Joint State and Parameter Estimation of Two Land Surface Models Using the Ensemble Kalman Filter and Particle Filter

Hongjuan Zhang[1,2], Harrie-Jan Hendricks-Franssen[1,2], Xujun Han[1,2], Jasper Vrugt[3,4] and Harry Vereecken[1,2]

[1]Forschungszentrum Jülich, Agrosphere (IBG 3), Jülich, Germany
[2]Centre for High-Performance Scientific Computing in Terrestrial Systems: HPSC TerrSys, Forschungszentrum Jülich, Jülich, Germany
[3]Department of Civil and Environmental Engineering, University of California, Irvine, USA
[4]Department of Earth Systems Science, University of California Irvine, USA

Correspondence to: Hongjuan Zhang (ho.zhang@fz-juelich.de)

Abstract. Land surface models (LSMs) contain a suite of different parameters and state variables to resolve the water and energy balance at the soil-atmosphere interface. Many of the parameters of these models cannot be measured directly in the field, and require calibration against flux and soil moisture data. In this paper, we use the Variable Infiltration Capacity Hydrologic Model (VIC) and the Community Land Model (CLM) to simulate temporal variations in soil moisture content at 5, 20 and 50 cm depth in the Rollesbroich experimental watershed in Germany. Four different data assimilation (DA) methods are used to jointly estimate the spatially distributed water content values, and hydraulic and/or thermal properties of the resolved soil domain. This includes the Ensemble Kalman Filter (EnKF) using state augmentation or dual estimation, the Residual Resampling Particle Filter (RRPF) and Markov chain Monte Carlo Particle Filter (MCMCPF). These four DA methods are tuned and calibrated for a five month data period, and subsequently evaluated for another five month period. Our results show that all the different DA methods improve the fit of the VIC and CLM model to the observed water content data, particularly if the maximum baseflow velocity (VIC), soil hydraulic (VIC) properties and/or soil texture (CLM) are jointly estimated along with the model states. In the evaluation period, the augmentation and dual estimation method performed slightly better than RRPF and MCMCPF. The differences in simulated soil moisture values between the CLM and VIC model were larger than variations among the data assimilation algorithms. The best performance for the Rollesbroich site was observed for the CLM model. The strong underestimation of the soil moisture values of the third VIC-layer are likely explained by an inadequate parameterization of groundwater drainage.

1 Introduction

Land surface models use a suite of different parameters to characterize adequately a myriad of different fluxes and state variables that determine the water and energy status of the land surface. Generally, water balance involves water processes from soil (evaporation, infiltration, surface runoff, etc.), canopy (interception, evapotranspiration, etc.), aquifer (discharge and recharge of groundwater) and atmosphere (precipitation) (Schaake, et al., 1996); energy balance includes latent and sensible heat fluxes from soil, snow, surface water and vegetated surface (Bertoldi, 2004). All these processes are characterized by parameters which are based on global or regional distributions of vegetation and soil properties (Milly and Shmakin, 2002). These parameters differ from one model to the next, however all land surface models need soil hydraulic parameters (e.g. saturated

hydraulic conductivity) to describe water process in soil, vegetation parameters (e.g. root profile) to calculate evaporation, soil thermal parameters (e.g. saturated thermal conductivity) to solve soil temperature, and surface albedo to estimate reflected shortwave radiation. Different models control these parameters in different ways. Some models estimate soil hydraulic and thermal parameters from soil texture on the basis of pedotransfer

5   functions. An example is the Community Land Model (CLM) (Vereecken et al., 2008; Oleson et al., 2013; Han et al., 2014). Other models require as input values for the hydraulic and thermal parameters. An example is the Variable Infiltration Capacity Model (VIC) (Liang et al., 1994; Gao et al., 2010).

