Agricultural and Forest Meteorology 230–231 (2016) 142–156

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Assimilating multi-source data into land surface model to simultaneously improve estimations of soil moisture, soil temperature, and surface turbulent fluxes in irrigated fields

Chunlin Huang[a,e,*], Weijing Chen[b], Yan Li[a,c], Huanfeng Shen[b], Xin Li[a,d]

[a] Key Laboratory of Remote Sensing of Gansu Province, Heihe Remote Sensing Experimental Research Station, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, Gansu, China
[b] School of Resource and Environmental Science, Wuhan University, Wuhan, Hubei, China
[c] University of Chinese Academy of Sciences, Beijing, China
[d] CAS Center for Excellence in Tibetan Plateau Earth Sciences, Chinese Academy of Sciences, Beijing, China
[e] Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing, China

ARTICLE INFO

Article history:
Received 19 October 2015
Received in revised form 4 March 2016
Accepted 22 March 2016
Available online 1 April 2016

Keywords:
Data assimilation
Ensemble Kalman Smoother
Soil moisture
Soil temperature
Surface turbulent fluxes
Common Land Model

ABSTRACT

The optimal estimation of soil moisture, soil temperature, and surface turbulent fluxes in irrigation fields is restricted by a lack of accurate irrigation information. To resolve the input uncertainty from imprecise irrigation quantity, an improved data assimilation scheme that is EnKS (Ensemble Kalman Smoother) implemented with inflation and localization (referred to as ESIL) is proposed to estimate soil moisture, soil temperature, and surface turbulent fluxes for irrigated fields by assimilating multi-source observations. The Daman station, which is located at an irrigated maize farmland in the middle reaches of the Heihe River Basin (HRB), is selected in this study to investigate the performance of the proposed assimilation scheme. The measured land surface temperature (LST) and surface soil moisture (SSM) in the first soil layer are taken as observations to conduct a series of data assimilation experiments to analyze the influence of a lack of irrigation information and combinations of multi-source observations on estimations of soil moisture, soil temperature, and surface turbulent fluxes. This study demonstrates the feasibility of ESIL in improving the estimation of hydrothermal conditions under unknown irrigation. The coefficient correlation (***R***) with the ESIL method increases from 0.342 and 0.703 to 0.877 and 0.830 for the soil moisture and soil temperature in the first layer, respectively. Meanwhile, the surface turbulent fluxes are significantly improved and the ***RMSE*** decreases from 173 W/m^2 and 186 W/m^2 to 97 W/m^2 and 111 W/m^2 for the sensible and latent heat fluxes, respectively.

1. Introduction

Accurate estimations of the soil moisture, soil temperature, and surface turbulent fluxes are crucial to agriculture and water management in irrigated fields. Land surface variables and fluxes can be acquired by modeling or observations. Modeling provides the evolution of continuous states in the time-space domain *via* parameterizing the inherent physical processes in the geo-sphere. However, the uncertainties that exist in forcing data, model parameters and model structure adversely affect the model output. Observations provide relatively accurate information, but the space gaps in *in-situ* measurements or the time gaps in remote sensing data cannot fulfill the requirements of practical application. Data assimilation takes full advantage of imperfect models and limited observations by merging the information embodied in observations into a dynamic model to correct the forecast trajectory. Numerous studies have assimilated both the surface soil moisture from ground-based networks or from microwave sensors and passive microwave brightness temperature data to improve soil moisture estimation and assimilate land surface temperatures from ground-based networks or satellite sensors to obtain more precise soil temperature profiles (Kumar and Kaleita, 2003; Gao et al., 2007; Huang et al., 2008a, b; Jia et al., 2009; Chen et al., 2011; Chen et al., 2015; Chu et al., 2015).

In the framework of a dynamic model, the soil moisture and soil temperature mutually influence to constitute the water and energy balance in solums. The soil temperature is a function of the soil moisture. Subsurface moisture influences the heat conductivity

* Corresponding author.
E-mail address: huangcl@lzb.ac.cn (C. Huang).

http://dx.doi.org/10.1016/j.agrformet.2016.03.013
0168-1923/

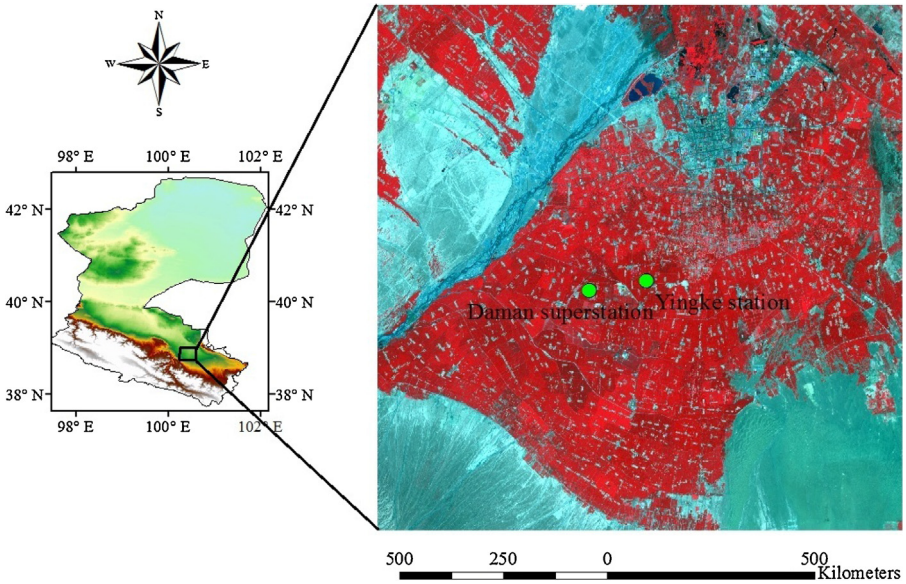

Fig. 1. Location of the study area in an SRTM 1-km DEM map and the locations of the field observation stations in an ASTER false color composite map with bands 3, 2, and 1.

at the interfaces of layers and the heat storage in different layers. In addition, the soil temperature determines the phase of the soil water content, including the transformation between frozen and unfrozen. Furthermore, the surface temperature affects the partitioning of incoming radiant energy into the ground (sensible heat flux and latent heat flux) and thus changes the delivery of the soil moisture, soil temperature, and surface turbulent fluxes. Yang et al. (2007) validated the feasibility of correcting the surface temperature and surface energy budget by assimilating soil moisture. Lakshmi (2000) demonstrated that surface temperature assimilation can reduce the effect of errors in precipitation and/or shortwave radiation data on simulated soil moisture. Given the internal positive interactions in the modification of soil moisture and soil temperature, simultaneously assimilating temperature and moisture observations is important to produce more exact soil moisture, soil temperature and surface turbulent fluxes.

A vital factor in retrieving accurate soil moisture and temperature profiles in arid and semi-arid agricultural areas is irrigation. Irrigation, which is a human intrusion into soil systems, alters the hydrothermal conditions of soil and changes the water transfer mechanisms between soil layers. Abundant irrigation rapidly increases the soil moisture and dramatic decreases the soil temperature, even in deeper layers. Moreover, irrigation applications in agricultural practices normally occur under dry conditions, in which the soil moisture is highly sensitive to irrigation applications. However, the irrigation schedules (when and how much to irrigate) that have been recorded in government departments are usually not sufficiently inclusive to be prescribed as model inputs. Several scientists have emphasized the impact of irrigation on hydrologic processes and relevant variables (Ozdogan and Salvucci, 2004; Ines et al., 2006; Moiwo and Tao, 2015; Lawston et al., 2015). Wang and Cai (2007) attempted to investigate the information of irrigation schedules for specific crops by using predefined empirical criteria to determine irrigation actions and an optimization algorithm to invert the irrigation quantification.

The objective of this article is to investigate how to assimilate the SSM in the first layer and/or the LST to improve the profiles of soil moisture and soil temperature and surface turbulent fluxes for different situations of known and unknown irrigation. We establish a data assimilation framework that is based on the CoLM (Common Land Model) and consider the SSM in the first layer and/or the LST that was measured at Daman station as observations. First, we distinguish the influence of assimilating different observations on the state variables and fluxes to ascertain the capability of combinations of multi-source information (SSM and LST) in improving the forecast accuracy of model outputs. Second, the aforementioned experiments are repeated under two postulations—known irrigation and unknown irrigation. The former is implemented through substituting the soil moisture and temperature profiles with the *in-situ* measurements at the irrigation moments, which are defined by the mutational volume of the soil water in the deeper layer, while the latter is implemented without any modifications to soil moisture and temperature profiles. We compare the discrepancies of the experimental results from these circumstances to determine the competence of assimilating observation information in improving soil moisture, soil temperature, and surface turbulent fluxes under unknown irrigation circumstances. Finally, considering the application of satellite data in the future work, we retest the combinations of multi-source information (SSM and LST) with various observation intervals or standard deviations. In addition, an ESIL that was modified to retrieve parameters is compared to the original ESIL to investigate the necessity of parameter estimation.

This article is structurally organized as follows. The land data assimilation scheme is briefly described in Section 2, including the hydrothermal process of the CoLM, experimental design and assimilation algorithms. The study area and data are also introduced in Section 2. Sections 3 and 4 present the results and discussion regarding the experiments. We finish this paper with some conclusions in Section 5.

2. Land data assimilation scheme

2.1. Model operator

The Common Land Model (CoLM) (Dai et al., 2003) is the improved version of the Community Land Model (version 2.0) with one vegetation layer, 10 unevenly spaced vertical soil layers, and up to 5 snow layers (depending on the total snow depth). Every surface grid cell can be subdivided into any number of tiles, and each tile contains a single land cover type. We employ the CoLM as a dynamic model (model operator) to maintain prognostic state variables, such as the soil moisture and soil temperature.

The vertical soil moisture transport is governed by infiltration, runoff, gradient diffusion, gravity, and soil water extraction through

