Hydrol. Earth Syst. Sci. Discuss., 8, 1433–1468, 2011 www.hydrol-earth-syst-sci-discuss.net/8/1433/2011/ doi:10.5194/hessd-8-1433-2011 © Author(s) 2011. CC Attribution 3.0 License.



This discussion paper is/has been under review for the journal Hydrology and Earth System Sciences (HESS). Please refer to the corresponding final paper in HESS if available.

Simultaneous estimation of land surface scheme states and parameters using the ensemble Kalman filter: idealized twin experiments

S. Nie¹, J. Zhu², and Y. Luo¹

¹National Climate Center, China Meteorological Administration, Beijing, China ²Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

Received: 20 December 2010 - Accepted: 13 January 2011 - Published: 28 January 2011

Correspondence to: S. Nie (niesp@cma.gov.cn)

Published by Copernicus Publications on behalf of the European Geosciences Union.





Abstract

The performance of the ensemble Kalman filter (EnKF) in soil moisture assimilation applications is investigated in the context of simultaneous state-parameter estimation in the presence of uncertainties from model parameters, initial soil moisture condition and atmospheric forcing. A physically-based land surface model is used for this purpose. Using a series of idealized twin experiments, model generated near-surface soil moisture observations are assimilated to estimate soil moisture state and three hydraulic parameters (the saturated hydraulic conductivity, the saturated soil moisture suction and a soil texture empirical parameter) in the model. The single imperfect parameter can be successfully estimated using the EnKF. Results show that all the three estimated parameters converge toward their respective true values, while the root mean squared errors (RMSE) of soil moisture associated with these parameters is on average reduced by 54% and 53% comparing with the non-parameter-estimation benchmark RMSE for near-surface layer and root zone layer, respectively. The performance of simultaneous

- ¹⁵ multi-parameter estimation is significant degraded, mainly because the inherent balance relationship of these parameters is broken and the degree of freedom increases in assimilation processes. By introducing constraints between estimated parameters, the performance of the constraint-based simultaneous multi-parameter estimations are as good as that of single-parameter cases even assimilating temporal-sparse obser-
- vations. In terms of the relative root mean squared error (RRE), the constraint-based estimation cases can achieve 36% to 53% in near-surface layer and 25% to 50% in root zone layer for different assimilation intervals ranging from 1-day to 40-days. This result suggests that the greatest advantage of this method can be displayed with a proper temporal-sparse assimilation interval of 10-days as actual measurement inter-
- val of conventional in situ soil moisture observations. As these obtained constraints are mostly in statistical sense, this constraint-based simultaneous state-parameter estimation scheme is supposed to be suitable for other land surface models in soil moisture assimilation applications.





1 Introduction

Soil moisture is a key state variable which controls the partitioning of water and energy fluxes at the land surface. It has an important influence on the surface water cycle, thereby influencing the latent heat flux and hence the surface energy balance. As

a numerical simulation to realistic land surface state, land surface model (LSM) is a popular instrument used to provide proper soil moisture initial conditions for numerical weather prediction models and climate models. However, as been quite simplified in physical and mathematical processes, LSM can only represent actual processes in nature approximately. Uncertainties in hydrodynamic processes, model variables and model parameters lead to large errors in the simulation of soil moisture condition. How to properly initialize soil moisture condition in LSMs is still an open and classical issue

in meteorology and hydrology research. Modern data assimilation technique is an effective approach to account for this issue. Merging information from uncertain soil moisture observations and uncertain land

- ¹⁵ model predictions optimally, this technique can improve the estimation of soil moisture state in LSMs (Houser et al., 1998; Reichle et al., 2001a,b). As a new-emerging sequential assimilation method, the ensemble Kalman filter (EnKF) has received an increasing amount of research attention in recent years. It is a Monte Carlo approximation to the traditional Kalman filter (Kalman and Bucy, 1961) which was first introduced
- ²⁰ by Evensen (1994). By propagating an ensemble of state vectors in parallel such that each state vector represents a particular realization of generated model replicates, it provides flow-dependent estimates of background error and can more optimally adjust the background to newly available observations. In recent years, the EnKF has been successfully applied to different soil moisture assimilation problems (Walker and
- Houser, 2001; Reichle et al., 2002a,b, 2008; Reichle and Koster, 2005; Crow and Van Loon, 2006; Crow and van den Berg, 2010; Ni-Meister et al., 2006; Zhang et al., 2010). In most of these studies, however, it is noticeable that the EnKF was only used for estimating time-varying state variables under the presumption that model parameters





were to be specified in advance by calibrations. In common calibration methods, model parameters are adjusted in such a way that the behavior of model approximates, as closely and consistently as possible, the observed represents of the real land system over some period of time using a historical batch of measurements (Niyogi et al., 2002;

- Xia et al., 2002; Coudert et al., 2006). There are two main weaknesses exist in such approach: (i) it can not include information from new observations, and (ii) it only address parameter error while errors from initial condition and atmospheric forcing data are ignored. As the EnKF is able to account for a wide range of possible model errors easily (Evensen, 2003), it has the potential to overcome these two drawbacks by
 explicitly taking all sources of uncertainty into account and developing a simultaneous
- treatment of state and parameter estimation to refine its assimilation performance.

By the means of state augmentation technique (Anderson, 2001), model parameter estimation can easily be included in the framework of the EnKF. The principle of state augmentation is that model parameters can be considered as parts of model

- states beside conventional state variables, and then the error covariance sampled by ensemble members can be used directly to update those model parameters in exactly the same manner as for the conventional state variables. Recently, the simultaneous state-parameter estimations using the EnKF have been successfully applied in atmospheric, oceanic, hydrologic and ecologic assimilation fields (Aksoy et al., 2006; Vrugt
- et al., 2005; Annan et al., 2005; Moradkhani et al., 2005; Chen et al., 2008). All these studies provide encouraging results and show that the EnKF is a robust and effective algorithm in simultaneous state-parameter estimation. However, similar studies in soil moisture assimilation are few at present. The study of Montaldo et al. (2007) shows that soil moisture assimilation approaches may fail when key LSM hydraulic parame-
- ²⁵ ter is estimated poorly. Therefore, in this paper we aim to thoroughly investigate the application of EnKF-based parameter estimation method in soil moisture assimilation.

