

Research Article

An Evapotranspiration Assimilation Method Based on Ensemble Kalman Filter and À Trous Wavelet

Shaohui Chen,¹ Jianwei Qi,² Xiaomin Sun,³ Xiangzheng Deng,⁴ and Jing Tian¹

¹ Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research (IGSNRR), Chinese Academy of Sciences (CAS), Beijing 100101, China

² China Aero Geophysical Survey & Remote Sensing Center for Land and Resources, Beijing 100101, China

³ Key Laboratory of Ecosystem Network Observation and Modeling, IGSNRR, CAS, Beijing 100101, China

⁴ Center for Chinese Agricultural Policy, IGSNRR, CAS, Beijing 100101, China

Correspondence should be addressed to Shaohui Chen; chensh@igsnr.ac.cn

Received 31 May 2013; Accepted 26 June 2013

Academic Editor: Hongbo Su

Copyright © 2013 Shaohui Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

It is challenging to assimilate the evapotranspiration product (EP) retrieved from satellite data into land surface models (LSMs). In this paper, a perturbed ensemble Kalman filter (PEKF) and à trous wavelet transform (AWT) integrated method are proposed to implement the evapotranspiration assimilation. In this method, the AWT is used to decompose the EPs into multiple channels since it is very powerful in fusing high frequency spatial information of multisource data, and then the Kalman filter is performed in the AWT domain. The proposed method combines the advantages of the PEKF that is capable of accommodating model error and observation error, and the AWT can effectively perform multiresolution fusion. Assimilation experiment conducted with the Noah model and the EP retrieved from the MODIS data shows that the proposed method performs better than the traditional ensemble Kalman filter (EnKF) and PEKF methods. The analysis results fit well with the evapotranspiration observation at two field sites with different land surface conditions. These indicate that the proposed method is promising for assimilating regional scale satellite retrieved EP into LSMs.

1. Introduction

Evapotranspiration (ET) is an important component of the water and energy exchanges between the atmosphere and land surface. It is crucial to accurately estimate ET for studying global or regional water and energy balances. Hence, good quality of spatial and temporal ET production (EP) can help to improve comprehension of water and energy cycle. However, this kind of EP is generally difficult to obtain in both dimensions of space and time because ET is influenced by many factors, such as air and skin temperatures, soil moisture, vegetation fraction, and horizontal advection. Up to now, there are two approaches to estimate the ET. One is site observations or remote sensing retrievals. Site observations have high spatial resolutions, but can only provide the EP for limited spatial locations [1]. Remote sensing retrievals have high spatial resolutions and can cover large range, but can only retrieve the instantaneous EP. The other is land

surface models (LSMs). LSMs are probably the most efficient approach for continuously estimating ET on a large range [1]. Because of the imperfection of the physics of LSMs and the uncertainties of input and driving data, the EP of the LSMs may contain significant errors. Hence, data assimilation (DA) has been applied to integrate observational ET into LSMs [2].

DA provides a framework for improving the LSMs by updating the state variables of the LSMs (SVLs) with observations and can combine the high spatial resolution of the observation with the high temporal resolution of the LSM. DA can be realized by two kinds of schemes: continuous assimilation and sequential assimilation. In continuous assimilation, the SVLs are modulated to be close to the observations. Continuous assimilation methods may cause the abrupt change of the SVLs before and after the DA, which will make the subsequent simulation of the LSMs to easily produce obvious errors. In sequential assimilation, the SVLs are updated according to some forecast principle.

The update is usually completed by adding the product of the gain that is estimated from model errors and observation errors and the difference between the observations and the outputs of the LSM to the SVLs. Sequential assimilation methods can produce a statistically optimal and dynamically consistent state estimate of the land surface by considering observation errors and model errors. Among sequential assimilation methods, Kalman filter- (KF-) derived methods yield accurate and consecutive ET estimate and have been applied widely in recent years [3].

Up to now, many KF-derived DA methods have been proposed. The KF method explicitly computes the error covariances through an additional matrix equation that propagates error information from one update time to the next, subject to possibly uncertain model dynamics [4, 5]. When the LSMs are linear, the KF method not only can get the optimal estimate of the SVLs, but also can estimate the error of the optimal estimate. However, the LSMs are usually highly nonlinear, and in order to apply the KF to the nonlinear LSMs, the KF method is extended to extended KF (EKF) method by using the first derivative of Taylor formula to linearize the LSMs. For the complicated, discontinuous, and nonlinear LSMs, the performance of these methods derived from the EKF is not very steady [6]. In order to overcome the limits of the KF- and EKF-based methods, Evensen [7, 8] proposed ensemble KF (EnKF) according to random dynamic forecast theory proposed by Epstein [9]. Burgers et al. [10] improved the EnKF to perturbed EnKF (PEKF) by randomly perturbing the SVLs, the observations, and the forcings using respective uncertainties. Because the DA methods derived from the EnKF and PEKF avoid the high dimensional nonlinear matrix operation, it is easy to integrate the DA methods into the LSMs. Hence, this branch of the DA methods has become an active research front in recent years.

Qin et al. [11] developed a variational data assimilation scheme based on the weak constraint concept. Actually, it assimilated surface skin temperature into a simple land surface model for estimation of ET. It is inconvenient to use the automatic differentiation technique which derived the adjoint codes to evaluate the gradient of the cost function when one had the ET retrievals. Pipunic et al. [12] confirmed that assimilating remotely sensed latent and sensible heat flux could potentially produce better heat flux predictions than assimilating soil moisture, or skin temperature observations from which they are derived. However, this experiment is only tested using synthetic data, and it needs further study about how the assimilation of ET using the actual remotely sensed ET is completed. Jang et al. [13] used MODIS data to calculate ET during clear sky conditions, while the MODIS-MM5 four-dimensional data assimilation system provided input variables for the calculation of ET under cloudy sky conditions, which means that the two types of ET are not merged in fact from the meaning of DA. Xu et al. [14] estimated turbulent fluxes through assimilation of geostationary operational environmental satellites data using the EnKF method. Though the use of geostationary operational environmental satellites data tackles the problem of remotely sensing data sparseness, the potential application of the experiment is constrained by its execution at site scale rather

than domain scale. French et al. [15] forecasted spatially distributed cotton ET by assimilating remotely sensed and ground-based observations. However, this method could not be feasible to provide continuous ET estimates that are better than can be achieved with either data alone if temporally continuous point observations are not available. However, more research is required to determine if and how well assimilation of ET can improve heat flux, soil moisture, and soil temperature predictions from the LSMs.

The main purpose of the DA is to combine the complementary information from measurements and models of the LSMs into an optimal estimate of the geophysical fields of interest [4]. The idea is similar to data fusion. By taking the fusion of low spatial and high spectral resolution multispectral data (LRMD) and high spatial and low spectral resolution panchromatic data (HRPD) as an example, the purpose of data fusion is to inject the details of the HRPD into the LRMD with the spectral properties of the LRMD reliably preserved. From the data fusion viewpoint, the KF, EnKF, and PEKF methods indirectly inject the high spatial information of the observations into the LSMs, and the indirect injection loses many details of the observations that cannot be extracted by the three methods. In data fusion, many studies have showed that à trous wavelet transform (AWT) is very powerful in injecting the details of the HRPD into the LRMD with the minimum influence to the latter.

The AWT was introduced by Holdschneider et al. in 2002 using an à trous (holes) algorithm and can preserve the translation invariance; that is, a translation of the original signal necessarily implies a translation of the corresponding wavelet coefficients [16]. The AWT is a computationally easy, dyadic, redundant, undecimated, nonorthogonal, and symmetric decomposition and provides good localization in both frequency and space domains by decomposing the data into multiple channels with the same size and decreasing resolutions [17]. These advantages make the AWT suitable for fusing multisource data. Therefore, it yields a better integration of the spatial and spectral quality than other methods. Recently, many fusion procedures based on the AWT have been proposed [18, 19].

The aim of using the AWT to ET assimilation is to improve ET prediction and hence to indirectly correct the heat flux predictions, meanwhile achieving physically correct soil moisture and temperature estimates through DA. Hence, this study tests the application of the AWT to assimilate remote sensed EP for producing better heat fluxes. Seldom papers could be found about the use of the AWT in DA [20–22] which shows promising results. However, more research is required to unearth the potential of the AWT to improve heat flux predications along with soil moisture and temperature predictions from the LSMs.

In Section 2, these advantages of the AWT are introduced into the DA. A novel assimilation method is proposed for improving the EP of the LSMs by means of the MODIS data-retrieved EP (MDRE), based on the joint use of the AWT and the PEKF. We use the AWT to decompose the MDRE and the EP of the LSMs into multiple channels and then perform the traditional PEKF in the AWT domain. The proposed AWT-PEKF method combines the advantages of the AWT

and the PEKF and can fully inject the details of the MDRE into the EP of the LSMs.

