SCIENCE CHINA
Earth Sciences

• RESEARCH PAPER •

February 2015 Vol.58 No.2: 211–230
doi: 10.1007/s11430-014-4964-7

A dual-pass data assimilation scheme for estimating surface fluxes with FY3A-VIRR land surface temperature

XU TongRen[1], LIU ShaoMin[1*], XU ZiWei[1], LIANG ShunLin[2,3] & XU Lu[1,4]

[1] *State Key Laboratory of Remote Sensing Science and School of Geography, Beijing Normal University, Beijing 100875, China;*
[2] *State Key Laboratory of Remote Sensing Science, and College of Global Change and Earth System Science, Beijing Normal University, Beijing 100875, China;*
[3] *Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA;*
[4] *Information Technology Department, National Library of China, Beijing 100081, China*

Received October 12, 2013; accepted June 26, 2014; published online October 27, 2014

In this work, a dual-pass data assimilation scheme is developed to improve predictions of surface flux. Pass 1 of the dual-pass data assimilation scheme optimizes the model vegetation parameters at the weekly temporal scale, and Pass 2 optimizes the soil moisture at the daily temporal scale. Based on ensemble Kalman filter (EnKF), the land surface temperature (LST) data derived from the new generation of Chinese meteorology satellite (FY3A-VIRR) are assimilated into common land model (CoLM) for the first time. Six sites, Daman, Guantao, Arou, BJ, Miyun and Jiyuan, are selected for the data assimilation experiments and include different climatological conditions. The results are compared with those from a dataset generated by a multi-scale surface flux observation system that includes an automatic weather station (AWS), eddy covariance (EC) and large aperture scintillometer (LAS). The results indicate that the dual-pass data assimilation scheme is able to reduce model uncertainties and improve predictions of surface flux with the assimilation of FY3A-VIRR LST data.

sensible heat flux, latent heat flux, ensemble Kalman filter, common land model

Citation: Xu T R, Liu S M, Xu Z W, et al. 2015. A dual-pass data assimilation scheme for estimating surface fluxes with FY3A-VIRR land surface temperature. Science China: Earth Sciences, 58: 211−230, doi: 10.1007/s11430-014-4964-7

Accurate monitoring of surface fluxes over land surfaces is necessary for global climate change research, ecological environment management, and agricultural and water resource planning. The magnitude of the surface flux is largely determined by vegetation parameters, soil moisture and other states of the land surface. The major methods used for estimating surface fluxes are field measurement, remote-sensing-based methods, and land surface modeling.

The variability in field measurements of surface flux has been documented over diurnal, seasonal, and inter-annual time scales (Liu et al., 2011). Certain permanent observation networks also have been constructed, such as the First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment (FIFE) (Kanemasu et al., 1992), FLUXNET (Baldocchi et al., 2001), WATER (Li et al., 2009) and HiWATER (Li et al., 2013). Direct observations from lysimeters, eddy covariance systems, Bowen ratio methods, and large aperture scintillometers, are necessary for increasing our understanding of water and energy balance at the land surface. However, these measurements are difficult to use for monitoring surface fluxes at the regional scale because they only produce either point or patch-scale data.

Spatially distributed estimates of surface flux can be obtained from remote-sensing-based methods (Bastiaanssen

*Corresponding author (email: smliu@bnu.edu.cn)

et al., 1998; Su, 2002; Liu et al., 2007). However, these methods are difficult to use for continuous monitoring of surface fluxes because the data derived from satellites are instantaneous and often contaminated by the presence of clouds. Land surface models have been rapidly developed to predict surface fluxes on continuous spatial and temporal scales with physical constraints (Dickinson et al., 1986; Sellers et al., 1996; Liou et al., 1999; Dai et al., 2003).

Regardless of their specific model structures, all land surface models require observational data to calibrate their parameters and adjust their states. New techniques such as data assimilation are needed to integrate either field or remotely sensed observations with the models, thus improving model accuracy by correcting the model state variables and parameters (Liang, 2004; Liang and Qin, 2008). Data assimilation has played an increasingly important role in improving predictions of such land surface state variables as the leaf area index (Xiao et al., 2011), soil temperature profile (Huang et al., 2008), soil moisture profile (Margulis et al., 2002), and other related variables, e.g., surface fluxes (Xu et al., 2011).

To model surface fluxes at the land surface, the acquisition of accurate land surface temperatures is highly important. The land surface temperature can control the water and energy balances of the land surface. For the same magnitude of solar radiation, lower land surface temperatures result from a wetland surface associated with relatively high latent heat flux and low sensible heat flux. Inaccurate surface flux predictions are highly correlated with errors in land surface temperature estimates. Therefore, assimilation of land surface temperatures has become an important method for improving surface flux predictions from land surface models. Huang et al. (2008) improved the ground heat flux predictions in a land surface model with the assimilation of MODIS land surface temperature products. The surface fluxes also can be obtained based on variational techniques and relatively simple models with the assimilation of field-measured land surface temperatures (Boni et al., 2001; Caparrini et al., 2004). Bateni et al. (2012) estimated soil and canopy surface fluxes with a dual-source data assimilation scheme. Xu et al. (2011) improved surface flux predictions by assimilating remotely sensed land surface temperatures. Moreover, the surface energy balance algorithm for land (SEBAL) or surface energy balance system (SEBS) can calculate instantaneous surface fluxes using meteorology data coupled with remotely sensed land surface parameters. These estimates can, in turn, be assimilated into land surface models. Schuurmans et al. (2003, 2011) showed that the assimilation of remotely sensed evapotranspiration derived from SEBAL is useful in hydrological model calibration. Pipunic et al. (2008) preformed experiments with the assimilation of land surface temperature, land surface soil moisture and surface fluxes derived from synthetic remote sensing data, and the greatest improvements in the land surface fluxes were found with the assimilation of remotely sensed surface fluxes. However, these researchers did not pay much attention to improving surface flux predictions using remotely sensed land surface temperatures.

In a land surface model, the sensible and latent heat fluxes are diagnostic variables, and they are affected by many factors (e.g., model states and parameters). Thus, new data assimilation strategies should be introduced to estimate surface fluxes by optimizing both model states and parameters. Data assimilation strategies that can simultaneously optimize the model states and parameters have been developed to estimate soil moisture profiles (Moradkhani et al., 2005; Qin et al., 2009; Tian et al., 2009). However, the model parameters and states may not change at the same temporal scales. Thus, dual-pass data assimilation scheme was developed to optimize model parameters and states at different temporal scales. Yang et al. (2007) developed a dual-pass data assimilation scheme to improve soil moisture simulations by optimizing both soil hydrological parameters and soil moisture. With the assimilation of microwave brightness temperatures, the soil hydrological parameters related to soil moisture are calibrated first at the monthly temporal scale and the soil moisture data are subsequently updated at the daily temporal scale. The dual pass scheme was constructed with the SCE_UA algorithm (Duan et al., 1993) and has been applied by others (Lu et al., 2012; Zhao et al., 2012). The current study constructed a dual-pass data assimilation scheme with an ensemble Kalman filter (EnKF) (Evensen, 1994) to improve the surface flux predictions with assimilation of remotely sensed land surface temperature. Using the dual-pass data assimilation scheme, the soil moisture profiles and model vegetation parameters can be optimized with separate loops at different temporal scales. Pass 1 of the dual-pass data assimilation scheme optimizes the model vegetation parameters related to surface fluxes at a long temporal scale, similar to a parameter calibration procedure. Pass 2 of the scheme optimizes the soil moisture at a short temporal scale.

As a new-generation Chinese polar orbiting meteorological satellite, the FY-3 series consists of two experimental (FY-3A/B) and at least four operational satellites (FY-3C/D/E/F). The FY-3 series is expected to have a service life that lasts until 2020. The main tasks of FY-3 satellites include providing meteorology parameters for numerical weather prediction, monitoring of global natural disasters, and monitoring global distributions of ice, snow and ozone. The FY-3A was launched in the year 2008. Visible and infrared radiometer (VIRR) is one of the 11 sensors mounted on the FY-3A satellite and has a nominal spatial resolution of approximately 1 km×1 km at the nadir. The FY-3A land surface temperature data will become the new data source for global water and energy balance research. This study carried out the first experiment to assimilate remotely sensed land surface temperatures (LST) derived from FY3A-VIRR into a land surface model.

Six Chinese flux observation sites with different land cover types (alpine meadow, grassland, cropland, and orchard) were selected to conduct the data assimilation experiments. The six sites are located in three major nature zones (eastern monsoon area, western arid and semi-arid area, and Qinghai-Tibet alpine zone) which is representative in China. Moreover, the results from the experiments are compared with those from data generated by a multi-scale surface-flux observation system that includes an eddy covariance (EC) system, a large aperture scintillometer (LAS) and an automatic weather station (AWS).

1 Methodology

The dual-pass data assimilation technique employed in this study optimizes the model states and parameters independently. The data assimilation scheme is illustrated in Figure 1. The dual-pass data assimilation scheme includes the model operator (CoLM), the data assimilation algorithm (EnKF), and forcing and ancillary data. The two passes of the scheme assimilate the remotely sensed land surface temperatures (LST) derived from the FY3A satellite. With the assimilation of the FY3A-VIRR LST data, Pass 1 is used to optimize the model parameters at the weekly temporal scale, and Pass 2 is used to optimize the soil moisture at the daily temporal scale. For Pass 1, the default model parameters are used in CoLM for the first week. At the end of the first week, the model parameters are updated by minimizing the difference between the predicted LST and FY3A-VIRR LST. In the second week, the optimal model parameters estimated from the previous week are used throughout the entire week, and the parameters are updated at the end of the week and used in the following week. In

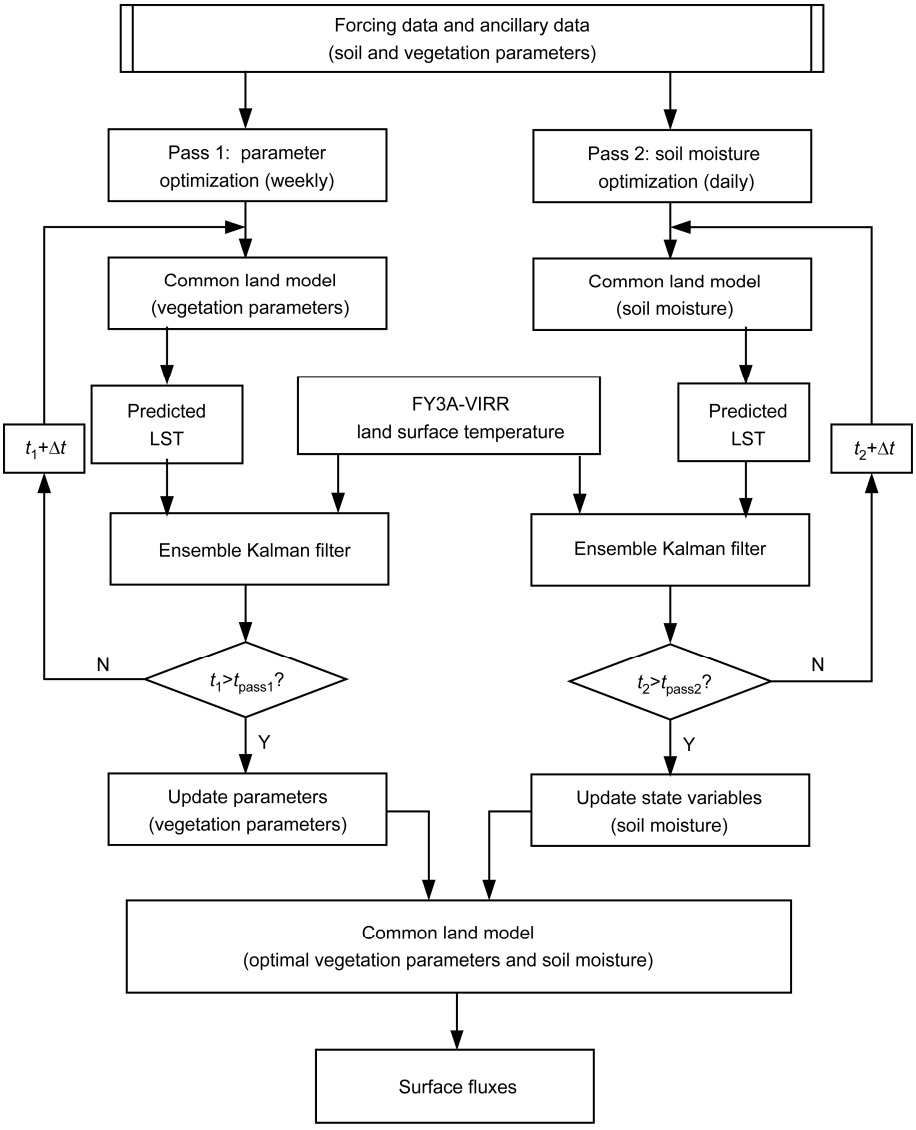

Figure 1 Flowchart of dual-pass data assimilation scheme.