Despite promising results are obtained from the applications of the EnKF in parameter estimation, some deficiencies still exist in these studies. A noticeable one is the decline of estimation performance when multiple imperfect parameters are estimated





simultaneously (Aksoy et al., 2006; Jung et al., 2010; Moradkhani et al., 2005). When the number of estimated parameters increases to a certain extent, some parameters could not converge to their benchmark "true" values even with long enough estimation periods. Recently, this problem is still a challenge in the application of the EnKF in parameter estimation. Therefore, the second objective of this paper is to impose a new "constraint-based" parameter estimation procedure for solving this problem and subsequently to improve the EnKF performance in simultaneous multi-parameter estimation of soil moisture and model parameters.

The organization of the paper is as follow. The model and parameter estimation framework used in this study are explained in Sect. 2. The background and approach of idealized twin experiments are presented in Sect. 3. The results and analyses are presented in Sect. 4. Finally, we conclude the paper in Sect. 5.

2 Land surface model and parameter estimation framework

2.1 Land surface model

5

- ¹⁵ The land surface model used is the Atmosphere-Vegetation Interaction Model (AVIM) (Ji, 1995), which contains a physical process mode and a vegetation biological process mode. Detailed descriptions of this model are given by Ji and Hu (1989) and Ji (1995). The version used in this study only considers the physical process mode which is a typical soil-vegetation-atmosphere (SVAT) type model developed by Ji and Hu (1989).
- Hu (1989). This model includes three soil layers with thicknesses of 0.1, 0.9, and 3.6 m from ground. The layer-averaged soil moisture and temperature are modeled for each of the three soil layers. In the deepest layer, both soil water flux and heat flux are assumed to be zero with constant soil moisture and temperature condition in this layer. The change of soil moisture in near-surface layer and root zone layer over a time step
- is controlled by the change in water flux over these two layers. Richards' equation for unsaturated flow is used for the simulation of this flux. It is expressed as:





$$F(z) = -k(z) \frac{\mathrm{d}\psi(z)}{\mathrm{d}z} \bigg|_{z} + k(z) \quad z \neq 0$$

where z is the depth and F(z) is the soil water flux. The unsaturated hydraulic conductivity k(z) and unsaturated soil water suction $\psi(z)$ are defined in Clapp and Hornberger (1978):

5
$$k(z) = k_{sat} \left(\frac{\delta}{\delta_{sat}}\right)^{2b+3}$$
 (2)
 $\psi(z) = \psi_{sat} \left(\frac{\delta}{\delta_{sat}}\right)^{-b}$ (3)

where δ and δ_{sat} are the unsaturated and saturated soil moisture; k_{sat} and ψ_{sat} are the saturated hydraulic conductivity and soil moisture suction respectively; andb is a soil texture empirical parameter. In this paper, parameters k_{sat} , ψ_{sat} and b are chosen to be estimated.

10

15

20

2.2 Parameter estimation framework using the EnKF

Parameter estimation frameworks used in this paper are based on the EnKF. As the comprehensive presentation of the standard EnKF is given by Evensen (2003), this subsection represents mainly the modifications to the standard EnKF after considering simultaneous state-parameter estimation in its framework.

State-parameter estimation without constraint 2.2.1

To extend the applicability of the EnKF to simultaneous state-parameter estimation, it is needed to build an evolution of parameter similar to that of model state variable. By adding mean-zero Gaussian random noise τ_{t-1}^{i} with covariance Q_{t-1}^{θ} to parameter θ^{i} , the evolution of parameter can be expressed in the form of:

$$\theta_{t}^{i-} = \theta_{t-1}^{i+} + \tau_{t-1}^{i}, \ \tau_{t-1}^{i} \sim N\left(0, \ Q_{t-1}^{\theta}\right)$$
1438

(1)

(4)

The superscripts "–" and "+" refer to states in forecast step and update step respectively. When multiple parameters are to be estimated simultaneously, perturbations on different parameters are considered as mutually independent for simplicity.

With artificially perturbed parameters, time evolution for each ensemble member of state vector \mathbf{x}^{i} in the EnKF can be expressed as follows:

$$x_t^{i-} = f_t \left(x_{t-1}^{i+}, u_{t-1}^{i}, \theta_{t-1}^{i+} \right), \quad i = 1, 2, ..., n.$$
(5)

where x_t^{i-} is the *i*-th forecast ensemble member at time *t* and x_{t-1}^{i+} is the *i*-th updated ensemble member at time t-1. The nonlinear operator f(.) denotes the land surface model processes which contain state vectors \mathbf{x}^i , forcing data vectors \mathbf{u}^i , and model parameter vectors $\boldsymbol{\theta}^i$. The forcing data perturbations are made by adding mean-zero Gaussian noise μ_{t-1}^i with covariance Q_{t-1}^u to the forcing data at each time step:

$$u_{t-1}^{i} = u_{t-1} + \mu_{t-1}^{i}, \ \mu_{t-1}^{i} \sim N\left(0, \ Q_{t-1}^{u}\right)$$

When observations are available, each ensemble member of state vector is updated as follows:

15
$$X_t^{i+} = X_t^{i-} + \mathbf{K}_t^{\theta, x} \left(y_t^i - \mathbf{H}_t x_t^{i-} \right)$$
 (7)

where \mathbf{H}_t is the measurement operator and y_t^i is the *i*-th member of observation ensemble generated by adding mean-zero random measurement error η_t^i with covariance Q_t^y to actual observation (Burgers et al., 1998):

$$y_t^i = y_t + \eta_t^i, \ \eta_t^i \sim N \left(0, \ Q_t^{\gamma}\right)$$
(8)

²⁰ $\mathbf{K}_{t}^{\theta,x}$ is the Kalman gain matrix that considers simultaneous state-parameter estimation. It is obtained by:

$$\mathbf{K}_{t}^{\theta,x} = \mathbf{P}_{t}^{\theta,x,-} \mathbf{H}_{t}^{\mathsf{T}} \left(\mathbf{H}_{t} \mathbf{P}_{t}^{\theta,x,-} \mathbf{H}_{t}^{\mathsf{T}} + \mathbf{R}_{t} \right)^{-1}$$
1439

10



(6)

(9)

where $\mathbf{P}_{t}^{\theta,x,-}$ and \mathbf{R}_{t} are forecast error covariance matrix and observation error covariance matrix respectively. $\mathbf{P}_{t}^{\theta,x,-}$ is computed as the sample covariance from forecast ensemble of model state variables and parameters. It is defined as an ensemble covariance matrix around the ensemble mean:

where, $\mathbf{X}_t = [x_t^{1-} - \bar{x}_t^{-}, ..., x_t^{n-} - \bar{x}_t^{-}; \theta_t^{1-} - \overline{\theta}_t^{-}, ..., \theta_t^{n-} - \overline{\theta}_t^{-}]$ and $\bar{x}_t^{-} = \frac{1}{n} \sum_{i=1}^n x_t^{i-}, \overline{\theta}_t^{-} = \frac{1}{n} \sum_{i=1}^n \theta_t^{i-}$ denote the ensemble mean of forecast state variables and parameters, respectively.