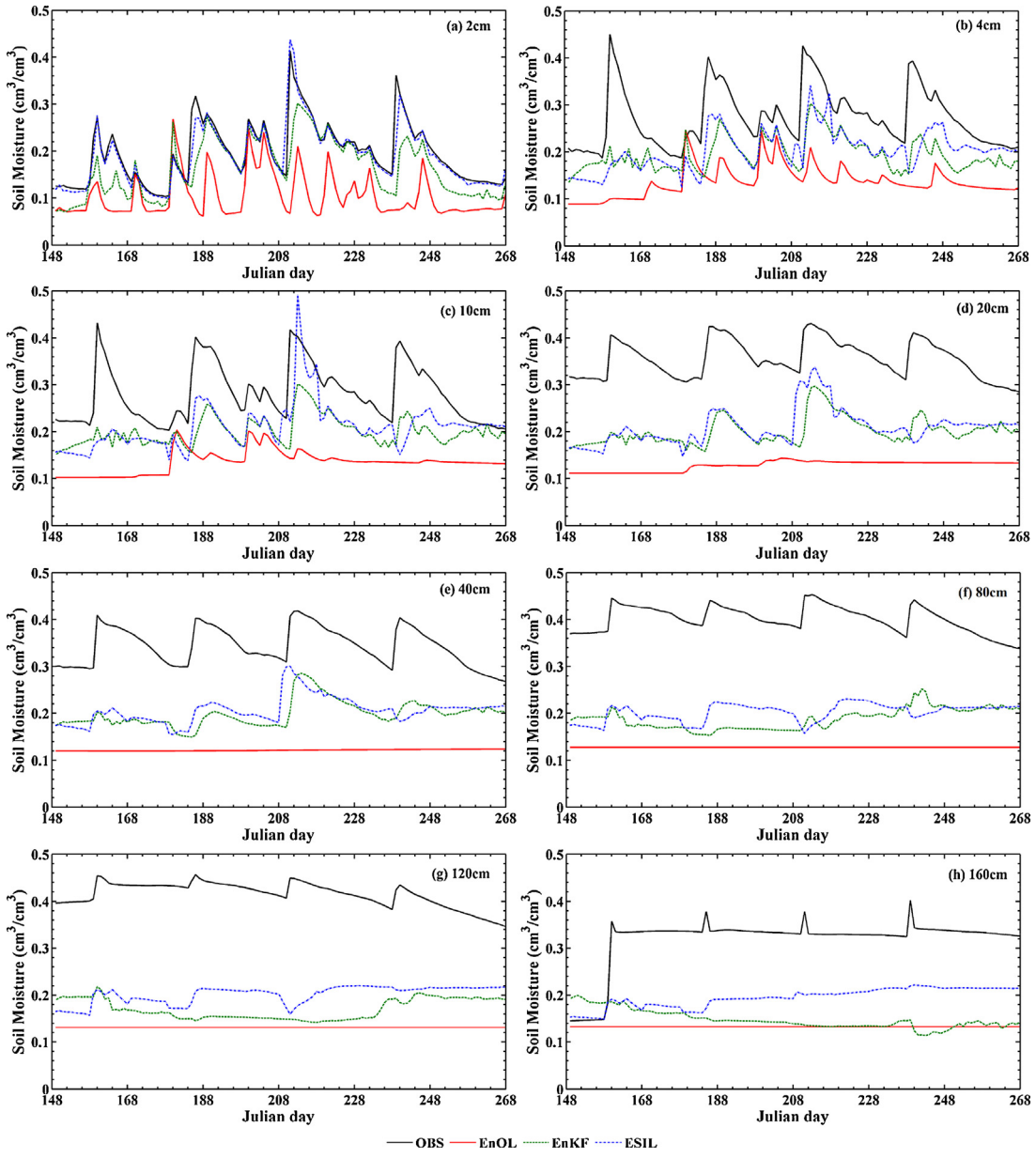

Fig. 2. Daily-average soil moisture from observations, simulations, and multi-source observations assimilation under unknown irrigation.

roots for canopy transpiration. The equations for liquid soil water can be written as

$$\frac{\Delta z_j}{\Delta t}\Delta\theta_j = [q_{j-1} - q_j] - f_{\text{root, } j^*} \times E_{tr}. \quad (1)$$

$\Delta\theta_j$ is the water change since the last time step in layer j. Δz_j and Δt are the soil thickness of layer j and the time step, respectively. $f_{root,\ j^*}$ and E_{tr} are the effective root fraction and transpiration, respectively. q_j is the water flow at the depth of the interface between layer j and layer $j+1$, which is calculated by Darcy's law:

$$q = -K\left(\frac{\partial\psi}{\partial z} - 1\right) \quad (2)$$

K and ψ are the hydraulic conductivity and matric potential of the soil, respectively, which vary with the soil water content and soil texture based on the scheme of Clapp and Hornberger (1978). The net water flux in the surface layer is provided by snowmelt, precipitation, and canopy dew, minus the surface runoff and evaporation.

The temperature variation in a soil layer can be described as a discretization form by the Crank–Nicholson scheme:

$$c_j \Delta z_j \frac{T_j^{k+1} - T_j^k}{\Delta t} = \frac{1}{2}\left(F_j^k - F_{j-1}^k + F_j^{k+1} - F_{j-1}^{k+1}\right), \quad (3)$$

where c_j is the volumetric soil heat capacity in the layer j. T_j^k is the layer-average temperature in the layer j at the time k; F_j^k is heat flux across the interface between layer j and $j+1$ at the time k, which is computed as follows:

$$F_j = \lambda\left(z_{h,j}\right)\frac{T_{j+1} - T_j}{z_{j+1} - z_j}, \quad (4)$$

where $\lambda\left(z_{h,j}\right)$ is the thermal conductivity at the interface $Z_{h,j}$, which is derived from the constraint that the heat flux across the interface is equal to that from the node j to the interface and the flux from the interface to the node $j+1$. The heat flux F at the surface is taken as $F = R_{n,g} - H_g - LE_g$. $R_{n,g}$ is the net radiation that is absorbed by the ground surface. H_q and LE_g are the sensible and

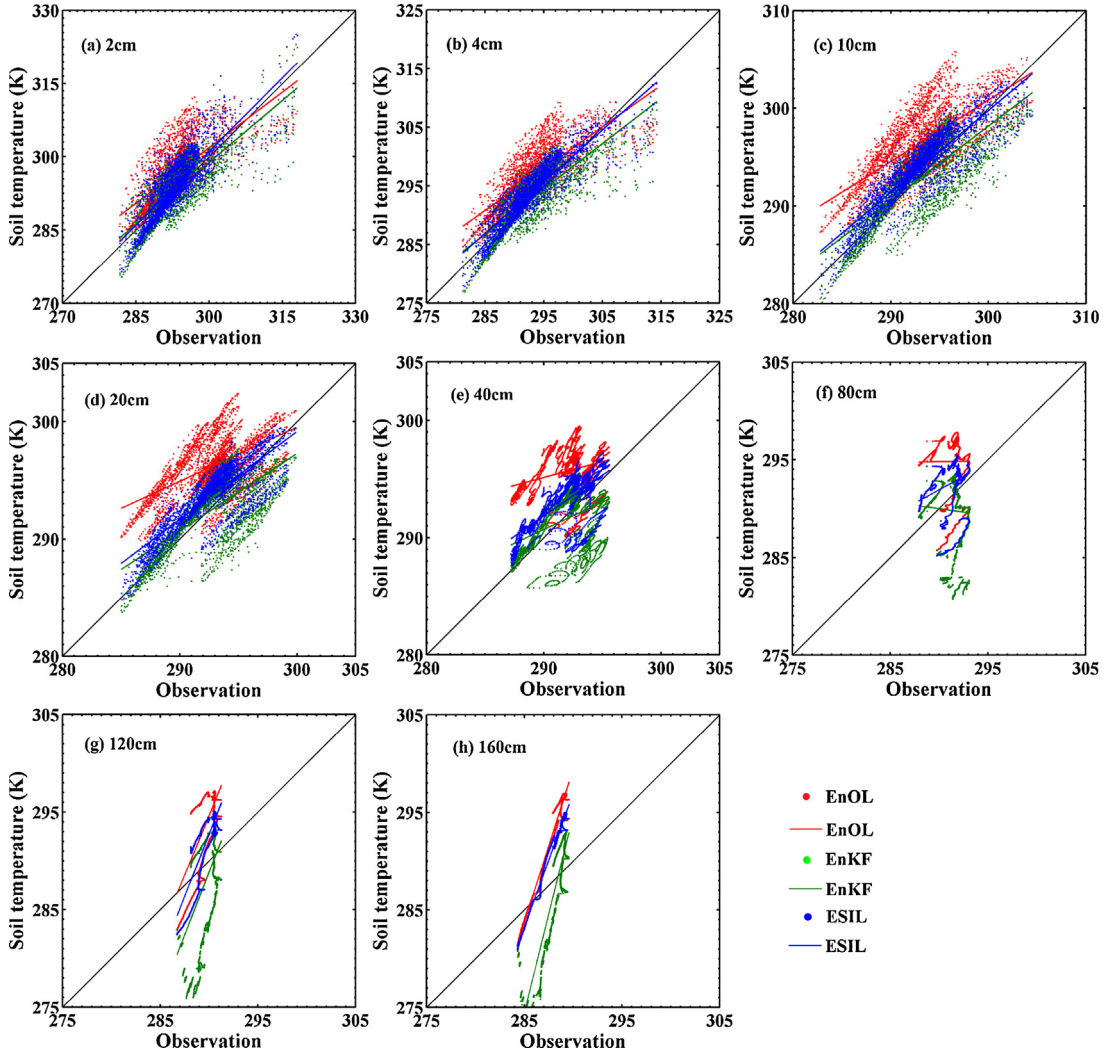

Fig. 3. Scatterplots of the soil temperature from simulations, observations, and multi-source observations assimilation under unknown irrigation.

latent heat fluxes. The heat flux is assumed to be zero at the bottom of the soil column.

2.2. Data assimilation algorithms

2.2.1. Ensemble Kalman Filter

The EnKF (Evensen, 1994) is divided into two steps: forecast and analysis. An overview of the EnKF procedure is presented below.

The initial state can be defined as X_0^a, and the ith member $X_{i,0}^a$ of the initial state ensemble is obtained by adding random noise to X_0^a:

$$X_{i,0}^a = X_0^a + u_i u_i \sim N(0, P_0) \tag{5}$$

where u_i is the background error vector, which conforms to a Gaussian distribution with a zero mean and the covariance matrix P_0.

Forecast: Each member of the state realizations is propagated according to

$$X_{i,k+1}^f = M(X_{i,k}^a, \alpha_{k+1}, \beta_{k+1}) + w_i w_i \sim N(0, Q). \tag{6}$$

Here, $M(\bullet)$ is a model operator and represents the CoLM in this case. The superscripts a and f refer to the analysis and forecast states, respectively. α and β are atmospheric forcing data and model parameters, which are used to run the CoLM. The model error is indicated by w_i with a zero mean and the covariance matrix Q and represents all the uncertainties that are related to the forcing data and model structure.

Analysis: When an observation is available, the observation vector is assimilated into the model. A linear correction equation is used according to a standard Kalman filter to update the forecasted state ensemble members:

$$X_{i,k+1}^a = X_{i,k+1}^f + K_{k+1}(Y_{i,K+1} - \hat{Y}_{i,k+1}). \tag{7}$$

Here, Y_{k+1} is generated by adding the stochastic perturbation into the actual observation Y_{k+1} at the time $k+1$ with a zero mean and the covariance matrix R according to the following:

$$Y_{i,k+1} = Y_{k+1} + v_i v_i \sim N(0, R). \tag{8}$$

$\hat{Y}_{i,k+1}$ is the projection of the model state in the observational space *via* the conversion of the observation operator $H(\bullet)$, which establishes a relationship between the model states and observations as $\hat{Y}_{i,k+1} = H(X_{i,k+1}^f)$. v_i represents independently and identically distributed Gaussian observation errors with the covariance matrix $\boldsymbol{R}$.

K_{k+1} is the Kalman gain matrix at the time $k+1$, which is calculated as

$$K_{k+1} = P_{k+1}^f H^T (H P_{k+1}^f H^T + R)^{-1}. \tag{9}$$

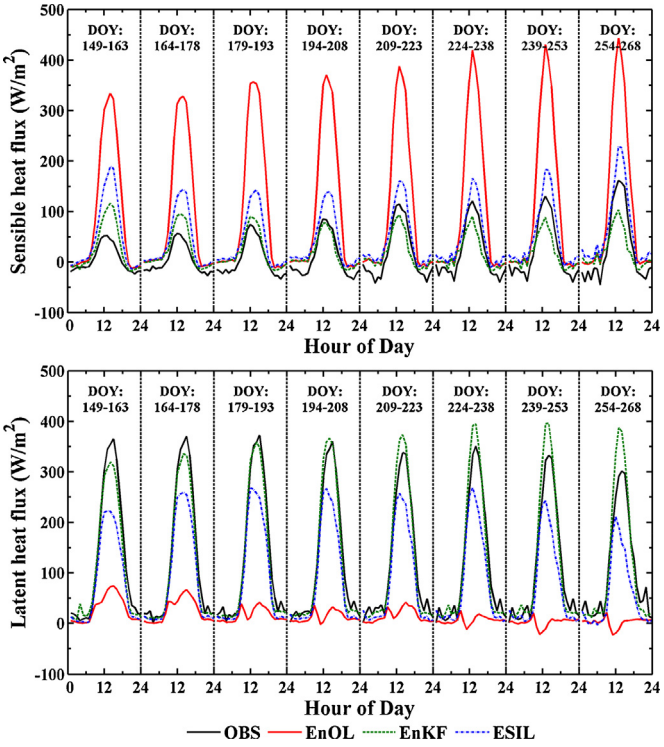

Fig. 4. Diurnal variation in the surface turbulent fluxes every 15 days from observations, simulation, and multi-source observations assimilation under unknown irrigation.

