Estimation of sensible and latent heat flux by assimilating MODIS LST products

XU Tong-ren[1], LIU Shao-min[1], QIN Jun[2], LIANG Shun-lin[3]

1. *State Key Laboratory of Remote Sensing Science, Beijing Normal University, Beijing* 100875, *China*;
2. *Institute of Tibetan Plateau Research of Chinese Academy of Sciences, Beijing* 100085, *China*;
3. *Department of Geography*, 2181 *Lefrak Hall, University of Maryland, College Park, MD*20742, *USA*

Abstract: In this paper, a land surface temperature data assimilation scheme is developed based on Ensemble Kalman Filter (EnKF) and Common Land Model version 1.0 (CLM), which is mainly used to improve the estimation of the sensible and latent heat fluxes by assimilating MODIS land surface temperature (LST) products. Leaf area index (LAI) is also updated dynamically by MODIS LAI products. In this study, the relationship between the MODIS LST and the CLM surface temperature is determined and taken as the observation operator of the assimilation scheme. Meanwhile, the MODIS LST is compared with the ground-measured surface temperature, and the Root Mean Square Error (RMSE) is taken as the observation error. The scheme is tested and validated based on measurements in three observation stations (Blackhill, Bondville and Brookings) of Ameriflux. Results indicate that data assimilation method improves the estimation of surface temperature and sensible heat flux. The RMSE of sensible heat flux reduced from 81.5W·m^{-2} to 58.4W·m^{-2} at the Blackhill site, from 47.0W·m^{-2} to 31.8W·m^{-2} at the Bondville site, from 46.5W·m^{-2} to 45.1W·m^{-2} at the Brookings site. The RMSE of latent heat fluxes reduced from 88.6W·m^{-2} to 57.7W·m^{-2} at the Bondville site, from 53.4W·m^{-2} to 47.2W·m^{-2} at the Blackhill site. In addition, it is a practical way to improve the estimation of sensible and latent heat flux by assimilating MODIS LST into land surface model.

Key words: MODIS land surface temperature products, common land model, ensemble kalman filter, sensible and latent heat flux

CLC number: TP701 **Document code:** A

1 INTRODUCTION

Estimating sensible and latent heat flux accurately is critical to climate change research, water resources planning and management, as well as water-saving agriculture applications. There are lots of methods to estimate sensible and latent heat flux. Meteorology and climatology methods are relatively accurate to in-situ estimation, but difficult to obtain regional scale fluxes. Although hydrological method can calculate evapotranspiration at regional scale, it only fits to a long time period (typically an annual cycle). The remote sensing method can acquire land surface information at a large scale in real-time, which has created new opportunities to monitor land surface sensible and latent heat flux (Bastiaanssen *et al.*, 1998; Su, 2002; Liu *et al.*, 2007). But as the remote sensing data is instantaneous, remote sensing method lacks the ability in describing surface energy and water process, and estimating daily, monthly and annual flux values. Over the past 20 years, with the developments of land surface models, it is able to simulate a continuous regional energy and water process of the soil - vegetation - atmosphere (Dickinson *et al.*, 1986; Sellers *et al.*, 1996; Dai *et al.*, 2001, 2003), which has great significance to understand the regional energy and water balance. But land surface models are generally complex, and always require a large number of regional vegetation, soil and meteorological parameters, which are difficult to determine. The above limitations facilitate development and application of land surface models.

The remote sensing method can acquire land surface information accurately at regional scale in real-time, while the land surface models can simulate a continuous regional energy and water process. The data assimilation method can combine the advantages of both these two techniques, which can produce high resolution soil temperature and moisture accurately by combining observations and land surface models. Recently, land surface data assimilation methods have experienced a great resurgence of interest (Li *et al.*, 2004; Yan *et al.*, 2006; Liang *et al.*, 2008). It has become a major tool to estimate soil temperature, soil moisture and surface fluxes. In the soil moisture data assimilation area, researchers assimilated in-situ soil moisture data (Zhang & Qiu, 2004; Huang & Li, 2006) and microwave data (Galantowiez *et al.*, 1999; Reichle *et al.*, 2001, 2002; Crosson, 2002; Crow & Wood, 2003; Yang *et al.*, 2007; Huang

Received: 2008-06-19; **Accepted**: 2009-01-07

Foundation: Hi-tech Research and Development Program of China (No. 2007AA12Z175), National Natural Science Foundation of China (No. 40671128), and the National Basic Research Program of China (No. 2007CB714401).

First author biography: XU Tong-ren (1982—), male, PhD candidate, Beijing Normal University, Mainly engaged in land surface data assimilation research.

Corresponding author: LIU Shao-min, E-mail: smliu@bnu.edu.cn

et al., 2008a) into land surface model.

Surface temperature is a key variable used to estimate sensible and latent heat flux. Boni *et al.* (2001) assimilated land surface temperature into a relatively simple model using the variational method, and acquire surface energy balance components and surface evapotranspiration. Schuurmans *et al.* (2003) assimilated remote sensing retrieved latent heat flux. Kumar and Kaleita (2003) found that there is a good correlation between land surface temperature retrieved from AVHRR and soil temperature measured at 5cm depth. He took the regression error as the observation error, assimilated the AVHRR retrieved land surface temperature into a land surface model with extended kalman filter method and estimated soil temperature profiles. Caparrini *et al*. (2004) combined force-restore equations with variational method, and estimated sensible and latent heat flux by assimilating remote sensing retrieved surface temperature. Qin *et al*. (2005), Huang *et al*. (2008b) took the regression equation of MODIS land surface temperature products and in-situ observation surface temperature as the observation operator, and took the RMSE (root mean square error) of the regression equation as observation error. They combined CLM model with ensemble kalman filter algorithm, and estimated soil temperature profiles by assimilating MODIS temperature products.

At the same time, the optimization algorithms suitable for land surface data assimilation were developed rapidly. Kalman filter method was first proposed by Kalman (1960). The standard kalman filter is an optimal data assimilation method for linear dynamic systems. As most systems are nonlinear, ensemble kalman filter was developed. Ensemble kalman filter was proposed by Evensen (1994, 2003) according to epstein's theory of stochastic dynamic prediction. It calculates the state prediction error covariance with the Monte-Carlo method. Ensemble kalman filter overcomes the shortcomings of standard kalman filter which need linear model operator and observation operator. In ensemble kalman filter, determination of model error is essential to the prediction. Some researches made a series of synthetic experiments to analyze the influences of model error (type, size) based on the assimilation results (Evensen, 2003; Crow, 2006). Reichle *et al*. (2002) conducted assimilation research by using time-related model error. Crosson *et al*. (2002) carried out assimilation research by taking a fixed value as model error. Qin *et al*. (2005) performed some research with model error related to the size of state variables.

The main purpose of this paper is to develop a land data assimilation algorithm which combines the CLM and the EnKF. Decomposition of component temperature method is used to establish observation operator. MODIS LST products will be assimilated into the assimilation algorithm to improve the prediction accuracy of sensible and latent heat flux. Meanwhile, three observation stations in different land cover types (forest, grassland, and cropland) of Ameriflux are selected to validate the assimilation results.

2 DATA ASSIMILATION ALGORITHM

Land surface models can not simulate surface temperature and fluxes accurately. Date assimilation method is able to improve estimations of surface temperature and fluxes by assimilating remote sensing radiometric temperature, as surface temperature is the input variables of land surface model. Fig. 1 is a flow chart of the land data assimilation system, and specific flow is as follows:

(1) Obtain meteorological data based on ground measurements;

(2) Obtain surface parameter sets based on remote sensing technology and ground measurements;

(3) Read forcing data and surface parameters into land surface model, and generate current state variables;

(4) Assimilate current MODIS LST products, and optimize state variables;

(5) Run land surface model continually to generate background of next step;

(6) Obtain optimal output through adjusting state variables continually.

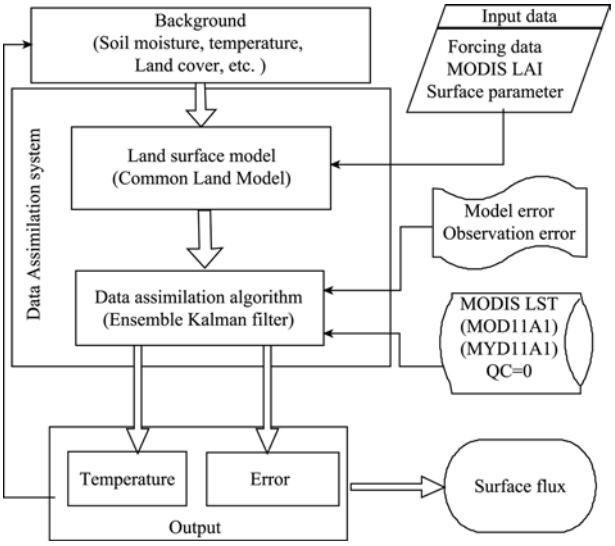

Fig. 1 The flowchart of data assimilation system

2.1 Land surface model

In this paper, CLM1.0 (Dai *et al.*, 2001, 2003) is selected as the experiment land surface model. CLM is constituted mainly by physical, hydrological, biological process, which can simulate surface temperature, soil temperature, soil moisture, sensible heat flux, latent heat flux, and so on.

Surface temperature calculation is the core part of this assimilation algorithm. Surface temperature can be obtained by solving the soil-atmosphere energy balance equation iteratively. In soil-atmosphere energy balance equation, solar radiation is the major energy source of land surface and atmospheric temperatures. For bare land, energy balance equation can be expressed as follows:

$$-S_g + H_g(T_g) + \lambda E_g(T_g) + G_g(T_g) = 0 \quad (1)$$

where, S_g is net radiation of soil; H_g, λE_g and G_g are soil sensible heat flux, soil latent heat flux, and ground heat flux respectively; T_g is ground temperature.

As land surface covered by vegetation, the energy balance equation can be described as follows:

$$-S_v + H_v(T_g, T_v) + \lambda E_v(T_g, T_v) = 0 \quad (2)$$

$$-S_g + H_g(T_g, T_v) + \lambda E_g(T_g, T_v) + G_g(T_g, T_v) = 0 \quad (3)$$

where, S_v is net radiation of vegetation canopy; H_v and λE_v are canopy sensible heat flux, canopy latent heat flux respectively; T_v is canopy temperature. When snow exists, heat flux for snow melting must be added to Eq. (1) and Eq. (3).

As can be seen from above equations, net radiation, sensible heat flux, latent heat flux, and ground heat flux are all influenced by surface temperature. Surface fluxes can be improved by improving CLM surface temperature. In order to assimilate surface temperature, we must find a starting point. Relatively speaking, the heat diffusion equation is more suitable for surface temperature assimilation. Vertical soil column of CLM is divided into 10 layers which are shown in Fig.2.

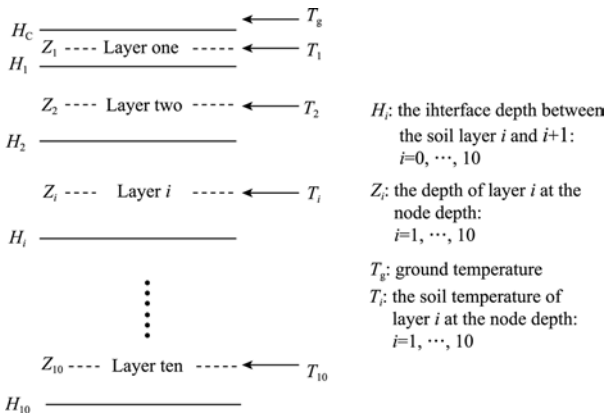

Fig. 2 Schematic diagram of soil layers in CLM

In CLM, dynamics of soil temperature are assumed to obey the following heat diffusion equation:

$$c\frac{\partial T}{\partial t} = -\frac{\partial}{\partial z}\left(k\frac{\partial T}{\partial z}\right) \quad (4)$$

where, c is volumetric soil heat capacity, T is soil temperature, t is time, k is thermal conductivity of soil, and z is soil depth. The depth of 10 soil layers is 0.00710, 0.02793, 0.06226, 0.11887, 0.21219, 0.36607, 0.61976, 1.03803, 1.72764, 2.86461m, respectively. Temperature variation in soil layer can be described as:

$$c_j \Delta z_j \frac{T_j^{n+1} - T_j^n}{\Delta t} = \frac{1}{2}(F_j^n - F_{j-1}^n + F_j^{n+1} - F_{j-1}^{n+1}) \quad (5)$$

where, c_j is volumetric soil heat capacity of layer j, T_j is mean soil temperature of layer j, Δt is time step, Δz_j is soil thickness of layer j, and F_j is heat flux across the interface between layer j and j+1.

In inner soil layers,

$$F_j = -\lambda\left[Z_{h,j}\right]\left(\frac{T_j - T_{j+1}}{Z_{j+1} - Z_j}\right) \quad (6)$$

where, $\lambda[Z_{h,j}]$ is thermal conductivity at the interface $Z_{h,j}$, $Z_{h,j}$ is interface soil depth of soil layer j and j+1, Z_{j+1} is soil depth of layer j+1. Assuming that heat flux from node j to interface between j and j+1 equals the heat from interface to node j+1, the following equation can be obtained:

$$\lambda_j \frac{T_m - T_j}{Z_{h,j} - Z_j} = \lambda_{j+1}\frac{T_{j+1} - T_m}{Z_{j+1} - Z_{h,j}} = \lambda(Z_{h,j})\frac{T_{j+1} - T_j}{Z_{j+1} - Z_j} \quad (7)$$

where, λ_j is thermal conductivity of layer j, and T_m is soil temperature of depth $Z_{h,j}$. Eq. (8) can be obtained by solving Eq. (7):

$$\lambda(Z_{h,j}) = \frac{\lambda_j \lambda_{j+1}[Z_{j+1} - Z_j]}{\lambda_j[Z_{j+1} - Z_{h,j}] + \lambda_{j+1}[Z_{h,j} - Z_j]} \quad (8)$$

Assuming that soil heat flux of soil bottom is zero, and soil temperature of all layers can be obtained by solving above equations.