At many locations, the information of soil properties (soil texture, saturated hydraulic conductivity or porosity) is not available or not accurate. Another important source of uncertainty for calculations with LSMs are the

10   meteorological input data, even if data from locally available measurements are used. Predictions with LSMs are strongly affected by the large uncertainty of model parameters and forcings (Kitanidis and Bras, 1980). Data assimilation provides a way to take advantage of all available ground-based, airborne or spaceborne observations to improve the compliance between numerical models and corresponding data. This approach allows for joint estimation of the states and parameters while taking into explicit consideration model structural error and forcing

15   data errors (Liu and Gupta, 2007). Several published studies have shown the merits of parameter estimation in the context of data assimilation involving soil moisture characterization (e.g., Montzka et al., 2011), rainfall-runoff modeling (e.g., Moradkhani et al., 2005a; Vrugt et al., 2005) and land surface modeling (e.g., Pauwels et al., 2009).

All data assimilation methods merge observations and models yet the degree of sophistication varies widely.

20   Much previous work has appeared on the topic of joint parameter-state estimation in the hydrologic/land-surface literature. The majority of these contributions involves assimilation of synthetic observations including (among others) groundwater table depth or piezometric head (Franssen and Kinzelbach, 2008; Bailey and Bau, 2012; Kurtz et al., 2014; Shi et al., 2014; Song et al., 2014; Tang et al., 2015), discharge (Rasmussen et al., 2015), groundwater temperature (Kurtz et al., 2014), soil moisture (Wu and Margulis, 2011; Plaza et al., 2012; Erdal et

25   al., 2014; Shi et al., 2014; Song et al., 2014; Pasetto et al., 2015), streamflow (Bailey and Bau, 2012; Moradkhani et al., 2012; Vrugt et al., 2013), brightness temperature from passive remote sensing (Montzka et al., 2011; Montzka et al., 2013; Han et al., 2014), and contaminant concentration (Gharamti et al., 2013). These published papers include use of the Particle Filter (PF) (Montzka et al., 2011; Plaza et al., 2012; Montzka et al., 2013), Markov Chain Monte Carlo Particle Filter (MCMCPF) (Moradkhani et al., 2012; Vrugt et al., 2013),

30   Ensemble Kalman Filter (EnKF) (Franssen and Kinzelbach, 2008; Wu et al., 2011; Gharamti et al., 2013; Erdal et al., 2014; Kurtz et al., 2014; Shi et al., 2014; Pasetto et al., 2015), iterative EnKF (Song et al., 2014), Extended Kalman Filter (Pauwels et al., 2009), Local Ensemble Transform Kalman Filter (LETKF) (Han et al., 2014), Ensemble Transform Kalman Filter (ETKF) (Rasmussen et al., 2015), and Normal Score Ensemble Kalman Filter (NS-EnKF) (Tang et al., 2015). General conclusion of these papers is that joint parameter and state

35   estimation by data assimilation significantly enhances the ability of the model to mimic the observed data, yet the findings of these papers might not necessarily apply to real-world data involving significant errors in the model structure, input and calibration data.

Some previous work also applied joint parameter-state estimation with real-world data. These works considered the assimilation of electrical conductivity data (Wu and Margulis, 2013), piezometric head data from wells

40   (Kurtz et al., 2014; Shi et al., 2015), groundwater temperature data (Kurtz et al., 2014), streamflow

measurements (Moradkhani et al., 2012), discharge measurements (Shi et al., 2015), active remote sensing data (Pauwels et al., 2009), passive brightness temperature information (Qin et al., 2009), soil moisture observations from lysimeter (Lue et al., 2011; Wu and Margulis, 2013; Erdal et al., 2014; Shi et al., 2015), land surface temperature observations (Bateni and Entekhabi, 2012) and sensible and latent heat fluxes (Shi et al., 2015). The

5 methods used were PF (Qin et al., 2009), MCMCPF (Moradkhani et al., 2012), EnKF (Bateni and Entekhabi, 2012; Wu and Margulis, 2013; Erdal et al., 2014; Kurtz et al., 2014; Shi et al., 2015) and Extended Kalman Filter (Pauwels et al., 2009; Lue et al., 2011). These papers also concluded that joint parameter and state estimation worked well in real-world cases. However, this overview indicates that few real-world applications involved the evaluation of soil moisture content in the context of joint state-parameter estimation with land

10 surface models (Lue et al., 2011; Shi et al., 2015), even although soil moisture plays a critical role in the partitioning of energy and water fluxes at the land surface.