P^f_{k+1} is the forecast error covariance matrix at the time $k+1$. $P^f_{k+1}H^T$ is the cross covariance between the model state forecasts X^f_{k+1} and their projections $H(X^f_{k+1})$ in the observation space, while $HP^f_{k+1}H^T$ is the error covariance of $H(X^f_{k+1})$.

Finally, the analysis state variable that is estimated at the time $k+1$ is given by the averaged value of the ensemble members. The analyzed ensemble is then integrated forward until the next observation appears, and the process is repeated.

2.2.2. Ensemble Kalman Smoother

The EnKS (Evensen and Van Leeuwen, 2000), an extension of the EnKF, maintains the main procedure of the EnKF but examines the relationship of states or observations that are distributed in time or space *via* a predefined calculation window. The smoother solution of the EnKS can be found by including the impact of measurements backward in time (Evensen, 2003; Lei et al., 2014) which is a functioning means to solving the "sawtooth effect" of state curves from the EnKF.

The EnKS requires only forward model runs and no backward integrations of the model equations. The CoLM runs once throughout the current assimilation window to obtain the overall forecasting states in the smoother window, and the augmented state vector X contains states at all time steps as follows:

$$X = [x_1 x_2 \cdots x_k]^T. \tag{10}$$

Correspondingly, the augmented observation vector Y contains all the observations in the smoother window with the temporally and spatially independent hypothesis:

$$Y = [y_1 y_2 \cdots y_k]^T. \tag{11}$$

The Kalman gain matrix to update the state vector is calculated by using the same formula as for the above-mentioned EnKF. The implementation of the EnKS requires that the ensemble during the prior times must be stored and able to be updated whenever new observations become available (Dunne and Entekhabi, 2005).

2.2.3. Inflation

The utility of a data assimilation scheme heavily relies on the evolution of the background error statistics because the inverse of the specified error covariance determines the weights that are given to the observations. However, the background error covariance tends to be underestimated partly because of the inappropriate presence of model errors; thus, the filter gives too much preference to the background field. Moreover, this factor can compound the underestimation in the next cycle, resulting in filter divergence. Wu et al. (2013) developed an adaptive procedure that was combined with a second-order least squares method to estimate the inflated forecast and adjust observational error covariance matrices. In this article, only one step is performed to calculate the inflation factor without adjusting the observational error covariance, which is displayed as follows:

$$\gamma_{k+1} = \frac{Tr\left[HP^f_{k+1}H^T\left(d_{k+1}d^T_{k+1} - R\right)\right]}{Tr\left[HP^f_{k+1}H^THP^f_{k+1}H^T\right]} \tag{12}$$

$$d_{k+1} = Y_{k+1} - H\left(\overline{X^f_{k+1}}\right), \tag{13}$$

where the over-bar in the equation represents the average over many cases or statistical expectation, and $Tr[\cdot]$ denotes the trace of a matrix. Therefore, the forecast error covariance matrix is adjusted to $\gamma_{k+1} \cdot P^f_{k+1}$.

2.2.4. Localization

The multiplicative inflation factor effectively manages filter divergence but also induces over-inflation when only sparse and irregular observations are acquirable or the connection is implicit among states. We introduce the localization scheme to avoid this issue and form a proper relationship from the top soil to the bottom, which is applied to rectify the horizontal and vertical covariance in atmospheric data assimilation from Houtekamer and Mitchell (2001). The correlation function is constructed *via* a fifth-order piecewise rational function developed by Gaspari and Stephen (1999):

$$\rho(i,j) = \begin{cases} -\frac{\delta^5}{4} + \frac{\delta^4}{2} + \frac{5\delta^3}{8} - \frac{5\delta^2}{3} + 1, 0 \le \delta \le 1 \\ \frac{\delta^5}{12} - \frac{\delta^4}{2} + \frac{5\delta^3}{8} + \frac{5\delta^2}{3} - 5\delta + 4 - \frac{2}{3\delta}, 1 < \delta \le 2 \\ 0, \delta > 2 \end{cases} \tag{14}$$

$$\delta = D_{i,j}/F \tag{15}$$

$$F = \sqrt{3/10} \times L, \tag{16}$$

where L is an influence length scale and set to 3 in this study. $D_{i,j}$ is the Euclidean distance between two state points x_i and x_j.

Thus, the Kalman gain K is subsequently refreshed:

$$K = \left[(\rho \circ P)H^T\right] \left[H(\rho \circ P)H^T + R\right]^{-1} \tag{17}$$

where the notation $\circ$ denotes element-by-element matrix multiplication (also called the Schur product). The EnKS applied in this study includes states in different time points that are tightly connected with adjacent ones. However, this type of relationship gradually vanishes with an increasing time interval. Thus, localization is also employed to readjust the temporal scale's impact between states and filter out redundant or invalid correlations.

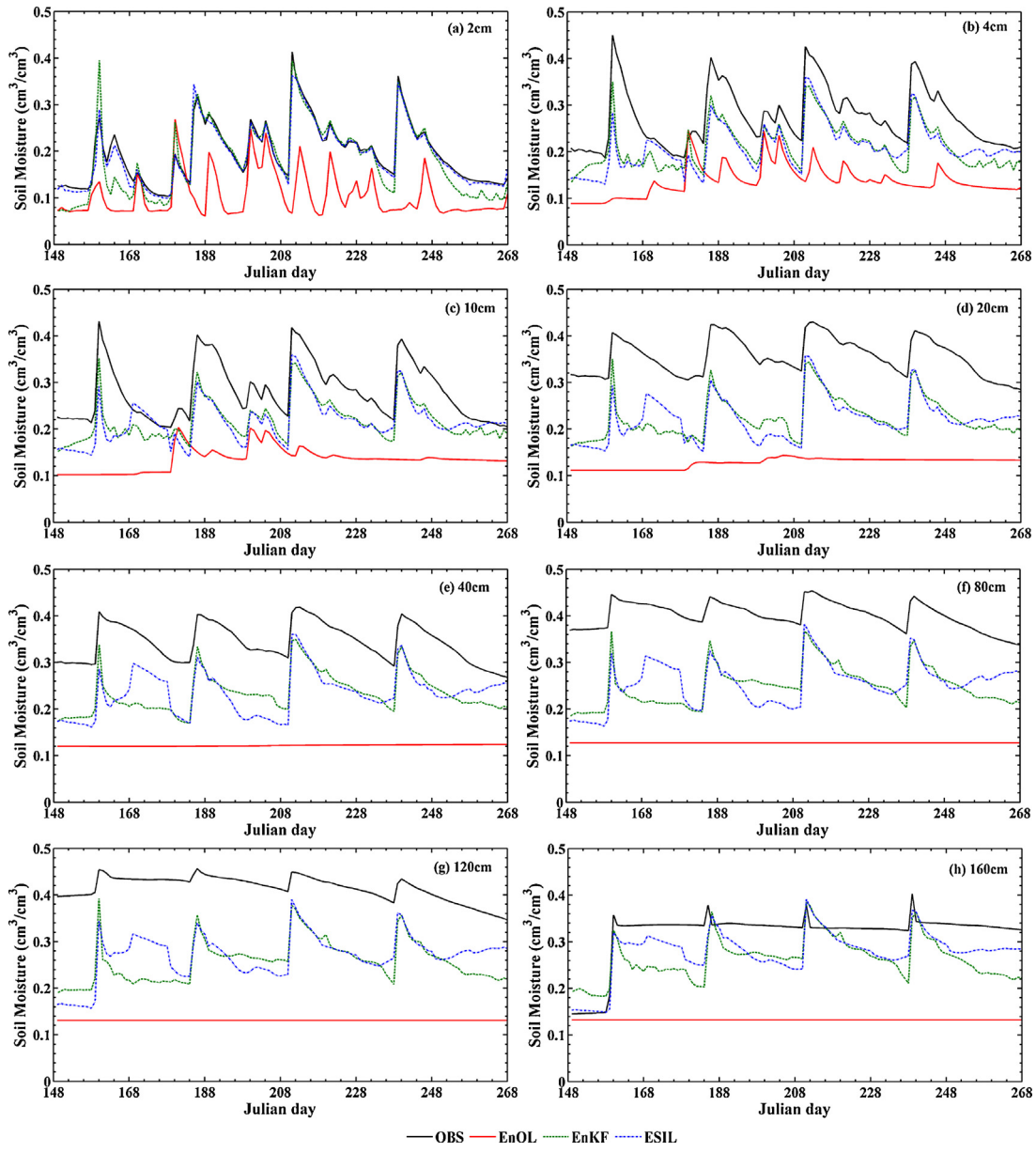

Fig. 5. Daily-average soil moisture from observations, simulations, and multi-source observations assimilation under known irrigation.

2.3. Study area and data

2.3.1. Description of the study area

The study area is located in an artificial oasis in the middle stream of the Heihe River Basin (100°14′E–100°32′E and 38°42′N–39°0′N), which is the second largest inland river basin in the arid region of northwestern China. This region has a typical continental climate, with an annual mean air temperature from 2.8 to 7.6 °C. The annual precipitation in this region is approximately 100–250 mm, but the potential evaporation is as high as 1200–1800 mm. The land cover in the study area is complex. Most of the study area consists of irrigated farmland, which primarily includes cultivated maize and wheat, and depends on irrigation water that is extracted from the Heihe River and groundwater. A highly developed irrigation system was constructed in this artificial oasis over the last few decades. Water bodies, sandy desert, desert steppes, and the Gobi Desert surround the oasis region. The average altitude of this region is approximately 1550 m, slightly higher in the southwest and lower in the northeast. In recent years, comprehensive eco-hydrological experiments such as the Watershed Allied Telemetry Experimental Research (WATER, Li et al., 2009) and the Heihe Watershed Allied Telemetry Experimental Research (HiWATER, Li et al., 2013a,b) have been successfully performed in the Heihe River Basin. The primary meteorological and hydrological parameters and the land surface variables in the study area were obtained through these projects.

2.3.2. Ground-based measurements

The *in-situ* measurements used in this study were collected at two observation stations in the study area: Yingke station from WATER and the Daman superstation from HiWATER. Furthermore, we processed the relevant data from ground stations as hourly inputs for the CoLM.