2.2.2 State-parameter estimation with constraint

Actually, some statistical relationships exist between different model parameters (e.g. Cosby et al., 1984; Rawls et al., 1982; Schaap and Leij, 2000; van Genuchten, 1980; Zhuang et al., 2001). As additional information, these statistical constraints between parameters are also needed to be taken into account in the framework of the EnKF to perform better state-parameter estimation.

In the general case the constraints are nonlinear, which can be expressed as:

15

20

$$g_t^* = G_t \left(\theta_t^* \right)$$

where g_t^* denotes nonlinear constraints between different model parameters θ_t^* at time t. Without losing any generality, the parameters which do not contained in g_t^* are defined as $\theta^{''}$. The post-constrained update to each ensemble member of state vectors x^i is computed as follows:

$$x_{t}^{i+} = x_{t}^{i-} + \mathbf{K}_{t}^{\theta^{-},g^{*}} \left(y_{t}^{i} - \mathbf{H}_{t} x_{t}^{i-} \right)$$
(12)



(11)



where $\mathbf{K}_{t}^{\theta^{^{\prime}},g^{^{*}}}$ is the Kalman gain matrix including constraints. It is obtained as follows:

$$\mathbf{K}_{t}^{\theta^{''},g^{*}} = \mathbf{P}_{t}^{\theta^{''},g^{*},-} \mathbf{H}_{t}^{\mathsf{T}} \left(\mathbf{H}_{t} \mathbf{P}_{t}^{\theta^{''},g^{*},-} \mathbf{H}_{t}^{\mathsf{T}} + \mathbf{R}_{t} \right)^{-1}$$
(13)

where $\mathbf{P}_t^{\theta^{-},g^{*,-}}$ is the post-constrained error covariance matrix of states ensemble and parameters ensemble. Note that each ensemble member will satisfy the constraints.

5 3 Experiments background and approach

The study in this paper is based on a series of idealized twin experiments taking soil moisture in top two layers and parameters k_{sat} , ψ_{sat} and *b* as state variables in the EnKF. The design of idealized twin experiment is similar to that of Crow and Van Loon (2006), with the assumptions that the "true" states are model-generated and the source and magnitude of model errors and observation errors are perfectly known in a statistical sense. This approach avoids a number of key complexities facing to assimilate actual soil moisture observations and makes the parameter estimation behavior of the EnKF more transparent. However, it needs to be noted that little information about the statistical properties of errors may degrade the performance of the EnKF in parameter estimation in realistic soil moisture assimilation.

Because the objective of this study is to investigate the feasibility of EnKF-based parameter estimation in soil moisture assimilation, all experiments here are conducted at point scale for computational simplification. Jiangji station (the outlet of Shiguanhe sub-basin in the Huaihe River Basin) of the HUaihe river Basin EXperiment (HUBEX,

²⁰ China's contribution to GEWEX Asian Monsoon Experiment; Fujiyoshi et al., 2006) is chosen as experiment site for having comprehensive meteorological forcing data sets. In this station, soil texture is sandy loam and vegetation type is broadleaf shrubs with bare soil. The experiment period covers the entire year in 1998. During the Intensive Observation Period in 1998 (from 21 May to 31 August), hourly gauge-based





precipitation, once daily air temperature, humidity, surface pressure and wind speed data sets were available in Jiangji. During other period in the year, daily observations of these meteorological forcing from the Gushi meteorological site (about 15 km from Jiangji) was used. Not having incoming radiation observation, the radiation forcing data from the NCEP (National Centers for Environmental Prediction) reanalysis dataset version 1 was used as a substitute. All these forcing data sets were used to force the AVIM in Jiangji in all idealized twin experiments with time step of half hours for the model and one-day frequency for assimilating soil moisture "observations".

The "true" soil moisture state is obtained by integrating the AVIM from a 2-yr spinup initial condition on 1 January 1998 to 31 December 1998 with standard AVIM parameters (Ji and Hu, 1989; Ji, 1995) and atmospheric forcing data described above. To get "prior" state of soil moisture, model error from three sources (i) parameters (k_{sat} , ψ_{sat} and b), (ii) soil moisture initial conditions, and (iii) precipitation and short-wave (longwave) radiations (Margulis et al., 2002; Reichle et al., 2002b) are considered. In prior integration, errors in parameters and initial soil moisture condition are generated by replacing the "truth states" values with assumed imperfect values; errors in precipitation

- and radiation are imposed by adding mean-zero Gaussian random noises (as shown in Eq. 6) to the true forcing fields. Specific differences between "true" and "prior" integrations are listed in Table 1. Collectively, these differences in parameter, initial condition
- and forcing data are considered as "actual errors" and represent our imperfect understanding to the true soil moisture states. In all these idealized twin experiments, the "actual observation" to be assimilated is the near-surface soil moisture. It is derived from the true state by adding mean-zero Gaussian random errors with a standard deviation of 5% volumetric moisture percent once a day. The precipitation forcing, the "true"
- ²⁵ and "prior" soil moisture for top two layers and the "actual observation" of near-surface soil moisture used in idealized twin experiment are displayed in Fig. 1. For errors in hydraulic parameters, initial condition, and atmospheric forcing data, there are significant deviations from prior states to true states of soil moisture in both layers.