This paper focuses therefore on the evaluation of joint state-parameter estimation in the context of soil moisture characterization with land surface models. The comparison in this paper includes four sequential data assimilation algorithms in combination with two different land surface models. The four data assimilation

15 algorithms which are compared are variants of the commonly used data assimilation algorithms Ensemble Kalman filter (EnKF) and particle filter (PF). For EnKF the state augmentation approach (Chen and Zhang, 2006) and the dual estimation approach (Moradkhani et al., 2005a) are compared. In the state augmentation approach, the state vector is augmented by parameters and then states and parameters are jointly updated over time. In the dual estimation approach, states and parameters are stored in two separate vectors. Parameters are

20 updated first and then the updated parameters are used to update states. PF updates states and parameters simultaneously, as states and parameters are jointly related to a certain particle with specific weight (Moradkhani et al., 2005b). The PF used in this study was the Residual Resampling Particle Filter (RRPF) (Douc et al., 2005) and Markov Chain Monte Carlo Particle Filter (MCMCPF) which alleviates the particle degeneration by adding a move step on particles after resampling to generate proposal particles (Moradkhani et al., 2012; Vrugt et al.,

25 2013). A Metropolis ratio is then calculated to decide whether the proposal particle is accepted or not. Relatively few papers (Dechant and Moradkhani, 2012; Dumedah and Coulibaly, 2013; Chen et al., 2015) compared sequential data assimilation algorithms for joint state-parameter estimation problems. Only Chen et al. (2015) made a comparison of the data assimilation algorithms for a LSM, the other two papers were concerned with rainfall-runoff modeling.

30 The main objectives of this study are as follows: (1) to test and evaluate the merits of joint parameter and state estimation for LSMs using real-world data; (2) to compare the performance of the four commonly used data assimilation methods in their ability to characterize adequately the soil moisture profiles of the experimental site; (3) to compare the simulation results of the CLM and VIC model and explain the differences in performance of these models.

35 The remainder of this paper is organized as follows. In section 2, we briefly review the VIC and CLM models used herein to simulate the soil moisture dynamics of the Rollesbroich experiment site. In this section we are especially concerned with parameter selection, and a description of the experimental site and data. Section 3 then introduces the basic concepts of the four different data assimilation algorithms used herein. This is followed in section 4 with a detailed explanation of the numerical setup of each data assimilation method and results of our

experiment. Section 5 discusses the main findings of our assimilation studies. Finally, this paper concludes in section 6 with a summary of our main findings.

2 Land Surface Models

We now discuss the two different land surface modeling schemes (models) used herein. The appendix provides
5   further details on each of the models.

2.1 Variable Infiltration Capacity Model (VIC)

The VIC model is a semi-distributed macro-scale hydrological model and takes account of vegetation variations within a grid cell. Accordingly, each grid cell is divided into land cover tiles (Liang et al., 1994; Liang et al., 1996; Cherkauer and Lettenmaier, 1999). On the other hand, soil properties (e.g., soil texture, hydraulic
10   conductivity, thermal conductivity) are held constant within each grid cell. VIC considers both the water and energy balance for the grid cell. For each grid cell, the total evapotranspiration, sensible heat flux, effective land surface temperature and runoff are obtained by summing over all the land cover tiles (vegetation types and bare soil) weighted by the fractional coverage (Gao et al., 2010). The VIC model can either be run in a water balance mode or a water-and-energy balance mode. In this paper, the water-and-energy balance mode was used.

15   In this study, VIC-3L was used, which is a three layer version of the VIC model. The soil column has a very thin surface layer (first layer), an upper layer (second layer) and a lower layer (third layer). The surface layer captures rapid dynamics related to rainfall events and bare soil evaporation. The upper layer is strongly influenced by the response to rainfall. The lower layer is affected by seasonal dynamics of deep soil moisture and base flow. In this study, the thicknesses of the 3 layers are 10cm, 20cm and 40cm respectively.

20   VIC-3L requests as input meteorological data (precipitation, wind speed, air temperature, longwave/shortwave radiation, relative humidity), soil properties like soil bulk densities and soil hydraulic parameters (saturated hydrologic conductivity k_s, residual water content of a soil layer, parameters for the soil-water characteristic curve, and parameters for the baseflow). Further model inputs to VIC-3L are the vegetation types and their characteristics, and the fractions of the different vegetation types in each grid cell. More details about the
25   parameterization are presented in Appendix A.

