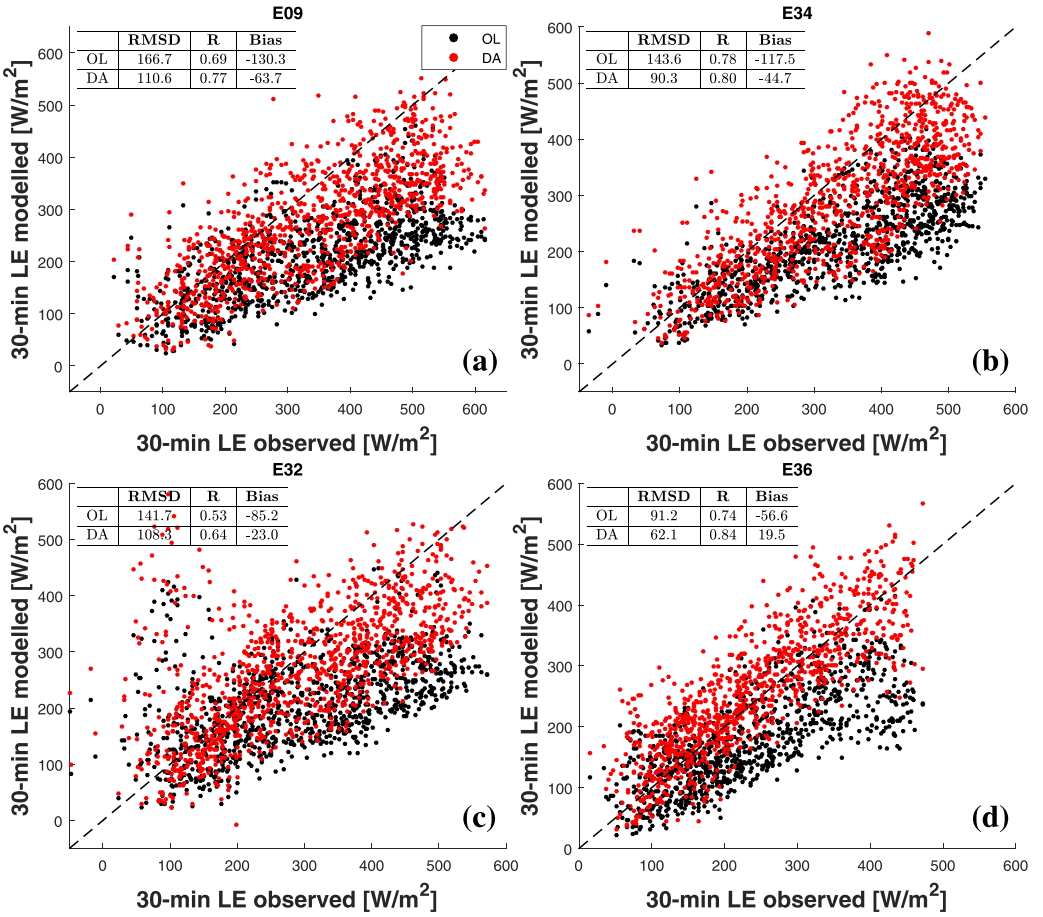


Figure 7. As in Figure 6 but for 30 min *LE* estimates.

evident than in *H* estimates. After assimilation, the RMSD and bias for *LE* are greatly reduced in all the stations. On average, the RMSD is reduced by 37% for *LE* estimates after assimilation. Daytime *LE* is generally underestimated from OL at all four stations except E36, indicated by the large negative bias values. After assimilation, the bias is reduced by 63% on average.

Time series of daytime *H* and *LE* estimates are plotted in Figure 8. Modeled 5 cm soil moisture and the corresponding SMAP observations are also plotted as a reference for soil wetness condition. The dynamics of observed *H* and *LE* agree very well with soil moisture evolution, which is consistent with the soil wetness control on surface energy partitioning. When data assimilation is performed, the magnitude of soil moisture is updated toward SMAP observations, which improves the characterization of EF dynamics. For example, during DOY 192 and 197, the OL soil moisture deviates significantly from SMAP observations in station E09 (Figure 8c), leading to substantial underestimation of *LE*. This may be caused by the failure in the forcing data to record some precipitation events. After assimilation, soil moisture gets much higher, and the high *LE* values are better modeled. The improvement from assimilation is also revealed in the better estimates of parameters. For example, from DOY 200 to 230, the modeled soil moisture at station E34 (Figure 8f) from OL and DA are almost the same, but the *LE* estimates from DA are in much better agreement with in situ observations. Since parameters are also updated during assimilation, more reasonable ranges are obtained for the parameters, and EF values are estimated more accurately.

Figure 9 shows the comparison between average modeled fluxes (R_n, *H*, and *LE*) and average in situ observations during daytime. The R_n estimates from OL are generally much lower than in situ observations. This is caused by the dry bias in soil moisture estimates (Figure 3) and warm bias in LST estimates (Figure 5) from OL. As a result, the outgoing shortwave and longwave radiation are higher, and the R_n estimates become lower, which explains the large negative bias in *LE* estimates from OL in Figure 7. The R_n estimates

