Mapping Regional Turbulent Heat Fluxes via Assimilation of MODIS Land Surface Temperature Data into an Ensemble Kalman Smoother Framework

Xinlei He1, Tongren Xu2, Sayed M. Bateni3, Christopher M.U. Neale4, Shaomin Liu1, Thomas Auligne4,5, Kaicun Wang6, and Shoudong Zhu7

1State Key Laboratory of Earth Surface Processes and Resource Ecology, School of Natural Resources, Faculty of Geography, Beijing Normal University, Beijing, China, 2Department of Civil and Environmental Engineering and Water Resources Research Center, University of Hawaii at Manoa, Honolulu, HI, 3Daugherty Water for Food Global Institute, University of Nebraska, Lincoln, NE, 4Joint Center for Satellite Data Assimilation, College Park, MD, 5University Corporation for Atmospheric Research, Boulder, CO, 6College of Global Change and Earth System Science, Beijing Normal University, Beijing, China, 7State Key Laboratory of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China

Abstract

Estimation of turbulent heat fluxes via variational data assimilation (VDA) approaches has been the subject of several studies. The VDA approaches need an adjoint model that is difficult to derive. In this study, remotely sensed land surface temperature (LST) data from the Moderate Resolution Imaging Spectroradiometer (MODIS) are assimilated into the heat diffusion equation within an ensemble Kalman smoother (EnKS) approach to estimate turbulent heat fluxes. The EnKS approach is tested in the Heihe River Basin (HRB) in northwest China. The results show that the EnKS approach can estimate turbulent heat fluxes by assimilating low temporal resolution LST data from MODIS. The findings indicate that the EnKS approach performs fairly well in various hydrological and vegetative conditions. The estimated sensible (H) and latent (LE) heat fluxes are compared with the corresponding observations from large aperture scintillometer systems at three sites (namely, Arou, Daman, and Sidaqiao) in the HRB. The turbulent heat flux estimates from EnKS agree reasonably well with the observations, and are comparable to those of the VDA approach. The EnKS approach also provides statistical information on the H and LE estimates. It is found that the uncertainties of H and LE estimates are higher over wet and/or densely vegetated areas (grassland and forest) compared to the dry and/or slightly vegetated areas (cropland, shrubland, and barren land).

1. Introduction

Sensible (H) and latent (LE) heat fluxes play a crucial role in the exchange of energy, water, and carbon between the land surface and overlying atmosphere. These fluxes are required in many disciplines such as meteorology, hydrology, ecology, agronomy, and water resources planning and management (Dickinson, 1987; Mascart et al., 1991; Murray et al., 2012; Wang, & Dickinson, 2012).

Flux tower networks have been used to measure turbulent heat fluxes in many projects (e.g., FLUXNET, AmeriFlux, FIFE, BEAREX08, HiWATER, etc.) (Baldocchi et al., 2001; Baldocchi, & Ma, 2011; Evett et al., 2011; Liu et al., 2018; Sellers et al., 1988). However, these measurements are sparse and available only over limited time periods, causing difficulty to characterize the spatiotemporal patterns in the turbulent heat fluxes over large-scale domains (Baldocchi et al., 2001; Serreze et al., 2005; Xu, Guo, et al., 2018, 2019).

Hence, a number of methods have been developed to estimate turbulent heat fluxes from remotely sensed land surface temperature (LST) data. Broadly speaking, these approaches fall into two main categories: retrieval-based (e.g., Anderson et al., 1997; Bastiaanssen, Menenti, et al., 1998; Bastiaanssen, Pelgrum, et al., 1998; Carlson, 2007; Jia et al., 2009; Jiang & Islam, 2003; Kustas et al., 2012; Liang et al., 2010; Liu et al., 2007; Ma et al., 2018; Mallick et al., 2013, 2014; Moran et al., 1994; Norman et al., 1995; Song et al., 2018; Su 2002; Sun et al., 2013; Tang et al., 2010; Wang et al., 2006; Yao et al., 2013; Zhu et al., 2017), and data assimilation approaches (e.g., Abdolghafoorian et al., 2017; Bateni, Entekhabi, & Jeng, 2013; Bateni,
Due to the abovementioned drawbacks of VDA approaches, ensemble schemes cannot directly provide information on the uncertainty of their estimates. Data assimilation approaches can be categorized into two main groupings: 1) variational schemes, and 2) ensemble schemes. Variational data assimilation (VDA) methods estimate the key unknowns of the surface energy balance (SEB) equations (i.e., neutral bulk heat transfer coefficient, $C_{HN}$, and evaporative fraction, EF) by assimilating LST observations into the heat diffusion or force-restore equations (Abdolghafoorian et al., 2017; Bateni, Entekhabi, & Jeng, 2013; Bateni, Entekhabi, & Castelli, 2013; Bateni et al., 2014; Bateni & Entekhabi, 2012a; Bateni & Liang, 2012; Crow & Kustas, 2005; He et al., 2018; Qin et al., 2007; Sini et al., 2008; Xu et al., 2014, 2016; Xu, Bateni, et al., 2015; Xu, He, et al., 2019). The unknown parameters of the VDA approaches (i.e., $C_{HN}$ and EF) are obtained by minimizing the misfit between the LST observations and estimations from the force-restore or heat diffusion equation. VDA approaches need an adjoint term, which is difficult to derive and code. They also cannot directly provide information on the uncertainty of their estimates.

Due to the abovementioned drawbacks of VDA approaches, ensemble-based data assimilation techniques (e.g., ensemble Kalman smoother, EnKS) have been used widely in hydrology (Bateni & Entekhabi, 2012b; Dunne & Entekhabi, 2005, 2006; Lei et al., 2014; Reichle et al., 2008; Xu, Bateni, et al., 2018). Compared to the VDA approaches, the EnKS methods have some unique characteristics: (1) they do not need an adjoint model, and thus they are relatively easy to implement; (2) they can easily generate background-error covariance and directly provide uncertainty of their estimates, and (3) they are able to account for a wide range of possible model and measurement errors (Bateni & Entekhabi, 2012b; Kalnay et al., 2007; Reichle et al., 2002; Whitaker et al., 2009).

Bateni and Entekhabi (2012b) assimilated half-hourly ground-based LST measurements into the EnKS system to estimate turbulent heat fluxes at the First International Satellite Land Surface Climatology Project Field Experiment (FIFE) site. In a follow-up study, Xu, Bateni, et al. (2018) estimated $H$ and LE by assimilating hourly LST data from the Geostationary Operational Environmental Satellite into the EnKS system. LST observations with high temporal resolution were used in both of the abovementioned works because they can capture the diurnal cycle of LST. However, high frequency LST data from Geostationary Operational Environmental Satellite have relatively low spatial resolution and are often unavailable in high latitudes (Zhang et al., 2014). In contrast, LST data from the Moderate Resolution Imaging Spectroradiometer (MODIS) have high spatial resolution (1 km × 1 km) as well as global coverage.

In this study, MODIS LST data (from Aqua and Terra platforms) are assimilated into the heat diffusion equation via the EnKS system to estimate turbulent heat fluxes. The two key unknown parameters of the EnKS approach are the neutral heat transfer coefficient ($C_{HN}$) and evaporative fraction (EF). $C_{HN}$ governs the sum of turbulent heat fluxes ($H + LE$), and EF represents their partitioning. The EnKS approach is tested in the Heihe River Basin (HRB), which samples various hydrological conditions. The turbulent heat fluxes observations from the large aperture scintillometer (LAS) at three sites (namely, Arou, Daman, and Sidaqiao) in the HRB are used to validate the EnKS estimates. The uncertainty of estimated turbulent heat fluxes from EnKS is also evaluated under different environmental conditions. Finally, performance of EnKS is compared with those of the ensemble open loop (EnOL) (without assimilating MODIS LST data) and VDA approaches. The readers are referred to Xu, He, et al. (2019) for detailed information on the VDA approach.