Experimental results conducted in Section 3 illustrate that there is consistency between the improvement analysis and the quality report of the assimilated EP. The performance of the AWT-PEKF method is quantified in improving the LSMs to predict the time varying EP. Intercomparisons are also made to show the advantage of the AWT-PEKF method over the conventional approaches based on the EnKF and PEKF. Estimation results indicate that the AWT-PEKF method can effectively retrieve the EP with a satisfactory accuracy. Section 4 concludes the paper.

2. Combined AWT-PEKF Data Assimilation Method

The PEKF can be understood as a purely statistical Monte Carlo method, where the ensemble of the SVLs evolves in state space with considering the mean of the ensemble as the best estimate and the spreading of the ensemble as the error variance [9]. The PEKF has gained popularity because of its simple conceptual formulation and relative ease of implementation; for example, it requires no derivation of a tangent linear or adjoint models and no integrations backward in time. Further, the computational requirements are affordable and comparable with other popular assimilation methods [9]. Hence, in this paper, the PEKF is adopted to assimilate the EP. The following briefly describes the PEKF.

Given the I SVLs of the LSMs at time $n - 1$, denoted by $X^{n-1} = (x_1^{n-1}, \dots, x_i^{n-1}, \dots, x_I^{n-1})^T$ where the superscript T denotes transpose, randomly perturb X^{n-1} K times using Gaussian noise to produce the ensemble $X_{(1,\dots,K)}^{n-1} = (X_1^{n-1}, \dots, X_k^{n-1}, \dots, X_K^{n-1})$. The PEKF consists of a forecast step:

$$X_{bk}^n = M(X_k^{n-1}), \quad 1 \leq k \leq K. \quad (1)$$

In (1), M denotes the nonlinear LSM, the subscript b denotes the background, and n denotes the simulation time.

The forecast error can be analyzed as

$$B^n H^T = \left\langle (X_b^n - \overline{X_b^n}) (H(X_b^n) - \overline{H(X_b^n)})^T \right\rangle, \quad (2)$$

$$HB^n H^T = \left\langle (H(X_b^n) - \overline{H(X_b^n)}) (H(X_b^n) - \overline{H(X_b^n)})^T \right\rangle, \quad (3)$$

$$\overline{X_b^n} = \frac{1}{K} \sum_{k=1}^K X_{bk}^n, \quad (4)$$

$$X_{b(1,\dots,K)}^n = (M(X_1^{n-1}), \dots, M(X_k^{n-1}), \dots, M(X_K^{n-1})), \quad (5)$$

$$\overline{H(X_b^n)} = \frac{1}{K} \sum_{k=1}^K H(X_{bk}^n), \quad (6)$$

$$H(X_{b(1,\dots,K)}^n) = (H(X_{b1}^{n-1}), \dots, H(X_{bk}^{n-1}), \dots, H(X_{bK}^{n-1})), \quad (7)$$

where H denotes the observation operator. Then the gain matrix G is obtained as

$$G = B^n H^T (HB^n H^T + O^n)^{-1}. \quad (8)$$

In (8), O denotes the observation error matrix.

Finally, the assimilation is performed as

$$X_a^n = X_b^n + G(Y_o^n - H(X_b^n)). \quad (9)$$

In (9), Y_o^n denotes the observation at time n .

The PEKF method uses the ensemble to describe the covariance matrix of the SVLs and avoids the explicit calculation of covariance matrix needed in the EKF. A modest misspecification of the initial ensemble normally does not influence the results very much over time. The PEKF allows for a wide range of noise models and one is not restricted to using Gaussian distributed noise.

It can be found from (9) that the assimilation is completed based on single grid and does not consider the interrelation between grids. The interrelation between grids is mainly reflected by the detail and textural information. The reason that the spatial resolution of the observation is higher than the simulation of the LSM is that the observation has more detail and textural information than the simulation of the LSM. The assimilation scheme based on single grid cannot fully inject the detail and textural information of the observation into the simulation of the LSM. The main purpose of the DA is to integrate the high accuracy of the observation and the temporal evolution of the LSMs together.

From the perspective of data fusion, the assimilation process can be considered as constructing one coefficient with both the same temporal response as the simulation of the LSM and the same spatial response as the observation at a particular grid location. With the development of the AWT, we expect much room for improvement over the traditional single-grid-based assimilation scheme to merge the observation and the simulation of the LSM in the AWT domain since the wavelet filter can consider the interrelation between the grids during the detail extraction of the observation.

The AWT can provide good localization in both frequency and space domains in terms of decomposing the data with finite energy into multiple channels, each one of them with a different degree of resolution [23]. Because of its shift invariance property, the AWT has been successfully used for fusing the data with different resolutions [24]. The AWT representation of the data can be obtained by using à trous algorithm [17].

This algorithm consists basically in the application of consecutive convolutions between the data under analysis and a scaling function at distinct degradation levels [25]. One of the most widely used scaling functions for the execution of the à trous algorithm is the b_3 -spline [23] as follows:

$$\frac{1}{256} \begin{pmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{pmatrix}. \quad (10)$$

If the original data is represented by D , the wavelet coefficients, C_j , for the level j are obtained by the difference between two consecutive degraded data, D_{j+1} and D_j , as shown in the following equation:

$$C_j = D_j - D_{j+1}. \quad (11)$$

To carry out the data synthesis from a degradation level J , an additive criterion should be applied in which all the coefficients obtained are added to the last degradation level of the original data, as shown in the following equation:

$$D = D_J + \sum_{j=1}^J C_j. \quad (12)$$

By manipulating independently the approximation components and the wavelet planes of the data, the AWT shows excellent performance in integrating the spectral information of the multispectral data and the spatial information of the panchromatic data. For the DA, the purpose is to integrate the high spatial resolution information of the observation into the simulation of the LSMs. Hence, the purposes of the data fusion and the DA have similarity.

It can be seen from (9) that the update is completed by adding the product of the gain and the difference to the SVLs. The high spatial information of the observation used to improve the SVLs is contained in the difference and is modulated into the SVLs by the gain. From the perspective of the AWT-based fusion, the additive DA cannot consider the difference between the wavelet planes. For different wavelet planes, the gain may be different and is varied with the wavelet plane. It will be beneficial to complete the DA in the AWT domain. What is more, the additive DA cannot consider the interrelation between the grids, which leads to the under injection of the detail information of the observation. With the AWT, much room for improvement over the traditional DA can be expected.

After the forecast of the LSM using the ensemble, see (1). Apply the AWT to the forecast ensemble and the observation as

$$X_{bk}^n = \sum_{j=0}^J X_{bkj}^n, \quad (13)$$

$$Y_o^n = \sum_{j=0}^J Y_{oj}^n. \quad (14)$$

In (13), X_{bk0}^n denotes the approximation component of X_{bk}^n , and $\{X_{bkj}^n\}_{j=1}^J$ denotes the J wavelet planes of X_{bk}^n . In (14), Y_{o0}^n denotes the approximation component of Y_o^n , and $\{Y_{oj}^n\}_{j=1}^J$ denotes the J wavelet planes of Y_o^n .

Apply the AWT to H as

$$H(X_{bk}^n) = \sum_{j=0}^J H(X_{bkj}^n). \quad (15)$$

In (15), $H(X_{bk0}^n)$ denotes the approximation component of $H(X_{bk}^n)$, and $\{H(X_{bkj}^n)\}_{j=1}^J$ denotes the J wavelet planes of $H(X_{bk}^n)$.

Then, (2)–(4) and (6) can be, respectively, rewritten as:

$$B_j^n H_j^T = \left\langle \left(X_{bj}^n - \overline{X_{bj}^n} \right) \left(H(X_{bj}^n) - \overline{H(X_{bj}^n)} \right)^T \right\rangle, \quad (16)$$

$$H_j B_j^n H_j^T = \left\langle \left(H(X_{bj}^n) - \overline{H(X_{bj}^n)} \right) \left(H(X_{bj}^n) - \overline{H(X_{bj}^n)} \right)^T \right\rangle, \quad (17)$$

$$\overline{X_{bj}^n} = \frac{1}{K} \sum_{k=1}^K X_{bkj}^n, \quad (18)$$

$$\overline{H(X_{bj}^n)} = \frac{1}{K} \sum_{k=1}^K H(X_{bkj}^n). \quad (19)$$

Then the gain matrix G_j at decomposition level j is obtained as

$$G_j = B_j^n H_j^T \left(H_j B_j^n H_j^T + O_j^n \right)^{-1}. \quad (20)$$

In (20), O_j^n denotes the observation error covariance matrix at decomposition level j at time n . O_j^n can be obtained using the perturbed observations [10].