Given the statistical properties of model errors and observation errors, the EnKF attempts to modify prior state back to the true state by assimilating "actual observations" (Crow and Van Loon, 2006). In the filter, the number of ensemble size is set to 100 in all experiments, which achieved a balance between the computational effort of processing

- a large number of runs and the need for having a sufficiently large set of ensembles to characterize the ensemble distribution. The ensemble of soil moisture initial values is generated by adding zero mean Gaussian noise with a standard deviation of 50% to the prior values at the first time step. The ensemble of forcing data is generated by perturbing prior forcing data with the same statistical properties as the actual forcing
- ¹⁰ data errors once a day. The random perturbation method was also applied to obtain the ensemble of model parameters. A noticeable issue here is the magnitude of standard deviation of perturbation on model parameters since no straightforward guidance exists for proper range of deviation of parameters to be estimated. Because parameters are not dynamical variables, the variances of them are reduced at the update step
- ¹⁵ but remain constant at the forecast step. This causes the variances of parameters to decrease progressively and may lead to filter divergence in parameters. To avoid filter divergence, therefore, perturbation on parameters is implemented in a similar way as that on forcing data according to Eq. (4) with a time interval of 10 days. The standard deviations of the perturbations on k_{sat} , ψ_{sat} , and *b* are chosen as $1.7 \, {\rm e}^{-6} \, {\rm ms}^{-1}$, 0.02 m,
- and 0.25 respectively, which are much smaller than the orders of parameters themselves. Small standard deviation and sparse perturbation interval on these parameters in the filter processes can avoid the behavior of the model being shocked for sharp change of parameters in model integration.

In non-constrained parameter estimation framework, perturbations on different parameters are considered as mutually independent. In constrained estimation framework, statistical relationships between these hydraulic parameters are appropriate to be taken into account in assimilation processes. As some literatures (e.g. Rawls et al., 1982; Zhuang et al., 2001) did not have all relationships between all these three hydraulic parameters and others (e.g. Schaap and Leij, 2000; van Genuchten, 1980) did





not have enough soil classifications as that in the AVIM model, therefore, the literature of Cosby et al. (1984) is chose here for having unified soil samples to get statistical relationships between parameters k_{sat} , ψ_{rmsat} , and b. The Table 5 in Cosby's paper can be formulized as follows to explicitly display constraints between these three parameters:

$${}_{5} \Delta k_{\text{sat}} = \left(\beta_{1}^{\Delta b} - 1\right) k_{\text{sat}}$$
 (14)

$$\Delta \psi_{\text{sat}} = \left(\beta_2^{\Delta b} - 1\right) \psi_{\text{sat}}$$

where Δk_{sat} , $\Delta \psi_{sat}$, and Δb are the perturbations of parameters k_{sat} , ψ_{sat} , and b respectively, and β_1 , β_2 are constrained coefficients assigned as 1.2474 and 0.827 respectively according to statistical relationships obtained from Cosby et al. (1984). These two equations will be included in the error covariance matrix in Eq. (13) to constrain random perturbations of parameters in post-constrained estimation experiments.

4 Results

10

4.1 Single-parameter estimation results

Results from the individual estimation of these three hydraulic parameters are presented in Figs. 2 and 3. In each experiment only one such parameter is perturbed around its imperfect mean value while other parameters are kept unperturbed at their true values.

Figure 2 shows the one year evolution (one-daily analyses) of the ensemble mean parameter values along with the true parameter values stay constant in time. The area between two dash gray lines around the estimated mean parameter value represents the 1- σ (one standard deviation) limits of the parameter spread. These standard deviation limits are computed by averaging the standard deviations of each 100-member ensembles at forecast step. Successful parameter estimation should be that the error



(15)



of the estimated parameter is smaller than or very close to the 1- σ limit. It can be seen that estimated mean parameter values of all three parameters converge to their true values within several months and the true values stay stably within the 1- σ limit subsequently. When these model parameters are included in the augmented state vectors of the FeV/F the methods are for a several back to the seve

tors of the EnKF, the perturbation of parameter can lead to the update of soil moisture state to a certain extent. We assume the general form of the relationship between the perturbations of these parameters and the update of soil moisture as follows:

 $\Delta sm \sim f_1 (\Delta k_{sat}) + f_2 (\Delta \psi_{sat}) + f_3 (\Delta b)$

10

where Δk_{sat} , $\Delta \psi_{sat}$, and Δb are the perturbations of parameters k_{sat} , ψ_{sat} , and b, and Δsm is the update of soil moisture. Nonlinear operators $f_1(\cdot)$, $f_2(\cdot)$ and $f_3(\cdot)$ denote the sampled relationships between Δsm and Δk_{sat} , $\Delta \psi_{sat}$, Δb respectively. In each single-parameter estimation experiment, one of these operators in Eq. (16) is considered in

- the EnKF update process. In assimilation processes, therefore, available near-surface soil moisture observation information can be transferred by the operator to correct the error in corresponding imperfect parameter and make it to be estimated successfully.
- Further analysis finds that the convergence rate of approach to the true values is different among these three parameters. Here, we define "approach time" as the time taken for a true parameter value to first fall within the $1-\sigma$ limit around the estimated mean parameter. It can be seen that the approach times are about 4 months for parameters
- k_{sat} and *b* and about 10 months for parameter ψ_{sat} . As the approach time can scale the efficiency of the EnKF to estimate each parameter to a certain extent, this result implies that errors in parameter k_{sat} and *b* are easier to be corrected than that in parameter ψ_{sat} by assimilating the same soil moisture observation in individual parameter estimation experiments.
- In addition to the mean parameter values, the evolutions of the root mean squared error (RMSE) of one-day-ahead soil moisture forecasting in top two layers are displayed in Fig. 3. In all panels, the RMSE from respective estimation experiments are plotted along with the RMSE from "non-parameter-estimation" benchmark experiments. The