The objectives of our study are to (1) estimate $H$ and LE by assimilating low temporal resolution LST data from polar orbiting satellites into an EnKS framework; (2) investigate performance of the EnKS framework in the HRB and validate the turbulent heat fluxes estimates with the large aperture scintillometer (LAS).
observations; (3) compare the results of the EnKS approach with those of Xu, He, et al. (2019) VDA approach; and (4) assess the uncertainty of EnKS H and LE estimates in different vegetative and hydrologic conditions.

This paper is organized as follows. Sections 2 and 3 explain the methods and models including the heat diffusion equation, SEB equation, and the EnKS algorithm. Section 4 describes the experiment sites and data sets. Section 5 provides the results and the discussion. Finally, conclusions are reported in section 6.

2. System Model

The SEB equation can be written as

$$R_n = H + LE + G,$$  \hspace{1cm} (1)

where $H$ and LE represent the sensible and latent heat fluxes (W m$^{-2}$), and $G$ is the ground heat flux (W m$^{-2}$). $R_n$ is the net radiation (W m$^{-2}$) and can be defined as

$$R_N = (1-\alpha)R_s^d + R_l^u - R_l^d,$$  \hspace{1cm} (2)

where $R_s^d$ is the downward shortwave radiation; $\alpha$ is the surface albedo ($-$), and $R_l^d$ and $R_l^u$ are the downward and upward longwave radiations, respectively (W m$^{-2}$).

Sensible heat flux can be represented in terms of the gradient between the LST ($T$) and air temperature ($T_a$),

$$H = \rho C_p C_H U (T - T_a),$$  \hspace{1cm} (3)

where $\rho$ is the air density (kg m$^{-3}$); $C_p$ is the air heat capacity (1,012 J kg$^{-1}$ K$^{-1}$), and $U$ is the reference-height wind speed (m s$^{-1}$). The bulk heat transfer coefficient ($C_H$) is related to the neutral bulk heat transfer coefficient ($C_{HN}$) and the atmospheric stability correction function ($f$) via (Abdolghafoorian et al., 2017; Bateni, Entekhabi, & Castelli, 2013; Xu et al., 2014, 2016)

$$C_H = C_{HN} f(R_i),$$  \hspace{1cm} (4)

where $R_i$ is the Richardson number. $C_{HN}$ depends on the vegetation phenology and is set to vary monthly (Caparrini et al., 2004a, 2004b; Crow & Kustas, 2005). $C_{HN}$ constitutes the first unknown parameter of the EnKS system.

The second unknown parameter of EnKS is EF, which is defined as

$$EF = LE/(H + LE).$$  \hspace{1cm} (5)

EF is nearly constant for near-peak radiation hours on days without precipitation (Tang & Li, 2017). In this study, EF is assumed to be constant during the assimilation window (0900–1800 LT) in each day, but it can vary from day to day.

The EnKS approach begins with the a priori EF estimate, and improves it via assimilation of LST. According to Xu, He, et al. (2019), the a priori EF estimate is obtained by

$$EF = EF_{\text{min}} + \frac{2(\text{EF}_{\text{max}} - \text{EF}_{\text{min}})}{\pi} \text{Arctan}(\varphi \tau),$$  \hspace{1cm} (6)

where $\text{EF}_{\text{max}}$ and $\text{EF}_{\text{min}}$ are the maximum and minimum EF values for a specific land cover type (e.g., barren land, grassland, cropland, forest, etc.); $\varphi$ is the calibration coefficient, and $\tau$ is the environmental index. Following Xu, He, et al. (2019), $\tau$ is obtained from a leaf area index (LAI) (for vegetated land) or apparent thermal inertia (ATI) (for barren land). The readers are referred to Xu, He, et al. (2019) for more information on $\tau$. The three unknown parameters in equation (6) (i.e., $\text{EF}_{\text{max}}$, $\text{EF}_{\text{min}}$, and $\varphi$) are estimated by the least-square method using EF measurements from 12 eddy covariance (EC) flux towers in the HRB, and corresponding $\tau$ values from LAI or ATI. The estimated $\text{EF}_{\text{max}}$, $\text{EF}_{\text{min}}$, and $\varphi$ values for four land cover types (i.e., cropland, grassland, forest, and barren land) in the HRB are reported in Table 1. The cropland is mainly covered by seeded corns. The grassland consists mainly of alpine meadow. The forest is a mixture of deciduous broadleaf forest, evergreen needleleaf forests, mixed forests,
and shrub. The barren land consists of barren or sparsely vegetated lands. Detailed information on the 12 EC flux towers in the HRB can be found in Xu, He, et al. (2019).

3. State-Augmented Ensemble Kalman Smoother

The state-augmented EnKS technique contains a forecast step and an update step. In the forecast step, the heat diffusion equation is used to estimate the dynamics of LST. Then, in the update step, the augmented state [i.e., estimated LST values from the heat diffusion equation, and the a priori EF estimates from equation (6)] is improved by assimilating MODIS LST data.

3.1. State and Parameter Propagation Models (Forecast Step)

In the forecast step, EnKS propagates an ensemble of LSTs by using the heat diffusion equation. The dynamics of soil temperature at depth $z$ and time $t (T(z, t))$ is governed by the heat diffusion equation, which is given by

$$C \frac{\partial T(z, t)}{\partial t} = K \frac{\partial^2 T(z, t)}{\partial z^2} + \omega,$$

where $C$ and $K$ are the soil volumetric heat capacity ($\text{J m}^{-3} \text{K}^{-1} \text{s}^{-1}$) and thermal conductivity ($\text{J m}^{-1} \text{K}^{-1}$), respectively, and $\omega$ is the model error.

Applying the heat diffusion equation requires specification of boundary conditions at the top and bottom of the soil column. The boundary condition at the top of the soil column is obtained by $-KdT(z = 0, t)/dz = G(t)$.

Table 1

| Land cover type | $E_{F_{\text{max}}}$ | $E_{F_{\text{min}}}$ | $\varphi$ |
|----------------|----------------------|----------------------|-----------|
| Barren land    | 0.67                 | 0.17                 | 2         |
| Grassland      | 0.86                 | 0.35                 | 7         |
| Cropland       | 0.96                 | 0.45                 | 7         |
| Forest         | 0.98                 | 0.06                 | 13        |

Figure 1. The land cover map in the Heihe River Basin.
The bottom boundary condition is \( dT(z = l, t) / dz = 0 \) (where \( l \) is the depth of bottom boundary). In this study, \( l \) is taken equal to 0.5 m because the soil temperature at the depth of 0.3–0.5 m is almost constant (Hu & Islam, 1995). The initial soil temperature (i.e., \( T(z, t = \tau_0) \), where \( \tau_0 \) is the initial time) in the soil column is obtained from reanalysis product (Shi et al., 2014).

To characterize the model error term, the uncertain input variables are perturbed by adding normally distributed random errors (with zero mean and specified variance) to them and generating ensembles. Random errors are created in a physically reasonable range to disturb uncertain inputs when creating ensembles, and reflect uncertainties in measurements, boundary conditions, and propagated states. It is hypothesized that the uncertain inputs are initial profile of soil temperature, heat diffusion coefficient (\( D = k/C \)), air temperature, wind speed, downward shortwave radiation, and augmented state vector variables (LST and EF) (Bateni & Entekhabi, 2012b). Following Bateni and Entekhabi (2012b) and Xu, Bateni, et al. (2018), normally distributed errors with a mean zero and a standard deviation of 2 K, 0.1 m s\(^{-1}\), 30 W m\(^{-2}\), and 0.05 are added to the initial soil temperature in the soil column, heat diffusion coefficient, air temperature, wind speed, downward shortwave radiation, and EF, respectively. These numbers are chosen based on the order of magnitude of uncertain inputs (i.e., initial profile of soil temperature, heat diffusion coefficient (\( D = k/C \)), air temperature, wind speed, downward shortwave radiation, and augmented state vector variables (LST and EF). For example, uncertainties in the heat diffusion coefficient (\( D = 0.63/2.58 \times 10^{-6} = 0.24 \times 10^{-6} \)) controls the heat diffusivity through the soil slab. To account for uncertainties in \( D \) and consequently heat diffusivity through the soil slab, a normally distributed noise with a mean of zero and a standard deviation of \( 0.1 \times 10^{-6} \text{ m}^2 \text{ s}^{-1} \) is added to \( D \) (Bateni & Entekhabi, 2012b; Xu, Bateni, et al., 2018).