Our data assimilation scheme only care for soil temperature of CLM, so CLM can be considered as "black-box" system as follows:

$$X_{k+1} = M(X_k, \theta_{k+1}, \beta_{k+1}) \quad (9)$$

where, X_{k+1}, X_k represent every soil layer temperature at the time k+1 and k respectively; $M(\cdot)$ represents land surface model, θ_{k+1} represents forcing data at time k+1, β_{k+1} represents surface parameters at time k+1.

From Eq. (9), we know that surface temperature at the time k+1 can be improved by adjusting surface temperature at the time k, then sensible and latent heat flux at the time k+1 can also be improved accordingly.

2.2 Ensemble Kalman filter

Ensemble Kalman filter was proposed by Evensen (1994, 2003). This algorithm can handle non-linear problems and is easy to implement. It has been widely applied in the field of data assimilation recently. Here a brief algorithm summary is provided to understand the implementation of Ensemble Kalman filter.

2.2.1 Model initialization

X_0 is defined as the initial value of model,

$$X_{i,0}^a = X_0^a + u_i \quad u_i \sim N(0, P) \quad (10)$$

where, $X_{i,0}^a$ is the analyzed state variable of the i^{th} member(i represents the ensemble number) at the initial time, the superscripts 'a' refer to state variables of analysis, and u_i is the background error vector, which conforms to Gaussian distribution with zero mean and covariance matrix P.

2.2.2 Model state update

The model state update equation can be expressed as follows:

$$X_{i,k+1}^f = M(X_{i,k}^a, \theta_{k+1}, \beta_{k+1}) + w_i \qquad w_i \sim N(0, Q) \quad (11)$$

where, $X_{i,k+1}^f$ represents the forecasted state variables of the

i^{th} member (i represents the ensemble member) at the time k+1, superscript 'a' and 'f' means state variables of analysis and forecasted, $X_{i,k}^{a}$ represents the analyzed state variables of i^{th} member at the time k and w_i is model error vector which conforms to Gaussian distribution with zero mean and covariance matrix Q.

2.2.3 Observation operator

The observation operator can be expressed as follows

$$Y_{i,k+1} = H(X_{i,k+1}^{f}) + v_i \quad v_i \sim N(0,R) \tag{12}$$

where, $Y_{i,k+1}$ is simulated observation of i^{th} member at the time k+1, $H(\cdot)$ represents the observation operator related model state variables to observations, v_i is observation error vector which conforms to Gaussian distribution with zero mean and covariance matrix R, R is observation error which can be obtained by comparing MODIS LST products with ground measurement.

When the data assimilation system confronted with observations, each number of state variables will be updated as follows:

$$X_{i,k+1}^{a} = X_{i,k+1}^{f} + K_{k+1}(Y_{k+1}^{o} - H(X_{i,k+1}^{f}) + v_i) \tag{13}$$

$$K_{k+1} = P_{k+1}^{f} H^{T} (HP_{k+1}^{f} H^{T} + R)^{-1} \tag{14}$$

$$P_{k+1}^{f} = \frac{1}{N-1}\sum_{i=1}^{N}(X_{i,k+1}^{f} - \bar{X}_{k+1}^{f})(X_{i,k+1}^{f} - \bar{X}_{k+1}^{f})^{\mathrm{T}} \tag{15}$$

$$P_{k+1}^{f} H^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}[X_{i,k+1}^{f} - \bar{X}_{k+1}^{f}][H(X_{i,k+1}^{f}) - H(\bar{X}_{k+1}^{f})]^{\mathrm{T}} \tag{16}$$

$$HP_{k+1}^{f} H^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}[H(X_{i,k+1}^{f}) - H(\bar{X}_{k+1}^{f})] \times [H(X_{i,k+1}^{f}) - H(\bar{X}_{k+1}^{f})]^{\mathrm{T}} \tag{17}$$

where, $X_{i,k+1}^{a}$ represents the analyzed state variable of i^{th} member at the time k+1 (it refers to surface temperature in this paper), K_{k+1} is kalman gain matrix at the time k+1, Y_{k+1}^{o} is observation at the time k+1 (it refers to MODIS LST products in this paper), P_{k+1}^{f} is the forecasted background error covariance matrix at the time k+1, $\bar{X}_{i,k+1}^{f}$ is mean value of forecasted state variable at the time k+1 (it refers to mean value of surface temperature in this paper), $[\cdot]^{\mathrm{T}}$ represents transposed matrix, $H(X_{i,k+1}^{f})$ is state variable of i^{th} member at the time k+1 (it refers to surface temperature in this paper), $H(\bar{X}_{k+1}^{f})$ is the mean value of simulated state variable at the time k+1 (it refers to mean value of surface temperature in this paper).

3 DATA COLLECTION AND DATA PROCESSING

Wang *et al*. (2008) validated the MODIS LST products at some sites of Ameriflux. Three sites (Bondville, Brookings and Blackhill) were selected to conduct the assimilation experiments following Wang's analysis (2008). Table 1 summarizes the test sites which were covered with crop, grass and forest respectively. Surface of the test sites is smoothness and evenness, in which Bondville site is validation site of MODIS data products. MODIS satellite data products and ground-measured meteorological data are available in the test sites. Ground-measured meteorological and flux data are measured every half an hour.

In this paper, the spatial resolution of MODIS LST products is 1km. Meteorological and flux data are ground measurements. The meteorological and flux data can represent the range of several kilometers due to the smoothness and evenness of these sites (Mathias, 2004).

3.1 MODIS data

In this paper, standard MODIS data products are selected (Table 2). MODIS products are stored in format of HDF, sinusoidal projection. In addition, MODIS products also provide satellite view time and quality control flag data. MODIS products can be obtained through the internet (https://wist.echo.nasa.gov/api/).

In this paper, only high quality MODIS LST and LAI products (QC=0) were selected to conduct the experiments. The low quality LAI data (QC >0) was filled with high quality data (QC=0) in previous years. Table 3 shows the MODIS LAI data at three test sites from January to December.

Table 1 Three observation sites from AmeriFlux

Site name	Latitude/(°N)	Longitude/(°W)	Land cover	Soil type	LAI	Canopy height/m	Time	Tower height/m
Bondville	40.01	88.29	Crop	Silt Loam	0—5.0	0.9	2005—2006	10.0
Brookings	44.35	96.84	Grass	Sandy Clay	0.2—3.0	0—0.4	2004—2005	4.0
Blackhill	44.15	103.64	Conifer	Clay	2.0—5.0	13—15	2004—2005	24.0

Table 2 MODIS standard data products

MODIS standard products	parameters	Spatial resolution/m	temporal resolution/d
MYD11A1	Land surface temperature	1000	1
MOD11A1	Land surface temperature	1000	1
MOD15A2	Leaf area index	1000	8

Table 3 MODIS standard LAI products

Site	Year	1	2	3	4	5	6	7	8	9	10	11	12
Bondville	2006	0.4	0.3	0.3	0.5	0.5	2.3	4.9	2.4	0.9	0.5	0.2	0.3
Brookings	2005	0.1	0.2	0.1	0.9	1.4	1.6	1.6	2.3	1.5	0.9	0.4	0.1
Blackhill	2005	2.2	1.8	1.8	1.9	3.8	3.6	5.1	4.8	3.8	1.9	2.1	2.2

3.2 Meteorology data

Meteorological forcing data is available through the Internet www.fluxnet.ornl.gov/fluxnet/index.cfm. In this paper, ground-measured meteorological data of three sites was selected as CLM forcing data, including: air temperature, air humidity, atmospheric pressure, wind speed, precipitation, incoming shortwave radiation, and incoming longwave radiation. In addition, outgoing shortwave radiation, outgoing longwave radiation, net radiation, soil heat flux, surface infrared temperature, soil temperature (observation depth of 2, 4, 8, 16, 32, 64, 128cm), and soil moisture (observation depth of 10, 20, 30, 40, 60, 100cm) were also used in the experiment.

3.3 Flux data

Flux data mainly includes sensible heat flux and latent heat flux. A series of preliminary treatment were carried out to Ameriflux eddy covariance (EC) data, which included: reject the data when sensor is malfunction, coordinating rotation, and WPL-correction, which guaranteed the quality of observational data. In addition the processing steps above, the flux data were screened rigorously, including: reject the data during periods (when and before or after 1h) with precipitation, and reject the data when friction velocity is lower than 0.1m/s.

Generally speaking, EC system has “energy imbalance” phenomenon (Wilson *et al.*, 2002), that is, energy closure is less than 1.0. Table 4 shows the energy closure calculated at the experiment sites in all-day and day. The day results were calculated with data from 8:00 to 17:00 in a day, and the all-day results were calculated with the whole day data. We know that energy closure is lower than 1.0 at Brookings, Bondville and Blackhill sites.

Table 4 Energy closure at three observation sites

Site	Bondville		Brookings		Blackhill	
	All-day	Day	All-day	Day	All-day	Day
Sample	9956	5638	8454	4828	11543	5269
	0.76	0.76	0.91	0.85	0.91	0.86

4 RESULTS

4.1 Quantifying model error

Forecast error mainly depends on the size of model error and observation error. However, model errors are often determined arbitrarily with the assumption that their influence will diminish quickly. But this assumption is mainly dependent on observations with high frequency and small errors; otherwise, forecast error will be governed by model errors.

In this paper, model errors are specified as static value, and not correlated with other state variables. With statistical method, a look up table was made to specify the model error.

4.2 Observation operator and observation error

Gao (2003) compared surface temperature from CLM with remote sensing radiometric temperature. Remote sensing radiometric temperature is obtained through Plank’s law. For bare soil land surface, surface temperature is ground temperature. For dense-vegetated land surface, surface temperature is canopy temperature. For sparse-vegetated land surface, surface temperature is a hybrid temperature of canopy and ground temperature. Considering the spatial resolution of remote sensing pixels, remote sensing radiometric temperature is hybrid temperature with different land covers. Surface temperature in CLM can be expressed as:

$$T^4 = f_{\mathrm{veg}}(T_{\mathrm{v}})^4 + (1 - f_{\mathrm{veg}})(T_{\mathrm{g}})^4 \tag{18}$$

where, T is CLM simulated surface temperature, T_{v} is canopy temperature, T_{g} is ground temperature, f_{veg} is fractional vegetation cover.

It is very important to establish observation operator successfully for a data assimilation system. In this paper, observation operator needs to be established by relating MODIS LST products to CLM state variables. Many researchers establish observation operator through regressing remote sensing radiometric temperature with ground-measured surface temperature (Kumar *et al.*, 2003; Qin *et al.*, 2005; Huang *et al.*, 2008b), which achieved good results in local area, but can not be applied at the regional scale. In this paper, component temperature decomposition method was selected to establish the observation operator (Anderson *et al.*, 2005):

$$T_{\mathrm{RAD}}^4(\theta) \approx f(\theta)\, T_{\mathrm{c}}^4 + (1 - f(\theta)) T_{\mathrm{s}}^4 \tag{19}$$

where, T_{RAD} is surface radiometric temperature, T_{c} is canopy radiometric temperature, T_{s} is ground radiometric temperature, f is fractional vegetation cover, θ is view zenith angle of sensors. The fractional vegetation cover can be expressed as follows (Anderson *et al.*, 2005):

$$f(\theta) = 1 - \exp(-0.5F / \cos(\theta)) \tag{20}$$

where, F is leaf area index which can be obtained from MODIS LAI products. Surface temperature is comprised by canopy and ground temperature, so Eq. (19) can be rewritten as:

$$T_{\mathrm{LST}}(\theta) = [(f(\theta)\varepsilon_{\mathrm{C}} T_{\mathrm{Creal}}{}^4 + (1 - f(\theta))\varepsilon_{\mathrm{S}} T_{\mathrm{Sreal}}{}^4) / \varepsilon_{\mathrm{CS}}]^{1/4} \tag{21}$$

where, T_{LST} is surface temperature, T_{Creal} is canopy temperature,

T_{Sreal} is ground temperature, ε_{cs} is site emissivity, the value referred to Wang (2008), and ε_c, ε_s is canopy and ground emissivity, the value is 0.985 and 0.96 respectively (Coll *et al.*, 2005; Jin *et al*., 2006).

In this paper, Eq. (21) was defined as observation operator of the assimilation system. MODIS LST products were regressed with ground-measured surface temperature, and the RMSE was taken as observation error. This observation operator can be applied to both local and regional assimilation test.

In order to validate MODIS LST products, ground-measured surface temperatures require a correction for emissivity effect as they are brightness temperature in nature (Huang *et al.*, 2005). The emissivity is 0.987 at Brookings and Bondville, 0.993 at Blackhill (Wang & Liang, 2008). MODIS LST products were compared with ground-measured surface temperatures (Fig. 3).

As can be seen from Fig. 3, the correlation coefficient (R^2) between MODIS LST and ground-measured surface temperature are 0.96, 0.944, 0.968 at daytime, while 0.955, 0.851, and 0.908 at night in Brookings, Bondville and Blackhill sites. In addition, MODIS LST is lower than ground-measured surface temperature, which is particularly evident when temperature is higher than 290K at Brookings site. The accuracy of MODIS is lower at daytime (RMSE=2.58) than at nighttime (RMSE=2.11) in Brookings site. The accuracy of MODIS is higher at daytime (RMSE=2.34 at Bondville, RMSE=1.86 at Blackhill) than at nighttime (RMSE=3.01 at Bondville, RMSE=3.25 at Blackhill) in Bondville and Blackhill.

