Hydrol. Earth Syst. Sci. Discuss., 9, 8493–8534, 2012
www.hydrol-earth-syst-sci-discuss.net/9/8493/2012/
doi:10.5194/hessd-9-8493-2012

Hydrology and Earth System Sciences Discussions

A dual-pass data assimilation scheme for estimating turbulent fluxes with FY3A data

T. R. Xu[1], S. M. Liu[1], Z. W. Xu[1], S. Liang[2], and L. Xu[1]

[1]State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University, Beijing 100875, China
[2]Department of Geography, University of Maryland, College Park, MD 20742, USA

Received: 8 June 2012 – Accepted: 19 June 2012 – Published: 11 July 2012

Correspondence to: T. R. Xu (xutr@bnu.edu.cn)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

A dual-pass data assimilation scheme is developed to improve predictions of turbulent fluxes with FY3A land surface temperature (LST) data. This scheme is constructed based on the ensemble Kalman filter (EnKF) and common land model (CoLM). Pass 1 of the dual-pass data assimilation scheme optimizes model vegetation parameters at a long temporal scale and pass 2 optimizes soil moisture at a short temporal scale. Four sites are selected for the data assimilation experiments, namely Arou, BJ, Guantao, and Miyun in the People's Republic of China (PRC) that include grass, alpine meadow, crop, and orchard land cover types. The results are compared with data generated by a multi-scale turbulent flux observation system that includes an eddy covariance (EC) and a large aperture scintillometer (LAS) system. Results indicate that the CoLM can simulate the diurnal variations of turbulent flux, but usually underestimates the latent heat flux and evaporation fraction (EF) and overestimates sensible heat flux. With the assimilation of FY3A LST data, the dual-pass data assimilation scheme can improve the predictions of turbulent flux. The average root mean square error (RMSE) values drop from 81.2 to 39.6 $W m^{-2}$ and from 101.7 to 58.9 $W m^{-2}$ (the RMSE values drop 51.2 % and 42.1 %) for sensible and latent heat fluxes, respectively. To compare the results with LAS measurements, the source areas are calculated using a footprint model and overlaid with FY3A pixels since the LAS cover more than one FY3A pixel. The comparisons show that the assimilation results are closer to LAS measurements. With the dual-pass data assimilation scheme, the estimated soil moistures are generally closer to observations. Furthermore, the vegetation parameters are retrieved and incorporated into CoLM which enhanced the model's predictive abilities.

HESSD

9, 8493–8534, 2012

1 Introduction

Accurate modeling and estimation of turbulent fluxes at the land surface are necessary for climate modeling; for predicting the impact of land-use changes; and for agricultural and water resource planning. The magnitude of turbulent fluxes is largely determined by vegetation parameters and the soil moisture and temperature of the land surface. The major methods used for estimating turbulent fluxes are ground measurements, remote sensing based method, and land surface modeling.

Field measurements of turbulent flux have documented their variability over diurnal, seasonal, and inter-annual time scales (Liu et al., 2011). Some 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) and the FLUXNET (Baldocchi et al., 2001). Direct observations such as lysimeter, eddy covariance systems, Bowen ratio methods, and large aperture scintillometer, are necessary for increasing our understanding of water and energy balance at the land surface. However, these measurements are difficult to use for monitoring turbulent fluxes at the regional scale, since they only produce either point or patch-scale data.

Spatially-distributed estimates of turbulent flux can be obtained by remote sensing based method (Bastiaanssen et al., 1998; Su, 2002; Liu et al., 2007). However, these methods are difficult to use to monitor turbulent fluxes continuously, since data derived from satellite are instantaneous and often contaminated by the presence of clouds. Land surface models have been developed rapidly in the past two decades to predict turbulent 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 structure, all land surface models need observational data to calibrate their parameters and adjust their states. New techniques such as data assimilation, is needed to integrate either field or remotely-sensed observations with models, and in so doing, improve model accuracy by correcting model state variables and parameters (Liang, 2004; Liang and Qin, 2008). Data assimilation has

HESSD
9, 8493–8534, 2012

played an increasingly important role in improving predictions of land surface state variables such as leaf area index (Xiao et al., 2011), soil temperature (Huang et al., 2008), soil moisture (Margulis et al., 2002), and other related variables such as turbulent fluxes (Xu et al., 2011a,b).

At the same time, the sequential data assimilation techniques have been developed for data assimilation. The Kalman filter (KF) is an optimal recursive data processing technique for linear dynamic systems first proposed by Kalman (1960). Later, the nonlinear version namely, extended Kalman filter (EKF), was developed by Jazwinski (1970). The Ensemble Kalman filter (EnKF) proposed by Evensen (1994) is another extension of the traditional KF based on Monte Carlo sampling and recursive data processing. The Easy implementation of the EnKF method has led to its widespread application in land surface researches such as hydrology and ecology (Reichle et al., 2002; Mo et al., 2008). Recently, researchers have begun to use EnKF to estimate model state variables and parameters together (Moradkhani et al., 2005; Xu et al., 2011b).

In order to model turbulent fluxes at the land surface, the acquisition of accurate land surface temperatures is very important. Land surface temperature can reflect the relative humidity at the land surface. For the same magnitude of solar radiation, the lower land surface temperatures results from a wet land surface associated with relatively high latent heat flux and low sensible heat flux. The inaccurate turbulent flux predictions are highly correlated with biases in land surface temperature estimates. Therefore, assimilation of land surface temperatures has become an important way for improving turbulent flux predictions from land surface models. Huang et al. (2008) improved the surface ground heat flux predictions in the common land model (CoLM) (Dai et al., 2003) with the assimilation of radiometric temperature measurements derived from remote sensing data. The turbulent fluxes can also be obtained on the basis of variational techniques and relatively simple models with the assimilation of field measured land surface temperatures (Boni et al., 2001; Caparrini et al., 2004). Remotely-sensed land surface temperatures have also been assimilated into relatively complicated land surface models, and improved the modeling accuracy of turbulent

HESSD
9, 8493–8534, 2012

fluxes (Xu et al., 2011a,b). The surface energy balance algorithm for land (SEBAL) or surface energy balance system (SEBS) can calculate instantaneous turbulent fluxes using field 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) did experiments with the assimilation of land surface temperature, land surface soil moisture and turbulent fluxes derived from synthetic remote sensing data, and the most obvious improvements were found with the assimilation of remotely-sensed turbulent fluxes.

A number of previous studies have sought to improve turbulent flux predictions by assimilating remotely-sensed land surface temperatures. Xu et al. (2011a) have done a series of experiments that improved turbulent flux predictions with the assimilation of MODIS land surface temperature products, and the soil temperatures or soil moisture levels are optimized independently. The results showed that soil moisture play more important roles than soil temperature in turbulent flux predictions. Furthermore, the model parameters such as soil parameters and vegetation parameters also play important roles in predicting turbulent fluxes at the land surface. Thus, the data assimilation strategy optimizing both model state variables and parameters should be developed. Yang et al. (2007) developed an auto-calibration data assimilation scheme to optimize both soil parameters and soil moisture by assimilating microwave brightness temperatures, and they proved both improvements in soil moisture and land surface energy budget predictions. Improvements in soil moisture predictions with the optimization of both soil moisture and parameters can also be found in Qin et al. (2009) and Tian et al. (2009). Turbulent flux predictions were improved through simultaneously optimization of soil moisture and model parameters by assimilating GOES land surface temperatures (Xu et al., 2011b). However, the model parameters such as vegetation parameters vary not so fast as model states. It is physically unreasonable to optimize both model states and model parameters at the same temporal scales. This study introduced a new technique

HESSD
9, 8493–8534, 2012

which optimized soil moisture and model parameters at different temporal scales in two passes. The first pass optimized model parameters at a long temporal scale (weekly), and the second pass optimized soil moisture in case satellite observations available. In this study, the land surface temperatures were derived from FY3A satellite launched in the year 2008 by People's Republic of China (PRC). This study did the first experiment to assimilate FY3A land surface temperatures into a land surface model. Four observation sites including different land cover types (grassland, alpine meadow, cropland, and orchard) in PRC are selected to validate the results.

In this paper, Sect. 2 introduces the methodology, including the assimilation method, and the land surface model (common land model, CoLM). Section 3 describes the experiment data, including the meteorology data from field measurements, and remotely-sensed land surface temperatures (LST) from FY3A. Section 4 tests model parameter sensitivities. Section 5 presents the results and discussions, including (i) comparisons of the simulation and assimilation results with EC data, (ii) comparisons of the simulation and assimilation results with LAS data, and (iii) retrievals of model parameters and the incorporations into CoLM. The conclusions are provided in Sect. 6.

2 Methodology

The dual-pass assimilation scheme consists of a land surface model namely common land model (CoLM) to simulate water and energy budget at the land surface by combing model states, land surface parameters and forcing data. Ensemble Kalman filter (EnKF) algorithm is selected as the data assimilation method. The dual-pass data assimilation technique is employed in this study which optimizes soil moisture and parameters independently.

Figure 1 shows the dual-pass data assimilation scheme, which includes land surface model, data assimilation algorithm, forcing and ancillary data, etc. The two passes of the scheme assimilate remotely-sensed land surface temperatures (LST) derived from FY3A satellite. In pass 1, the model parameters are optimized with the assimilation of

HESSD
9, 8493–8534, 2012

FY3A LST at a long temporal scale (weekly). During the running of land surface model, it will predict LST and compare with FY3A LST data. After running for the time about a week, the EnKF algorithm will be launched to minimize the difference between the predicted LST and FY3A LST by adjusting model parameters. Thus, the optimal model parameters are obtained and used in next time step. In pass 2, the EnKF algorithm will be launched once the FY3A LST available to minimize the difference between the predicted LST and FY3A LST by adjusting soil moisture. Thus the optimal soil moisture is also obtained. With the optimal model parameters and soil moisture, we assume the optimal turbulent fluxes are predicted.

