

## Research Article

# Soil Moisture Assimilation Using a Modified Ensemble Transform Kalman Filter Based on Station Observations in the Hai River Basin

Guocan Wu,<sup>1,2</sup> Bo Dan,<sup>1,2</sup> and Xiaogu Zheng<sup>3</sup>

<sup>1</sup>College of Global Change and Earth System Science, Beijing Normal University, Beijing 100875, China

<sup>2</sup>Joint Center for Global Change Studies, Beijing 100875, China

<sup>3</sup>Key Laboratory of Regional Climate-Environment Research for East Asia, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

Correspondence should be addressed to Bo Dan; danb2012@mail.bnu.edu.cn

Received 25 March 2016; Revised 31 May 2016; Accepted 10 July 2016

Academic Editor: Shaoqing Zhang

Copyright © 2016 Guocan Wu 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.

Assimilating observations to a land surface model can further improve soil moisture estimation accuracy. However, assimilation results largely rely on forecast error and generally cannot maintain a water budget balance. In this study, shallow soil moisture observations are assimilated into Common Land Model (CoLM) to estimate the soil moisture in different layers. A proposed forecast error inflation and water balance constraint are adopted in the Ensemble Transform Kalman Filter to reduce the analysis error and water budget residuals. The assimilation results indicate that the analysis error is reduced and the water imbalance is mitigated with this approach.

## 1. Introduction

As a key component of land underlying surface, soil plays an essential role in the exchange of energy, heat, water vapor and carbon dioxide between the land surface and atmosphere [1]. Soil moisture is an important land surface variable that can directly affect vegetation respiration, transpiration, and various chemical reactions. These actions will ultimately result in the redistribution of surface energy and water by changing the surface albedo, soil heat capacity, surface evaporation, and vegetation growth [2, 3].

Many studies have investigated soil moisture assimilation in different locations and on different time scales. For instance, a study assimilating the near-surface observations focused on the effect of ensemble size and observation depth [4]. The dual-based data assimilation system for assimilating satellite observations of soil moisture reproduces the temporal evolution of daily soil moisture, especially under freezing conditions [5]. The assimilation of satellite data to a land surface model can further improve soil moisture estimation

accuracy [6–9], which can also interactively or simultaneously estimate model parameters and surface variables [10, 11].

Generally speaking, soil moisture estimation is mainly from land surface model simulation and observation. Land surface models (e.g., Common Land Model, CoLM) [12–15], which use observations of surface meteorological conditions and precipitation, can characterize the exchange of energy and water within the surface layer. Soil moisture simulations from land surface models can provide temporally and spatially continuous series, but limited to the specification of the forcing variables, uncertain model parameters describing the exchange of energy and water with the atmosphere and models, and imperfect model physical processes [16].

The advantage of observation is that it can obtain the value of the observed variable on the represented time and space [17]. However, in situ soil moisture observations can only represent a very spatially limited range and cannot meet the needs of regional studies due to the heterogeneity of surface variables [18, 19]. Satellite remote sensing data are

only roughly available at shallow layers and cannot provide complete soil moisture profiles [20, 21].

Based on observations and model simulations, land surface data assimilation can provide an analysis of soil moisture with less uncertainty and better coverage. The assimilation results are largely limited to meteorological forcing error, which is one of the error sources in model simulation. Therefore, forcing data with long term and high precision are very important for improving the simulation capability of land surface models [22–24]. Many studies use reanalysis data directly or corrected reanalysis data with monthly observations [25]. However, observations from approximately 200 stations were only offered by China for international exchange to construct such datasets. Therefore, their quality over China mainland is questionable. Recently, a new China mainland forcing field (including near-surface air temperature, relative humidity, surface pressure, wind speed, precipitation, downward shortwave radiation, and longwave radiation) is constructed by fusing hourly observation, reanalysis, and remote sensing data [26], which is evaluated in this study.

On the other hand, the assimilation results crucially rely on the estimation accuracy of the forecast and observation error covariance matrices for any assimilation methods [27, 28]. The Ensemble Transform Kalman Filter (ETKF) is a commonly used assimilation method, that has been widely studied and applied in atmospheric science research [29, 30]. In ETKF, the forecast perturbations are transformed into analysis perturbations by multiplying a transformation matrix and the eigenvector decomposition of a matrix of the ensemble size is used to construct the transform matrix [31, 32]. Past research on ensemble based land data assimilation methods have found that the ensemble spread decreases during model integration because the ensemble forecasts use the same boundary and meteorological forcing. In ETKF, the forecast error is initially estimated as the perturbed forecast states minus their ensemble mean. The sampling errors in such estimations, resulting from the limited ensemble size as well as the poor initial perturbations and model error, can generally lead to an underestimate of the forecast error covariance matrix and can eventually result in filter divergence [33, 34]. Therefore, using inflation techniques to address the underestimation of forecast error becomes increasingly important [35].

Basically, there are two approaches for estimating the multiplicative inflation factors based on the innovation statistic: the moment estimation [36, 37] and the maximum likelihood estimation [38–40] which is applied in this study. Liang et al. [38] compared the two methods using Lorenz-96 model as a test bed and found out that the maximum likelihood approach leads to a smaller analysis error than the moment estimation approach (in their Figure 3). This is a reason for us to use maximum likelihood estimation in this study. Moreover, the  $-2$  log-likelihood can be used as the cost function to estimate both inflation factor and threshold layer, while this cannot be done if the moment estimation is applied. The effect of forecast error inflation is investigated for reducing the analysis error.

In addition to reducing the analysis error in land surface data assimilation, obtaining a balanced or closed water

budget is extremely important for analyzing the terrestrial hydrological cycle. The terrestrial water budget simulated by land surface models (LSMs) is a key process of the global hydrologic cycle and an important part of the physical consistency [41]. The physical consistency (i.e., closure of the water budgets) can be maintained by constructing water balance constraint in land surface models [42]. A balanced or closed water budget is important for estimating runoff, developing and validating hydrological models, and better understanding land-atmosphere water exchange processes [43]. However, state updating in land surface data assimilation generally degrades the water budget balance, because the analysis increment compensates for system biases or errors but does not conserve water. If the degree of water imbalance is excessive, it is reasonable to develop a data assimilation system that can remove or reduce the imbalance of water. Therefore, the water balance constraint technique is often adopted in the assimilation procedure to compensate for this defect. Pan and Wood [42] derived a two-stage constrained Kalman filter solution in which the first stage is a traditional Kalman filter and the second stage imposes a water balance constraint in an optimal manner. Yilmaz et al. [43, 44] introduced a weak constraint solution to the conventional data assimilation systems to reduce the water budget imbalance. In this study, the water balance constraint is adopted for the ETKF assimilation scheme with forecast error inflation. The effect of the water balance constraint is investigated for reducing the water budget residual.

As a case study, the ETKF assimilation scheme with a water balance constraint is applied to the Automatic Weather Station (AWS) located in the Daxing district of Beijing ( $39^{\circ}37'N$ ,  $116^{\circ}25'E$ ). The new forcing data are interpolated to the in situ observation station and the corresponding accuracy is evaluated. Then, the interpolated forcing data are used to drive CoLM, and the shallow soil moisture observations are assimilated to estimate the soil moisture in different layers. The analysis errors and water budget residuals are compared among the different schemes. The assimilation results can provide some experiences for satellite data assimilation for future research.

The rest of the paper is organized as follows. The model and data are introduced in Section 2. The modified ETKF schemes with forecast error inflation and a water balance constraint are described in Section 3. The assimilation results of the observation station are presented in Section 4. Lastly, Section 5 provides discussions and the conclusions are in Section 6.

## 2. Model and Data

*2.1. Model.* The development of Common Land Model (CoLM) dates back to the mid-1990s, and it is based on land surface model (LSM) [45]. The biosphere-atmosphere transfer scheme (BATS) [46] and the Institute of Atmospheric Physics land surface model (IAP 94) [47] also account for different physical processes, such as glaciers, lakes, wetlands, and dynamic vegetation. It can be described as a community effort and that has been widely adopted by multiple climate models. Many researchers have studied land state simulation

TABLE 1: The node depths (cm) of 10 soil layers in the CoLM model.