An ensemble of initial conditions \( T_j(z, \tau_0) \) \( (j = 1, \ldots, N_e, \text{where } j \text{ denotes the } j \text{th ensemble member, and } N_e \text{ is the number of ensembles}) \) is generated by adding an ensemble of \( N_e \) model errors to the initial profile of soil temperature. This ensemble is integrated forward in time by the heat diffusion equation to generate the \( a \) priori estimates of LST (the so-called forecast or EnOL LST estimates) at each time step in the assimilation window (i.e., 0900–1800 LT). In this study, \( N_e \) is set to 100 (Bateni & Entekhabi, 2012b; Xu, Bateni, et al., 2018).

In addition to LST, an ensemble of EF is integrated forward in time. Since EF is assumed to be invariant in the assimilation window, it can be propagated by

\[
EF(t + 1) = EF(t) + \omega', \tag{8}
\]

where \( \omega' \) represents the uncertainty of EF, which is presumed to have a normal distribution. An ensemble of \( N_e \) model errors (\( \omega' \)) is added to the \( a \) priori EF estimate from equation (6) to generate an ensemble of the \( a \) priori EF estimates, \( EF_j \) \( (j = 1, \ldots, N_e, \) \( \text{This ensemble is propagated forward in time (starting from } t = \tau_0) \) by equation (8) to generate the \( a \) priori EF estimates (the so-called forecast EF estimates) at each time step in the assimilation window.

### 3.2. Joint State-Parameter Update Model (Update Step)

The state (LST) and model unknown parameters (\( C_{\text{UIN}} \) and EF) can be updated by the state-augmented EnKS approach in which EF is added into the system state vector (LST), and artificially treated as an additional state variable. The augmented state vector, \( \mathbf{T} \), is then defined as

\[
\mathbf{T}(t) = \begin{bmatrix}
T_1(t) & T_2(t) & \cdots & T_{N_e}(t) \\
EF_1(t) & EF_2(t) & \cdots & EF_{N_e}(t)
\end{bmatrix}, \tag{9}
\]

where the \( j \)th column of \( \mathbf{T}(t) \) contains the \( j \)th realization of the forecast LST [from equation (7)] and EF [from equation (8)] at time \( t \). The MODIS LST observation at time \( t \), \( T_{\text{obs}}(t) \), is related to the augmented state vector \( \mathbf{T}(t) \) by the operator \( \mathbf{H} \):

\[
T_{\text{obs}}(t) = \mathbf{HT}(t) + \varepsilon, \tag{10}
\]

where \( \mathbf{H} = [1, 0] \). The Gaussian measurement error \( \varepsilon \) with a mean of zero and a covariance of \( R \) is created and added to LST observations to account for the contribution of observations error to the posterior
covariance and avoid correlation among the ensembles (Burgers et al., 1998). Based on Li et al. (2014) and Lu et al. (2018), the standard deviation of 3 K is chosen for perturbing MODIS LST data.

Each ensemble member of the augmented state vector can be updated by (Evensen, 2009)

$$T_{a}^{j}(t) = T_{f}^{j}(t) + K_{j}^{T} T_{obs}(t) + \epsilon_{j} - H T_{f}^{j}(t),$$

where the different terms are defined by

$$K = P^{j} H^{T} (H P^{j} H^{T} + R)^{-1},$$

$$P^{j} H^{T} = \frac{1}{N_{c} - 1} \sum_{i=1}^{N_{c}} [X_{f}^{j}(t') - \bar{X}_{f}(t)] H (X_{f}^{j}(t) - \bar{X}_{f}(t))^{T},$$

$$H P^{j} H^{T} = \frac{1}{N_{c} - 1} \sum_{i=1}^{N_{c}} [H (X_{f}^{j}(t') - \bar{X}_{f}(t))] [H (X_{f}^{j}(t) - \bar{X}_{f}(t))].$$

The EnKS framework [i.e., equation (11)] updates the forecast (or the a priori) LST and EF (denoted by superscript f) and generate the so-called analysis (or the a posteriori) LST and EF (denoted by superscript a). $T_{f}^{j}$ and $T_{a}^{j}$ represent the jth ensemble member of forecast and analysis augmented state vector.

The MODIS LST observation at update time $t$ is used in the EnKS system to update the augmented state vector not only at that update time but also at previous times, $t'$ (Evensen, 2009; Lei et al., 2014; Li, Ryu, et al., 2013). $K$ is the Kalman gain matrix, and $R$ is the measurement error covariance. $P^{j}$ represents the error covariance matrix of the forecast model state (LST) and parameter (EF); the superscript $T$ denotes transpose, and $\bar{X}_{f}(t)$ represents the mean of forecasted augmented state variables at time $t$.

The state-augmented EnKS algorithm is run with a number of reasonable $C_{HN}$ values during each monthly period. The $C_{HN}$ value that yields the minimum difference between the estimated and observed LST, and the
corresponding EF estimates are chosen as optimum values, Herein, $C_{HN}$ is varied from 0.001 to 0.05 with an increment of 0.001 (Bateni et al., 2013; He et al., 2018; Xu, Bateni, et al., 2018; Xu, He, et al., 2019).

4. Study Domain and Data

The HRB is the second largest basin in northwest China (approximately 37.7°N, and 97.1 to 102.0°E). It has complex biological and environmental characteristics, and various landscapes including ice/frozen
soil/snow, forests, grasslands, oases, deserts, and lakes. Air temperature and precipitation show a north-south gradient over the HRB.

The “Heihe Watershed Allied Telemetry Experimental Research” (HiWATER) was established in 2012 (Li, Cheng, et al., 2013, 2018; Liu et al., 2011, 2018; Xu, Guo, et al., 2018). The half-hourly sensible heat flux was measured by the LAS instrument. The net radiation and ground heat flux were measured every 30 min by the four-component radiometer and ground heat flux plate, respectively. Finally, latent heat flux observations were obtained as the residual of the SEB equation (LE = RN – G – H). The HiWATER experiment data can be downloaded from the Heihe Data Archive (http://card.westgis.ac.cn/).

The hourly micrometeorological data including wind speed, air temperature and humidity, atmospheric pressure, and incoming shortwave and longwave radiation are generated by the Weather Research and Forecasting (WRF) model (5 × 5 km) and used as forcing data in the EnKS approach (Pan et al., 2012). Using the nearest method, the forcing data were resampled from 5 to 1 km to match the size of computational pixels (1 × 1 km) over the study domain. The MODIS LST data are available on a daily basis since the launch of the MODIS sensor in 1999 for Terra platform and in 2002 for Aqua platform. In this study, MODIS LST data from Aqua (MYD11A1) and Terra (MOD11A1) platforms with the spatial resolution of 1 × 1 km and 2-revisits during daytime are assimilated into the EnKS system to estimate turbulent heat fluxes. Because of cloud contamination, 65.33% of MODIS LSTs are available during the modeling period. They can be downloaded from the Level-1 and Atmosphere Archive and Distribution System Distributed Active Archive Center archive (https://ladsweb.nascom.nasa.gov/search/). The land cover data were provided by Zhong et al. (2014). LAI and Albedo are obtained from the Global Land Surface Satellite product (Xiao et al., 2014, 2016) (http://glass-product.bnu.edu.cn/). C and K are determined from the soil type and moisture (Chen, 2008; de Vries, 1963). The soil type is found from the HRB Digital Soil Mapping product (Song et al., 2016). The C and K vary from 1.52 × 10⁶ to 3.36 × 10⁶ J m⁻³ K⁻¹ s⁻¹ and from 0.37 to 1.21 J m⁻¹ K⁻¹ in the HRB. Soil moisture data are obtained from the Soil Moisture Active Passive (SMAP) L4 product with the spatial resolution of 9 × 9 km (https://search.earthdata.nasa.gov/).