Pass 2, together with the optimal model parameter estimates (default model parameters for the first week), the soil moisture is updated every day (once the FY3A-VIRR LST data are available) by minimizing the difference between the predicted LST and the FY3A-VIRR LST. This section will introduce the model operator and data assimilation algorithm used in the dual-pass data assimilation scheme.

1.1 Common land model

The common land model (CoLM) is a state-of-the-art model developed by many groups and validated with extensive field data sites. The CoLM can combine many processes, i.e., physical, hydrological, and biological, to simulate land state variables (e.g., soil temperature and soil moisture) and diagnostic variables (e.g., surface radiation and surface fluxes) (Dai et al., 2003). The CoLM contains one vegetation layer, ten unevenly spaced vertical soil layers and up to five snow layers (the number of snow layers can be changed with the total snow depth). A two-big-leaf model was built in 2004 for leaf temperature, photosynthesis, and stomatal conductance and divides the vegetation canopy into sunlit and shaded leaves (Dai et al., 2004).

In CoLM, the sensible and latent heat fluxes can be calculated using the land-atmosphere energy balance equation. For a bare ground land surface, the equation is as follows:

$$G = R_{n,g} - H_g - LE_g , \tag{1}$$

where G is soil heat flux at the soil surface (W m^{-2}), $R_{n,g}$ is the net radiation absorbed by the ground surface (W m^{-2}), and H_g and LE_g are the sensible and latent heat fluxes at the soil surface, respectively (W m^{-2}). The soil heat transfer is assumed to obey the following equation:

$$c\frac{\partial T}{\partial t} = -\frac{\partial F}{\partial z} , \tag{2}$$

where c is the volumetric heat capacity, T is the soil temperature, t denotes time, F is the soil heat flux, and z is the vertical coordinate. For vegetated land surfaces, the leaf temperatures are determined by the energy balance equation for the sunlit and shaded fractions of the canopy, and the sunlit and shaded leaf temperatures can be calculated as follows:

$$C_c \frac{\partial T_{csun}}{\partial t} = 0 = R_{n,csun} - H_{csun} - LE_{csun} , \tag{3}$$

$$C_c \frac{\partial T_{csha}}{\partial t} = 0 = R_{n,csha} - H_{csha} - LE_{csha} , \tag{4}$$

where C_c is the canopy heat capacity (J m^{-2} K^{-1}) and is assumed to be negligible, T_{csun} and T_{csha} are the sunlit and shade leaf temperatures (K), respectively, t denotes time, $R_{n,csun}$ and $R_{n,csha}$ are the net radiation amounts absorbed by the sunlit and shaded canopies (W m^{-2}), respectively, H_{csun} and H_{csha} are the sensible heat fluxes from the sunlit and shaded canopies (W m^{-2}), respectively, and LE_{csun} and LE_{csha} are the latent heat fluxes from sunlit and shaded canopies (W m^{-2}), respectively.

The total surface fluxes from the land surface to the atmosphere can be calculated as follows:

$$H = H_g + H_{csun} + H_{csha} , \tag{5}$$

$$LE = LE_g + LE_{csun} + LE_{csha} , \tag{6}$$

where H and LE are the sensible and latent heat fluxes from the land surface to the atmosphere (W m^{-2}), respectively. The details of the ground, sunlit and shaded flux calculations can be found in Dai et al. (2003, 2004).

The land surface temperature is a prognostic variable in CoLM and can be calculated using the following equation:

$$T_s = \left(F_u / \varepsilon\sigma\right)^{0.25} , \tag{7}$$

where T_s is the simulated land surface temperature (K), F_u is the surface upward long-wave radiation emitted from soil and canopy (W m^{-2}), σ is the Stefan-Boltzmann constant (5.67×10^{-8} W m^{-2} K^{-4}), and ε is the broadband emissivity. This equation can be considered as the observational operator of the dual-pass data assimilation scheme.

In CoLM, soil moisture is important for dividing the net radiation into the ground heat, sensible, and latent heat fluxes. A large soil moisture value will lead to a large latent heat flux and small sensible heat flux and surface temperature values (and vice versa). The equation for liquid soil water and ice can be expressed as (Dai et al., 2003):

$$\frac{\partial W_{liq}}{\partial t} = -\frac{\partial q}{\partial z} - E_t + M_{il} , \tag{8}$$

where W_{liq} is the mass of the soil water (kg), t is time, q is the water flow within the soil (kg m^{-2} s^{-1}), E_t is evapotranspiration (kg m^{-2} s^{-1}), and M_{il} is the mass rate of melting (positive) or freezing (negative) of soil ice.

Ten unevenly vertical soil layers are set up in CoLM, and the depth of soil layer jth at the node depth is defined as z_j=0.025×{exp[0.5(j–0.5)–1]}.

The input data of CoLM include ancillary data and forcing data. A variety of ancillary data is required for modeling and includes land cover type, soil and vegetation parameters. The Land Use and Land Cover (LULC) in CoLM can be based on the International Geosphere-Biosphere Programme (IGBP) or the United States Geological Survey (USGS) classification system. Shangguan et al. (2012) developed a conterminous China soil texture (i.e., sand, silt and clay content) dataset with 1 km×1 km resolution for land modeling use. The leaf area index (LAI) can characterize vegetation growth conditions and is important for surface flux estimates; the LAI was derived from the MODIS LAI

products and directly incorporated into CoLM in this study. The other vegetation parameters, such as surface roughness length, are optimized using the dual-pass data assimilation scheme. In this study, the forcing data were taken from a continuous series of meteorological data measured by automatic weather stations with a temporal resolution of a half hour. The data include wind speed, air temperature, relative humidity, air pressure, precipitation, incoming shortwave radiation, and incoming long-wave radiation. The model is run using the same time steps as the measured meteorological data, and the field measurements are used for model state variable initialization, i.e., soil moisture and soil temperature. The spatial scale of CoLM depends on the spatial scale of the forcing data and land surface parameters. In this study, the ground-measured meteorology data are used as the forcing data with a representative area larger than 1 km. The land surface parameters occur primarily at the 1 km spatial scale, and thus, we assume that the spatial scale of CoLM is approximately 1 km.

1.2 Assimilation method

In this study, the EnKF algorithm is employed as the data assimilation method. Due to easy implementation, the EnKF algorithm has been broadly applied in the construction of data assimilation schemes. The EnKF is based on ensemble generations in which the approximation of the predicted state error covariance matrix is calculated by spreading an ensemble of model states using the states from the previous time step. The key concept in the performance of the EnKF is the generation of the ensemble of model states or parameters and observations at each update time by introducing noise drawn from a distribution with a zero mean and covariance equal to the model states and observation error covariance matrix.

Consider $X=[w_1, w_2, \ldots, w_{10}]^{\mathrm{T}}$ as the state vector of the data assimilation scheme, where w_1, w_2, and w_{10} denote the soil moisture of the first layer, second layer and tenth layer, respectively. The first time the algorithm is run, the first guest value X_0 is used to create a series of state vector ensemble members by adding random noises:

$$X_{i,0} = X_0 + \eta_i, \quad \eta_i \sim N(0, B_0), \tag{9}$$

where $X_{i,0}$ is the state variable of the ith member at the beginning time, and η_i is the background error vector that conforms to a Gaussian distribution with a zero mean and covariance matrix of B_0.

In the forecast step, the each soil moisture ensemble member is predicted according to:

$$X^f_{i,t+1} = M(X^f_{i,t}, \alpha_{t+1}, \beta_{t+1}) + \mu_i, \quad \mu_i \sim N(0, Q), \tag{10}$$

where $X^f_{i,t}$, $X^f_{i,t+1}$ represent the forecasted state variables of the ith member at times t and t+1, the superscript "f" represents the forecasted state variables, $M(\cdot)$ is the model operator (CoLM in this study), α_{t+1} and β_{t+1} indicate forcing data and model parameters at time t+1, respectively, and μ_i is the model error vector, which conforms to Gaussian distribution with a zero mean and covariance matrix Q (Q: The model error matrix).

When FY3A-VIRR LST is available, the observation operator will predict LST as given by the following equation:

$$T^f_{si} = H(X^f_i, \alpha, \beta) + v_i, \quad v_i \sim N(0, R), \tag{11}$$

where T^f_{si} is the model-predicted land surface temperature of the ith member, $H(\cdot)$ is the observation operator that relates the model state variables to observations, and v_i is the observation error that conforms to Gaussian distributions with a zero mean and covariance matrix R (R: The observation error). The state variable of each member is updated as follows:

$$X^a_i = X^f_i + K(T^o_s - T^f_{si}), \tag{12}$$

$$K = P^f H^T (HP^f H^T + R)^{-1}, \tag{13}$$

$$P^f = \frac{1}{N-1}\sum_{i=1}^{N}\left(X^f_i - \bar{X}^f\right)\left(X^f_i - \bar{X}^f\right)^{\mathrm{T}}, \tag{14}$$

$$P^f H^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}\left[X^f_i - \bar{X}^f\right]\left[H\left(X^f_i\right) - H\left(\bar{X}^f\right)\right]^{\mathrm{T}}, \text{ and} \tag{15}$$

$$HP^f H^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}\left[H\left(X^f_i\right) - H\left(\bar{X}^f\right)\right]\left[H\left(X^f_i\right) - H\left(\bar{X}^f\right)\right]^{\mathrm{T}}, \tag{16}$$

where X^a_i represents the analyzed state variables of the ith member, K is the Kalman gain matrix, T^o_s is the FY3A-VIRR LST observations, P^f is the forecasted background covariance matrix, N is the number of ensembles; and $\bar{X}^f_i$ is the mean of the forecasted state vector ensemble members.

As a type of quantitative global sensitivity analysis method, the Extended Fourier Amplitude Sensitivity Test (EFAST) (Saltelli and Bolado, 1999) is used to rank the model vegetation parameters and identify factors that could be considered optimized in CoLM. According to the sensitivity test, the five most important vegetation parameters are the surface roughness length (z0m), quantum efficiency at 25°C (effcon), maximum rate of carboxylation at 25°C (vmax25), conductance-photosynthesis slope parameter (gradm), and conductance-photosynthesis intercept (binter). The five most important parameters are selected and optimized using EnKF algorithm in this study. Using the dual-pass assimilation technique, the EnKF updates the model states and parameters separately. The soil moisture and model parameters are updated separately using eqs. (12)–(16). In Pass 1, the state variables X in eq. (9) include the

selected vegetation parameters, and X includes the ten-layer model soil moisture in Pass 2.

