Investigating the impact of surface soil moisture assimilation on state and parameter estimation in SWAT model based on the ensemble Kalman filter in upper Huai River basin

Yongwei Liu[1, 2], Wen Wang[1, 2 *], Yiming Hu[1]

[1] College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China. E-mail: yongweiliu@163.com
[2] State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China.
* Corresponding author. E-mail: w.wang@126.com

Abstract: This paper investigates the impact of surface soil moisture assimilation on the estimation of both parameters and states in the Soil and Water Assessment Tool (SWAT) model using the ensemble Kalman filter (EnKF) method in upper Huai River basin. The investigation is carried out through a series of synthetic experiments and real world tests using a merged soil moisture product (ESA CCI SM) developed by the European Space Agency, and considers both the joint state-parameter updating and only state updating schemes. The synthetic experiments show that with joint state-parameter update, the estimation of model parameter SOL_AWC (the available soil water capacity) and model states (the soil moisture in different depths) can be significantly improved by assimilating the surface soil moisture. Meanwhile, the runoff modeling for the whole catchment is also improved. With only state update, the improvement on runoff modeling shows less significance and robustness. Consistent with the synthetic experiments, the assimilation of the ESA CCI SM with joint state-parameter update shows considerable capability in the estimation of SOL_AWC. Both the joint state-parameter update and the only state update scheme could improve the streamflow modeling although the optimal model and observation error parameters for them are quite different. However, due to the high vegetation coverage of the study basin, and the strong spatial mismatch between the satellite and the model simulated soil moisture, it is still challenging to significantly benefit the runoff estimates by assimilating the ESA CCI SM.

Keywords: Soil moisture; EnKF; SWAT; ESA CCI SM.

INTRODUCTION

In hydrological simulation and forecasting, soil moisture (SM) produces significant impact on rainfall-runoff process. Over the past years, a large body of studies have been carried out to discuss the approaches to improving the estimation of SM, in turn to improve the streamflow simulation and forecast in hydrological modeling (Alvarez-Garreton et al., 2014, 2015; Brocca et al., 2012; Crosson et al., 2002; Crow and Ryu, 2009; Laiolo et al., 2015; Lunt et al., 2005; Reichle et al., 2008; Walker et al., 2001).

A promising approach to improving SM estimation and rainfall-runoff modeling is to integrate the observations of SM into hydrological modeling by the means of data assimilation (DA) techniques. This is due to the fact that the DA techniques, especially the newly developed sequential DA techniques (e.g., the ensemble Kalman filter, EnKF), provide an integrated framework to analyze all sources of uncertainties in hydrological modeling including model inputs, model parameters and model structures and the output observations (Wang et al., 2009). The SM data utilized in the assimilation can be obtained by two ways: field measurement and remote sensing observation. Aubert et al. (2003) assimilated the in-situ SM and streamflow in a lumped conceptual rainfall-runoff model using an extended Kalman filter, and demonstrated that the assimilation of SM data is particularly effective during flood events. Lee et al. (2011) assimilated the in situ soil moisture and streamflow data into a gridded Sacramento and kinematic-wave routing model and found that the prediction of soil moisture and streamflow significantly improved. Other studies on in-situ SM assimilation include Heathman et al. (2003), Lü et al. (2011) and Yu et al. (2012, 2014). However,

the field measurements can only provide SM on point scale, which usually fails to fully represent the spatial variability of SM. With the development of the remote sensing techniques since the 1990s, a large body of studies on assimilating satellite SM observations to improve hydrological modeling have been carried out from two aspects. One is mainly to evaluate the performance of SM assimilation in improving rainfall-runoff modeling by designing synthetic experiments (Chen et al., 2011; Crow and Ryu, 2009; Han et al., 2012; Reichle et al., 2008; Xie and Zhang, 2010). The other is to assimilate the satellite SM retrievals to improve the hydrological modeling (Alvarez-Garreton et al., 2014, 2015; Laiolo et al., 2015; Lievens et al., 2015). The popularity of satellite SM assimilation is due to the fact that the satellite observations could provide SM estimates with good spatial coverage and regular frequent time intervals, which is a valuable for rainfall-runoff modeling, especially on large scales and in sparsely monitored catchments.

To date, the assimilation of SM mainly focuses on improving the estimation of hydrological states and assumes that the parameters are relatively accurate and reliable in the model propagating process. This is mainly attributed to the present data quality of the remote sensing SM and its large uncertainties that are quite difficult to quantify. Besides, it can be related to the fact that the parameter estimation is more difficult than state estimation in hydrology as the relation between parameters and states are nonlinear in most hydrological models and the model parameters cannot be directly measured like states (Xie and Zhang, 2013). On the other hand, the parameters are hardly to be accurately tuned to represent the real physical coupling of the system based on the hydrological measurements (e.g., runoff measured at gauges) in

automatic calibration with certain objectives (Moradkhani et al., 2005). The biased parameters may lead to system bias in hydrological modeling (Chen et al., 2011; Troch et al., 2003), which will deteriorate the state update in the assimilation process (Moradkhani et al., 2005). Therefore, the joint state-parameter update are necessary in the assimilation process. In terms of SM assimilation, the parameters closely related to the soil property or soil water routing process can be expected to be estimated by assimilating the SM measurements considering the development of the remote sensing techniques (Barre et al., 2008; Das et al., 2014; Entekhabi et al., 2010) and the great efforts on SM retrievals in recent years (Das et al., 2011; Njoku et al., 2003). However, to our knowledge, the joint state-parameter estimation via the assimilation of the surface SM is mainly carried out through the synthetic experiments (e.g., Chen et al., 2015).

In this study, we not only use the synthetic experiments but also attempt to use the real world tests to investigate the performance of assimilating surface SM for the joint estimation of parameters and states (SM in different depths) in SWAT model using the ensemble Kalman filter (EnKF) method in the upper Huai River basin. As a comparison, we also analyze the impacts of updating only the state (SM in different depths) on rainfall-runoff modeling, where the parameters are not updated. By comparing the two cases, we want to check if the states and parameters can be estimated simultaneously by assimilating the surface SM, and if the joint state-parameter update scheme is superior to the scheme where only the state is updated.

Through the following sections of this article, the SWAT model and the EnKF approach for state and parameter estimation are described first. Then the study basin, the data utilized and the implementation of SM assimilation are introduced. After that, the results of surface SM assimilation are analyzed. Finally, the conclusions are drawn in the last section.

METHODOLOGY
Soil and water assessment tool (SWAT)

SWAT is a physically based basin scale distributed model developed by the USDA (United States Department of Agriculture) Agricultural Research Service (USDA-ARS) (Neitsch et al., 2011). It has been widely used in rainfall-runoff modeling in the field of hydrology in recent years (Chen et al., 2011; Xie et al., 2014). For hydrological modeling, the catchment is geographically divided into subbasins; then, each subbasin is further delineated into several Hydrological Response Units (HRUs) with the same information on land cover, soil and slope steepness. HRUs are the basic calculation units for the land phase of hydrologic cycle including the surface runoff generation, evapotranspiration, soil water routing, and groundwater generation processes.

Soil moisture plays a central role in the above processes in SWAT simulations. The water infiltrated or percolated to soil profile after surface runoff generation is redistributed using a storage routing technique with its field capacity as a threshold. The water balance equation for each soil layer can be expressed as:

$$SW'_{ly} = SW_{ly} + \Delta w_{perc,ly} - Q_{lat,ly} - E_{a,ly} \quad (1)$$

where SW_{ly} and SW'_{ly} are the soil water content (mm) at the start and end of the day respectively, $\Delta w_{perc,ly}$ is the net percolation received in the layer ly (i.e. the percolation or infiltration from the overlying layer minus that to the next layer), $Q_{lat,ly}$ is the

lateral flow generated from layer ly, $E_{a,ly}$ is the evapotranpiration drawn from the layer ly.

The percolation ($w_{perc,ly}$) from layer ly is calculated as:

$$w_{perc,ly} = (SW_{ly} - FC_{ly}) \cdot (1.0 - \exp\left[\frac{-24 \cdot K_{sat,ly}}{SAT_{ly} - FC_{ly}}\right]) \quad (2)$$

where FC_{ly} is the soil water content of layer ly at field capacity (mm), $K_{sat,ly}$ is the saturated hydraulic conductivity (mm/h), SAT_{ly} is the amount of water in layer ly when completely saturated (mm). The percolation generated from the bottom of the sol profile is the source of ground water recharge.