The WATER project was applied to the study area to research the ecological and hydrological processes of agricultural systems. Yingke station was installed in the central area of the oasis and covered with maize (100°24′37.2″E/38°51′25.7″N, Fig. 1). An automatic meteorological station (AMS) mounted on flux tower was

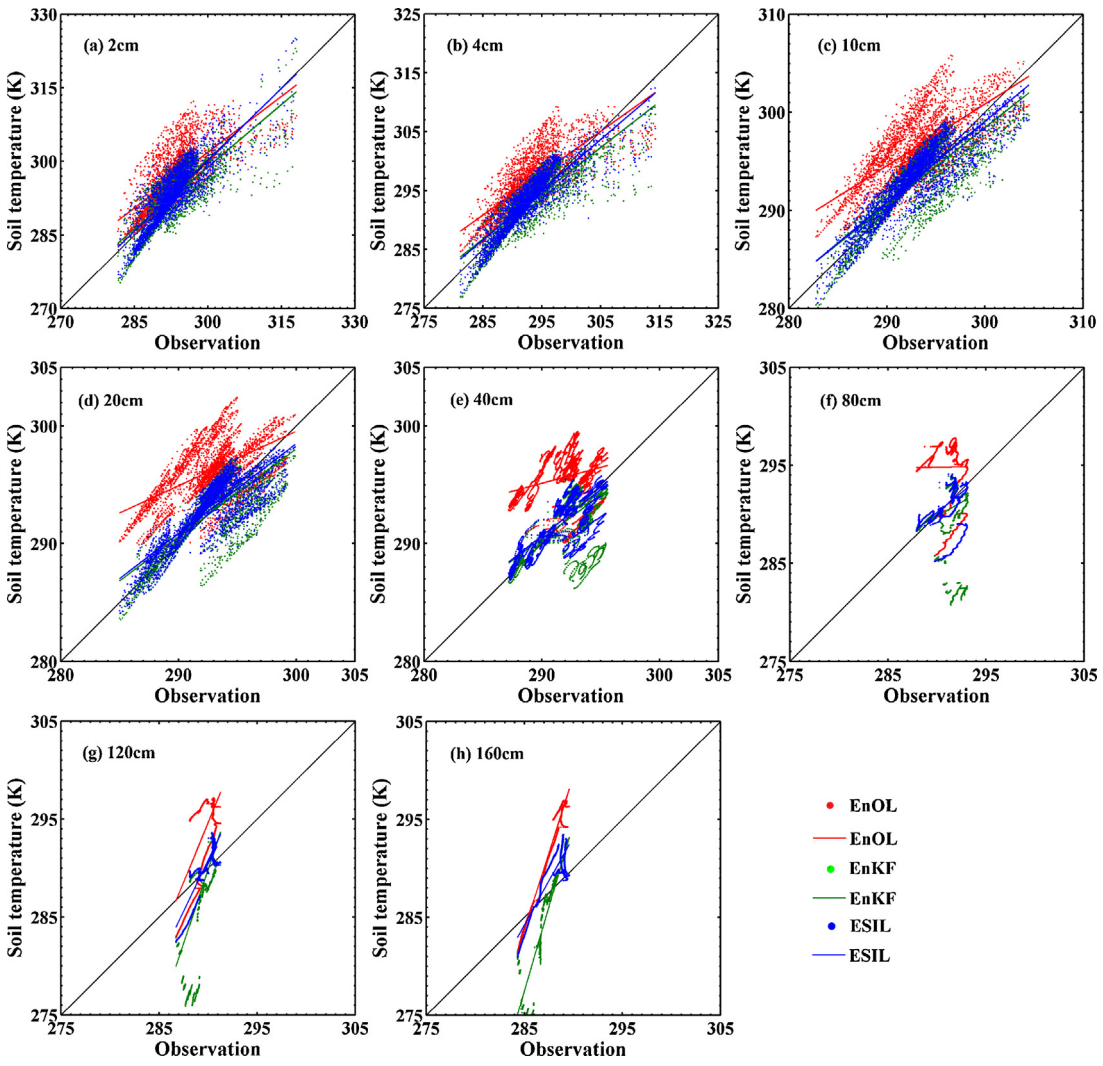

Fig. 6. Scatterplots of the soil temperature from observations, simulations, and multi-source observations assimilation under known irrigation.

applied to measure the wind speed/direction, air temperature and humidity at heights of 2 m and 10 m. In addition, the AMS included sensors for air pressure, precipitation and the four components of radiation, which were installed on the flux tower at heights of 0.5 m, 4 m and 4 m, respectively. The meteorological data were processed in 30-min intervals from November 2007 to December 2011 during the WATER experiment.

HiWATER is a comprehensive eco-hydrological experiment under the framework of the Heihe Plan, which is based on the diverse needs of interdisciplinary research and existing observational infrastructures in the Heihe River Basin. A multi-scale observational experiment on evapotranspiration over heterogeneous land surfaces (MUSOEXE), the first thematic experiment in the HiWATER project, was conducted in the study area from May to September 2012 (Li et al., 2013a,b; Liu et al., 2013). The Daman superstation, a 40-m boundary layer tower, was located in an irrigated farmland in the artificial oasis that was covered with maize (100°22′20.09″E/38°51′20.04″N, Fig. 1). Wind speed/direction, air temperature, and humidity sensors were installed at seven levels (3 m, 5 m, 10 m, 15 m, 20 m, 30 m, and 40 m) on the tower. The meteorological data included air pressure, precipitation and the four components of radiation were measured by the sensors on the tower at heights of 2 m, 2.5 m, and 12 m, respectively. The soil moisture profiles (0.02 m, 0.04 m, 0.1 m, 0.2 m, 0.4 m, 0.8 m, 1.2 m, and 1.6 m) and soil temperature profiles (0 m, 0.02 m, 0.04 m, 0.1 m, 0.2 m, 0.4 m, 0.8 m, 1.2 m, and 1.6 m) were measured by soil moisture sensors and soil temperature probes, which were buried south of the tower at a depth of 2 m. The surface temperature was recorded by an infrared thermometer mounted at a height of 12 m on the flux tower. The meteorological data and the surface flux data were processed in 10-min and 30-min intervals, respectively.

2.3.3. Satellite-based measurements

The Heihe Leaf Area Index (LAI) production, produced by Fan (2014), was applied in this study to replace the default value from the CoLM. The Heihe LAI production is produced based on a canopy Bidirectional Reflectance Distribution Function (BRDF) model that characterizes the surface reflectance as a function of a series of parameters (Liao et al., 2013) and is available at the Ecological and Environmental Science Data Center for Western China (http://westdc.westgis.ac.cn). The LAI images were processed at 1000-m spatial resolution and 8-day temporal resolution, with a Universal Transverse Mercator (UTM) Projection based on the WGS84 datum by Fan. We used ArcGIS to extract the LAI at the Daman superstation.

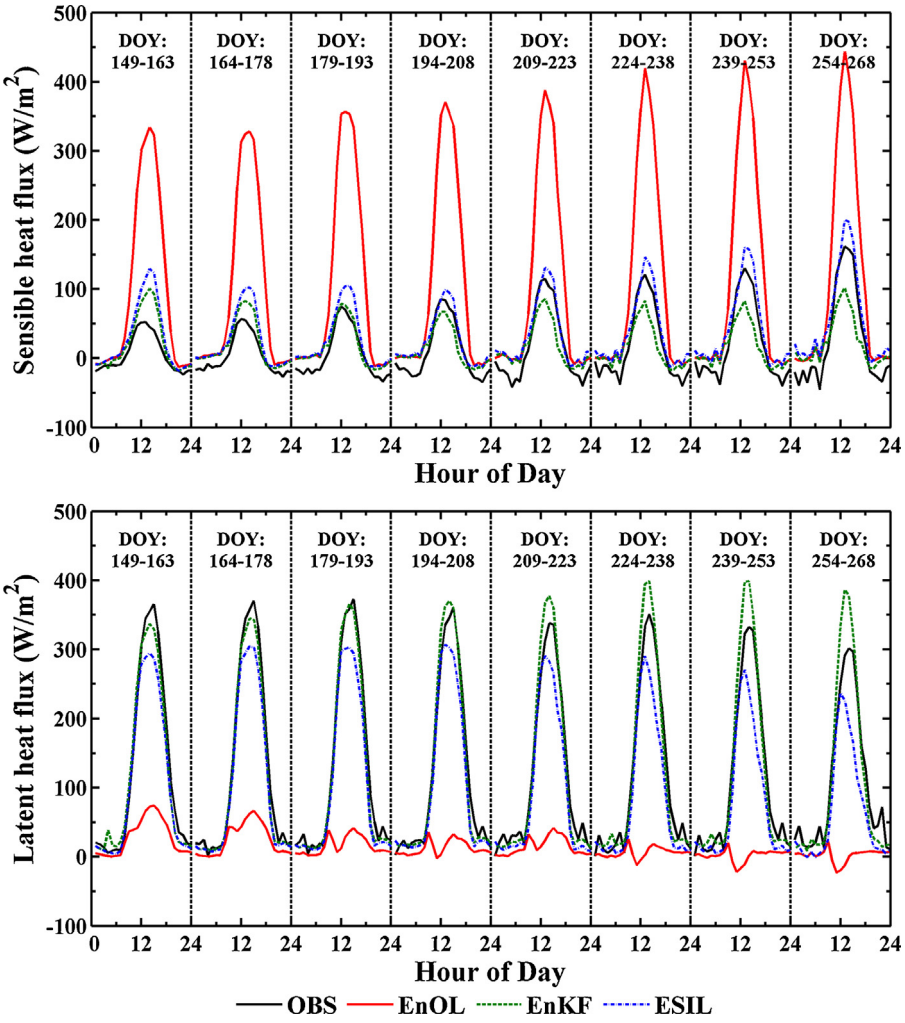

Fig. 7. Diurnal variation in the surface turbulent fluxes every 15 days from observations, simulations, and multi-source observations assimilation under known irrigation.

2.4. Assimilation experiments

2.4.1. Experimental design

One target of this study was to explore the influence of irrigation information on assimilation; thus, we focused on the crop growth period and conducted the assimilation experiment over 120 days commencing on May 27, 2012 (Julian day 148). The Daman superstation was chosen as the experimental site because of its more adequate measurement data among all the AWSs in the HiWATER project with an underlying crop surface. Considering the short temporal coverage of the HiWATER experiment, the most proximate one-year meteorological data (from May 27, 2010 to May 26, 2011) from Yingke station were used to spin-up the CoLM. The soil layers of the model were redistricted to conform to the observation depths at Daman station before the assimilation experiments (new soil nodes: 0.02, 0.04, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, and 1.6; unit: m).

In-situ SSM in the first layer and/or LST were introduced as observations to determine the functions of single and multi-source observations on some specific variables. The observation ensembles were acquired by disturbing the *in-situ* measured values with 3% multiplicative error for the SSM and 2 K additive error for the LST. The observation frequencies of the SSM and LST were set as once a day and twice a day respectively, according to the accessibility of the acquisition of different variables in reality. The three assimilation experiments (assimilation of the SSM in the first layer, the LST, and both types of observations) were implemented under two circumstances (unknown irrigation and known irrigation) to investigate the influence of irrigation information. Assuming known irrigation, we identified the irrigation moment based on the mutational volume of soil water in the deeper layer and substituted the soil moisture profile and soil temperature profile with the *in-situ* measurements at the irrigation moment.

Both the SSM in the first layer and the LST were assimilated into the CoLM to examine the influence of the observation interval and standard deviation. The assimilation frequencies were 0.5 days, 1 day, 3 days, 5 days and 8 days with a fixed 3% multiplicative standard deviation for the SSM and 0.25 days, 0.5 days, 1 day, 2 days and 3 days with a fixed 2 K additive standard deviation for the LST. In addition, the standard deviations were 1%, 3%, 5%, 7% and 9% multiplicative error with a fixed 1-day observation interval for the SSM and 1 K, 2 K, 3 K, 4 K and 5 K additive errors with a fixed 0.5-day observation interval for the LST.