2.1 Assimilation method

In this study, the EnKF algorithm is employed as the data assimilation method. For the easy implementation, the EnKF algorithm has been widely used to construct data assimilation systems. The EnKF is based on ensemble generations where the approximation of predicted state error covariance matrix is made by spreading an ensemble of model states using states from the previous time step. The key point in the performance of the EnKF is to generate the ensemble of model states or parameters and observations at each update time by introducing noise drawn from a distribution with zero mean and covariance equal to the model states and observation error covariance matrix. Xu et al. (2011b) have summarized the EnKF algorithm relating to remotely-sensed LST and land surface model as follows:

The model operators which can predict state variables can be described as follows:

$$X_{k+1} = M(X_k, \alpha_{k+1}, \beta_{k+1}), \quad (1)$$

where X_{k+1} and X_k represent model state variables at time $k+1$ and k, respectively; $M(\cdot)$, model operator; α_{k+1}, forcing data at time $k+1$; and β_{k+1}, model parameters at time $k+1$.

At the beginning of the algorithm, the first-guest value X_0, model parameter β_0, and background error covariance P_0 are determined on the basis of prior knowledge. The

HESSD
9, 8493–8534, 2012

initial state variables ensemble can be obtained by adding random noises to X_0 as follows:

$$X_{i,0} = X_0 + u_i \qquad u_i \sim N(0, P_0), \tag{2}$$

where μ_i is the background error vector that conforms to the Gaussian distribution with zero mean and covariance matrix of P_0. The state variables then proceed by adding i (i represents the ensemble member) number of random noises, which conform to Gaussian distribution. It is expressed as the following equation:

$$X^{\mathrm{f}}_{i,1} = M(X_{i,0}, \alpha_1, \beta_1) + w_i \qquad w_i \sim N(0, Q), \tag{3}$$

where $X^{\mathrm{f}}_{i,1}$ represents the forecasted state variables of the ith member at time 1; the superscript "f" represents the forecasted state variables; w_i is the model error vector, which conforms to Gaussian distribution with zero mean and covariance matrix **Q**; and **Q** represents the model error. The soil moisture errors are set according to Xu et al. (2011a), and the model parameter uncertainties are set to 10 % of its default value, according to Mölders (2005).

When observations are unavailable, the model state variables will proceed using the following equation:

$$X^{\mathrm{f}}_{i,k+1} = M(X^{\mathrm{f}}_{i,k}, \alpha_{k+1}, \beta_{k+1}) + w_i \qquad w_i \sim N(0, \mathbf{Q}), \tag{4}$$

where $X^{\mathrm{f}}_{i,k}$, $X^{\mathrm{f}}_{i,k+1}$ represent the forecasted state variables of the ith member at times k and $k+1$.

When observations are available, the observation operator will predict LST as given by the following equation:

$$Y_i = H(X^{\mathrm{f}}_i) + v_i \qquad v_i \sim N(0, R), \tag{5}$$

where Y_i is the surface temperature of ith member; $H(\cdot)$, the observation operator that relates model state variables to observations; and v_i, the observation error that

HESSD
9, 8493–8534, 2012

conforms to Gaussian distributions with zero mean and covariance matrix **R** (R: the observation error). Each state variable is updated as follows:

$$X_i^{\mathrm{a}} = X_i^{\mathrm{f}} + \mathbf{K}(Y^o - Y_i), \tag{6}$$

$$\mathbf{K} = \mathbf{P}^{\mathrm{f}}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}^{\mathrm{f}}\mathbf{H}^{\mathrm{T}} + R)^{-1}, \tag{7}$$

$$\mathbf{P}^{\mathrm{f}} = \frac{1}{N-1}\sum_{i=1}^{N}(X_i^{\mathrm{f}} - \bar{X}^{\mathrm{f}})(X_i^{\mathrm{f}} - \bar{X}^{\mathrm{f}})^{\mathrm{T}}, \tag{8}$$

$$\mathbf{P}^{\mathrm{f}}\mathbf{H}^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}[X_i^{\mathrm{f}} - \bar{X}^{\mathrm{f}}][\mathbf{H}(X_i^{\mathrm{f}}) - \mathbf{H}(\bar{X}^{\mathrm{f}})]^{\mathrm{T}}, \quad \text{and} \tag{9}$$

$$\mathbf{H}\mathbf{P}^{\mathrm{f}}\mathbf{H}^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}[\mathbf{H}(X_i^{\mathrm{f}}) - \mathbf{H}(\bar{X}^{\mathrm{f}})][\mathbf{H}(X_i^{\mathrm{f}}) - \mathbf{H}(\bar{X}^{\mathrm{f}})]^{\mathrm{T}}. \tag{10}$$

where X_i^{a} represents the analyzed state variables of the ith member; **K**, Kalman gain matrix; Y^0, observations; $\mathbf{P}^{\mathrm{f}}$, the forecasted background error covariance matrix; $\mathbf{H}^{\mathrm{T}}$, the transposed matrix of observation operator; N, the number of ensembles; $\bar{X}_i^{\mathrm{f}}$, the mean value of forecasted state variables; and $[\cdot]^{\mathrm{T}}$, the transposed matrix.

After the state variables are updated, the analysis error (AE) can be obtained as follows:

$$\mathrm{AE} = \frac{1}{N-1}\sum_{i=1}^{N}(X_i^{\mathrm{a}} - \overline{X^{\mathrm{a}}})(X_i^{\mathrm{a}} - \overline{X^{\mathrm{a}}})^{\mathrm{T}}, \tag{11}$$

where $\overline{X^{\mathrm{a}}}$ is the mean value of the analyzed state variables.

Since the dual-pass assimilation technique is employed, EnKF will update model states and parameters separately. Xu et al. (2011a) have proved that soil moisture play an important role in predicting turbulent flux in CoLM, and the sensitivity of model

HESSD
9, 8493–8534, 2012

parameters is tested in Sect. 4. Thus, the soil moisture and model parameters are updated separately using Eqs. (6)–(10).

2.2 Common land model

Common land model (CoLM) is a state-of-the-art model developed by many groups and validated with extensive field data sites. CoLM can combine many processes such as physical, hydrological, and biological processes together to simulate land states such as surface temperature, soil moisture, surface radiation, turbulent fluxes (Dai et al., 2001, 2003). A two-big-leaf model was built in 2004 for canopy temperature, photosynthesis, and stomatal conductance (Dai et al., 2004). The structure of CoLM has been summarized by Xu et al. (2011b).

In CoLM, turbulent fluxes are calculated by solving the soil-vegetation-atmosphere energy balance equation. In the case of a non-vegetated surface, the energy balance equation is as follows:

$$R_{n,g} - H_g - LE_g - G_g = 0, \quad (12)$$

where $R_{n,g}$ is the net radiation absorbed by the ground surface ($W m^{-2}$) and H_g, LE_g, and G_g, the sensible heat flux, latent heat flux, and soil heat flux at the soil surface, respectively ($W m^{-2}$). In the case of a vegetated surface, the energy balance equation of the canopy is as follows:

$$R_{n,c} - H_c - LE_c = 0, \quad (13)$$

where $R_{n,c}$ is the net radiation absorbed by the canopy ($W m^{-2}$), and H_c and LE_c are the sensible heat flux and latent heat flux from the leaves, respectively ($W m^{-2}$).

Turbulent fluxes from the land surface can be obtained using the following equations:

$$H = \rho_a c_p(\theta_s - \theta_a)/r_{ah}, \quad (14)$$

$$LE = \rho_a(q_s - q_a)/r_{aw}. \quad (15)$$

HESSD
9, 8493–8534, 2012

where H and LE are the sensible and latent heat fluxes from the land surface, respectively; ρ_a, the density of atmospheric air ($kg\,m^{-3}$); c_p, the specific heat of air at constant pressure ($1012\,J\,kg^{-1}\,K^{-1}$); θ_s and θ_a, the air temperature at land surface and reference height, respectively (K); r_{ah}, the aerodynamic resistance for sensible heat ($s\,m^{-1}$); q_s and q_a, the air specific humidity at land surface and reference height, respectively ($kg\,kg^{-1}$); and r_{aw}, the aerodynamic resistance for water vapor ($s\,m^{-1}$). The resistances are considered over pathways between the land surface and the reference height. The sensible heat and latent heat fluxes between the atmosphere at reference height and the canopy top (or bare ground) are derived from the Monin-Obukhov similarity theory, which is solved by an iterative numerical method.

In CoLM, the land surface temperature can be calculated using the following equation:

$$T_s = (F_u/\varepsilon\sigma)^{0.25} \tag{16}$$

where T_s is simulated land surface temperature (K); F_u, surface upward longwave radiation from CoLM ($W\,m^{-2}$); σ, the Stefan–Boltzmann constant ($5.67 \times 10^{-8}\,W\,m^{-2}\,K^{-4}$); ε, the broadband emissivity (–). This equation can be considered as the observational operator of the data assimilation system.

The input data of CoLM include ancillary data and forcing data. CoLM is designed to handle a variety of data sources. This includes land surface type, soil and vegetation parameters. Land surface types are based on the International Geosphere-Biosphere Programme (IGBP) classification system. Soil texture is sourced from a database in accordance with the percentage of sand, silt, and clay, which are available at a spatial resolution of 1 km (Shangguan et al., 2012). Leaf area index (LAI) is a key vegetation parameter of CoLM, which was derived from the MODIS LAI products, and directly incorporated into CoLM. 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 the temporal resolution of half-hour. The data includes

HESSD
9, 8493–8534, 2012

wind speed, air temperature, relative humidity, air pressure, precipitation, incoming shortwave radiation, and incoming longwave radiation. The model is run with the same time steps as the measured meteorological data. The field measurements are used for the model state variables initialization such as soil moisture and soil temperature.