The lateral flow ($Q_{lat,ly}$) of layer ly is calculated by:

$$Q_{lat,ly} = \frac{0.048 \cdot (SW_{ly} - FC_{ly}) \cdot K_{sat,ly} \cdot slp}{(\phi_{por,ly} - \phi_{fc,ly}) \cdot L_{hill}} \quad (3)$$

where slp is the slope steepness (m/m), L_{hill} is the hill slope length, $\phi_{por,ly}$ is the total porosity of the soil layer (mm/mm) and $\phi_{fc,ly}$ is the porosity of the soil layer ly filled with water when it is at field capacity.

The evapotranpiration drawn from the soil is mainly performed by soil evaporation and plant uptake. The potential evapotranspiration and the potential plant uptake are simulated first using the Penman-Monteith equation (Monteith, 1965). The plant uptake under ideal conditions is a function of the potential plant uptake and the leaf area index. The potential soil water evaporation and the ideal plant water uptake in each layer on a profile are estimated using the depth distribution function. The actual soil water evaporation and plant uptake are constrained by the available soil water of a given layer and are not allowed to be compensated by other layers. However, this deficiency can be made up by adjusting the soil compensation factor (ESCO) and the plant compensation factor (EPCO) to modify the depth distribution of the potential soil water evaporation and the ideal plant uptake.

The surface runoff, lateral flow and ground water generated on each HRU are aggregated to the main channel. SWAT incorporates a storage feature to calculate the surface runoff and lateral flow that can reach the main channel on current day, while a technique similar to the linear reservoir method is used to route the ground water to the main channel. The water routing through the channel network is realized by the variable storage routing method (Williams, 1969).

The ensemble Kalman filter for state and parameter estimation

The Ensemble Kalman filter (EnKF) is a sequential DA approach, which was first introduced by Evensen (1994). It incorporates the Monte Carlo method to generate a state ensemble to represent the probability distribution of the state. The state forecast is then achieved by propagating the ensemble members forward in time using the model independently with a stochastic term representing the model errors (Evensen, 1994). The forecasted state ensemble θ_t^f at time t is given by:

$$\theta_t^f = F(\theta_{t-1}^u, u_t) + w_t \quad w_t \sim N(0, r^2) \quad (4)$$

where F represents the SWAT model in this study. θ_{t-1}^u is the updated state ensemble at time $t-1$. In this study, it includes the SM of four layers for all HRUs of the watershed. u_t is the mod-

el forcing input at time t, which is assumed to follow a normal distribution. w_t is the stochastic perturbation of the forecast states, which reflects the model structure error and is assumed to be normally distributed with mean 0 and covariance of r^2.

The state update is obtained by:

$$\theta_t^u = \theta_t^f + K_t[Z_t - H(\theta_t^f)] \tag{5}$$

where Z_t is the observation matrix at time t, which includes the SM observations with a stochastic perturbation being assumed to be normal distribution with mean 0 and standard deviation of ε. H is observation operator, which is used to map the model states to the observations. In the synthetic experiments, Z_t includes the surface SM of all subbasins, and H is constructed by 0 and the area proportions of HRU in subbasins as the simulated SM are on HRU level while the synthetic surface SM observations are on subbasins. In the real world test, Z_t are the surface SM on the remote sensing grids covering the whole basin, so, H is constructed by 0 and the area proportions of HRU in the grids as the simulated SM are on HRU level while the SM observations are on the grids. K_t is the Kalman gain at time t, which defines the weight of modeling and observation and is calculated by the forecast error covariance and the observation error covariance:

$$K_t = P_{ms,t}\left(P_{s,t} + R_{s,t}\right)^{-1} \tag{6}$$

where $P_{ms,t}$ is the cross error covariance between the predicted states θ_t^f and the measurement prediction $H(\theta_t^f)$ at time t, $P_{s,t}$ is the error covariance of the measurement prediction at time t, $R_{s,t}$ is the error covariance of the observations at time t.

The joint state-parameter updating in state augmentation technique is done by concatenating the states (θ_t^f) with the parameters (δ_t^f) in the state vector, see equation (7), and the parameter update is the same with that of the state update described in equations (5) and (6)):

$$\begin{bmatrix}\theta\\\delta\end{bmatrix}_t^u = \begin{bmatrix}\theta\\\delta\end{bmatrix}_t^f + K_t^*(Z_t - H^*\begin{bmatrix}\theta\\\delta\end{bmatrix}_t^f) \tag{7}$$

Since the parameters (δ_t^f) are not observed, the observation operator for the joint state-parameter updating is:

$$H^* = \begin{bmatrix}H & 0\end{bmatrix} \tag{8}$$

K_t^* is the Kalman gain for joint state-parameter update. Its calculation is similar to the Kalman gain (K_t) in equation (6) but with $P_{ms,t}$ changing to the cross error covariance between the predicted states and parameters $(([\theta\ \delta]_t^f)^T)$ and the measurement prediction $H(\theta_t^f)$ at time t.

It is worth noting that the joint state-parameter updating requires an evolution of the parameters to avoid ensemble shrinkage and filter degeneracy. In this study, the conditional covariance inflation method (Aksoy et al., 2006) is adopted as described by:

$$\begin{aligned}\delta_t^f &= \delta_{t-1}^u & Var(\delta_{t-1}^u) \geq T\\\delta_t^f &= \delta_{t-1}^u + v_t & Var(\delta_{t-1}^u) \leq T, \ v_t \sim N(0, T - Var(\delta_{t-1}^u))\end{aligned} \tag{9}$$

where δ_t^f is the forecast parameter ensemble at time t and δ_{t-1}^u

is the updated parameter ensemble at time t–1. T is the prescribed threshold set as the initial variance multiplied by a scaling factor γ in this study. Considering that model blow-up commonly occurred in the assimilation process due to over updating to the parameters, a smoothing factor β (between 0 and 1.0) is applied to dampen the rapid temporal variation of parameters (Yang and Delsole, 2009):

$$\delta_t^f = (1-\beta)\delta_{t-1}^u + \beta\delta_{t-1}^f \tag{10}$$

where $\beta = 0$ reflects that there is no smoothing in the parameter updating process, the larger value of β indicates the stronger smoothing on parameter update.

STUDY AREA AND DATA USED

The study catchment is located in the upper Huai River basin within $113°15'E \sim 113°55'E$ and $32°10'N \sim 32°45'N$ in China (Figure 1). It covers an area of about 1640 km² with the elevation ranging from 109 m to 953 m. The basin is located in the transition zone between the northern subtropical and the warm temperate zones. Average annual precipitation is about 900 mm, of which 50% to 80% falls during June to September. The dominant land cover is rice (56%) and forest (22%). Sandy loams and silt loams are the main soil types.

The SWAT model requires the meteorological forcing data and the underlying land surface data. The meteorological data are the precipitation, maximum/minimum air temperature, solar radiation, wind speed and relative humidity. The precipitation is provided by 12 local rainfall gauges in the basin (Figure 1). Other five meteorological data above are collected from Xinyang meteorological station near the basin. The land surface data includes the digital elevation data (DEM), soil category and the land cover data. The DEM data is downloaded from the Shuttle Radar Topography Mission (SRTM) with a spatial resolution of 90 m (http://datamirror.csdb.cn/index.jsp). The land use data is resampled from a year-1995 land use map at a scale of 1:210000 provided by the government of Xinyang city. The soil data is resampled from a soil map at a scale of 1:100000 collected from Soil Handbook of Henan province. According to this handbook, the soil of this basin can be classified into 5 categories, of which the soil texture and the corresponding USDA (United States Department of Agriculture) classification as well as the area proportions are given in Table 1. Based on the above data sets, the catchment is divided into 15 subbasins. The subbasins are further divided into 55 HRUs. Besides, the soil on profile is dived into four layers (0–50 mm, 50–200 mm, 200–400 mm and 400–800 mm) from top to bottom, except for Cugutu with two layers (0–50 mm, 50–200 mm).