However, the deviation between MODIS LST and ground-measured surface temperature is caused not only accuracy of MODIS LST, but also difference of temporal scales (MODIS surface temperature is a instantaneous value, and ground-measured temperature is the mean value of 30 min) and spatial scales (MODIS LST is mean temperature of 1km×1km area, and ground-measured temperature is about tens of square meters).

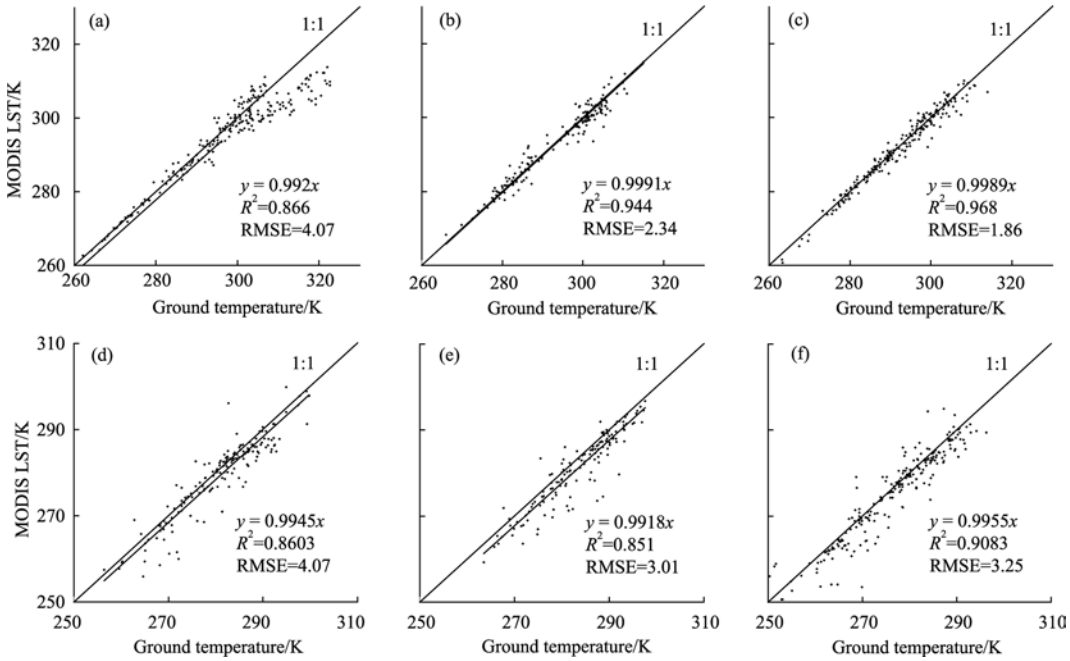

Fig. 3 Comparisons between the MODIS LST products and ground temperatures measured at Bondville, Brookings and Blackhill
(a) Brookings (daytime); (b) Bondville (daytime); (c) Blackhill (daytime); (d) Brookings (nighttime); (e) Bondville (nighttime); (f) Blackhill (nighttime)

4.3 Influence of LAI on surface temperature and flux

Surface solar radiation absorption and reflection are influenced by leaf area index (LAI). In CLM, LAI is an important surface parameter. It is determined by empirical equation to describe vegetation seasonal variation, which cannot describe the vegetation growing exactly. Some sensitivity tests of LAI will be conducted in this section.

The sensitivity tests were conducted at Bondville site in 2006. The model output (surface temperature, sensible and latent heat flux, when LAI = 2.0) were selected as the reference value. The increment of 1.5 was defined for the variation of LAI, and the influence of LAI variation to the model output was discussed according to the following situation: (1) LAI=0.5; (2) LAI=3.5; (3) LAI=5.0.

Table 5 summarizes the RMSE of surface temperature and flux comparing to the reference value.

Table 5 Sensitivity analysis of LAI

State variables		LAI=0.5	LAI=3.5	LAI=5.0
Surface temperature/K	RMSE	2.72	1.71	2.73
Sensible heat flux/(W·m^{-2})	RMSE	7.39	5.58	9.45
Latent heat flux/(W·m^{-2})	RMSE	12.53	10.71	18.73

($\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2}$, N is sample number, P_i is the calculation value, O_i is reference value)

Table 6 indicates that the RMSE of state variables (surface temperature, sensible and latent heat flux) when LAI = 0.5 is larger than LAI = 3.5, indicating that CLM is more sensitive when LAI is smaller. The RMSE of state variables (surface temperature, sensible and latent heat flux) become larger while LAI ranges from 3.5 to 5.0. As LAI increased, the increment of RMSE becomes smaller.

LAI can influence the simulation of surface temperature, sensible and latent heat flux in CLM. In order to describe the vegetation variations accurately, LAI in CLM is updated dynamically by MODIS LAI products.

4.4 Assimilation of MODIS LST results

The assimilation results were shown in Table 6.

Table7 indicates that the estimation of sensible heat flux has obvious improvement in different land cover types. The most obvious improvement takes place at Blackhill site (forest land) with RMSE decrease from 81.5W·m^{-2} to 58.4W·m^{-2}. RMSE at Bondville (crop land) decrease from 47.0W·m^{-2} to 31.8W·m^{-2}. RMSE at Brookings (grass land) decrease from 46.5W·m^{-2} to 45.1W·m^{-2}. The estimation of latent heat flux at Bondville, Blackhill has a slight improvement. RMSE at Bondville decreased from 88.6W·m^{-2} to 57.7W·m^{-2}. RMSE at Blackhill decreased from 53.4W·m^{-2} to 47.2W·m^{-2}.

Fig. 4—Fig. 6 shows the simulation and assimilation results at Brookings, Bondville and Blackhill sites. Taking into account that summer is a season with flourishing vegetation, and also taking into account the continuity of flux observation data, Julian day 181-210 is selected at Brookings in 2005, Julian day 154-184 is selected at Bondville in 2006, Julian day 228-258 is selected at Blackhill in 2005 (many flux data is missed from May to July). As can be seen from Fig.4—Fig.6, the assimilation results are closer to the observation than simulation results. It is very important to produce more accurate MODIS LST products for data assimilation. Such as Julian day 164—168 at Bondville site (Fig.5), the accuracy of MODIS LST products is high, and the assimilation results are also closer to observation than simulation. In Julian day 155, as MODIS temperature (303.78K, 309K) is higher than ground-measured temperature (300.68K, 306.4K), the assimilation results are no better than simulation. Surface temperature simulation results are not good under forest surface (Fig.6), while data assimilation method is able to improve the estimation accuracy of sensible and latent heat fluxes.

5 SUMMARY AND DISCUSSIONS

This paper summarizes a point-scale land data assimilation scheme based on EnKF algorithm and CLM. MODIS LST products were assimilated into the system. The estimation

Table 6 Summary of data assimilation results at Brookings, Bondville, and Blackhill sites

		Brookings		Bondville		Blackhill	
		H	LE	H	LE	H	LE
Samples		6517	8452	10348	9965	11813	11564
Simulation	RMSE/(W·m^{-2})	46.5	68.7	47.0	88.6	81.5	53.4
Assimilation	RMSE/(W·m^{-2})	45.1	73.3	31.8	57.7	58.4	47.2

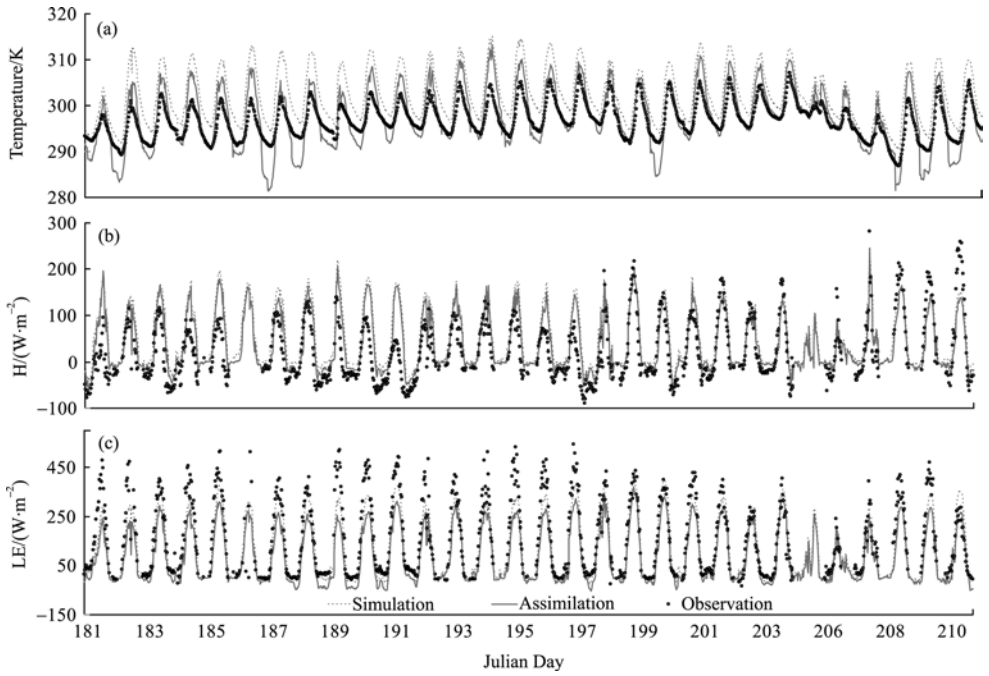

Fig. 4 Simulation and assimilation results of the surface temperature, sensible heat and latent heat fluxes at Brookings site
(a) Surface temperature; (b) Sensible heat flux; (c) Latent heat flux

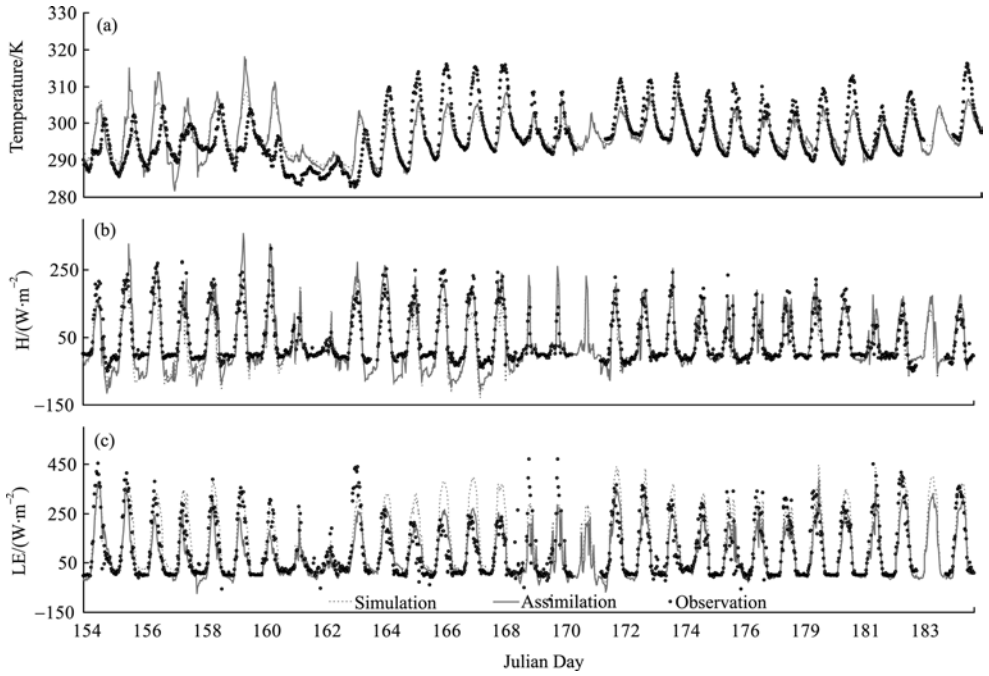

Fig. 5 Simulation and assimilation results of the surface temperature, sensible heat and latent heat fluxes at Bondville site
(a) Surface temperature; (b) Sensible heat flux; (c) Latent heat flux

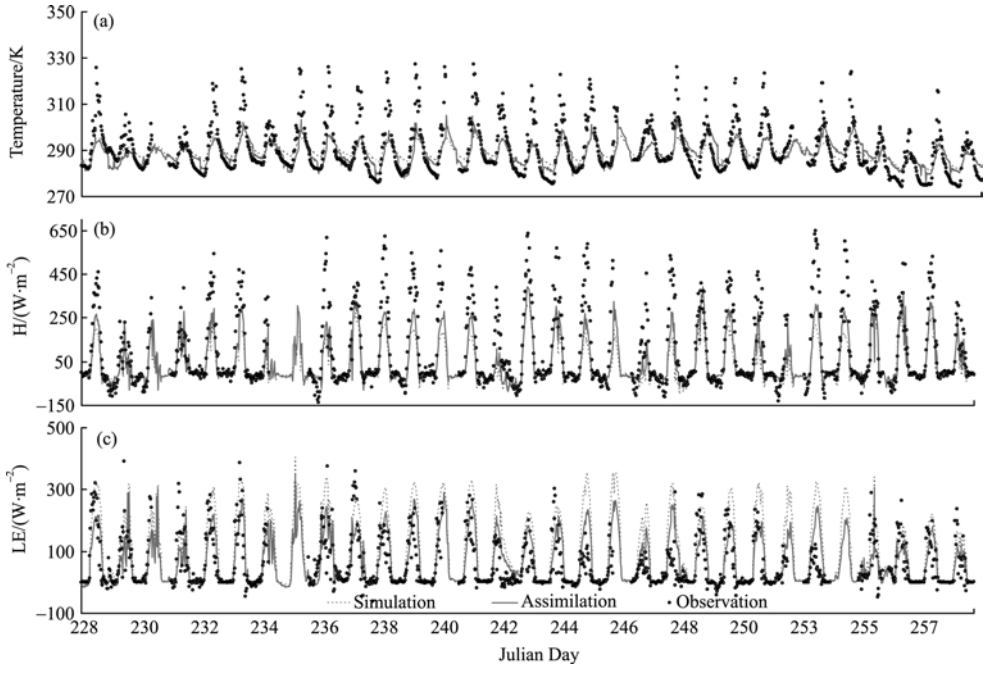

Fig. 6 Simulation and assimilation results of the surface temperature, sensible heat and latent fluxes at Blackhill site
(a) Surface temperature; (b) Sensible heat flux; (c) Latent heat flux

accuracy of surface temperature, sensible and latent heat flux were improved. MODIS LAI products were also used to update LAI in CLM, which also improved the model simulation results.