Finally, the assimilation is performed as

$$X_a^n = X_b^n + \sum_{j=0}^J G_j \left(Y_{oj}^n - H(X_{bj}^n) \right). \quad (21)$$

Compared with (9), equation (21) has two obvious advantages. First, in (9), the execution of the assimilation is completed grid by grid, and the relationship between grids is not considered; in (21), the interrelation between the local grids is considered and is quantified by filtering the grids equivalent to the length of the filter during the decomposition process. Second, in (9), the same gain is used to extract the new information absent in the SVLs from the observation; in (21), the gain is calculated in the AWT domain and is obtained according to the wavelet planes. It is more effective to update the SVLs using wavelet plane varying gains according to the observation.

For assimilating the MDRE into the EP of the LSM, the DA procedure based on the AWT-PEKF method comprises the following steps (Figure 1):

- (1) run the LSM in open time loop;
- (2) given the Gaussian probability density function, produce the ensemble of the SVLs according to their probability characteristic. Run the LSM during ensemble loop;
- (3) integrate the LSM to get the simulation result over the simulation range, namely, complete range loop;
- (4) end ensemble loop. Obtain the mean EP by averaging the simulated EPs as

$$EP_0^n = \sum_{k=1}^K EP_k^n. \quad (22)$$

In (22), K denotes the size of the ensemble of the SVLs. EP_0^n denotes the mean EP at the current simulation time n ;

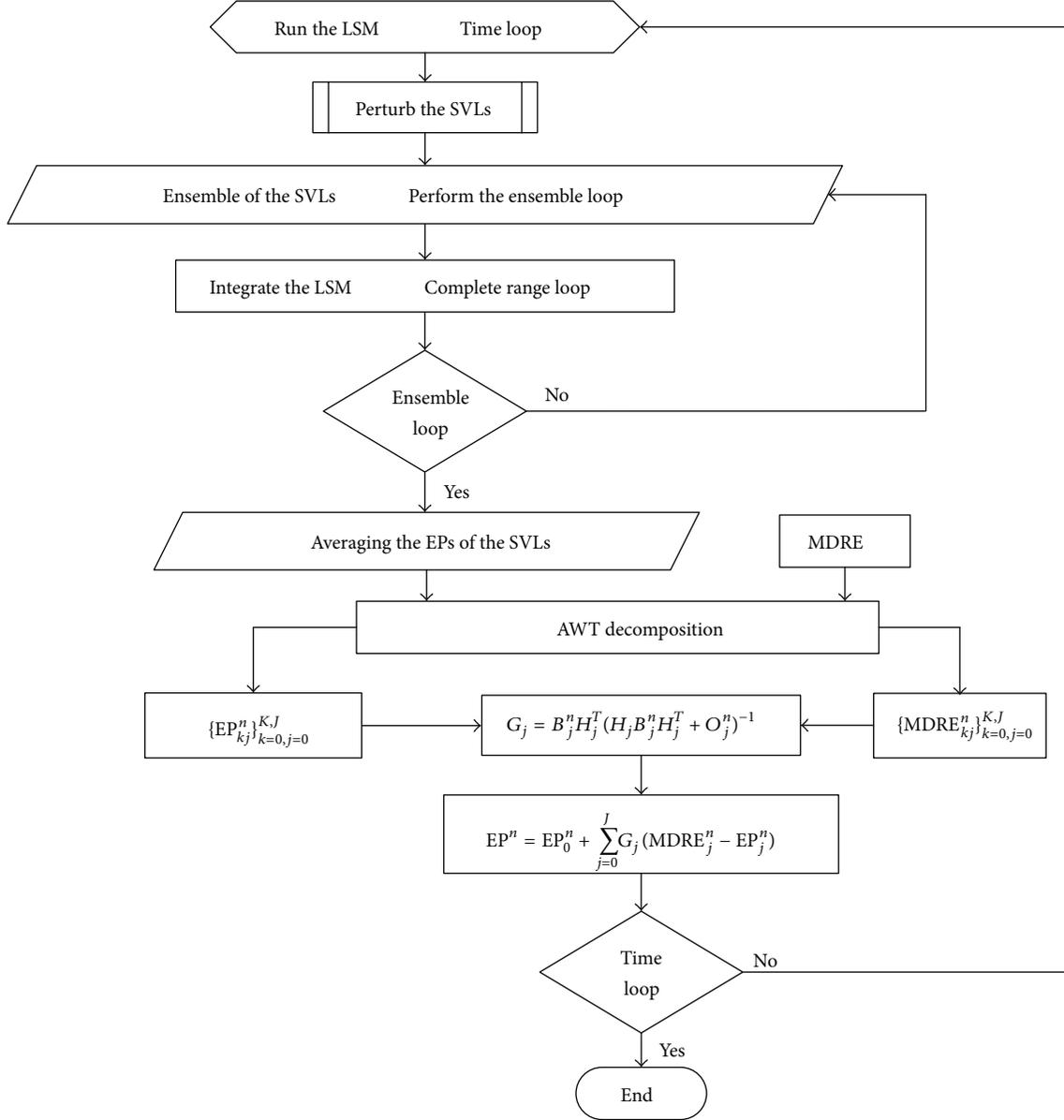


FIGURE 1: The schematic flowchart of the AWT-PEKF method.

- (5) perturb the MDRE K times and produce the ensemble of the MDRE, denoted by $\{MDRE_k^n\}_{k=0}^K$ [9]. $MDRE_0^n$ denotes the original MDRE;
- (6) apply the AWT on the MDRE and EP. Respectively, obtain the decomposition results as

$$EP_k^n = \sum_{j=0}^J EP_{kj}^n, \quad 0 \leq k \leq K, \quad (23)$$

$$MDRE_k^n = \sum_{j=0}^J MDRE_{kj}^n, \quad 0 \leq k \leq K. \quad (24)$$

In (23) and (24), J is the number of the AWT decomposition levels;

- (7) calculate the gain at decomposition level j , G_j , using the decomposition results as (25):

$$G_j = B_j^n H_j^T (H_j B_j^n H_j^T + O_j^n)^{-1}, \quad (25)$$

$$B_j^n H_j^T = \langle (EP_j^n - EP_{0j}^n) (EP_j^n - EP_{0j}^n)^T \rangle, \quad (26)$$

$$H_j B_j^n H_j^T = \langle (EP_j^n - EP_{0j}^n) (EP_j^n - EP_{0j}^n)^T \rangle. \quad (27)$$

In (25)–(27), the observation operator, H , is 1 because it does not need transformation between the assimilated variable and the observation;

(8) innovate the EP of the LSM as

$$EP^n = EP_0^n + \sum_{j=0}^J G_j (MDRE_j^n - EP_j^n); \quad (28)$$

(9) integrate the LSM until ending open time loop.

3. Experimental Section

In this section, the AWT-PEKF method is tested through assimilating the EP of the Noah model with the MDRE retrieved using one operational two-layer method (OTLM) [26], and the assimilated result is validated with the site observations from two field stations in Chinese Ecosystem Research Network (Cern), each of which has a different underlying land cover condition. The Noah model is developed by the environmental modeling center of national center for environmental prediction [27]. In addition, we also compare the assimilated results with those of the EnKF and PEKF methods.

3.1. Site Observations. Cern was established in 1988 for studying the problems related with Chinese ecology. The network provides continuous observations of ecosystem level exchanges of biogeochemical, water, energy, and momentum at diurnal, monthly, seasonal, and annual time scales. Cern currently comprises forty-six sites, which distribute over the whole of China. The data of the site observations used in this study are obtained from the Yucheng and Lawn sites. At the two sites, ET has been continuously measured by eddy correlation system which is composed of a 3D sonic anemometer and an open path CO₂/H₂O analyzer since 2001. The ET data that were obtained by averaging the original data of 10 Hz over 30 min are used in the experiment as the validating data.

The Yucheng site was installed on Huabei plain of China. The field site is located near the Yucheng county, Shandong province, China, and the geographical coordinate is approximate 36.96°N, 116.63°E, and the altitude is 20 m above sea level. The climate of the area is warm and semi-humid with annual mean temperature 13.1°C, mean precipitation 600 mm per year, and the average monthly relative humidity 66.44%. The soil is mainly classified as sandy-clay loam and sandy loam. The characteristic of this site is typical of agricultural plot. The farm has been continuously alternated between wheat and corn. Lawn site was temporally installed about 54 kilometers from the Yucheng site in 2005. The latitude and longitude are 36.46°N and 116.13°E. The underlying vegetation of this site is grass. The leaf area index is about 2.0. The average canopy height is about 20–40 cm. The area is 10 km². Figure 2 shows the schematic positions of the two sites, and the research range is mentioned in the following section.