Layer	1	2	3	4	5	6	7	8	9	10
Depth (cm)	0.7	2.8	6.2	11.9	21.2	36.6	62.0	103.8	172.8	286.5

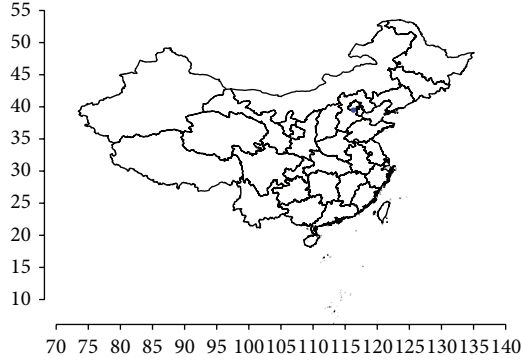


FIGURE 1: The blue circle is the location of the observation site in Daxing district of Beijing (39°37N, 116°25E). The dominant vegetation type in the area is corn/wheat.

and assimilation based on this generally used model. For instance, assimilating the MODIS-based albedo and snow cover fraction can improve snow depth simulation [48], and the estimation of water and energy fluxes during the growing season is useful for data-limited heterogeneous farmland [49].

The model performance has been validated in very extensive field data, including sites adopted by the Project for the Intercomparison of Land-Surface Parameterization Schemes and others [13]. There are 10 unevenly spaced soil layers (see Table 1), one vegetation layer, and 5 snow layers (depending on snow depth) in CoLM. The model parameters, including global terrain, elevation, land use/vegetation, a land-water mask, and soil types are available from the United States Geological Survey (USGS) at 30-arc-second resolution. The CoLM is set as the forecast model to simulate soil moisture in this study.

**2.2. Site Data.** The Automatic Weather Station (AWS) studied in this paper is located in the Daxing district of Beijing (39°37N, 116°25E; see Figure 1) and the dominant vegetation type in the area is corn/wheat. The observation items conclude air temperature/humidity/pressure, wind speed/direction, precipitation, radiation, soil temperature/moisture, and so on, and the collection frequency is every 10 minutes from 2008 to 2010. The depths of the soil moisture sensor are 5, 10, 20, 40, 60, and 100 cm. The data that are obviously beyond the range of physical possibility are rejected, and the gaps are filled using interpolation [50]. The shallow soil moisture observations (at 5 cm) are assimilated to CoLM in this study and the other soil moisture observations are used for validation.

**2.3. Forcing Data.** The atmospheric forcing data used to run CoLM mainly included the near-surface air temperature,

relative humidity, surface pressure, wind speed, precipitation, downward shortwave radiation, and longwave radiation. The accuracy of atmospheric forcing data is very important to the simulation of the land surface model [22, 23, 51]. In many studies, global atmospheric forcing datasets have been mainly derived from reanalysis datasets, such as the ERA-interim dataset [52], the Climate Forecast System Reanalysis (CFSR) dataset [25], and the Princeton forcing dataset [53]. However, only observations from approximately 200 stations in China have been used in the international exchanges to construct such datasets. Therefore, their quality over China mainland is questionable. To compensate for this defect, a method of applying monthly observations to correct the reanalysis data is widely used, which may lead to insufficient accuracy at hourly scale and low spatial resolution.

Recently, a new technique is applied to construct an atmospheric forcing dataset for China mainland with a resolution of  $1 \times 1$  km [26]. Firstly, a partial thin-plate smoothing spline with orography and reanalysis data as explanatory variables to ground-based observations is used to estimate the trend surface. Secondly, a simple kriging procedure is used to correct the residual. In Section 4, this dataset is accessed through realistic meteorological observation data.

### 3. Methodology

**3.1. Forecast and Observation System.** Using notations similar to those of Ide et al. [54], a nonlinear discrete-time forecast system is written as

$$\mathbf{y}_i^f = M_{i-1}(\mathbf{y}_{i-1}^a), \quad (1)$$

where  $i$  is the time step index;  $\mathbf{y}$  is the  $n$ -dimensional state vector ( $n = 126$  in CoLM); the superscripts “f” and “a” specify forecast and analysis respectively; and  $M_i$  is the nonlinear forecast operator (CoLM in this study). The major objective of this study is to assimilate the shallow soil moisture observations to improve the estimation of the variables related to the water budget; that is, the 22-dimensional vector  $\mathbf{x} = (\mathbf{SM}^T, \mathbf{SIC}^T, \text{CWC}, \text{SWE})^T$  in which  $\mathbf{SM}$  and  $\mathbf{SIC}$  are the 10-dimensional vectors specifying the soil moisture and the soil ice, respectively, at the 10 vertical levels shown in Table 1, and the units are volume percentage % ( $1\% = 0.01 \text{ m}^3/\text{m}^3$ ); the scalars CWC and SWE specify the canopy water content and the snow water equivalent, respectively.

The observation system can be written as

$$o_i = \mathbf{H}_i \mathbf{x}_i^t + \varepsilon_i, \quad (2)$$

where  $o_i$  is the soil moisture observations that are only available at the depth of 5 cm at 0000 UTC every day;  $\mathbf{H}_i$  is a 22-dimensional vector that linearly interpolates the soil moisture at depths of 2.8 cm and 6.2 cm to the depth of

5 cm; and  $\varepsilon_i$  is the observation error which is assumed to be statistically independent from the forecast error, is time-uncorrelated, and has mean zero and covariance  $R_i$ .

**3.2. ETKF with Forecast Error Inflation and Water Balance Constraint.** To improve the assimilation efficiency, the forecast error inflation technique is usually applied to the assimilation procedure. In addition to reducing analysis error, land data assimilation of soil moisture generally degrades the water balance as a result of state updates. Following Yilmaz et al. [43, 44], the water budget balance residual at time  $i$  can be expressed as

$$\begin{aligned} r_i &\equiv \beta_i - \mathbf{c}^T \mathbf{x}_i, \\ \beta_i &= \mathbf{c}^T \mathbf{x}_{i-1} + \text{Pr}_i - \text{Ev}_i - \text{Rn}_i, \end{aligned} \quad (3)$$

where  $\mathbf{c}$  is the 22-dimensional unit vector weighted by 1/3600 to unify the dimensions,  $\text{Pr}_i$  specifies precipitation from atmospheric forcing data, and  $\text{Ev}_i$  and  $\text{Rn}_i$  specify the diagnostic states of evapotranspiration and runoff. The water budget balance residual  $r_i$  represents the difference of the soil moistures at the current time step and previous time step, considering precipitation, evapotranspiration, and runoff. The detailed procedure of ETKF assimilation with the proposed forecast error inflation and water balance constraint scheme is described as follows:

**3.2.1. Forecast Step.** Calculate the  $j$ th perturbed forecast state at time  $i$  as

$$\mathbf{y}_{i,j}^f = M_{i-1,j} (\mathbf{y}_{i-1,j}^a), \quad (4)$$

where  $\mathbf{y}_{i,j}^f$  is the  $j$ th perturbed forecast state at time  $i$  and  $\mathbf{y}_{i-1,j}^a$  is the  $j$ th perturbed analysis state at time  $i-1$  ( $j = 1, 2, \dots, m$ );  $M_{i-1,j}$  is the CoLM forced by the  $j$ th perturbed atmospheric forcing data; and  $m$  is the total number of ensemble members (select as 100 in this study).

The mean forecast state of the variables related to the water budget is defined as

$$\mathbf{x}_i^f = \frac{1}{m} \sum_{j=1}^m \mathbf{x}_{i,j}^f. \quad (5)$$

The forecast error matrix and the forecast error covariance matrix are initially estimated as

$$\mathbf{X}_i^f = (\mathbf{x}_{i,1}^f - \mathbf{x}_i^f, \dots, \mathbf{x}_{i,m}^f - \mathbf{x}_i^f), \quad (6)$$

$$\mathbf{P}_i = \frac{1}{m-1} \mathbf{X}_i^f \mathbf{X}_i^{fT}, \quad (7)$$

respectively.