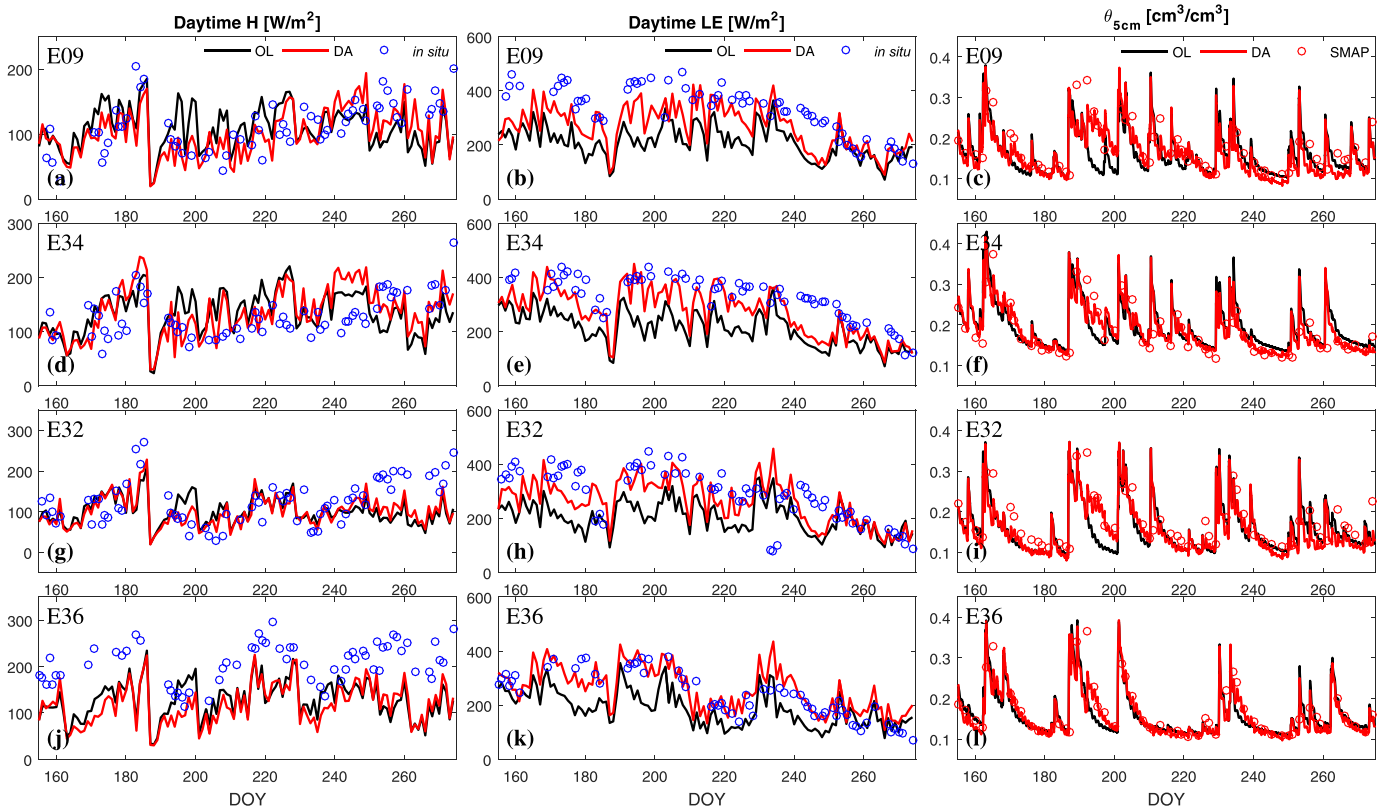


Figure 8. Time series of daytime flux estimates from OL, DA plotted against in situ measurements on days with assimilation at the four stations. The 5 cm soil moisture data from OL, DA, and the corresponding SMAP grid cells are also plotted in the column on the right.

from DA are in better agreement with in situ observations, which further proves the validity of the R_n estimation approach. *H* and *LE* estimates from OL hardly fall in the gray bands, which indicate 20% uncertainties around in situ observations caused by measurement uncertainty and surface heterogeneity. After assimilation, the flux daytime evolution is largely improved, especially for *LE* estimates.

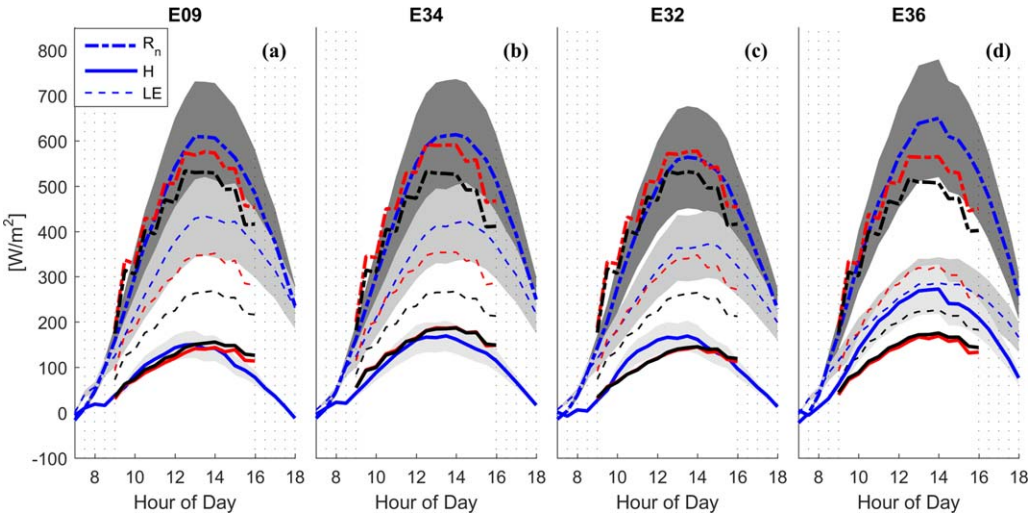


Figure 9. Modeled and measured average daytime cycles of R_n (dash-dot line), *H* (solid line), and *LE* (dashed line). The colors indicate results from DA (red), OL (black), and in situ observations (blue). The gray bands represent observed fluxes ±20% uncertainties.