(16)



results of non-parameter-estimation benchmark experiments are obtained by considering imperfect parameters but no parameter estimation in assimilation processes. For all three parameters, the RMSE of the estimation experiments is lower than that of the non-parameter-estimation benchmark experiments for both two soil layers. In these
 ⁵ idealized twin experiments, parameter error is considered as one of the main error sources of soil moisture simulation. With parameter errors been reduced by estimation processes in the EnKF, therefore, less error contributions from imperfect parameters can refine the performance of soil moisture forecasting in each estimation experiment. Further analysis reveals that the decrease of RMSE from non-parameter-estimation
 ¹⁰ experiments to estimation experiments varies among parameters. To quantify relative estimation performance, we define the "relative root mean squared error", which is computed as follows:

$$RRE = \frac{\langle RMSE_{No-Estimation} \rangle - \langle RMSE_{Estimation} \rangle}{\langle RMSE_{No-Estimation} \rangle} \times 100\%$$
(17)

where the operation ⟨·⟩ denotes time average over the entire experiment period. The
 RRE is a relative measure of how much error has been reduced by parameter estimation comparing to non-parameter-estimation benchmark experiments. The results of time average RMSE and RRE in three single-parameter estimation experiments are summarized in Table 2. It is shows that the RMSE of near-surface layer is larger than that of root zone layer for all experiments. One possible reason might be that imperfect

- ²⁰ atmospheric forcing data has larger effects on soil moisture in near-surface layer than that in deeper layer. Among these three parameters, the largest RRE is exhibited by the parameter *b* (72% for near-surface layer and 66% for root zone layer). This result implies that the soil moisture forecasting is more sensitive to the error in parameter *b* than that in other two parameters, which is consistent with the sensitivity analysis of
- ²⁵ Wen et al. (1998). Combining the analyses of ensemble mean parameter values and RMSE above, it can be concluded that the EnKF-based single-parameter estimation perform successfully in soil moisture assimilation.





4.2 Multi-parameter estimation results

To obtain a comprehensive picture of the EnKF's capability and limits in parameter estimation when multiple imperfect parameters are involved, the results from simultaneous three-parameter estimation experiments are presented here.

- Firstly, the case with mutually independent parameter perturbation is discussed. Figure 4 shows the evolution of the ensemble mean parameter values. Significant degrade of simultaneous estimation performance for all the three parameters can be observed. The estimated ensemble mean values of all three parameters can not converge to their true values throughout the entire experiment period. Similar results are also obtained
 from simultaneous dual-parameter estimation cases (figures are not shown). Differ-
- ent from single-parameter estimation case, sampled relationships of all three parameters in Eq. (16) need to be considered simultaneously in the EnKF for the increase of imperfect parameters here. Therefore, the degree of freedom in assimilation processes increases and makes the simultaneous multi-parameter estimation unstable
- and intractable. In the update process of assimilation, moreover, independent random perturbations added on different parameters may break inherent balance relationships between them and deteriorate the performance of parameter estimation. For these two reasons, the EnKF fails in the simultaneous multi-parameter estimation case with mutually independent parameter perturbations, despite soil moisture state can still be actimated excession (figure are not above). For despite with eigenvectors are used.
- estimated successfully (figures are not shown). For dealing with simultaneous multiparameter estimation properly, constraints in Eqs. (14) and (15) (obtained from Cosby et al., 1984) should be taken into account to improve estimation performance.

Next, the simultaneous multi-parameter estimation case with constrained parameter perturbation is discussed. The evolutions of the ensemble mean parameter val-

²⁵ ues are shown in Fig. 5. It can be seen that the estimated ensemble mean values for all three parameters converge to their true values successfully with the same approach time of nine month. From the viewpoint of adding information, imposing these constraints between model parameters also add new information to the assimilation





system in addition to the available soil moisture observation. Therefore, even if the degree of freedom in the assimilation system increase for more imperfect parameters been considered in simultaneous multi-parameter estimation, these additional correlation between parameters can be used effectively to closure the Eq. (16) and offset the

- ⁵ degrade of the EnKF's performance. Moreover, constraints can also make observation information to be transferred to all estimated imperfect parameters in a balanced way and keep these parameters being corrected in a coordinated and consistent way during the whole update processes of assimilation. For these reasons, this constrained simultaneous multi-parameter estimation method displays a good performance in soil
- ¹⁰ moisture assimilation. Because these constrained relationships between parameters are always in statistical sense and can be obtained from literature or from standard parameter tables of the model, this constraint-based parameter estimation method has potentialities to be used in other land surface models even with more imperfect parameters to be estimated.
- The RMSE evolutions of constrained estimation experiment, non-constrained estimation experiment, and non-parameter-estimation benchmark experiment are displayed in Fig. 6. It can be seen that the RMSE of constrained and non-constrained estimation experiments are both lower than that of non-parameter-estimation benchmark experiment. The RREs are 44% and 25% for near-surface and root zone layer respec-
- tively in the constrained estimation case, while 35% and 12% in the non-constrained estimation case. With error contributions from imperfect parameter been reduced by estimation processes, the RRE in constrained estimation is largeer than that in non-constrained estimation for both layers. It also can be noted that, the impact of considering constraints in parameter estimation on the model soil moisture state are not
- significant enough even after the approach time. As the movement of water in soil is a slow process, it needs some time to embody the effects of parameter errors on soil moisture state in land model. Therefore, we guess that frequent corrections to soil moisture state by assimilating daily observations may weaken the advantage of the constrained parameter estimation in soil moisture forecasting. To verify this guess,





we explore further the behaviors of this constraint-based simultaneous multi-parameter estimation method with temporal-sparse observation in the following section.

4.3 Sparse observation assimilation results

Actual conventional in situ soil moisture observations in China and other areas in the world are sparse in time with a measurement interval of about 10-days (Robock et al., 2000; Nie et al., 2008). For applying the constrained simultaneous multi-parameter estimation method in assimilating actual in situ soil moisture in the future, the performance of it with temporal-sparse observation conditions are tested in this subsection by concerned idealized twin experiments.

- Results from the case with 10-days assimilation interval are shown in Figs. 7 and 8. It can be seen from Fig. 7 that, the performances of ensemble mean values for all parameters are comparable to that of one-day assimilation interval experiments in Fig. 5. The estimated mean values of all three parameters can successfully converge toward their true values. Figure 8 shows that the RMSEs of soil moisture of top two lay-
- ers in both constrained and non-constrained estimation experiments are lower than that in non-parameter-estimation benchmark experiment. However, different from the one-day assimilation interval experiments, there are significant reductions of RMSE from non-constrained estimation to constrained estimation here especially after the approach time. Because assimilation interval is large enough in this case, errors in
- different sources (especially in model parameters) have more time to accumulate into soil moisture state in the forecast step of the EnKF. Therefore, when errors in imperfect parameters are corrected better by the EnKF in constrained estimation experiment than that in non-constrained estimation experiment, model integration with little parameter error can produce significantly improvement in soil moisture forecasting. These
- results indicate that this EnKF-based constrained simultaneous multi-parameter estimation method has strong applicability in soil moisture assimilation, even if available observations are as sparse in time as actual conventional in situ soil moisture observations.