To apply the EnKF technique, generation of ensemble members of the state vector is necessary, and noise should be added to the forcing data, model parameters, and model state variables. Because the in-situ meteorology data are used for forcing data, they are not perturbed by addition of noise. In this study, the variable ensemble members can be obtained by adding a series of Gaussian-distributed noises to the model parameters of Pass 1 and the model soil moisture of Pass 2. The sizes of the ten-layer soil moisture noises are 0.036, 0.033, 0.033, 0.034, 0.032, 0.030, 0.032, 0.024, 0.024 and 0.024 $m^3 m^{-3}$ (from the top to the bottom layer), according to Xu et al. (2011). The sizes of the model parameter noises are set to 10% of the range of the value.

Because Pass 2 of the dual-pass data assimilation scheme updates the soil moisture when the FY3A-VIRR LST data are available, eqs. (12)–(16) can be used directly. For Pass 1, the analyzed model parameters are calculated when the FY3A-VIRR LST data are available but are not updated. At the end of the week, the calculated model parameter analysis are averaged and updated as follows:

$$\overline{\beta_i^a} = \frac{1}{D}\sum_{k=1}^{D}(\beta_{i,t}^a), \tag{17}$$

where $\overline{\beta_i^a}$ represents the averaged model parameter analysis of the ith member, $\beta_{i,t}^a$ is the model parameter analysis of the ith member at the time t in one week, and D is the number of FY3A-VIRR LST observations in one week. The averaged model parameter analysis is transferred to Pass 2 of the data assimilation scheme.

2 Experimental data

2.1 Site description

Data from six sites located in three major nature zones (western arid and semi-arid area, eastern monsoon area, and Qinghai-Tibet alpine meadow) of People's Republic of China (PRC) are used for the data assimilation experiments. The data assimilation experiments are conducted during the vegetation growth seasons to capture the partition of available energy in the six sites with different climatological conditions. The Daman site is a superstation of the Heihe Watershed Allied Telemetry Experimental Research (HiWATER) (Li et al., 2013). The experimental time period for the Daman site is DOY 148–273 in 2012 and DOY 121–273 in 2010 for the other five sites. The Daman is a cropland site covered with seed corn located in Gansu province, the Guantao is cropland site coved with corn and cotton located in Hebei province, the Arou is an alpine meadow site covered with dense grass located in Qinghai province, the BJ site is a grassland site covered with sparsely distributed short grass during the rainy season and is located in the Tibet Plateau, the Miyun site is located in the northern mountain area of the Beijing city and has a surface primarily covered with orchard and maize, and Jiyuan is a forest site located in the mountain area of Henan province. Table 1 summarizes the surface conditions of the six experimental sites. A multi-scale surface flux observation system consisting of an eddy covariance (EC) and a large aperture scintillometer (LAS) is set up at each site to acquire the surface fluxes at two spatial scales simultaneously. Automatic weather station (AWS) is also equipped in the six flux observation sites, and the AWS can provide the necessary forcing data and auxiliary data model. The EC-measured sensible and latent heat fluxes and the LAS-measured sensible heat flux are used to validate the dual-pass data assimilation scheme. In this type of observation system, the source area of LAS system is generally larger than one FY3A-VIRR pixel (1 km resolution). The EC and AWS are located next to the center of the LAS optical path.

All of the observed data from the AWS, EC, and LAS were collected in half-hour time steps. The processing method for these data can be found in Liu et al. (2011, 2013). The LAS system consists of a transmitter and a receiver installed on a pair of towers approximately 500–5000 m apart that can measure the average value of sensible heat flux along the optical path of the instrument. Generally, the source area of the LAS measurements is generally larger than EC measurements, and covers more than one FY3A-VIRR pixel. Thus, source area of LAS measurements must be calculated using a footprint model. The LAS footprint is

Table 1 Summary of the surface characteristics of the six experimental sites

Site	Latitude	Longitude	Location	Nature zones	Land cover	Elevation (m)
Daman	38.37°N	100.85°E	Gansu	Western arid and semi-arid	Crop land	1560
Guantao	36.52°N	115.13°E	Hebei	Eastern monsoon	Crop land	30
Arou	38.04°N	100.91°E	Qinghai	Qinghai-Tibet alpine	Alpine meadow	3033
BJ	31.37°N	91.89°E	Tibet	Qinghai-Tibet alpine	Grass land	4520
Miyun	40.63°N	117.32°E	Beijing	Eastern monsoon	Orchard	350
Jiyuan	35.03°N	112.47°E	Henan	Eastern monsoon	Forest	400

calculated according to Liu et al. (2011, 2013):

$$f_{LAS}(x, y, z_{eff}) = \int_{x_2}^{x_1} W(x') f(x' - x, y' - y, z_{eff}) \mathrm{d}x', \quad (18)$$

where $W(x')$ is the path-weighting function of the LAS; x_1, x_2 are the locations of the LAS transmitter and receiver, respectively; x', y' are the points along the optical length of the LAS; x, y are the coordinates upwind of each point (x', y'); and z_{eff} is the effective measurement height of the LAS. The monthly LAS footprints are used in this study and are determined by averaging each half-hour footprint when the sensible heat fluxes are larger than zero; footprint values from the time period of 22∶00 to 6∶00 are also excluded.

2.2 FY3A-VIRR land surface temperature (LST)

The land surface temperature (LST) can be retrieved from the visible and infrared radiometer (VIRR) sensor mounted on the FY3A satellite, and the products can be downloaded from the website at http://satellite.cma.gov.cn/PortalSite/Ord/Satellite.aspx. The VIRR is a multi-channel instrument designed for comprehensive detection of the earth environment. The VIRR instrument contains 10 channels ranging from 0.58 to 12.5 μm. The fourth (10.3–11.3 μm) and fifth (11.5–12.5 μm) channels are infrared channels with little water vapor absorption and have a nominal spatial resolution of 1 km×1 km at the nadir. With these two infrared channels, the land surface temperature is obtained based on a local split window method (Becker and Li, 1990) as follows:

$$T_s = A_0 + P(T_4 + T_5)/2 + M(T_4 - T_5)/2, \quad (19)$$

where T_s is the FY3A-VIRR LST (K), A_0 is a constant, T_4 and T_5 are the brightness temperatures of the fourth and fifth channels (K), respectively, and P and M are the functions of the land surface emissivity, which can be regressed from the simulated data. Yang et al. (2006) recalculated the parameters in this algorithm based on the spectral response function of the FY3A-VIRR sensor. The FY3A-VIRR LST products can provide the LST and emissivity of each pixel, and the data are stored in a hierarchical data format (HDF), which is a sinusoidal projection with a spatial resolution of 1 km (Yang and Dong, 2011).

The LST derived from the FY3A-VIRR must be validated using *in-situ* measurements. The ground-measured surface temperatures can be calculated using the upward long- wave radiation at the land surfaces, the land surface emissivity, and the downward long-wave radiation according to thermal radiative transfer theory (Liang, 2004).

$$T_s = \left\{ \left[F_u - (1-\varepsilon) F_d \right] / \varepsilon\sigma \right\}^{0.25}, \quad (20)$$

where T_s is the land surface temperature (K), F_u is the surface upward long-wave radiation (W m^{-2}), F_d is the surface downward long-wave radiation (W m^{-2}), σ is the Stefan-Boltzmann constant (5.67×10^{-8} W m^{-2} K^{-4}), and ε is the broadband emissivity, which is 0.987 for grasslands and croplands and 0.993 for orchard and forest, according to Wang et al. (2008).

The FY3A-VIRR LST data are compared with the ground measurements (Figure 2). As shown in Figure 2, the FY3A-VIRR LST and ground measurements follow the same trend; the correlation coefficients (R) are 0.90, 0.81, 0.62, 0.71, 0.85 and 0.85 at the Daman, Guantao, Arou, BJ, Miyun and Jiyuan sites, respectively. The FY3A-VIRR LST data are larger than the field measurements at the Daman and Arou sites and smaller at the other four sites. The root mean square error (RMSE) values between the FY3A-VIRR LST data and the field measurements are 2.7, 3.5, 3.8, 5.0, 3.7 and 3.0 K at the Daman, Guantao, Arou, BJ, Miyun and Jiyuan sites, respectively, and are used as the observational errors in the dual-pass data assimilation scheme. The deviations between the FY3A-VIRR LST data and the ground measurements are determined by many factors. The terrain effect can affect the accuracy of the FY3A-VIRR LST retrievals. The mismatch of the spatial and temporal scales between the FY3A-VIRR and field-measured LST can also cause these biases. Remote sensing data such as the FY3A-VIRR LST are instantaneous values, whereas the ground measurement is a mean value over approximately 30 min. The spatial resolution of the FY3A-VIRR is approximately 1 km×1 km, whereas the footprint of the ground measurements is approximately dozens of square meters as determined by mounting level of the radiometer.

3 Results and discussions

In this section, the FY3A-VIRR LSTs are assimilated into CoLM with the developed dual-pass data assimilation scheme, and the results are compared with the multi-scale surface flux observations at the six experimental sites. The data assimilation scheme is conducted during the vegetation growth season. The statistics of the model biases (BIAS= model predictions–observations), the root mean square error (RMSE) and the correlation coefficients (R) are selected to assess the performance of the data assimilation scheme. In Section 3.1, the results are validated using the EC-derived sensible and latent heat fluxes. In Section 3.2, the results are compared with the LAS-derived sensible heat flux. In Section 3.3, the retrieved soil moisture and model parameters are shown, and the error sources in surface flux predictions are analyzed.

3.1 Comparisons of the simulation and assimilation results with EC data

In this section, the ground-measured land surface temperature and the EC-derived sensible and latent heat fluxes are

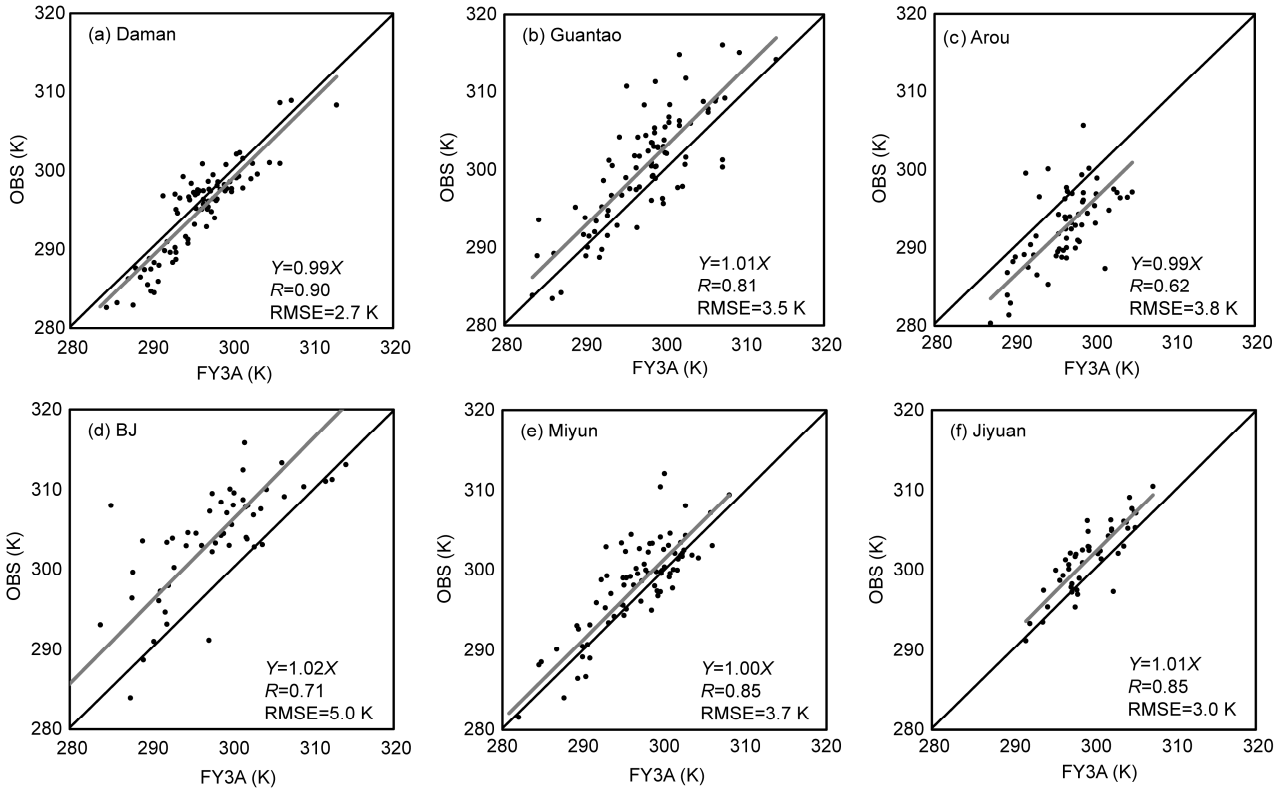

Figure 2 Comparison between the FY3A-VIRR LST data and the ground-measured surface temperatures (OBS) at the six experimental sites.

used to validate the results, which are shown in Figures 3, 4 and Table 2.