3 Experiment data

3.1 Site description

Four observation sites are selected for the data assimilation experiments in the People's Republic of China (PRC) in 2010. Arou site is a grassland site, located in Qinghai province; BJ site is an alpine meadows site, located in Tibet Plateau; GT site is a cropland, located in Hebei province; MY site is located in the northern mountain area of the Beijing city, with a surface mainly covered with orchard and maize. At each site, an automatic weather station (AWS) and a multi-scale turbulent flux observation system consisting of eddy covariance (EC), large aperture scintillometer (LAS) are set up to acquire turbulent fluxes at two spatial scales simultaneously. As we all know, the AWS can provide forcing data and auxiliary data model needed. EC measured sensible and latent heat flux and LAS measured sensible heat flux are used to validate the dual-pass data assimilation scheme. In this type of observation system, the LAS system generally covers more than one FY3A pixels (1 km resolution) with an ellipsoid shape footprint. EC and AWS are located beside the center of LAS optical path. Table 1 summarizes the instruments and surface characteristics of the experiment sites.

All of the above observed data of AWS, EC, and LAS were collected in half-hour time step. The processing method of these data can be found in Liu et al. (2011). Since the EC measured turbulent fluxes surfer from an energy-imbalance problem, energy balance ratio (EBR, EBR = (sensible heat flux + latent heat flux)/(net radiation – surface soil heat flux)) is used to assess the energy-imbalance of each site. The surface soil heat flux is calculated using the method proposed by Yang and Wang. (2008).

HESSD
9, 8493–8534, 2012

The EBR is 0.79, 0.83, 0.91, and 0.80 for Arou, BJ, Guantao and Miyun sites (soil heat flux at BJ site is neglected since the soil heat flux is not obtained).

The LAS system consists of a transmitter and a receiver installed on a pair of towers approximately 500–5000 m apart, which can measure the average value of sensible heat flux along the optical path of the instrument. Generally, the source area of LAS measurements cover more than one FY3A pixels, and need to be calculated using a footprint model. The LAS footprint is calculated follows Meijininger et al. (2002):

$$f_{\mathrm{LAS}}(x, y, z_{\mathrm{eff}}) = \int_{x_2}^{x_1} W(x') f(x' - x, y' - y, z_{\mathrm{eff}}) \mathrm{d}x' \quad (17)$$

where $W(x')$ is the path-weighting function of the LAS, x_1, x_2 are the locations of the LAS transmitter and receiver; 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 LAS. Monthly LAS footprints are used in this study, determined by averaging every half-hourly footprint when the sensible heat fluxes were larger than zero, and LAS measured sensible heat flux ranging from time period of 22:00 to 06:00 LT are also excluded.

3.2 FY3A land surface temperature (LST)

FY3A satellite has been launched in the year 2008 by the Chinese government. The main task of FY3A satellite includes providing meteorology parameters for numerical weather prediction, monitoring natural disasters globally, monitoring distributions of ice, snow and ozone globally, etc. The satellite overpass time is approximately 1030 (local solar time) in its descending mode and 2230 (local solar time) in its ascending mode. FY3A satellite carries about 11 instruments, which are divided into 5 groups: imaging group, atmospheric sounding group, atmospheric composition detection group, earth radiation budget measurement group, and space environment monitor group. The

HESSD
9, 8493–8534, 2012

Earth emitted and reflected radiations from land surfaces, atmosphere, ocean, space environment can be measured by the instruments onboard the satellite using different sensor groups. These data are transmitted to ground application system in real time where the data are processed and release to all the users around the world.

LST can be retrieved by using the visible and infraRed radiometer (VIRR) instrument. The VIRR is a multi-channel instrument designed for comprehensive detection of earth environment. The VIRR instrument has 10 channels ranged from 0.58 to 12.5 μm. The 10.3 ~ 11.3 and 11.5 ~ 12.5 μm channels are infrared windows with little water vapor absorption, and they have a nominal spatial resolution of 1 km × 1 km at the nadir. With the two infrared channels, land surface temperature is obtained using a local split window method (Becker and Li, 1990). The FY3A LST products can provide LST and emissivity of each pixel, and the data are stored in the hierarchical data format (HDF), which is a sinusoidal projection with a spatial resolution of 1 km.

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

$$T_{\mathrm{sfc}} = \{[F_{\mathrm{ulr}} - (1-\varepsilon)F_{\mathrm{dlr}}]/\varepsilon\sigma\}^{0.25}, \tag{18}$$

where T_{sfc} is land surface temperature (K); F_{ulr} is surface upward longwave radiation ($\mathrm{W\,m^{-2}}$); F_{dlr} is surface downward longwave radiation ($\mathrm{W\,m^{-2}}$); σ is the Stefan–Boltzmann constant ($5.67 \times 10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}$); ε is the broadband emissivity (–), which is 0.987 for grasslands and croplands, and 0.993 for forests/orchard according to Wang et al. (2008).

The FY3A LST data are compared with ground measurements (Fig. 2). As shown in Fig. 2, the FY3A LST and ground measurements follows the same trend – the correlation coefficients (R) are 0.64, 0.77, 0.81, and 0.85 at the Arou, BJ, Guantao, and Miyun site, respectively. The FY3A LST data are higher than field measurements at Arou site and lower at the other three sites. The root mean square error (RMSE) values between

HESSD
9, 8493–8534, 2012

FY3A LST data and field measurements are approximately 1–5 K that are used as the observational errors in the dual-pass data assimilation scheme. The deviations between FY3A LST data and the ground measurements are determined by many factors. The terrain effect can affect the accuracy of FY3A LST retrievals. The mismatch of spatial and temporal scales between FY3A and field-measured LST can also cause these biases. As we all know, the remote sensing data such as FY3A LST is an instantaneous value, while the ground measurements is a mean value of about 30 min. The footprint of FY3A is approximately 1 km × 1 km, while ground measurements is an approximately dozens of square meters determined by mount level of the radiometer.

4 Model parameter sensitivity analysis

In order to determine the parameters which should be optimized afterwards, sensitivity analysis is applied to determine the non-influential factors. Extended Fourier Amplitude Sensitivity Test (EFAST) (Saltelli and Bolado, 1999), as a kind of quantitative global sensitivity analysis methods, is used here to rank model parameters and identify factors which could be considered to be optimized in CoLM. EFAST measures first-order sensitivity index and total effect index to represent the main effect contribution and the total contribution of each input factor to the variance of the outputs. EFAST is widely applied in crop models (Confalonieri et al., 2010; Drouet et al., 2011), ecological models (Crosetto and Tarantola, 2001; Miao et al., 2011) and hydrological models (Crosetto and Tarantola, 2000; Reusser et al., 2011).

Compared with local sensitivity analysis, the global sensitivity analysis has two advantages: (1) parameters are explored within the entire interval; (2) the variation of output is induced by factors globally which means the sensitivity of this factor includes the sole parameter effects as well as the interaction between parameters (Saltelli et al., 2000). For a complex non-linear model like CoLM, it's better to use global sensitivity analysis to do sensitivity analysis for model output. The computation steps of EFAST are as follows:

HESSD
9, 8493–8534, 2012

A search-curve with random term to generate random samples for every parameter (Saltelli et al., 2000)

$$x_i(s) = G_i(\sin \omega_i s) = \frac{1}{2} + \frac{1}{\pi} \arcsin(\sin(\omega_i s + \phi_i)), \tag{19}$$

where s is an independent variable between $-\pi/2$ and $\pi/2$, ω_i is the frequencies of the interested parameter, φ_i is a random phase-shift parameter taking values in $[0, 2\pi)$, so the starting point of search-curve can be anywhere within the space of parameters.

By calculating the Fourier spectrum of model output, we can get the Fourier amplitude which represents the contribution to the variance of model output. According to decomposition of variance (Sobol, 1993), the total variance of model output is

$$V(Y) = \sum_i V_i + \sum_i \sum_{j>i} V_{ij} + \ldots + V_{12\ldots k}, \tag{20}$$

Based on decomposition of variance, sensitivity measures $S_{i1,i2,\ldots}$, is given by

$$S_{i_1,i_2,\ldots,i_s} = \frac{V_{i_1,\ldots,i_s}}{V} \qquad 1 \le i_1 < \cdots < i_s \le k, \tag{21}$$

where S_i is called first-order sensitivity index which estimates the fractional contribution of x_i, S_{ij} $(i \neq j)$ is called the second-order sensitivity index which explains the variance of model output due to x_i and x_j, and total sensitivity index includes all the sensitivity indices that contain this factor. EFAST method computes S_{Ti} by dividing all the parameters into x_i and $x_{\sim i}$ as follows,

$$S_{Ti} = 1 - V_{\sim i}/V \tag{22}$$

The parameters used to calculate turbulent fluxes and their range of value is listed in Table 2. Since it is hard to define the type of probability distribution function (PDF) and different types of PDF would not lead to considerable difference in the result of

HESSD
9, 8493–8534, 2012

sensitivity analysis (Rodriguez-Camino and Avissar, 1998), the PDF of parameters is defined as uniform probability distribution. The range of value is set to 90 % ~ 110 % of default value, except for *z0m*, *displa* which are determined by observational canopy height with $z0m = 0.07h_c$ and $displa = 2/3h_c$. Guantao site is selected to conduct the sensitivity analysis from April to September, 2010.

According to sensitivity analysis of 10 parameters for sensible, latent heat fluxes and surface temperature, the five most important parameters are z0m, vmax25, hhti, gradm, and binter (Table 3). Apart from four parameters, other factors have no obvious effects on output variables so they can be set to default value. The hhti parameter means 1/2 point of high temperature inhibition function, and the value is above 300 K. The sensitivity results may be unreasonable after add about 30 K random temperature to this parameter. Thus, this parameter is not considered to be updated in the data assimilation scheme. The z0m, vmax25, gradm, and binter are selected and optimized using EnKF in this study. Among the four important model parameters, vmax25, gradm, and binter are factors related to canopy photosynthesis or stomata resistance function.