The ESA CCI SM product was developed in the Climate Change Initiative (CCI) projects by the European Space Agency (ESA). It combines the soil moisture retrievals from six passive (SMMR, SSM/I, TMI, AMSR2, AMSR-E and Windsat) and two active (ERS AMI and ASCAT) microwave sensors into a global data set spanning over 30 years from 1978 to 2014 (product version 02.2). The ESA CCI SM product consists of three surface SM datasets: the active, passive and combined product. The active and passive datasets were created by fusing scatterometer and radiometer SM products respectively, while the combined product is a merged product of both the active and passive products. We use the combined ESA CCI SM product in this study, which is downloaded from
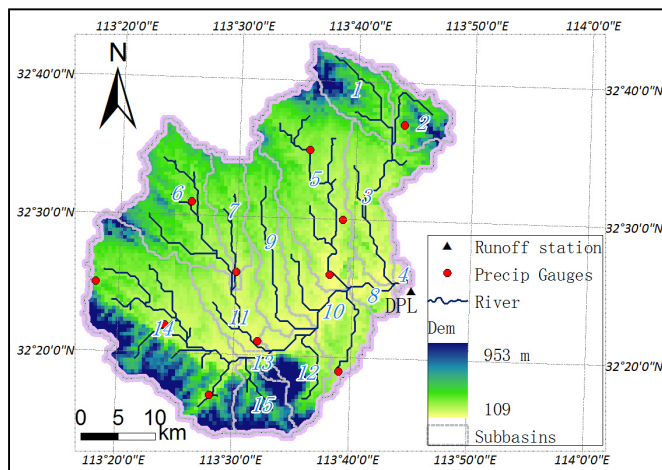
Fig. 1. The upper Huai River catchment. Note: the number 1 to 15 represents the subbasin number. DPL means Dapoling runoff station.

(http://www.esa-soilmoisture-cci.org/node/145) with a spatial resolution of 0.25° and a temporal resolution of 1 day with the reference time at 0:00 UTC.

IMPLEMENTATION OF SURFACE SOIL MOISTURE ASSIMILATION

The uncertainties of hydrological modeling are included in the meteorological forcing inputs, model parameters and model structures. These uncertainties are quantified by the variance of the ensemble members in EnKF, and each ensemble member is generated by the perturbations on model forcing inputs (including the precipitation and the minimum and maximum temperature), sensitive model parameters and the predicted SM presented in Table 2. Precipitation at each site is perturbed with a multiplicative error (Clark et al., 2008; Xie and Zhang, 2010) to avoid negative values after perturbation. The standard deviations (STD) of the multiplicative error (σ_p) reflect the instrument error and the error of the spatial representativeness of the rainfall gauges. Both the maximum and minimum temperature are perturbed by an additive error with a STD of 1.0℃, which is also adopted by Chen et al. (2011). The model states (the SM at 4 layers) are perturbed using Gaussian multiplicative error to account for the uncertainties in model structures. This small error design of 0.01 is to avoid rapid changes of soil water content between continuous time steps. Besides, it prevents the spurious correlations between the surface SM and the SM in depths, especially when the model has weak vertical coupling in soil water routing process. In addition, the parameters sensitive to SM are assumed to be uniformly distributed within their potential value ranges. Regarding to the uncertainty of the surface SM observation, an additive error of 0.04 mm/mm is adopted (Crow and Wood,

2003; Reichle et al., 2008). In the conditional covariance inflation approach, the scaling factor γ used to determine the threshold T is set to 0.01 and the smoothing factor β is assumed 0.8 (Yang and Delsole, 2009) to dampen the rapid temporal variation of parameters during parameter evolution process.

In the synthetic experiment, three sets of data are generated, i.e., the synthetic truth run, ensemble open loop (EnOL) run and EnKF run. The synthetic truth is one realization of SWAT model running by adding perturbations on model inputs and model parameters. The synthetic surface SM observations (0–50 mm) on subbasins are generated by adding the observation errors (Table 2) to the synthetic "true" subbasin average SM from the assumed true model running process. The EnOL run is an ensemble running of SWAT model with perturbations on model inputs, model parameters and model states (SM in different depths) described in Table 2. The EnKF run is an ensemble running of SWAT model with the same error assumption to EnOL, but with the integration of the synthetic surface SM observations during the model propagation process. Here, two schemes are considered in the EnKF run: (1) the EnKF_SP, where both the SM and parameters sensitive to SM are simultaneously updated via EnKF method; (2) the EnKF_S, where only the SM are updated after the bias caused by the parameters is removed by the joint state-parameter update scheme of the first 200 days. The ensemble size of EnOL and EnKF is set as 100 considering the requirement on the accuracy and the computational burden. The performance of SM assimilation is illustrated by comparing the results of EnKF to that of EnOL based on the synthetic truth. In order to ensure the robustness of the analysis results, ten synthetic truths are generated considering the possible different performance of SM assimilation in different conditions (dry or wet situations). The model initialization is from Jan 1, 2005 to Dec 31, 2005, and the assimilation is from Jan 1, 2006 to Dec 31, 2008.

In the real world test, the combined ESA CCI SM product is assimilated. The assimilation of satellite soil moisture is a complex task, the success of which highly relies on the appropriate assumptions on the rescaling techniques and the model and observation errors (Massari et al., 2015). Here, the cumulative distribution function (CDF) matching method (Reichle and Koster, 2004) is used with 3 year training window (from 2003-1-1 to 2005-1-1) to remove the biases between the satellite SM and the model simulated SM before assimilation, which have been proved to be effective in many previous researches (e.g., Alvarez-Garreton et al., 2014). Then the biased removed ESA CCI SM data are assimilated for the period from Jan 1, 2006 to Dec 31, 2008.

In terms of the error set for model and observations, we will assume an appropriate value first, and then attempt to explore the sensitivity of the assimilation results to the error parameters of model and observations. For the model error, a Gaussian multiplicative perturbation with the standard deviation (σ_p) of 0.4 is put on the precipitation; the parameters sensitive to SM (i.e., SOL_AWC for different soil types) are perturbed by a

Table 1. Soil classification and its physical properties for the study basin.

Chinese soil type	Clay (%)	Silt (%)	Sand (%)	Rock (%)	USDA soil texture	Area coverage (%)	Φ_{por} (mm/mm)	Φ_{wp} (mm/mm)
Cugutu	7.05	34	35.9	23.1	Sandy loam	40.6	0.38	0.05
Huanghetu	23.43	65.1	11.5	0	Silt loam	23.64	0.42	0.14
Huangzongrang	17.03	39.4	43.6	0	Loam	10.18	0.39	0.11
Huichaotu	12.86	51.8	35.3	0	Silt loam	1.01	0.38	0.08
Shuidaotu	16.46	71	12.5	0	Silt loam	24.57	0.41	0.11

Table 2. Generation of perturbations on model and observations for each ensemble member.

Variables	Error distribution	Mean	Standard deviation	Bound
recipitation	Gaussian, Multiplicative	1	0.2	(0, −)
Max temp	Gaussian, Additive	0	1.0 (℃)	−
Min temp	Gaussian, Additive	0	1.0 (℃)	−
Predicted SM	Gaussian, multiplicative	1	0.01	(0, −)
Parameters (SOL_AWC)	Uniform distribution	−	−	given in Section Parameter sensitivity
Observed SM	Gaussian, Additive	0	0.04 (m^3/m^3)	(0, −)

Note: "Max", "Min" and "temp" means maximum, minimum, and temperature respectively.

Gaussian additive noise with the standard deviation of 0.05 mm/mm to ensure that the perturbed parameters stay in physical domain detailed in the Section of Parameter sensitivity. Besides, the error set for model inputs of the maximum and minimum temperatures and the model states are identical to that of the synthetic experiment. For the observation error, an upper bound of the rescaled ESA CCI SM error is firstly determined, and a scaling factor (SF) of 80% is firstly attempted (Alvarez-Garreton et al., 2014). Similarly, the two schemes (EnKF_SP and EnKF_S) in the EnKF run are considered. In EnKF_SP both the parameters and SM in different depths are updated. However, in EnKF_S only the SM in different depths are updated while the perturbed parameters with Gaussian distribution remains unchanged.