Generally, the diurnal variation trends of the land surface temperature and surface flux can be predicted correctly using the CoLM (Figure 3). However, the land surface temperatures and the sensible heat flux are overestimated and the latent heat flux is underestimated with this model. The obvious errors in the surface temperature and surface flux modeling are corrected with the assimilation of the FY3A-VIRR LST data. The curves portrayed by the dual-pass data assimilation scheme are generally closer to the EC measurements than the model predictions. From Figure 3, the model-simulated latent heat fluxes reach their peak relatively early (at approximately 10∶00 a.m.) and are at times nearly zero during the daytime, especially in the afternoon at the BJ site. With the assimilation of the FY3A-VIRR LST data, the diurnal variations of the latent heat flux agree well with observations. From eqs. (5)–(6), the surface fluxes can be produced by the vegetation canopy and ground. Using evapotranspiration as an example, the errors in the transpiration from the canopy and the errors in the evaporation from the ground can be caused by the vegetation parameters and soil moisture. In this study, the MODIS LAI products are incorporated directly into the model, and the vegetation parameters are optimized at the weekly temporal scale with Pass 1 of the dual-pass data assimilation scheme. At the daily temporal scale, the soil moisture is optimized with Pass 2 of the dual-pass data assimilation scheme. With the optimized vegetation parameters and soil moisture, the CoLM can predict the latent heat flux more accurately.

The daily averaged evaporation fraction (EF, ratio between latent heat flux and available energy) can be used to describe the splitting of available energy into sensible and latent heat fluxes. In CoLM, the EF is important for energy balance and describes the magnitude of available energy for evapotranspiration. Thus, the simulation and assimilation results are also assessed with the EC-derived EF. In this study, the daily averaged EF is calculated using *EF=latent heat flux/(latent heat flux+sensible heat flux)*. Figure 4 shows comparisons of the simulation and assimilation results with the EC-derived EF at the six experimental sites. From Figure 4, the CoLM usually underestimates the EF, which means that the model splits relatively less available energy for evapotranspiration. Precipitation data are also shown in Figure 4; the model-simulated EF is increased quickly after precipitation occurs, which indicates that soil moisture is important for predictions of surface fluxes. However, the EF decreases after a short period of precipitation, and the EF is increased and is closer to the observations with the assimilation of FY3A-VIRR LST data.

Table 2 summarizes the BIAS, RMSE and R-values of the model simulation and assimilation results compared with the EC-derived surface fluxes at the six sites. According to this table, the developed dual-pass data assimilation scheme can improve the predictions of surface flux and surface temperature. For the sensible and latent heat fluxes,

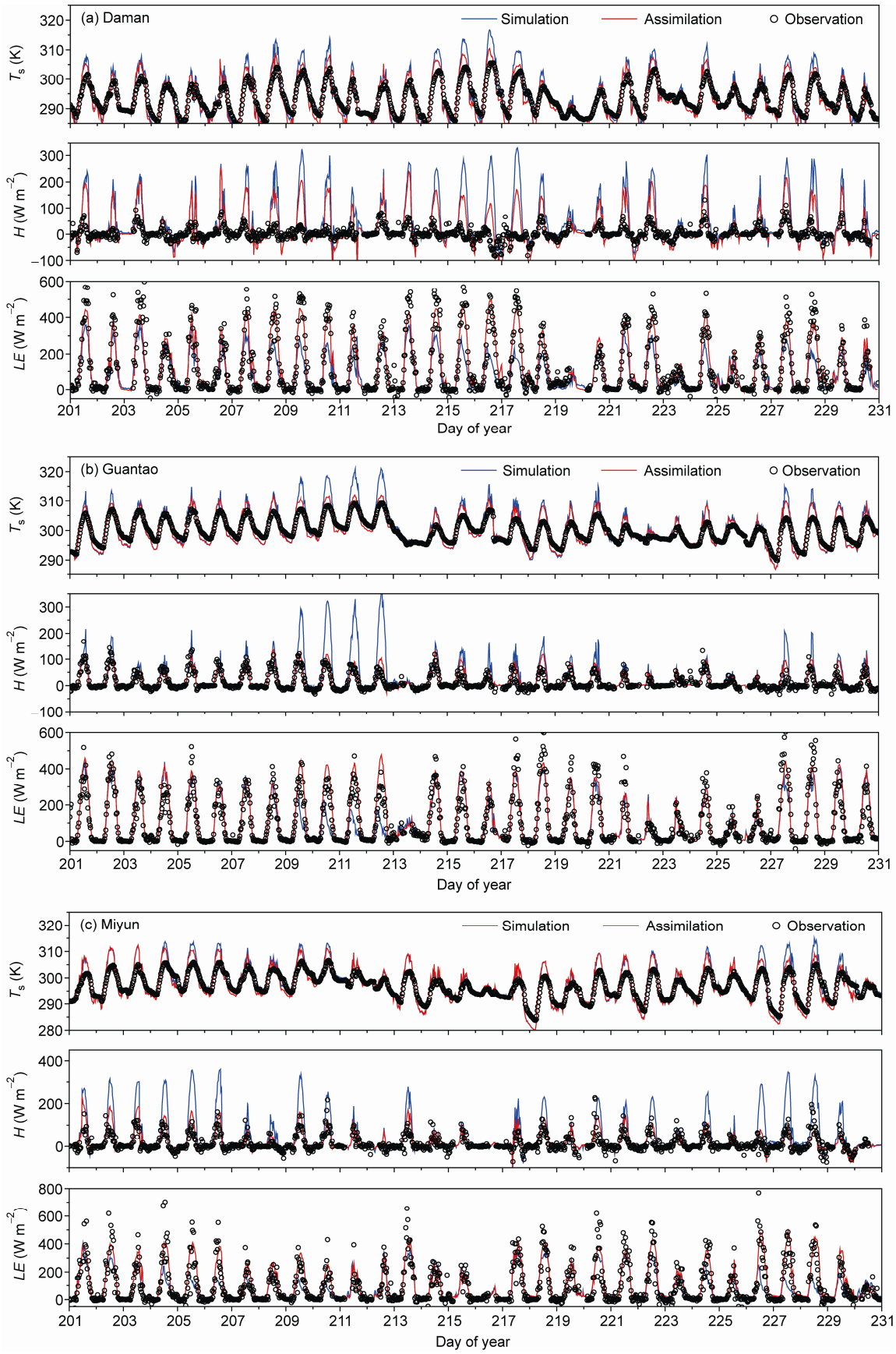

Figure 3 Comparisons of EC-derived sensible and latent heat fluxes with estimates from simulation and assimilation at the Daman, Guantao and Miyun sites.

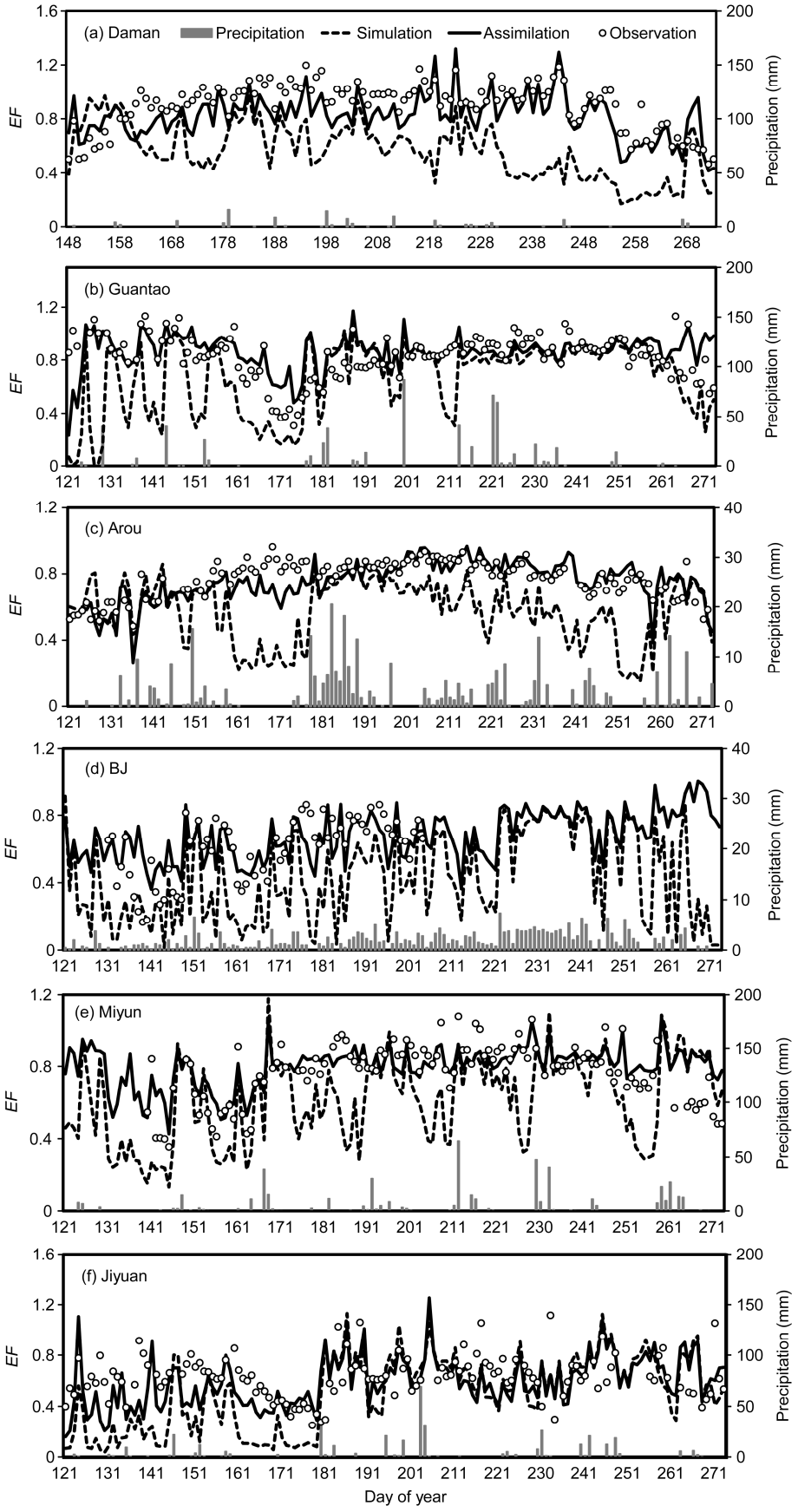

Figure 4 Comparisons of the EC-derived evaporation fraction (EF) with estimates from simulation and assimilation at the six experimental sites.

the average BIAS values of the six sites change from 28.7 to 5.3 W m^{-2} and from −16.9 to 16.1 W m^{-2}; the average RMSE values drop from 71.7 to 46.8 W m^{-2} and from 93.7 to 70.8 W m^{-2}, and the RMSE values drop by 34.7% and 24.4%, respectively. The correlations between the model and observations are increased (the average R values increase from 0.75 to 0.78 and from 0.66 to 0.82). For the evaporation fraction (EF), the average BIAS values of the assimilation EF results drop from –0.19 to 0.04, the average RMSE values drop from 0.31 to 0.15, and the average R-values increase from 0.43 to 0.73, respectively. All of the statistics convey the message that the dual-pass data assimilation scheme reduces the model uncertainties and improves the prediction abilities of the model.