5 Results and discussions

In this section, the model simulation and data assimilation results are compared with multi-scale turbulent flux observations at the four sites. In Sect. 5.1, the FY3A LST is assimilated into CoLM with the developed dual-pass data assimilation scheme. The model outputs are compared with EC-derived sensible and latent heat fluxes. The RMSE values are selected to assess the performance of the dual-pass data assimilation scheme. At the same time, the model uncertainties are also assessed using analysis error (AE) values. In Sect. 5.2, the model outputs are compared with LAS-measured sensible heat flux. In Sect. 5.3, the model parameters are retrieved and incorporated into CoLM.

HESSD
9, 8493–8534, 2012

5.1 Comparisons of the simulation and assimilation results with EC data

Tables 4–5 and Fig. 3 show the assimilation results of the four sites (continuously 30 days at the vegetation growing season). Since some turbulent flux data from an EC/LAS system are missing in the measurement and quality control processes. The data of Julian day from 231 to 260 is selected for the comparison for the Arou site; from 175 to 204 for the BJ site; from 201 to 230 for the Guantao and Miyun sites. This section compares the model results with EC measurement.

Generally, the diurnal variations trend of land surface temperature and turbulent flux can be predicted correctly by the CoLM as shown in Fig. 3. However, the land surface temperatures and the sensible heat flux are usually overestimated and the latent heat flux is underestimated with the model. The obvious biases in surface temperature and turbulent flux modeling are corrected with the assimilation of FY3A LST data. The curves portrayed by the dual-pass data assimilation scheme are generally closer to the EC measurements than the model predictions. From Fig. 3, the model simulated latent heat flux reach the peak relative early (about 10:00 a.m.), and sometimes nearly zero in the day time, especially in the afternoon at the BJ site. With the assimilation of FY3A LST, the model can simulate the diurnal variations of latent heat flux like observations. From Yang et al. (2009), the soil surface resistance was not parameterized in CoLM that caused biases in simulation of latent heat flux at BJ site. The soil surface resistance can be negligible for a relative wet surface, and the latent heat flux is mainly controlled by solar radiation. However, the resistance can play a major role in the modeling of latent heat flux for a relative dry surface. If the resistance is not taking into consideration in a land surface model, the latent heat flux will vary frequently after the surface dry up rapidly.

Table 4 summarizes the RMSE values of the model simulation and assimilation results compared with the EC-derived turbulent fluxes at the four sites. From this table, the developed dual-pass data assimilation scheme can improve the predictions of turbulent flux and surface temperature. For sensible and latent heat fluxes, the

HESSD
9, 8493–8534, 2012

average RMSE values of the four sites drop from 81.2 to 39.6 $W m^{-2}$ and from 101.7 to 58.9 $W m^{-2}$ (the RMSE values drop 51.2 % and 42.1 %), respectively. The most significant improvements are found at Arou site, and the RMSE values drop from 108.2 to 54.9 $W m^{-2}$ and from 131.1 to 52.8 $W m^{-2}$ (the RMSE drop 49.3 % and 59.7 %) for sensible and latent heat fluxes, respectively.

The AE values are summarized in Table 5 that can describe model uncertainties. The larger AE value means the higher model uncertainties, and vice versa. The AE values can be calculated using forecast state variables for the non-assimilation case in Eq. (11). By using this dual-pass data assimilation scheme, the model uncertainties are reduced with the assimilation of FY3A LST data. The average AE values of the four site drop from 29.2 to 21.0 $W m^{-2}$ and from 47.6 to 34.1 $W m^{-2}$ (the AE values drop 28.1 % and 28.4 %) for sensible and latent heat fluxes, respectively. The most significant reductions in model uncertainties are found at Miyun site, and the AE values drop from 19.1 to 4.0 $W m^{-2}$ and from 35.0 to 7.6 $W m^{-2}$ (the AE values drop 79.1 % and 78.3 %) for sensible and latent heat fluxes, respectively.

Daily averaged evaporation fraction (EF, ration between latent heat flux and available energy) can be used to describe the splitting of available energy into sensible and latent heat flux. In CoLM, EF is important for energy balance and it describes the magnitude of available energy for evapotranspiration. Thus, the simulation and assimilation results are also assessed with EC-derived EF. In this study, EF is calculated using EF = latent heat flux/(latent heat flux + sensible heat flux), and the daily averaged EF is obtained by averaging the data from 10:00 to 15:00. Figure 4 shows the comparisons of simulation and assimilation results with EC derived EF at the four experiment sites. From Fig. 4, the CoLM usually underestimates the EF during the growing season which means the model splits relative less available energy for evapotranspiration. With the assimilation of FY3A LST data, EF is closer to observations, and the RMSE drops down.

HESSD
9, 8493–8534, 2012

5.2 Comparisons of the simulation and assimilation results with LAS data

The representative of EC is usually within 1 km, however, the model usually predict turbulent fluxes at a larger scale, the different spatial scales of them may cause problems in the validation. Furthermore, EC derived turbulent fluxes surfer from the energy-imbalance problem. The LAS instrument can measure flux at the larger scale than EC, and avoid the energy-imbalance problem. Thus, LAS measured sensible heat flux is compared with model results in this section.

Since the source area of LAS measurements can cover more than one FY3A pixel, the weight of each covered pixel should be determined for the comparison. The monthly source areas are calculated using a footprint model (Eq. 17) and overlaid with FY3A pixel at the experiment sites shown in Fig. 5. From this figure, The LAS source area did not show obvious variation, and extend from the transmitter to the receiver point with the main contributing source areas of approximately half pixel width and 2 or 3 pixel length. Obviously, the weight of each pixel covered by LAS source areas can be determined as (Jia et al., 2012):

$$H_{\text{average}} = \sum_{i=1}^{n} (w_i \times H_i) \tag{23}$$

where H_{average} is the model results with the same spatial representativeness as the LAS observation, w_i is the relative weight of each pixel, H_i is the model results of each remote sensing pixel, and n is the number of pixels within the source area.

The comparison results at the four experiment sites are shown in Fig. 6. From this figure, the sensible heat flux with the assimilation of FY3A LST is closer to LAS observations than model simulation, and the scatters with the assimilation of FY3A LST are closer to 1 : 1 line, especially at Guantao site. The larger R indicates the assimilation results have higher correlation with LAS measurements than model simulation. The RMSE values of sensible heat flux drop from 108.8 to 55.1 $\mathrm{W\,m^{-2}}$, from 91.1 to 57.8 $\mathrm{W\,m^{-2}}$, from 89.0 to 31.9 $\mathrm{W\,m^{-2}}$, and from 70.0 to 63.3 $\mathrm{W\,m^{-2}}$ (the RMSE

HESSD
9, 8493–8534, 2012

values drop 49.4 %, 36.6 %, 64.2 %, and 9.6 %) for Arou, BJ, Guantao, and Miyun site, respectively.

5.3 Retrievals of model parameters and the incorporations into CoLM

As we all know, soil moisture plays an important role in the terrestrial water cycle, and the vegetation parameters play a significant role in water and energy movement among land surface, canopy and atmosphere. Thus, they may affect the turbulent flux estimates and need to be retrieved using the dual-pass data assimilation system in addition to turbulent fluxes.

Figure 7 shows soil moisture retrievals at BJ site (Alpine meadow) from Julian day 101 to 301. As is shown in Fig. 7, the assimilation results are closer to the observations than model predictions during this time period. The RMSE values of soil moisture retrievals drop from 0.157 $m^3 m^{-3}$ to 0.145 $m^3 m^{-3}$ (drop 7.6 %) and drop from 0.041 $m^3 m^{-3}$ to 0.021 $m^3 m^{-3}$ (drop 48.8 %) for the 4 cm and 20 cm depth through the assimilation of FY3A LST data. At the 4 cm depth, the soil moisture from both model simulation and assimilation has big biases against observations. As described in Yang et al. (2009), in Tibet Plateau, the top soil contains dense grass roots which make high porosity and high water-holding capacity. This caused the phenomenon that soil moisture in 4 cm is high and led to big bias against observations.

Figure 8 shows the seasonal variations of the retrieved four vegetation parameters (z0m, vcmax25, gradm, and binter) at the Miyun site (Orchard) from Julian day 101 to 301. With the assimilation of FY3A LST, the parameter uncertainties are within a relative stable range. After a time period of assimilation with FY3A LST data, the parameters tend to be a stable value, z0m and binter increase to a high value, and vcmax25 and gradm decreases to a low value. The stable parameters are averaged at the four experiment sites, and the values are 0.28, 65.5×10^{-6}, 6.5, and 0.02 for z0m, vcmax25, gradm, and binter, respectively.

The retrieved model parameters are incorporated into CoLM and compared with model results with default values (Table 6). From this table, the CoLM with retrieved

HESSD
9, 8493–8534, 2012

model parameters can produce high accuracy surface temperature and turbulent fluxes. The average RMSE values of the four sites drop from 4.4 to 3.6 K, from 81.2 to 71.5 $W m^{-2}$, and from 101.7 to 91.5 $W m^{-2}$ (the RMSE values drop 18.4 %, 12.0 % and 10.1 %) for surface temperature, sensible and latent heat fluxes, respectively. The turbulent fluxes get obvious improvements at Guantao and Miyun site, while get little improvements at Arou and BJ site. Guantao and Miyun are crop land and orchard land, respectively, the leaf area index is relative larger than grass land (Arou and BJ site). Thus, the vegetation parameters play more important role at Guantao and Miyun site, and get clear improvements.