The two sites represent the main land covers of the studied area. The underlying surfaces of the two sites are relatively homogeneous. Therefore, the observation data of the ET from the two sites can represent the ET statuses of the two homogeneous areas around the two sites. The two areas

correspond at least to nine grids of the MDRE and the EP of the Noah model at the spatial resolution of 1 km. Hence, it is sufficient to use these observations as the validating data.

3.2. The MDRE. In order to retrieve the MDRE, the OTLM proposed by Zhang et al. [26] is used. In the OTLM, pixel component arranging and comparing algorithm (PCACA) and layered energy-separating algorithm (LESA) are two key procedures. The two procedures are used to partition mixed surface temperature and mixed surface albedo according to vegetation fraction to obtain those of bare soil and vegetation. Then, net radiation of bare soil and vegetation is calculated using the partitioned results. Finally, the MDRE is retrieved with a spatial resolution of 1 km.

The OTLM has several advantages. First, the PCACA is based on vegetation fraction and ground temperature trapezoid relation theory and initiates a new method of decomposing mixed pixels. Second, it is very convenient because only single angle remote sensing data are required which can be obtained from most of the satellite sensors. Third, the core of the LESA is Bowen-ratio energy balance method, which reduces the uncertainties in surface energy partition based on the Beer law. Fourth, key nonremote sensing parameters that influence regional ET can also be obtained by using the OTLM. The OTLM has been successfully applied to Huabei area, China, [28]. Hence, the OTLM is used to retrieve the MDRE. Detailed descriptions about the OTLM can be found in [26, 28].

3.3. The Temporal Extension of the MDRE. Because the MODIS is easily influenced by the clouds, the data are not good enough to retrieve the MDRE every day. Figure 3 shows the available MDREs retrieved from the Aqua data at 14:30 local solar time in May, 2005. The unit for Figure 3, the following mentioned figures, and the site observation is W/m².

On the contrary, the better the assimilated results are, the more there are the MDREs. Hence, it first requires to temporally extend the MDRE by the available MDREs and the EPs of the Noah model. The series of the extended MDRE is then used as the synthetic observations. In order to extend the MDRE, an intensity modulation method is employed because this method is very simple and can be easily implemented into the Noah model to perform high speed real-time MDRE extension for a long temporal span. This method can interpolate the absent MDRE with the fidelity to the evolvment trend of the Noah EP. This unavailable MDRE_{*n*} at the simulation time *n* is obtained as

$$MDRE_n = \sum_{s=1}^S \left(\frac{|s-n|}{\sum_{t=1}^S |t-n|} \frac{EP_n MDRE_s}{EP_s} \right). \quad (29)$$

In (29), *s*, *n*, and *t* denote the time. {MDRE_{*s*}}_{1≤*s*≤*S*} is the set of the available good MDREs, and *S* is the size of the set. EP_{*s*} is the Noah EP corresponding to the MDRE_{*s*}, and EP_{*n*} is the Noah EP at the simulation time *n*. *s* - *n* describes the departure of the EP_{*s*} from the EP_{*n*} and characterizes the temporal difference of the EP_{*s*} from the EP_{*n*}, while *t* - *n* describes the departure of the MDRE_{*t*} from the MDRE_{*n*}, and

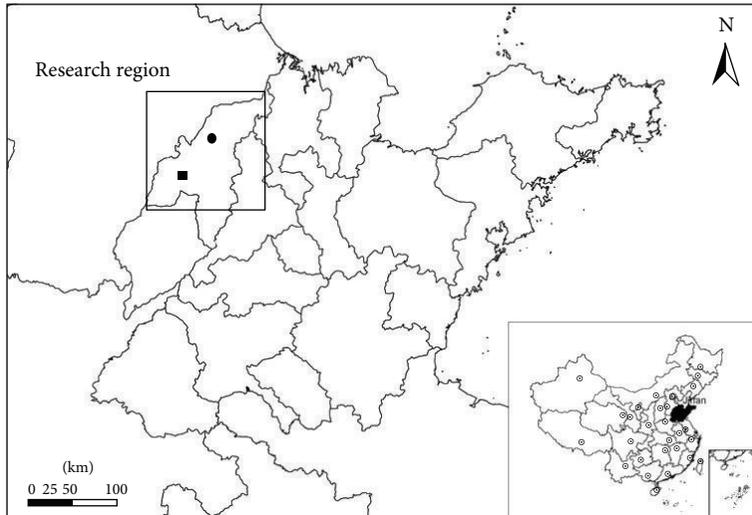


FIGURE 2: The schematic positions of Yucheng site (●), Lawn site (■), and research range (□).

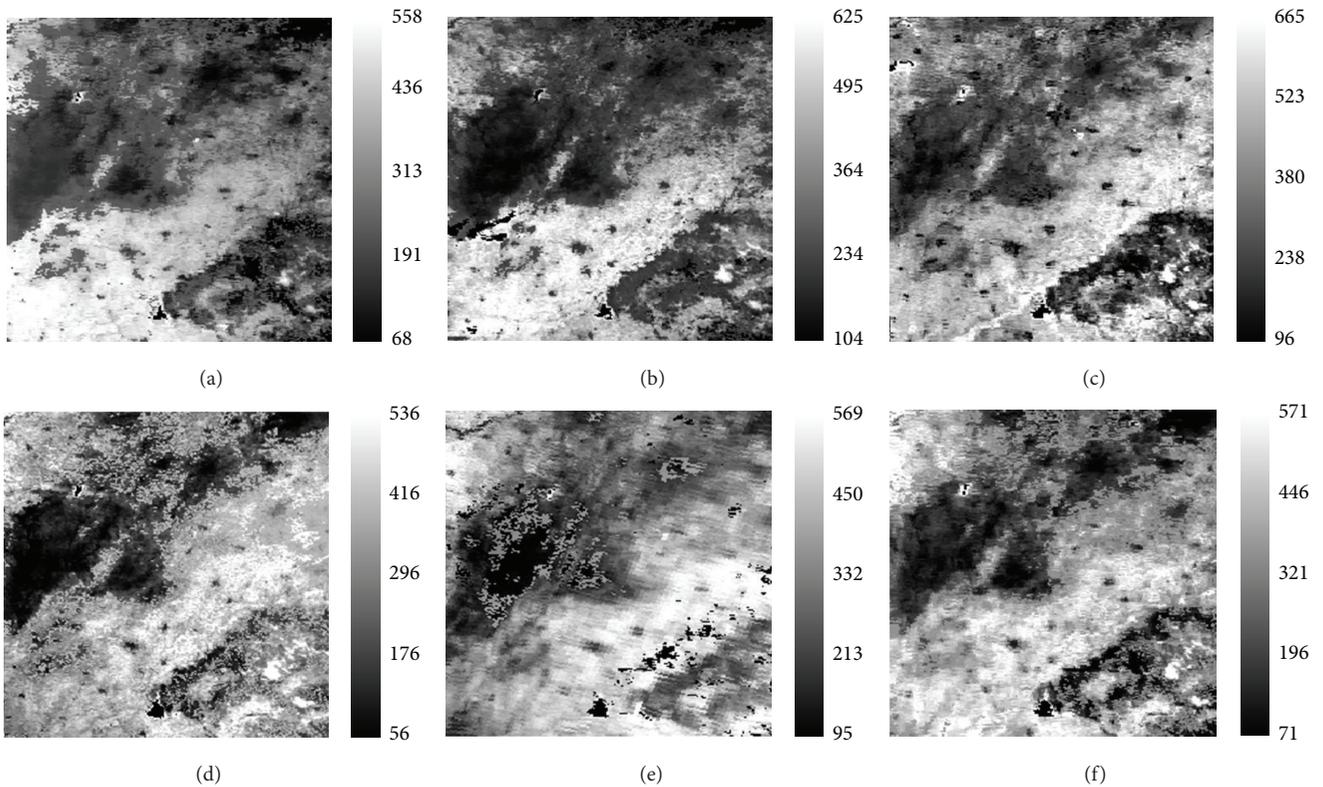


FIGURE 3: The available MDREs retrieved from the Aqua data at 14:30 local solar time in May, 2005. (a) The MDRE on May 2, 2005; (b) the MDRE on May 7, 2005; (c) the MDRE on May 14, 2005; (d) the MDRE on May 19, 2005; (e) the MDRE on May 25, 2005; (f) the MDRE on May 31, 2005.

quantifies the contribution of the $MDRE_t$ to the $MDRE_n$. By using a ratio between EP_n and EP_s , spatial details of $MDRE_s$ are modulated to the predicted $MDRE_n$ without altering its spatial properties and contrast.