The smoother window of the EnKS was set to be 10 days in all the assimilation experiments to assure that at least one observation exists in a smoother window when testing the influence of the observation interval. The observations were designed to update the corresponding states: the soil moisture in ten soil layers comprised a state vector only when the SSM in the first layer served as an observation, and the soil temperature in ten layers comprised a state vector only when the LST served as an observation. This approach guarantees the explicit relationship between states vari-

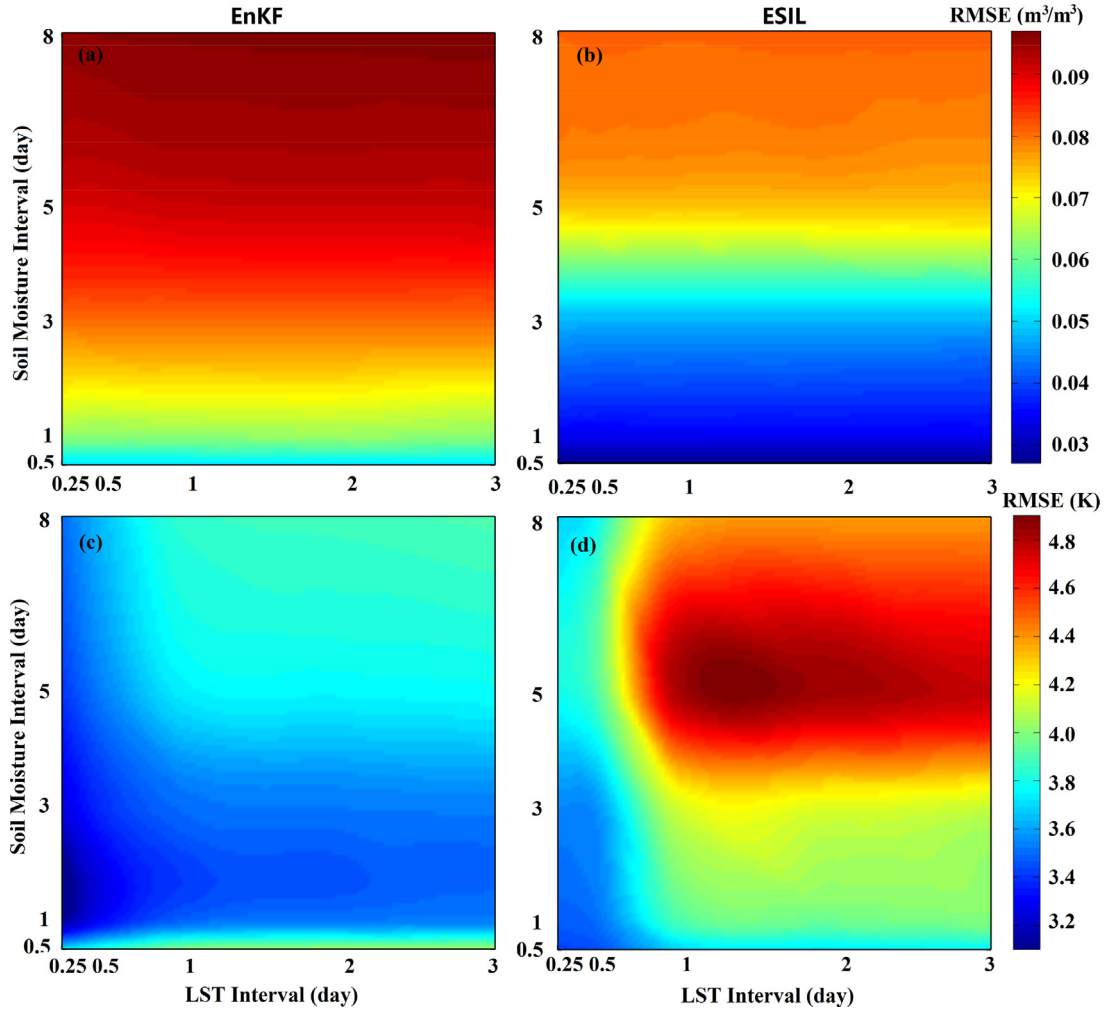

Fig. 8. RMSE distributions with varying observation intervals. (a) and (b) represent the RMSE of the soil moisture in the first soil layer from EnKF and ESIL, respectively; (c) and (d) represent the RMSE of the soil temperature in the first soil layer from EnKF and ESIL, respectively.

ables and observations and the updated states from assimilation feedback into the CoLM to impact other variables.

2.4.2. Ensemble generation

One-year forcing data from Yingke station were repeated in a 20-year period to spin-up the CoLM to obtain a stable and reasonable distribution of the initial state ensemble. Three important parameters (percentage of sand, percentage of clay and porosity) for soil moisture and soil temperature were uniformly resampled in a specific range according to their uncertainty, which was determined from values that fluctuated approximately 10% from the default values in the CoLM, to ensure a sufficiently dispersed state ensemble without excessively underestimating the background error.

The meteorological data imported into the CoLM is one vital factor that greatly influences the output variables. Normally distributed additive perturbations or log-normally distributed multiplicative perturbations were applied depending on the variable to account for errors from the forcing data. Thus, a positive perturbation of the downward shortwave radiation tends to be associated with negative perturbations to the longwave radiation and the precipitation, and vice versa. Table 1 displays the standard deviations and cross-correlations for the perturbations in precipitation, shortwave radiation, longwave radiation and air temperature. The mean values for the perturbed factors were equal to zero for the additive case and one for the multiplicative case.

Table 1
Summary of the perturbation parameters for the atmospheric forcing data and the cross correlation coefficients that were used to generate random perturbations for the different variables.

Variables	Noise type	Standard deviation	Cross correlation
Precipitation	Multiplicative	0.5	[1.0 −0.8 0.5 0.0,
Shortwave radiation	Multiplicative	0.3	−0.8 1.0 −0.5 0.4,
Longwave radiation	Additive	30 W/m^2	0.5 −0.5 1.00.4,
Air temperature	Additive	2 K	0.00.4 0.4 1.0]

2.4.3. Evaluation metrics

To assess the performance of the soil moisture and/or soil temperature assimilation, we defined several evaluation measures, including the root mean square error (RMSE), the mean bias error (MBE), the correlation coefficient (R), and the normalized error reduction (NER), which are described as follows:

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(X_t - X_{true,t})^2} \tag{18}$$

$$MBE = \frac{1}{T}\sum_{t=1}^{T}(X_t - X_{true,t}) \tag{19}$$

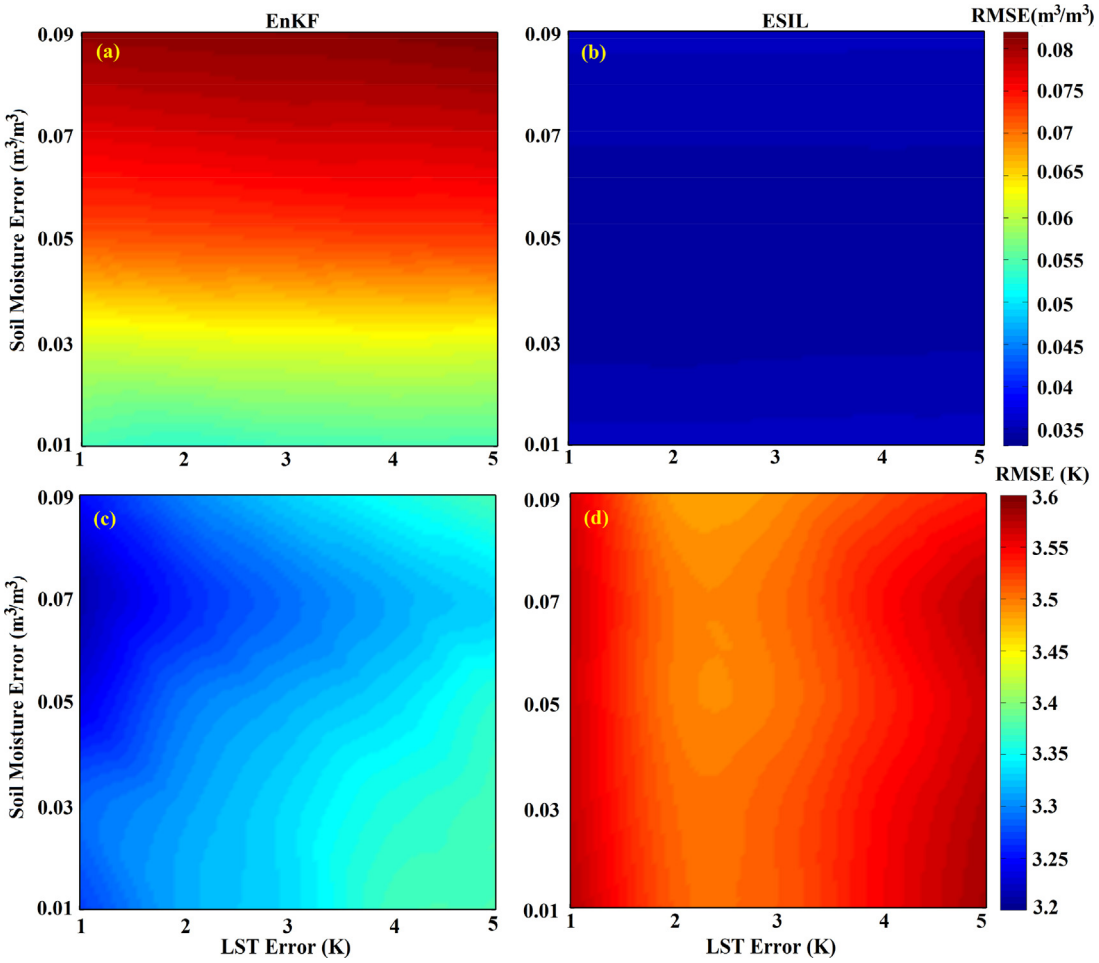

Fig. 9. RMSE distributions with assumed different observation errors. (a) and (b) represent the RMSE of the soil moisture in the first soil layer from EnKF and ESIL, respectively; (c) and (d) represent the RMSE of the soil temperature in the first soil layer from EnKF and ESIL, respectively.

$$R = \frac{\sum_{t=1}^{T}(X_t - \overline{X})(X_{true,t} - \overline{X_{true}})}{\sqrt{(X_t - \overline{X})^2 (X_{true,t} - \overline{X_{true}})^2}} \tag{20}$$

$$NER = 1 - \frac{RMSE_a}{RMSE_o}, \tag{21}$$

where T is the total number of steps and X_t and $X_{true,t}$ represent the simulation or assimilation values of states in different cases and the true values at step t, respectively. The horizontal line above the expression indicates the mean value. $RMSE_0$ and $RMSE_a$ represent the RMSE of the simulation and assimilation scenarios, respectively.

3. Results

The EnKS applied in this research was combined with inflation and localization to validly estimate the background error and expand the observational influence in the time domain and vertical space, which is referred to as 'ESIL' hereafter. Additionally, the ensemble mean value of the open loop simulations is referred to as 'EnOL'. The standard 'EnKF' is employed as a reference object to illustrate the effectiveness of 'ESIL' in different scenarios. All the simulation/assimilation results are compared to the *in-situ* hourly measurements.

Table 2
Statistical error metrics for the simulation/assimilation results of the soil moisture in the upper four layers with different observations under unknown irrigation.

		SM1			SM2	SM3	SM4
		MBE	RMSE	R	RMSE	RMSE	RMSE
	EnOL	−0.091	0.115	0.342	0.150	0.157	0.230
OBS_T	EnKF	−0.091	0.115	0.356	0.151	0.157	0.229
	ESIL	−0.092	0.116	0.348	0.151	0.157	0.229
OBS_M	EnKF	−0.034	0.06	0.686	0.098	0.097	0.158
	ESIL	−0.005	0.035	0.875	0.089	0.088	0.148
OBS_B	EnKF	−0.032	0.062	0.693	0.095	0.095	0.156
	ESIL	−0.005	0.034	0.877	0.090	0.090	0.149

3.1. Soil moisture

Table 2 presents the statistical error metrics of the soil moisture in the upper four layers (2 cm, 4 cm, 10 cm and 20 cm), which are labelled as SM1, SM2, SM3 and SM4 and were obtained from the EnOL, EnKF and ESIL under the condition of unknown irrigation. OBS_T, OBS_M and OBS_B indicate the three scenarios of only the LST assimilation, only the SSM assimilation and both types of observations assimilation, respectively. As shown in Table 2, both EnKF and ESIL fail to improve the soil moisture estimation in all four layers when only the LST is assimilated into the CoLM. In this scenario, the soil moisture is not the member of the state vector, and changes in the soil moisture only depend on the feedback into the CoLM for updating the soil temperature according to the experimental