6 Conclusions

In this study, a dual-pass data assimilation scheme was developed to estimate the turbulent fluxes with FY3A LST data by the independent optimization of soil moisture and vegetation parameters. This scheme was build based on the EnKF algorithm and CoLM. The first pass of the data assimilation scheme optimizes model parameters at a long temporal scale, and the second pass optimizes soil moisture at a short temporal scale. The results were compared with multi-scale turbulent flux observations (derived from EC and LAS) at four observation sites in PRC. Ultimately, the retrieved vegetation parameters are incorporated into CoLM.

Through the comparisons with EC-derived sensible and latent heat flux, the assimilation curves match well with observations (Fig. 3) and the average RMSE values of the four sites drop from 81.2 to 39.6 $W m^{-2}$ and from 101.7 to 58.9 $W m^{-2}$ (the RMSE values drop 51.2 % and 42.1 %) for sensible and latent heat fluxes, respectively (Table 4). The AE values are also reduced, and the average AE values of the four site drop from 29.2 to 21.0 $W m^{-2}$ and from 47.6 to 34.1 $W m^{-2}$ (the AE values drop 28.1 % and 28.4 %) for sensible and latent heat fluxes, respectively (Table 5). This indicates the dual-pass data assimilation scheme reduced the model uncertainties. The evaporation fraction (EF) is also used to assess the performance of the dual-pass data assimilation

HESSD
9, 8493–8534, 2012

scheme. With the assimilation of FY3A, the scheme improves the underestimation of EF simulation during vegetation growing season, and improves the distribution of available energy into sensible and latent heat flux (Fig. 4). Furthermore, the results are compared with LAS-derived sensible heat flux, which can measure flux at the larger scales and avoid the energy-imbalance problems of EC data. The comparisons show that the assimilation results match well against LAS measurements than model simulations, and the correlations between assimilation results and LAS measurements are higher than that of model simulations (Fig. 6).

In addition to the estimations of turbulent fluxes, model soil moisture was retrieved using the dual-pass data assimilation scheme. The retrieved soil moisture was compared with in-situ measurements, and the assimilation curve is generally closer to observations (Fig. 7). Furthermore, the four vegetation parameters namely, z0m, vcmax25, gradm, and binter, were also retrieved. With the assimilation of FY3A LST, the retrieved parameters reach the stable values in a short time, and the uncertainties are within a relatively small range (Fig. 8). Finally, the retrieved parameters at the four sites were averaged and incorporated into CoLM, which improved the estimates of surface temperature and turbulent fluxes (Table 6). Improvements of turbulent flux estimates were clear for crop land (Guantao and Miyun site), while little improvements were found for grass land (Arou and BJ site).

To enhance the effects of the dual-pass data assimilation scheme, a few problems need to be resolved. First, there are some biases between FY3A LST and ground measurements (Fig. 2) that produce errors in FY3A LST retrievals. Second, the land surface model used in data assimilation scheme should be localized using in-situ measurements, especially in ecologically-unique areas such as the Tibet plateau. In some areas of Tibet plateau, the top soil contains dense root systems which leads to the high water capacity and therefore cannot be calculated with empirical formula of soil characteristics. Third, model parameterization is also important for turbulent flux prediction such as soil surface resistance. The missing parameterization of soil surface resistance in CoLM leads to large errors in the model turbulent flux predictions at BJ

HESSD
9, 8493–8534, 2012

site, especially for latent heat flux. To make the data assimilation scheme more effective and obtain more accurate turbulent flux predictions, a number of developments are needed: (1) FY3A land surface temperature products retrieval algorithms should be improved; (2) the land surface model should be calibrated using local measurements in some special areas, such as, Tibet plateau; and (3) studies on the key processes for water and energy exchanges in land surface model should enhanced.

The retrievals of model states and parameters using land surface information, especially the remote sensing data, attract much attention of researchers since they play important roles in the determination of land surface energy budget. Obviously, data assimilation is a new technique that can integrate model and observations together to produce more accurate and continuous land surface states. Since the high variability of land surface, to get more stable and reliable land surface states, multi-source observations need to be assimilated into land surface models in the future.

Acknowledgement. This work was funded by the National Natural Science Foundation of China (91125002), the Special Research Foundation of the Public Benefit Industry (GYHY200706046) and the Fundamental Research Funds for the Central Universities.

References

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

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

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

Boni, G., Entekhabi, D., and Castelli, F.: 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, 2001.

HESSD

9, 8493–8534, 2012

Caparrini, F., Castelli, F., and Entekhabi, D.: Estimation of surface turbulent fluxes through assimilation of radiometric surface temperature sequences, J. Hydrometeorol., 5, 145–159, 2004.
Confalonieri, R., Bellocchi, G., Bregaglio, S., Donatelli, M., and Acutis, M.: Comparison of sensitivity analysis techniques: a case study with the rice model WARM, Ecol. Model., 221, 1897–1906, 2010.
Crosetto, M. and Tarantola, S.: Uncertainty and sensitivity analysis: tools for GIS-based model implementation, Int. J. Geogr. Inf. Sci., 15, 415–437, 2001.
Dai, Y. J., Zeng, X. B., and Dickinson, R. E.: Common Land Model (CoLM): Technical Documentation and User's Guide, 69 pp., 2001.
Dai, Y. J., Zeng, X. B., Dickinson, R. E., Baker, I., Bonan, G. B., Bosilovich, M. G., Dennning, A. S., Dirmeyer, P. A., Houser, P. R., Niu, G. Y., Oleson, K. W., Schlosser, C. A., and Yang, Z. L.: The common land model, B. Am. Meteorol. Soc., 84, 1013–1023, 2003.
Dai, Y. J., Dickinson, R. E., and Wang, Y. P.: A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance, J. Climate, 17, 2281–2299, 2004.
Dickinson, R. E., Kennedy, P. J., Henderson-Sellers, A., and Wilson, M.: Biosphere–Atmosphere Transfer Scheme (BATS) version 1E as coupled to the NCAR Community Climate Model, NCAR Tech. Rep. TN-275+STR, 72 pp., 1986.
Drouet, J. L., Capian, N., Fiorelli, J. L., Blanfort, V., Capitaine, M., Duretz, S., Gabrielle, B., Martin, R., Lardy, R., Cellier, P., and Soussana, J. F.: Sensitivity analysis for models of greenhouse gas emissions at farm level. Case study of N_2O emissions simulated by the CERES-EGC model, Environ. Pollut., 159, 3156–3161, 2011.
Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics, J. Geophys. Res., 99, 10143–10162, 1994.
Huang, C. L., Li, X., and Lu, L.: Retrieving soil temperature profile by assimilating MODIS LST products with ensemble Kalman filter Remote Sens. Environ., 112, 1320–1336, 2008.
Jazwinski, A. H.: Stochastic Processes and Filtering Theory, Elsevier, New York, 376 pp., 1970.
Jia, Z., Liu, S., Xu, Z., Chen, Y., and Zhu, M.: Validation of remotely sensed evapotranspiration over the Hai River Basin, China, J. Geophys. Res., in press, 117, doi:10.1029/2011JD017037, 2012.
Kalman, R. E.: A new approach to linear filtering and prediction problems, Trans. ASME J. Basic Eng., 82, 35–45, 1960.

HESSD
9, 8493–8534, 2012

Kormann, R. and Meixner, F.: An analytical footprint model for non-neutral stratification, Bound.-Lay. Meteorol., 99, 207–224, 2001.
Liang, S.: Quantitative Remote Sensing of Land Surfaces, John Wiley and Sons, Inc., New York, 534 pp., 2004.
Liang, S. and Qin, J.: Data assimilation methods for land surface variable estimation, in: Advances in Land Remote Sensing: System, Modeling, Inversion and Applications, edited by: Liang, S., 319–339, Springer, 2008.
Liou, Y., Galantowicz, J. F., and England, A. W.: A land surface process/radiobrightness model with coupled heat and moisture transport for prairie grassland, IEEE T. Geosci. Remote., 37, 1848–1859, 1999.
Liu, S. M., Hu, G., Lu, L., and Mao, D. F.: Estimation of regional evapotranspiration by TM/ETM+ data over heterogeneous surfaces, Photogramm. Eng. Rem. S., 73, 1169–1178, 2007.
Liu, S. M., Xu, Z. W., Wang, W. Z., Jia, Z. Z., Zhu, M. J., Bai, J., and Wang, J. M.: 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, doi:10.5194/hess-15-1291-2011, 2011.
Margulis, S. A., McLaughlin, D., Entekhabi, D., and Dunne, S.: Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment, Water Resour. Res., 38, 1229, doi:10.1029/2001WR001114, 2002.
Miao, Z., Lathrop, R. G., Xu, M., La Puma, I. P., Clark, K. L., Hom, J., Skowronski, N., and Van Tuyl, S.: Simulation and sensitivity analysis of carbon storage and fluxes in the New Jersey Pinelands, Environ. Modell. Softw., 26, 1112–1122, 2011.
Meijninger, W. M. L., Hartogensis, O. K., Kohsiek, W., Hoedjes, J. C. B., Zuurbier, R. M., and De Bruin, H. A. R.: Determination of area-averaged sensible heat fluxes with a large aperture scintillometer over a heterogeneous surface – Flevoland field experiment, Bound.-Lay. Meteorol., 105, 37–62, 2002.
Mo, X., Chen, J. M., Ju, W., and Black, T. A.: Optimization of ecosystem model parameters through assimilating eddy covariance flux data with an ensemble Kalman filter, Ecol. Model., 217, 157–173, 2008.
Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P. R.: Dual state-parameter estimation of hydrological models using ensemble Kalman filter, Adv. Water Resour., 28, 135–147, 2005.