The estimation of sensible heat flux has obvious improvement by using data assimilation method over different land cover types. At Blackhill site (forest land), surface temperature simulations has large deviation with ground-measured temperature (about 10K during the daytime), while the RMSE between MODIS LST products and the ground-measured temperature is only about 1.86K. So the assimilation results (temperature and sensible heat flux estimates) have obviously improvement (RMSE decreased from 81.5W·m^{-2} to 58.4W·m^{-2} compared with the simulation results), and the estimation of latent heat flux also has improvement (RMSE decreased from 53.4W·m^{-2} to 47.2W·m^{-2} compared with the simulation results). At Bondville station (crop land), RMSE of sensible heat flux decreased from 47.0W·m^{-2} to 31.8W·m^{-2}, and RMSE of latent

heat flux decrease from 88.6W·m^{-2} to 57.7W·m^{-2}. At Brookings station (grass land), RMSE of sensible heat flux decreased from 46.5W·m^{-2} to 45.1W·m^{-2}.

The results also indicate that the accuracy of MODIS LST products is one of the key factors for data assimilation. The accuracy of MODIS LST products is relatively high, and rigorous quality control was implemented on the selection of MODIS LST products (only use the data with QC=0). Therefore, the estimation of surface temperature has greatly improvement compared with simulation in most assimilation cases. However, there is still a big deviation between MODIS LST products and ground-measured temperature at some time. The deviation between MODIS LST and ground-measured temperature is not only the accuracy of MODIS LST, but also the difference of temporal scales (MODIS surface temperature is an instantaneous value, and ground-measured temperature is mean value of 30 min) and spatial scales (MODIS LST spatial scale is 1km × 1km, and ground-measured temperature spatial scale is about tens of square meters). In addition, simulation of surface temperature has systematic errors, which would make the assimilation results worse than simulation. For example, Julian day 155 (Fig.5) at Bondville site, MODIS LST is about 3K higher than ground-measured temperature, and surface temperature simulation also has phase deviations, which make the assimilation result 10k higher than the observation during the daytime. Therefore, it will be helpful for data assimilation to develop more accurate MODIS LST products. In this paper, the relationship equation between surface temperature simulation and MODIS LST was defined as the observation operator, and MODIS LAI products were also used in this observation operator. The operator can be used in a regional data assimilation system, but its applicability needs further validation.

The process described by CLM is relatively complex, and requires lots of vegetation and soil parameters which are difficult to acquire. The forcing data also have errors. All the limitations contaminated model outputs. In addition, there is a systematic error in model simulated surface-air temperature differences, and surface temperature change fast in one day, which make data assimilation particularly difficult. Therefore, better assimilation strategy need to be proposed in next step.

There is an "energy imbalance" phenomenon in eddy covariance system which can measure sensible and latent heat fluxes. Due to this phenomenon, about 10%—25% energy (about 20—50W·m^{-2}) disappear (Table 4). The validation results may be easier to understand after the observations are corrected. In addition, the CLM simulated surface temperature is mean value of a few kilometers, while the ground-measured surface temperature only represent about a few square meters. There are also some measurement errors of observation instrument, which may influence the validation results.

In land data assimilation system, the magnitude of model error and observation error determine the direction that the data assimilation results go around. If model error is large, the assimilation results may be closer to the observations, and vice versa. However, model error and observation error are difficult to determine, and they may change with the variation of time and space. In this paper, with statistical methods, fixed values were determined as the model and observation error which may smooth the model and observation error and reduce the performance of assimilation system. Therefore, assimilation results may be improved with the development of new methods to determine model and observation error.

REFERENCES

Anderson M C and Norman J M. 2005. Effects of vegetation clumping on two-source model estimates of surface energy fluxes from an agricultural landscape during SMACEX. *Journal of Hydrometeorology*, **6**: 892—909

Bastiaanssen W G M, Menenti M, Feddes R A and Holtslag A A M. 1998. The surface energy balance algorithm for land (SEBAL): Part 1 Formulation. *Journal of Hydrology*, **212—213**: 198—212

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

Caparrini F, Castelli F and Entekhabi D. 2004. Estimation of surface turbulent fluxes through assimilation of radiometric surface temperature sequences. *Journal of Hydrometeorology*, **5**: 145—159

Coll C, Caselles V, Joan M, Valor E, Niclos R, Sanchez J M and Rivas R. 2005. Ground measurements for the validation of land surface temperatures derived from AATSR and MODIS data. *Remote Sensing of Environment*, **97**: 288—300

Crosson W L, Laymon C A, Inguva I and Schamschula M P. 2002. Assimilating remote sensing data in a surface flux–soil moisture model. *Hydrological Processes*, **16**: 1645—1662

Crow W T and Wood E F. 2003. The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble kalman filtering: a case study based on ESTAR measurements during SGP97. *Advances in Water Resources*, **26**: 137—149

Crow W T. 2006. Impact of incorrect model error assumptions on the sequential assimilation of remotely sensed surface soil moisture. *Journal of hydrology*, **7**: 421—431

Dai Y J, Zeng X B and Dickinson R E. 2001. Common Land Model (CLM): Technical Documentation and User's Guide. < http://219.142.99.7/models.do > (10 Oct, 2007)

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. 2003. The common land model. *Bulletin of the American Meteorological Society*, **84**: 1013—1023

Dickinson R E, Henderson S A and Kennedy P J. 1986. Biosphere–Atmosphere Transfer Scheme (BATS) for the NCAR Community Climate Model, NCAR Technical Note NCAR/TN-275CSTR, National Center for Atmospheric Research, Boulder, CO

Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast

error statistics. *Journal of Geophysical Research*, **99**: 10143—10162

Evensen G. 2003. The ensemble kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, **53**: 343—367

Galantowiez J F, Entekhabi D and Njoku E G. 1999. Test of sequential data assimilation for retrieving profile soil moisture and temperature from observed l-band radio brightness. *IEEE Transactions on Geoscience and Remote Sensing*, **37**: 1860—1870

Gao Z Q and Liu J Y. 2003. The comparison of land surface temperature with CLM and split window retrieving method. *Acta Geographica Sinica*, **58**(4): 494—502

Huang C L and Li X. 2006. Experiments of soil moisture data assimilation system based on ensenble kalman filter. *Plateau Meteorology*, **25**(4): 665—671

Huang C L, Li X, Lu L and Gu J. 2008a. Experiments of one-dimensional soil moisture assimilation system based on ensemble Kalman filter. *Remote Sensing of Environment*, **112**(3): 888—900

Huang C L, Li X and Lu L.2008b.Retrieving soil temperature profile by assimilating MODIS LST products with ensemble Kalman filter. *Remote Sensing of Environment*, **112**(4): 1320—1336

Huang M F, Liu S M, Liu S H and Zhu Q J. 2005. A study of the difference between true surface temperature and radiometric surface temperature. *Advances in Earth Science*, **20**(10): 1075—1082

Jin M L and Liang S. 2006. An improved land surface emissivity parameter for land surface models using global remote sensing observations. *Journal of Climate*, **19**: 2867—2881

Kalman R E. 1960. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, **82** (Series D): 35—45

Kumar P and Kaleita A L. 2003. Assimilation of near-surface temperature using extended kalman filter. *Advances in Water Resources*, **26**: 79—93

Li X, Toshio K and Mahadevan P. 2004. A very fast simulated re-annealing (VFSA) approach for land data assimilation. *Computers &Geosciences*, **30**: 239—248

Liang S and Qin J. 2008. Advances in Land Remote Sensing: System, Modeling, Inversion and Applications. Springer Verlag

Liu S M, Hu G, Lu L and Mao D F. 2007. Estimation of regional evapotranspiration by TM/ETM+ data over heterogeneous surfaces. *Photogrammetric Engineering and Remote Sensing*, **73**(10): 1169—1178

Mathias G, Corinna R and Thomas F. 2004. A combination of quality assessment tools for eddy covariance measurements with footprint modelling for the characterisation of complex sites. *Agricultural and Forest Meteorology*, **127**: 175—188

Qin J, Liang S L, Liu S M, Wang J D and Li X W. 2005. Estimation of land surface energy balance components by assimilating MODIS skin temperature into common land model using the ensemble kalman filter method, The 9th International Symposium on Physical Measurements and Signatures in Remote Sensing. Beijing: ISPRS

Reichle H R, Mclaughlin D B and Entekhabi D. 2001. Variational data assimilation of microwave radiobrightness observations for land Surface hydrology applications. *IEEE Transactions on Geoscience and Remote Sensing*, **39**: 1708—1718

Reichle H R, Mclaughlin D B and Entekhabi D. 2002. Hydrologic data assimilation with the ensemble kalman filter. *American Meteorological Society*, **130**: 103—114

Schuurmans J M, Troch P A, Veldhuizen A A, Bastiaanssen W G M and Bierkens M F P. 2003. Assimilation of remotely sensed latent heat flux in a distributed hydrological model. *Advances in Water Resources*, **26**: 151—159

Sellers P J, Randall D A, Collatz G J, Field C B, Dazlich D A, Zhang C, Collelo G D and Bounoua L.1996. A revised land surface parameterization (SiB2) for atmosphere GCMs, part I: model formulation. *Journal of Climate*, **9**: 676—705

Su Z. 2002.The surface energy balance system (SEBS) for estimation of turbulent heat fluxes. *Hydrology & Earth System Sciences*, **6**(1): 85—99

Wang W H, Liang S L and Tilden M. 2008. Validating MODIS land surface temperature products using long-term nighttime ground measurements. *Remote Sensing of Environment*, **112**(3): 623—635

Wilson K, Goldstein A, Falge E, Aubinet M, Baldocchi D, Berbigier P, Bernhofer C, Ceulemans R, Dolman H, Field C, Grelle A, Ibrom A, Law B E, Kowalski A, Meyers T, Moncrieff J, Monson R, Walter O, Tenhunen J, Valentini R and Verma S. 2002. Energy balance closure at FLUXNET sites. *Agricultural and Forest Meteorology*, **113**: 223—243

Yan Y, Liu Q H and Liu Q. 2006. Methodology of winter wheat yield prediction based on assimilation of remote sensing data with crop growth model. *Journal of Remote Sensing*, **10**(5): 804—811

Yang K, Takahiro W and 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. *Journal of the Meteorological Society of Japan*, **85**: 229—242

Zhang S W and Qiu C J. 2004. Estimating soil water contents from soil temperature measurements by using an adaptive kalman filter. *Journal of Applied meteorology*, **43**: 379—389

同化 MODIS 温度产品估算地表水热通量

徐同仁[1], 刘绍民[1], 秦 军[2], 梁顺林[3]

1. 北京师范大学 遥感科学国家重点实验室, 地理学与遥感科学学院, 北京 100875;
2. 中国科学院 青藏高原研究所, 北京 100085;
3. Department of Geography,2181 Lefrak Hall, University of Maryland, College Park, MD20742, USA

摘 要: 基于集合卡尔曼滤波和通用陆面模型(CLM 1.0)发展了一个地表温度的同化系统。这个系统同化了 MODIS 温度产品, 并将 MODIS 的叶面积指数引入 CLM 模型中, 主要用于改进地表水热通量的估算精度。将 CLM 输出的地表温度与 MODIS 地表温度建立关系, 并作为同化系统的观测算子。将 MODIS 地表温度与实测地表温度进行了比较, 将其均方差(Root Mean Square Error, RMSE)作为观测误差。选取 3 个美国通量网站点(Blackhill、Bondville、Brookings)作为实验数据, 结果表明: 同化结果中地表温度、显热通量的估算精度均有提高。其中 Blackhill 站的估算精度改进最大, 均方差由 81.5W·m^{-2} 减小到 58.4W·m^{-2}, Bondville 站均方差由 47.0W·m^{-2} 减小到 31.8W·m^{-2}, Brookings 站均方差由 46.5W·m^{-2} 减小到 45.1W·m^{-2}。潜热通量估算精度在 Bondville 站均方差由 88.6W·m^{-2} 减小到 57.7W·m^{-2}, Blackhill 站均方差由 53.4W·m^{-2} 减小到 47.2W·m^{-2}。总之, 结合陆面过程模型同化 MODIS 温度产品估算地表水热通量是可行的。