Table 2

| Study sites     | Bias (K) | MAPE (%) | RMSE (K) | R² (–) |
|-----------------|----------|----------|----------|--------|
|                 | EnOL     | EnKS     | EnOL     | EnKS   | EnOL   | EnKS   | EnOL     | EnKS   |
| Arou            | −0.89    | −0.63    | 15.65    | 12.75  | 4.12   | 3.70   | 0.88     | 0.90   |
| Daman           | 0.52     | 0.01     | 10.23    | 5.6    | 3.25   | 2.62   | 0.89     | 0.92   |
| Sidaqiao        | 0.79     | 0.28     | 12.36    | 9.23   | 2.99   | 2.14   | 0.90     | 0.93   |
| Three-sites-average | 0.14   | −0.11    | 12.75    | 9.19   | 3.45   | 2.82   | 0.89     | 0.92   |
The HRB is discretized into square computational pixels. The size of each computational pixel is 1 km, resulting in 237,573 pixels. The modeling period covers 5 months from 1 May to 30 September 2015 (Day of Year, DOY 121–273), and the daily assimilation window ranges from 0900 to 1800 LT.

5. Results and Discussions

The retrieved monthly $C_{HN}$ maps from the EnKS approach and corresponding LAI values are shown in Figure 2. As indicated, $C_{HN}$ varies with the vegetation phenology, and its changes are consistent with those of LAI. The $C_{HN}$ values increase from May to July, and decrease from July to September due to the seasonal variations in the vegetation phenology (i.e., LAI). The $C_{HN}$ estimates are generally higher at the vegetated areas (i.e., grassland, cropland, and forest) in the south of HRB than the barren land in the north. In the south of HRB, $C_{HN}$ increases with the growth of the crop and reaches its maximum in July in which LAI is at its peak. There is a sharp drop in LAI in September due to the crop harvest. Analogously, $C_{HN}$ reduces in September. While there is almost no vegetation in the north of HRB and LAI is invariant, the corresponding retrieved $C_{HN}$ values vary in different monthly modeling periods. This is because $C_{HN}$ depends not only on the vegetation phenology (LAI) but also to a lesser extent on the wind speed, friction velocity, and solar elevation (Duynkerke, 1992; Smedman et al., 2007).

Soil moisture controls latent heat flux and is the key indicator of EF (Bateni, Entekhabi, & Castelli, 2013). Therefore, the variations of EF estimates should be consistent with those of soil moisture. Figure 3 shows the spatial patterns of the a priori [obtained from equation (6)] and the a posteriori [updated by the EnKS approach, equation (11)] EF estimates, standard deviation of EnKS EF retrievals, and SMAP soil moisture data over the HRB for DOYs 121, 153, 204, 257, and 273. As shown, the a posteriori EF estimates can capture variations in the SMAP soil moisture more robustly than the a priori EF estimates. For example, the estimated EF values are highest in the southeast of HRB on DOY 204 that soil moisture is at its peak. The higher EF estimates in the south of HRB are due to the high precipitation and dense vegetation. The EF estimates decrease toward the center and north of HRB because of low precipitation and sparse vegetation.

The standard deviations of EnKS EF estimates are shown in Figure 3 as well. The standard deviations of EF retrievals are generally lower in the north of HRB. In the north, evaporation is in its second phase (i.e., water limited); the drying rate is mainly controlled by land surface state variable (i.e., LST) rather than atmospheric factors. Therefore, the coupling between EF and LST is vigorous and the uncertainty of EF estimates is reduced. In contrast, because of the dense vegetation cover and heavy precipitation in the south of HRB, evaporation is in its first phases (i.e., energy limited) and is mainly controlled by the atmospheric state variables (i.e., air temperature and specific humidity). Therefore, the coupling between EF and LST becomes weak, resulting in a higher uncertainty in the EF estimates. By directly providing the standard deviation of the EF estimates, the EnKS approach shows the relative dependence of evaporation on atmospheric factors and surface properties in different environmental conditions. In addition to the different environmental conditions, the standard deviation of the EF estimates also caused by the noise in the forcing data as well as parametric model errors.

Figure 4 shows the hourly LST estimates from EnOL (green points) and EnKS (red points) versus the corresponding measurements at the Arou, Daman, and Sidaqiao sites. The predicted LSTs from the EnKS model are closer to the observations compared to those of the EnOL model. This is because the EnKS approach shifts the LST ensemble members toward the MODIS LST measurements in each update step, and reduces their difference. The statistical indices of hourly LST estimates from the EnKS and EnOL approaches at the three sites are summarized in Table 2. The root mean square errors (RMSEs) of EnOL LST estimates at the Arou, Daman, and Sidaqiao sites are 4.12, 3.25, and 2.99 K, respectively. The EnKS approach reduces the
abovementioned RMSEs by 10.19%, 19.38%, and 28.43%, respectively. It is worth mentioning that the EnKS LST estimates have the spatial resolution of 1 × 1 km (i.e., the size of computational pixels in the HRB), while LST data at the Arou, Daman, and Sidaqiao site are in situ point-scale measurements. The abovementioned inconsistency in the spatial resolutions, simplistic assumptions (e.g., constant daily EF, constant monthly CHN, constant soil thermal conductivity and heat capacity, etc.), and measurement errors cause misfits

Figure 6. Time series of daytime-averaged (0900–1800 LT) sensible (H), and latent (LE) heat fluxes estimates from EnKS and EnOL at the Arou, Daman, and Sidaqiao sites in the HRB. H and LE observations are shown by open circles. Blue or Pink bands represent the standard deviations of turbulent heat fluxes from EnOL and EnKS, respectively.
between the EnKS LST estimates and corresponding ground-based measurements at the Arou, Daman, and Sidaoqiao sites (Figure 4).

The normalized innovations of LST are shown in Figure 5. The normalized innovations of LST are defined as the ratio of the misfit between the MODIS LST observations and the corresponding estimates from the EnKS.
approach, \( T_{obs}(t) - HT(t) \), and the standard deviation term \( (HP^{-1/2} + R)^{-1/2} \) (Zupanski et al., 2005). The histogram of normalized innovations can be used as a means to evaluate the performance of EnKS, and realize whether the observation and model errors are selected properly (De Lannoy et al., 2010; Kumar et al., 2008). If the EnKS’s Gaussian assumption is met, the normalized innovations will have the standard normal distribution with the zero mean and unit variance, \( N(0, 1) \) (Crow, 2003; Margulis et al., 2002). The deviation of normalized innovations from the standard normal distribution implies that EnKS is suboptimal (Bateni & Entekhabi, 2012b).