HESSD
9, 8493–8534, 2012

Mölders, N.: Plant and soil parameter caused uncertainty of predicted surface fluxes, Mon. Weather Rev., 133, 3498–3516, 2005.
Pipunic, R. C., Walker, J. P., and Western, A.: Assimilation of remotely sensed data for improved latent and sensible heat flux prediction: a comparative synthetic study, Remote Sens. Environ., 112, 1295–1305, 2008.
Qin, J., Liang, S., Yang, K., Koike, T., Liu, R., and Kaihotsu, I.: Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal, J. Geophys. Res., 114, D15103, doi:10.1029/2008JD011358, 2009
Reichle, H. R., Mclaughlin, D. B., and Entekhabi, D.: Hydrologic data assimilation with the ensemble kalman filter, Mon. Wea. Rev., 130, 103–114, 2002.
Reusser, D. E., Buytaert, W., and Zehe, E.: Temporal dynamics of model parameter sensitivity for computationally expensive models with the Fourier amplitude sensitivity test, Water Resour. Res., 47, W07551, doi:10.1029/2010WR009947, 2011
Rodriguez-Camino, E. and Avissar, R.: Comparison of three land-surface schemes with the Fourier amplitude sensitivity test (FAST), Tellus A, 50, 313–332, 1998.
Schuurmans, J. M., Troch, P. A., Veldhuizen, A. A., Bastiaanssen, W. G. M., and Bierkens, M. F. P.: Assimilation of remotely sensed latent heat flux in a distributed hydrological model, Adv. Water Resour., 26, 151–159, 2003.
Schuurmans, J. M., van Geer, F. C., and Bierkens, M. F. P.: Remotely sensed latent heat fluxes for model error diagnosis: a case study, Hydrol. Earth Syst. Sci., 15, 759–769, doi:10.5194/hess-15-759-2011, 2011.
Sellers, P. J., Randall, D. A., Collatz, G. J., Field, C. B., Dazlich, D. A., Zhang, C., Collelo, G. D., and Bounoua, L.: A revised land surface parameterization (SiB2) for atmosphere GCMs, part I: model formulation, J. Climate, 9, 676–705, 1996.
Shangguan, W., Dai, Y., Liu, B., Ye, A., and Yuan, H.: A soil particle-size distribution dataset for regional land and climate modelling in China, Geoderma, 171–172, 85–91, 2012.
Su, Z.: The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes, Hydrol. Earth Syst. Sci., 6, 85–100, doi:10.5194/hess-6-85-2002, 2002.
Saltelli, A. and Bolado, R.: An alternative way to compute Fourier amplitude sensitivity test (FAST), Comput. Stat. Data An., 26, 445–460, 1999
Saltelli, A., Chan, K., and Scott, E. M.: Sensitivity Analysis, John Wiley & Sons, Ltd., Chichester, New York, 2000.

HESSD
9, 8493–8534, 2012

Sobol, I. M.: Sensitivity estimates for nonlinear mathemetical models, Math. Model. Comput. Exp., 1, 407–414, 1993

Tian, X., Xie, Z., Dai, A., Shi, C., Jia, B., Chen, F., and Yang, K.: A dual-pass variational data assimilation framework for estimating soil moisture profiles from AMSR-E microwave brightness temperature, J. Geophys. Res., 114, D16102, doi:10.1029/2008JD011600, 2009.

Wang, W. H., Liang, S. L., and Tilden, M.: Validating MODIS land surface temperature products using long-term nighttime ground measurements, Remote Sens. Environ., 112, 623–635, 2008.

Xiao, Z. Q., Liang, S., Wang, J. D., Jiang, B., Li, X. J.: Real-time retrieval of leaf area index from MODIS time series data, Remote Sens. Environ., 115, 97–106, 2011.

Xu, T., Liu, S. M, Liang, S., and Qin, J.: Improving predictions of water and heat fluxes by assimilating MODIS land surface temperature products into common land model, J. Hydrometeorol., 12, 227–244, 2011a.

Xu, T., Liang, S., and Liu, S.: Estimating turbulent fluxes through assimilation of geostationary operational environmental satellites data using ensemble Kalman filter, J. Geophys. Res., 116, D09109, doi:10.1029/2010JD015150, 2011b.

Yang, K. and Wang, J. M.: A temperature prediction-correction method for estimating surface soil heat flux from soil temperature and moisture data, Sci. China Ser. D, 51, 721–729, 2008.

Yang, K., Takahiro, W., and Toshio, K.: An auto-calibration system to assimilate AMSR-E data into a land surface model for estimating soil moisture and surface energy budge, J. Meteor. Soc. Japan, 85, 229–242, 2007.

Yang, K., Chen, Y.-Y., and Qin, J.: Some practical notes on the land surface modeling in the Tibetan Plateau, Hydrol. Earth Syst. Sci., 13, 687–701, doi:10.5194/hess-13-687-2009, 2009.

HESSD

9, 8493–8534, 2012

Table 1. Summary of equipment and surface characteristics.

Instrument	Variable	Arou Height/depth (m)	BJ Height/depth (m)	Guantao Height/depth (m)	Miyun Height/depth (m)
EC	Sensible and latent heat flux	3.15	3.0	15.6	26.66
LAS	Sensible heat flux	9.5 (path length 2390 m)	8.6 (path length 1560 m)	15.6 (path length 2760 m)	35.86 (path length 2420 m)
AWS	Air temperature/humidity	2.0 and 10.0	8.2 and 10.0	4 and 12.5	10.66 and 30.56
	Wind speed/direction	2.0 and 10.0	1.0, 5.0 and 10.0	12.7	10.66 and 30.56
	Radiation	1.5	1.5	15.7	30.76
	Soil heat flux	0.05	–	0.02	0.02
	Soil temperature	0.1, 0.2, 0.4, 0.8, 1.2, 1.6	0, 0.04, 0.1, 0.2, 0.4	0, 0.02, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 1	0, 0.02, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 1
	Soil moisture	0.1, 0.2, 0.4, 0.8, 1.2, 1.6	0.04, 0.2	0.02, 0.05, 0.1, 0.2, 0.4, 0.6, 1	0.02, 0.05, 0.1, 0.2, 0.4, 0.6, 1
	Precipitation	–	–	–	–
	Air pressure	–	–	–	–
Location		100.91° E, 38.04° N	91.89° E, 31.37° N	115.13° E, 36.52° N	117.32° E, 40.63° N
Elevation (m)		2990	4520	30	350
Landscape		Grass	Alpine meadow	Crop	Orchard

HESSD
9, 8493–8534, 2012

Table 2. Some soil and vegetation parameters in common land model (CoLM).

Parameter	Description	Unit	Default value
z0m	Surface roughness length	m	0.1
displa	Zero plane displacement	m	0.667
effcon	Quantum efficiency at 25 °C	$\mu mol\, mol^{-1}$	0.08
vmax25	Maximum rate of carboxylation at 25 °C	$\mu mol\, m^{-2}\, s^{-1}$	0.0001
hlti	1/2 point of low temperature inhibition function	K	281.16
hhti	1/2 point of high temperature inhibition function	K	308.16
gradm	Conductance-photosynthesis slope parameter	–	9.0
binter	Conductance-photosynthesis intercept	–	0.01
d50	Coefficient of root profile	–	19.0
beta	Coefficient of root profile	–	−1.738

HESSD
9, 8493–8534, 2012

Table 3. First order and total sensitivity indices of model parameters from April to September.

Parameter	H		LE		T_s	
	S_i	S_{Ti}	S_i	S_{Ti}	S_i	S_{Ti}
z0m	7.61×10^{-1}	0.88	2.25×10^{-1}	0.43	8.83×10^{-1}	0.97
displa	2.00×10^{-3}	0.04	4.14×10^{-4}	0.01	1.30×10^{-3}	0.03
effcon	1.30×10^{-3}	0.04	2.71×10^{-4}	0.02	9.51×10^{-4}	0.03
vmax25	4.66×10^{-4}	0.04	4.30×10^{-3}	0.05	3.13×10^{-4}	0.02
hlti	8.43×10^{-4}	0.03	3.13×10^{-4}	0.01	5.12×10^{-4}	0.02
hhti	7.88×10^{-2}	0.18	4.07×10^{-1}	0.78	2.49×10^{-2}	0.08
gradm	3.50×10^{-3}	0.09	5.90×10^{-3}	0.16	1.20×10^{-3}	0.04
binter	2.20×10^{-3}	0.04	6.53×10^{-4}	0.03	5.69×10^{-4}	0.03
d50	9.68×10^{-4}	0.04	3.73×10^{-4}	0.02	1.20×10^{-3}	0.03
beta	4.74×10^{-4}	0.03	4.12×10^{-4}	0.02	5.61×10^{-4}	0.03

HESSD
9, 8493–8534, 2012

Table 4. RMSE values of simulation and assimilation results compared with EC data*.

		Arou	BJ	Guantao	Miyun	Average
T_s (K)	Sim	6.8	3.2	4.2	3.2	4.4
	Ass	2.5	3.6	2.3	2.9	2.8
H (W m^{-2})	Sim	108.2	84.7	63.1	68.8	81.2
	Ass	54.9	52.7	24.7	26.1	39.6
LE (W m^{-2})	Sim	131.1	110.5	82.8	82.4	101.7
	Ass	52.8	70.3	54.3	58.3	58.9

* Here Sim means model simulation results; Ass means data assimilation results.

HESSD
9, 8493–8534, 2012

Table 5. AE values with and without data assimilation*.

		Arou	BJ	Guantao	Miyun	Average
T_s (K)	Non_ass	1.3	2.7	1.2	1.2	1.6
	Ass	1.0	2.7	0.8	0.3	1.2
H (W m^{-2})	Non_ass	32.0	37.8	27.8	19.1	29.2
	Ass	24.8	37.5	17.6	4.0	21.0
LE (W m^{-2})	Non_ass	38.8	75.6	40.9	35.0	47.6
	Ass	30.4	75.1	23.3	7.6	34.1

* Here Non_ass means results without data assimilation; Ass means results with data assimilation.