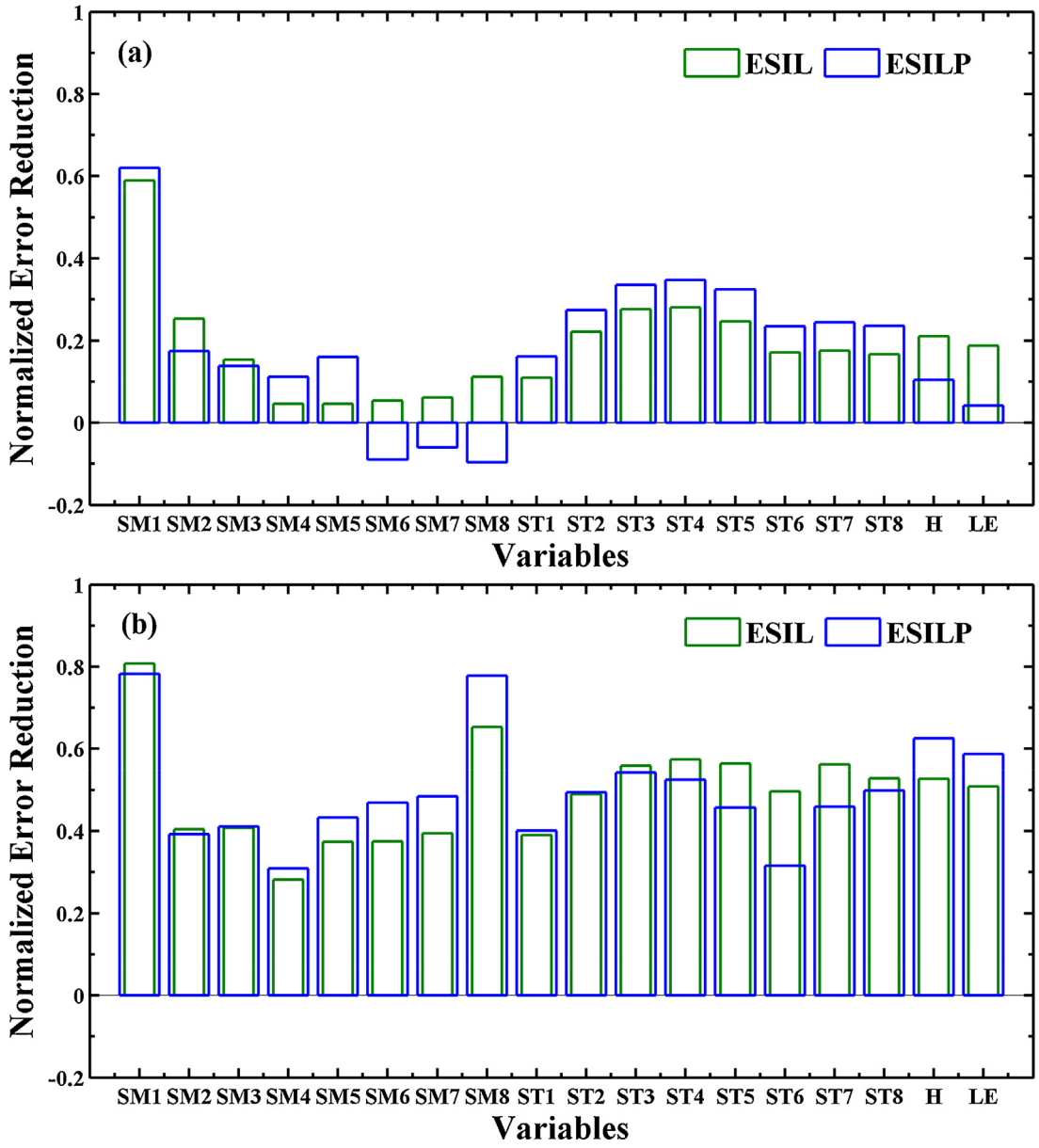

Fig. 10. Measurements of the improvements from EnOL via ESILP (implemented with parameter estimation) and ESIL from multi-source observations assimilation. ((a): unknown irrigation; (b): known irrigation).

design. However, the implied functional relationship between the soil temperature and soil moisture in the CoLM is unidirectional, which indicates that the soil moisture is almost immune to soil temperature modifications and results in an unapparent improvement in soil moisture. In the other two scenarios, the ***MBE*** and ***RMSE*** for the soil moisture in the first layer from the EnKF drop from −0.091 and 0.115 (ensemble simulation) to −0.034 and 0.063 (OBS_M) and to −0.032 and 0.062 (OBS_B). The ESIL further reduces the corresponding statistical metrics to approximately −0.005 and 0.034, which demonstrates vast superiority to the EnKF. The assimilation of both the SSM and LST improves the algorithm accuracy of EnKF to a certain extent, which also reflects the positive impact of the observation quantity and the positive interaction between the soil moisture and soil temperature. In terms of the ESIL, a minimal difference exists between OBS_M and OBS_B, which is mainly due to the excellent performance with OBS_M and the bare contribution from LST assimilation. Moreover, the ***R*** values for soil moisture in the first layer from the EnKF and ESIL double from 0.342 to approximately 0.69 and 0.87, respectively, in both OBS_M and OBS_B. The ***RMSE*** values for the soil moisture in deeper layers (SM2, SM3 and SM4) from the EnKF and ESIL also dramatically decrease by more than 40 percent from 0.150, 0.157 and 0.230, and the ESIL performs slightly better in all three layers. A comparison of the influence of all three assimilation scenarios on the soil moisture in deeper layers shows similar trends to those in the first layer.

The assimilation results of the soil moisture for all eight observations depths compares to the CoLM simulations and *in-situ* measurements (OBS) in the third scenario (OBS_B) are presented in Fig. 2. Fig. 2(a)–(f) indicate the curve graphs for different nodes (2, 4, 10, 20, 40, 80, 120 and 160 cm). The estimated soil moisture in the first layer from the EnKF and ESIL enhances the features capture with the black curve and maintains good agreement with the true values (Fig. 2(a)). In particular, the ESIL revises both the underestimated and overestimated portions and coincides well with the OBS.

Table 3
Statistical error metrics for the simulation/assimilation results of the soil temperature in the upper four layers with different observations under unknown irrigation.

		ST1			ST2	ST3	ST4
		MBE	RMSE	R	RMSE	RMSE	RMSE
	EnOL	3.522	5.322	0.703	4.741	4.388	4.287
OBS_T	EnKF	2.520	4.9573	0.738	3.885	3.313	3.002
	ESIL	2.026	4.652	0.751	3.580	3.097	3.011
OBS_M	EnKF	0.006	3.560	0.748	3.212	2.711	2.493
	ESIL	1.920	4.231	0.774	3.276	2.885	2.772
OBS_B	EnKF	−0.071	3.310	0.793	2.798	2.298	2.163
	ESIL	1.038	3.507	0.830	2.393	1.975	1.909

In the intermediate layers, the soil moisture curves from the EnKF and ESIL are affected by improvements in the surface soil moisture and appropriately reflect the trajectories of OBS, especially when there is rare variety existing in the EnOL (Fig. 2(b)–(e)). Judging from the hourly *in-situ* measurement data, the peaks in deeper layers (Fig. 2(f)–(h)) respond to irrigation actions rather than precipitation processes. Meanwhile, the steady tendency of the simulation curves (EnOL) reveal that the inherent water flow delivery mechanism of the CoLM in this area for deeper layers is unreasonable. The soil moisture in the deeper layers may be insusceptible to water quantity changes from natural processes in the upper layers. In addition, irrigation exacerbates the original underestimation phenomenon of the soil moisture. Although unable to grasp the peak values, the EnKF and ESIL still make a few improvements.

3.2. Soil temperature

Table 3 describes the statistical error metrics of the soil temperature corresponding to Table 2. ST1, ST2, ST3 and ST4 refer to quantitative analysis data for the soil temperature at 2 cm, 4 cm, 10 cm and 20 cm. The ESIL (EnKF) reduces the ***RMSE*** for the soil temperature in the first layer to 4.65 (4.65) from 5.32 (EnOL), and the deeper layers show more encouraging ***RMSE*** reduction (more than 20%) in the OBS_T scenario. The unnoticeable improvement can be ascribed to the discrepancy of land surface temperature and soil temperature in the first layer in arid and semi-arid regions, especially during early crop growth periods. The ***RMSE*** values display more obvious decreases once the SSM in the first layer is included as an observation. The considerable ***RMSE*** reduction in the ESIL (EnKF) ranges from over 20% (30%) for ST1 to over 35% (50%) for ST4 in both OBS_M and OBS_B. The ***MBE*** and ***R*** values confirm the ***RMSE*** evaluation, while the extraordinarily minor ***MBE*** values of the EnKF in OBS_M and OBS_B are an expression of positive and negative cancellation according to the corresponding ***RMSE*** value, which also indicates the correction of deviation. This achievement may be induced by revisions in the soil moisture profile when the SSM in the first layer is assimilated. Unsurprisingly, the soil moisture has a positive impact on the soil temperature because of the implied functional relationship between the soil moisture and soil temperature in the CoLM. The soil moisture as an input element for formulas decides the value of the heat capacity, thermal conductivity and Kersten number in each soil layer, which affects the calculation of soil temperature. Under the experimental design with an SSM observation for one day, the EnKF feedbacks a modified soil moisture profile once a day, while the ESIL feedbacks every ten days at the end of every smoother window. Thus, the ESIL does not consider the absolute predominance to the EnKF because of the amelioration of the soil temperature in OBS_M. According to the combination assimilation of two types of observations, both the EnKF and ESIL inevitably reach the best position compared to the other two scenarios. The ESIL is still slightly inferior to the EnKF, which also reveals that the correction of soil moisture profiles plays a somewhat more significant role than LST assimilation in the improvement of soil temperature estimation.

Fig. 3 depicts scatterplots that compare assimilated/simulated soil temperatures with observations (*in-situ* measurements) in the third scenario (OBS_B). The straight lines from the EnOL, EnKF and ESIL are derived from linear regression with observations. Similarly, Fig. 3(a)–(f) indicate the curve graphs for different nodes (2, 4, 10, 20, 40, 80, 120 and 160 cm). The point clouds in Fig. 3(a)–(d) show that the fluctuation in soil temperature in the top four layers and the correlation coefficient (***R***) with observations is 0.703, 0.774, 0.718 and 0.571 for the EnOL. The EnKF (ESIL) inordinately raises the corresponding ***R*** values to 0.793, 0.829, 0.807 and 0.708 (0.830, 0.899, 0.889 and 0.833) by relieving overestimated soil temperatures. The discrete points produced by the EnOL concentrate around and approach the black line (1:1 line) *via* the implementation of the EnKF and ESIL, especially the latter. The concentrated points in Fig. 3(e)–(h), which almost assemble into lines, indicate the stability of soil moisture in deeper layers during the experimental period. In Fig. 3(e)–(g), the ESIL points draw closer to the 1:1 line than the EnOL points, as well as the fitting lines. The ***R*** values dramatically increase from 0.271, 0.006, 0.670 and 0.981 to 0.661, 0.295, 0.811 and 0.991 for the ESIL. The EnKF only shows improvement in Fig. 3(e), with a growth of 0.13 for ***R***, and diverges from the observations to an underestimated position in Fig. 3(e)–(h).

3.3. Surface turbulent fluxes

In the land surface model, the soil moisture is positively correlated with the latent heat flux and negatively correlated with the sensible heat flux and surface temperature. The combination assimilation of SSM and LST effectively improves the estimation accuracy of both the sensible and latent heat fluxes, as shown in Fig. 4.

Fig. 4 displays the diurnal variation every 15 days of the surface turbulent fluxes derived by simulation and assimilation of both the SSM and LST under unknown irrigation. Compared to the black curves (*in-situ* measurements), the evident underestimated latent heat flux and overestimated sensible heat flux are responses to the underestimated soil moisture and overestimated soil temperature in the first layer. The ESIL reduces the ***RMSE*** values of the sensible and latent heat fluxes to 97 W/m^2 and 111 W/m^2 from 173 W/m^2 and 186 W/m^2, respectively. The EnKF produces more remarkable improvements in both the sensible and latent heat fluxes, as seen in Fig. 4, with the corresponding ***RMSE*** values decreasing to 65 W/m^2 and 77 W/m^2. The advantage of the EnKF in terms of correcting both the sensible and latent heat fluxes can be ascribed to the higher feedback frequency of the soil moisture and temperature profile compared to the ESIL, which resembles the previous argument. The feasibility of assimilating the land surface temperature and soil moisture to revise the surface turbulent fluxes is consistent with the conclusions drawn by Xu et al. (2011) and Hain et al. (2012).