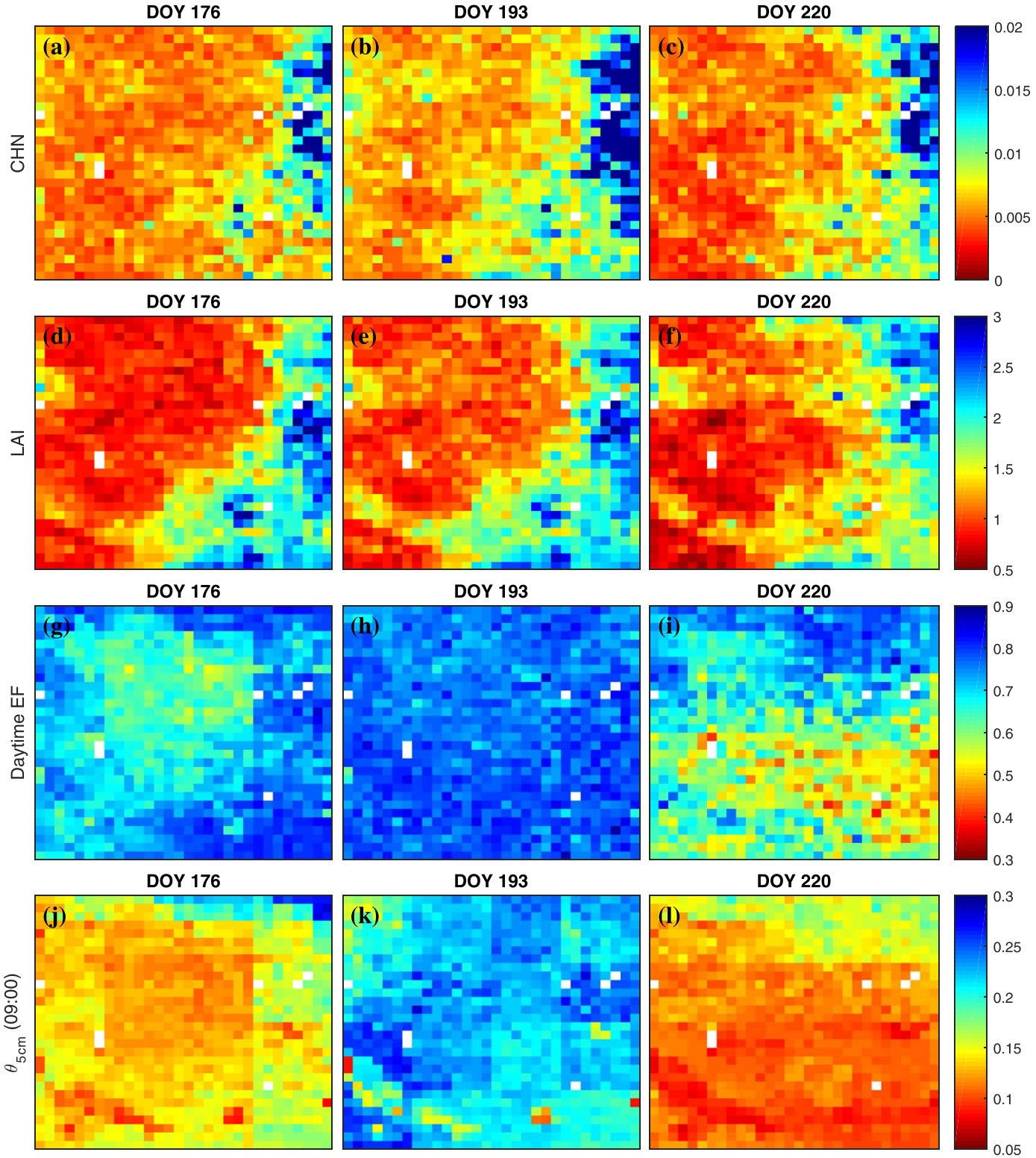


Figure 10. Spatial distribution of C_{HN}, LAI, EF, and 5 cm modeled soil moisture on 3 days (DOY 176, 193, 220).

3.3. Parameter Estimation

Figure 10 shows the C_{HN} and EF estimates and the corresponding LAI and modeled 5 cm soil moisture on 3 days with different wetness conditions. A positive gradient of LAI from the northwest to the southeast is clearly seen in Figures 10d–10f, which influences the spatial pattern of C_{HN}. During the modeling period, the area average LAI value increases from about 1.22 on DOY 155 to about 1.35 on DOY 170 and decreases to about 1.28 on DOY 185, and then increases quickly to about 1.48 on DOY 201, followed by a sharp decrease to about 0.8 on DOY 274. During DOY 176 and 193, C_{HN} increases dramatically, particularly in the western part of the area, which is largely caused by the increase in LAI. From DOY 193 to 220, LAI decreases greatly, especially in the eastern part. As a result, a decrease is observed in C_{HN} estimates. The magnitude of estimated C_{HN} is comparable to the reported values in literature (Bateni et al., 2014; Caparrini et al., 2004b; Xu et al., 2014).

The estimated EF distribution is shown in Figures 10g–10i. On DOY 176, the area is in the process of a long dry-down event from DOY 170 to 187, and the eastern part is much wetter than the western part. As a result, the EF estimates demonstrate an increasing trend eastward. Several precipitation events occurred shortly before DOY 193, therefore the area is very wet, and EF values are generally very high. From DOY 200 until DOY 225, the precipitation pattern is very different in the north and south. As is demonstrated in Figures 3 and 8, precipitation in the north features several large precipitation events, while the south experiences a long dry-down. On DOY 220, the soil moisture in the north is much higher than that in the south, and the same pattern is seen in the EF map. Some grid cells with EF values greater than 0.5 exist in the south, which is caused by the heterogeneity in the soil properties.