HESSD
9, 8493–8534, 2012

Table 6. RMSE values of model results with default and retrieved parameter values*.

		Arou	BJ	Guantao	Miyun	Average
T_s (K)	Def	6.8	3.2	4.2	3.2	4.4
	Ret	4.9	3.2	2.6	3.5	3.6
H (W m^{-2})	Def	108.2	84.7	63.1	68.8	81.2
	Ret	125.9	90.0	24.7	45.3	71.5
LE (W m^{-2})	Def	131.1	110.5	82.8	82.4	101.7
	Ret	129.9	110.2	54.3	71.4	91.5

* Here Def means model results with default model parameter values; Ret means model results with the retrieved model parameter values.

HESSD
9, 8493–8534, 2012

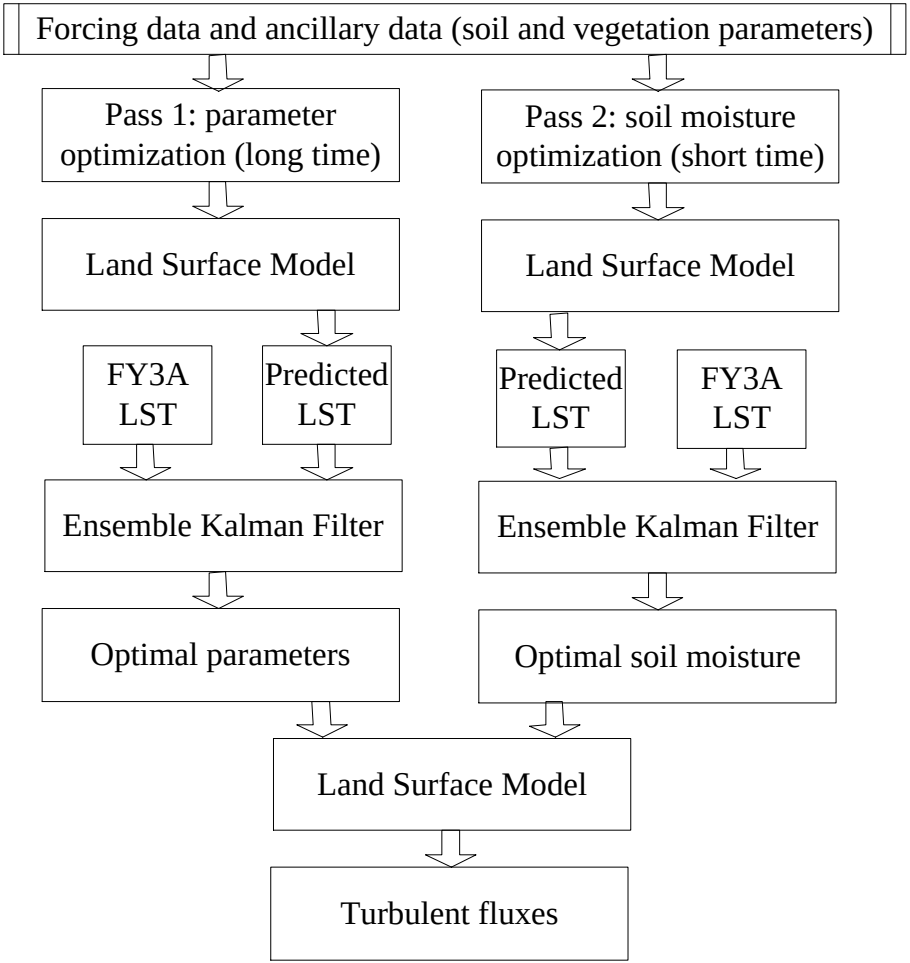

Fig. 1. The flowchart of dual-pass data assimilation scheme.

HESSD
9, 8493–8534, 2012

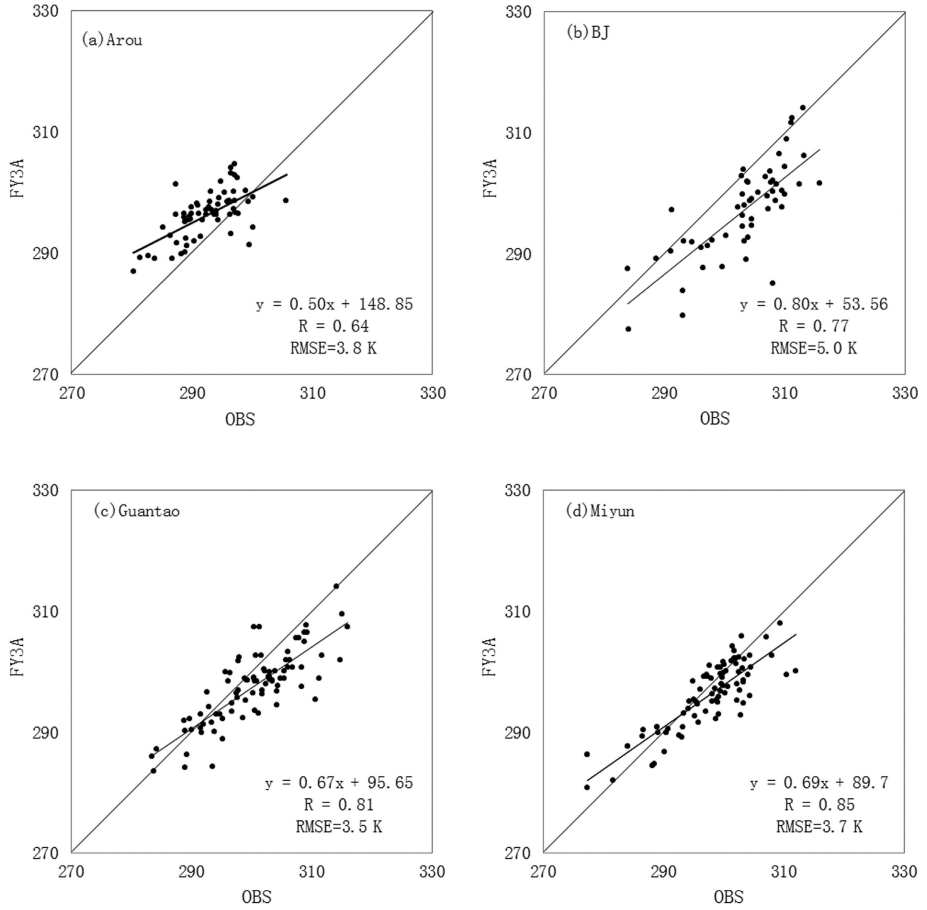

Fig. 2. Comparisons between the FY3A LST data and ground-measured surface temperatures (OBS) at four test sites (RMSE and *R*: root mean square error and correlation coefficient between FY3A LST and ground-measured surface temperatures, respectively).

HESSD
9, 8493–8534, 2012

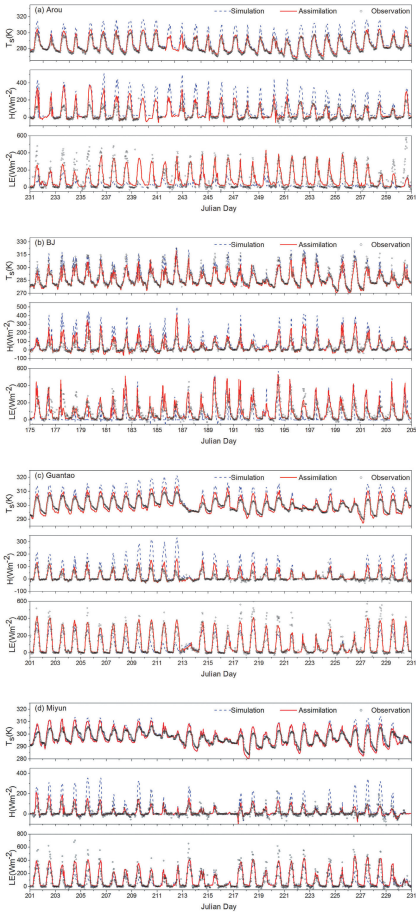

Fig. 3. Comparisons of EC observations with estimates from simulation and assimilation at the four experiment site.

HESSD
9, 8493–8534, 2012

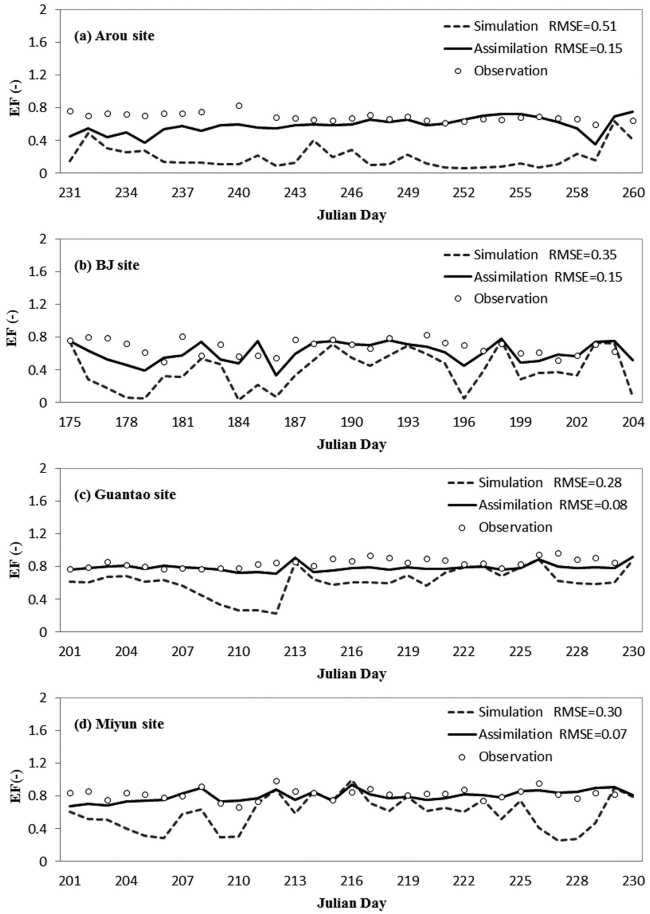

Fig. 4. Comparisons of EC system derived evaporation fraction (EF) with estimates from simulation and assimilation at the four experiment sites.