4. Discussion

4.1. Influence of irrigation information

We duplicated the above experiments, including OBS_T, OBS_M and OBS_B, under the condition of known irrigation to explore the influence of irrigation information. The soil moisture and soil temperature profiles were substituted with the *in-situ* measurements at irrigation moments, which were determined by steep rises in the water content in deeper layers. Table 4 and Fig. 5 show the experimental results for the soil moisture. First, the intake of irrigation information tremendously improves the utility of LST assimilation. Unlike the invalid performance in Table 2, the EnKF and ESIL

Table 4
Statistical error metrics for the simulation/assimilation results of the soil moisture in the upper four layers with different observations under known irrigation.

		SM1			SM2	SM3	SM4
		MBE	RMSE	R	RMSE	RMSE	RMSE
	EnOL	−0.091	0.115	0.342	0.150	0.157	0.230
OBS_T	EnKF	0.006	0.051	0.872	0.060	0.056	0.113
	ESIL	0.007	0.051	0.873	0.059	0.055	0.113
OBS_M	EnKF	−0.014	0.037	0.921	0.072	0.069	0.134
	ESIL	−0.005	0.018	0.970	0.075	0.076	0.137
OBS_B	EnKF	−0.017	0.038	0.920	0.0711	0.069	0.133
	ESIL	−0.005	0.018	0.970	0.075	0.076	0.137

Table 5
Statistical error metrics for the simulation/assimilation results of the soil temperature in the upper four layers with different observations under known irrigation.

		ST1			ST2	ST3	ST4
		MBE	RMSE	R	RMSE	RMSE	RMSE
	EnOL	3.5220	5.322	0.703	4.741	4.388	4.287
OBS_T	EnKF	0.4748	3.173	0.821	2.474	1.995	1.754
	ESIL	0.5723	3.166	0.836	2.373	1.850	1.659
OBS_M	EnKF	−0.2344	3.433	0.763	3.060	2.462	2.102
	ESIL	0.5572	3.465	0.779	2.811	2.292	1.964
OBS_B	EnKF	−0.0544	3.266	0.797	2.749	2.176	1.929
	ESIL	0.5308	3.130	0.842	2.348	1.829	1.642

reduce the ***RMSE*** (***MBE***) to 0.051 (0.006) and 0.051 (0.007) in OBS_T, while the performance of the EnKF in OBS_T unexpectedly surpasses that in OBS_B with unknown irrigation. Moreover, the improvements in the ***RMSE*** from the EnKF and ESIL in deeper layers exceed 50%, reaching 65%. Second, the assimilation of soil moisture further boosted the effectiveness of the EnKF and ESIL. Remarkably, the discrepancies in the statistical error metrics are still negligible in OBS_M and OBS_B for the same assimilation algorithm. The EnKF (ESIL) enhances the surface soil moisture estimations by reducing the ***RMSE*** to approximately 0.037 (0.017) and raising the ***R*** to approximately 0.92 (0.96). However, the improvements in the ***RMSE*** values in deeper layers are inferior to those in OBS_T because LST assimilation has very little impact on the soil moisture; thus, the improvement in OBS_T is mainly related to the introduction of irrigation information. However, irrigation action alters the hydrothermal conditions and disrupts the water transfer mechanisms between soil layers, which complicates the process of using surface observations to improve deep estimation. Thus, the improvements in the SSM are far superior in OBS_T, while those in the deeper soil moisture are not. The ***RMSE*** values of SM2, SM3, and SM4 from the EnKF and ESIL decline to various degrees and all surpass the corresponding values in any scenario with unknown irrigation. The comparisons of the soil moisture profile estimations (both unknown irrigation and known irrigation) from the different algorithms suggest that the ESIL had a dominant advantage over the EnKF. Fig. 5(a) shows the high coincidence degree between the blue curve (ESIL) and the black curve (OBS), while the green curve (EnKF) approaches closer to OBS compared to Fig. 2(a). The soil moisture curves in deeper layers approximate both the black curve and its variation trajectory (Fig. 5(b)–(h)). The maintenance of water content from irrigation moments in the time domain substantially improves the severe underestimation of the soil moisture and reflects the peaks that are created by irrigation actions in deeper layers.

Table 5 and Fig. 6 show the experimental results for the soil temperature. All the statistical error metrics for the EnKF and ESIL in the three scenarios surpass the corresponding values in Table 3. Irrigation actions alters the hydrothermal conditions, which increases the water amount and decreases the soil temperature in the solum. The soil temperature and soil moisture information is corrected by replacing the profiles with *in-situ* measurements, which successfully facilitates the effectiveness of the assimilation. The EnKF and ESIL exhibit similar performance in a given scenario according to Table 5. A comparison of the ***MBE*** and ***RMSE*** values of the EnKF and ESIL (0.475 and 3.173, 0.572 and 3.166) in OBS_T shows the diminution of deviation error in the soil temperature, which is not seen in Table 3. Hence, the introduction of irrigation information has a homologous impact on the soil temperature through the feedback of SSM assimilation into the CoLM (seen in Table 3), which may have weakened the contribution of SSM assimilation and comparatively degraded the performance by excessively correcting the deviation error in OBS_B. This phenomenon occurs in the EnKF, but the ESIL produces the best results with OBS_B because of the different feedback frequencies of the soil moisture when assimilating the SSM. Meanwhile, the EnKF induces minor changes in OBS_M and OBS_B, in contrast to the unknown irrigation condition, while the ESIL slightly upgrades the ***RMSE*** values. Thus, the irrigation information boosts the performance of the ESIL more remarkably than that of the EnKF. Fig. 6(a)–(d) show the scatter points of the soil temperature in the upper four layers. Distinguishing improvements from Fig. 3(a)–(d) are difficult; the correlation coefficients increased by no more than 0.03. However, compared to Fig. 3(e)–(h), Fig. 6(e)–(h) show obviously aggregated scatter points that do not diverge from the 1:1 line, as with the EnOL. In particular, the correlation coefficients of the EnKF and ESIL for Fig. 6(f) increase by more than 0.2 based on the corresponding ***R*** under unknown irrigation, while those for the other three plots increase unequally by less than 0.1.

The estimated sensible and latent heat fluxes from the assimilation of both the SSM and LST after the addition of irrigation information are displayed in Fig. 7. The mismatch between the simulations and measurements of the two surface turbulent fluxes further decreased for unknown irrigation. The ***RMSE*** values of the sensible and latent heat fluxes for the EnKF decline to 57 W/m^2 and 72 W/m^2, while the ESIL show more noticeable progress with a reduction to 66 W/m^2 and 79 W/m^2. Although assimilations *via* the EnKF and ESIL under unknown irrigation decrease the ***MBE*** values of the two surface turbulent fluxes to approximately 50 percent from 116 W/m^2 and −121 W/m^2, the addition of irrigation reduces these values by over 75%.

4.2. Influence of observation intervals

Fig. 8 presents the ***RMSE*** distributions for the surface soil moisture from the EnKF and ESIL with changing observation intervals under the condition of unknown irrigation, while the bottom two panels are for the surface soil temperature. The labelled points on the coordinate axes comprise the observation interval vectors from experiments with a fixed 0.03 multiplicative standard deviation for the SSM and a fixed 2 K additive standard deviation for the LST. The remainder in the two-dimensional space is replenished *via* interpolation.

As seen in Fig. 8(a) and (b), the surface soil moisture seems to be impervious to the observation frequency of the LST. The limited capability of LST assimilation to improve soil moisture estimation is also shown in Table 2. For the observation intervals of the SSM, the decrease in observation quantity degrades the assimilation performance but influences the EnKF and ESIL to different extents. The ESIL outperforms the EnKF for all combinations of observation intervals in terms of the surface soil moisture. Despite only one observation available in the smoother window as the observation interval expanded to 8 days, the ESIL takes the preferential position with a smaller worst ***RMSE*** (0.08 for the ESIL and 0.96 for the EnKF). The variation in the ***RMSE*** values from the EnKF drastically increases, along with a reduction in the SSM quantity, and tends to stabilize as the observation interval of the SSM exceeded more than five days. The ESIL shows a similar variation pattern with more slug-

gish processes and produces a more satisfying ***RMSE*** values than the EnKF.

The influence of the observation interval on the surface soil temperature is complicated, as shown in Fig. 8(c) and (d). LST assimilation eliminates the stochastic error of the soil temperature, and SSM assimilation removes the deviation error of the soil temperature *via* correcting the soil moisture profile. The instantaneous correction property of the EnKF improves the interactions of the deviation error corrections, which produces a slight advantage compared to the ESIL in terms of estimating the surface soil temperature with an increasing observation interval. This result also reveals that correcting the deviation error plays a more significant role than the stochastic error when observations are sparse. Thus, the ***RMSE*** values of the EnKF show little discrepancy with varying observational intervals and visibly rise along the bottom and top-right corners. The relatively larger ***RMSE*** values with dense observations illustrate that the observation quantity is not always positively correlated to the algorithm performance. The ***RMSE*** values of the ESIL for the surface soil temperature present a typical change trend, with the lowest value in the bottom-left corner and the highest value in the top-right corner, which indicates that an increase in the ***RMSE*** value accompanies the mutual growth of the observation interval of LST and SSM.

4.3. Influence of observation errors

Fig. 9 plots the ***RMSE*** distributions with changing observation errors under the condition of unknown irrigation. The standard deviations that are labelled on the coordinate axes were tested in experiments with one observation a day for the SSM and two observations a day for the LST. The remainder in the two-dimensional space was replenished *via* interpolation. Fig. 9(a) and (b) shows that the ***RMSE*** values for the surface soil moisture obtained by EnKF monotonically increased along with the standard deviation of the SSM, while the growth trend slows down with increasing standard deviation. Similar to Fig. 8, the varying standard deviations of the LST show definite contributions to the soil moisture estimation. The ***RMSE*** distribution of the ESIL for the surface soil moisture is nearly resistant to the variation in the standard deviation of observation, which can be ascribed to the sufficient quantity of observations and the superiority of implementing the EnKS with inflation and localization in the maximum usage of observation information. A comparison of the two top plots shows an accelerated discrepancy between the EnKF and ESIL with increasing standard deviation. The vertical strips in the bottom-left plot indicate that the EnKF is prone to changes in the standard deviation of the LST and that the ***RMSE*** values of the EnKF for the surface soil temperature progressively converge to a minimum point as the standard deviation of the LST decreases. In addition, the ***RMSE*** values of the ESIL for the surface soil temperature exhibit little variation with changing standard deviations for both the LST and SSM, while the negligible divergence is possibly caused by the sampling error of the observations, which makes valid instructive information difficult to determine.

4.4. Influence of parameter estimation

Parameter uncertainty can easily create systematic deviations in model states; thus, two parallel filters were designed to recursively estimate both states and parameters (Chen et al., 2015). The EnKF was employed to estimate parameters while the ESIL was used to update states. Four parameters (minimum soil suction, saturated hydraulic conductivity, porosity, and the "b" parameter) were chosen according to the parameter sensitivity analysis manipulated by Li et al. (2013a,b). The initial parameter sets were uniformly sampled in specific ranges that were defined by Li et al. (2013a,b). The ***NER*** used to measure the magnitude of improvement from the EnOL *via* ESILP (implemented with parameter estimation) and ESIL are plotted in Fig. 10. SM1-SM8 and ST1-ST8 represent the results of soil moisture and temperature profiles, while H and LE stand for the sensible heat flux and latent heat flux, respectively.