The histogram of normalized innovations over the HRB is compared with \( N(0, 1) \) in Figure 5. The distribution of LST normalized innovations has a small negative bias over the HRB. This nonzero mean indicates that the input and output error statistics are not consistent. Hence, it suggests that the defined model and observation errors are not optimal. It might be also related to auto-correlated observation errors. The histogram of normalized innovations in Figure 5 also shows that the forecast error covariance \( (HP^{-1/2} + R)^{-1/2} \) is slightly overestimated. The overestimated forecast error covariance leads to a slightly higher peak for normalized innovations compared to the normal distribution (Reichle et al., 2002; Zupanski et al., 2005). This is most likely due to the inclusion of nonoptimal measurements and model error parameters. Overall, the normalized innovations of LST are close to the white noise, \( N(0, 1) \), over the HRB, implying that the spread of ensembles is adequate, the model estimates are reliable and almost near optimal (Hol et al., 2008; De Lannoy et al., 2010; Zupanski et al., 2005). At the same time, the small inconsistency between the normalized innovations histogram and the standard normal distribution indicates that the EnKS model is not truly optimal.

Table 3
Statistical indices of hourly \( H \) and \( LE \) estimates from the EnOL, EnKS, and VDA approaches at the three experimental sites

| Study sites | EnOLS \( H \) (W m\(^{-2}\)) | EnKL \( H \) (W m\(^{-2}\)) | VDA \( H \) (W m\(^{-2}\)) | EnOLS \( LE \) (W m\(^{-2}\)) | EnKL \( LE \) (W m\(^{-2}\)) | VDA \( LE \) (W m\(^{-2}\)) |
|---|---|---|---|---|---|---|
| Arou | Bias (W m\(^{-2}\)) | 26.31 | 9.01 | 14.07 | 49.13 | 36.58 | 36.15 |
| | MAPE (%) | 50.36 | 40.23 | 35.30 | 50.12 | 38.31 | 35.08 |
| | RMSE (W m\(^{-2}\)) | 55.51 | 40.71 | 30.19 | 160.33 | 119.8 | 115.85 |
| | \( R^2 \) (-) | 0.51 | 0.61 | 0.66 | 0.49 | 0.58 | 0.59 |
| Daman | Bias (W m\(^{-2}\)) | -30.98 | -21.99 | 21.89 | 68.52 | 51.61 | 46.69 |
| | MAPE (%) | 57.21 | 48.63 | 45.19 | 49.13 | 32.52 | 26.52 |
| | RMSE (W m\(^{-2}\)) | 59.61 | 46.72 | 39.42 | 136.55 | 95.55 | 84.43 |
| | \( R^2 \) (-) | 0.46 | 0.58 | 0.61 | 0.52 | 0.63 | 0.66 |
| Sidaoqiao | Bias (W m\(^{-2}\)) | 40.22 | 35.71 | 12.70 | 99.41 | 76.56 | 68.01 |
| | MAPE (%) | 49.53 | 38.55 | 33.84 | 55.68 | 40.12 | 31.44 |
| | RMSE (W m\(^{-2}\)) | 68.35 | 51.39 | 42.72 | 110.63 | 96.11 | 82.61 |
| | \( R^2 \) (-) | 0.51 | 0.63 | 0.67 | 0.51 | 0.63 | 0.68 |
| Three-sites-average | Bias (W m\(^{-2}\)) | 11.85 | 7.58 | 16.22 | 72.35 | 54.92 | 50.28 |
| | MAPE (%) | 52.37 | 42.47 | 38.11 | 51.64 | 36.98 | 31.01 |
| | RMSE (W m\(^{-2}\)) | 61.16 | 46.27 | 37.44 | 135.84 | 103.82 | 94.30 |
| | \( R^2 \) (-) | 0.49 | 0.60 | 0.65 | 0.51 | 0.61 | 0.64 |

Table 4
Standard deviations of hourly \( H \) and \( LE \) estimates from the EnOL and EnKS approaches at the three experimental sites

| Study sites | \( H \) (W m\(^{-2}\)) | \( LE \) (W m\(^{-2}\)) |
|---|---|---|
| | EnOLS | EnKL | EnOLS | EnKL |
| Arou | 95.70 | 57.57 | 130.58 | 96.65 |
| Daman | 88.39 | 47.74 | 115.94 | 87.87 |
| Sidaoqiao | 56.22 | 31.52 | 96.27 | 69.88 |
| Three-sites-average | 80.10 | 45.61 | 114.26 | 84.80 |
121–160) at the three sites. The LE estimates have a larger uncertainty (i.e., thicker blue bands) compared to the $H$ retrievals. This is because the uncertainty of $H$ is mainly dependent on the standard deviation of LST [see equation (3)]. While, LE is estimated by $H$ and EF [see equation (5)]. Thus, the uncertainty of LE depends on the uncertainties of EF and LST. More sources of errors increase the uncertainty of estimated LE values.

Figure 7 shows the scatter plot of hourly sensible and latent heat fluxes estimates from EnKS versus the corresponding measurements at the Arou, Daman, and Sidaqqiao sites. $H$ and LE estimates from EnKS agree

Figure 8. Maps of monthly mean $H$ (first row) and LE (second row) estimates.

Figure 9. The standard deviations of $H$ (first row) and LE (second row) estimates over the HRB.
fairly well with the observations and mainly fall around the 1:1 line. The misfit between the model estimates and observations might be due to the errors in the forcing data (e.g., air temperature, humidity, incoming solar radiation, and wind speed), misspecification of model parameters (e.g., albedo, emissivity, soil thermal conductivity, soil heat capacity, etc.), and simplistic model assumptions (i.e., daily constant EF and monthly constant CHN). Figure 7 also shows that LE estimates are more scattered compared to H retrievals. This is because of more sources of errors in the LE estimates as discussed above. LE observations are also obtained as the residual of the SEB equation (i.e., LE = RN − H − G), which can yield erroneous LE data (Liu et al., 2011).

The statistical indices of hourly turbulent heat fluxes estimates from the EnOL, EnKS, and VDA approaches at the three sites are summarized in Table 3. For sensible heat flux, the three-site-averaged Bias (RMSE) from the EnKS approach is 7.58 W m\(^{-2}\) (46.27 W m\(^{-2}\)), which is 36.03% (24.35%) lower than the Bias (RMSE) of 11.85 W m\(^{-2}\) (61.16 W m\(^{-2}\)) from EnOL. For latent heat flux, the three-site-averaged Bias (RMSE) is 54.92 (103.82) W m\(^{-2}\) for EnKS, and 72.35 (135.84) W m\(^{-2}\) for EnOL. The relatively low RMSE values indicate that the EnKS approach can provide fairly accurate turbulent heat fluxes estimates over various vegetative and hydrological conditions. Table 3 also compares the performance of EnKS with the VDA approach of Xu, He, et al., (2019) at the three sites. The three-site-averaged Bias (RMSEs) of sensible and latent heat fluxes estimates from VDA are 16.22 W m\(^{-2}\) (37.44 W m\(^{-2}\)) and 50.28 W m\(^{-2}\) (94.30 W m\(^{-2}\)), respectively. Because of the small suboptimality of EnKS, the results of VDA are slightly better than those of EnKS.
Figure 11. Box plots of the ensemble distribution for the area-averaged standard deviations of evaporative fraction (EF Std.), LST (LST Std.), sensible heat flux (H Std.), and latent heat flux (LE Std.) at five land cover types in the HRB. The corresponding box plots for LAI and soil moisture data are also shown in third row.

Figure 12. Scatterplot of 8-day ET retrievals from the EnKS approach and MODIS product versus in situ LAS measurements at the Arou, Daman, and Sidaoqiao sites for DOY 121–273 in 2015.
The standard deviation of $H$ and LE estimates from the EnOL and EnKS approaches at the three experimental sites are summarized in Table 4. As indicated, the standard deviation of estimated $H$ and LE values from the EnKS approach is lower than that of EnOL, implying that the EnKS approach is able to use information in the MODIS LST data in order to generate less uncertain $H$ and LE values. Table 4 also shows that the uncertainty of $H$ and LE estimates is larger at the Arou site (wet) compared to the Sidaoqiao site (semiarid).