3.4. Added Value of SMAP Soil Moisture Data

To evaluate the added value of assimilating the coarse SMAP soil moisture data in flux estimation, the model is also run with only GOES LST data assimilated (hereafter DA$_T$) and with only SMAP soil moisture data assimilated (hereafter DA$_\theta$). The comparison of 30 min flux estimates from DA, DA$_T$, and DA$_\theta$ is shown in Figure 11.

When only SMAP soil moisture data are assimilated, the soil moisture estimates from DA$_\theta$ will be similar to those from DA. As a result, DA$_\theta$ is able to yield comparable R values in both *H* and *LE* estimates with DA,

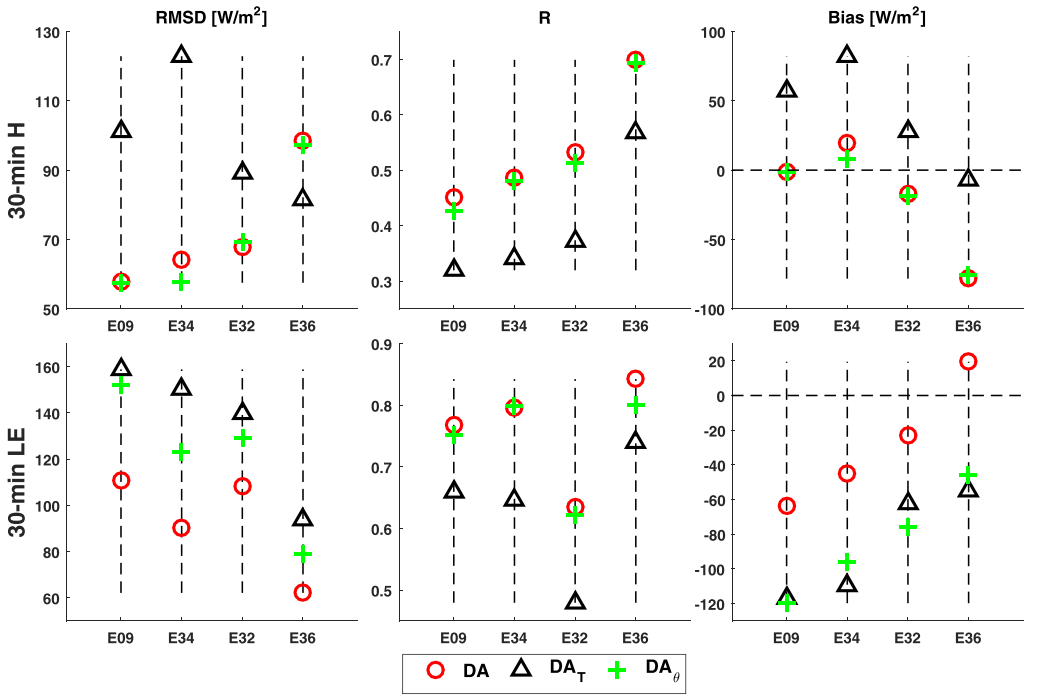


Figure 11. Comparison of 30-min *H* and *LE* estimates from DA, DA$_T$ (only GOES LST assimilated) and DA$_\theta$ (only SMAP soil moisture assimilated).

which again proves the tight coupling between soil moisture and heat flux partitioning. However, since the parameters and LST states cannot be updated through soil moisture assimilation, the LST evolution from DA_θ is similar with LST from OL. As a result, large RMSD and bias exist in the *LE* estimates as is in OL estimates.

When only GOES LST data are assimilated, DA_T performs worse than DA in flux estimates in station E09, E34, and E32 based on all the metrics. In station E36, DA_T produces smaller RMSD and bias in *H* estimates. This is caused by the large underestimation of R_n at this station (Figure 9). As a result, the RMSD and bias in *LE* estimates become larger in DA_T. In general, assimilating SMAP soil moisture data leads to significant improvement in flux estimates, despite the coarse resolution.

It should be noted that in station E09 and E34, the flux estimates from DA_T can be even worse than OL results, featured by a significant overestimation of *H* and underestimation of *LE*. This is caused by the poor characterization of soil moisture evolution. As soil moisture data are not assimilated, the soil moisture estimates from DA_T are similar to those from OL. As is discussed in section 3.2, when some precipitation events are not recorded in the forcing data, soil moisture estimates from DA_T will exhibit a dry bias, and EF will be very low accordingly. Since the model is still constrained by GOES LST data in DA_T, the particles with very large C_{HN} values will get larger weights to fit the LST observations. Consequently, the *H* estimates from DA_T will be much higher and the *LE* estimates much lower compared to OL results, leading to worse flux estimates. This demonstrates that despite the very coarse resolution, assimilating SMAP soil moisture data is not only beneficial but also crucial for successful and robust flux estimation, particularly when the uncertainties in forcing data, model structure or parameters are large.

4. Conclusions

This study proposes a methodology to estimate surface heat fluxes by assimilating SMAP soil moisture data and GOES LST data using a hybrid particle assimilation strategy. The methodology is based on a dual-source (DS) surface energy balance model, in which the contributions from soil and canopy are calculated separately. Fluxes are derived by estimating four parameters which are used to calculate a neutral bulk heat transfer coefficient (C_{HN}) and an evaporative fraction (EF). In the hybrid particle assimilation strategy, SMAP soil moisture data are assimilated using the particle filter, and GOES LST data are assimilated within the daytime window (09:00–16:00 LT) using an adaptive particle batch smoother. In this way, the resolution gap between SMAP and GOES data is bridged, and the flux estimation becomes independent of in situ flux observations. The methodology is applied over an area in the U.S. Southern Great Plains, and the results are compared with in situ observations.