Fig. 10(a) shows a comparison of ***NER*** values that are derived from both SSM and LST assimilation under unknown irrigation. Notwithstanding the promising results observed from soil temperature estimation, only trivial improvements in the soil moisture in the first layer are produced with parameter estimation. Moreover, obvious deterioration occurs in the ***NER*** values of soil moisture in deeper layers and surface turbulent fluxes. Equally unsurprisingly, the parameters that are retrieved by the assimilation of surface observations are not expected to always improve the state accuracy in deeper layers, especially with the vertical heterogeneity of the solum and model structure error. The retrieved parameters, called "effective parameters", contribute to the consistency between the surface states with the observations but rarely confirm the simultaneous improvement of multiple states. In addition, unknown irrigation aggravates this phenomenon because the bias caused by irrigation is compensated by the parameters. The results of ESILP and ESIL from assimilating the SSM and LST under known irrigation are displayed in Fig. 10(b). No improvement in the surface states is produced, mainly because of the extremely small bias indicated by ***MBE*** values in Tables 4 and 5 after adding irrigation information. Irrigation information enormously modifies the system deviation; thus, such unapparent bias limits the function of parameter estimation for the surface states. The soil moisture and temperature in deeper layers and the two surface turbulent fluxes performed differently. The updated parameters are imported into the CoLM and replaced the original values to verify the validity of these parameters. The ***RMSE*** values of the surface soil moisture and temperature simulated by the new parameters decreased by approximately 25% and 5%, which confirms the definition of effective parameters.

5. Conclusions

We conducted a series of experiments to demonstrate the influence of the assimilation of multi-source observations, irrigation information, observation intervals, the standard deviations of observations and parameter estimation on various states. First, the assimilation of multi-source observations improves multiple variables, surpassing the performance of the assimilation of single observations. Second, adding irrigation information can improve the estimation accuracy of the soil moisture, soil temperature, and sensible and latent heat fluxes derived from the EnKF and ESIL, especially the soil moisture and temperature in deeper layers. However, the ESIL produces quite encouraging outcomes under unknown irrigation, especially in shallow layers, which expands the prospect of retrieving the soil moisture and temperature profiles and surface turbulent fluxes without detailed irrigation records. The inflation and localization implemented in the original EnKS improves the estimation of the background error matrix and the extension of the observation information, which are vital to account for the limited ensemble size, inaccurate model error specifications, and deficient observations. Finally, the stabilized and promising effectiveness of the ESIL with varying observation intervals and standard deviations broadens its reliability in practical applications. Compared to the parameter estimation results, the ESIL provides comprehensive improvements when we examine the estimation accuracy of multiple states.

With the successful launch of diverse satellites, remote sensing has turned into a widespread approach to monitor the earth environment and resource in global region. Data acquired by sensors such as AMSR-E (Advanced Microwave Scanning Radiometer for Earth Observing System), MODIS, ASTER (Advanced Space-

borne Thermal Emission and Reflection Radiometer) and ASCAT (Advanced Scatterometer) extremely enrich soil moisture and land surface temperature observations. The ESIL, tested with various observation intervals and standard deviations, demonstrates superiority when dealing with some sparse satellite data or different observation errors from satellite data. Thus, our work in the next stage will take full advantage of such plentiful remote sensing data under the data assimilation framework of the ESIL.

Acknowledgements

This work is supported by the National Science Foundation of China under Grants 91325106 and 41271358, the Major State Basic Research Development Program under Grant 2011CB707103, the Hundred Talent Program of the Chinese Academy of Sciences under grant 29Y127D01, and the Cross-disciplinary Collaborative Teams Program for Science, Technology and Innovation of the Chinese Academy of Sciences.

References

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. Adv. Water Resour. 34 (4), 526–536, http://dx.doi.org/10.1016/j.advwatres.2011.01.011.

Chen, W.J., Huang, C.L., Shen, H.F., Li, X., 2015. Comparison of ensemble-based state and parameter estimation methods for soil moisture data assimilation. Adv. Water Resour. 86, 425–438, http://dx.doi.org/10.1016/j.advwatres.2015.08.003.

Chu, N., Huang, C.L., Li, X., Du, P.J., 2015. Simultaneous estimation of surface soil moisture and soil properties with a dual ensemble Kalman smoother. Sci. China Earth Sci. 58, 2327–2339, http://dx.doi.org/10.1007/s11430-015-5175-6.

Clapp, R.B., Hornberger, G.M., 1978. Empirical equations for some soil hydraulic properties. Water Resour. Res. 14 (4), 601–604, http://dx.doi.org/10.1029/WR014i004p00601.

Dai, Y., Zeng, X., Dickinson, R.E., Baker, I., Bonan, G.B., Bosilovich, M.G., et al., 2003. The common land model. Bull. Am. Meteorol. Soc. 84, 1013–1023, http://dx.doi.org/10.1175/BAMS-84-8-1013.

Dunne, S., Entekhabi, D., 2005. An ensemble-based reanalysis approach to land data assimilation. Water Resour. Res. 41 (2), W02013 http://dx.doi.org/10.1029/2004WR003449.

Evensen, G., Van Leeuwen, P.J., 2000. An ensemble Kalman smoother for nonlinear dynamics. Mon. Weather Rev. 128 (6), 1852–1867, http://dx.doi.org/10.1175/1520-0493(2000)128<1852:AEKSFN>2.0.CO;2.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. 99 (C5), 10143–10162, http://dx.doi.org/10.1029/94jc00572htt://dx.org/.

Evensen, G., 2003. The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dyn. 53 (4), 343–367, http://dx.doi.org/10.1007/s10236-003-0036-9.

Fan, W., 2014. Heihe 1 Km LAI Production. Heihe Plan Science Data Center.

Gao, H., Wood, E.F., Drusch, M., McCabe, M.F., 2007. Copula-derived observation operators for assimilating TMI and AMSR-E retrieved soil moisture into land surface models. J. Hydrometeorol. 8 (3), 413–429, http://dx.doi.org/10.1175/JHM570.1.

Gaspari, G., Stephen, E.C., 1999. Construction of correlation functions in two and three dimensions. Q. J. R. Meteorol. Soc. 125 (554), 723–757, http://dx.doi.org/10.1002/qj.49712555417.

Hain, C.R., Crow, W.T., Anderson, M.C., Mecikalski, J.R., 2012. An ensemble Kalman filter dual assimilation of thermal infrared and microwave satellite observations of soil moisture into the Noah land surface model. Water Resour. Res. 48 (11), http://dx.doi.org/10.1029/2011wr011268, W11517.

Houtekamer, P.L., Mitchell, H.L., 2001. A sequential ensemble Kalman filter for atmospheric data assimilation. Mon. Weather Rev. 129 (1), 123–137, http://dx.doi.org/10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2.

Huang, C.L., Li, X., Lu, L., 2008a. Retrieving soil temperature profile by assimilating MODIS LST products with ensemble Kalman filter. Remote Sens. Environ. 112 (4), 1320–1336, http://dx.doi.org/10.1016/j.rse.2007.03.028.

Huang, C.L., Li, X., Lu, L., Gu, J., 2008b. Experiments of one-dimensional soil moisture assimilation system based on ensemble Kalman filter. Remote Sens. Environ. 112 (3), 888–900, http://dx.doi.org/10.1016/j.rse.2007.06.026.

Ines, A.V.M., Honda, K., DasGupta, A., Droogers, P., Clemente, R.S., 2006. Combining remote sensing-simulation modeling and genetic algorithm optimization to explore water management options in irrigated agriculture. Agric. Water Manage. 83 (3), 221–232, http://dx.doi.org/10.1016/j.agwat.2005.12.006.

Jia, B., Xie, Z., Tian, X., Shi, C., 2009. A soil moisture assimilation scheme based on the ensemble Kalman filter using microwave brightness temperature. Sci. China Ser. D Earth Sci. 52 (11), 1835–1848, http://dx.doi.org/10.1007/s11430-009-0122-z.

Kumar, P., Kaleita, A.L., 2003. Assimilation of near-surface temperature using extended Kalman filter. Adv. Water Resour. 26 (1), 79–93, http://dx.doi.org/10.1016/S0309-1708(02)00098-2.

Lakshmi, V., 2000. A simple surface temperature assimilation scheme for use in land surface models. Water Resour. Res. 36 (12), 3687–3700, http://dx.doi.org/10.1029/2000WR900204.

Lawston, P.M., Santanello Jr., J.A., Zaitchik, B.F., Rodell, M., 2015. Impact of irrigation methods on land surface model spinup and initialization of WRF forecasts. J. Hydrometeorol. 16 (3), 1135–1154, http://dx.doi.org/10.1175/JHM-D-14-0203.1.

Lei, F.N., Huang, C.L., Shen, H.F., Li, X., 2014. Improving the estimation of hydrological states in the SWAT model via the ensemble Kalman smoother: synthetic experiments for the Heihe River Basin in northwest China. Adv. Water Resour. 67, 32–45, http://dx.doi.org/10.1016/j.advwatres.2014.02.008.

Li, X., Li, X., Li, Z., Ma, M., Wang, J., et al., 2009. Watershed allied telemetry experimental research. J. Geophys. Res. 114, D22103, http://dx.doi.org/10.1029/2008JD011590.

Li, X., Cheng, G., Liu, S., Xiao, Q., Ma, M., Jin, R., et al., 2013a. Heihe watershed allied telemetry experimental research (HiWATER): Scientific objectives and experimental design. Bull. Am. Meteor. Soc. 94 (8), 1145–1160, http://dx.doi.org/10.1175/BAMS-D-12-00154.1.

Li, J., Duan, Q.Y., Gong, W., Ye, A., Dai, Y., Miao, C., et al., 2013b. Assessing parameter importance of the Common Land Model based on qualitative and quantitative sensitivity analysis. Hydrol Earth Syst Sci 17 (8), 3279–3293, http://dx.doi.org/10.5194/hess-17-3279-2013.

Liao, Y., Fan, W., Xu, X., 2013. Algorithm of Leaf Area Index Product for HJ-CCD over Heihe River Basin , Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International, pp. 169–172.

Liu, S., Xu, Z., Zhu, Z., Jia, Z., Zhu, M., 2013. Measurements of evapotranspiration from eddy-covariance systems and large aperture scintillometers in the Hai River Basin. China. J Hydrol. 487, 24–38, http://dx.doi.org/10.1016/j.jhydrol.2013.02.025.

Moiwo, J.P., Tao, F.L., 2015. Contributions of precipitation, irrigation and soil water to evapotranspiration in (semi)-arid regions. Intel. J. Climat. 35 (6), 1079–1089, http://dx.doi.org/10.1002/joc.4040.

Ozdogan, M., Salvucci, G.D., 2004. Irrigation-induced changes in potential evapotranspiration in southeastern Turkey: test and application of Bouchet's complementary hypothesis. Water Resour. Res. 40 (4), W04301, http://dx.doi.org/10.1029/2003WR002822.

Wang, D., Cai, X., 2007. Optimal estimation of irrigation schedule—an example of quantifying human interferences to hydrologic processes. Adv. Water Resour. 30 (8), 1844–1857, http://dx.doi.org/10.1016/j.advwatres.2007.02.006.

Wu, G., Zheng, X., Wang, L., Zhang, S., Liang, X., Li, Y., 2013. A new structure for error covariance matrices and their adaptive estimation in EnKF assimilation. Q. J. R. Meteorol. Soc. 139 (672), 795–804, http://dx.doi.org/10.1002/qj.2002.

Xu, T., Liu, S., Liang, S., Qin, J., 2011. Improving predictions of water and heat fluxes by assimilating MODIS land surface temperature products into the common land model. J. Hydrometeorol. 12 (2), 227–244, http://dx.doi.org/10.1175/2010JHM1300.1.

Yang, K., Watanabe, T., Koike, T., Li, X., FUJII, H., Tamagawa, K., et al., 2007. Auto-calibration system developed to assimilate AMSR-E data into a land surface model for estimating soil moisture and the surface energy budget. J. Meteorol. Soc. Jpn. 85 (A), 229–242, http://dx.doi.org/10.2151/jmsj.85a.229.