3.4. The Noah Model. The Noah model is originated from a physically based land surface-vegetation-atmosphere-transfer scheme. During the past ten years, it underwent

substantial upgrades, including modifications to the formulations of canopy conductance, bare soil evaporation, vegetation phenology, ground heat flux, and so forth. These model enhancements significantly improve its performance, and are physically more faithful to nature and thus most likely the route for more improvements in the future [29, 30].

The Noah model was chosen for assimilation test for several reasons. It can simulate many states of the land

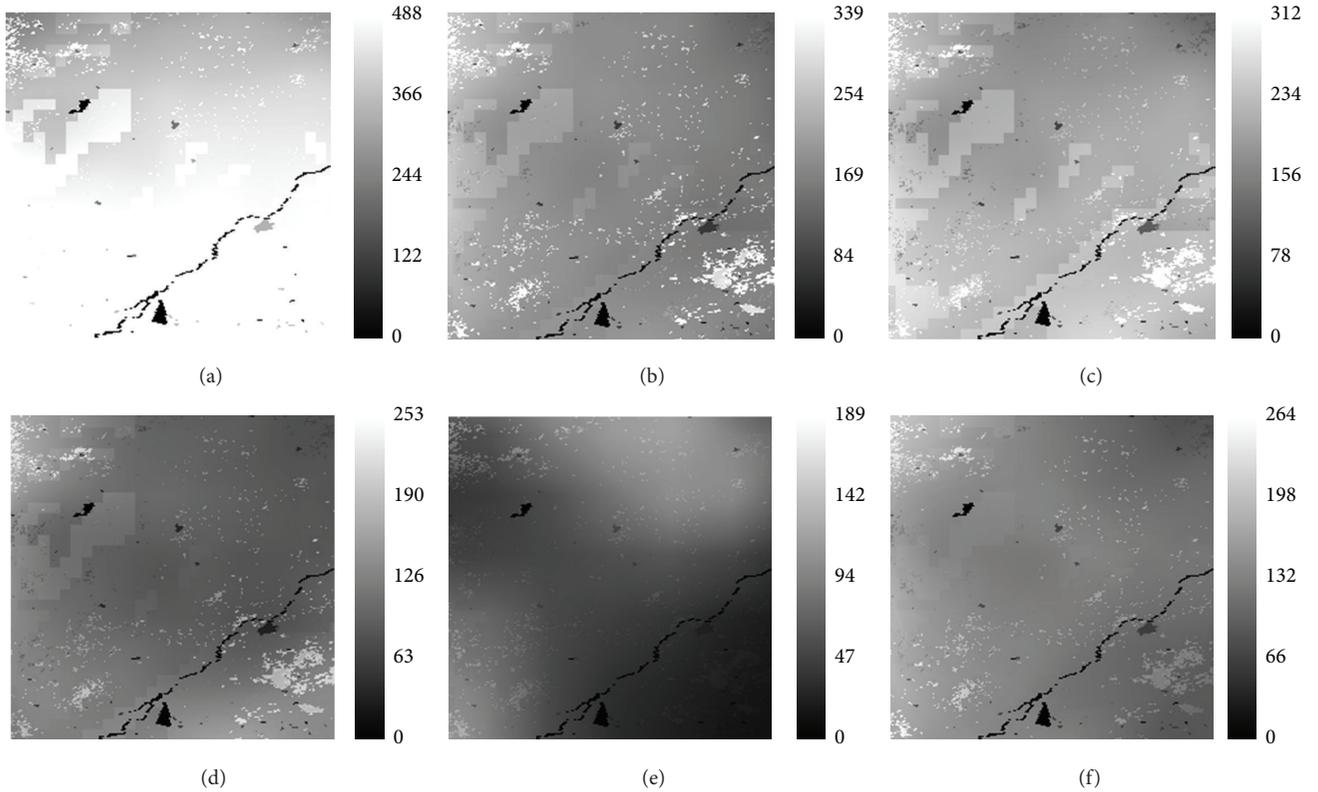


FIGURE 4: The EPs of the Noah model at 15:00 solar time closest to the available MDREs in Figure 3. (a) The EP on May 2, 2005; (b) the EP on May 7, 2005; (c) the EP on May 14, 2005; (d) the EP on May 19, 2005; (e) the EP on May 25, 2005; (f) the EP on May 31, 2005.

surface, including soil temperature, skin temperature, and the energy and water fluxes of the land surface energy and water balances. In various coupled and uncoupled assessments, it has been proven to have the ability to reproduce the observed land surface energy, and water budgets effectively [29, 30]. Because of its comprehensive nature, the model yields more output variables. The model is updated periodically on the NCEP website (<ftp://ftp.emc.ncep.noaa.gov/mmb/gcp/ldas/noahsm/>, accessed on April 8, 2013) and version 2.7.1 is used in this experiment (release date: February 9, 2005).

The Noah model contains four vertical soil layers: a thin 10 cm top layer, a second root zone layer of 30 cm, a deep root zone of 60 cm, and a subroot zone of 100 cm. It can be run for 13 vegetation covers (2 of which use the same parameter values) and nine different soil types (two of which also use the same parameters). It has 33 parameters: 10 related to vegetation and 23 that describe soil properties. It also has 16 initial states (when run with four root layers). The model uses a local greenness fraction from the normalized difference vegetation index (NDVI) to establish seasonality in the model for each of the 13 vegetation types.

In this assimilation experiment, the Noah model was configured to have dimensions of 250×250 at $1 \text{ km} \times 1 \text{ km}$ spanning a domain bounded by 35.77°N to 38.26°N , 114.81°E to 117.3°E . Time step is 10800 seconds. Assimilation time span is between 1 January, 2005 and 31 May, 2005 after the Noah model is run from 1 June, 2004 because the MODIS data in this span are less affected by clouds, and it is beneficial

for retrieving the MDRE. On this grid, the elevation was derived from the 1 km digital elevation of the GTOPO30 database [31]. The vegetation classification was derived from the global, 1 km, Advanced Very High Resolution Radiometer (AVHRR) based, 13 class vegetation database [32]. The 9-class soil texture data were derived from the top layer of the 1 km, 11-layer soil dataset, state soil geographic database [33]. The monthly vegetation cover fraction data are taken from the satellite-based AVHRR 5-year global monthly climatology of green vegetation fraction [34]. Six EPs of the Noah model at 15:00 solar time corresponding to the nearest time of the six MDREs are shown in Figure 4.

3.5. Results. The data from the two sites are first used to evaluate the ET values of the available MDREs and the Noah model at the corresponding locations before DA. For validation, we utilize root-mean-square error (RMSE) as the estimation index. RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{1}{M} \sum_{m=1}^M (o - d)^2}. \quad (30)$$

In (30), o and d , respectively, denote the observation and validated data. M denotes the number of the data. The unit for the RMSE is W/m^2 in this paper. The estimation results are, respectively, shown in Table 1.

It can be found from Table 1 that the Noah EP has large RMSE values than the MDREs compared with the two site