RESULTS
Parameter sensitivity analysis

Theoretically, the assimilation is only effective when there is a strong correlation between the model states being updated and the observations being assimilated (Xie and Zhang, 2013). This principle is also appropriate for parameter estimation, which means that it is impossible to correct parameters that cause smaller scale errors in the predicted observation equivalent of the model states than can be observed (Smith, 2010). Therefore, it is crucial to evaluate the sensitivity of the given parameters to the measurement model predictions. Here, four parameters related to soil property and soil water routing processes are pre-selected and described in Table 3. The sensitivity analysis of the four parameters to the surface SM is conducted based on the theory of Latin Hypercube (McKay et al., 1979) and One-factor-At-a-Time (LH-OAT) sampling approach (Morris, 1991). For each parameter, an ensemble of 100 members are sampled from its value range (Table 3) using a stratified sampling method at the start of modeling. Other remaining parameters are kept at their calibrated values. After running the model, the corresponding standard deviation (STD) of the surface SM for 100 members is calculated. The larger STD values (reflect the large ensemble spread) of the surface SM ensemble indicates higher sensitivity of the corresponding parameter. Figure 2 demonstrates the sensitivity results of the four parameters on HRU-1 as an example. It is clear that SOL_AWC is the most sensitive among the four parameters. The other three parameters only present certain sensitivity occasionally: the saturated hydraulic conductivity (K_{sat}) is sensitive only during rainy periods,

while the soil evaporation compensation factor (ESCO) and the plant uptake compensation factor (EPCO) are only sensitive in dry periods. Therefore, only SOL_AWC is finally selected and will be updated in this study. When taking samples of SOL_AWC for doing the synthetic experiment or perturbing SOL_AWC for doing the real world test, we should ensure that the value falls within its physical domain. As SOL_AWC is the difference between field capacity Φfc and permanent wilting point Φwp, and Φfc is less than the total soil porosity Φpor, therefore, the physical domain of SOL_AWC should be between zero and Φpor − Φwp, that is, (0, 0.33), (0,0.28), (0, 0.28), (0, 0.3) and (0, 0.3) for Cugutu, Huanghetu, Huangzongrang, Huichaotu and Shuidaotu, respectively, according to the values of Φpor and Φwp given in Table 1.
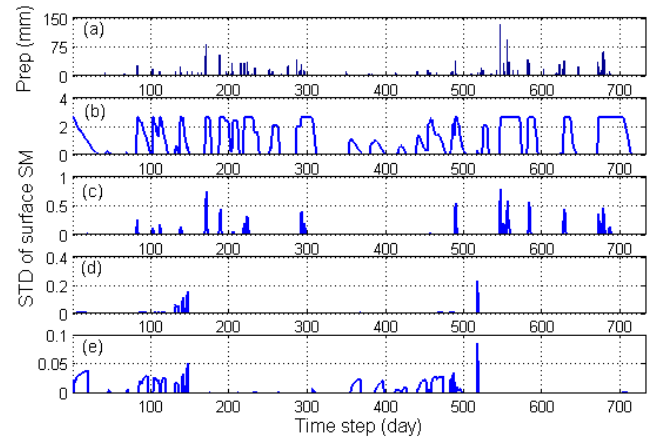

Fig. 2. Variations of the catchment average precipitation (Prep) and the standard deviation (STD) of surface soil moisture (SM) for four parameters (Table 3) on HRU-1 from 1994-1-1 to 1995-12-31. (a) precipitation (b) SOL_AWC (c) K_{sat} (d) ESCO (e) EPCO.

Results for the synthetic experiments
Parameter and state estimation

Figure 3 shows the assimilation process of SOL_AWC for the 5 types of soil in one experiment randomly chosen from the ten experiments. It is found that the SOL_AWC for 4 types of soil (except for Huichaotu, Figure 3(d)) can be accurately estimated by assimilating the synthetic surface SM observations.

Table 3. Parameters closely related to soil moisture in SWAT.

| Parameter	Units	Description	Value range
SOL_AWC	mm /mm	Available water capacity of the soil, which is calculated by subtracting the fraction of water present at permanentwilting point from that present at field capacity	(0.01, 1.0)
K_{sat}	mm/hour	Saturated hydraulic conductivity	(0.0, 250)
ESCO	−	Soil evaporation compensation factor	(0.01, 1.0)
EPCO	−	Plant uptake compensation factor	(0.01, 1.0)
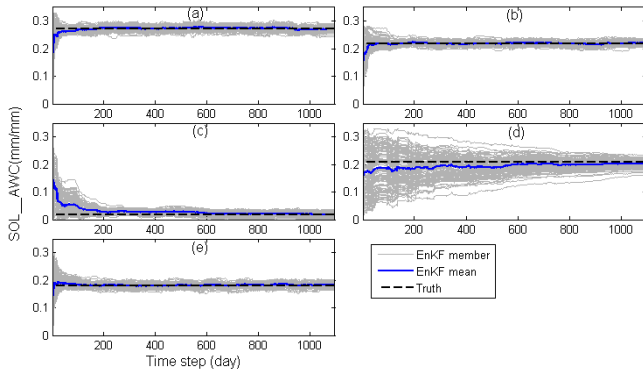
Fig. 3. Estimation of the parameter SOL_AWC for the 5 types of soil from 2006-1-1 to 2008-12-31 by assimilating the synthetic surface soil moisture observations. (a) Cugutu (b) Huanghetu (c) Huangzongrang (d) Huichaotu (e) Shuidaotu.

Although there is a bias (i.e., a difference between the ensemble mean and the synthetic truth) at the start of assimilation, after about 200 days at maximum, all the ensemble members of SOL_AWC converge to the truth and the ensemble mean matches the truth well. This indicates the significant capability of surface SM assimilation to eliminate parameter bias and improve the estimation of SOL_AWC. The non-convergence and underestimation of SOL_AWC for soil Huichaotu is partly caused by its little area proportion (1.01%) in the study basin. The change of SOL_AWC for Huichaotu cannot cause large difference in the surface SM modeling on subbasins, thus its estimation by assimilating the surface SM on subbasin scale contains certain difficulty. The results are similar for other nine experiments.

Figure 4 provides the catchment area average SM of four layers (0–50 mm, 50–200 mm, 200–400 mm and 400–800 mm) from top to bottom in EnOL, EnKF_SP, EnKF_S and the synthetic truth, which are obtained from the same experiment mentioned above. The SM estimated by EnKF_SP is significantly improved compared with that by EnOL, especially for the upper two layers, as it significantly approaches the synthetic truth during the assimilation process. This can be illustrated from the decease of the root mean square error (RMSE) by EnKF_SP compared with that by EnOL (Table 4). Nevertheless, it seems that the effects of EnKF_SP on SM estimation show a decreasing trend with the increase of soil depth, which can be revealed from the decreasing degree of RMSE reduction in Table 4. This decreased efficiency is certainly influenced by the weak vertical coupling in SWAT model (Chen et al., 2011), which weakens the correlation between the surface SM and the SM in depths and deteriorates the SM update in deep layers in the assimilation. Besides, it can be seen in Figure 4 that the EnKF_S obtains similar performance on SM estimation to EnKF_SP, which indicates that the joint state-parameter update after 200 days does not make large difference to that with only state update on SM estimation. This is because the parameter SOL_AWC has gradually converged to stable values during the first 200 days by joint state-parameter update in EnKF_S. It makes the observations having little effect on the parameters after 200 days in EnKF_SP. Thus, the EnKF_S and EnKF_SP show similar influence on SM estimation. However, in terms of the RMSE in Table 4, the EnKF_S produces a little inferior impacts on the SM estimation of the fourth layer compared with that of the EnKF_SP (Table 4). This might relate to the lack of the possible temporal evolution of the model parameters in
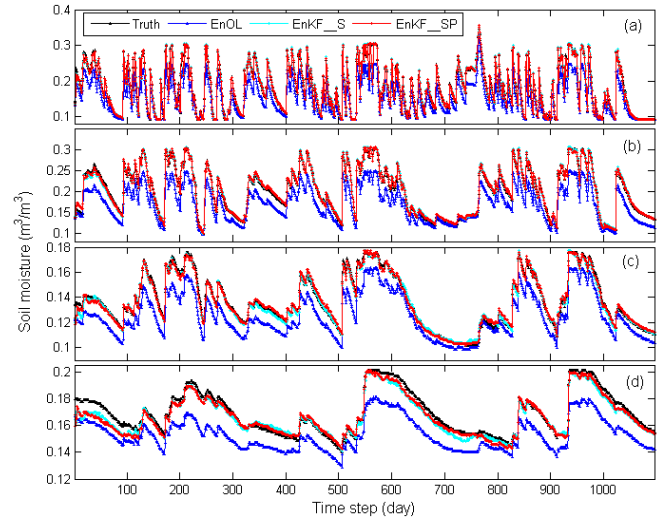

Fig. 4. The catchment area average soil moisture for the four layers from top to bottom estimated in the EnKF and EnOL from 2006-1-1 to 2008-12-31. (a) the first layer (0–50 mm), (b) the second layer (50–200 mm), (c) the third layer (200–400 mm) and (d) the forth layer (400–800 mm).