2.2 Community Land Model (CLM)

CLM is the land model for the Community Earth System Model (CESM) (Oleson et al., 2013). It includes the hydrological cycle, biogeochemical cycles, biogeophysics and dynamic vegetation. Unlike the VIC-3L model, a grid cell in CLM has multiple subgrid levels. The first subgrid level is defined by land units (vegetated, lake, urban, glacier, and crop), and each land unit has a number of columns (second subgrid level). For the vegetated
30   land unit, as well as for lakes and glaciers, there is one column; for the urban land use, there are five columns; for crop land, there is a distinction between irrigated and unirrigated columns with one single crop occupying one column. The third subgrid level is the Plant Functional Type (PFT) level, including bare soil. The vegetated column has 16 possible PFTs besides bare soil. For the crop column, several crop types are available. Processes like canopy evaporation and transpiration are calculated for each available PFT. Processes related to soil or snow
35   are calculated for each column, which requires PFT level properties to be aggregated to the column level. The

aggregation is computed by a weighted sum of the desired quantities over all PFTs whose weights depend on the PFT area relative to the complete area. This aggregation in CLM is the same as for VIC-3L.

Soil temperature is calculated for 15 soil layers, while hydrology is calculated for the top 10 soil layers. CLM input includes atmospheric forcing data, land surface data including information on PFTs, and adjustable

5   parameters and physical constants. CLM uses soil properties like soil texture and organic matter density in combination with model internal pedotransfer functions to derive soil hydraulic and thermal parameters like saturated hydraulic conductivity. More details about the parameterization are presented in Appendix B.

2.3 Differences between VIC-3L and CLM

VIC-3L and CLM show a number of important differences concerning their calculations of the water and energy

10   balances:

(1) The two models use a different approach for solving flow in the unsaturated zone. CLM uses a modified Richards' equation, which includes coupling with an unconfined aquifer. VIC-3L uses a bucket type approach which takes into account the variable infiltration capacity.

(2) In VIC-3L, the unsaturated and saturated zones are treated in a lumped sense and the impact of groundwater

15   is not taken into account. In CLM, the interaction between an unconfined aquifer and the unsaturated soil column is considered. Changes in water table depth are calculated and included as boundary condition for solving flow in the unsaturated zone.

(3) Soil hydraulic parameters like saturated hydraulic conductivity, parameters used to calculate baseflow and soil thermal information like average soil temperature (and other parameters) are the direct input information in

20   VIC-3L. On the contrary, hydraulic conductivity, saturated soil matric potential, the Clapp-Hornberger exponent B and soil thermal conductivity are calculated by model internal pedotransfer functions, using soil texture and soil organic matter density as input information in CLM.

(4) The depths of the three soil layers are user-defined in VIC-3L, while in CLM, the thicknesses of the 15 soil layers are internally defined. All the calculations are based on these thicknesses.

25   ### 2.4 Selection of parameters to be updated

The sensitivity of land surface parameters of VIC-3L was investigated in the past by other authors using Monte Carlo Analysis, Generalized Likelihood Uncertainty Assessment (GLUE), or different calibration approaches (Demaria et al., 2007; Xie et al., 2007; Troy et al., 2008). The results revealed that parameter sensitivity was dependent on climate. For CLM, only sand fraction, clay fraction, and organic matter density are direct input

30   data, and soil hydraulic and thermal parameters are calculated by pedotransfer functions which are hard coded in CLM (Oleson et al., 2013; Han et al., 2014). Table 1 shows the parameters chosen to be updated during the assimilation period in our work for both the VIC model and CLM. The definition of these parameters can be seen in Appendix A and B.

3. Assimilation Algorithms

35   Data assimilation algorithms combine observations and model predictions together and update model states and parameters. Commonly used data assimilation algorithms are EnKF, PF and variants of them.

3.1 EnKF

EnKF was proposed by Evensen (1994) and follows a Monte Carlo approach to generate stochastic realizations for estimating the forecast-error statistics.