3.2 Comparisons of the simulation and assimilation results with the LAS data

Depending on the spatial representativeness of the ground-measured meteorology data and surface parameters, the spatial scale of the CoLM is approximately 1 kilometer, whereas the representativeness of EC is usually within 1 kilometer. The different spatial scales of these categories may cause problems in the validation. The LAS instrument can measure the sensible heat flux at a larger scale than the EC system, and the LAS-measured sensible heat flux is compared with the dual-pass data assimilation results in this section.

To demonstrate the accuracy of the LAS data, the sensible heat flux derived from the LAS system is compared with data from the EC system (Figure 5). The LAS and EC sensible heat flux values larger than 50 W m^{-2} are used for this comparison. From Figure 5, the sensible heat flux from the LAS system shows the same trend as the EC system, and the sensible heat flux from the LAS is slightly larger than that from the EC. The correlations between them are 0.79, 0.74, 0.87, 0.78, 0.44, and 0.74 at the Daman, Arou, BJ, Guantao, Miyun and Jiyuan sites, respectively. The differences between the sensible heat fluxes measured by the EC and LAS system are primarily caused by the energy imbalance of the EC, the heterogeneity of the underlying surfaces, and the differences between the source areas of the EC and LAS measurements (Liu et al., 2011). The Miyun site is located in a mountainous area with complex land surface

Table 2 BIAS, RMSE and *R*-values of the simulation and assimilation results compared with the half-hourly EC-derived sensible and latent heat fluxes[a)]

Site	Statistics	*H* (W m^{-2})		*LE* (W m^{-2})		*EF*	
		Sim	Ass	Sim	Ass	Sim	Ass
Daman	BIAS	42.7	16.8	–44.6	4.7	–0.32	–0.12
	RMSE	90.9	60.8	112.2	56.6	0.39	0.19
	R	0.60	0.59	0.76	0.93	0.58	0.88
Guantao	BIAS	17.6	–2.3	–9.9	19.4	–0.17	0.04
	RMSE	52.3	30.4	73.6	56.5	0.31	0.15
	R	0.71	0.81	0.79	0.91	0.36	0.56
Arou	BIAS	36.3	10.3	–13.7	26.0	–0.21	–0.02
	RMSE	77.1	34.2	87.3	60.8	0.31	0.09
	R	0.76	0.87	0.76	0.93	0.04	0.72
BJ	BIAS	38.5	–1.6	–12.3	29.6	–0.22	0.09
	RMSE	78.5	44.2	95.8	76.0	0.28	0.16
	R	0.88	0.87	0.50	0.78	0.54	0.72
Miyun	BIAS	19.6	0.8	–8.3	16.5	–0.16	0.03
	RMSE	56.2	43.3	75.3	68.4	0.25	0.11
	R	0.70	0.68	0.77	0.86	0.46	0.76
Jiyuan	BIAS	17.2	7.8	–12.8	0.5	–0.16	–0.06
	RMSE	75.2	67.9	117.9	106.2	0.32	0.21
	R	0.84	0.85	0.39	0.49	0.59	0.76
Average	BIAS	28.7	5.3	–16.9	16.1	–0.21	–0.01
	RMSE	71.7	46.8	93.7	70.8	0.31	0.15
	R	0.75	0.78	0.66	0.82	0.43	0.73

a) In this work, BIAS, RMSE and *R* indicate the bias, root mean square error and correlation between the model results and observations, respectively; *H*, *LE,* and *EF* are the sensible and latent heat flux (W m^{-2}) and evapotranspiration fraction, respectively; *Sim* denotes model simulation results; and *Ass* is the data assimilation results with the dual-pass data assimilation scheme.

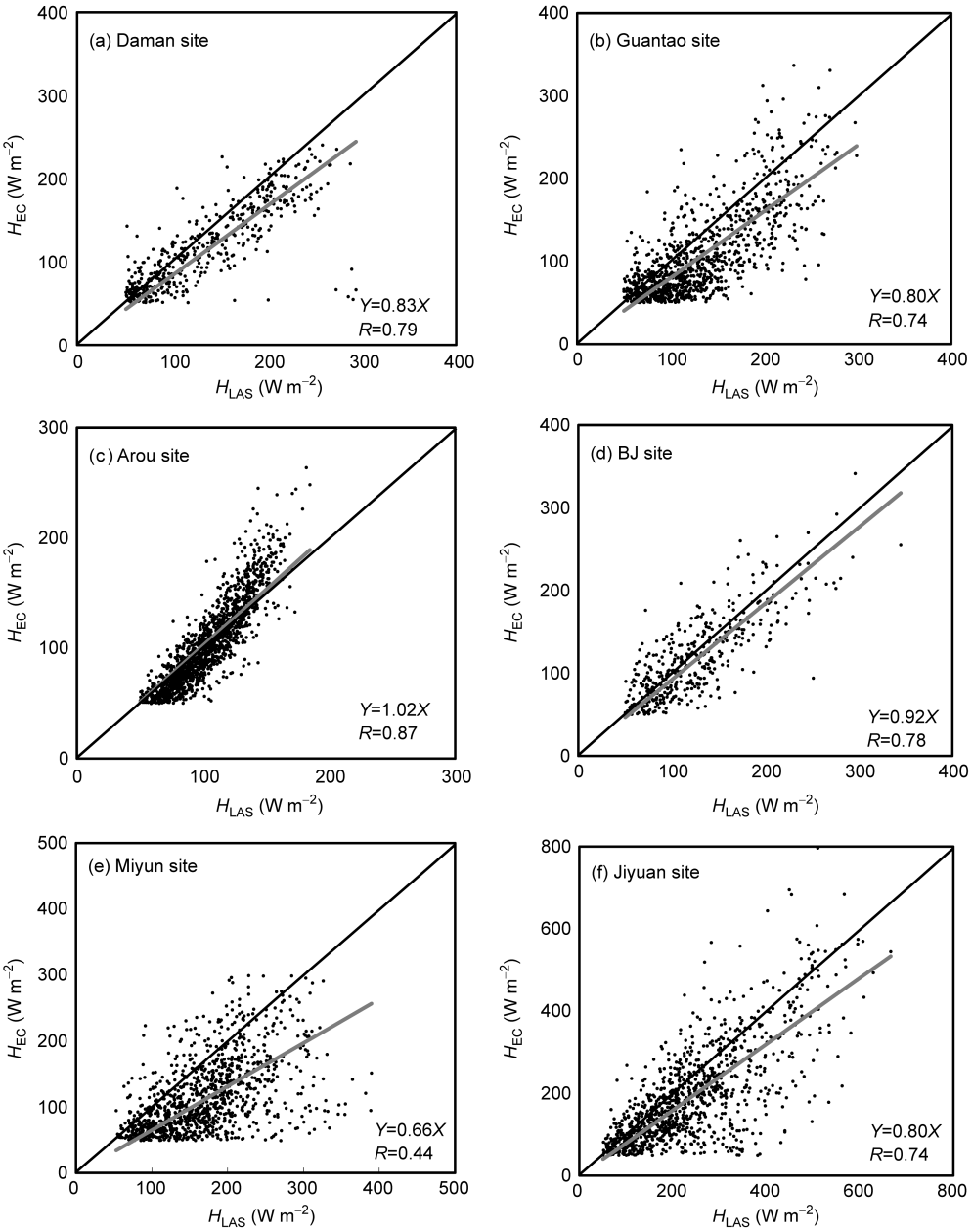

Figure 5 Comparisons of sensible heat flux derived from the LAS (H_{LAS}) and EC (H_{EC}) when H_{LAS} and H_{EC} are greater than 50 W m^{-2} at the six experimental sites.

characteristics that may lead to large differences between the LAS and EC-derived sensible heat fluxes.

Because the source area of the LAS measurements is larger than one FY3A-VIRR pixel, the weight of each covered pixel should be determined for the comparison. The source areas are calculated using a footprint model (eq. (18)) and are overlaid with the FY3A-VIRR pixel at the experiment sites shown in Figure 6. The source area of 50% contribution to the measured fluxes is used for Miyun site (Miyun site is located in a valley of mount areas), and 80% contribution to the measured fluxes is used for the other five sites. The figure shows that the LAS source area did not show obvious variation and extended from the transmitter to the receiver point with main contributing source areas with a width of approximately half a pixel and a length of 2 to 3 pixels. The averaged sensible heat flux can be determined as in (Jia et al., 2012):

$$H_{\text{average}} = \sum_{i=1}^{n} \left(w_i \times H_i \right), \tag{21}$$

where H_{average} represents the model results with the same spatial representativeness as the LAS observation, w_i is the relative weight of the LAS footprint in each pixel, H_i is the model result of each remote sensing pixel, and n is the