关键词: MODIS 温度产品, 通用陆面模式, 集合卡尔曼滤波, 地表水热通量

中图分类号: TP701 文献标识码: A

1 引 言

地表水热通量的准确估算在气候变化研究、水资源规划与管理以及节水农业中非常关键。地表水热通量的估算有许多种方法。气象学方法和气候学方法在单点估算上比较准确, 但是不容易在空间上扩展。水文学方法虽然能够计算大面积的蒸散量, 但时间周期较长(一般以年为周期)。遥感方法具有宏观、实时的特点, 这给地表水热通量的监测带来了新的机遇(Bastiaanssen 等, 1998; Su, 2002; Liu 等, 2007), 但这种方法缺乏对地表过程机理的描述, 而且遥感估算的水热通量是瞬时的, 由瞬时值推算到日、月、年时会遇到困难。近 20 年来发展的陆面过程模型能够模拟土壤-植被-大气连续体中能量和水分的连续变化过程(Dickinson 等, 1986; Sellers 等, 1996; Dai 等, 2001, 2003), 对于了解区域能量和水分平衡意义重大, 但陆面过程模型所描述的过程一般比较复杂, 需要众多的区域植被、土壤和气象参数, 而很多参数是很难确定的, 这使得陆面过程模型在描述地表各变量时不够精确, 也阻碍了陆面过程模型的发展和应用。

遥感数据具有宏观、实时的特点, 陆面过程模型可描述地表过程的连续变化, 数据同化则可结合两者的优势。利用不同来源、不同空间和时间分辨率的观测数据(包括各种地面观测数据以及卫星遥感数据), 将它们有机融合, 并结合陆面过程模式发展同化算法以提供高精度、高分辨率的土壤温、湿度数据, 对于理解陆气相互作用机理和改进气候系统的模拟和预测具有重要意义。近年来, 用于估计陆面参数的同化方法得到很大发展(Li 等, 2004; Yan 等, 2006; Liang 等, 2008)。使用同化技术将陆面过程模型和遥感数据相结合成为土壤水分、地表温度与地表水热通量估算的热点。在土壤水分同化方面, 目前主要有同化台站的土壤湿度观测数据(Zhang & Qiu, 2004; Huang & Li, 2006)和微波遥感数据(Galantowiez 等, 1999; Reichle 等, 2001, 2002;

收稿日期: 2008-06-19; 修订日期: 2009-01-07

基金项目: 国家高技术研究发展计划课题(编号: 2007AA12Z175)、国家自然科学基金项目(编号: 40671128)和国家重点基础研究发展规划项目(编号: 2007CB714401)。

第一作者简介: 徐同仁(1982—), 男, 在读博士研究生, 北京师范大学, 主要从事陆面数据同化方面的研究。

通讯作者: 刘绍民, E-mail: smliu@bnu.edu.cn。

Crosson, 2002; Crow & Wood, 2003; Yang 等, 2007; Huang 等, 2008a), 并且取得了较好的效果。

在地表水热通量的估算中, 地表温度是一个关键变量。一些学者对陆面温度的同化做了很多的研究。Boni 等(2001)使用较为简单的模型, 以变分的方式同化了卫星传感器得到的地表温度数据, 同时考虑了地表温度的观测误差与地表能量平衡的物理约束, 得到了地表能量平衡组分和地表蒸散的估算。Schuurmans 等(2003)同化了遥感反演的潜热通量。Kumar 等(2003)发现 AVHRR 反演的温度与 5cm 实测土壤温度有很好的相关性, 以拟合的误差作为观测误差利用扩展卡尔曼滤波算法将 AVHRR 反演的地表温度同化到陆面模型中, 对温度廓线和地表水热通量进行了估算。Caparrini 等(2004)使用 force-restore 方程结合变分方法同化了遥感反演的地表温度, 并对地表水热通量进行了估算。Qin(2005)、Huang 等(2008b)将 MODIS 温度产品与站点实测地表温度进行拟合作为观测算子, 拟合的均方差作为观测误差, 结合集合卡尔曼滤波算法和 CLM 模型同化了 MODIS 温度产品, 较好地反演了各层土壤温度。

随着陆面数据同化研究的深入, 适于陆面数据同化的最优化算法也迅速发展起来。卡尔曼滤波最早由 Kalman(1960)首先提出。对于线性动态系统, 标准卡尔曼滤波无疑是最优的数据同化方法。由于大部分系统是非线性的, 所以人们发展了集合卡尔曼滤波。集合卡尔曼滤波由 Evensen(1994,2003)根据 Epstein 的随机动态预报理论提出的, 通过 Monte-Carlo 法来计算状态预报误差的协方差。将模型状态预报看成近似随机动态预报, 用一个状态总体(设数目为 N)来代表随机动态预报中的概率密度函数。集合卡尔曼滤波克服了卡尔曼滤波要求线性化的模型算子和观测算子的缺点。在集合卡尔曼滤波中, 模型误差的确定对预报结果至关重要。一些学者做了一系列的理想试验, 详尽地分析了错误的模型误差估计(类型、大小)对同化结果的影响(Evensen,2003;Crow,2006)。Reichle(2002)采用了与时间相关的模型误差进行了同化研究。Crosson(2002)采用固定值作为模型误差进行了同化实验。Qin(2005)采用模型误差与状态变量大小相关的方法实现了 CLM 模型同化 MODIS 温度产品。

2 同化算法

目前陆面过程模型对于地表温度和地表水热通量的模拟并不十分准确。由于模型中地表温度的计算是前后关联的, 将瞬时的遥感地表温度产品同化到陆面过程模型中, 不仅改变当前时刻的地表温度和土壤温度, 而且对地表温度和土壤温度的日变化也有较大影响, 从而也影响到地表水热通量的估算。图 1 是同化算法的流程图。具体流程如下:

(1) 利用地面观测站获取陆面过程模型所需要的气象驱动数据;

(2) 应用遥感技术和地面观测得到陆面过程模型所需的地表参数集;

(3) 驱动数据与地表参数进入陆面过程模型, 生成当前时刻的状态变量;

(4) 同化当前时刻的 MODIS 的地表温度产品, 优化状态变量;

(5) 陆面同化继续向前运行, 生成下一时刻的背景场;

(6) 不断调整, 最后得到最优的输出结果。

2.1 陆面过程模型

本文选用的陆面过程模型是通用陆面模式 CLM1.0(Dai 等, 2001,2003)。CLM 模型主要由物理、水文、生物过程构成, 可以模拟地表温度、土壤温度、土壤水分、感热通量、潜热通量等状态变量。

本文中同化算法的核心是 CLM 模型中地表温度的计算。CLM 模型模拟地表温度是根据地气之间能量平衡方程迭代得到的。太阳辐射是地表和大气的主要能源, 地表是能量接收、储存和转换的主要场所。对于裸地, 能量平衡方程为:

$$-S_{\mathrm{g}}+H_{\mathrm{g}}(T_{\mathrm{g}})+\lambda E_{\mathrm{g}}(T_{\mathrm{g}})+G_{\mathrm{g}}(T_{\mathrm{g}})=0 \tag{1}$$

式中, S_{g} 为土壤的净辐射; H_{g}, λE_{g}, G_{g} 分别为土壤的感热通量, 潜热通量和土壤热通量; T_g 是土壤温度。

对于有植被覆盖的地表, 用土壤温度和植被冠层温度来描述的能量平衡方程可表示为:

$$-S_{\mathrm{v}}+H_{\mathrm{v}}(T_{\mathrm{g}},T_{\mathrm{v}})+\lambda E_{\mathrm{v}}(T_{\mathrm{g}},T_{\mathrm{v}})=0 \tag{2}$$

$$-S_{\mathrm{g}}+H_{\mathrm{g}}(T_{\mathrm{g}},T_{\mathrm{v}})+\lambda E_{\mathrm{g}}(T_{\mathrm{g}},T_{\mathrm{v}})+G_{\mathrm{g}}(T_{\mathrm{g}},T_{\mathrm{v}})=0 \tag{3}$$

式中, S_{v} 为植被冠层的净辐射; H_{v}, λE_{v} 分别为植被冠层的感热通量和潜热通量; T_{v} 是植被冠层温度。如果有积雪, 式(1)、式(3)还需加一项融雪耗热量。

由上述公式可以看出, 净辐射、感热通量、潜热通量、土壤热通量与地表温度有关, 通过改变地表温度的预报, 即可改变地表水热通量的预报。同化地表温度, 必须找一个切入点。相比较而言, 土壤的热传导方程比较适合。CLM 模型中土壤的分层如图 2。

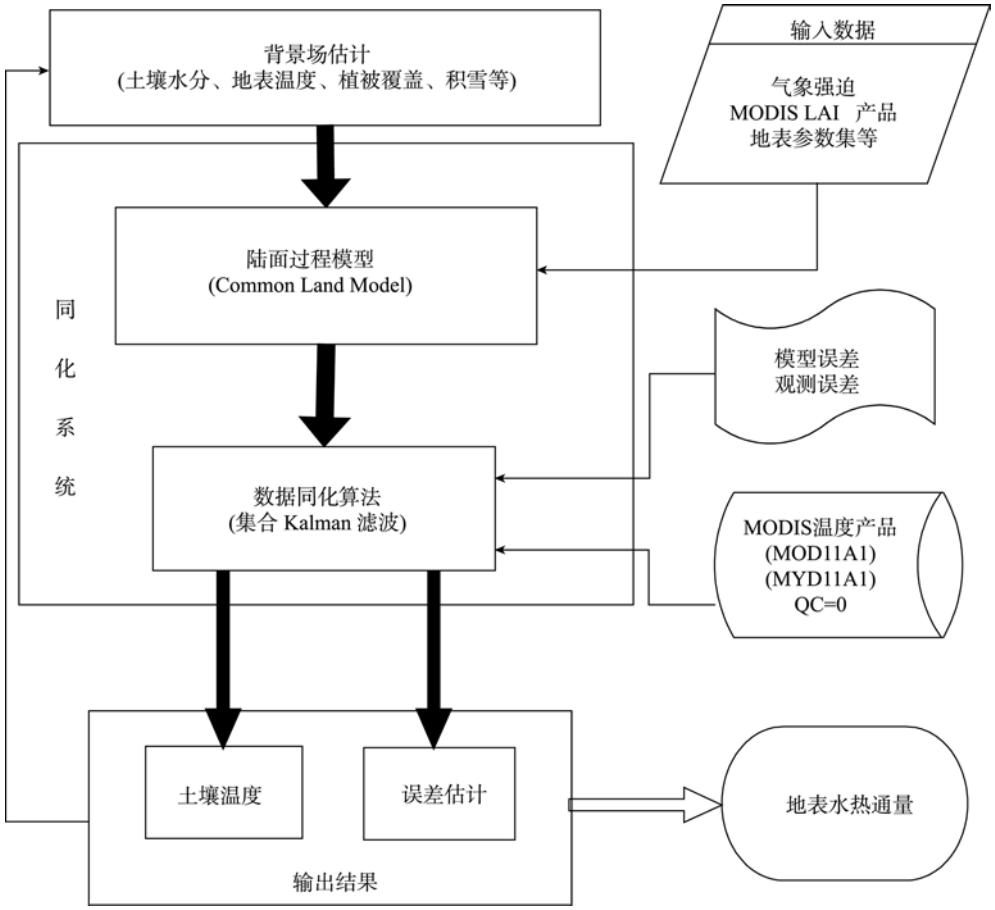

图 1 同化系统流程图

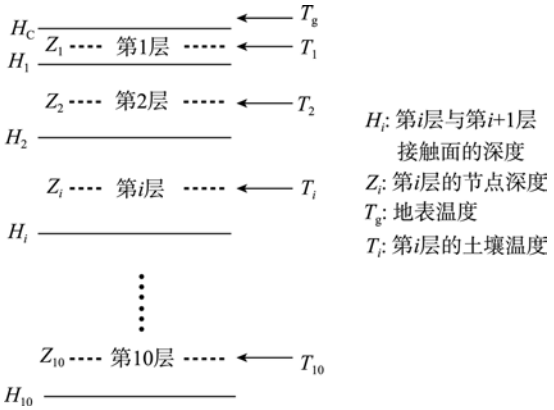

图 2 CLM 模型中土壤分层示意图

在CLM模型中, 土壤温度的变化特征可用热传导方程表示:

$$c\frac{\partial T}{\partial t}=-\frac{\partial}{\partial Z}\left(k\frac{\partial T}{\partial Z}\right) \tag{4}$$

式中, c 为土壤体积热容量, T 是土壤温度, t 是时间, k 为土壤热传导系数, Z 是土壤厚度。模型中土壤分为 10 层, 厚度分别为 0.00710, 0.02793, 0.06226, 0.11887, 0.21219, 0.36607, 0.61976, 1.03803, 1.72764, 2.86461m。将公式(4)转化成差分形式:

$$c_j\Delta Z_j\frac{T_j^{n+1}-T_j^n}{\Delta t}=\frac{1}{2}(F_j^n-F_{j-1}^n+F_j^{n+1}-F_{j-1}^{n+1}) \tag{5}$$

式中, c_j 表示 j 层土壤体积热容量, T_j 表示 j 层土壤平均温度, n 和 n+1 分别表示第 n 和 n+1 时刻, Δt 为时间步长, ΔZ_j 是 j 层土壤层厚度, F_j 是 j 到 j+1 层的土壤热通量。

在土壤层内部,

$$F_j=-\lambda\left[Z_{h,j}\right]\left(\frac{T_j-T_{j+1}}{Z_{j+1}-Z_j}\right) \tag{6}$$

式中, $\lambda[Z_{h,j}]$ 表示 $Z_{h,j}$ 深度的土壤热传导系数, $Z_{h,j}$ 表示第 j 层与第 j+1 层界面的土壤深度, Z_{j+1} 表示第 j+1 层的土壤深度, Z_j 表示第 j 层的土壤深度。假定 Z_j 到 $Z_{h,j}$ 的土壤热通量等于 $Z_{h,j}$ 到 Z_{j+1} 的土壤热通量, 则得到下式:

$$\lambda_j\frac{T_m-T_j}{Z_{h,j}-Z_j}=\lambda_{j+1}\frac{T_{j+1}-T_m}{Z_{j+1}-Z_{h,j}}=\lambda(Z_{h,j})\frac{T_{j+1}-T_j}{Z_{j+1}-Z_j} \tag{7}$$