Then the forecast error matrix is inflated to the following form:

$$\mathbf{X}_i^{f(s)} = \left[ \sqrt{\lambda_i^{(s)}} \right] \mathbf{X}_i^f, \quad (8)$$

and the forecast error covariance matrix is adjusted as

$$\mathbf{P}_i^{(s)} = \left[ \sqrt{\lambda_i^{(s)}} \right] \mathbf{P}_i \left[ \sqrt{\lambda_i^{(s)}} \right], \quad (9)$$

where  $\left[ \sqrt{\lambda_i^{(s)}} \right]$  is a diagonal matrix with the diagonal elements  $\sqrt{\lambda_i^{(s)}}$  and  $s$  is the inflation layer that is selected by (11). For the soil moistures in the layers shallower than  $s$ ,  $\lambda_i^{(s)}$  is a scalar  $\lambda_i^{(s)}$  that inflates the forecast errors; the soil moisture in the layers deeper than  $s$  are not updated, and then the corresponding elements of  $\lambda_i^{(s)}$  (i.e., the  $\lambda_i^{(s)}$  related to  $\mathbf{SM}$  in the layer from  $s+1$  to 10) are set to zero; the forecast errors of the water budget variables other than soil moistures are not inflated, and then the corresponding elements of  $\lambda_i^{(s)}$  (i.e., the  $\lambda_i^{(s)}$  related to  $\mathbf{SIC}$ ,  $\mathbf{CWC}$ , and  $\mathbf{SWE}$ ) are set to unit. To cover the observations at 5 cm depth, the threshold layer  $s$  should be larger than 3.

Because the observation and forecast errors are assumed to be statistically independent of each other, time-uncorrelated, and with mean zero and covariance  $\mathbf{P}_i^{(s)}$  and  $R_i$ , the innovation statistic  $o_i - \mathbf{H}_i \mathbf{x}_i^f$  is normally distributed with mean zero and covariance  $\mathbf{H}_i \mathbf{P}_i^{(s)} \mathbf{H}_i^T + R_i$ . Therefore, the inflation factor  $\hat{\lambda}_i^{(s)}$  can be estimated by minimizing the following  $-2$  log-likelihood of the innovation statistic [38, 39]; that is,

$$\begin{aligned} L_i^{(s)}(\lambda_i^{(s)}) &= \ln \left( \mathbf{H}_i \mathbf{P}_i^{(s)} \mathbf{H}_i^T + R_i \right) \\ &+ (o_i - \mathbf{H}_i \mathbf{x}_i^f)^T \left( \mathbf{H}_i \mathbf{P}_i^{(s)} \mathbf{H}_i^T + R_i \right)^{-1} (o_i - \mathbf{H}_i \mathbf{x}_i^f). \end{aligned} \quad (10)$$

The total objective function value  $L^{(s)}$  during the entire assimilation period is

$$L^{(s)} \equiv \sum_{i=1}^N \left( L_i^{(s)}(\hat{\lambda}_i^{(s)}) \right), \quad (11)$$

which is treated as a function of the threshold layer  $s$ . For a given  $s$ , the inflation factors  $\lambda_i^{(s)}$  are estimated by minimizing (10) and the value of  $L^{(s)}$  is also calculated through (11). Then the threshold layer  $s$  is selected as the smallest number such that  $L^{(s)}$  is the smallest value among  $(L^{(3)}, L^{(4)}, \dots, L^{(s+1)})$ . The forecast error inflation is conducted using the selected  $s$  and the corresponding inflation factors  $\lambda_i^{(s)}$ .

**3.2.2. Analysis Step.** Calculate the analysis state as

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{X}_i^{f(s)} \mathbf{w}_i^{a(s)}, \quad (12)$$



where  $\mathbf{w}_i^{a(s)}$  is the weight vector, which is estimated by minimizing the following objective function with a weak water balance constraint [43]

$$\begin{aligned} J_i(\mathbf{w}) = & (m-1)\mathbf{w}^T\mathbf{w} + \left(o_i - \mathbf{H}_i(\mathbf{x}_i^f + \mathbf{X}_i^{f(s)}\mathbf{w})\right)^T \\ & \cdot R_i^{-1}\left(o_i - \mathbf{H}_i(\mathbf{x}_i^f + \mathbf{X}_i^{f(s)}\mathbf{w})\right) \\ & + \left(\beta_i - \mathbf{c}^{(s)T}(\mathbf{x}_i^f + \mathbf{X}_i^{f(s)}\mathbf{w})\right)^T \\ & \cdot \varphi_i^{-1}\left(\beta_i - \mathbf{c}^{(s)T}(\mathbf{x}_i^f + \mathbf{X}_i^{f(s)}\mathbf{w})\right), \end{aligned} \quad (13)$$

where  $\mathbf{c}^{(s)}$  is a 22-dimensional vector similar to vector  $\mathbf{c}$  but where components corresponding to the soil moisture in the layers deeper than  $s$  are modified to zero, which is different from Yilmaz et al. [43, 44]. One has that

$$\varphi_i = \frac{1}{m-1} \sum_{j=1}^m (\beta_{i,j} - \beta_i)^2 \quad (14)$$

is the error variance of  $\beta_i$ , and  $\beta_{i,j}$  is the  $j$ th perturbed realizations of  $\beta_i$ . For the linear observation operator, the estimated weight vector had the following analytic form:

$$\begin{aligned} \mathbf{w}_i^{a(s)} = & \tilde{\mathbf{P}}_i^{(s)} \left(\mathbf{X}_i^{f(s)T} \mathbf{H}_i^T R_i^{-1} (o_i - \mathbf{H}_i \mathbf{x}_i^f)\right) \\ & + \mathbf{X}_i^{f(s)T} \mathbf{c}^{(s)T} \varphi_i^{-1} (\beta_i - \mathbf{c}^{(s)T} \mathbf{x}_i^f), \end{aligned} \quad (15)$$

where  $\tilde{\mathbf{P}}_i^{(s)}$  is the posterior variance of  $\mathbf{w}_i^{a(s)}$ ; that is,

$$\begin{aligned} \tilde{\mathbf{P}}_i^{(s)} = & \left( (m-1)\mathbf{I} + \mathbf{X}_i^{f(s)T} \mathbf{H}_i^T R_i^{-1} \mathbf{H}_i \mathbf{X}_i^{f(s)} \right. \\ & \left. + \mathbf{X}_i^{f(s)T} \mathbf{c}^{(s)T} \varphi_i^{-1} \mathbf{c}^{(s)} \mathbf{X}_i^{f(s)} \right)^{-1}. \end{aligned} \quad (16)$$

Then, calculate a perturbed analysis state as

$$\mathbf{x}_{i,j}^a = \mathbf{x}_i^a + \mathbf{X}_i^{f(s)} \mathbf{w}_{i,j}^{a(s)}, \quad (17)$$

where  $\mathbf{W}_{i,j}^{a(s)}$  is the  $j$ th column of the matrix  $\mathbf{W}_i^{a(s)} = ((m-1)\tilde{\mathbf{P}}_i^{(s)})^{1/2}$  [31]. The analysis states of the variables other than those related to the water budget are set as the corresponding forecast states. Lastly, set  $i = i + 1$  and return to the Forecast Step for the next iteration. The complete assimilation procedure is shown in Figure 2.

**3.3. Validation Statistics.** To compare the estimated 10 layers of soil moisture with the 6 levels of soil moisture observations, a root-mean-square-error statistic is used:

$$\text{RMSE}(\text{lev}) = \sqrt{\frac{1}{N} \sum_{i=1}^N ((\mathbf{h}\mathbf{x}_i^a)(\text{lev}) - o_i(\text{lev}))^2}, \quad (18)$$

where  $\mathbf{h}$  is the linear interpolation operator that interpolates the soil moisture analysis to all of the depths of the observations and lev specifies the level of observations.