In general, the modeled soil moisture from OL and DA follows the dynamics of in situ observations very well. After assimilation, modeled surface soil moisture is updated toward the in situ observations, which is most evident in dry-down events. In addition, the spatial heterogeneity of soil moisture within a SMAP grid cell is preserved, since the model grids are of finer resolution than SMAP grids. The improvement in LST estimates is also significant after assimilation. Here no bias correction is applied to GOES LST data. Some studies have suggested that the lack of prior scaling of LST data may lead to unrealistic surface heat flux estimates (Reichle et al., 2010). Future work will consider the influence of LST bias correction on flux estimates.

H and *LE* estimates are improved after assimilation at both daytime and 30 min scales. The improvement is more evident in *LE* estimates. The time series of daytime *H* and *LE* as well as their mean daytime cycles are also better characterized after assimilation.

Despite the coarse resolution, assimilating SMAP soil moisture data leads to large improvement in flux estimates. As the estimation is influenced by the uncertainties in forcing data, model structure and parameters, modeled soil moisture may be biased. By assimilating SMAP soil moisture data, the bias is reduced, which facilitates the estimation of EF. Therefore, assimilating SMAP soil moisture data is not only beneficial but also crucial for successful and robust flux estimation, particularly when the uncertainties are large. Here the 36 km SMAP soil moisture data are used, so the methodology should also be applicable for other soil moisture products with coarse resolution (e.g., SMOS, AMSR-E, AMSR2, and ASCAT). Further improvement is expected if soil moisture data with finer resolution (e.g., SMAP 9 km enhanced soil moisture product) are used.

The forcing data used in this study are from remote sensing and reanalysis products, and in situ flux observations are not required to calibrate the model. As the dependence on in situ data is minimized, the methodology can be easily applied to other areas.

Acknowledgments
The first author was financially supported for his PhD research by the China Scholarship Council (CSC) with the project reference number 201306040112. Some colormaps used in this manuscript are from Thyng et al. (2016). The data sets used in this study can be accessed through the sources provided in section 2.2. The authors thank the providers for the use of their data.

References

Abdolghafoorian, A., Farhadi, L., Bateni, S. M., Margulis, S., & Xu, T. (2017). Characterizing the effect of vegetation dynamics on the bulk heat transfer coefficient to improve variational estimation of surface turbulent fluxes. *Journal of Hydrometeorology*, *18*(2), 321–333.

Allen, R. G., Tasumi, M., & Trezza, R. (2007). Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)-Model. *Journal of Irrigation and Drainage Engineering*, *133*(4), 380–394.

Anderson, M., Kustas, W., Norman, J., Hain, C., Mecikalski, J., Schultz, L., . . . Gao, F. (2011). Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. *Hydrology and Earth System Sciences*, *15*(1), 223–239.

Anderson, M., Norman, J., Diak, G., Kustas, W., & Mecikalski, J. (1997). A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. *Remote Sensing of Environment*, *60*(2), 195–216.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., . . . Wofsy, S. (2001). FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, *82*(11), 2415–2434.

Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J., & Anderson, C. (2007). Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT). *Geophysical Research Letters*, *34*, L20401. https://doi.org/10.1029/2007GL031088

Bastiaanssen, W., Menenti, M., Feddes, R., & Holtslag, A. (1998a). A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. *Journal of Hydrology*, *212*, 198–212.

Bastiaanssen, W., Pelgrum, H., Wang, J., Ma, Y., Moreno, J., Roerink, G., & Van der Wal, T. (1998b). A remote sensing surface energy balance algorithm for land (SEBAL). Part 2: Validation. *Journal of Hydrology*, *212*, 213–229.

Bateni, S., & Entekhabi, D. (2012). Surface heat flux estimation with the ensemble Kalman smoother: Joint estimation of state and parameters. *Water Resources Research*, *48*, W08521. https://doi.org/10.1029/2011WR011542

Bateni, S., Entekhabi, D., & Castelli, F. (2013b). Mapping evaporation and estimation of surface control of evaporation using remotely sensed land surface temperature from a constellation of satellites. *Water Resources Research*, *49*, 950–968. https://doi.org/10.1002/wrcr.20071

Bateni, S., Entekhabi, D., & Jeng, D.-S. (2013a). Variational assimilation of land surface temperature and the estimation of surface energy balance components. *Journal of Hydrology*, *481*, 143–156.

Bateni, S., Entekhabi, D., Margulis, S., Castelli, F., & Kergoat, L. (2014). Coupled estimation of surface heat fluxes and vegetation dynamics from remotely sensed land surface temperature and fraction of photosynthetically active radiation. *Water Resources Research*, *50*, 8420–8440. https://doi.org/10.1002/2013WR014573

Bateni, S., & Liang, S. (2012). Estimating surface energy fluxes using a dual-source data assimilation approach adjoined to the heat diffusion equation. *Journal of Geophysical Research*, *117*, D17118. https://doi.org/10.1029/2012JD017618

Bell, J. E., Palecki, M. A., Baker, C. B., Collins, W. G., Lawrimore, J. H., Leeper, M. E., . . . Diamond, H. J. (2013). US Climate Reference Network soil moisture and temperature observations. *Journal of Hydrometeorology*, *14*(3), 977–988.

Boni, G., Entekhabi, D., & Castelli, F. (2001). Land data assimilation with satellite measurements for the estimation of surface energy balance components and surface control on evaporation. *Water Resources Research*, *37*(6), 1713–1722.