式中, λ_j 表示第 j 土壤的土壤热传到系数, T_m 表示 $Z_{h,j}$ 深度的土壤温度。通过公式(7)$\lambda(Z_{h,j})$可以表示为:

$$\lambda(Z_{h,j})=\frac{\lambda_j\lambda_{j+1}[Z_{j+1}-Z_j]}{\lambda_j[Z_{j+1}-Z_{h,j}]+\lambda_{j+1}[Z_{h,j}-Z_j]} \tag{8}$$

假定最底层的土壤热通量为 0，可以通过上述公式计算得到各层的土壤温度。

在这个同化系统中，本文只关心 CLM 中各层的温度，所以可将 CLM 看成一个“黑盒子”系统，表示为:

$$X_{k+1} = M(X_k, \theta_{k+1}, \beta_{k+1}) \tag{9}$$

式中，X_{k+1}, X_k 表示 k+1, k 时刻的各层土壤温度; $M(\cdot)$ 表示模型算子; θ_{k+1} 表示 k+1 时刻的驱动参数; β_{k+1} 表示 k+1 时刻的地表参数。

由公式(9)可以看出，在同化过程中通过调整 k 时刻的地表温度，可以改变 k+1 时刻的地表温度的预报，从而改变 k+1 时刻的地表水热通量的估算。

2.2 集合卡尔曼滤波算法

Evensen(1994, 2003)提出了集合卡尔曼滤波算法。该算法可以处理模型的非线性问题，而且该算法简单易行，近年来被数据同化领域广泛应用。下面将对该算法做简单的介绍。

2.2.1 初始化模型

将 X_0 定义为模型的初始值，

$$X^a_{i,0} = X^a_0 + u_i \quad u_i \sim N(0, P) \tag{10}$$

式中，$X^a_{i,0}$ 代表初始时刻的第 i (i 为集合数)分析变量，上标“a”表示分析变量，u_i 是均值为零、方差为 P 的高斯分布的背景误差向量，P 为背景误差。

2.2.2 模型状态预报

预报方程为:

$$X^f_{i,k+1} = M(X^a_{i,k}, \theta_{k+1}, \beta_{k+1}) + w_i \quad w_i \sim N(0, Q) \tag{11}$$

式中，$X^f_{i,k+1}$ 代表第 k+1 时刻的第 i(i 为集合数)预报变量，其中上标“f”表示预报变量，$X^a_{i,k}$ 代表第 k 时刻的第 i(i 为集合数)分析变量，w_i 是均值为零、方差为 Q 的高斯分布的模型误差向量，Q 是模型误差。

2.2.3 观测算子

观测算子可以表示为:

$$Y_{i,k+1} = H(X^f_{i,k+1}) + v_i \quad v_i \sim N(0, R) \tag{12}$$

式中，$Y_{i,k+1}$ 是模拟的第 k+1 时刻的第 i 个观测值(本文中为 MODIS 温度产品)，$H(\cdot)$代表与模型状态变量相对应的观测算子，v_i 是均值为零、方差为 R 的高斯分布的观测误差向量，R 为观测误差，可通过 MODIS 的地表温度与地面实测温度比较得到。

当没有观测值的时候，模型一直向前预报。当遇到观测值以后，模型的状态变量就会得到如下的更新:

$$X^a_{i,k+1} = X^f_{i,k+1} + K_{k+1}(Y^o_{k+1} - H(X^f_{i,k+1}) + v_i) \tag{13}$$

$$K_{k+1} = P^f_{k+1} H^T (H P^f_{k+1} H^T + R)^{-1} \tag{14}$$

$$P^f_{k+1} = \frac{1}{N-1}\sum_{i=1}^{N}(X^f_{i,k+1} - \bar{X}^f_{k+1})(X^f_{i,k+1} - \bar{X}^f_{k+1})^{\mathrm{T}} \tag{15}$$

$$P^f_{k+1} H^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}[X^f_{i,k+1} - \bar{X}^f_{k+1}][H(X^f_{i,k+1}) - H(\bar{X}^f_{k+1})]^{\mathrm{T}} \tag{16}$$

$$H P^f_{k+1} H^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}[H(X^f_{i,k+1}) - H(\bar{X}^f_{k+1})]\times [H(X^f_{i,k+1}) - H(\bar{X}^f_{k+1})]^{\mathrm{T}} \tag{17}$$

式中，$X^a_{i,k+1}$ 代表第 k+1 时刻的第 i(i 为集合数)状态变量(本文指地表温度)，K_{k+1} 是 k+1 时刻的观测增量的系数矩阵，Y^o_{k+1} 是 k+1 时刻的观测变量(本文指 MODIS 温度产品)，P^f_{k+1} 是 k+1 时刻的背景误差协方差矩阵，$\bar{X}^f_{i,k+1}$ 表示 k+1 时刻的预报的状态变量均值(本文指地表温度的均值)，$[\cdot]^{\mathrm{T}}$ 表示矩阵的转置，$H(X^f_{i,k+1})$ 是模拟的 k+1 时刻的第 i(i 为集合数)状态变量(本文指地表温度)，$H(\bar{X}^f_{k+1})$ 是 k+1 时刻模拟的状态变量的均值(本文指地表温度的均值)。

3 数据获取与处理

Wang 等(2008)分析了 MODIS 温度产品在美国通量网部分站点的精度。本文参考其分析结果，选择了 Bondville、Brookings 和 Blackhills 站进行同化试验，如表 1，涉及农作物、草地和森林等下垫面。这些站点所在的实验区是下垫面均一，四周开阔，其中 Bondville 站点还是 MODIS 数据产品的地面验证点。这 3 个站的地面资料齐全，并有对应的 MODIS 卫星数据产品。地面气象、通量数据均为 0.5h/次。

本文所用的 MODIS 温度产品空间分别率为 1km，地面气象、通量数据均为站点观测数据。由于所选站点所在的实验区是下垫面均一，四周开阔，认为地面气象数据、通量数据能够代表数公里的范围(Mathias, 2004)。

3.1 MODIS 数据

本文选择 MODIS 标准数据产品作为试验数据，如表 2 所示。MODIS 数据产品主要以 HDF 格式存储，空间投影方式为 Sinusoidal。此外还存储了相应的过境时间和质量控制说明。数据产品可以通过网站(https://wist.echo.nasa.gov/api/)获得。

本文选取精度较高的 MODIS 温度、叶面积指数数据(QC=0)进行试验。对于精度不高(QC>0)的数据，则根据该像元往年的数据来填补。表 3 为各站点 1—12 月 MODIS 的 LAI 数据。

表 1 AmeriFlux 站点的概况

观测站名称	北纬/(°)	西经/(°)	植被类型	土壤类型	叶面积指数	冠层高度/m	时间	通量塔高度/m
Bondville	40.01	88.29	农作物	粉砂土壤	0—5.0	0.9	2005—2006 年	10.0
Brookings	44.35	96.84	稀疏草地	砂质黏土	0.2—3.0	0—0.4	2004—2005 年	4.0
Blackhill	44.15	103.64	常绿针叶林	黏土	2.0—5.0	13—15	2004—2005 年	24.0

表 2 MODIS 标准数据产品

MODIS 标准产品	地表参数	空间分辨率/m	时间分辨率/d
MYD11A1	地表温度 LST	1000	1
MOD11A1	地表温度 LST	1000	1
MOD15A2	叶面积指数 LAI	1000	8

表 3 MODIS 标准叶面积指数产品

站点	年份	1	2	3	4	5	6	7	8	9	10	11	12
Bondville	2006	0.4	0.3	0.3	0.5	0.5	2.3	4.9	2.4	0.9	0.5	0.2	0.3
Brookings	2005	0.1	0.2	0.1	0.9	1.4	1.6	1.6	2.3	1.5	0.9	0.4	0.1
Blackhill	2005	2.2	1.8	1.8	1.9	3.8	3.6	5.1	4.8	3.8	1.9	2.1	2.2

3.2 气象数据

气象驱动数据通过美国通量网站: www.fluxnet.ornl.gov/fluxnet/index.cfm 下载。本文采用 3 个站观测的气象数据作为 CLM 模型驱动数据, 包括: 空气温度、空气湿度、大气压、风速、降水量、太阳总辐射、大气下行辐射。另外用到的气象数据还包括: 反射辐射、地表向上长波辐射、净辐射、土壤热通量、地表红外温度、土壤温度(观测深度为 2, 4, 8, 16, 32, 64, 128cm)、土壤湿度(观测深度为 10, 20, 30, 40, 60, 100cm)。

3.3 通量数据

本文采用的通量数据包括: 显热通量和潜热通量。美国通量网站涡动相关仪观测数据经过一系列基本处理, 包括: 野点值的剔除、坐标转换处理、WPL 订正等, 这些处理保证了观测数据的质量。此外本文对通量观测数据进行严格的筛选, 剔除了降雨天的数据和夜间摩擦风速小于 0.1m/s 的通量数据。

一般情况下, 涡动相关仪观测的地表能量存在不闭合现象(Wilson 等, 2002), 即能量平衡的闭合率小于 1.0。本文对所选站点进行了能量平衡闭合率的计算。表 4 表示了各个站点的能量平衡闭合状况。

表 4 3 个站的能量平衡闭合率

观测站	Bondville		Brookings		Blackhills	
	全天	白天	全天	白天	全天	白天
样本数	9956	5638	8454	4828	11543	5269
能量平衡闭合率	0.76	0.76	0.91	0.85	0.91	0.86

从表 4 看出: Brookings、Bondville 和 Blackhills 站的能量平衡闭合率均小于 1, 存在能量平衡不闭合现象。

4 结果分析

4.1 模型误差

预报误差的大小主要取决于模型误差和观测误差的大小。通常认为模型误差可以随便给定, 因为它会随着时间的推移逐渐消失, 但这是建立在观测时间间隔不大, 而且观测误差不大的基础上。否则, 预报误差主要取决于模型误差。

本文将模型误差看成静止的, 不跟其他变量相关, 并且利用统计方法将模型误差做成查找表的形式。

4.2 观测算子与观测误差

Gao 等(2003)比较了 CLM 模型模拟的地表温度与遥感反演得到的地表温度。卫星遥感反演的地表温度是通过 Plank 定律得到的。传感器获得的辐射包括地表和大气辐射两部分。地表温度对裸土, 就是土壤表面温度; 对于稠密的植被, 可以说是植被冠层温度; 而对稀疏的地表, 就是植被冠层和土壤表面的一个混合温度。考虑遥感数据的空间分辨率, 卫星遥感反演的地表温度就是像元尺度上不同地表类型对应的一个混合温度。CLM 中的地表温度可以通过下面公式得到的:

$$T^4 = f_{\mathrm{veg}}(T_{\mathrm{v}})^4 + (1 - f_{\mathrm{veg}})(T_{\mathrm{g}})^4 \qquad (18)$$

式中, T 为 CLM 模拟的地表温度, T_v 为植被冠层温度, T_g 为土壤表面温度, f_{veg} 为植被覆盖度。

观测算子的建立是同化系统构建成败的关键之一, 在本文问题则转化为如何将 MODIS 温度产品与 CLM 模型的状态变量建立关系。很多学者将遥感反演的温度与地表实测的温度建立拟合关系并作为观测算子(Kumar 等, 2003; Qin 等, 2005; Huang 等, 2008b), 这样做在单点或小区域范围会取得较好的结果, 但同样的拟合关系并不能应用到更多的点或区域上。本文采用了组分温度分解的方法将模型状态变量与 MODIS 的地表温度建立关系式(Anderson 等, 2005):

$$T_{RAD}^4(\theta) \approx f(\theta)\, T_c^4 + (1-f(\theta))T_s^4 \tag{19}$$

式中, T_{RAD} 是模型模拟的地表辐射温度, T_c 是模型输出的冠层辐射温度, T_s 是模型输出的第 1 层土壤辐射温度。f 是植被覆盖度, θ 表示传感器的观测天顶角。植被覆盖度表示为(Anderson 等, 2005):

$$f(\theta) = 1 - \exp(-0.5F/\cos(\theta)) \tag{20}$$

式中, F 是叶面积指数, 由 MODIS 的 LAI 产品获取。CLM 模型模拟的地表温度分为冠层和土壤温度, 所以式(19)可以写为:

$$T_{LST}(\theta) = [(f(\theta)\varepsilon_C T_{Creal}{}^4 + (1-f(\theta))\varepsilon_S T_{Sreal}{}^4)/\varepsilon_{CS}]^{1/4} \tag{21}$$

式中, T_{LST} 是模型模拟的地表混合温度, T_{Creal} 是模型输出的冠层温度, T_{Sreal} 是模型输出的第 1 层土壤温度, ε_{CS} 是站点的比辐射率, 参考 Wang(2008)的取值。ε_C, ε_S 分别是植被和土壤的比辐射率; 取值为 0.985 和 0.96(Coll 等, 2005; Jin 等, 2006)。

利用公式(21)作为同化系统的观测算子, 并把 MODIS 的地表温度与实测地表温度建立关系, 将均方差 RMSE 作为观测误差。该观测算子既可以应用于单点同化试验, 也可以应用在区域同化试验中。