Table 4. RMSE of the catchment area average soil moisture obtained by EnOL, EnKF_S and EnKF_SP in the synthetic experiment (based on the data for the whole data assimilation period except for its first 200 days).

RMSE (m^3/m^3)	Layer 1	Layer 2	Layer 3	Layer 4
EnOL	0.0323	0.0353	0.0125	0.0175
EnKF_S	0.0025	0.0031	0.0021	0.0038
EnKF_SP	0.0022	0.0032	0.0020	0.0029

Note: Layer 1, Layer 2, Layer 3 and Layer 4 represent the first (0–50 mm), second (50–200 mm), third (200–400 mm) and the forth layer (400–800 mm) respectively.

EnKF_S as the parameters remain unchanged after 200 days (Moradkhani et al., 2005).

Effectiveness of data assimilation on ainfall-runoff modeling

The improvement of SM estimation can be expected to improve the rainfall-runoff simulation and prediction (Alvarez-Garreton et al., 2014, 2015; Brocca et al., 2012). In order to analyze the performance of surface SM assimilation on rainfall-runoff modelling, the normalized error reduction of RMSE (NER) is used:

$$NER = 1.0 - \frac{RMSE_{EnKF}}{RMSE_{EnOL}} \tag{11}$$

where $RMSE_{EnKF}$ and $RMSE_{EnOL}$ are the RMSE of the EnKF and EnOL mean based on the synthetic truth. The larger values of NER indicate the relatively better performance of EnKF.

Figure 5 provides the NER of the ensemble average states and flux (Table 5) for the ten synthetic experiments. It can be seen that the total runoff at the catchment outlet (V_{out}) is considerably improved by the EnKF_SP and EnKF_S in the soil moisture assimilation. It seems that the V_{out} obtained by EnKF_SP has a little better estimation than that by EnKF_S as it obtains a bit higher median value of NER. The improvement of the total runoff is contributed by the improvement of the three
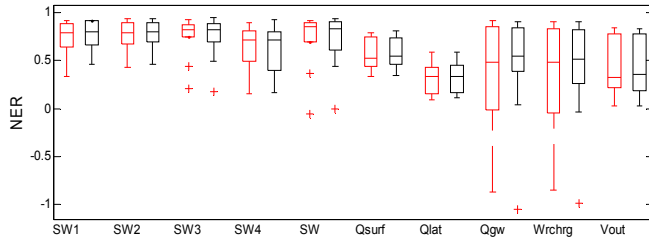
Fig. 5. NER (equation 11) of the EnKF_S (red box) and EnKF_SP (black box) ensemble averaged variables (Table 5) for the ten synthetic experiments (based on the data for the whole data assimilation period except for its first 200 days). The upper limb, lower limb and the red or black line in the box represent the upper quartile (here, represented by h1), lower quartile (represented by h2) and the median value of NER for the ten experiments. The dotted line stretched from the box indicates the range of h1-1.5(h2-h1) and h2 +1.5(h2-h1), while the red "+" is the strange values beyond this range.

different streamflow components including the surface runoff (Q_{surf}), lateral flow (Q_{lat}) and ground water (Q_{gw}), in which Q_{surf} and Q_{gw} take the major part (about 80%) of the total runoff. The surface runoff improvement is caused by the improved SM on profile, which obtains a large degree of improvement by surface SM assimilation. The ground water primarily benefits from the improvement of the soil water recharge (w_{rchrg}) to the shallow and deep aquifers. w_{rchrg} is the only source of ground water, which is influenced by the improvement of the soil water distribution on profile and the field water capacity (equation 2) that is controlled by SOL_AWC. Compared to the EnKF_SP, the worse and less robust performance of EnKF_S on ground water resulted from the bad performance of that on w_{rchrg}, which also contribute the less significant improvement of the total runoff. This unsatisfactory performance of EnKF_S might be related to the fact that the parameter ensemble of SOL_AWC will not be propagated and corrected after 200 days, which might make the lack consideration of the possible temporal evolution of the model parameters in the model propagating process (Moradkhani et al., 2005). Overall, the EnKF_SP outperforms the EnKF_S in both the improvement level and stability of the rainfall-runoff process by the surface soil moisture assimilation. This also indicates the importance of the joint state-parameter update in soil moisture assimilation.

Results for the real world tests
Model calibration and validation

The model is calibrated using the daily runoff records from Jan 1, 2002 to Dec 31, 2005 observed at the basin outlet (Dapoling station, Figure 1). The parameter optimization is achieved by a combination of auto calibration using the Sequential Uncertainty Fitting (SUFI2) (Abbaspour et al., 2004) with the Nash-Sutcliffe coefficient of efficiency (NSE,
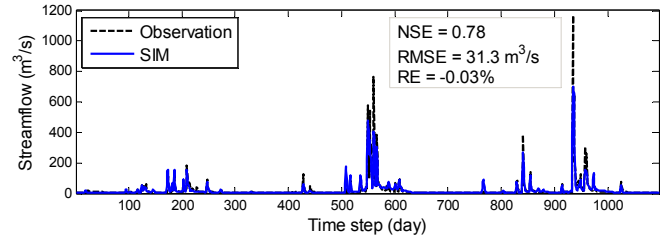

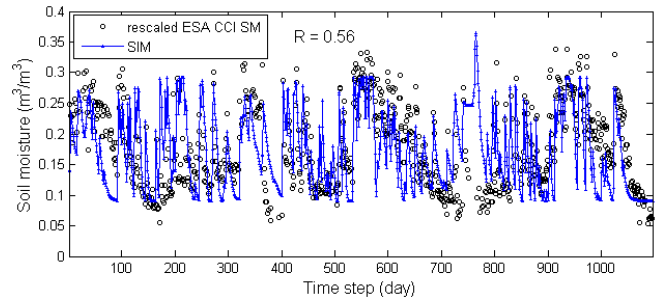**Fig. 6.** Validation of streamflow at Dapoling station from 2006-1-1 to 2008-12-31.

Fig. 7. The model simulated soil moisture and rescaled ESA CCI SM on catchment scale from 2006-1-1 to 2008-12-31.

Nash and Sutcliffe, 1970) as the objective criteria and the manual fine-tuning method. The NSE, the root mean square error (RMSE) and the relative error (RE = (1.0 – sum(SIM)/sum(OBS)) * 100%, where sum(SIM) and sum(OBS) represent the summarized value of the simulated and the observed streamflow series for the whole calculation period separately) at Dapoling station for the calibration period are 0.74, 44.38 m³/s and 2.43% respectively. Figure 6 shows the validation of streamflow at Dapoling station from Jan 1, 2006 to Dec 31, 2008. The NSE, RMSE and RE are 0.78, 31.3 m³/s and –0.03% respectively for this period. The model simulated soil moisture and the rescaled ESA CCI SM on the basin in this validation period are present in Figure 7. It can be seen that the model simulated soil moisture varies more frequently than that of the ESA CCI SM. The Pearson's correlation coefficient between them is 0.56.