Figure 8 shows the monthly mean $H$ and LE estimates from the EnKS method over the HRB during the growing season (May-September). LE increases from May to July and then decreases from July to September over the HRB, which is consistent with the temporal variations in precipitation and LAI. As can be seen, there is a sharp variation in the $H$ and LE estimates from north to south of HRB. The $H$ (LE) values in barren areas in the north of HRB are higher (lower) than those of grasslands, croplands, and forest in the south. LE is low in the north and center of HRB due to rare precipitation and sparse vegetation cover, but it increases significantly toward the south because of high precipitation and dense vegetation cover. High LE values in the center of HRB are associated with crop growth and irrigation in the cropland areas.

Figure 9 shows the standard deviations of $H$ and LE estimates from EnKS over the HRB during the growing season. As expected, the standard deviations of $H$ and LE estimates are lower in the north and center of the HRB due to low precipitation and sparse vegetation cover, but they increase toward the south of the HRB because of high precipitation and dense vegetation cover.

Figure 10 shows the monthly mean estimated LST from EnKS (first row) and observed LST from MODIS (second row) over the HRB during the growing season (May-September). Monthly mean MODIS LST maps are obtained by averaging MODIS LST observations. MODIS LST data are available twice daily only in clear-sky conditions (Bisht et al., 2005). The estimated LST values are consistent with the observations. Because of seasonal variation in solar radiation (not shown here), LST increases from May to July and decreases from July to September. The higher latent heat flux in the south keeps the surface relatively cool whereas the areas to the center and north with lower latent heat flux show higher temperature.

The third row in Figure 10 shows the relative difference of monthly mean estimated and observed LST. In the north and center of the HRB (dry and slightly vegetated conditions), the relative difference of estimated and observed LST is small, but it increases toward the south of the HRB (wet and densely vegetated conditions). The higher relative difference in the south is because of the heavy precipitation and dense vegetation that make the retrieval of EF from MODIS LST data more uncertain.

The HRB contains various vegetative and hydrological conditions, which make it possible to assess the uncertainty of EnKS estimates in different environmental conditions. Figure 11 shows the statistical boxplots of area-averaged standard deviations of evaporative fraction (EF Std.), LST (LST Std.), sensible heat flux ($H$ Std.), and latent heat flux (LE Std.) over five land cover types (cropland, forest, grassland, shrub land, and barren land) in the HRB. The statistical boxplots of LAI and soil moisture data over the same five land cover types are also shown in Figure 11. The lower and upper box edges are the 1st and 3rd quartiles in the sample distribution, respectively. The median position is marked within the box. Lines extending from the box ends indicate the 5th and 95th percentiles. As expected, the standard deviations (uncertainties) of EF, LST, $H$, and LE estimates for grassland (with higher soil moisture) and forest (with higher LAI) are larger than those of cropland, shrubland, and barren land. The lowest standard deviations can be found in the barren land (with the lowest soil moisture and LAI).

The MODIS global ET product provides ET data with the spatial resolution of 1 × 1 km every 8 days (Mu et al., 2007). Figure 12 compares the 8-day ET estimates from EnKS (red points) and MODIS (green

Table 5
| Study sites | Bias (mm) MODIS | MAPE (%) MODIS | RMSE (mm) MODIS | R^2 (−) MODIS | Bias (mm) EnKS | MAPE (%) EnKS | RMSE (mm) EnKS | R^2 (−) EnKS |
|------------|----------------|----------------|----------------|---------------|----------------|----------------|----------------|---------------|
| Arou       | 4.58           | −0.48          | 17.70          | 1.86          | 6.95           | 5.24           | 0.44           | 0.59          |
| Daman      | 9.11           | −2.52          | 37.43          | 10.37         | 10.28          | 5.72           | 0.70           | 0.73          |
| Sidaoqiao  | 23.70          | 9.99           | 88.07          | 37.12         | 24.63          | 10.59          | 0.28           | 0.77          |
| Three-sites-average | 12.46 | 2.33       | 47.73          | 16.45         | 13.95          | 7.18           | 0.47           | 0.70          |
6. Conclusion

Remotely sensed LST data from Moderate Resolution Imaging Spectroradiometer (MODIS) are assimilated into the heat diffusion equation within the ensemble Kalman smoother (EnKS) approach to estimate turbulent heat fluxes. The EnKS approach is tested over the HRB with contrasting vegetative and hydrological conditions. The key unknown parameters of the EnKS approach are the neutral heat transfer coefficient (\( C_{\text{HN}} \)) and evaporative fraction (EF). The results show that the spatiotemporal patterns of \( C_{\text{HN}} \) estimates agree well with those of vegetation phenology. The patterns in the estimated evaporative fraction (EF) maps resemble those of the SMAP soil moisture product.

The EnKS H and LE estimates are validated with the LAS observations at the Arrou (grassland), Daman (cropland), and Sidaqiao (shrub-forest) sites in the south, center, and north of HRB, respectively. The hourly and day-time averaged sensible and latent heat flux estimates from the EnKS model agree relatively well with observations at the three experimental sites. For the EnKS approach, the three-site-averaged RMSEs of hourly sensible and latent heat fluxes are 46.27 W m\(^{-2}\) and 103.82 W m\(^{-2}\), respectively, which are 24.35% and 23.57% lower than those of EnOl. These outcomes imply that the EnKS approach can take advantage of implicit information in the MODIS LST data to improve the EnOl H and LE estimates. The performance of EnKS is compared with the VDA approach. It is observed that the VDA approach performs slightly better than EnKS.

EnKS can directly generate the uncertainties of its estimates. The uncertainties of H and LE estimates over the wet and densely vegetated areas in the south of HRB are higher than the dry and slightly vegetated areas in the center and north. In dry and slightly vegetated areas, the drying rate is mainly controlled by the land surface state variable (i.e., LST). Therefore, the coupling between EF and LST becomes vigorous and the uncertainties of turbulent heat fluxes estimates are reduced. In contrast, in the wet and densely vegetated sites, the drying rate is mainly affected by atmospheric state variables (i.e., air temperature and specific humidity) and coupling between EF and LST becomes weak, resulting in higher H and LE uncertainties.

Acknowledgments
This work is supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant XDA20100101) and the National Natural Science Foundation of China (41671335), the U. S. Department of Agriculture-Natural Resources Conservation Service (Grant 69-3A75-17-54), the U. S. Geological Survey (Grant 2017Hi440B), and the Project Supported by State Key Laboratory of Earth Surface Processes and Resource Ecology (2017-KF-16). We would like to thank all the scientists, engineers, and students who participated in the HiWATER field campaigns. The ground-measured turbulent heat fluxes, meteorological variables, and land cover data are downloaded freely via the Heihe Data Center (http://card.westgis.ac.cn/). The MODIS LST data can be downloaded from the Level-1 and Atmosphere Archive and Distribution System Distributed Active Archive Center archive (https://ladsweb.nascom.nasa.gov/search/). LAI and Albedo data are available on the Beijing Normal University Data Center (http://glass-product.bnu.edu.cn/). SMAP products are available on the Earthdata Search Center (https://search.earthdata.nasa.gov/).

References

Abdolghaforian, A., Farhadi, L., Bateni, S. M., Margulis, S., & Xu, T. R. (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, 321–333. https://doi.org/10.1175/JHM-D-16-0097.1

Anderson, M. C., Norman, J. M., Diak, G. R., Kustas, W. P., & Meckikalski, J. R. (1997). A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. Remote Sensing of Environment, 60(2), 195–216. https://doi.org/10.1016/S0034-4257(96)00121-5

Baldocchi, D., Falge, E., Gu, L. H., & Olson, R. (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, 2415–2434. https://doi.org/10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2

Baldocchi, D., & Ma, S. (2011). Tonzi Ranch Tower FLUXNET L3 and L4 data, ftp://cdiac.ornl.gov/pub/ameri/data/Level4/Sites_ByName/Tonzi_Ranch/, Carbon Dioxide Inf. Anal. Cent., Oak Ridge Natl. Lab., Oak Ridge, Tenn.