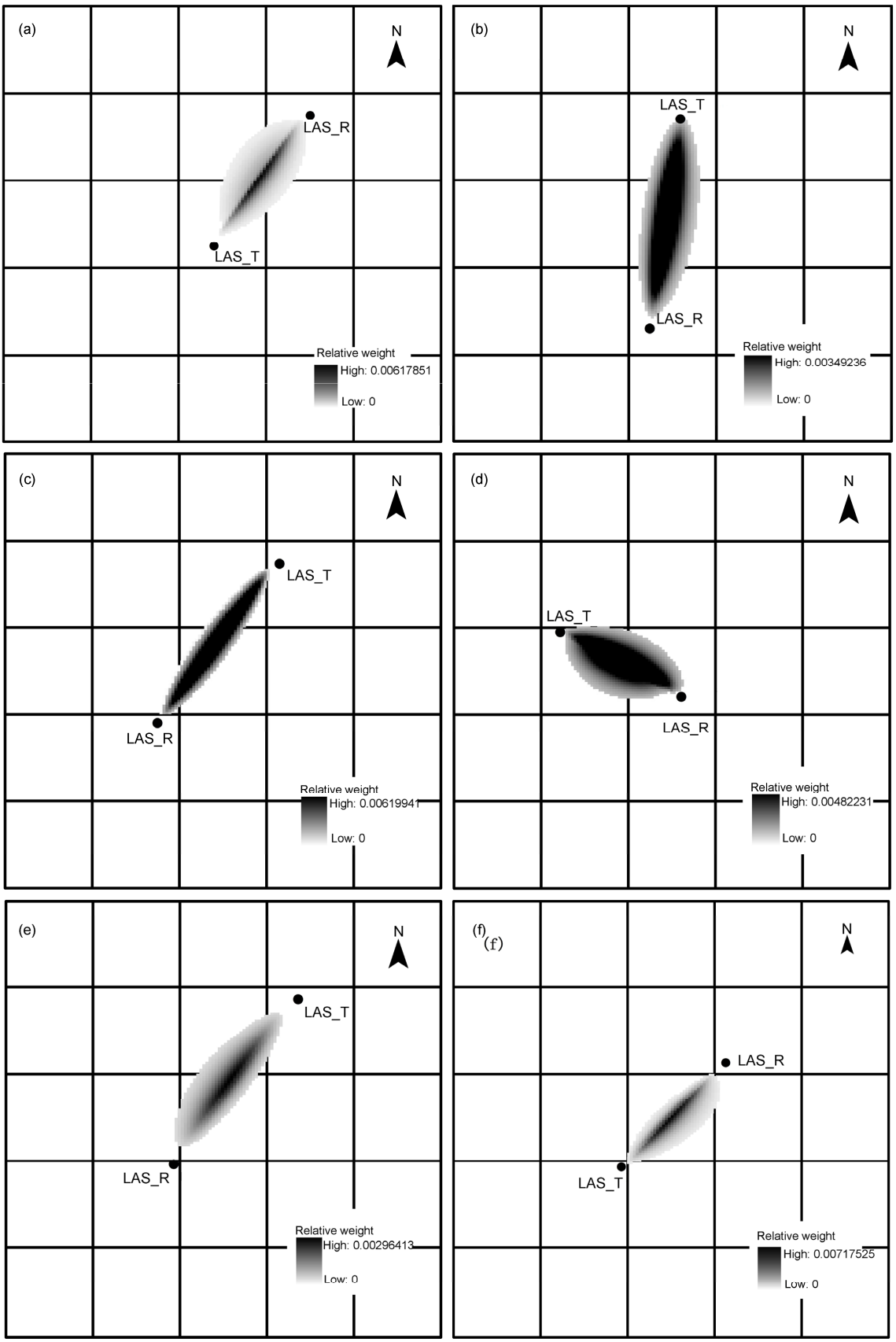

Figure 6 The monthly source areas of the LAS measurements overlaid with the FY3A-VIRR pixels at the six experimental sites: (a) Daman site; (b) Guantao site; (c) Arou site; (d) BJ site; (e) Miyun site; (f) Jiyuan site.

number of pixels within the source area.

The comparison results at the six experimental sites are shown in Figure 7. From this figure, the sensible heat flux with the assimilation of the FY3A-VIRR LST is closer to the LAS observations than the model simulation, and the scatter with the assimilation of the FY3A-VIRR LST is closer to the 1∶1 line (with a smaller BIAS). The larger *R*-values indicate that the assimilation results show a better correlation with the LAS measurements than the model simulation. The averaged RMSE for the six sites values drop from 110.3 to 76.3 W m^{-2} (the RMSE values drop by 30.8%). At the Miyun site, the sensible heat flux from the

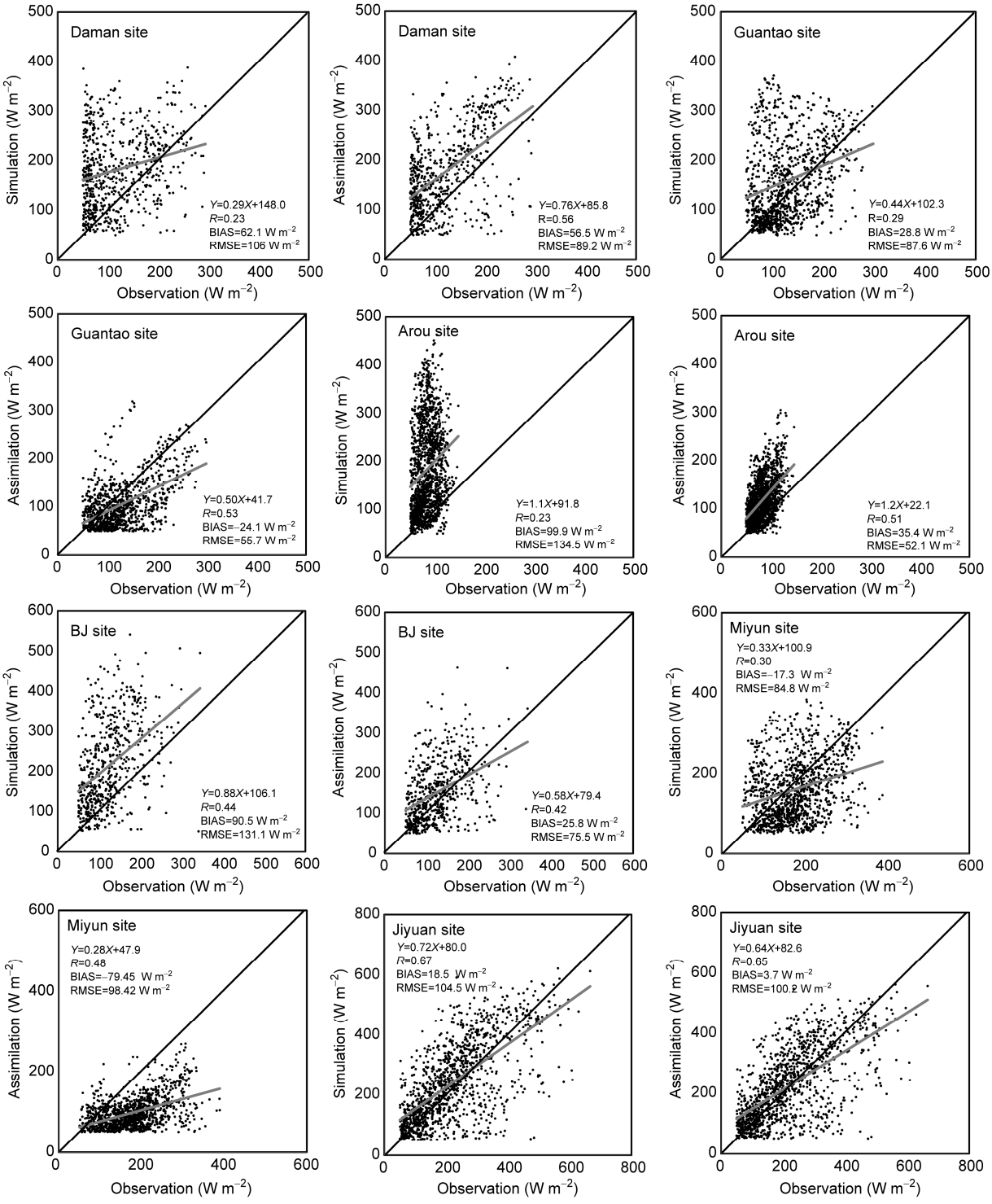

Figure 7 Comparisons of the LAS-derived sensible heat flux with estimates from simulation and assimilation when H_{LAS}>50 W m^{-2} at the six experiment sites.

LAS exhibits a large difference with data from the EC (Figure 5), and the R-value of the data assimilation result increases from 0.30 to 0.48 with the LAS validation, whereas the BIAS value changes from −17.3 to −79.5 W m^{-2} and the RMSE value increases from 84.8 to 98.4 W m^{-2}.

3.3 Error sources in surface flux predictions

Because soil moisture plays an important role in the terrestrial water cycle, the vegetation parameters play a significant role in water and energy movement among the land surface, canopy and atmosphere. The soil moisture and vegetation parameters are selected and optimized with the dual-pass data assimilation scheme to improve the predictions of surface flux. This section first presents the retrieved soil moisture and vegetation parameters from the dual-pass data assimilation scheme. Next, the error sources in the surface flux predictions are investigated and analyzed.

Figure 8 shows the soil moisture retrievals at depths of 0.05, 0.1, 0.2, 0.4, 0.6 and 1.0 m at the Miyun site (Orchard) from day 121 to 273 in the year 2010. Generally, the soil moisture estimates at certain soil depths after the data assimilation show better agreement with their corresponding in situ observations than the model simulation without the assimilation. At other depths, the assimilation may degrade the soil moisture estimates. As shown in Figure 8, at depths of 0.05, 0.1, 0.2 and 0.6 m, the assimilation results are closer to the observations than the model predictions during this time period, and the RMSE values of the soil moisture retrievals decrease through the assimilation of the FY3A-VIRR LST data. At depths of 0.4 and 1.0 m, the data assimilation scheme degrades the soil moisture estimates. To obtain accurate surface flux predictions, the Pass 2 of the scheme may sacrifice the model soil moisture prediction accuracies through minimizing the difference between simulated and observed land surface temperature. The soil moisture data assimilation results at the other sites show performance similar to that of the Miyun site; thus, the results are not shown.

Table 3 summarizes the averaged values of the land surface temperature, sensible and latent heat fluxes, and soil moisture for the model simulation and assimilation with Pass 2 (optimization of soil moisture). This table shows the relationships among soil moisture, land surface temperature, and surface fluxes. According to this table, with increased soil moisture, the latent heat flux increased accordingly, and the land surface temperature and sensible heat flux decreased. For the averaged values of the six sites, the soil moisture increased from 0.17 to 0.24 m^3 m^{-3} (increased 38%), the land surface temperature decreased from 291.9 to 290.9 K, the sensible heat flux decreased from 46.2 to 28.5 W m^{-2} (decreased 38%), and latent heat flux increased from 69.6 to 94.5 W m^{-2} (increased 36%).

Figure 9 shows the seasonal variations of the five retrieved vegetation parameters at the Miyun site (Orchard) from day 121 to 273 of the year 2010. From Figure 9, the parameters are retrieved at the weekly temporal scales with the dual-pass data assimilation scheme, which indicates that the vegetation parameters are updated once per week. With the assimilation of the FY3A-VIRR LST data, the parameter uncertainties are within a relatively stable range (the error bars of Figure 9). The vegetation parameters tend to show stable but not constant values as the vegetation parameters change with the seasonal variations of the vegetation conditions. The surface roughness length (z0m), maximum rate of carboxylation at 25°C (vmax25), conductance-photosynthesis slope parameter (gradm) and conductance-photosynthesis intercept (binter) increase to high values, and the quantum efficiency at 25°C (effcon) does not change significantly compared with the initial value, which means that this parameter agreed well with the real condition.

The errors in the surface flux predictions are partly model biases (Figures 3–4), and the dual-pass data assimilation can reduce the model biases significantly (the EF decreases from −0.21 to −0.01 in Table 3). The evaporation fraction (EF) BIAS between the predicted and EC observations are shown in Figure 10, which presents the abilities of Pass 1 and Pass 2 of the dual-pass data assimilation scheme to reduce the biases at the six sites. Generally, the dual-pass data assimilation scheme reduces the model biases and performs

Table 3 Averaged values of simulation and assimilation results[a)]

	T_s (K)		H (W m^{-2})		LE (W m^{-2})		SM (m^3 m^{-3})	
Site	Sim	Pass 2	Sim	Pass 2	Sim	Pass 2	Sim	Pass 2
Daman	292.7	290.4	51.2	36.5	69.7	101.8	0.31	0.36
Guantao	297.7	296.5	34.1	15.2	70.1	97.5	0.21	0.26
Arou	285.8	285.2	55.7	38.9	77.5	95.7	0.11	0.24
BJ	284.9	283.6	63.6	42.9	62.4	93.8	0.06	0.07
Miyun	293.9	294.3	28.4	26.5	80.4	82.5	0.13	0.21
Jiyuan	296.5	295.8	43.9	10.7	57.7	95.7	0.20	0.27
Average	291.9	290.9	46.2	28. 5	69.6	94.5	0.17	0.24

a) In this work, T_s is the land surface temperature (K); H and LE are the sensible and latent heat fluxes (W m^{-2}), respectively; SM is the soil moisture (m^3 m^{-3}), *Sim* is the model simulation results; and Pass 2 is the data assimilation results with Pass 2 (optimize soil moisture) of the dual-pass data assimilation scheme.