Parameter estimation and validation

Figure 8 presents the SOL_AWC for 5 types of soil by assimilating the combined ESA CCI SM products from Jan 1, 2006 to Dec 31, 2007 in EnKF_SP. After about 200 days, the ensemble trajectories are nearly stable with little variations, and the ensemble samples can be approximated by a Gaussian distribution with a small standard deviation (STD). This indicates that the uncertainty of the parameter estimation can be decreased with the joint state-parameter update in EnKF_SP. The relatively stable ensemble trajectory and the little ensemble spread

Table 5. The variables in SWAT model marking the rainfall-runoff process.

| Variables	Units	Description
SW_{ly}	mm	Watershed average soil water content in layer *ly* on a given day
SW	mm	Watershed average profile soil water content on a given day
Q_{surf}	mm	Amount of surface runoff discharged to the main channel on a given day
Q_{lat}	mm	Amount of lateral flow discharged to the main channel on a given day
w_{rchrg}	mm	Amount of water entering both shallow and deep aquifers on a given day
Q_{gw}	mm	Amount of ground water flow discharged to the main channel on a given day
V_{out}	m³/s	The volume of water flowing out of the channel during a time step
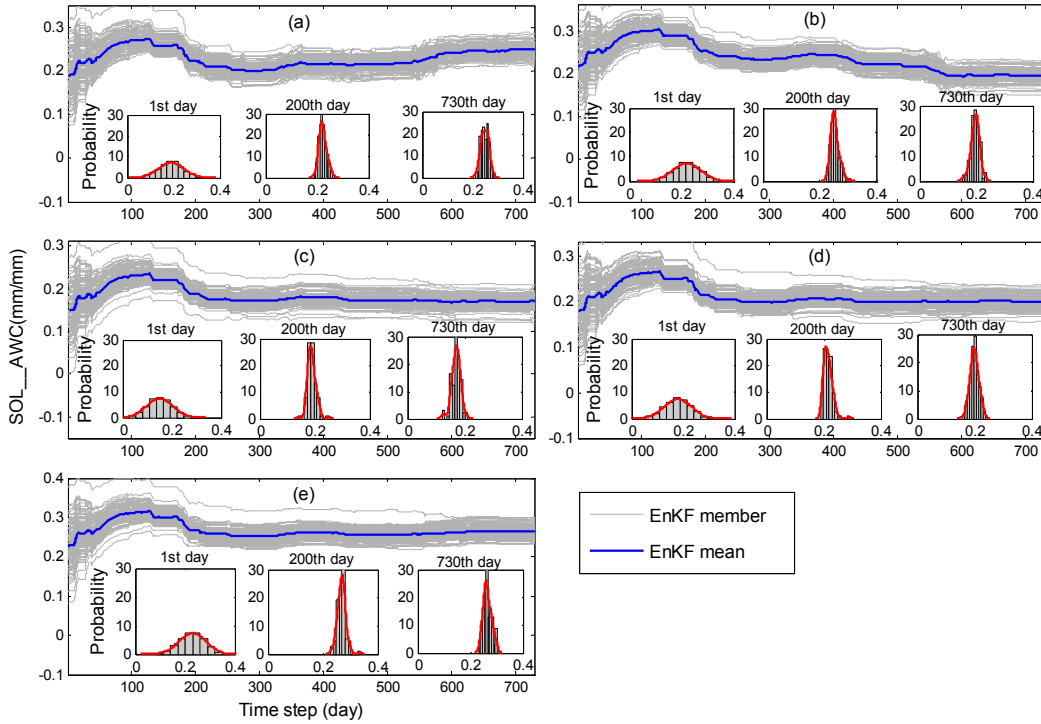
Fig. 8. Estimation of the parameter SOL_AWC for the 5 types of soil from 2006-1-1 to 2007-12-31 by assimilating the combined ESA CCI SM. The histograms in each plot, fitted with the Gaussian distribution function, represent the distribution of the ensemble members at three time steps (i.e. 1st, 200th and 730th day). (a) Cugutu (b) Huanghetu (c) Huangzongrang (d) Huichaotu (e) Shuidaotu.
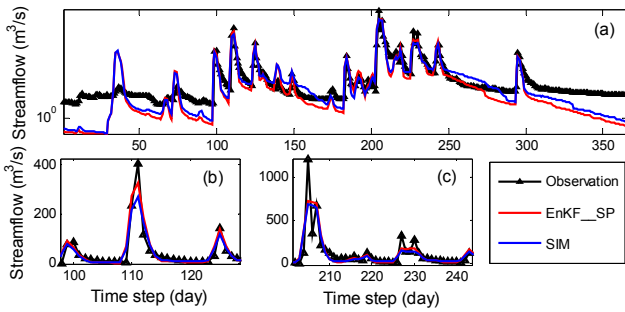

Fig. 9. The predicted streamflow using different parameter sets from 2008-1-1 to 2008-12-31.

(reflected by the STD of the parameter ensemble) imply the potential that the SOL_AWC can be estimated probably by assimilating the combined ESA CCI SM. However, it is not sure whether the SOL_AWC is appropriately estimated, because we cannot obtain its the true value. Here, we attempt to use the streamflow measurements at the catchment outlet for parameter validation. We perform a single run of SWAT model from Jan 1, to Dec 31, 2008, using the parameter estimates from EnKF_SP, and evaluate the predicted streamflow against the observed streamflow at the basin outlet (Dapoling station). Besides, the streamflow in the control run, i.e. the simulated streamflow using the prior calibrated parameters in the section of Model calibration and Validation, is adopted as a comparison. Figure 9 shows the predicted streamflow based on the above two parameter sets, the EnKF_SP calibrated parameter (marked by EnKF_SP) and the prior calibrated parameter (marked by SIM). It can be seen that the high flow obtained by the EnKF_SP calibrated parameters is considerably improved compared to the streamflow in control run (e.g., the day around
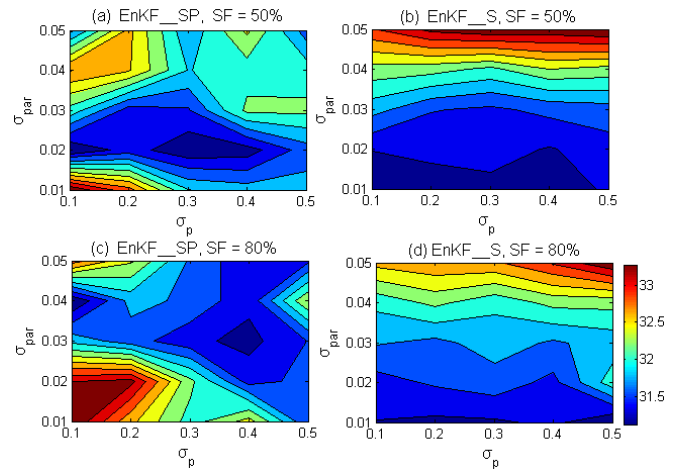

Fig. 10. Contour plot of the RMSE for the outlet streamflow estimated by EnKF_SP and EnKF_S from Jan 1, 2006 to Dec 31, 2008, as a function of the standard deviation (STD) of the multiplicative Gaussian perturbations to the precipitation (σ_p), the STD of the Gaussian perturbations on SOL_AWC (σ_{par}), and the scaling factor (SF) to the upper bound of the observation error for the rescaled ESA CCI SM.

110). However, the improvement was not obtained for the low flow. Simulated streamflow become even worse for some periods (e.g., the day over 300). For the whole period, the streamflow by EnKF_SP only obtains a slight improvement, as the RMSE is decreased from 39.80 to 39.72 m^3/s and the NSE increased from 0.77 to 0.78. Although this improvement is not significant, it still shows the capability of the satellite soil moisture assimilation in parameter optimization.