To test the effect of the forecast error inflation, a Chi-square statistic [55] is used in this study. The statistic at time  $i$  is defined as

$$\chi_i^2 = (o_i - \mathbf{H}_i \mathbf{x}_i^f)^T (\mathbf{H}_i \mathbf{P}_i^{(s)} \mathbf{H}_i^T + R_i)^{-1} (o_i - \mathbf{H}_i \mathbf{x}_i^f). \quad (19)$$

If the forecast and observation error covariance matrix are correctly estimated, the innovation statistic  $o_i - \mathbf{H}_i \mathbf{x}_i^f$  is normally distributed with mean zero and covariance  $\mathbf{H}_i \mathbf{P}_i^{(s)} \mathbf{H}_i^T + R_i$ ; then  $\chi_i^2$  will follow a Chi-square distribution with the number of observations as its degrees of freedom [56]. The value of  $\chi_i^2$  in this study can be easily calculated and should be close to 1. Therefore the inflation effect can be effectively evaluated.

To illustrate the water budget balance at a location, the water budget balance residual is calculated as follows:

$$r = \frac{1}{N} \sum_{i=1}^N r_i. \quad (20)$$

If an assimilation scheme maintains a water budget balance, the water balance residual should be close to 0.

## 4. Experimentation

**4.1. Evaluation of the Atmospheric Forcing Data.** In the following experiments at the Daxing Automatic Weather Station, the atmospheric forcing data with high precision are obtained from the observations. To evaluate the accuracy of interpolated forcing data (described in Section 2.3), CoLM is driven using the interpolated data, the CFSR data, and the ERA-interim data as the meteorological forcing for one year (from June 1 2009 to May 31 2010). The simulated results are then interpolated to the 6 observation layers and the corresponding RMSE are plotted in Figure 3. Because the CFSR data are used to estimate the trend surface when generating the interpolated forcing data, the latter performs better than the former in almost all 6 layers. In the shallow layers (5 cm and 10 cm) and deep layers (60 cm and 100 cm), the RMSE of the interpolated data is smaller than the ERA-interim data but is converse in the middle layers (20 cm and 40 cm). This may be due to the soil moisture in the deep layers fluctuating slower than in the middle and shallow layers, and also the interpolation errors in the shallow layers being smaller than in the middle and deep layers. On the whole, the interpolated data perform better than the CFSR data and the ERA-interim data. Whether this is a general principal or is site dependent will be validated further using more observations in the near future.

Next, the interpolated forcing data are used to drive CoLM from June 1 2009 to August 31 2010 to simulate the soil moisture at the station. The results in the first year are treated as spin-up and the last three months are used for the assimilation experiments. The time-mean values of observed and interpolated forcing data, as well as the corresponding RMSEs of the seven variables, are listed in Table 2. The simulated soil moistures using interpolated forcing data are interpolated to the observation layer and are compared with

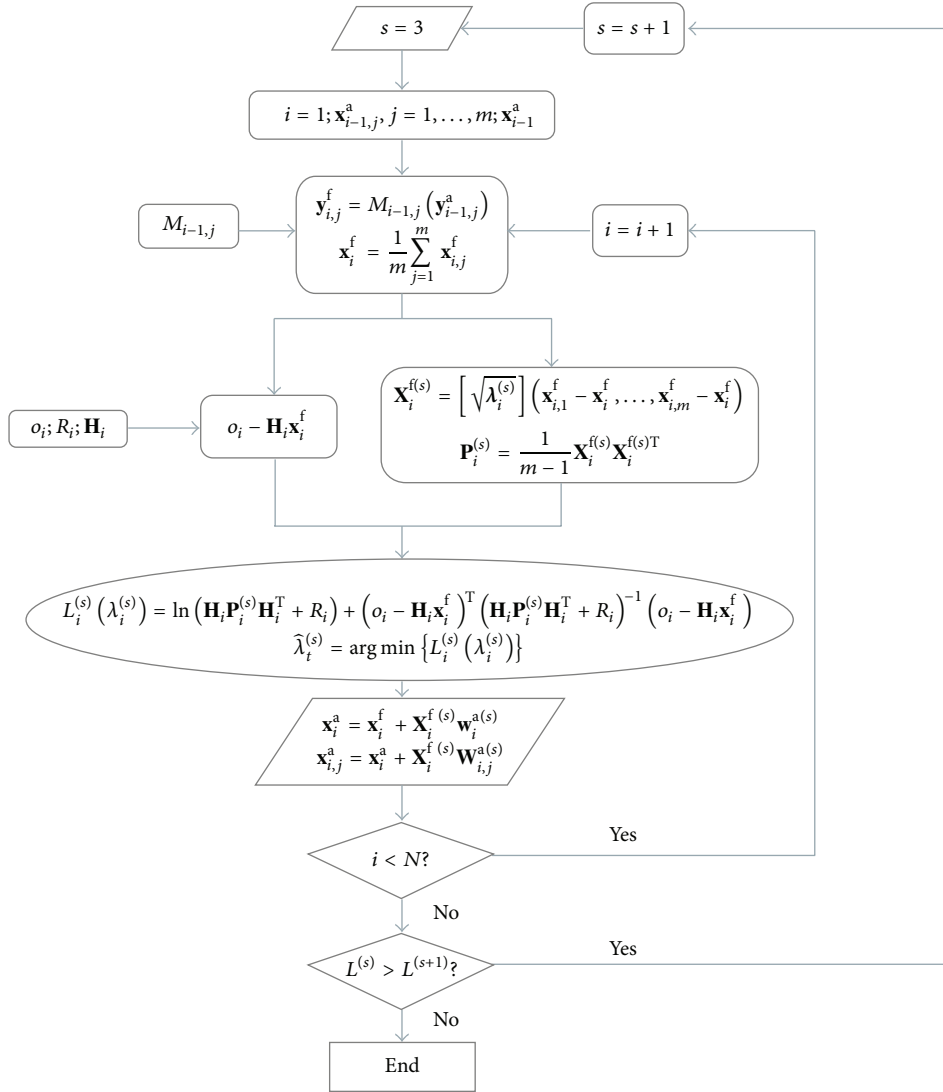


FIGURE 2: The assimilation procedure of the experiments.

TABLE 2: The time-mean values of observed and interpolated forcing data, as well as the corresponding time-mean RMSEs: near-surface air temperature, relative humidity, wind speed, precipitation, downward shortwave radiation, and longwave radiation.

	Near-surface air temperature (K)	Relative humidity (Kg/Kg)	Surface pressure (Pa)	Wind speed (m/s)	Precipitation (mm/s)	Downward shortwave radiation (W/m <sup>2</sup> )	Downward longwave radiation (W/m <sup>2</sup> )
Observed forcing data	287.92	0.0080	101555	2.38	$1.65e - 5$	153.41	331.19
Interpolated forcing	287.91	0.0081	101017	2.26	$1.91e - 5$	180.53	353.81
RMSE	1.87	0.0019	618.36	1.52	0.0002	126.69	44.27

the soil moisture observations. The time-mean RMSEs for the 6 observation layers are listed in Table 3. Generally speaking, the interpolated forcing data are comparable with the observed forcing data, especially for near-surface air temperature and relative humidity. The simulated soil moisture using interpolated forcing data is different from the observed values but can be improved by assimilating the observation to CoLM, which is introduced in the last section.

**4.2. Assimilation Results.** In this section, the interpolated data are used to drive CoLM and the soil moisture observation at the depth of 5 cm at 0000 UTC is assimilated every day from June 1 2010 to August 31 2010. The four variables strongly related to soil moisture (near-surface air temperature, precipitation, downward shortwave radiation, and longwave radiation) are perturbed with a 5% error scale to generate the ensemble forecast states. The initial values used during

TABLE 3: The time-mean values of observed and simulated soil moisture in the observation layers, as well as the corresponding time-mean RMSEs.

	5 cm	10 cm	20 cm	40 cm	60 cm	100 cm	Mean
Observed soil moisture	18.51	19.92	26.92	28.99	29.44	29.06	25.47
Simulated soil moisture	26.67	27.58	28.19	32.39	34.19	34.72	30.62
RMSE	5.80	4.29	7.53	11.38	4.88	3.27	6.19

TABLE 4: The time-mean RMSEs of estimated soil moistures in observation layers of different assimilation schemes.