The stochastic EnKF scheme includes the following steps (Burgers et al., 1998):

$$x_t^{i-} = f\left(x_{t-1}^i, p_{t-1}^i, u_t^i\right) + v_t \tag{21}$$

where i refers to the ith ensemble member (i = 1, ..., N), f to a simulation model (in our case the VIC-3L model or CLM), t to the time step, x_t^{i-} to the predicted state vector at time t (in our case soil moisture), p to the parameter vector, u to the forcing data, and v_{t-1} to model error at time step t. From the ensemble of state vectors at time t, the background error covariance matrix C is obtained according to:

$$C = \frac{1}{N-1}\sum_{i=1}^{N}(x_t^{i-} - \overline{x_t})(x_t^{i-} - \overline{x_t}) \tag{22}$$

where N is the number of ensemble members, and $\overline{x_t}$ indicates the ensemble mean at time step t. The observation equation is given by:

$$y_t^i = y_t + w_t^i \tag{23}$$

where y is the vector with observations and w is the observation error, which is generated from a normal distribution N(0, σ) and σ is the expected measurement standard deviation. The ensemble members of state vectors are updated with the help of observations according to:

$$x_t^i = x_t^{i-} + K(y_t^i - Hx_t^{i-}) \tag{24}$$

where x_t^i is the updated state vector, and H is an observation operator that connects measurements and model states, it is the identity matrix if y refers to in-situ ground measurements available at all grid cells. K is Kalman gain and R is the observation error covariance matrix calculated by:

$$R = \frac{1}{N-1}\sum_{i=1}^{N}(y_t^i - \overline{y_t})(y_t^i - \overline{y_t}) \tag{25}$$

where $\overline{y_t}$ is the average over the perturbed observations. However, R is usually defined a priori on the basis of expected measurement errors. Finally, the Kalman gain K is calculated by:

$$K = CH^T(HCH^T + R)^{-1} \tag{26}$$

3.1a EnKF with state augmentation

There are two often applied approaches for joint estimation of states and parameters in EnKF: state augmentation and dual estimation. In the state augmentation approach, the state and parameter vector are combined into a single joint state vector (Franssen and Kinzelbach, 2008), and the states and parameters are estimated simultaneously.

In state augmentation, the state vector x, the model error covariance matrix C, the measurement operator H, and the Kalman gain K consist of two blocks:

$$x^i = \begin{bmatrix} s^i \\ p^i \end{bmatrix} \tag{27}$$

$$C = \begin{bmatrix} C_{ss} & C_{ps}^T \\ C_{ps} & C_{pp} \end{bmatrix} \tag{28}$$

$$H^* = \begin{bmatrix} H_s, H_p \end{bmatrix} \tag{29}$$

where s refers to model states and p to parameters. The model error covariance matrix C now includes four parts: C_{ss}, C_{ps}^T, C_{ps}, and C_{pp}. The measurement operator H is also augmented to H^* which includes Hs and Hp. The Kalman gain K is now given by:

$$K = CH^{*T}(H^*CH^{*T} + R)^{-1} = \begin{bmatrix} K_s \\ K_p \end{bmatrix} \qquad (30)$$

5   The updating Eq. (24) is now given by:

$$\begin{bmatrix} s_t^i \\ p^i \end{bmatrix} = \begin{bmatrix} s_t^{i-} \\ p^{i-} \end{bmatrix} + \begin{bmatrix} K_s \\ K_p \end{bmatrix} [y_t^i - Hs_t^{i-}] \qquad (31)$$

3.1b EnKF with dual estimation

In the dual estimation approach, states and parameters are stored in two vectors which are modified in two separate operations (Moradkhani et al., 2005a). The parameter ensemble members are updated in a first step

10   according to:

$$p_t^i = p_t^{i-} + K_p(y_t^i - Hs_t^{i-}) \qquad (32)$$

Next, the updated parameters are used to update the ensemble of model state predictions according to Eq. (21) and (24). The model has to be run twice for the dual estimation approach and therefore the CPU-time approximately doubles compared to the state augmentation approach.