Sensitivity of the model and observation error

Figure 10 shows the root mean square error (RMSE) of the outlet streamflow obtained by EnKF_SP and EnKF_S with varying standard deviation (STD) of the Gaussian multiplicative perturbations on precipitation (σ_p) from 0.1 to 0.5 along with the varying STD of the Gaussian additive perturbations on parameter SOL_AWC (σ_{par}) from 0.01 to 0.05 mm/mm. At the same time, the scaling factors (SF) to the upper bound (0.067, 0.066, 0.070 and 0.065 m^3/m^3 for the four remote sensing grids covering the basin) of the observation error for the rescaled ESA CCI SM are considered at 50% and 80%, while the error of the model inputs for the maximum and minimum temperature and the model states (soil moisture in different depths) remain unchanged at 1.0°C and 0.01 respectively. Figures 10 (a) and (c) show that, in terms of the RMSE for the outlet streamflow, the performance of EnKF_SP is sensitive to both the error of SOL_AWC (σ_{par}) and the precipitation input (σ_p). The optimal performance of EnKF_SP falls in different domain of σ_p and σ_{par} with different observation errors (i.e. 50% and 80% of SF). Note that the RMSE of the simulated streamflow in the control run is 31.3 m^3/s (see the section of Model calibration and Validation). The good impacts of EnKF_SP on streamflow modeling can be obtained when σ_{par} is around 0.02 and σ_p being limited within 0.43 under the case that SF kept at 50% (Figure 10 (a)), and when SF equals 80%, the optimal σ_{par} and σ_p are around 0.03 and 0.4 respectively (Figure 10 (c)). Figures 10 (b) and (d) show that, the EnKF_S performance is sensitive to the magnitude of the parameter perturbation (σ_{par}). With the increase of σ_{par}, the performance of EnKF_S on streamflow modeling becomes worse. The EnKF_S will produce negative influence on the streamflow modeling once σ_{par} exceeds 0.02 when SF is kept at 50% (Figure 10(b)), and the EnKF_S effects will be negative once σ_{par} exceeds 0.01 when SF is kept at 80% (Figure 10(d)). The bad performance of EnKF_S is due to the overfitting of the model simulated soil moisture (SM) to that of the satellite soil moisture, as the large perturbation to SOL_AWC will lead to large ensemble spread of the predicted SM. That in turn results in large forecast error covariance between the model predicted SM and the measurement predictions, which makes very large (or unreasonable) SM updates in the assimilation process. The different sensitivity of the EnKF_SP and EnKF_S performance to the model and observation error parameters indicate the capability of the remote sensing SM on parameter estimation. Overall, the EnKF_SP and EnKF_S have the capability to improve the streamflow modeling by assimilating the combined ESA CCI SM although the optimal model and observation error parameters for them are quite different. This also indicates the significance of an appropriate error assumption/quantification on the model and observations in satellite soil moisture assimilation.

DISCUSSION AND CONCLUSIONS

The impact of surface soil moisture assimilation on the estimation of both model parameters and model states (soil moisture in different depths) are investigated in SWAT model using the ensemble Kalman filter (EnKF) method. The investigation is based on a series of synthetic experiments and a real world test using the combined ESA CCI SM product in the upper Huai River basin. The following two schemes are considered in EnKF: (1) both the model parameter and soil moisture in different depths are updated simultaneously by assimilating the surface soil moisture (EnKF_SP), (2) only the soil moisture in different depths are updated by assimilating the surface soil moisture (EnKF_S).

In the synthetic experiments, the results of EnKF_SP indicate that the surface soil moisture assimilation with joint state-parameter update has significant capability to improve the estimation of parameter SOL_AWC (the available soil water capacity) and soil moisture itself in different depths, even though this improvement is decreasing with the increase of soil depth. For EnKF_S, the surface soil moisture assimilation with only state update with unbiased parameters (the parameter bias is removed by joint parameter-state update for the first 200 days) also obtains satisfactory SM estimation in different depths on soil profile. The less significant influence of EnKF_S on the SM estimation of the fourth layer might be related to the lack of consideration for the possible temporal evolution of the parameters in EnKF_S as the parameters remain unchanged after 200 days in the EnKF_S based data assimilation process. In terms of the impact of surface soil moisture assimilation on the rainfall-runoff process, the performance of EnKF_SP is relatively better and robust than that of the EnKF_S, especially for the ground water. This also indicates the importance of both state and parameter update on the improvement of hydrological modeling in soil moisture assimilation.

In the real world tests, the assimilation of the combined ESA CCI SM by EnKF_SP shows that it has considerable capability in the estimation of parameter SOL_AWC. The uncertainty of SOL_AWC (represented by the standard deviation of the parameter ensemble) is largely decreased and the ensemble trajectory remains stable after assimilating the satellite soil moisture for about 200 days. Moreover, the predicted steamflow using the parameters from EnKF_SP is slightly better than that from the prior calibration, which also reflects the capability of the satellite soil moisture in parameter optimization. The sensitivity analysis of the EnKF_SP and EnKF_S performance on streamflow modeling to the model and observation errors shows that both of them have the capability to improve the streamflow modeling by assimilating the combined ESA CCI SM although the optimal model and observation error parameters for them are quite different. However, the improvement of streamflow modeling by EnKF_SP and EnKF_S in the combined ESA CCI SM assimilation is not significant in this study. This result can be analyzed from the following three aspects: (1) The underlying conditions of the study catchment might influence the usefulness of the combined ESA CCI SM product. The dominant land cover of this basin is rice (56%), which is usually wet and with high density of vegetation cover. This will deteriorate the data quality of the satellite soil moisture retrievals. Besides, the mountainous region takes about 30% in this basin, of which the undulating terrain might also produce bad influence on the utility of the remote sensing data. Also, Brocca et al. (2013) did not find significant improvement in the mountainous region by assimilating different soil moisture products into rainfall-runoff modeling using a simple nudging scheme. (2) The spatial mismatch between the satellite soil moisture and the model simulated soil moisture might influence the assimilation performance. The grid-based ESA CCI SM has a spatial resolution of about 25 km, which is much larger than that of the model simulated soil moisture. The coarse spatial resolution of the soil moisture observation might weaken its sensitivity to the changes of hydrological modeling, thus influence its assimilation performance. (3) The simplification on the model and observation error assumption might affect the optimality of the assimilation scheme. The model error is assumed a Gaussian additive or multiplicative distribution, and the spatial and temporal correlations of the error (e.g., the model input error for precipi-

tations) are not considered. Besides, the non-stationary error characteristics for the satellite soil moisture are also ignored in this study. Therefore, based on the knowledge that the surface soil moisture assimilation has the capability to improve the state and parameter estimates, a lot of future work needs to be done to improve the data quality of the satellite soil moisture products and their utilization by optimizing the data assimilation scheme to improve the rainfall-runoff modeling.

REFERENCES

Abbaspour, K., Johnson, C., Van Genuchten, M.T., 2004. Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. Vadose Zone Journal, 3, 4, 1340–1352.

Aksoy, A., Zhang, F., Nielsen-Gammon, J.W., 2006. Ensemble-based simultaneous state and parameter estimation in a two-dimensional sea-breeze model. Monthly Weather Review, 134, 10, 2951–2970.

Alvarez-Garreton, C., Ryu, D., Western, A., Crow, W.T., Robertson, D.E., 2014. The impacts of assimilating satellite soil moisture into a rainfall–runoff model in a semi-arid catchment. Journal of Hydrology, 519, 2763–2774.

Alvarez-Garreton, C., Ryu, D., Western, A., Su, C.H., Crow, W.T., Robertson, D.E., Leahy, C., 2015. Improving operational flood ensemble prediction by the assimilation of satellite soil moisture: comparison between lumped and semi-distributed schemes. Hydrology and Earth System Sciences Discussions, 11, 9, 10635–10681.

Aubert, D., Loumagne, C., Oudin, L., 2003. Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall–runoff model. Journal of Hydrology, 280, 1, 145–161.

Barre, H.M.J., Duesmann, B., Kerr, Y.H., 2008. SMOS: The Mission and the System. Geoscience and Remote Sensing, IEEE Transactions on 46(3): 587–593.

Brocca, L., Moramarco, T., Melone, F., Wagner, W., Hasenauer, S., Hahn, S., 2012. Assimilation of Surface-and Root-Zone ASCAT Soil Moisture Products Into Rainfall–Runoff Modeling.Geoscience and Remote Sensing, IEEE Transactions on 50(7): 2542–2555.

Brocca, L., Moramarco, T., Dorigo, W., Wagner, W., 2013. Assimilation of satellite soil moisture data into rainfall-runoff modelling for several catchments worldwide. Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International, IEEE.

Chen, F., Crow, W.T., Starks, P.J., Moriasi, D.N., 2011. Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture. Advances in Water Resources, 34, 4, 526–536.

Chen, W., Huang, C., Shen, H., Li, X., 2015. Comparison of ensemble-based state and parameter estimation methods for soil moisture data assimilation. Advances in Water Resources, 86, 425–438.

Clark, M.P., Rupp, D.E., Woods, R.A., Zheng, X., Ibbitt, R.P., Slater, A.G., Schmidt, J., Uddstrom, M.J., 2008. Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. Advances in Water Resources, 31, 10, 1309–1324.

Crosson, W.L., Laymon, C.A., Inguva, R., Schamschula, M.P., 2002. Assimilating remote sensing data in a surface flux–soil moisture model.Hydrological processes, 16, 8, 1645–1662.

Crow, W.T., Ryu, D., 2009. A new data assimilation approach for improving runoff prediction using remotely-sensed soil moisture retrievals.Hydrology and Earth System Sciences, 13, 1, 1–16.