为了检验 MODIS 温度产品的精度, 将地面实测的红外辐射温度订正为真实地表温度(Huang 等, 2005)与 MODIS 的地表温度进行比较和分析(图 3)。其中在 Brookings 站和 Bondville 站, 比辐射率取 0.987; 在 Blackhill 站点, 比辐射率取 0.993(Wang & Liang, 2008)。

从图 3 可以看出: 在 Brookings、Bondville 和 Blackhill 站白天的 MODIS 地表温度与地面实测温度的相关系数(R^2)分别是 0.96, 0.944, 0.968, 而夜间是 0.955, 0.851, 0.908。白天的相关性高于夜间。另外 MODIS 的地表温度比地面实测值低。在 Brookings 站, 高于 290K 时这种现象尤为明显。在 Brookings 站, MODIS 的地表温度白天精度(RMSE=2.58)比夜间(RMSE=2.11)低。在 Bondville 和 Baclkhill 站, MODIS 的地表温度白天精度(Bondville 站 RMSE=2.34, Blackhill 站点 RMSE=1.86)比夜间(Bondville 站点 RMSE=3.01, Blackhill 站点 RMSE=3.25)高。

当然, MODIS 的地表温度和实测地表温度的差异不仅与 MODIS 温度产品的反演精度有关, 而且

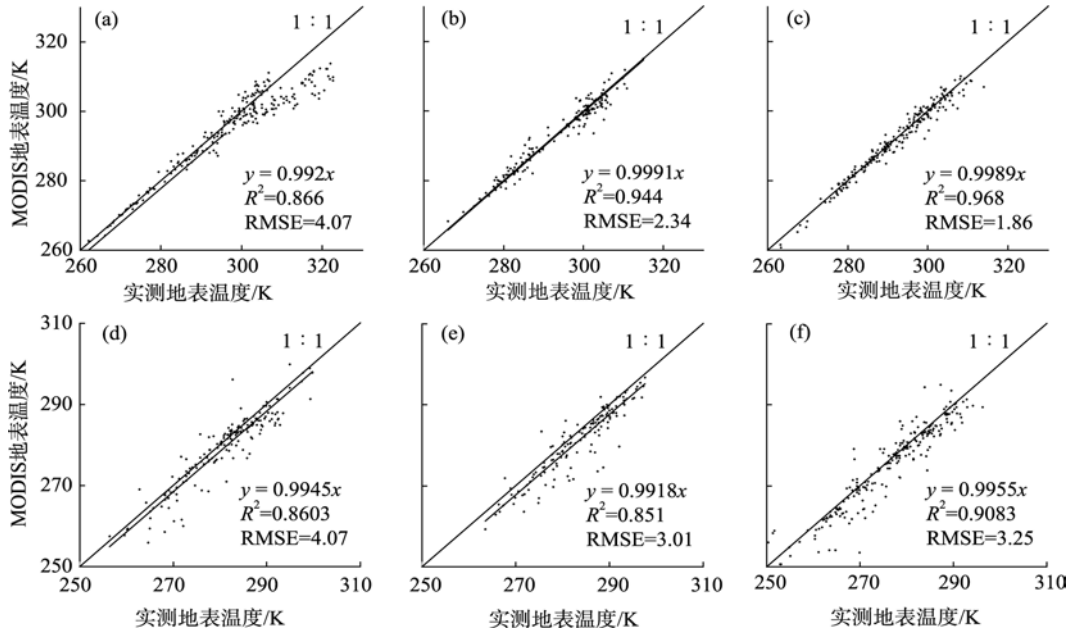

图 3 Brookings 站、 Bondville 站、Blackhill 站 MODIS 地表温度与实测地表温度的比较

(a)Brookings 站(白天); (b) Bondville 站(白天); (c) Blackhill 站(白天); (d) Brookings 站(夜间); (e)Bondville 站(夜间); (f) Blackhill 站(夜间)

还与两者的时间尺度差异(MODIS 地表温度是瞬时值, 而实测地表温度是 30min 平均值)和空间尺度差异(MODIS 地表温度的代表范围是 1km×1km, 而实测地表温度的代表范围约为几至十平方米左右)有关。

4.3 叶面积指数 LAI 对 CLM 模型模拟的地表温度和水热通量的影响

叶面积指数(LAI)影响地表对太阳辐射的吸收和反射。在 CLM 模型中, LAI 是一个重要的地表参数。CLM1.0 模型中, LAI 是凭借个人经验以及作物的生长规律进行输入的, 这样会造成较大误差, 本节将对 CLM 模型参数 LAI 进行敏感性分析。

利用 2006 年 Bondville 站的实测数据作为模型所需要的驱动数据和验证数据, 选取 LAI＝2.0 时的地表温度、显热通量和潜热通量为参考值, 选取 LAI 的变化量为 1.5, 分别讨论以下情况对地表温度、显热通量和潜热通量的影响: (1)LAI=0.5; (2)LAI=3.5; (3)LAI=5.0。

表 5 是 3 种条件下得到的地表温度、地表通量与参考值比较的均方差 RMSE。

表 5 LAI 敏感性分析

状态变量		LAI=0.5	LAI=3.5	LAI=5.0
地表温度/K	RMSE	2.72	1.71	2.73
显热通量/(W·m^{-2})	RMSE	7.39	5.58	9.45
潜热通量/(W·m^{-2})	RMSE	12.53	10.71	18.73

($\mathrm{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i-O_i)^2}$, N 为样本点个数, P_i 为计算值, O_i 为参考值)

由表 5 看出, LAI=0.5 时较 LAI=3.5 时地表温度、显热通量、潜热通量的 RMSE 变化幅度大, 说明当 LAI 较小时模型对它更加敏感。LAI 从 3.5 变化到 5.0 时, 地表温度、显热通量、潜热通量的 RMSE 逐渐变大。随着 LAI 的增大, RMSE 的变化逐渐变小。

从以上分析可以看出, CLM 中 LAI 的不精确输入对地表温度及地表水热通量的模拟有一定的影响。为了精确地描述 CLM 模型中 LAI 的变化, 本文将 MODIS 的 LAI 产品加入到模型中。

4.4 同化效果

本文选用了均方差(RMSE)来评价同化 MODIS 温度产品的效果。各个站的同化效果见表 6。

从表 6 中可看出: 与模拟结果相比, 同化后显热通量的估算精度在不同下垫面上均有提高: Blackhill 站(森林下垫面)改进最大, 均方差 RMSE 由 81.5W·m^{-2} 减小到 58.4W·m^{-2}, Bondville 站(农作物下垫面)RMSE 由 47.0W·m^{-2} 减小到 31.8W·m^{-2}, Brookings 站(草地下垫面)RMSE 由 46.5W·m^{-2} 减小到 45.1W·m^{-2}。而潜热通量估算精度在 Bondville、Blackhill 站有一定的提高: Bondville 站 RMSE 由 88.6W·m^{-2} 减小到 57.7W·m^{-2}, Blackhill 站 RMSE 由 53.4W·m^{-2} 减小到 47.2W·m^{-2}。

图 4—图 6 分别是 Brookings、Bondville 与 Blackhill 站的地表温度、显热通量和潜热通量模拟与同化结果的比较。考虑到夏季是植被生长旺盛期, 也是地表与大气之间能量和水分交换的旺盛期, 再考虑到通量观测资料的连续性, 图 4 选择 Brookings 站 2005 年的第 181 到 210 天, 图 5 选择 Bondville 站 2006 年的第 154 到 184 天, 图 6 选择 Blackhill 站 2005 年的第 228 到 258 天(5 到 7 月份通量资料缺失较多)。从图 4—图 6 中可以看出: Brookings、Bondville 与 Blackhill 站的地表温度和显热通量的同化效果较好。提高地表温度模拟效果的关键是 MODIS 温度产品的精度, 如在 Bondville 站(图 5), 第 164—168 天 MODIS 温度产品精度较高, 地表温度同化效果就比较好, 同样显热通量和潜热通量估算精度也有提高。而在第 155 天, 由于白天 MODIS 温度(303.78K, 309K)比实测地表温度(300.68K, 306.4K)偏高较多, 所以导致地表温度的同化结果偏高, 从而导致显热通量的估算结果也偏高。CLM 在森林下垫面(图 6)模拟地表温度效果不佳, 而同化 MODIS 温度产品可弥补这一缺点, 同时也提高了地表水热通量的估算精度。

表 6 Brookings、Bondville、Blackhill 站的同化结果

		Brookings		Bondville		Blackhill	
		显热通量	潜热通量	显热通量	潜热通量	显热通量	潜热通量
样本数		6517	8452	10348	9965	11813	11564
模型结果	RMSE/(W·m^{-2})	46.5	68.7	47.0	88.6	81.5	53.4
同化结果	RMSE/(W·m^{-2})	45.1	73.3	31.8	57.7	58.4	47.2

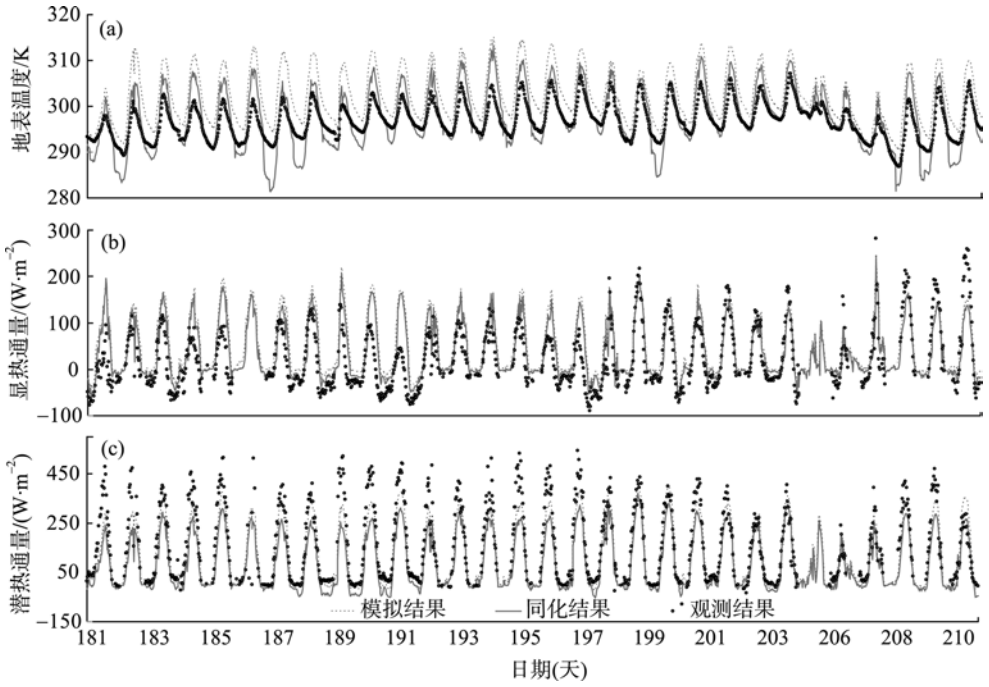

图 4 Brookings 站点的地表温度、地表水热通量的模型模拟和同化结果
(a)地表温度; (b)显热通量; (c)潜热通量

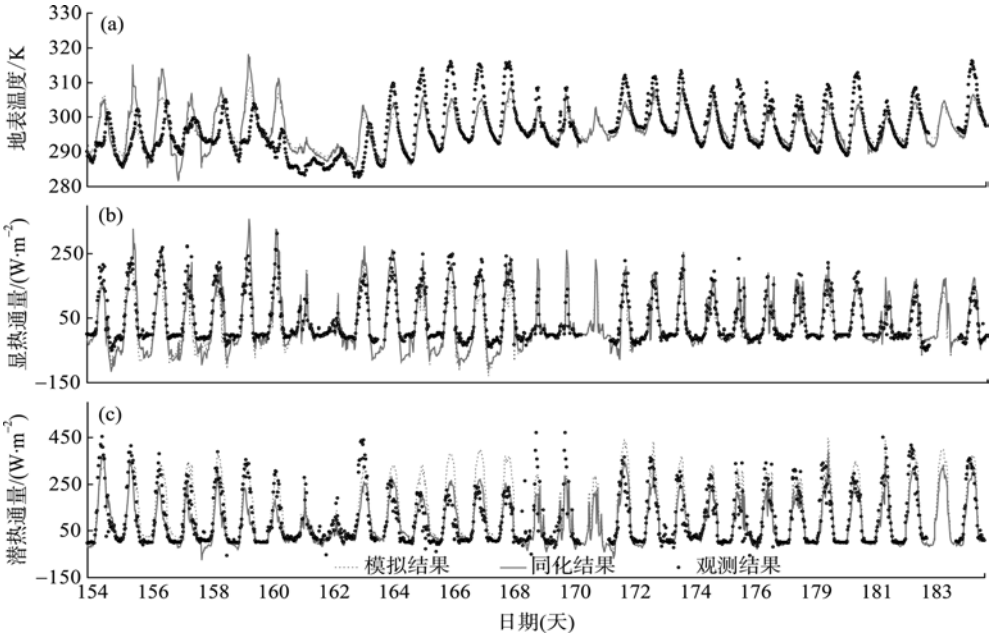

图 5 Bondville 站点的地表温度、地表水热通量的模型模拟和同化结果
(a)地表温度; (b)显热通量; (c)潜热通量

5 结论和讨论

本文发展了一个基于 CLM 模型和集合卡尔曼滤波的同化系统，将 MODIS 温度产品同化到该系统中，从而改进了地表温度、水热通量的模拟精度，同时将 MODIS LAI 产品应用到模型中，也提高了模型的模拟效果。