	5 cm	10 cm	20 cm	40 cm	60 cm	100 cm	Mean
ETKF	5.89	4.85	3.98	2.12	3.93	5.47	4.37
ETKF-Inf ( $s = 6$ )	5.81	3.30	3.22	6.49	2.27	2.20	3.88
WCETKF	4.38	3.32	6.18	10.89	5.58	5.79	6.02
WCETKF-Inf ( $s = 5$ )	5.44	3.02	2.54	6.76	4.36	4.23	4.39

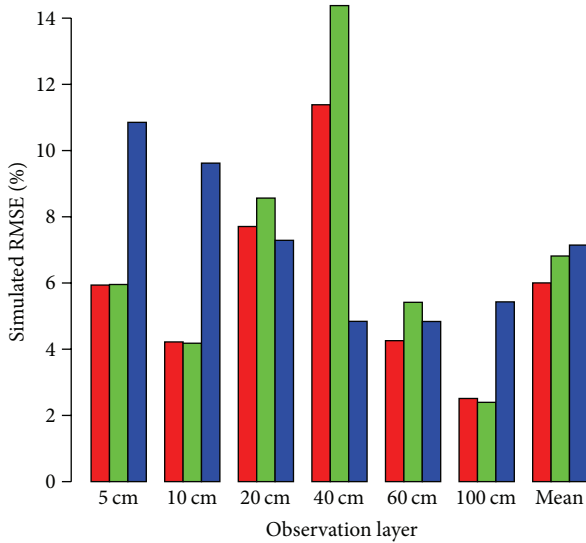


FIGURE 3: The simulation results of soil moisture in the 6 observation layers using different meteorological forcing. Red bars: the interpolated forcing; green bars: the CFSR forcing; blue bars: the ERA-interim forcing.

the assimilation period are the model forecasts after one year's spin-up. The assimilated soil moisture observations are plotted in Figure 4. The observation error scale is set as 0.5%. The following four ETKF assimilation settings are examined.

ETKF: traditional ETKF without forecast error inflation and water balance constraint.

ETKF-Inf: ETKF with forecast error inflation but without water balance constraint.

WCETKF: ETKF with water balance constraint but without forecast error inflation.

WCETKF-Inf: ETKF with forecast error inflation and water balance constraint.

The selected inflation layer is 5 for WCETKF-Inf and 6 for ETKF-Inf using the criterion of (11) described in Section 3.2.1. The estimated 10-layer soil moistures are interpolated to the

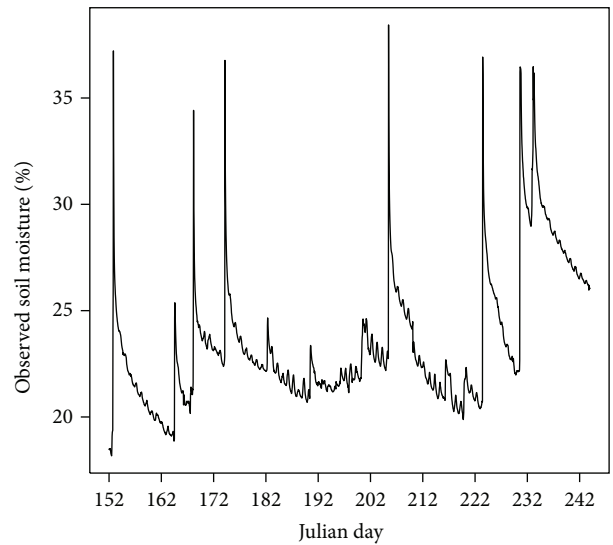


FIGURE 4: The assimilated soil moisture observation from June 1 2010 to August 31 2010.

6 observation layers and the corresponding RMSEs of these assimilation schemes as well as the corresponding mean values are listed in Table 4. For all of the four different assimilation schemes, the RMSEs are generally smaller than those of the simulations (the second row in Table 3), except in separate layers. The mean RMSE of ETKF-Inf (3.88) is smaller than that of ETKF (4.87), whereas the mean RMSE of WCETKF-Inf (4.39) is smaller than that of WCETKF (6.02), indicating that the forecast error inflation can improve the assimilation accuracy. However, the mean RMSE of WCETKF (6.02) is larger than that of ETKF (4.87), whereas the mean RMSE of WCETKF-Inf (4.39) is larger than that of ETKF-Inf (3.88), indicating that the water balance constraint may result in assimilation accuracy loss to some extent.

The ensemble members of the ETKF-Inf scheme are plotted in Figure 5, which indicates the uncertainty of our analysis state to some extent. It shows that the ensemble spreads in the very shallow and deep layers are relatively small whereas those in the middle layers are slightly large, which

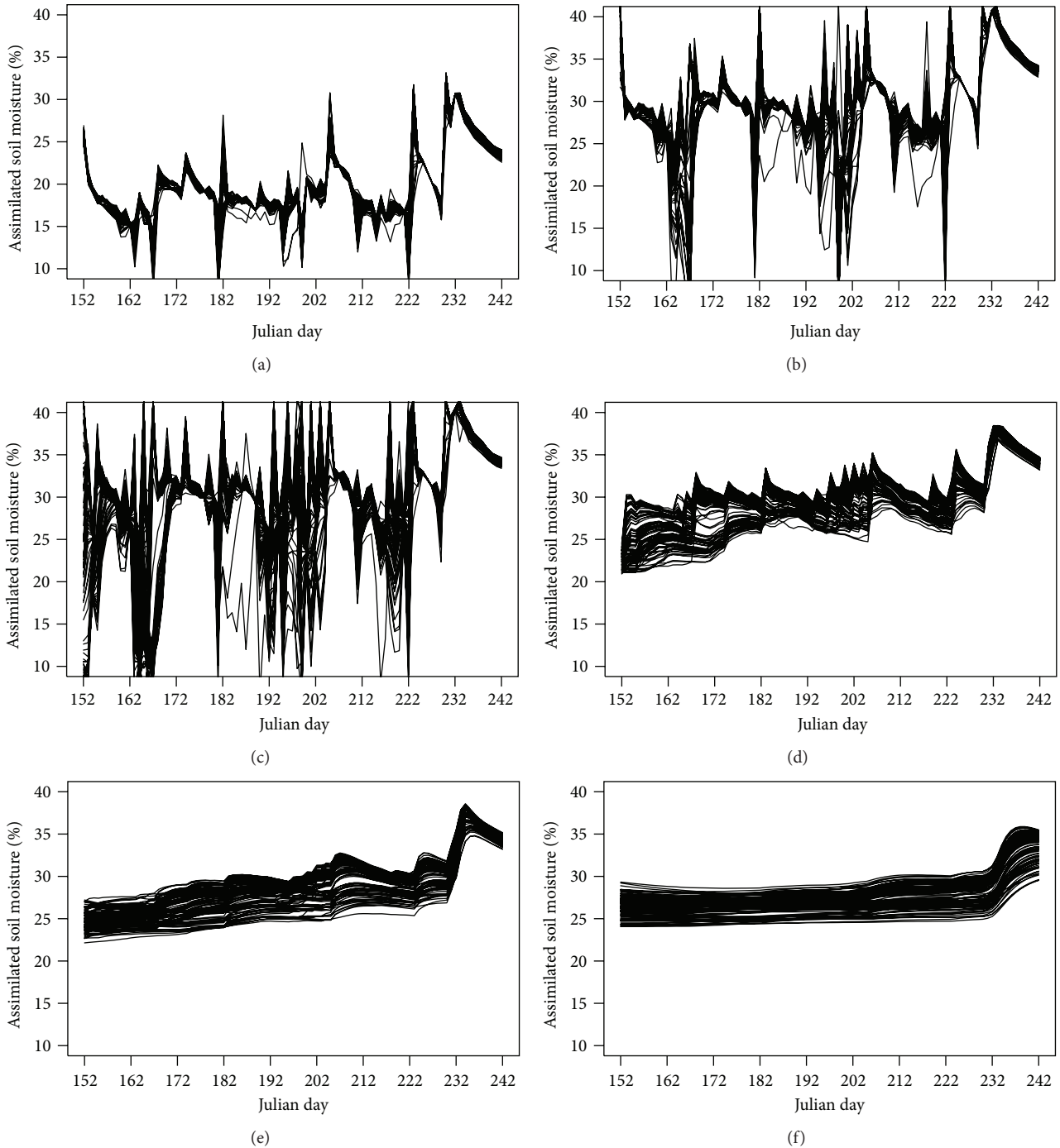


FIGURE 5: The ensemble members of the ETKF-Inf ( $s = 6$ ) scheme in the 6 observation layers from June 1 2010 to August 31 2010. (a)–(f) are for the observation layers at 5 cm, 10 cm, 20 cm, 40 cm, 60 cm, and 100 cm, respectively.

is consistent with the fluctuation of the used forcing data. This may be due to the assimilated observation being in the shallow layer (at the depth of 5 cm) and the soil moisture in the deep layers fluctuating slower than in the shallow layers. Therefore the corresponding uncertainty in the very shallow and deep layers seems to be small, whereas that in the middle layers may be larger.