Results from further experiments with 20-, 30- and 40-days assimilation intervals show that the estimated ensemble mean values of all three parameters can converge to their true values within the experiment period (figures are not shown), despite the approach times and the 1- σ limits of estimated parameters increase with the increase of assimilation intervals. The RMSE evolution results of these extremely sparse assim-

- ⁵ of assimilation intervals. The RMSE evolution results of these extremely sparse assimilation intervals experiments are displayed in Fig. 9. With the increase of assimilation intervals from 10 days to 40 days, the RMSE of soil moisture in both two layers increase correspondingly. Table 3 gives a summary of RREs for constrained simultaneous threeparameter estimation experiments with different assimilation intervals. It can be seen
- that the largest RRE is exhibited in the 10-days-interval experiment. When assimilation interval is too small, frequent corrections from soil moisture observations to model state may weaken the effects of proper parameter estimation in soil moisture forecasting. Whereas in the extremely sparse observation cases (≥20 days here), too little available information to update model state from soil moisture observations over a cer-
- tain period (e.g. one month) makes it different to overcome the accumulation of errors in soil moisture state in the process of model integration. Therefore, it can be concluded that proper assimilation frequency (once 10-days here) can display the greatest advantages of this constrained simultaneous multi-parameter estimation method in soil moisture assimilation using the EnKF. As conventional soil moisture in situ observa-
- tions in China are always with 10-days measurement interval (Nie et al., 2008), this method is supposed to be a good choice for assimilating these in situ observations. The application of this constrained simultaneous multi-parameter estimation method in assimilating temporal-sparse in situ soil moisture observations over China region will be studied in the future.





5 Conclusions

This study explores the applicability of the EnKF-based simultaneous state-parameter estimation in soil moisture data assimilation using a physical process land surface model by a series of idealized twin experiments. Uncertainties in model parameters,

initial soil moisture condition, and atmospheric forcing data are considered as primary sources of model errors. By the means of state augmentation, model-based pseudo near-surface soil moisture observations are assimilated to estimate model parameter and soil moisture state simultaneously. Three key hydraulic parameters: the saturated hydraulic conductivity, the saturated soil moisture suction and a soil texture empirical
 parameter, are subjected to estimation attempts in various experiments.

The estimation of single imperfect parameter is in general successful for all three estimated parameters. The ensemble mean value of each estimated parameters converge to its true value successfully. Moreover, with parameter errors been reduced in estimation process, the RMSE in estimation experiments are lower than that in non-

- ¹⁵ parameter-estimation benchmark experiments. In simultaneous multi-parameter estimation experiments, However, significant degrades can be seen for the estimation performance of all parameters. The estimated ensemble mean values of all three parameters can not converge to their true values despite soil moisture state can still be estimated successfully. The failure of estimation is ascribed to independent pertur-
- ²⁰ bations on different estimated parameters in assimilation processes, which cause the increase of degree of freedom of assimilation system and the breakage of inherent balance relationships between these parameters.

A strategy of considering constraints between estimated parameters in the filter is introduced to improve the performance of simultaneous multi-parameter estimation. The constraints used are obtained from the study of Cosby et al. (1984). The performance of the constraint-based multi-parameter estimation is successful, even if observations are available with temporal-sparse intervals such as 10-days or much longer. From the viewpoint of adding information, imposing these constraints between model parameters





also add new information to the assimilation system in addition to the available observations, which is the correlation between parameters. For this reason, the constraintbased estimation method effectively overcomes the negative impacts of non-closure problem in multi-parameter estimation and behaves successfully in soil moisture assimilation. As these constraints are always in statistical sense and can be obtained from literature or from standard parameter tables of the model, it is reasonable to apply this constraint-based parameter estimation method to other land surface models even with more imperfect parameters to be estimated.

5

Comparing to non-parameter-estimation assimilation, the constraint-based multiparameter estimation case can reduce much more RMSE in soil moisture state. Moreover, with proper temporal-sparse assimilation interval, this method has the best performance in improving soil moisture model state using the EnKF. Although obtained from idealized twin experiments, these results can still provide an instructive analysis of how to take the greatest advantage of this constraint-based multi-parameter estimation method in soil moisture assimilation. For its superiorities to non-parameter-estimation method in soil moisture assimilation, it is believed that this constraint-based simultane-

ous state-parameter estimation method might become a good choice for assimilating actual temporal-sparse in situ soil moisture observations over China in the future.

Acknowledgements. The authors wish to thank the National Meteorological Information Cen ter of China Meteorological Administration for the meteorological forcing data used in this study. This research was supported by the National Natural Science Foundation of China (Contract No. 40905046, 40437017 and 40221503), the Chinese Academy of Science (Contract No. KZCX2-YW-202), the Key Technologies R&D Program of China under Grant No. 2009BAC51B0x, the National High Technology Research and Development Program of China (863 Program: 2009AA1220005), and the National Basic Research Program of China (973 Program: 2010CB951902).





References

15

- Aksoy, A., Zhang, F., and Nielsen-Gammon, J. W.: Ensemble-based simultaneous state and parameter estimation in a two-dimensional sea-breeze model, Mon. Weather Rev., 134, 2951–2970, 2006.
- 5 Anderson, J. L.: An ensemble adjustment Kalman filter for data assimilation, Mon. Weather Rev., 129, 2884–2903, 2001.
 - Annan, J. D., Hargreaves, J. C., Edwards, N. R., and Marsh, R.: Parameter estimation in an intermediate complexity earth system model using an ensemble Kalman filter, Ocean Modell., 8, 135–154, 2005.
- ¹⁰ Burgers, G., van Leeuwen, P. J., and Evensen, G.: Analysis scheme in the ensemble Kalman filter, Mon. Weather Rev., 126, 1719–1724, 1998.
 - Chen, M., Liu, S., Tieszen, L. L., and Hollinger, D. Y.: An improved state-parameter analysis of ecosystem models using data assimilation, Ecol. Model., 219, 317–326, 2008.
 - Clapp, R. B. and Hornberger, G. M.: Empirical equations for some soil hydraulic properties, Water Resour. Res., 14, 601–604, 1978.
 - Cosby, B. J., Hornberger, G. M., Clapp, R. B., and Ginn, T. R.: A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils, Water Resour. Res., 20(6), 682–690, 1984.