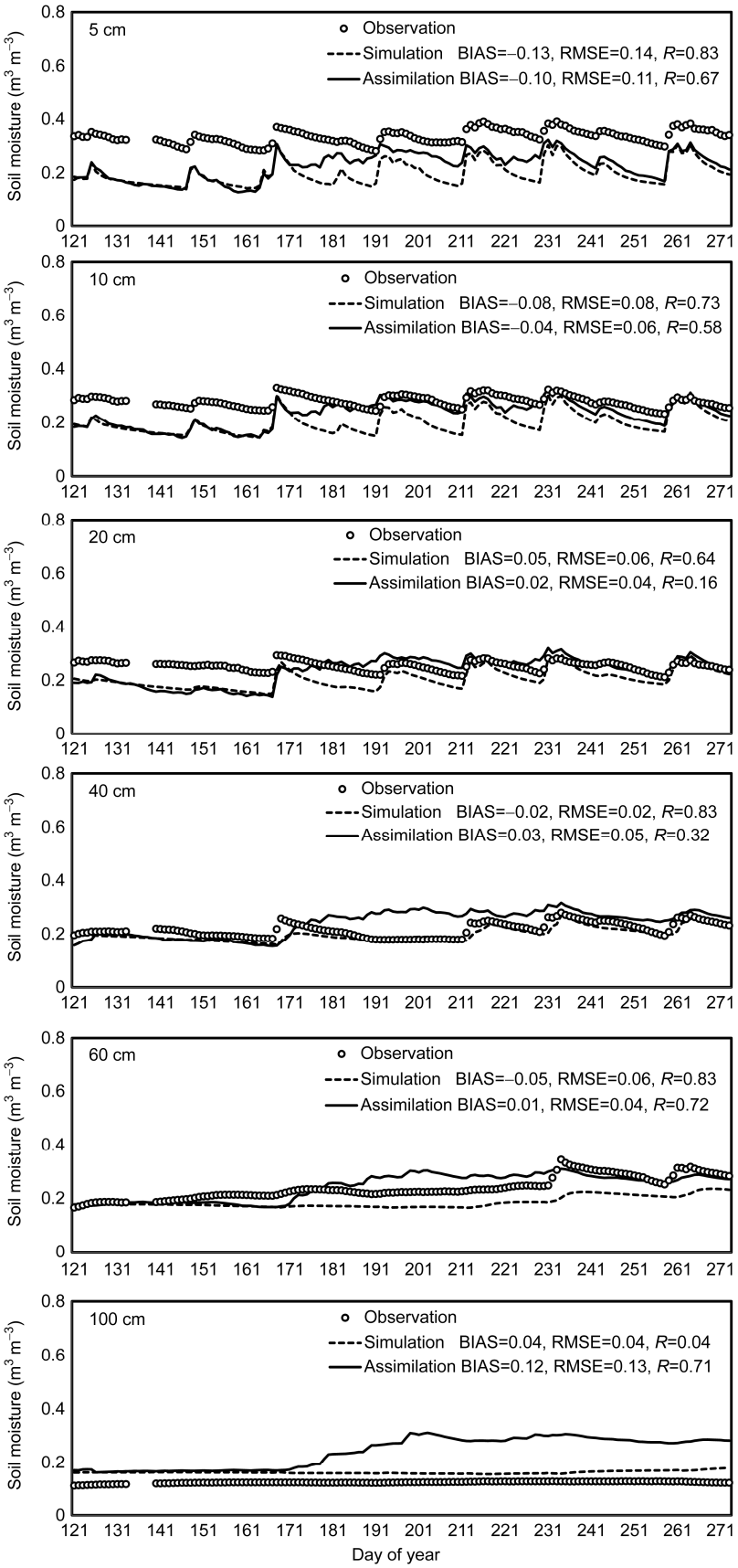

Figure 8 Comparisons of the soil moisture measurements with estimates from the simulation and assimilation at the Miyun site from day 121 to 273 of year 2010.

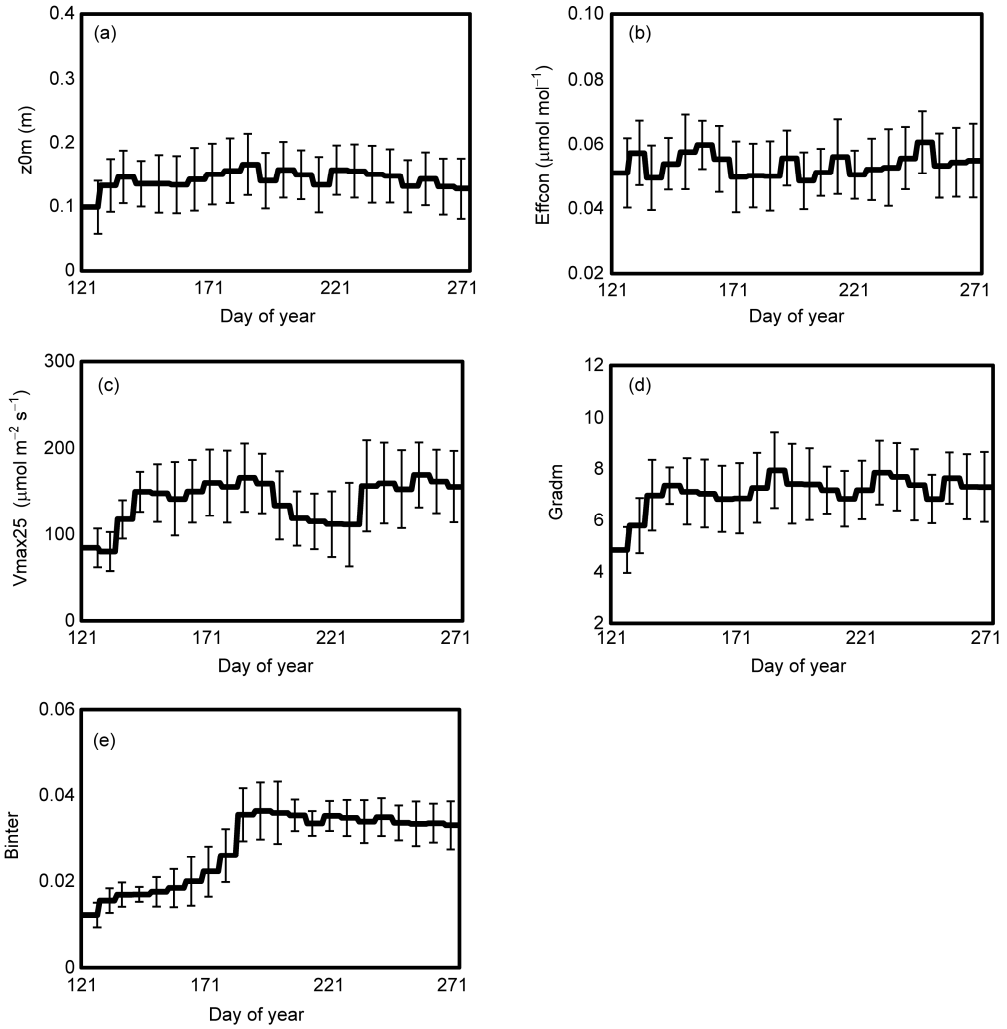

Figure 9 Model parameter retrievals at the Miyun site from day 121 to 273 of the year 2010. error bars indicate the uncertainty of the retrieved parameter.

better than the use of Pass 1 and Pass 2 independently. Both Pass 1 (optimization of the vegetation parameter) and Pass 2 of the scheme (optimization of the soil moisture) play important roles in the surface flux predictions. Pass 2 performs better than Pass 1, which means that soil moisture plays an important role in predictions of surface flux. Because Pass 1 optimizes the vegetation parameters for the each week, it is better able to cope with the high LAI values. Pass 1 performs better at a densely vegetated site (e.g., the Daman site) than at sparsely vegetated sites (e.g., the BJ site). With Pass 1 and Pass 2, the dual-pass data assimilation scheme can reduce most of the surface flux prediction biases and produce more stable results.

4 Conclusions

In this study, a dual-pass data assimilation scheme was constructed to estimate the surface fluxes via the independent optimization of the soil moisture and vegetation parameters. Pass 1 of the data assimilation scheme optimized the vegetation parameters at the weekly temporal scale, and Pass 2 optimized the soil moisture at the daily temporal scale. The land surface temperature data from the new generation Chinese meteorology satellite FY3A-VIRR were assimilated into CoLM for the first time based on the EnKF algorithm. The results were validated using the multi-scale surface flux observations (derived from the EC and LAS) at six sites. Ultimately, the soil moisture and vegetation parameters were retrieved, and the error sources in surface flux predictions were analyzed.

According to comparisons with the EC-derived sensible and latent heat flux, the assimilation curves matched well with observations (Figure 3). The dual-pass data assimilation scheme reduced the model uncertainties and improved the prediction abilities of the model. The average BIAS values of the six sites changed from 28.7 to 5.3 W m^{-2} and from –16.9 to 16.1 W m^{-2}; the average RMSE values dropped from 71.7 to 46.8 W m^{-2} and from 93.7 to 70.8 W m^{-2}; and the average R-values increased from 0.75 to 0.78 and from

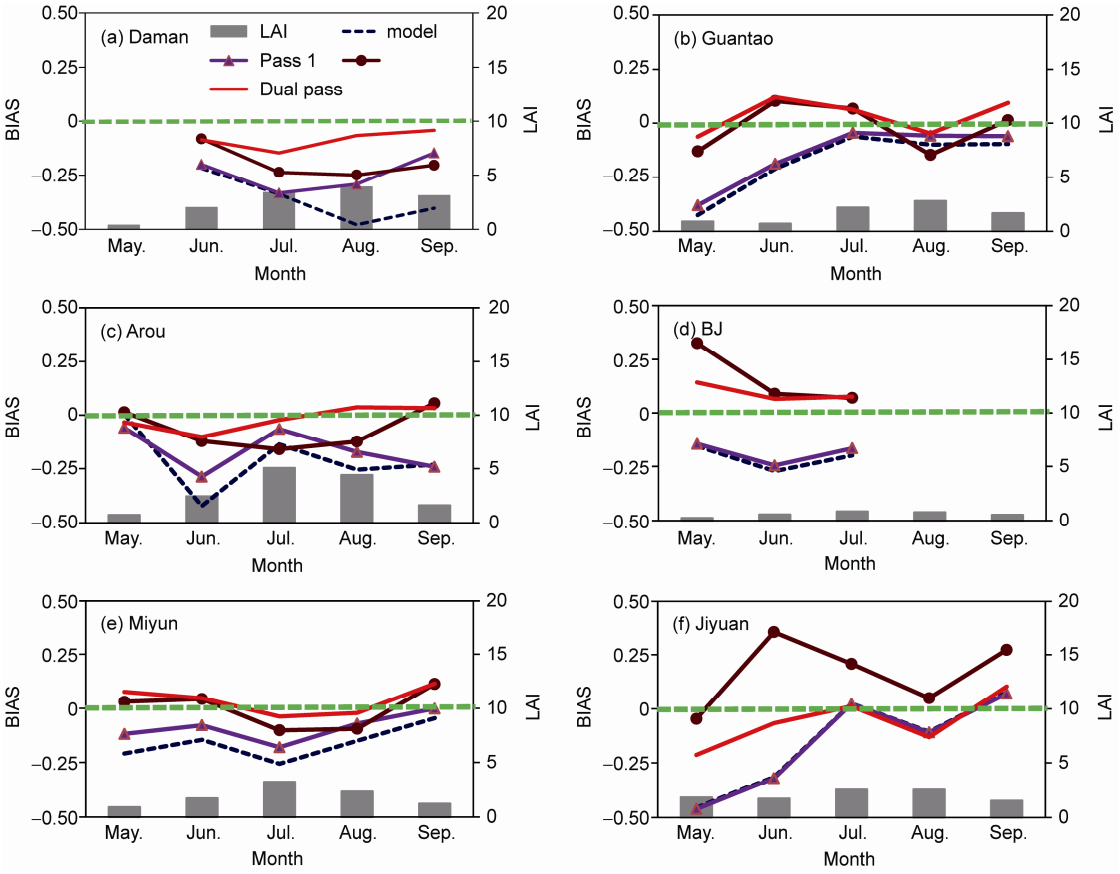

Figure 10 Evaporation fraction (EF) biases (BIAS) at the experimental sites.