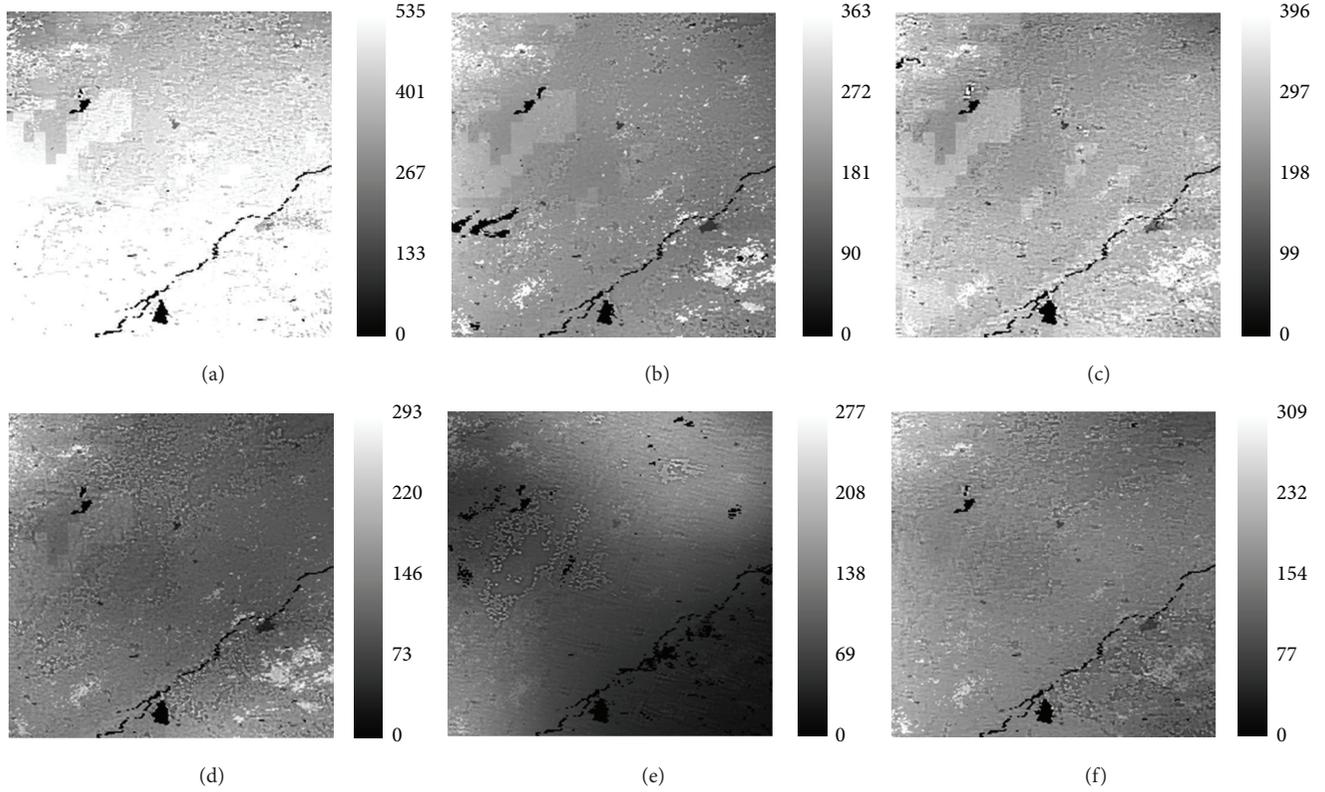


FIGURE 5: The assimilated EPs from the EnKF method at 15:00 solar time. (a) The assimilated EP on May 2, 2005; (b) the assimilated EP on May 7, 2005; (c) the assimilated EP on May 14, 2005; (d) the assimilated EP on May 19, 2005; (e) the assimilated EP on May 25, 2005; (f) the assimilated EP on May 31, 2005.

observations. The main reason is that the spatial resolution of the MDRE is higher than that of the Noah EP, and more details can be found from the MDRE by comparing Figures 3 and 4. The configuration of the Noah model to the spatial resolution of the MDRE did not produce new information in the EP because of the coarse resolution of the driving data. It can be seen from Table 1 that the EP can be improved significantly with the MDREs through DA.

When performing the AWT-PEKF method, the SVLs including skin temperature, soil temperature and volumetric liquid soil moisture for the first soil layer, canopy water content, and the MDRE are perturbed fifty times according to their error Gaussian distributions, respectively. The four SVLs are perturbed because they are the key variables in the calculation of the ET in the Noah model. For comparisons, the EnKF, and PEKF methods are also performed. The assimilated EPs from the EnKF, PEKF, and proposed methods are, respectively, shown in Figures 5, 6, and 7.

In order to evaluate the three methods, the two site observations are employed to validate the DA results by the RMSE. Table 2 and Figure 8 show the results.

The RMSE reveals the accurate degree of the EPs produced by each method. The lower the RMSE is, the better the assimilation effect is, and vice versa. It can be seen from Table 2 and Figure 8 that all methods improve significantly the accuracy of the EPs compared with the EPs without assimilation: the RMSE values are reduced from 51.58

TABLE 1: The RMSE results of the MDRE and the Noah EP using the two site observations.

	Yucheng site		Lawn site	
	MDRE	EP	MDRE	EP
RMSE	18.74	51.58	23.47	68.69

TABLE 2: The RMSE results of the three methods using the two site observations.

	EnKF	PEKF	AWT-PEKF
Yucheng site	28.75	22.54	16.81
Lawn site	25.18	21.97	18.16

and 68.69 in Table 1 to 28.75, 22.54, and 16.81 and 25.18, 21.97, and 18.16, respectively, for the Yucheng and Lawn sites. The three methods allow an obvious accuracy improvement of the EPs when increasing the spatial details of the EPs. Table 2 and Figure 8 show that the RMSE values of the AWT-PEKF method are the lowest among the three methods. Therefore, we can draw a conclusion that the EPs of the AWT-PEKF method have the least error. The EPs of the PEKF method have slight error. The EPs of the EnKF method have the largest error. This is expected because the ET equation used in the Noah model is a highly nonlinear equation and the perturbation only for the SVLs used in the EnKF can only

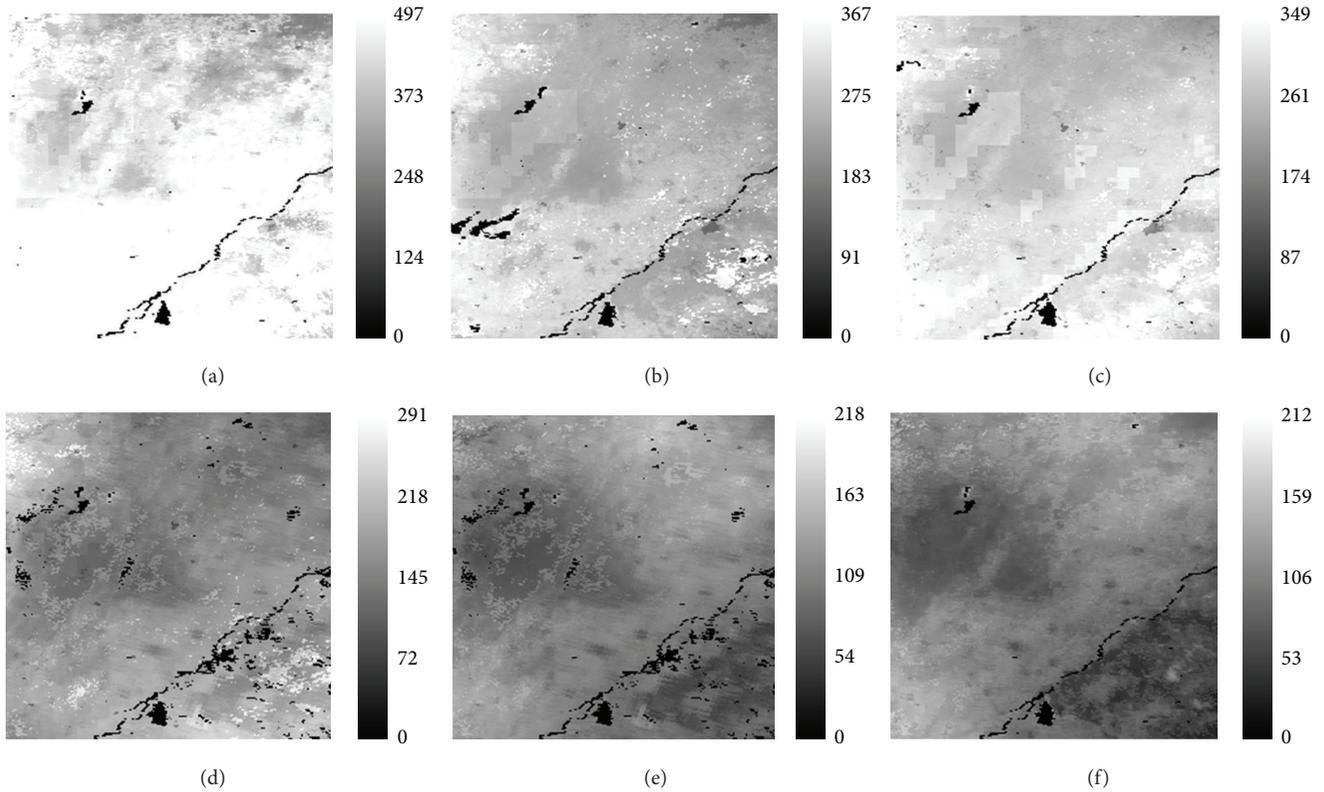


FIGURE 6: The assimilated EPs from the PEKF method at 15:00 solar time. (a) The assimilated EP on May 2, 2005; (b) the assimilated EP on May 7, 2005; (c) the assimilated EP on May 14, 2005; (d) the assimilated EP on May 19, 2005; (e) the assimilated EP on May 25, 2005; (f) the assimilated EP on May 31, 2005.

propagate analytically with finite precision, which introduces large errors in the EP estimation and leads to suboptimal performance.