Coudert, B., Ottlé, C., Boudevillain, B., Demarty, J., and Guillevic, P.: Contribution of thermal

- ²⁰ infrared remote sensing data in multiobjective calibration of a dual-source SVAT model, J. Hydrometeorol., 7, 404–420, 2006.
 - Crow, W. T. and van den Berg, M. J.: An improved approach for estimating observation and model error parameters in soil moisture data assimilation, Water Resour. Res., 46, W12519, doi:10.1029/2010WR009402, 2010.
- ²⁵ Crow, W. T. and Van Loon, E.: Impact of incorrect model error assumptions on the sequential assimilation of remotely sensed surface soil moisture, J. Hydrometeorol., 7, 421–432, 2006. Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res., 99(C5), 10143–10162, 1994.
- ³⁰ Evensen, G.: The ensemble Kalman filter: Theoretical formulation and practical implementation, Ocean Dynam., 53, 343–367, 2003.





Discussion Paper S. Nie et al. **Title Page** Abstract Introduction Discussion Conclusions References **Figures Tables** Paper 14 Back Close **Discussion** Paper Full Screen / Esc **Printer-friendly Version** Interactive Discussion

HESSD

8, 1433–1468, 2011

Land surface scheme

states and

parameters using the

Kalman filter

Discussion

Paper

Fujiyoshi, Y., Ding, Y. H., and Zhang, Y.: Outline of GAME/HUBEX, in: Final Report of GAME/HUBEX, edited by: Fujiyoshi, Y. and Ding, Y. H., GAME/HUBEX Project Office, Beijing, 1–6, 610 pp., 2006.

Houser, P. R., Shuttleworth, W. J., Famiglietti, J. S., Gupta, H. V., Syed, K. H., and Goodrich, D.

- C.: Integration of soil moisture remote sensing and hydrologic modeling using data assimila-5 tion, Water Resour. Res., 34, 3405-3420, 1998.
 - Ji, J. J.: A climate-vegetation interaction model: Simulating physical and biological processes at the surface, J. Biogeogr., 22, 445-451, 1995.
- Ji, J. J. and Hu, Y.: A simple land surface process model for use in climate studies, Acta Meteorol. Sin., 3, 342–351, 1989. 10
 - Jung, Y., Xue, M., and Zhang, G.: Simultaneous estimation of microphysical parameters and the atmospheric state using simulated polarimetric radar data and an ensemble Kalman filter in the presence of an observation operator error. Mon. Weather Rev., 138, 539-562. doi:10.1175/2009MWR2748.1.2010.
- Kalman, R. and Bucy, R.: New results in linear prediction and filtering theory, T. AMSE J. Basic 15 Eng., 83D, 95-108, 1961.
 - Margulis, S. A., McLaughlin, D., Entekhabi, D., and Dunne, S.: Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment, Water Resour. Res., 38, 1299, doi:10.1029/2001WR001114, 2002.
- Montaldo, N., Albertson, J. D., and Mancini, M.: Dynamic calibration with an ensemble Kalman 20 filter based data assimilation approach for root-zone moisture predictions, J. Hydrometeorol., 8, 910–921, 2007.
 - Moradkhani, H., Hsu, K., Gupta, H. V., and Sorooshian, S.: Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, Water Resour. Res., 41, W05012, doi:10.1029/2004WR003604, 2005.
 - Nie, S., Luo, Y., and Zhu, J.: Trends and scales of observed soil moisture variations in China, Adv. Atmos. Sci., 25(1), 43–58, 2008.

25

Ni-Meister, W., Houser, P. R., and Walker, J. P.: Soil moisture initialization for climate prediction: Assimilation of scanning multifrequency microwave radiometer soil moisture data into a land

1454

surface model, J. Geophys. Res., 111, D20102, doi:10.1029/2006JD007190, 2006. 30 Niyogi, D. S., Xue, Y., and Raman, S.: Hydrological land surface response in a tropical regime and a midlatitudinal regime, J. Hydrometeorol., 3, 39-56, 2002.

- Reichle, R. H., Entekhabi, D., and McLaughlin, D. B.: Downscaling of radio brightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach, Wate Resour. Res., 37, 2353–2364, 2001a.
- Reichle, R. H., McLaughlin, D. B., and Entekhabi, D.: Variational data assimilation of microwave radiobrightness observations for land surface hydrologic applications, IEEE T. Geosci. Remote,, 39, 1708–1718, 2001b.

Rawls, W. J., Brakensiek, D. L., and Saxton, K. E.: Estimation of soil water properties, T. Am.

- Reichle, R. H., McLaughlin, D. B., and Entekhabi, D.: Hydrologic data assimilation with the ensemble Kalman filter, Mon. Weather Rev., 130, 103–114, 2002a.
 - Reichle, R. H., Walker, J. P., Koster, R. D., and Houser, P. R.: Extended versus ensemble Kalman filtering for land data assimilation, J. Hydrometeorol., 3, 728–740, 2002b.
 - Reichle, R. H. and Koster, R. D.: Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model, Geophys. Res. Lett., 32, L02404, doi:10.1029/2004GL021700. 2005.
- ¹⁵ doi:10.1029/2004GL021700, 2005.

Soc. Agr. Eng., 25, 1316–1320, 1982.

5

25

Reichle, R. H., Crow, W. T., and Keppenne, C. L.: An adaptive ensemble Kalman filter for soil moisture data assimilation, Water Resour. Res., 44, W03423, doi:10.1029/2007WR006357, 2008.