15   A problem associated with EnKF is the filter inbreeding where the underestimation of ensemble variance becomes more severely after several data assimilation cycles. In extreme cases, the model ensemble variance is so small that the weights for the measurements are close to zero and observations are not able to correct the ensemble anymore. Filter inbreeding is aggravated by a low number of ensemble members which results in spurious correlations among state variables/parameters, and reduces the ensemble variance artificially. Another

20   reason for the underestimation of ensemble spread could be a too small prior uncertainty for parameters and/or model forcings, or an important model structural error. Ensemble inflation methods are an effective way to ameliorate the filter inbreeding (Anderson, 2007; Whitaker and Hamill, 2012). In our work, the inflation algorithm proposed by Whitaker and Hamill (2012) was applied to the ensemble of parameter values and the ensemble of each parameter increased or decreased its variance according to:

25    $$p_t^i = \overline{p_t} + \left(p_t^i - \overline{p_t}\right)\left(1 + \frac{\sigma_b - \sigma_a}{\sigma_a}\right) \qquad (33)$$

where $\overline{p_t}$ is the ensemble mean for a parameter pt at time step t, σ_b is the posterior ensemble standard deviation of the parameter and σ_a is the prior ensemble standard deviation. This method artificially keeps the ensemble standard deviation of parameters equal to the initial standard deviation for the parameters. This method is especially important for applications with small ensemble sizes.

30   **3.2 Residual Resampling Particle Filter (RRPF) with parameter resampling**

The particle filter was first suggested in the research area of object recognition, robotics and target tracking (Arulampalam et al., 2002). It was introduced in hydrology by Moradkhani et al. (2005a). PF solves the Bayesian recursion equations directly by using an ensemble based approach and a set of particles to represent the

samples from the probability density function (PDF). Each particle has a weight assigned to it that represents the probability of that particle being sampled from the PDF. The state-space model can be non-linear and the initial state and noise distributions can take any arbitrary PDF.

The posterior PDF at time t given the observations y_t is approximated by the PF according to:

$$p(x_t^i|y_t) \approx \sum_{i=1}^{N} w_t^i \, \delta(x_t - x_t^i) \tag{34}$$

where x_t^i is assumed to be the ith state sample (in our case soil moisture) drawn from the posterior pdf $p(x_t|y_t)$ with weight w_t^i and δ is the Dirac delta function. However, as it is impossible to sample from the true posterior PDF, a proposal distribution ($q(x_t^i|y_t)$) is an alternative. The weight for a particle i is calculated according to:

$$w_t^i \square \frac{p(x_t^i|y_t)}{q(x_t^i|y_t)} \tag{35}$$

For the sequential updating case, the recursive weight update equation is defined:

$$w_t^i = w_{t-1}^i \frac{p(y_t|x_t^i) p(x_t^i|x_{t-1}^i)}{q(x_t^i|x_{t-1}^i, y_t)} \tag{36}$$

The state estimated from the N particles is given by:

$$x_t = \sum_{i=1}^{N} w_t^i \, x_t^i \tag{37}$$

Particles tend to degenerate (particle degeneration (Carpenter et al., 1999)), especially for higher dimensional problems, which means that the weights become nearly zero for most particles and only a few particles receive a weight significantly larger than zero. The effective sample size N_{eff} is calculated after each updating step to detect particle degeneration:

$$N_{eff} = \frac{1}{\sum_{i=1}^{N} (w_t^i)^2} \tag{38}$$

If the effective sample size is less than a pre-defined threshold (typically N/2), this is considered particle degeneration.

To avoid a small effective sample size, resampling is necessary for the PF. Gordon et al. (1993) introduced the Sequential Importance Resampling (SIR). In SIR, N particles are drawn from the current particle set with probabilities proportional to their weights. The N samples receive now all a weight equal to 1/N. Other resampling algorithms have been suggested like Residual Resampling (RR) (Liu and Chen, 1998) which was used in our work. In RR, (a) $\hat{N}_i = [N w_t^i]$, and [ ] is the integer operator; (b) a SIR procedure is performed to select the remaining $N_j = N - \sum_{i=1}^{N} \hat{N}_i$ samples with new weights $w_t^i = (N w_t^i - \hat{N}_i)/N_j$. The variance of the particles is smaller than the variance given by SIR (Weerts and Serafy, 2006). The detailed schemes of SIR and RR are described in (Liu et al., 1998; Weerts and Serafy, 2006). When particles are resampled, the parameters generating the particles are also resampled by the vector containing the resample indices. Plaza et al. (2012) illustrated the importance of parameter resampling in PF by a series of data assimilation experiments.