0.66 to 0.82, respectively. The evaporation fraction (EF) was also used to assess the performance of the dual-pass data assimilation scheme (Figure 4). The EF simulations quickly decreased, and the values were lower than the observations after the precipitation occurred. With the assimilation of the FY3A-VIRR LST, the scheme improved the underestimation of the EF estimates and improved the distribution of available energy into sensible and latent heat fluxes. The average BIAS values of the assimilation EF results dropped from –0.21 to –0.01, the average RMSE values dropped from 0.31 to 0.15, and the average R-values increased from 0.43 to 0.73.

Furthermore, the sensible heat flux measured by the LAS with larger spatial representativeness was used to validate the results. The source areas of the LAS measurements were calculated using a footprint model and overlaid with a FY3A-VIRR pixel, which are larger than one FY3A-VIRR pixel (Figure 6). The sensible heat fluxes calculated from each covered FY3A-VIRR pixel were averaged and compared with the LAS-derived sensible heat flux. The comparisons showed that the assimilation results match well against the LAS measurements, and the correlations between the assimilation results and LAS measurements were higher than that of the model simulations at the Daman, Guantao, Arou, BJ, and Jiyuan sites (Figure 7).

In addition to the estimates of surface flux, the dual-pass data assimilation scheme also retrieved the model soil moisture and vegetation parameters. The soil moisture estimates at certain soil depths after the data assimilation show better agreement with their corresponding in-situ observations than the model simulation without the assimilation. At other depths, the assimilation may degrade the soil moisture estimates (Figure 8). Furthermore, the five vegetation parameters, i.e., surface roughness length (z0m), quantum efficiency at 25°C (effcon), maximum rate of carboxylation at 25°C (vmax25), conductance-photosynthesis slope parameter (gradm), and conductance-photosynthesis intercept (binter) were retrieved at the weekly temporal scales (Figure 9).

The aim of this study was to eliminate or reduce the errors between the model simulation and observation. Because the land surface temperature and surface fluxes are prognostic variables in CoLM, the simulation is quite complicated and is affected by many factors. From Figure 4, the EF simulations were increased relative to the observations after the precipitation occurred, which means that soil moisture is an important variable in CoLM for surface flux predictions. Furthermore, the accuracy of the vegetation parameters is vital for canopy transpiration. Thus, the dual- pass data assimilation is developed. Pass 1 optimizes the vegetation parameters at the weekly temporal scale, which is similar to

a parameter calibration procedure, and Pass 2 optimizes the soil moisture at the daily temporal scale. Both passes focus on reducing the errors of the model on the weekly and daily temporal scales and correcting the model predictions with the assimilation of the FY3A-VIRR LST data. The model errors in the surface flux predictions were investigated in Section 3.3 of this study (Figure 10). Figure 10 showed that the model biases in the surface flux predictions were caused by both the soil moisture and vegetation parameters. With Pass 1 and Pass 2, the dual-pass data assimilation scheme is able to reduce most of the surface flux prediction biases (BIAS).

Because of the short memory of the land surface temperature, it is difficult to assimilate the land surface temperatures. Directly assimilating land surface temperature to optimize soil moisture may cause model simulations that are more unstable (Figure 8). In this study, the predictions of the surface flux and land surface temperature can be improved by updating the soil moisture, but the updated soil moisture may not fit the observations well at all conditions (Figure 8). Thus, converting the land surface temperature into soil moisture or evapotranspiration (ET) using diagnostic models first and subsequently assimilating the converted variables into the model is one approach to avoid the model instability. Future studies are planned to assimilate the land-surface-temperature-converted variables (soil moisture or ET) into the land surface model and apply the data assimilation scheme over regions (e.g., Heihe River Basin, Haihe River Basin) located in China.

This work was funded by the National Natural Science Foundation of China (Grant Nos. 91125002, 41201330), the Fundamental Research Funds for the Central Universities, and the Special Foundation for Free Exploration of State Laboratory of Remote Sensing Science (Grant No. 13ZY-06). The authors thank Prof. Jinsong Zhang for providing Jiyuan site data used in this study.

Baldocchi D, Falge E L, Gu H, et al. 2001. FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. Bull Am Meteorol Soc, 82: 2415–2434

Bastiaanssen W G M, Menenti M, Feddes R A, et al. 1998. The surface energy balance algorithm for land (SEBAL): Part 1 Formulation. J Hydrol, 212-213: 198–212

Bateni S M, Liang S. 2012. Estimating surface energy fluxes using a dual-source data assimilation approach adjoined to the heat diffusion equation. J Geophys Res, 117: D17118

Becker F, Li Z L. 1990. Towards a local split window method over land surface. Int J Remote Sens, 11: 369–393

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 Resour Res, 37: 1713–1722

Caparrini F, Castelli F, Entekhabi D. 2004. Estimation of surface turbulent fluxes through assimilation of radiometric surface temperature sequences. J Hydrometeor, 5: 145–159

Dai Y J, Zeng X B, Dickinson R E, et al. 2003. The common land model. Bull Am Meteorol Soc, 84: 1013–1023

Dai Y J, Dickinson R E, Wang Y P. 2004. A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance. J Clim, 17: 2281–2299

Dickinson R E, Kennedy P J, Henderson-Sellers A, et al. 1986. Biosphere-atmosphere transfer scheme (BATS) version 1E as coupled to the NCAR Community Climate Model. NCAR Tech Rep TN-275+STR

Duan Q Y, Gupta V K, Sorooshian S. 1993. Shuffled complex evolution approach for effective and efficient global minimization. J Optimiz Theory App, 76: 501–521

Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. J Geophys Res, 99: 10143–10162

Huang C L, Li X, Lu L. 2008. Retrieving soil temperature profile by assimilating MODIS LST products with ensemble Kalman filter. Remote Sens Environ, 112: 1320–1336

Jia Z, Liu S, Xu Z, et al. 2012. Validation of remotely sensed evapotranspiration over the Hai River Basin, China. J Geophys Res, 117: D13113, doi: 10.1029/2011JD017037

Kanemasu E T, Verma S B, Smith E A, et al. 1992. Surface flux measurements in FIFE: An overview. J Geophys Res, 97: 18547–18555

Li X, Li X W, Li Z Y, et al. 2009. Watershed allied telemetry experimental research. J Geophys Res, 114: D22103

Li X, Cheng G D, Liu S M, et al. 2013. Heihe watershed allied telemetry experimental research (HiWATER): Scientific objectives and experimental design. Bull Am Meteorol Soc, 94: 1145–1160

Liang S. 2004. Quantitative Remote Sensing of Land Surfaces. New York: John Wiley and Sons

Liang S, Qin J. 2008. Data assimilation methods for land surface variable estimation. In: Liang S, ed. Advances in Land Remote Sensing: System, Modeling, Inversion and Applications. New York: Spriner-Verlag Inc. 319–339

Liou Y, Galantowicz J F, England A W. 1999. A land surface process/radiobrightness model with coupled heat and moisture transport for prairie grassland. IEEE Trans Geosci Remote Sensing, 37: 1848–1859

Liu S M, Hu G, Lu L, et al. 2007. Estimation of regional evapotranspiration by TM/ETM+ data over heterogeneous surfaces. Photogramm Eng Remote Sens, 73: 1169–1178

Liu S M, Xu Z W, Wang W Z, et al. 2011. A comparison of eddy-covariance and large aperture scintillometer measurements with respect to the energy balance closure problem. Hydrol Earth Syst Sci, 15: 1291–1306

Liu S M, Xu Z W, Zhu Z L, et al. 2013. Measurements of evapotranspiration from eddy-covariance systems and large aperture scintillometers in the Hai River Basin, China. J Hydrol, 487: 24–38

Lu H, Koike T, Yang K, et al. 2012. Improving land surface soil moisture and energy flux simulations over the Tibetan plateau by the assimilation of the microwave remote sensing data and the GCM output into a land surface model. Int J Appl Earth Observ Geoinf, 17: 43–54

Margulis S A, McLaughlin D, Entekhabi D, et al. 2002. Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 field experiment. Water Resour Res, 38: 1229

Moradkhani H, Sorooshian S, Gupta H V, et al. 2005. Dual state-parameter estimation of hydrological models using ensemble Kalman filter. Adv Water Resour, 28: 135–147

Pipunic R C, Walker J P, Western A. 2008. Assimilation of remotely sensed data for improved latent and sensible heat flux prediction: A comparative synthetic study. Remote Sens Environ, 112: 1295–1305

Qin J, Liang S, Yang K, et al. 2009. Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal. J Geophys Res, 114: D15103

Schuurmans J M, Troch P A, Veldhuizen A A, et al. 2003. Assimilation of remotely sensed latent heat flux in a distributed hydrological model. Adv Water Resour, 26: 151–159

Schuurmans J M, Geer F C, Bierkens M F P. 2011. Remotely sensed latent heat fluxes for model error diagnosis: Acase study. Hydrol Earth Syst Sci, 15: 759–769

Sellers P J, Randall D A, Collatz G J, et al. 1996. A revised land surface parameterization (SiB2) for atmosphere GCMs, part I: Model formulation. J Clim, 9: 676–705

Shangguan W, Dai Y, Liu B, et al. 2012. A soil particle-size distribution dataset for regional land and climate modelling in China. Geoderma,

171-172: 85–91
Su Z. 2002. The surface energy balance system (SEBS) for estimation of turbulent heat fluxes. Hydrol Earth Syst Sci, 6: 85–99
Saltelli A, Bolado R. 1999. An alternative way to compute Fourier amplitude sensitivity test (FAST). Comput Stat Data An, 26: 445–460
Tian X, Xie Z, Dai A, et al. 2009. A dual-pass variational data assimilation framework for estimating soil moisture profiles from AMSR-E microwave brightness temperature. J Geophys Res, 114: D16102
Wang W H, Liang S L, Tilden M. 2008. Validating MODIS land surface temperature products using long-term nighttime ground measurements, Remote Sens Environ, 112: 623–635
Xiao Z Q, Liang S, Wang J D, et al. 2011. Real-time retrieval of leaf area index from MODIS time series data. Remote Sens Environ, 115: 97–106
Xu T, Liu S M, Liang S, et al. 2011. Improving predictions of water and heat fluxes by assimilating MODIS land surface temperature products into common land model. J Hydrometeorol, 12: 227–244
Yang H, Yang Z D. 2006. A modified land surface temperature split window retrieval algorithm and its applications over China. Glob Planet Change, 52: 207–215
Yang J, Dong C H. 2011. Operational Products and Application for the New Generation of Polar-Orbiting FY Meteorological Satellites (in Chinese). Beijing: Science Press
Yang K, Takahiro W, Toshio K. 2007. An auto-calibration system to assimilate AMSR-E data into a land surface model for estimating soil moisture and surface energy budge. J Meteorol Soc Jpn, 85: 229–242
Zhao L, Yang K, Qin J, et al. 2012. Optimal exploitation of AMSR-E signals for improving soil moisture estimation through land data assimilation. Ieee Trans Geosci Remote Sensing, 51: 399–410