**4.3. Water Budget Residual.** The water balance constraint is conducted in this study to reduce the water budget residual. For the two assimilation schemes with inflating forecast error on selected inflation layers, the water budget residual (see (19)) of ETKF-Inf scheme is 0.1580 mm, whereas it is only 0.0662 mm of WCETKF-Inf scheme, indicating that the water balance constraint can reduce the water budget residual.



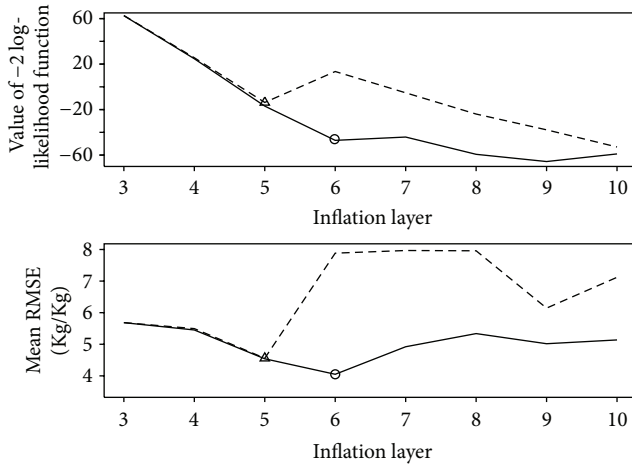


FIGURE 6: The value of objective function and the mean RMSE as a function of inflation layer. The solid line: ETKF-Inf; the dashed line: WCETKF-Inf.

## 5. Discussions

**5.1. Inflation.** It is widely recognized that the initial estimate of ensemble forecast errors should be inflated to improve the assimilated results. In this study, a method for estimating the multiplicative inflation factors by minimizing the  $-2$  log-likelihood of the innovation statistic is investigated. The forecast errors of soil moisture in the layers shallower than  $s$  are inflated, while the threshold layer  $s$  is selected by the criterion described in Section 3.2.1. To examine the effect of the select criterion, the assimilation experiments are conducted with the inflation layer from 3 to 10. The mean RMSE and the value of the objective function as a function of the inflation layer are plotted in Figure 6. For the assimilation schemes with and without a water balance constraint, the selected inflation layers are both consistent with the smallest mean RMSE, indicating that the selected criteria described in Section 3.2.1 are effective.

Another statistic to evaluate the effect of the forecast error inflation is the Chi-square statistic (see (20)) described in Section 3.3. For WCETKF-Inf and ETKF-Inf assimilation schemes with selected inflation layers (5 and 6, resp.), as well as the two assimilation schemes without a forecast error inflation (ETKF and WCETKF), the corresponding time series of  $\chi_i^2$  from June 1 2010 to August 31 2010 are plotted in Figure 7. It shows that the values are remarkably close to 1 if the forecast error is inflated. In contrast, if the initially estimated forecast error (i.e.,  $\lambda_i^{(s)} = 1$ ) is used in the assimilation procedure, the values of  $\chi_i^2$  depart extremely from 1. This indicates that the forecast error inflation method used in this study is effective for correctly estimating the error statistics.

The  $-2$  log-likelihood function of the innovation statistic seems like a good objective function to quantify the goodness-of-fit of the forecast error covariance matrix. In fact, suppose that the forecast and observation errors are statistically independent of each other, time-uncorrelated, and normally distributed with zero mean and covariance  $\mathbf{P}_i^{(s)}$

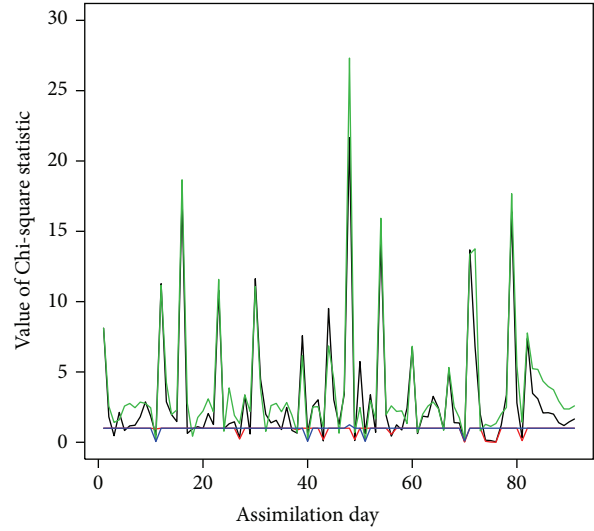


FIGURE 7: The Chi-square statistic (see (20)) at the assimilation time from June 1 2010 to August 31 2010 for the four assimilation schemes. The black line: ETKF; the red line: ETKF-Inf ( $s = 6$ ); the green line: WCETKF; the blue line: WCETKF-Inf ( $s = 5$ ).

and  $R_i$ ; then the innovation statistic will also follow a normal distribution with mean zero and covariance  $\mathbf{H}_i \mathbf{P}_i^{(s)} \mathbf{H}_i^T + R_i$ . Therefore the inflation factor estimation method based on the  $-2$  log-likelihood function will incorporate the information of the whole distribution other than that of the moment. In most cases of this study, the  $-2$  log-likelihood function corresponds to a smaller RMSE.

**5.2. Water Balance Constraint.** Assimilation of hydrological observations (e.g., soil moisture) can improve the estimates of hydrological variables but generally results in water balance degradation. To compensate for this defect, a water balance constraint in the assimilation procedure becomes increasingly important.

To evaluate the effect of the water balance constraint in this study, the water balance residual (see (19)) is calculated for the inflation layer from 3 to 10. The results of ETKF-Inf and WCETKF-Inf assimilation schemes are plotted in Figure 8. The figure clearly shows that, for each scheme, the water balance residual increases as the inflation layer increases. However, for a fixed inflation layer, the water balance residual of the WCETKF-Inf assimilation scheme is smaller than that of the ETKF-Inf assimilation scheme. This illustrates that the water balance constraint is effective for maintaining the water balance and the inflation layer selection is very important.

Another problem that requires attention is that, for a fixed inflation layer, the mean RMSE of the WCETKF-Inf assimilation is usually larger than that of the ETKF-Inf assimilation scheme. In land data assimilation systems, the state updates usually produce a water budget imbalance residual. The analysis error may increase if a water balance constraint is conducted, which may not be experiment dependent but may be the general case. In fact, the cost function of

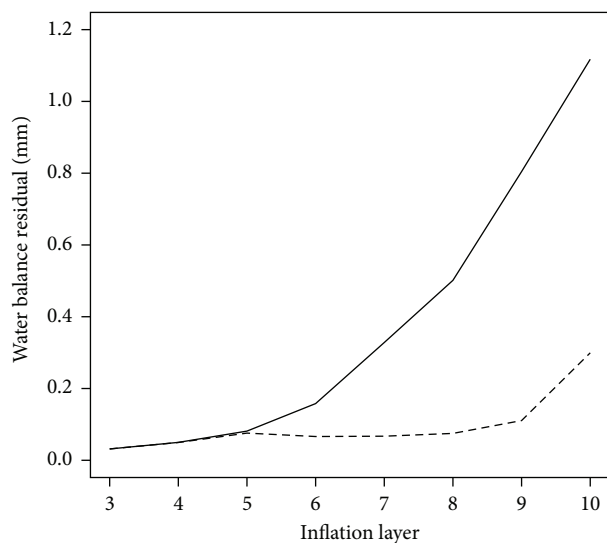


FIGURE 8: The water balance residual as a function of inflation layer. The solid line: ETKF-Inf; the dashed line: WCETKF-Inf.

the assimilation scheme with a water balance constraint is the summation of three terms that represent the forecast errors, observation errors, and the water budget imbalance, respectively. The corresponding assimilation results can be treated as a trade-off in reducing the analysis error and water budget residual. Although it may lead to the analysis error increasing, the water balance constraint can make the land surface model more consistent in the physical structure. A better understanding of the water budget can also help to facilitate model parameterization development, better understand the hydrological processes, improve our knowledge of land-atmosphere water exchange and related physical mechanisms and, therefore, help to improve our ability to develop models.