The disadvantage of resampling is that the diversity of particles is reduced as particles tend to cluster in state space which is often a poor representation of the posterior distribution. The ensemble inflation methods mentioned above could also be implemented to solve particle degeneration (Qin et al., 2009). In our work, the

method described by Plaza et al. (2012) and Moradkhani et al. (2005b) was used, in which the resampled
parameter values were perturbed by white Gaussian noise to increase the particle spread. Plaza et al. (2012)
concluded that resampling of replicating particles with larger weights would negatively affect the assimilation
performance, and that perturbation of resampled parameters would relieve this problem. The applied method can

5   be summarized as follows:

IF $N_{eff} < N/2$

- Residual Resampling step

  Calculate the resampling index vector j

  $\hat{x}_t = x_t(j)$

  $\hat{p}_t = p_t(j)$

- Perturb the resampled parameters

10          $p_t^i = \hat{p}_t^i + \varepsilon_t^i$ $\varepsilon_t^i \sim N(0, s^2 \sigma_{prior}^2)$

- Assign weights

  $w_t^i = 1/N$

END IF

Where s is a small tuning parameter and σ_{prior}^2 is the prior variance for parameter p. s was 0.1 in our work.

3.3 Markov Chain Monte Carlo PF (MCMCPF)

15   To achieve a higher variability in particles and to avoid particle degeneration, Moradkhani et al. (2012) and
Vrugt et al. (2013) used Markov Chain Monte Carlo methods (MCMC). In MCMC methods, after RR, it
becomes necessary to add a move step, creating a proposal distribution. The proposal distribution allows for a
relatively large move which probably jumps far away from the probability mass of the posterior distribution. In
this work, the formulation by Vrugt et al. (2013) was used to generate proposal state particles and parameter sets.

20   Details of the methodology can be found in Vrugt et al. (2013).

The Metropolis acceptance ratio α is calculated to determine whether the proposed state-parameter combination
is accepted.

$$\alpha = \min(1, \frac{p\left(x_{i,t-1}^{pro}\middle|x_{t-2}^i\right)p\left(y_{t-1}^i\middle|x_{i,t-1}^{pro}, p_{i,t}^{pro}\right)p\left(y_t^i\middle|x_{i,t}^{pro}, p_{i,t}^{pro}\right)}{p\left(x_{t-1}^i\middle|x_{t-2}^i\right)p\left(y_{t-1}^i\middle|x_{t-1}^i, p_t^i\right)p(y_t^i|x_t^i, p_t^i)}) \tag{39}$$

where $x_{i,t}^{pro}$ is the ith proposed state sampled from the proposal state distribution at time step t, $p_{i,t}^{pro}$ is the ith

25   proposed parameter sampled from the proposal parameter distribution at time step t, and y_t^i represents the ith
observation at time step t. The proposed state-parameter combination is accepted if (α > U(0,1)) where U(0,1) is
an uniformly distributed random number. Through this acceptance/rejection step, the algorithm ensures
variability of particles in the posterior density. After a single iteration, the algorithm moves to the next time step.
More iterations will lead to better results, but increase the needed CPU-time because it resamples proposal

30   particles and repeats model runs. The MCMC step can be summarized as follows:

IF $N_{eff} < N/2$

Residual Resampling step

    Calculate the resampling index vector j

$\hat{x}_t = x_t(j)$

$\hat{p}_t = p_t(j)$

MCMC Resampling

Create proposal x_{t-1}^{pro} based on x_{t-1}

5 Create proposal p_t^{pro} based on $\hat{p}_t$

Simulate proposal x_t^{pro} based on proposal x_{t-1}^{pro} and proposal p_t^{pro} using model

Calculate the Metropolis ratio $\alpha(x_{i,t}^{pro}, \hat{x}_t^i)$

Calculate the accept index vector j

$\hat{x}_t(j) = x_t^{pro}(j)$ if proposal x_t^{pro} is accepted, $\hat{x}_t^i$ will be replaced by proposal $x_{i,t}^{pro}$