As seen from Tables 1 and 2, the AWT-PEKF method produces the best assimilation effect. This is probably because the AWT-based assimilation scheme does good to improve the EP simulation of the Noah model. The AWT-PEKF method gets the advantage of the EnKF and PEKF methods, because the novel assimilation scheme injects the spatial details of the MDRE into the Noah EP by taking into account both the interrelationship between grids, as is the case of the AWT decomposition process, and ensemble-based innovation, as is the case of the PEKF forecast. The two procedures allow the AWT-PEKF method to produce the EPs closer to the ET observation. One can therefore conclude that the Noah model can predict the EP relatively accurate using the AWT-PEKF method.

By combining the quantitative estimation results and the intercomparison, it can be seen that the AWT-PEKF method gives the EPs closer to the measured EPs than the EnKF and PEKF methods when the assimilated EPs are compared with the observations from the Yucheng and Lawn sites.

3.6. Discussion. Though the AWT-PEKF method outperforms the EnKF and PEKF methods in the experiment, three points are needed to be studied further. The first is

that the validating data are sparse. In validation, there are only two available field sites. Because it is difficult to get the ET observation, it needs to confirm whether or not the assimilated EPs obtained by the AWT-PEKF method are also close to the ET observation at other locations. The second is that it needs to test whether or not the AWT-PEKF method is also effective in assimilating other variables. Other variables, such as skin temperature and soil moisture, are forecasted differently from the ET and are also influenced by many factors. Hence, the extension of the AWT-PEKF method to other variables is a big job. The third is that the assimilation of the ET cannot influence the consecutive simulation of the Noah model.

As for the first point, the potential solution is to perform the AWT-PEKF method in the area where the field sites are available, meanwhile the MDREs can be retrieved from cloud-free MODIS data. As for the second point, the AWT-PEKF method can be easily extended to assimilate skin temperature and soil moisture products only if the two products retrieved from the MODIS data are prepared in advance. As for the third point, in the following study, we will introduce another novel assimilation method, in which the SVLs relative with the ET are simultaneously updated in order to transfer the assimilation effect into the consecutive simulation of the LSM. The assimilation idea will be presented in another research paper.

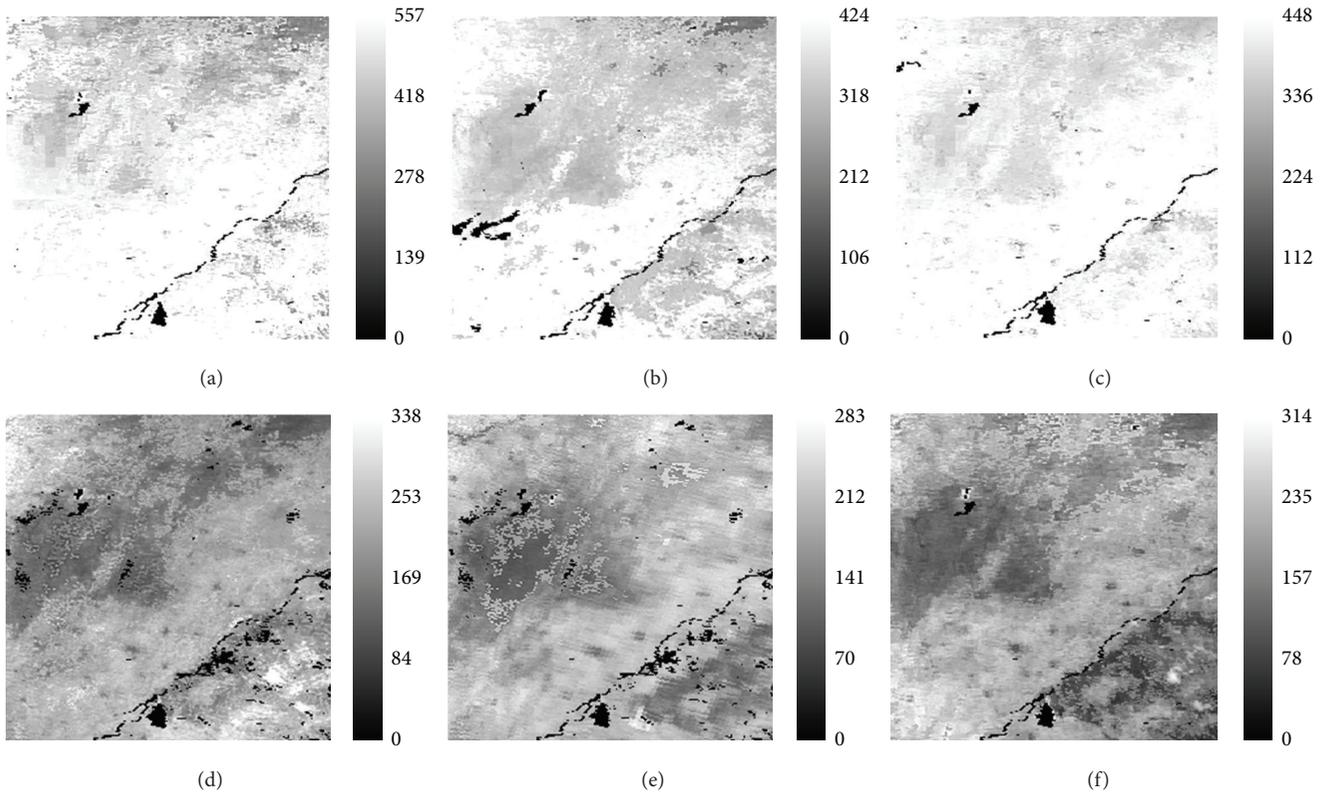


FIGURE 7: The assimilated EPs from the AWT-PEKF method at 15:00 solar time. (a) The assimilated EP on May 2, 2005; (b) the assimilated EP on May 7, 2005; (c) the assimilated EP on May 14, 2005; (d) the assimilated EP on May 19, 2005; (e) the assimilated EP on May 25, 2005; (f) the assimilated EP on May 31, 2005.

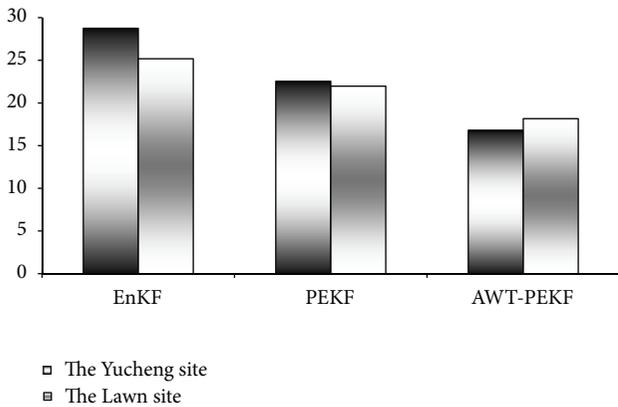


FIGURE 8: The RMSE histograms of the three methods using the two site observations.

4. Conclusions

In this paper, we study the hybrid use of the AWT and PEKF methods for assimilating the MDRE into the EP of the Noah model in order to improve the consecutive simulation of the Noah model. The AWT is used to decompose the MDRE for injecting its detail information represented by wavelet planes, while the PEKF is used to complete the assimilation by the model and observation uncertainties.

The AWT-PEKF method retains the respective advantages of the AWT and PEKF. Firstly, it is based on multigrid, and the interrelation between grids is considered using the wavelet filter during the filtering procedure. Secondly, according to the gain derived using the PEKF from the model and observation uncertainties, the details and textures of the MDRE are modulated into the EPs of the Noah model using the AWT from image fusion viewpoint.

The performance of the proposed method is compared with those of the EnKF and PEKF methods using one assimilation experiment. Intercomparison results of the RMSE confirm the effectiveness of the AWT-PEKF method in improving the spatial accuracy of the EPs. Overall evaluation shows that the AWT-PEKF method is promising and superior to the traditional EnKF and PEKF methods. Several issues are unresolved, such as the validation of the assimilated results, the effects of ensemble size, initial perturbation fields on assimilated results, and the actual performance of this new method in real other variable assimilations. These aspects require further investigation.