HESSD
9, 8493–8534, 2012

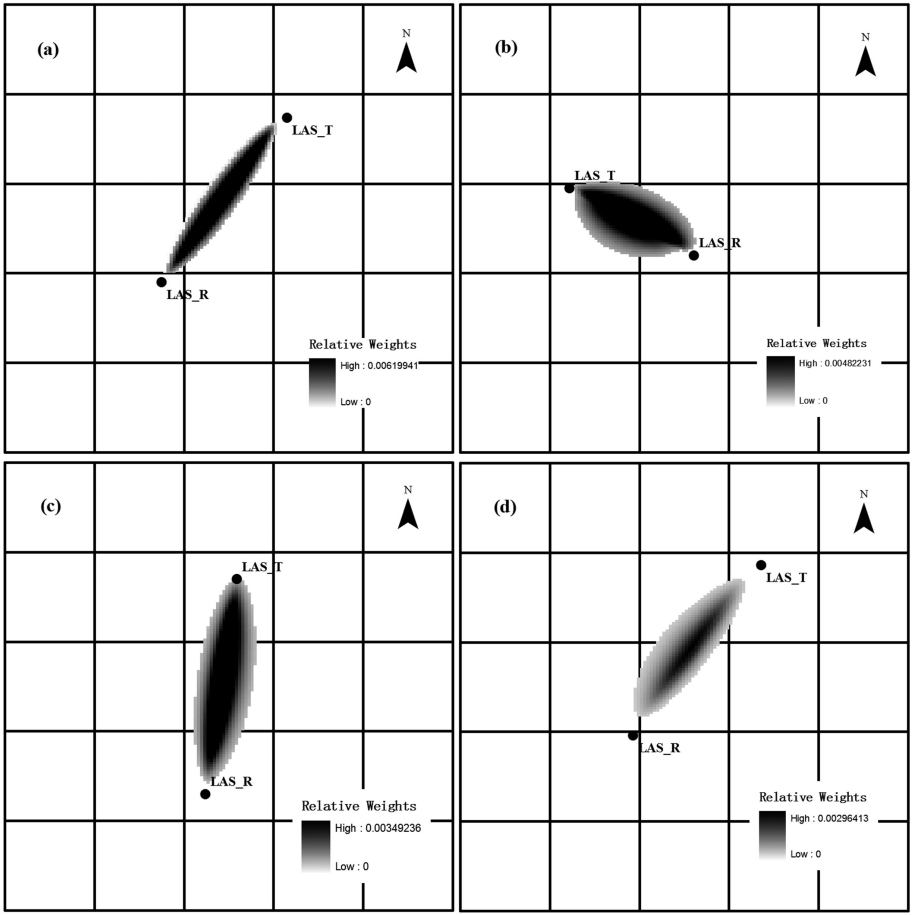

Fig. 5. The monthly source areas of LAS measurements overlaid with FY3A pixels at the experiment sites. **(a)** Arou site; **(b)** BJ site; **(c)** Guantao site; **(d)** Miyun site.

HESSD
9, 8493–8534, 2012

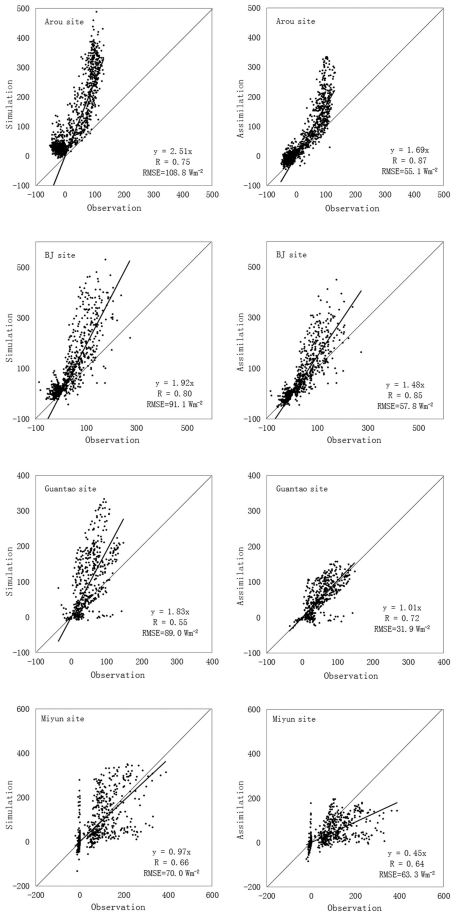

Fig. 6. Comparisons of LAS observations with estimates from simulation and assimilation at the four experiment sites.

HESSD
9, 8493–8534, 2012

A dual-pass data assimilation scheme for turbulent fluxes

T. R. Xu et al.

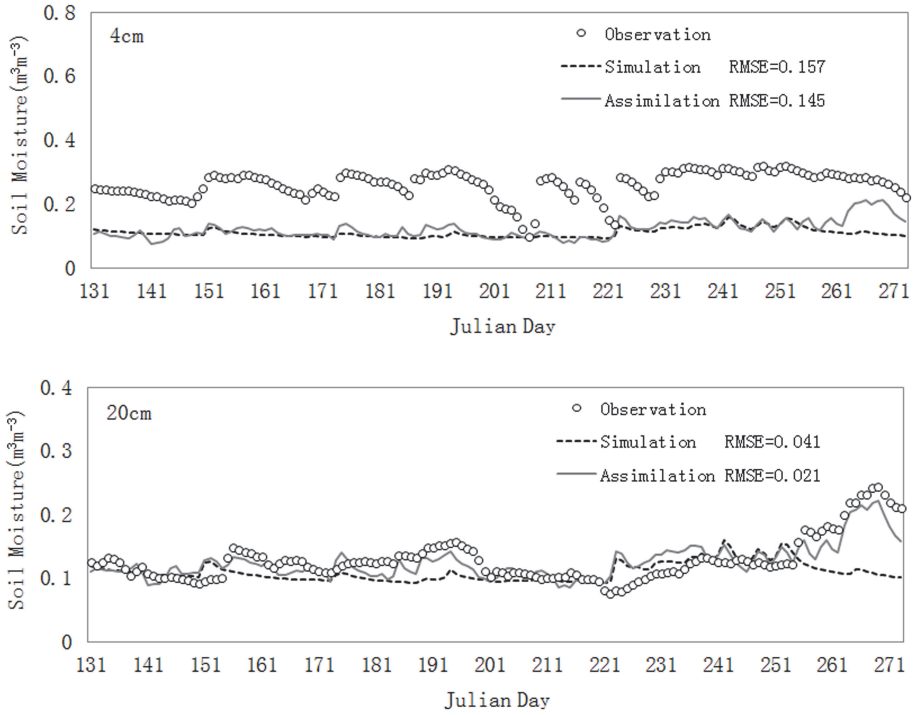

Fig. 7. Comparisons of soil moisture retrievals between simulation and assimilation at BJ site from Julian day 131 to 271, 2010. **(a)** shows the results in 4 cm depth; **(b)** shows the results in 20 cm depth.

HESSD
9, 8493–8534, 2012

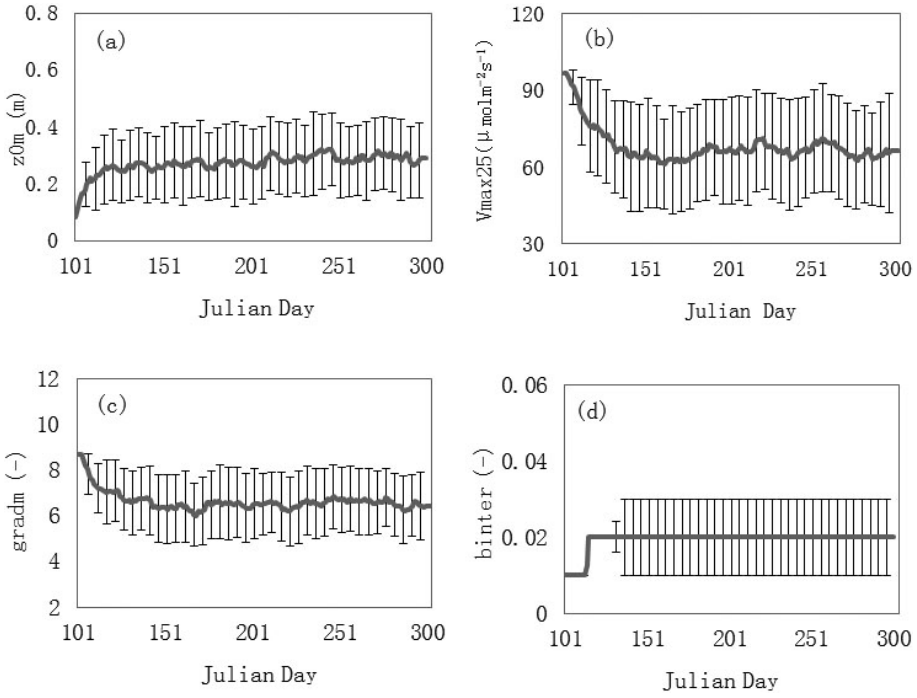

Fig. 8. Model parameter retrievals at Miyun site from Julian day 101 to 300, 2010.

HESSD
9, 8493–8534, 2012