通过一年的数据分析发现：同化结果与模型的模拟结果相比，显热通量的估算精度在不同下垫面上均有改进。在 Blackhill 站(森林下垫面)，由于 CLM 模拟的地表温度与实测地表温度相差较大(白天相差 10K 左右)，而 MODIS 温度产品与实测地表温度的 RMSE 仅为 1.86K，所以同化结果中地表温度和显热通量的估算结果(与模拟结果相比 RMSE 由

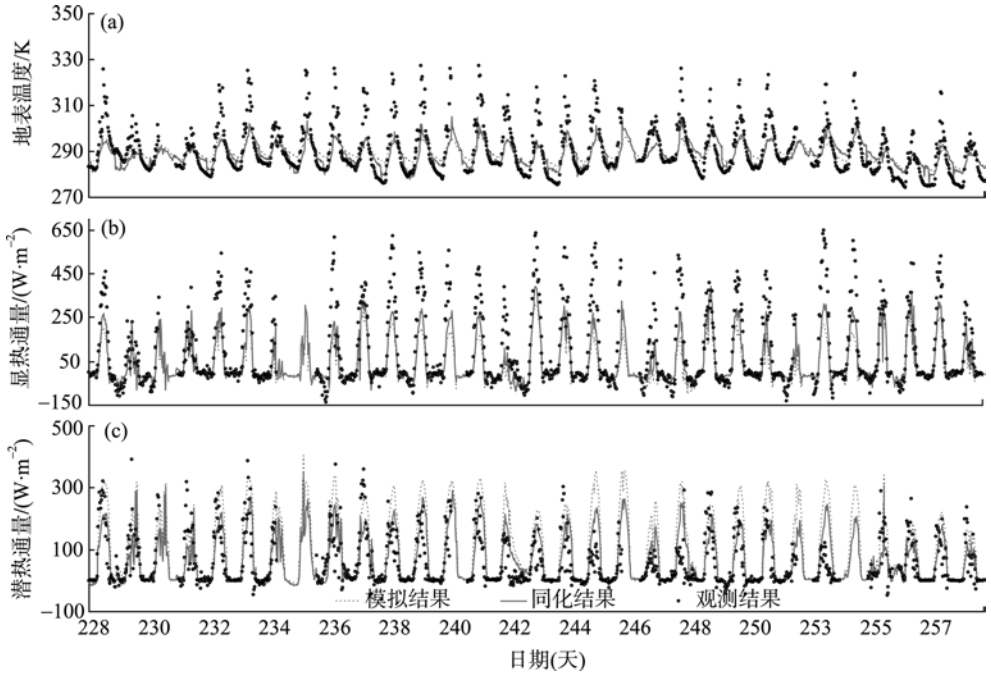

图 6 Blackhill 站点的地表温度、地表水热通量的模型模拟和同化结果
(a)地表温度; (b)显热通量; (c)潜热通量

81.5W·m^{-2}减小到 58.4W·m^{-2})有很大的改进，潜热通量的估算精度也有较大的提高(与模拟结果相比 RMSE 由 53.4W·m^{-2}减小到 47.2W·m^{-2})。在 Bondville 站(农作物下垫面)，显热通量同化结果与模拟结果相比，RMSE 由 47.0W·m^{-2}减小到 31.8W·m^{-2}，潜热通量 RMSE 由 88.6W·m^{-2}减小到 57.7W·m^{-2}。在 Brookings 站(草地下垫面)显热通量同化结果与模拟结果相比, RMSE 由 46.5W·m^{-2}减小到 45.1W·m^{-2}。

从同化结果可以看出，MODIS 温度产品的精度是同化的关键之一。MODIS 温度产品精度相对较高，且本文对 MODIS 温度产品进行了严格的质量控制(仅使用 QC=0 的值)。所以在大多数情况下，同化后的地表温度与模式模拟的地表温度相比精度有所提高。但有时 MODIS 温度产品与实测地表温度相比存在较大的偏差，MODIS 的地表温度和实测地表温度的差异不仅与 MODIS 温度产品的反演精度有关，而且与两者的时间尺度差异(MODIS 地表温度是瞬时值，而实测地表温度是 30min 平均值)和空间尺度差异(MODIS 地表温度的代表范围是 1km×1km，而实测地表温度的代表范围约为几至十平方米左右)有关，模式对地表温度的模拟也存在一定的系统偏差，导致同化后的地表温度与实测温度相比偏差较大。比如在 Bondville 站点第 155 天(图 5)，MODIS 温度产品比实测地表温度高 3K 左右，模式模拟的地表温度与实测地表温度相比也存在位相上的偏差，这样导致了白天同化的地表温度有时候比实测温度高 10K 左右。因此发展更精确的 MODIS 温度产品对于同化是十分有帮助的。在观测算子的确定方面，本文将 CLM 模型输出的地表温度与 MODIS 地表温度建立关系，其中用到了 MODIS 的 LAI 产品，这对于发展区域的二维同化系统有一定促进作用，但其适用性还需要进一步的验证。

CLM 模型所描述的过程比较复杂，需要许多的植被、土壤参数，而很多参数是很难准确地估算的，模型需要的气象驱动也存在一定的误差，这些使得模型在描述地表各变量时不够精确。此外，由于目前的陆面过程模式对地气温差的模拟有系统性的误差，而且地表温度的变化是一个快过程，对其同化尤其困难。因此下一步的工作需要针对这些特征，提出更好的同化策略。

由于观测站点显热通量、潜热通量数据是由涡动相关仪测量的，存在能量不闭合现象，一般约有 10%—25%左右(约 20—50W·m^{-2})的能量缺失(表 4)。如果对实测地表水热通量进行能量闭合订正，会改善地表水热通量的验证结果。此外模式模拟的地表温度是数公里范围的一个平均温度，而站点实测地表温度仅能代表几平米左右的范围，且仪器本身也存在一定的测量误差，这样会对地表温度验证造成一定影响。

在陆面数据同化系统中，模型误差和观测误差

的确定决定了同化系统结果向哪个方向靠近。如果模型误差较大，则同化结果向观测值靠近，反之，则向模型值靠近。但是模型误差和观测误差是很难确定的，它们会随着时间和空间而变化和发展。本文对它们均采用经验统计方法，将模型误差和观测误差看成固定值，这样会平滑了模型误差和观测误差，降低了同化结果的精度。因此发展新的模型误差和观测误差的估计方法可以提高同化结果的精度。

REFERENCES

Anderson M C and Norman J M. 2005. Effects of vegetation clumping on two-source model estimates of surface energy fluxes from an agricultural landscape during SMACEX. *Journal of Hydrometeorology*, **6**: 892—909

Bastiaanssen W G M, Menenti M, Feddes R A and Holtslag A A M. 1998. The surface energy balance algorithm for land (SEBAL): Part 1 Formulation. *Journal of Hydrology*, **212—213**: 198—212

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

Caparrini F, Castelli F and Entekhabi D. 2004. Estimation of surface turbulent fluxes through assimilation of radiometric surface temperature sequences. *Journal of Hydrometeorology*, **5**: 145—159

Coll C, Caselles V, Joan M, Valor E, Niclos R, Sanchez J M and Rivas R. 2005. Ground measurements for the validation of land surface temperatures derived from AATSR and MODIS data. *Remote Sensing of Environment*, **97**: 288—300

Crosson W L, Laymon C A, Inguva I and Schamschula M P. 2002. Assimilating remote sensing data in a surface flux–soil moisture model. *Hydrological Processes*, **16**: 1645—1662

Crow W T and Wood E F. 2003. The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble kalman filtering: a case study based on ESTAR measurements during SGP97. *Advances in Water Resources*, **26**: 137—149

Crow W T. 2006. Impact of incorrect model error assumptions on the sequential assimilation of remotely sensed surface soil moisture. *Journal of hydrology*, **7**: 421—431

Dai Y J, Zeng X B and Dickinson R E. 2001. Common Land Model (CLM): Technical Documentation and User's Guide. < http://219.142.99.7/models.do > (10 Oct, 2007)

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. 2003. The common land model. *Bulletin of the American Meteorological Society*, **84**: 1013—1023

Dickinson R E, Henderson S A and Kennedy P J. 1986. Biosphere–Atmosphere Transfer Scheme (BATS) for the NCAR Community Climate Model, NCAR Technical Note NCAR/TN-275CSTR, National Center for Atmospheric Research, Boulder, CO

Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research*, **99**: 10143—10162

Evensen G. 2003. The ensemble kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, **53**: 343—367

Galantowiez J F, Entekhabi D and Njoku E G. 1999. Test of sequential data assimilation for retrieving profile soil moisture and temperature from observed l-band radio brightness. *IEEE Transactions on Geoscience and Remote Sensing*, **37**: 1860—1870

Gao Z Q and Liu J Y. 2003. The comparison of land surface temperature with CLM and split window retrieving method. *Acta Geographica Sinica*, **58**(4): 494—502

Huang C L and Li X. 2006. Experiments of soil moisture data assimilation system based on ensenble kalman filter. *Plateau Meteorology*, **25**(4): 665—671

Huang C L, Li X, Lu L and Gu J. 2008a. Experiments of one-dimensional soil moisture assimilation system based on ensemble Kalman filter. *Remote Sensing of Environment*, **112**(3): 888—900

Huang C L, Li X and Lu L. 2008b. Retrieving soil temperature profile by assimilating MODIS LST products with ensemble Kalman filter. *Remote Sensing of Environment*, **112**(4): 1320—1336

Huang M F, Liu S M, Liu S H and Zhu Q J. 2005. A study of the difference between true surface temperature and radiometric surface temperature. *Advances in Earth Science*, **20**(10): 1075—1082

Jin M L and Liang S. 2006. An improved land surface emissivity parameter for land surface models using global remote sensing observations. *Journal of Climate*, **19**: 2867—2881

Kalman R E. 1960. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, **82** (Series D): 35—45

Kumar P and Kaleita A L. 2003. Assimilation of near-surface temperature using extended kalman filter. *Advances in Water Resources*, **26**: 79—93

Li X, Toshio K and Mahadevan P. 2004. A very fast simulated re-annealing (VFSA) approach for land data assimilation. *Computers &Geosciences*, **30**: 239—248

Liang S and Qin J. 2008. Advances in Land Remote Sensing: System, Modeling, Inversion and Applications. Springer Verlag

Liu S M, Hu G, Lu L and Mao D F. 2007. Estimation of regional evapotranspiration by TM/ETM+ data over heterogeneous surfaces. *Photogrammetric Engineering and Remote Sensing*, **73**(10): 1169—1178

Mathias G, Corinna R and Thomas F. 2004. A combination of quality assessment tools for eddy covariance measurements with footprint modelling for the characterisation of complex sites.

Agricultural and Forest Meteorology, **127**: 175—188

Qin J, Liang S L, Liu S M, Wang J D and Li X W. 2005. Estimation of land surface energy balance components by assimilating MODIS skin temperature into common land model using the ensemble kalman filter method. The 9th International Symposium on Physical Measurements and Signatures in Remote Sensing. Beijing: ISPRS

Reichle H R, Mclaughlin D B and Entekhabi D. 2001. Variational data assimilation of microwave radiobrightness observations for land Surface hydrology applications. *IEEE Transactions on Geoscience and Remote Sensing*, **39**: 1708—1718

Reichle H R, Mclaughlin D B and Entekhabi D. 2002. Hydrologic data assimilation with the ensemble kalman filter. *American Meteorological Society*, **130**: 103—114

Schuurmans J M, Troch P A, Veldhuizen A A, Bastiaanssen W G M and Bierkens M F P. 2003. Assimilation of remotely sensed latent heat flux in a distributed hydrological model. *Advances in Water Resources*, **26**: 151—159

Sellers P J, Randall D A, Collatz G J, Field C B, Dazlich D A, Zhang C, Collelo G D and Bounoua L.1996. A revised land surface parameterization (SiB2) for atmosphere GCMs, part I: model formulation. *Journal of Climate*, **9**: 676—705

Su Z. 2002. The surface energy balance system (SEBS) for estimation of turbulent heat fluxes. *Hydrology & Earth System Sciences*, **6**(1): 85—99

Wang W H, Liang S L and Tilden M. 2008. Validating MODIS land surface temperature products using long-term nighttime ground measurements. *Remote Sensing of Environment*, **112**(3): 623—635

Wilson K, Goldstein A, Falge E, Aubinet M, Baldocchi D, Berbigier P, Bernhofer C, Ceulemans R, Dolman H, Field C, Grelle A, Ibrom A, Law B E, Kowalski A, Meyers T, Moncrieff J, Monson R, Walter O, Tenhunen J, Valentini R and Verma S. 2002. Energy balance closure at FLUXNET sites. *Agricultural and Forest Meteorology*, **113**: 223—243

Yan Y, Liu Q H and Liu Q. 2006. Methodology of winter wheat yield prediction based on assimilation of remote sensing data with crop growth model. *Journal of Remote Sensing*, **10**(5): 804—811

Yang K, Takahiro W and 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. *Journal of the Meteorological Society of Japan*, **85**: 229—242

Zhang S W and Qiu C J. 2004. Estimating soil water contents from soil temperature measurements by using an adaptive kalman filter. *Journal of Applied meteorology*, **43**: 379—389

附中文参考文献

高志强, 刘纪远. 2003. 基于陆面模式和遥感技术的地表温度比较. 地理学报, **58**(4): 494—502

黄春林, 李新. 2006. 基于集合卡尔曼滤波的土壤水分同化试验. 高原气象, **25**(4): 665—671

黄妙分, 刘绍民, 刘素红, 朱启疆. 2005. 地表温度和地表辐射温度差值分析. 地球科学进展, **20**(10): 1075—1082

闫岩, 柳钦火, 刘强. 2006. 基于遥感数据与作物生长模型同化的冬小麦长势监测与评估方法研究. 遥感学报, **10**(5): 804—811