Cai, X., Pan, M., Chaney, N. W., Colliander, A., Misra, S., Cosh, M. H., . . . Wood, E. F. (2017). Validation of SMAP soil moisture for the SMAP-VEX15 field campaign using a hyper-resolution model. *Water Resources Research*, *53*, 3013–3028. https://doi.org/10.1002/2016WR019967

Caparrini, F., Castelli, F., & Entekhabi, D. (2003). Mapping of land-atmosphere heat fluxes and surface parameters with remote sensing data. *Boundary-Layer Meteorology*, *107*(3), 605–633.

Caparrini, F., Castelli, F., & Entekhabi, D. (2004a). Estimation of surface turbulent fluxes through assimilation of radiometric surface temperature sequences. *Journal of Hydrometeorology*, *5*(1), 145–159.

Caparrini, F., Castelli, F., & Entekhabi, D. (2004b). Variational estimation of soil and vegetation turbulent transfer and heat flux parameters from sequences of multisensor imagery. *Water Resources Research*, *40*, W12515. https://doi.org/10.1029/2004WR003358

Carlson, T. (2007). An overview of the "triangle method" for estimating surface evapotranspiration and soil moisture from satellite imagery. *Sensors*, *7*(8), 1612–1629.

Castelli, F., Entekhabi, D., & Caporali, E. (1999). Estimation of surface heat flux and an index of soil moisture using adjoint-state surface energy balance. *Water Resources Research*, *35*(10), 3115–3125.

Chan, S. K., Bindlish, R., O'neill, P. E., Njoku, E., Jackson, T., Colliander, A., . . . Kerr, Y. (2016). Assessment of the SMAP passive soil moisture product. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(8), 4994–5007.

Choudhury, B. J. (1987). Relationships between vegetation indices, radiation absorption, and net photosynthesis evaluated by a sensitivity analysis. *Remote Sensing of Environment*, *22*(2), 209–233.

Colliander, A., Jackson, T., Bindlish, R., Chan, S., Das, N., Kim, S., . . . Yueh, S. (2017). Validation of SMAP surface soil moisture products with core validation sites. *Remote Sensing of Environment*, *191*, 215–231.

Crago, R., & Brutsaert, W. (1996). Daytime evaporation and the self-preservation of the evaporative fraction and the Bowen ratio. *Journal of Hydrology*, *178*(1), 241–255.

Crago, R. D. (1996). Conservation and variability of the evaporative fraction during the daytime. *Journal of Hydrology*, *180*(1), 173–194.

Crow, W. T., & Kustas, W. P. (2005). Utility of assimilating surface radiometric temperature observations for evaporative fraction and heat transfer coefficient retrieval. *Boundary-Layer Meteorology*, *115*(1), 105–130.

Das, N. (2013). *SMAP ancillary data report: Soil attributes* (Rep. JPL D-53058). Pasadena: Jet Propulsion Laboratory, California Institute of Technology.

DeChant, C. M., & Moradkhani, H. (2012). Examining the effectiveness and robustness of sequential data assimilation methods for quantification of uncertainty in hydrologic forecasting. *Water Resources Research*, *48*, W04518. https://doi.org/10.1029/2011WR011011

De Lannoy, G. J., & Reichle, R. H. (2016). Assimilation of SMOS brightness temperatures or soil moisture retrievals into a land surface model. *Hydrology and Earth System Sciences*, *20*(12), 4895–4911.

Dirmeyer, P. A., Zeng, F. J., Ducharne, A., Morrill, J. C., & Koster, R. D. (2000). The sensitivity of surface fluxes to soil water content in three land surface schemes. *Journal of Hydrometeorology*, *1*(2), 121–134.