## 6. Conclusion

In this study, interpolated forcing data is evaluated and the observations in the shallow layer are assimilated to CoLM to better estimate soil moisture. The forecast error is inflated to improve the analysis state accuracy and the water balance constraint is adopted to reduce the water budget residual in the assimilation procedure. The experimental results illustrate that the adaptive forecast error inflation can reduce the analysis error, whereas the proper inflation layer can be selected based on the  $-2$  log-likelihood function of the innovation statistic. The water balance constraint can result in a substantially reduced water budget residual, at a low cost of assimilation accuracy loss. The assimilation scheme can potentially be applied to assimilate the remote sensing data and we plan to do this in the near future.

## Competing Interests

The authors declare that there are no competing interests regarding the publication of this paper.

## Acknowledgments

The datasets are provided by the Haihe Experiments of Professor Liu SM (<http://westdc.westgis.ac.cn/haihe>). This work is supported by the National Basic Research Program of China (Grants nos. 2012CB956203 and 2015CB953703), the National Natural Science Foundation of China (Grant no. 41405098), and the R&D Special Fund for Nonprofit Industry (Meteorology, Grant no. GYHY201206008).

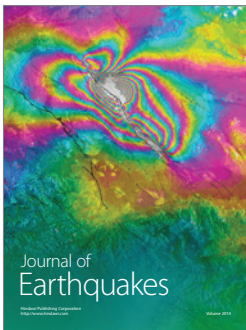
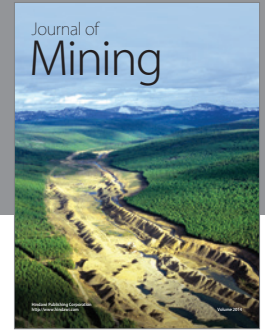
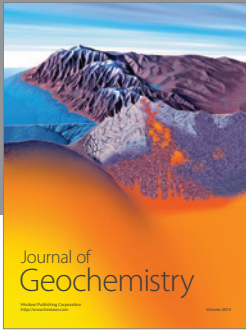
## References

- [1] D. A. Robinson, C. S. Campbell, J. W. Hopmans et al., "Soil moisture measurement for ecological and hydrological watershed-scale observatories: a review," *Vadose Zone Journal*, vol. 7, no. 1, pp. 358–389, 2008.
- [2] D. Entekhabi, I. Rodriguez-Iturbe, and F. Castelli, "Mutual interaction of soil moisture state and atmospheric processes," *Journal of Hydrology*, vol. 184, no. 1-2, pp. 3–17, 1996.
- [3] M. Rodell, P. R. Houser, U. Jambor et al., "The global land data assimilation system," *Bulletin of the American Meteorological Society*, vol. 85, no. 3, pp. 381–394, 2004.
- [4] S. W. Zhang, H. R. Li, W. D. Zhang, C. J. Qiu, and X. Li, "Estimating the soil moisture profile by assimilating near-surface observations with the ensemble Kalman filter (EnKF)," *Advances in Atmospheric Sciences*, vol. 22, no. 6, pp. 936–945, 2005.
- [5] X. Tian, Z. Xie, and A. Dai, "A land surface soil moisture data assimilation system based on the dual-UKF method and the community land model," *Journal of Geophysical Research Atmospheres*, vol. 113, no. 14, Article ID D14127, 2008.
- [6] X. Han, X. Li, H. J. H. Franssen, H. Vereecken, and C. Montzka, "Spatial horizontal correlation characteristics in the land data assimilation of soil moisture," *Hydrology and Earth System Sciences*, vol. 16, pp. 1349–1363, 2012.
- [7] K. Yang, T. Koike, I. Kaihotsu, and J. Qin, "Validation of a dual-pass microwave land data assimilation system for estimating surface soil moisture in semiarid regions," *Journal of Hydrometeorology*, vol. 10, no. 3, pp. 780–793, 2009.
- [8] K. Yang, L. Zhu, Y. Chen et al., "Land surface model calibration through microwave data assimilation for improving soil moisture simulations," *Journal of Hydrology*, vol. 533, pp. 266–276, 2016.
- [9] Y. Sawada, T. Koike, and J. P. Walker, "A land data assimilation system for simultaneous simulation of soil moisture and vegetation dynamics," *Journal of Geophysical Research: Atmospheres*, vol. 120, no. 12, pp. 5910–5930, 2015.
- [10] C. Montzka, H. Moradkhani, L. Weihermüller, H.-J. H. Franssen, M. Canty, and H. Vereecken, "Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter," *Journal of Hydrology*, vol. 399, no. 3-4, pp. 410–421, 2011.
- [11] H. Moradkhani, S. Sorooshian, H. V. Gupta, and P. R. Houser, "Dual state-parameter estimation of hydrological models using ensemble Kalman filter," *Advances in Water Resources*, vol. 28, no. 2, pp. 135–147, 2005.
- [12] Y. J. Dai, X. B. Zeng, and R. E. Dickinson, "Common Land Model (CoLM): Technical Documentation and User's Guide," 2001.
- [13] Y. Dai, X. Zeng, R. E. Dickinson et al., "The common land model," *Bulletin of the American Meteorological Society*, vol. 84, no. 8, pp. 1013–1023, 2003.