Acknowledgments

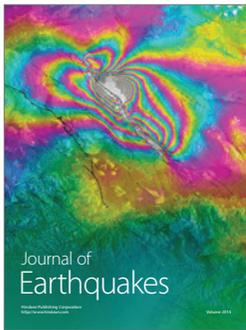
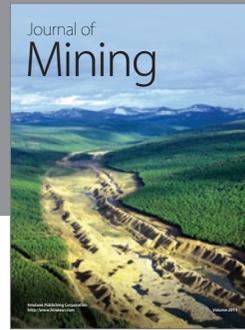
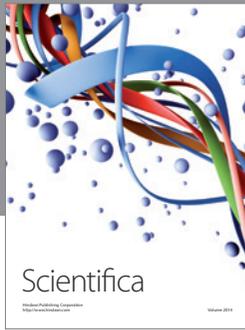
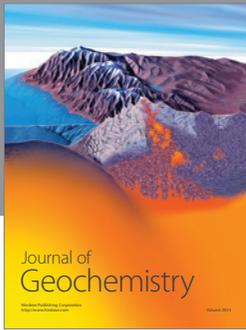
The authors thank the anonymous reviewers for their sincere suggestions which helped to improve the paper. This work is supported jointly by the Project of Natural Science Fund

of China (41101329) and National Basic Research Program of China (2010CB950904 and 2010CB428403).

References

- [1] J. Qin, S. Liang, K. Yang, I. Kaihotsu, R. Liu, and T. Koike, "Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal," *Journal of Geophysical Research D*, vol. 114, no. 15, Article ID D15103, 2009.
- [2] C. L. Meng, Z.-L. Li, X. Zhan, J. C. Shi, and C. Y. Liu, "Land surface temperature data assimilation and its impact on evapotranspiration estimates from the common land model," *Water Resources Research*, vol. 45, no. 2, Article ID W02421, 2009.
- [3] X. Tian, Z. Xie, and A. Dai, "An ensemble-based explicit four-dimensional variational assimilation method," *Journal of Geophysical Research*, vol. 113, article D21124, 2008.
- [4] R. H. Reichle, "Data assimilation methods in the Earth sciences," *Advances in Water Resources*, vol. 31, no. 11, pp. 1411–1418, 2008.
- [5] J. P. Pinto, M. C. Bernardino, and A. P. Silva, "A Kalman filter application to a spectral wave model," *Nonlinear Processes in Geophysics*, vol. 12, no. 6, pp. 775–782, 2005.
- [6] X. Li, T. Koike, and M. Pathmathevan, "A very fast simulated re-annealing (VFSA) approach for land data assimilation," *Computers and Geosciences*, vol. 30, no. 3, pp. 239–248, 2004.
- [7] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *Journal of Geophysical Research*, vol. 99, no. 5, pp. 10–162, 1994.
- [8] G. Evensen, "The Ensemble Kalman Filter: theoretical formulation and practical implementation," *Ocean Dynamics*, vol. 53, pp. 343–367, 2003.
- [9] E. S. Epstein, "Stochastic dynamic prediction," *Tellus*, vol. 21, pp. 739–759, 1969.
- [10] G. Burgers, P. J. Van Leeuwen, and G. Evensen, "Analysis scheme in the ensemble Kalman filter," *Monthly Weather Review*, vol. 126, no. 6, pp. 1719–1724, 1998.
- [11] J. Qin, R. Liu, S. Liang, H. Zhang, and B. Hu, "A new method based on remote sensing observations and data assimilation for estimation of evapotranspiration over field crops," *New Zealand Journal of Agricultural Research*, vol. 50, no. 5, pp. 997–1004, 2007.
- [12] R. C. Pipunic, J. P. Walker, and A. Western, "Assimilation of remotely sensed data for improved latent and sensible heat flux prediction: a comparative synthetic study," *Remote Sensing of Environment*, vol. 112, no. 4, pp. 1295–1305, 2008.
- [13] K. Jang, S. Kang, J. Kim et al., "Mapping evapotranspiration using MODIS and MM5 Four-Dimensional Data Assimilation," *Remote Sensing of Environment*, vol. 114, no. 3, pp. 657–673, 2010.
- [14] T. Xu, S. Liang, and S. Liu, "Estimating turbulent fluxes through assimilation of geostationary operational environmental satellites data using ensemble Kalman filter," *Journal of Geophysical Research D*, vol. 116, no. 9, Article ID D09109, 2011.
- [15] A. French, D. Hunsaker, and T. Clarke, "forecasting spatially distributed cotton evapotranspiration by assimilating remotely sensed and ground-based observations," *Journal of Irrigation and Drainage Engineering*, vol. 138, no. 11, pp. 984–992, 2012.
- [16] M. Holdschneider, R. Kronland-martinet, J. Morlet, and P. Tchamitchian, "A real time algorithm for signal analysis with the help of the wavelet transform," in *Proceedings of the International Conference on Time-Frequency Methods and Phase Space*, J. M. Combes, A. Grossman, and P. Tchamitchian, Eds., p. 286, Springer, Marseille, France, 1987.
- [17] P. Dutilleul, "An implementation of the "algorithme à trous" to compute the wavelet transform," in *Wavelet. Proceedings of the International Conference on Time-Frequency Methods and Phase Space*, J. M. Combes, A. Grossman, and P. Tchamitchian, Eds., pp. 298–304, Springer, Marseille, France, 1987.
- [18] T.-M. Tu, W.-C. Cheng, C.-P. Chang, P. S. Huang, and J.-C. Chang, "Best tradeoff for high-resolution image fusion to preserve spatial details and minimize color distortion," *IEEE Geoscience and Remote Sensing Letters*, vol. 4, no. 2, pp. 302–306, 2007.
- [19] S. Chen, R. Zhang, H. Su, J. Tian, and J. Xia, "SAR and multispectral image fusion using generalized IHS transform based on trous wavelet and EMD decompositions," *IEEE Sensors Journal*, vol. 10, no. 3, pp. 737–745, 2010.
- [20] A. Deckmyn and L. Berre, "A wavelet approach to representing background error covariances in a limited-area model," *Monthly Weather Review*, vol. 133, no. 5, pp. 1279–1294, 2005.
- [21] O. Pannekoucke, L. Berre, and G. Desroziers, "Filtering properties of wavelets for local background-error correlations," *Quarterly Journal of the Royal Meteorological Society*, vol. 133, no. 623, pp. 363–379, 2007.
- [22] A. Tangborn, "Wavelet approximation of error covariance propagation in data assimilation," *Tellus*, vol. 56, no. 1, pp. 16–28, 2004.
- [23] J. Núñez, X. Otazu, and O. Fors, "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 3, pp. 1204–1211, 1999.
- [24] Y. Chibani and A. A. Houacine, "Redundant versus orthogonal wavelet decomposition for multisensor image fusion," *Pattern Recognition*, vol. 36, no. 4, pp. 879–887, 2002.
- [25] M. González-Audiciana, X. Otazu, O. Fors, and A. Seco, "Comparison between the Mallat's and the à trous discrete wavelet transform based algorithms for the fusion of multispectral and panchromatic images," *International Journal of Remote Sensing*, vol. 26, no. 3, pp. 595–614, 2005.
- [26] R. Zhang, J. Tian, H. Su, X. Sun, S. Chen, and J. Xia, "Two improvements of an operational two-layer model for terrestrial surface heat flux retrieval," *Sensors*, vol. 8, no. 10, pp. 6165–6187, 2008.
- [27] F. Chen and J. Dudhia, "Coupling and advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system—part I: model implementation and sensitivity," *Monthly Weather Review*, vol. 129, no. 4, pp. 569–585, 2001.
- [28] J. Tian, R. Zhang, X. Sun, Z. Zhu, and Y. Zhou, "Study of a model for correcting the effects of horizontal advection on surface fluxes measurement based on remote sensing," *Science in China*, vol. 49, no. 2, pp. 273–280, 2006.
- [29] M. B. Ek, K. E. Mitchell, Y. Lin et al., "Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model," *Journal of Geophysical Research D*, vol. 108, no. 22, pp. 1–16, 2003.
- [30] K. E. Mitchell, D. Lohmann, and P. R. Houser, "The multi-institution North American Land Data Assimilation System (NLDAS): utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system," *Journal of Geophysical Research*, vol. 109, 2004.

- [31] K. L. Verdin and S. K. Greenlee, "Development of continental scale digital elevation models and extraction of hydrographic features," in *Proceedings of 3rd International Conference/Workshop on Integrating GIS and Environmental Modeling, National Center for Geographic Information and Analysis*, pp. 21–26, Santa Fe, NM, USA, 1996.
- [32] M. C. Hansen, R. S. Defries, J. R. G. Townshend, and R. Sohlberg, "Global land cover classification at 1 km spatial resolution using a classification tree approach," *International Journal of Remote Sensing*, vol. 21, no. 6-7, pp. 1331–1364, 2000.
- [33] D. A. Miller and R. A. A. White, "Conterminous United States multi-layer soil characteristics data set for regional climate and hydrology modeling," *Earth Inter*, vol. 2, no. 2, 1998.
- [34] G. Gutman and A. Ignatov, "The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models," *International Journal of Remote Sensing*, vol. 19, no. 8, pp. 1533–1543, 1998.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