Dong, J., Steele-Dunne, S. C., Judge, J., & van de Giesen, N. (2015). A particle batch smoother for soil moisture estimation using soil temperature observations. *Advances in Water Resources*, *83*, 111–122.
Dong, J., Steele-Dunne, S. C., Ochsner, T. E., Hatch, C. E., Sayde, C., Selker, J., . . . van de Giesen, N. (2016a). Mapping high-resolution soil moisture and properties using distributed temperature sensing data and an adaptive particle batch smoother. *Water Resources Research*, *52*, 7690–7710. https://doi.org/10.1002/2016WR019031
Dong, J., Steele-Dunne, S. C., Ochsner, T. E., & van de Giesen, N. (2016b). Estimating soil moisture and soil thermal and hydraulic properties by assimilating soil temperatures using a particle batch smoother. *Advances in Water Resources*, *91*, 104–116.
Entekhabi, D., Njoku, E. G., O'neill, P., Kellogg, K. H., Crow, W. T., Edelstein, W. N., . . . Van Zyl, J. (2010). The soil moisture active passive (SMAP) mission. *Proceedings of the IEEE*, *98*(5), 704–716.
Entekhabi, D., Yueh, S., O'neill, P. E., Kellogg, K. H., Allen, A., Bindlish, R., . . . West, R. (2014). *SMAP handbook* (pp. 400–1567). Pasadena, CA: JPL Publication.
Farhadi, L., Entekhabi, D., & Salvucci, G. (2016). Mapping land water and energy balance relations through conditional sampling of remote sensing estimates of atmospheric forcing and surface states. *Water Resources Research*, *52*, 2737–2752. https://doi.org/10.1002/2015WR017680
Farhadi, L., Entekhabi, D., Salvucci, G., & Sun, J. (2014). Estimation of land surface water and energy balance parameters using conditional sampling of surface states. *Water Resources Research*, *50*, 1805–1822. https://doi.org/10.1002/2013WR014049
Freitas, S. C., Trigo, I. F., Macedo, J., Barroso, C., Silva, R., & Perdigão, R. (2013). Land surface temperature from multiple geostationary satellites. *International Journal of Remote Sensing*, *34*(9–10), 3051–3068.
Gentine, P., Entekhabi, D., Chehbouni, A., Boulet, G., & Duchemin, B. (2007). Analysis of evaporative fraction diurnal behaviour. *Agricultural and Forest Meteorology*, *143*(1), 13–29.
Gillies, R., Kustas, W., & Humes, K. (1997). A verification of the'triangle'method for obtaining surface soil water content and energy fluxes from remote measurements of the Normalized Difference Vegetation Index (NDVI) and surface radiant temperature. *International Journal of Remote Sensing*, *18*(15), 3145–3166.
Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., . . . Iguchi, T. (2014). The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, *95*(5), 701–722.
Houldcroft, C. J., Grey, W. M., Barnsley, M., Taylor, C. M., Los, S. O., & North, P. R. (2009). New vegetation albedo parameters and global fields of soil background albedo derived from MODIS for use in a climate model. *Journal of Hydrometeorology*, *10*(1), 183–198.
Huffman, G. J., Bolvin, D. T., & Nelkin, E. J. (2015). Integrated Multi-satellitE Retrievals for GPM (IMERG) technical documentation (NASA/GSFC Code 612, 47). College Park: University of Maryland.
Idso, S., Jackson, R., Reginato, R., Kimball, B., & Nakayama, F. (1975). The dependence of bare soil albedo on soil water content. *Journal of Applied Meteorology*, *14*(1), 109–113.
Imaoka, K., Kachi, M., Fujii, H., Murakami, H., Hori, M., Ono, A., . . . Shimoda, H. (2010). Global Change Observation Mission (GCOM) for monitoring carbon, water cycles, and climate change. *Proceedings of the IEEE*, *98*(5), 717–734.
Jiang, L., & Islam, S. (2001). Estimation of surface evaporation map over Southern Great Plains using remote sensing data. *Water Resources Research*, *37*(2), 329–340.
Jung, M., Reichstein, M., & Bondeau, A. (2009). Towards global empirical upscaling of fluxnet eddy covariance observations: Validation of a model tree ensemble approach using a biosphere model. *Biogeosciences*, *6*(10), 2001–2013.
Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., . . . Williams, C. (2011). Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research*, *116*, G00J07. https://doi.org/10.1029/2010JG001566
Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Martinuzzi, J., Font, J., & Berger, M. (2001). Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(8), 1729–1735.
Kim, S.-B., Moghaddam, M., Tsang, L., Burgin, M., Xu, X., & Njoku, E. G. (2014). Models of L-band radar backscattering coefficients over global terrain for soil moisture retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(2), 1381–1396.
Knyazikhin, Y., Glassy, J., Privette, J. L., Tian, Y., Lotsch, A., Zhang, Y., . . . Running, S. W. (1999). *MODIS leaf area index (LAI) and fraction of photosynthetically active radiation absorbed by vegetation (FPAR) product (MOD15) algorithm theoretical basis document* (Theoretical Basis Document). Greenbelt, MD: NASA Goddard Space Flight Center.
Kumar, J., Hoffman, F. M., Hargrove, W. W., & Collier, N. (2017). Understanding the representativeness of FLUXNET for upscaling carbon flux from eddy covariance measurements. *Earth System Science Data Discuss*, 1–25. https://doi.org/10.5194/essd-2016-36
Kustas, W., Humes, K., Norman, J., & Moran, M. (1996). Single-and dual-source modeling of surface energy fluxes with radiometric surface temperature. *Journal of Applied Meteorology*, *35*(1), 110–121.
Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., . . . Sobrino, J. A. (2013). Satellite-derived land surface temperature: Current status and perspectives. *Remote Sensing of Environment*, *131*, 14–37.
Lu, S., Ju, Z., Ren, T., & Horton, R. (2009). A general approach to estimate soil water content from thermal inertia. *Agricultural and Forest Meteorology*, *149*(10), 1693–1698.
Lu, Y., Dong, J., Steele-Dunne, S. C., & van de Giesen, N. (2016). Estimating surface turbulent heat fluxes from land surface temperature and soil moisture observations using the particle batch smoother. *Water Resources Research*, *52*, 9086–9108. https://doi.org/10.1002/2016WR018943
Margulis, S. A., Girotto, M., Cortés, G., & Durand, M. (2015). A particle batch smoother approach to snow water equivalent estimation. *Journal of Hydrometeorology*, *16*(4), 1752–1772.
McCabe, M. F., & Wood, E. F. (2006). Scale influences on the remote estimation of evapotranspiration using multiple satellite sensors. *Remote Sensing of Environment*, *105*(4), 271–285.
Mladenova, I., Jackson, T., Njoku, E., Bindlish, R., Chan, S., Cosh, M., . . . Santi, E. (2014). Remote monitoring of soil moisture using passive microwave-based techniques-Theoretical basis and overview of selected algorithms for AMSR-E. *Remote Sensing of Environment*, *144*, 197–213.
Moradkhani, H., DeChant, C. M., & Sorooshian, S. (2012). Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-Markov chain Monte Carlo method. *Water Resources Research*, *48*, W12520. https://doi.org/10.1029/2012WR012144
Moradkhani, H., Hsu, K.-L., Gupta, H., & Sorooshian, S. (2005). Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resources Research*, *41*, W05012. https://doi.org/10.1029/2004WR003604
Nagler, P. L., Cleverly, J., Glenn, E., Lampkin, D., Huete, A., & Wan, Z. (2005b). Predicting riparian evapotranspiration from MODIS vegetation indices and meteorological data. *Remote Sensing of Environment*, *94*(1), 17–30.
Nagler, P. L., Scott, R. L., Westenburg, C., Cleverly, J. R., Glenn, E. P., & Huete, A. R. (2005a). Evapotranspiration on western US rivers estimated using the Enhanced Vegetation Index from MODIS and data from eddy covariance and Bowen ratio flux towers. *Remote Sensing of Environment*, *97*(3), 337–351.