- [14] Y. Dai, R. E. Dickinson, and Y.-P. Wang, "A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance," *Journal of Climate*, vol. 17, no. 12, pp. 2281–2299, 2004.
- [15] X. Zeng, M. Shaikh, Y. Dai, R. E. Dickinson, and R. Myneni, "Coupling of the common land model to the NCAR community climate model," *Journal of Climate*, vol. 15, no. 14, pp. 1832–1854, 2002.
- [16] A. Robock, L. Luo, E. F. Wood et al., "Evaluation of the North American Land Data Assimilation System over the southern Great Plains during the warm season," *Journal of Geophysical Research*, vol. 108, no. 22, 2003.
- [17] A. Robock, K. Y. Vinnikov, G. Srinivasan et al., "The global soil moisture data bank," *Bulletin of the American Meteorological Society*, vol. 81, no. 6, pp. 1281–1299, 2000.
- [18] G. R. Miller, D. D. Baldocchi, B. E. Law, and T. Meyers, "An analysis of soil moisture dynamics using multi-year data from a network of micrometeorological observation sites," *Advances in Water Resources*, vol. 30, no. 5, pp. 1065–1081, 2007.
- [19] K. Y. Vinnikov, A. Robock, N. A. Speranskaya, and C. A. Schlosser, "Scales of temporal and spatial variability of midlatitude soil moisture," *Journal of Geophysical Research*, vol. 101, no. 3, pp. 7163–7174, 1996.
- [20] Y. H. Kerr, "Soil moisture from space: where are we?" *Hydrogeology Journal*, vol. 15, no. 1, pp. 117–120, 2007.
- [21] J. C. Shi, L. M. Jiang, L. X. Zhang et al., "Physically based estimation of bare-surface soil moisture with the passive radiometers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3145–3152, 2006.
- [22] A. A. Berg, J. S. Famiglietti, J. P. Walker, and P. R. Houser, "Impact of bias correction to reanalysis products on simulations of North American soil moisture and hydrological fluxes," *Journal of Geophysical Research D: Atmospheres*, vol. 108, no. 16, pp. 1–15, 2003.
- [23] B. Fekete, C. J. Vörösmarty, J. O. Roads, and C. J. Willmott, "Uncertainties in precipitation and their impacts on runoff estimates," *Journal of Climate*, vol. 17, no. 2, pp. 294–304, 2004.
- [24] B. Nijssen and D. P. Lettenmaier, "Effect of precipitation sampling error on simulated hydrological fluxes and states: anticipating the Global Precipitation Measurement satellites," *Journal of Geophysical Research*, vol. 109, no. 2, Article ID D02103, 2004.
- [25] S. Saha, "The NCEP climate forecast system reanalysis," *Bulletin of the American Meteorological Society*, vol. 91, pp. 1015–1057, 2010.
- [26] T. Li, X. Zheng, Y. Dai et al., "Mapping near-surface air temperature, pressure, relative humidity and wind speed over Mainland China with high spatiotemporal resolution," *Advances in Atmospheric Sciences*, vol. 31, no. 5, pp. 1127–1135, 2014.
- [27] R. H. Reichle, "Data assimilation methods in the Earth sciences," *Advances in Water Resources*, vol. 31, no. 11, pp. 1411–1418, 2008.
- [28] G. Wu, X. Zheng, L. Wang, S. Zhang, X. Liang, and Y. Li, "A new structure for error covariance matrices and their adaptive estimation in EnKF assimilation," *Quarterly Journal of the Royal Meteorological Society*, vol. 139, no. 672, pp. 795–804, 2013.
- [29] C. H. Bishop and Z. Toth, "Ensemble transformation and adaptive observations," *Journal of the Atmospheric Sciences*, vol. 56, no. 11, pp. 1748–1765, 1999.
- [30] C. H. Bishop, B. J. Etherton, and S. J. Majumdar, "Adaptive sampling with the ensemble transform kalman filter part I: theoretical aspects," *Monthly Weather Review*, vol. 129, no. 3, pp. 420–436, 2001.
- [31] B. R. Hunt, E. J. Kostelich, and I. Szunyogh, "Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter," *Physica D: Nonlinear Phenomena*, vol. 230, no. 1–2, pp. 112–126, 2007.
- [32] X. G. Wang and C. H. Bishop, "A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes," *Journal of the Atmospheric Sciences*, vol. 60, no. 9, pp. 1140–1158, 2003.
- [33] J. L. Anderson and S. L. Anderson, "A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts," *Monthly Weather Review*, vol. 127, no. 12, pp. 2741–2758, 1999.
- [34] E. M. Constantinescu, A. Sandu, T. Chai, and G. R. Carmichael, "Ensemble-based chemical data assimilation. I: general approach," *Quarterly Journal of the Royal Meteorological Society*, vol. 133, no. 626, pp. 1229–1243, 2007.
- [35] G. Wu, X. Yi, L. Wang et al., "Improving the ensemble transform Kalman filter using a second-order Taylor approximation of the nonlinear observation operator," *Nonlinear Processes in Geophysics*, vol. 21, no. 5, pp. 955–970, 2014.
- [36] H. Li, E. Kalnay, and T. Miyoshi, "Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter," *Quarterly Journal of the Royal Meteorological Society*, vol. 135, no. 639, pp. 523–533, 2009.
- [37] X. Wang and C. H. Bishop, "A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes," *Journal of the Atmospheric Sciences*, vol. 60, no. 9, pp. 1140–1158, 2003.
- [38] X. Liang, X. Zheng, S. Zhang, G. Wu, Y. Dai, and Y. Li, "Maximum likelihood estimation of inflation factors on error covariance matrices for ensemble Kalman filter assimilation," *Quarterly Journal of the Royal Meteorological Society*, vol. 138, no. 662, pp. 263–273, 2012.
- [39] X. Zheng, "An adaptive estimation of forecast error covariance parameters for Kalman filtering data assimilation," *Advances in Atmospheric Sciences*, vol. 26, no. 1, pp. 154–160, 2009.
- [40] X. Zheng, G. Wu, S. Zhang, X. Liang, Y. Dai, and Y. Li, "Using analysis state to construct a forecast error covariance matrix in ensemble Kalman filter assimilation," *Advances in Atmospheric Sciences*, vol. 30, no. 5, pp. 1303–1312, 2013.
- [41] A. C. V. Getirana, E. Dutra, M. Guimberteau et al., "Water balance in the Amazon basin from a land surface model ensemble," *Journal of Hydrometeorology*, vol. 15, no. 6, pp. 2586–2614, 2014.
- [42] M. Pan and E. F. Wood, "Data assimilation for estimating the terrestrial water budget using a constrained ensemble Kalman filter," *Journal of Hydrometeorology*, vol. 7, no. 3, pp. 534–547, 2006.
- [43] M. T. Yilmaz, T. Delsole, and P. R. Houser, "Improving land data assimilation performance with a water budget constraint," *Journal of Hydrometeorology*, vol. 12, no. 5, pp. 1040–1055, 2011.
- [44] M. T. Yilmaz, T. DelSole, and P. R. Houser, "Reducing water imbalance in land data assimilation: ensemble filtering without perturbed observations," *Journal of Hydrometeorology*, vol. 13, no. 1, pp. 413–420, 2012.
- [45] G. B. Bonan, "A Land Surface Model (LSM Version 1. 0) for ecological, hydrological, and atmospheric studies," Technical Description and User's Guide, 1996.
- [46] R. E. Dickinson, A. Henderson-Sellers, and P. J. Kennedy, "Biosphere-atmosphere transfer scheme (BATS) version 1e as coupled to the NCAR community climate model," 1993.

- [47] Y. J. Dai and Q. C. Zeng, "A land surface model (IAP94) for climate studies part I: formulation and validation in off-line experiments," *Advances in Atmospheric Sciences*, vol. 14, no. 4, pp. 433–460, 1997.
- [48] J. Xu and H. Shu, "Assimilating MODIS-based albedo and snow cover fraction into the Common Land Model to improve snow depth simulation with direct insertion and deterministic ensemble Kalman filter methods," *Journal of Geophysical Research*, vol. 119, no. 18, pp. 10684–10701, 2014.
- [49] M. Choi, S. O. Lee, and H. Kwon, "Understanding of the common land model performance for water and energy fluxes in a farmland during the growing season in Korea," *Hydrological Processes*, vol. 24, no. 8, pp. 1063–1071, 2010.
- [50] S. M. Liu, Z. W. Xu, Z. L. Zhu, Z. Z. Jia, and M. J. Zhu, "Measurements of evapotranspiration from eddy-covariance systems and large aperture scintillometers in the Hai River Basin, China," *Journal of Hydrology*, vol. 487, pp. 24–38, 2013.
- [51] J. Sheffield, A. D. Ziegler, E. F. Wood, and Y. B. Chen, "Correction of the high-latitude rain day anomaly in the NCEP-NCAR reanalysis for land surface hydrological modeling," *Journal of Climate*, vol. 17, no. 19, pp. 3814–3828, 2004.
- [52] D. P. Dee, S. M. Uppala, A. J. Simmons et al., "The ERA-Interim reanalysis: configuration and performance of the data assimilation system," *Quarterly Journal of the Royal Meteorological Society*, vol. 137, no. 656, pp. 553–597, 2011.
- [53] J. Sheffield, G. Goteti, and E. F. Wood, "Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling," *Journal of Climate*, vol. 19, no. 13, pp. 3088–3111, 2006.
- [54] K. Ide, P. Courtier, M. Ghil, and A. C. Lorenc, "Unified notation for data assimilation: operational, sequential and variational," *Journal of the Meteorological Society of Japan*, vol. 75, no. 1, pp. 181–189, 1997.
- [55] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics, Philadelphia, Pa, USA, 2005.
- [56] S. Zhang, X. Zheng, J. M. Chen et al., "A global carbon assimilation system using a modified ensemble Kalman filter," *Geoscientific Model Development*, vol. 8, no. 3, pp. 805–816, 2015.





**Hindawi**

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