Robock, A., Vinnikov, K. Y., Srinivasan, G., Entin, J. K., Hollinger, S. E., Speranskaya, N. A.,

- Liu, S., and Namkhai, A.: The global soil moisture data bank, B. Am. Meteorol. Soc., 81, 1281–1299, 2000.
 - Schaap, M. G. and Leij, F. J.: Improved prediction of unsaturated hydraulic conductivity with the Mualem-van Genuchten model, Soil Sci. Soc. Am. J., 64, 843–851, 2000.
 - van Genuchten, M. T.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, Soil Sci. Soc. Am. J., 44, 892–898, 1980.
 - Vrugt, J. A., Diks, C. G. H., Gupta, H. V., Bouten, W., and Verstraten, J. M.: Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, Water Resour. Res., 41, W01017, doi:10.1029/2004WR003059, 2005.
 - Walker, J. P. and Houser, P. R.: A methodology for initializing soil moisture in a global cli-
- ³⁰ mate model: Assimilation of near surface soil moisture observations, J. Geophys. Res., 106, 11761–11774, 2001.
 - Wen, L., Gallichand, J., Viau, A. A., Delage, Y., and Benoit, R.: Calibration of the CLASS model and its improvement under agricultural conditions, T. ASAE, 41, 1345–1351, 1998.





Xia, Y., Pitman, A. J., Gupta, H. V., Leplastrier, M., Henderson-Sellers, A., and Bastidas, L. A.: Calibrating a land surface model of varying complexity using multicriteria methods and the Cabauw dataset, J. Hydrometeorol., 3, 181–194, 2002.

Zhang, S., Zeng, X., Zhang, W., and Barlage, M.: Revising the ensemble-based Kalman filter covariance for the retrieval of deep-layer soil moisture, J. Hydrometeorol., 11, 219-227, 2010.

5

Zhuang, J., Nakayama, K., Yu, G. R., and Miyazaki, T.: Predicting unsaturated hydraulic conductivity of soil based on some basic soil properties, Soil Till. Res., 59, 143-154, 2001.

Discussion Pa	HE 8, 1433–1	HESSD 8, 1433–1468, 2011				
per Discussion	Land surface scheme states and parameters using the Kalman filter S. Nie et al.					
Paper	Title	Title Page				
—	Abstract	Introduction				
Discussic	Conclusions Tables	References Figures				
n Paper	I4	×I				
	•	•				
	Back	Close				
iscussion Paper	Full Scru Printer-frier Interactive	Full Screen / Esc Printer-friendly Version Interactive Discussion				
Paper	Interactive Discussion					



Table 1. Specific differences of soil hydraulic parameters, initial soil moisture condition, and meteorological forcing data between "true" and "prior" model integrations in the idealized twin experiments.

Variables and Parameters	Units	True	Prior		
Saturated hydraulic conductivity	ms ⁻¹	$5.23 e^{-6}$	5.0 e ⁻⁵		
Empirical parameter b	-	4.74	12.0		
Saturated soil moisture suction	m	-0.218	-0.7		
Initial soil moisture	cm ³ cm ⁻³	2-yr spinup values	0.12 for both two layers		
Precipitation	mm (day) ⁻¹	Gauge-based data and NCEP dataset 1	Adding Gaussian noise with mean square deviation of 20% to the true values once daily and the minimum mean square deviation is limited to $2 \text{ mm} (\text{day})^{-1}$		
Long- and short-waves rediations	W m ⁻²	NCEP dataset 1	Adding Gaussian noise with mean square deviation of 30% to the true values once daily		





Table 2. Summary of time average root mean squared error (RMSE) and relative root mean
squared error (RRE) of soil moisture in near-surface layer (SM1) and root zone layer (SM2) in
three single-parameter estimation experiments and corresponding non-parameter-estimation
experiments with individual imperfect parameters of the saturated hydraulic conductivity k_{sat} ,
the saturated soil moisture suction Ψ_{sat} , and a soil texture empirical parameter <i>b</i> respectively.

Imperfect parameter	Non-estimation RMSE cm ³ cm ⁻³		Estimation RMSE cm ³ cm ⁻³		RI	RRE	
	SM1	SM2	SM1	SM2	SM1	SM2	
<i>k</i> _{sat}	0.026	0.023	0.014	0.009	46%	61%	
Ψ_{sat}	0.025	0.013	0.014	0.009	44%	31%	
b	0.053	0.053	0.015	0.018	72%	66%	







Table 3. Summary of relative root mean squared error (RRE) of soil moisture in both two layers in constrained simultaneous three-parameter estimation experiments with assimilation intervals of 1-day, 10-days, 20-days, 30-days and 40-days respectively.

Assimilation intervals	1-days	10-days	20-days	30-days	40-days
RRE in surface layer	44%	53%	47%	43%	36%
RRE in root zone layer	25%	50%	44%	39%	32%



Fig. 1. The **(a)** precipitation forcing and **(b)** soil moisture used in idealized twin experiment during the year 1998 in Jiangji station. For soil moisture in near-surface layer and root zone layer, blue and red solid lines are for "true" states as well as green and orange dash lines are for "prior" states. Gray solid line represents "actual observation" of near-surface soil moisture.















Fig. 3. The time evolution of the root mean squared error (RMSE) of near-surface layer soil moisture (a1), (a2), (a3), and root zone layer soil moisture (b1), (b2), (b3) of one-day-ahead soil moisture forecasting from single-parameter estimation results (solid black lines) compared with that of non-parameter-estimation benchmark experiments (solid gray lines). Parameters shown are (a1), (b1) the saturated hydraulic conductivity; (a2), (b2) the saturated soil moisture suction; (a3), (b3) a soil texture empirical parameter *b*.











Printer-friendly Version

Interactive Discussion



























Fig. 8. Same as in Fig. 3 but for the time evolution of RMSE of near-surface layer soil moisture **(a)** and root zone layer soil moisture **(b)** from the constrained (solid black lines), non-constrained (solid gray lines) simultaneous three-parameter estimation experiments, and non-parameter-estimation benchmark experiment (short-dashed gray lines) with 10-days assimilation interval.







Fig. 9. The time evolution of the RMSE for **(a)** near-surface layer and **(b)** root zone layer soil moisture in constrained simultaneous three-parameter estimation experiments with temporal-sparse assimilation intervals of 10-days (black lines), 20-days (gray lines), 30-days (red lines) and 40-days (blue lines) respectively.