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., . . . Xia, Y. (2011). The community Noah land surface model with multi-parameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research*, *116*, D12109. https://doi.org/10.1029/2010JD015139

Njoku, E. G., Jackson, T. J., Lakshmi, V., Chan, T. K., & Nghiem, S. V. (2003). Soil moisture retrieval from AMSR-E. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(2), 215–229.

Norman, J. M., Kustas, W. P., & Humes, K. S. (1995). A two-source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature. *Agricultural and Forest Meteorology*, *77*(3), 263–293.

Oleson, K. W., Lawrence, D., Gordon, M. B., Flanner, M., Kluzek, G., Peter, E., . . . Thornton, P. E. (2010). *Technical description of version 4.0 of the Community Land Model (CLM)* (NCAR Tech. Note NCAR/TN-478+STR). Boulder, CO: National Center for Atmospheric Research.

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

Reichle, R. H., Kumar, S. V., Mahanama, S. P., Koster, R. D., & Liu, Q. (2010). Assimilation of satellite-derived skin temperature observations into land surface models. *Journal of Hydrometeorology*, *11*(5), 1103–1122.

Schaap, M. G., Leij, F. J., & Van Genuchten, M. T. (2001). ROSETTA: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal of Hydrology*, *251*(3), 163–176.

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.

Sellers, P., Mintz, Y., Sud, Y., & Dalcher, A. (1986). A simple biosphere model (SiB) for use within general circulation models. *Journal of the Atmospheric Sciences*, *43*(6), 505–531.

Sini, F., Boni, G., Caparrini, F., & Entekhabi, D. (2008). Estimation of large-scale evaporation fields based on assimilation of remotely sensed land temperature. *Water Resources Research*, *44*, W06410. https://doi.org/10.1029/2006WR005574

Sobrino, J., & Romaguera, M. (2004). Land surface temperature retrieval from msg1-seviri data. *Remote Sensing of Environment*, *92*(2), 247–254.

Stordal, A. S., Karlsen, H. A., Nævdal, G., Skaug, H. J., & Vallès, B. (2011). Bridging the ensemble Kalman filter and particle filters: The adaptive Gaussian mixture filter. *Computational Geosciences*, *15*(2), 293–305.

Su, Z. (2002). The surface energy balance system (SEBS) for estimation of turbulent heat fluxes. *Hydrology and Earth System Sciences Discussions*, *6*(1), 85–100.

Sun, D., & Pinker, R. T. (2003). Estimation of land surface temperature from a Geostationary Operational Environmental Satellite (GOES-8). *Journal of Geophysical Research*, *108*(D11), 4326. https://doi.org/10.1029/2002JD002422

Tang, R., Li, Z.-L., & Tang, B. (2010). An application of the T*s*–VI triangle method with enhanced edges determination for evapotranspiration estimation from MODIS data in arid and semi-arid regions: Implementation and Validation. *Remote Sensing of Environment*, *114*(3), 540–551.

Thyng, K. M., Greene, C. A., Hetland, R. D., Zimmerle, H. M., & DiMarco, S. F. (2016). True colors of oceanography: Guidelines for effective and accurate colormap selection. *Oceanography*, *29*(3), 9–13.

Timmermans, W. J., Kustas, W. P., Anderson, M. C., & French, A. N. (2007). An intercomparison of the surface energy balance algorithm for land (SEBAL) and the two-source energy balance (TSEB) modeling schemes. *Remote Sensing of Environment*, *108*(4), 369–384.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., . . . Mocko, D. (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

Xu, T., Bateni, S. M., & Liang, S. (2015). Estimating turbulent heat fluxes with a weak-constraint data assimilation scheme: A case study (hiWATER-MUSOEXE). *IEEE Geoscience and Remote Sensing Letters*, *12*(1), 68–72.

Xu, T., Bateni, S., Liang, S., Entekhabi, D., & Mao, K. (2014). Estimation of surface turbulent heat fluxes via variational assimilation of sequences of land surface temperatures from Geostationary Operational Environmental Satellites. *Journal of Geophysical Research: Atmospheres*, *119*, 10–780. https://doi.org/10.1002/2014JD021814

Yan, H., DeChant, C. M., & Moradkhani, H. (2015). Improving soil moisture profile prediction with the particle filter-Markov chain Monte Carlo method. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(11), 6134–6147.

Yan, H., & Moradkhani, H. (2016). Combined assimilation of streamflow and satellite soil moisture with the particle filter and geostatistical modeling. *Advances in Water Resources*, *94*, 364–378.

Zeng, J., Chen, K.-S., Bi, H., & Chen, Q. (2016). A preliminary evaluation of the SMAP radiometer soil moisture product over United States and Europe using ground-based measurements. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(8), 4929–4940.

Zeng, X. (2001). Global vegetation root distribution for land modeling. *Journal of Hydrometeorology*, *2*(5), 525–530.

Zheng, D., van der Velde, R., Su, Z., Wang, X., Wen, J., Booij, M. J., . . . Chen, Y. (2015). Augmentations to the Noah model physics for application to the Yellow River source area. Part II: Turbulent heat fluxes and soil heat transport. *Journal of Hydrometeorology*, *16*(6), 2677–2694.