Bastiaanssen, W. G. M., Menenti, M., Feddes, R. A., & Holtslag, A. A. M. (1998). A remote sensing surface energy balance algorithm for land (SEBAL): 1. Formulation. Journal of Hydrology, 212–213, 198–212. https://doi.org/10.1016/S0022-1694(98)00253-4

Bastiaanssen, W. G. M., Pelgrum, H., Wang, J., Ma, Y., Moreno, J. F., Roerink, G. J., & van der Wal, T. (1998). A remote sensing surface energy balance algorithm for land (SEBAL): 2. Validation. Journal of Hydrology, 212–213, 213–229. https://doi.org/10.1016/S0022-1694(98)00254-6

Bateni, S. M., & Entekhabi, D. (2012a). Relative efficiency of land surface energy balance components. Water Resources Research, 48, W04510. https://doi.org/10.1029/2011WR011357

Bateni, S. M., & Entekhabi, D. (2012b). 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. M., Entekhabi, D., & Castelli, F. (2013). 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. M., Entekhabi, D., & Jung, D. S. (2013). Variational assimilation of land surface temperature and the estimation of surface energy balance components. *Journal of Hydrology*, 481, 143-156. https://doi.org/10.1016/j.jhydrol.2012.12.039

Bateni, S. M., 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. M., & Liang, S. (2012). Estimating surface turbulent fluxes from satellitedata in the Yellow River Delta wetland. *Remote Sensing of Environment*, 126, 1719–1724. https://doi.org/10.1016/j.rse.2012.07.005

Burgers, G., van Leeuwen, P. J., & Evensen, G. (1998). Analysis scheme in the ensemble Kalman filter. *Monthly Weather Review*, 126, 1719–1724. https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2

Caparrini, F., Castelli, F., & Entekhabi, D. (2004a). Estimation of surface turbulent fluxes and surface parameters with remote sensing data. *Boundary-Layer Meteorology*, 107(3), 605–633.

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, 1612–1629. https://doi.org/10.3390/s7081612

Carrera, M., Belair, S., & Bilodeau, B. (2015). The Canadian Land Data Assimilation System (CaLDAS): Description and synthetic evaluation study. *Journal of Hydrometeorology*, 16, 1293–1314. https://doi.org/10.1175/JHM-D-14-0089.1

Castelli, F., Entekhabi, D., & Caporali, E. (1999). Estimation of surface heat transfer and an index of soil moisture using adjoint-state surface energy balance. *Water Resources Research*, 35, 3115–3126. https://doi.org/10.1029/1999WR900140

Chen, S. X. (2008). Thermal conductivity of sands. *Heat and Mass Transfer, 44*(10), 1241–1246. https://doi.org/10.1007/s00231-007-0357-1

Crow, W. T. (2003). Correcting land surface model predictions for the impact of temporarily sparse rainfall rate measurements using an ensemble Kalman filter and surface brightness temperature observations. *Journal of Hydrometeorology*, 4, 960–973. https://doi.org/10.1175/1525-7541(2003)004<0960:CSMPFL>2.0.CO;2

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*, 119(1), 105–120. https://doi.org/10.1007/s10546-004-2121-6

De Lannoy, G. J. M., Reichle, R., Houser, H., Arsenault, P. R., Verhoest, K. R., & Fasoulas, V. (2010). Satellite-scale snow water equivalent assimilation into a high resolution land surface model. *Journal of Hydrometeorology*, 11(2), 352–369. https://doi.org/10.1175/2009JHM1192.1

vries, D. A. (1963). Thermal properties of soils, In W. R. van Wijk (Ed.), *Physics of Plant Environment*, (pp. 210–235). New York: North-Holland.

Dickinson, R. E. (1987). Evapotranspiration in global climate models. *Advances in Space Research*, 7(11), 17–26. https://doi.org/10.1016/0273-1223(87)90290-0

Donne, S., & Entekhabi, D. (2005). An ensemble-based reanalysis approach to land data assimilation. *Water Resources Research*, 41, W02013. https://doi.org/10.1029/2004WR003449

Donne, S., & Entekhabi, D. (2006). Surface state and flux estimation using the ensemble Kalman smoother during the Southern Great Plains 1997 field experiment. *Water Resources Research*, 42, W01407. https://doi.org/10.1029/2005WR004334

Duyckwerke, P. G. (1992). The roughness length for heat and other vegetation parameters for a surface of short grass. *Journal of Applied Meteorology*, 31, 579–586. https://doi.org/10.1175/1520-0450(1992)031<0579:TURLFG>2.0.CO;2

Evensen, G. (2009). *Data assimilation: The ensemble Kalman filter*. Berlin: Springer.

Evett, S. R., Kustas, W. P., Gowda, P. H., Anderson, M. C., Prueger, J. H., & Howell, T. A. (2011). Overview of the Bushland Evapotranspiration and Agricultural Remote sensing EXperiment 2008 (BEAREX08): A field experiment evaluating methods for quantifying ET at multiple scales. *Advances in Water Resources*, 30, 4–19. https://doi.org/10.1016/j.advwatres.2012.03.010

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

He, X. L., Xu, H. Y., Bateni, S. M., Neale, M. U., Auligne, T., Liu, S. M., et al. (2018). Evaluation of the weak constraint data assimilation approach for estimating turbulent heat fluxes at six sites. *Remote Sensing*, 10(12), 1994. https://doi.org/10.3390/rs10121994

Hol, J., D., Schoon, T. B., & Gustafsson, F. (2008). A new algorithm for calibrating a combined camera and IMU sensor unit, in: 10th International Conference on Control, Automation, Robotics and Vision (ICARCV). 1857–1862

Hu, Z., & Islam, S. (1995). Prediction of ground temperature and soil moisture content by the force-restore method. *Water Resources Research*, 31(10), 2531–2539. https://doi.org/10.1029/95WR01650

Jia, L., Yang, C., Yan, Y., & Liu, G. (2009). Regional estimation of daily to annual regional evapotranspiration with MODIS data in the Yellow River Delta wetland. *Hydrology and Earth System Sciences*, 13, 1775–1787. https://doi.org/10.5194/hess-13-1775-2009

Jiang, L., & Islam, S. (2003). An intercomparison of regional latent heat flux estimation using remote sensing data. *International Journal of Remote Sensing*, 24(11), 2221–2236. https://doi.org/10.1080/01431160210154821

Kalnay, E., Li, H., Mijolth, T., Yang, C. S., & Ballabrega-Poy, J. (2007). 4-D-Var or ensemble Kalman filter? *Tellus Series A*, 59, 758–773. https://doi.org/10.1111/j.1600-0870.2007.00261.x

Kumar, S. V., Reichen, R. H., & Peters-Lidard, C. D. (2008). A land surface data assimilation framework using the landinformation system: Description and applications. *Advances in Water Resources*, 31, 1419–1432. https://doi.org/10.1016/j.advwatres.2008.01.013

Kustas, W. P., Aliferi, J. G., Anderson, M. C., Colaizzi, P. D., Prueger, J. H., Evett, S. R., et al. (2012). Evaluating the two-source energy balance model using local thermal and surface flux observations in a strongly advective irrigated agricultural area. *Advances in Water Resources*, 35, 120–133. https://doi.org/10.1016/j.advwatres.2012.07.005
Smedman, A., Högström, U., Sahlee, E., & Johansson, C. (2007). Critical re-evaluation of the bulk transfer coefficient for heat over the ocean during unstable and neutral conditions. *Quarterly Journal of the Royal Meteorological Society*, 133, 227–250. https://doi.org/10.1002/qj6