Crow, W.T., Wood, E.F., 2003. The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble Kalman filtering: A case study based on ESTAR measurements during SGP97. Advances in Water Resources, 26, 2, 137–149.

Das, N.N., Entekhabi, D., Njoku, E.G., 2011. An Algorithm for Merging SMAP Radiometer and Radar Data for High-Resolution Soil-Moisture Retrieval.Geoscience and Remote Sensing, IEEE Transactions on 49, 5, 1504–1512.

Das, N.N., Entekhabi, D., Njoku, E.G., Shi, J.J. C., Johnson, J.T., Colliander, A., 2014. Tests of the SMAP Combined Radar and Radiometer Algorithm Using Airborne Field Campaign Observations and Simulated Data.Geoscience and Remote Sensing, IEEE Transactions on 52, 4, 2018–2028.

Entekhabi, D., Njoku, E.G., O'Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, N., McDonald, K.C., Moghaddam, M., Moran, S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Leung, T., Van Zyl, J. 2010. The Soil Moisture Active Passive (SMAP) Mission.Proceedings of the IEEE 98, 5, 704–716.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research, 99, 10143–10162.

Han, E., Merwade, V., Heathman, G.C., 2012. Implementation of surface soil moisture data assimilation with watershed scale distributed hydrological model. Journal of Hydrology, 416–417, 98–117.

Heathman, G.C., Starks, P.J., Ahuja, L.R., Jackson, T.J., 2003. Assimilation of surface soil moisture to estimate profile soil water content. Journal of Hydrology, 279, 1–4, 1–17.

Lü, H., Yu, Z., Zhu, Y., Drake, S., Hao, Z. and Sudicky, E.A., 2011. Dual state-parameter estimation of root zone soil moisture by optimal parameter estimation and extended Kalman filter data assimilation. Advances in Water Resources, 34, 3, 395–406.

Laiolo, P., Gabellani, S., Campo, L., Silvestro, F, Delogu, F, Rudari, R., Pulvirenti, L, Boni, G, Fascetti, F., Pierdicca, N., 2015. Impact of different satellite soil moisture products on the predictions of a continuous distributed hydrological model. International Journal of Applied Earth Observation and Geoinformation.

Lee, H., Seo, D.-J., Koren, V., 2011. Assimilation of streamflow and in situ soil moisture data into operational distributed hydrologic models: Effects of uncertainties in the data and initial model soil moisture states. Advances in Water Resources, 34, 12, 1597–1615.

Lievens, H., Tomer, S.K., Al Bitar, A., De Lannoy, G.J.M., Drusch, M., Dumedah, G., Hendricks Franssen, H.J. Hendricks, Kerr, Y.H., Martens, B., Pan, M., Roundy J.K., Vereecken, H., Walker, J.P., Wood E.F., Verhoest, N.E.C., Pauwels, V.R.N., 2015. SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia. Remote sensing of environment, 168, 146–162.

Lunt, I., Hubbard, S., Rubin, Y., 2005. Soil moisture content estimation using ground-penetrating radar reflection data. Journal of Hydrology, 307, 1, 254–269.

Massari, C., Brocca, L., Tarpanelli, A., Moramarco, T., 2015. Data Assimilation of Satellite Soil Moisture into Rainfall-Runoff Modelling: A Complex Recipe? Remote Sensing, 7, 9, 11403–11433.

McKay, M.D., Beckman, R.J., Conover, W.J., 1979. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics, 21, 2, 239–245.

Monteith, J.L., 1965. Evaporation and the environment. In: 19th Symposia of the Society for Experimental Biology: The state and movement of water in living organisms. Cambridge Univ. Press, London, pp. 205–234.

Moradkhani, H., Sorooshian, S., Gupta, H.V., Houser, P.R., 2005. Dual state-parameter estimation of hydrological models using ensemble Kalman filter. Advances in Water Resources, 28, 2, 135–147.

Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics, 33, 2, 161–174.

Nash, J., Sutcliffe, J., 1970. River flow forecasting through conceptual models part I—A discussion of principles. Journal of Hydrology, 10, 3, 282–290.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil and Water Assessment Tool Theoretical Documentation Version 2009. TR-406, Texas Water Resources Institute Technical Report No.406. Texax A&M University. (available at http://swat.tamu.edu/media/99192/swat2009-theory.pdf)

Njoku, E.G., Jackson, T.J., Lakshmi, V., Chan, T.K., Nghiem, S.V., 2003. Soil moisture retrieval from AMSR-E. Geoscience and Remote Sensing, IEEE Transactions on 41, 2, 215–229.

Reichle, R.H., Koster, R.D., 2004. Bias reduction in short records of satellite soil moisture. Geophysical Research Letters, 31, 19. DOI:10.1029/2004GL020938.

Reichle, R.H., Crow, W.T., Keppenne, C.L., 2008. An adaptive ensemble Kalman filter for soil moisture data assimilation. Water resources research, 44, 3.

Smith, P.J., 2010. Joint state and parameter estimation using data assimilation with application to morphodynamic modelling. University of Reading, Reading.

Troch, P.A., Paniconi, C., McLaughlin, D., 2003. Catchment-scale hydrological modeling and data assimilation. Advances in Water Resources, 26, 2, 131–135.

Walker, J.P., Willgoose, G.R., Kalma, J.D., 2001. One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: A comparison of retrieval algorithms. Advances in Water Resources, 24, 6, 631–650.

Wang, D., Chen, Y., Cai, X., 2009. State and parameter estimation of hydrologic models using the constrained ensemble Kalman filter. Water Resources Research, 45, 11.

Williams, J., 1969. Flood routing with variable travel time or variable storage coefficients.Trans. ASAE, 12, 1, 100–103.

Xie, X., Zhang, D., 2010. Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter. Advances in Water Resources, 33, 6, 678–690.

Xie, X., Zhang, D., 2013. A partitioned update scheme for state-parameter estimation of distributed hydrologic models based on the ensemble Kalman filter. Water Resources Research, 49, 11, 7350–7365.

Xie, X., Meng, S., Liang, S., Yao, Y., 2014. Improving streamflow predictions at ungauged locations with real-time updating: application of an EnKF-based state-parameter estimation strategy. Hydrology and Earth System Sciences, 18, 10, 3923–3936.

Yang, X., Delsole, T., 2009. Using the ensemble Kalman filter to estimate multiplicative model parameters. Tellus A 61, 5, 601–609.

Yu, Z., Liu, D., Lü, H., Fu, X., Xiang, L., Zhu, Y., 2012. A multi-layer soil moisture data assimilation using support vector machines and ensemble particle filter. Journal of Hydrology, 475, 53–64.

Yu, Z., Fu, X., Luo, L., Lü, H., Ju, Q., Liu, D., Kalin, A.D., Huang, D., Yang, C., Zhao, L., 2014. One-dimensional soil temperature simulation with Common Land Model by assimilating in situ observations and MODIS LST with the ensemble particle filter. Water Resources Research, 50, 8, 6950–6965.

NOMENCLATURE

SWAT	Soil and Water Assessment Tool
EnKF	the ensemble Kalman filter method
SM	soil moisture
ESA CCI	the soil moisture products in the Climate Change
SM	Initiative projects by the European Space Agency
HRU	Hydrological Response Units
STD	standard deviation
EnOL	open loop
EnKF_S	EnKF with only state update
EnKF_SP	EnKF with joint state-parameter update
SIM	simulation
OBS	observation
ESCO	soil evaporation compensation factor
EPCO	plant evaporation compensation factor
K_{sat}	saturated hydraulic conductivity
SOL_AWC	available soil water capacity
Φ_{wp}	fraction of water content at wilting point
Φ_{fc}	fraction of water content at field capacity
FC	field water capacity
Φ_{por}	total soil porosity
RMSE	root mean square error
NER	normalized reduction of the root mean square error
RE	relative error
NSE	Nash and Sutcliffe coefficient
σ_p	standard deviation of the Gaussian multiplicative perturbation on precipitation
σ_{par}	standard deviation of the Gaussian additive perturbation on parameters
SF	scaling factor to the upper bound of the soil moisture observation error
V_{out}	streamflow at the basin outlet
Q_{surf}	surface runoff
Q_{lat}	lateral flow
Q_{gw}	ground water
w_{rchrg}	water recharge to the shallow and deep aquifer from the bottom of soil profile