Song, L. S., Liu, S. M., Kustas, W. P., Nieto, H., Sun, L., Xu, Z. W., et al. (2018). Monitoring and validating spatially and temporally continuous daily evaporation and transpiration at river basin scale. *Remote Sensing of Environment*, 219, 72–88. https://doi.org/10.1016/j.rse.2018.10.002

Song, X., Liu, F., Zhang, G., Li, D., & Zhao, Y. (2016). Estimation of soil texture at a regional scale using local soil-landscape models. *Soil Science*, 181, 435–445. https://doi.org/10.1097/SS.0000000000000185

Su, Z. (2002). The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. *Hydrology and Earth System Sciences*, 6, 85–100. https://doi.org/10.5194/hess-6-85-2002

Sun, G., McNulty, S., Amatya, D., Liang, Y., & Kolka, R. (2005). Regional annual water yield from forest lands and its response to potential deforestation across the southeastern United States. *Journal of Hydrology*, 308, 258–268. https://doi.org/10.1016/j.jhydrol.2004.11.021

Sun, L., Liang, S., Yuan, W., & Chen, Z. (2013). Improving a Penman-Monteith evapotranspiration model by incorporating soil moisture control on soil evaporation in semiarid areas. *International Journal of Digital Earth*, 6(1), 134–156. https://doi.org/10.1080/17538947.2013.783635

Tang, R., & Li, Z. L. (2017). An improved constant evaporative fraction method for estimating daily evapotranspiration from remotely sensed instantaneous observations. *Geophysical Research Letters*, 55, 2319–2326. https://doi.org/10.1002/2017GL072621

Tang, R., Li, Z. L., & Tang, B. (2010). An application of the Ts-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, 540–551. https://doi.org/10.1016/j.rse.2009.10.012

Wang, K., Li, Z., & Cribb, M. (2006). Estimating of evaporative fraction from a combination of day and night land surface temperature and NDVI: A new method to determine the Priestley-Taylor parameter. *Remote Sensing of Environment*, 102, 293–305. https://doi.org/10.1016/j.rse.2006.02.007

Wang, K. C., & Dickinson, R. E. (2012). A review of global terrestrial evapotranspiration: observation, modeling, climatology, and climatic variability. *Reviews of Geophysics*, 50, RG2005. https://doi.org/10.1029/2011RG000373

Whitaker, J. S., Compo, G. P., & Thepaut, J. N. (2009). A comparison of variational and ensemble-based data assimilation systems for reanalysis of sparse observations. *Monthly Weather Review*, 137, 1991–1999. https://doi.org/10.1175/2008MWR2781.1

Xiao, Z., Liang, S., Wang, J., Xiang, Y., Zhao, X., & Song, J. (2016). Long-time-series global land surface satellite leaf area index product derived from MODIS and AVHRR surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 54(9), 5301–5318. https://doi.org/10.1109/TGRS.2016.2560522

Xiao, Z., Liang, S., Wang, J. D., Chen, P., Yin, X. J., Zhang, L. Q., & Song, J. (2014). Use of general regression neural networks for generating the GLASS leaf area index product from time-series MODIS surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 52(1), 209–223. https://doi.org/10.1109/TGRS.2013.2237780

Xu, T. R., Bateni, S. M., & Liang, S. (2015). Estimating turbulent heat fluxes with a weak-constraint data assimilation scheme: A case study (HIWATER-MUSOEXE). *IEEE Transactions on Geoscience and Remote Sensing*, 12, 68–72. https://doi.org/10.1109/LGRS.2014.2326180

Xu, T. R., Bateni, S. M., Liang, S., Entekhabi, S. 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*, 119, 10,780–10,798. https://doi.org/10.1002/2014JD021814

Xu, T. R., Bateni, S. M., Margulis, S. A., Song, L., Liu, S. M., & Xu, Z. W. (2016). Partitioning evapotranspiration into soil evaporation and canopy transpiration via a two-source variational data assimilation system. *Journal of Hydrometeorology*, 17, 2353–2379. https://doi.org/10.1175/JHM-D-15-0178.1

Xu, T. R., Bateni, S. M., Neale, C. M. U., Auligne, T., & Liu, S. M. (2018). Estimation of turbulent heat fluxes by assimilation of land surface temperature observations from GOES satellites into an ensemble Kalman smoother framework. *Journal of Geophysical Research: Atmospheres*, 123, 2409–2423. https://doi.org/10.1002/2017JD027732

Xu, T. R., Guo, Z. X., Liu, S. M., He, X. L., Meng, Y., Xu, Z. W., et al. (2018). Evaluating different machine learning methods for upscaling evapotranspiration from flux towers to the regional scale. *Journal of Geophysical Research: Atmospheres*, 123(16), 8674–8690. https://doi.org/10.1002/2018JD028447

Xu, T. R., Guo, Z. X., Xia, Y. L., Ferreira, V. G., Liu, S. M., Wang, K. C., et al. (2019). Evaluation of twelve evapotranspiration products from machine learning, remote sensing and land surface models over conterminous United States. *Journal of Hydrology*, 579, 124105. https://doi.org/10.1016/j.jhydrol.2019.124105

Xu, T. R., He, X. L., Bateni, S. M., Auligne, T., Liu, S. M., Xu, Z. W., et al. (2019). Mapping regional turbulent heat fluxes via variational assimilation of land surface temperature data from polar orbiting satellites. *Remote Sensing of Environment*, 211, 444–461. https://doi.org/10.1016/j.rse.2018.11.022

Xu, T. R., Liang, S., & Liu, S. (2011). Estimating turbulent fluxes through assimilation of geostationary operational environmental satellites data using ensemble Kalman filter. *Journal of Geophysical Research*, 116, D09109. https://doi.org/10.1029/2010JD015150

Xu, T. R., Liu, S. M., Liang, S., & Qin, J. (2011). Improving predictions of water and heat fluxes by assimilating MODIS land surface temperature products into common land model. *Journal of Hydrometeorology*, 12, 227–244. https://doi.org/10.1175/2010JHM1300.1

Xu, T. R., Liu, S. M., Xu, Z. W., Liang, S., & Xu, L. (2015). A dual-pass data assimilation scheme for estimating surface fluxes with FY3A-VIRR land surface temperature. *Science China Earth Sciences*, 58(2), 211–230. https://doi.org/10.1007/s11430-014-4964-7

Yao, Y., Liang, S., Cheng, J., Liu, S., Fisher, J., Zhang, X., et al. (2013). MODIS-driven estimation of terrestrial latent heat flux in China based on a modified Priestley-Taylor algorithm. *Agricultural and Forest Meteorology*, 171-172, 187–202. https://doi.org/10.1016/j.agrformet.2012.11.016

Zhang, X., Liang, S., Zhou, G., Wu, H., & Zhao, X. (2014). Generating Global Land Surface Satellite Incident shortwave radiation and photosynthetically active radiation products from multiple satellite data. *Remote Sensing of Environment*, 152, 318–332. https://doi.org/10.1016/j.rse.2014.07.003

Zhong, B., Ma, P., Nie, A., Yang, A., Yao, Y., Lü, W., et al. (2014). Land cover mapping using time series HF-1/CCD data. *Science China Earth Sciences*, 57(8), 1790–1799. https://doi.org/10.1007/s11430-014-4877-5

Zhu, W., Jia, S., & Lv, A. (2017). A universal Ts-VI triangle method for the continuous retrieval of evaporative fraction from MODIS products. *Journal of Geophysical Research: Atmospheres*, 122, 10,206–10,227. https://doi.org/10.1002/2017JD026964

Zupanski, M., Fletcher, S. J., Navon, I. M., Uzunoglu, B., Heikes, R. P., Randall, D. A., et al. (2005). Initiation of ensemble data assimilation. *Tellus*, 58(2), 159–170. https://doi.org/10.1111/j.1600-0870.2006.00173.x