10 $\hat{p}_t = p_t^{pro}(j)$ if proposal p_t^{pro} is accepted, $\hat{p}_t^i$ will be replaced by proposal $p_{i,t}^{pro}$

Assign weights

$w_t^i = 1/N$

END IF

4. Case study

15 ### 4.1 Rollesbroich site

The Rollesbroich site (50°37'27"N, 6°18'17"E) is a grassland site and a subcatchment of the TERENO Rur catchment in Germany (Bogena et al., 2010; Qu et al., 2014). It is located in the Eifel hills and the dominant soil texture is silty loam. It covers an area of 27 ha with an altitude ranging between 474 and 518m.a.s.l. The mean annual air temperature is 7.7 °C, the mean annual precipitation is 1033mm, and the mean slope is 1.63°. At the

20 site an eddy covariance tower (50°37'19"N, 6°18'15"E, height 514.7m.a.s.l) and a soil moisture and soil temperature sensor network (with measurements at 5, 20 and 50cm depth) are installed, amongst others. Soil moisture time series at 41 locations are being recorded. Figure 1 shows the locations of the measurement devices.

In this work, the Rollesbroich site is modelled as a single point and the data of the soil sensor network are

25 averaged to calculate areal averages of soil moisture content at 5cm, 20cm and 50cm depth. The forcing data in this study (hourly air temperature, air pressure, relative humidity, wind speed, incoming shortwave and longwave radiation), were measured at the eddy covariance tower. Precipitation was measured by a tipping bucket located close to the eddy covariance station. Figure 2 shows the daily precipitation and daily air temperature for the years 2011 and 2012. Soil texture was determined for the area based on 273 soil samples,

30 taken from three different depths, ranging between 5 and 11 cm, 11 and 35 cm, and 35 to 65 cm. The sample locations coincided with the location of the SoilNet sensors. The soil textural composition, organic carbon content, and bulk density were determined using standard laboratory procedures. Other soil hydraulic parameters were estimated from these data with help of pedotransfer functions. Finally, for each of the three depth ranges average values were calculated.

4.2 Experiment Setup

VIC-3L and CLM were spun-up with measured meteorological data from January 1, 2011 to February 29, 2012 using an hourly time step. The assimilation period was from March 1, 2012 to July 31, 2012. Daily soil moisture observations were assimilated in the assimilation period to update model states and possibly also parameters. The verification period was from August 1, 2012 to December 31, 2012. In this period, models were not informed by observations, but used the updated parameter values as input.

Soil moisture contents measured at 5cm, 20cm and 50cm depth were assimilated jointly. The definition of the model layers in VIC-3L was in correspondence with these data, the three layers extended from 0cm to 10cm, 10cm to 30cm and 30cm to 70cm. Parameters were also defined for the three layers. In CLM, the 10 predefined soil layers were involved in the hydrological calculations. Soil moisture content measurements at 5cm, 20cm and 50cm corresponded to the third, fifth and the sixth model layer in CLM. The parameters of the other layers were updated with help of the calculated spatial covariances in case of EnKF.

Figure 2 shows that the year 2012 had abundant rainfall, with some intensive precipitation events in the summer like the one on the 27th of July 2012 with 31mm precipitation in one hour. From our experience, if the rainfall intensity is too high, the parameter estimation is negatively affected. This is probably related to surface runoff which is not handled well by the model, and the reduced state-parameter correlation for very high soil moisture contents. Therefore, if the cumulative daily rainfall was more than 20mm no parameter updating was made for that day and the two next days. For those days, only states were updated.

In order to evaluate joint state-parameter estimation algorithms for the two land surface models and the four different data assimilation algorithms, the following experiments were carried out (see also Table 2):

(1) Open loop run. Model runs for an ensemble of stochastic realisations from March 1, 2012 to December 31, 2012 without data assimilation.

(2) State updating only. In this case, only soil moisture was updated (in the assimilation period) by the soil moisture observations.

(3) Joint state-parameter updating. In the assimilation period, soil moisture and selected parameters were updated by assimilating soil moisture observations. The updated parameter values from the final time step of the assimilation period were used in the verification period.

For each of these three groups, the following scenarios were studied:

(a) Type of algorithm. RRPF, MCMCPF and joint state-parameter estimation with EnKF using a dual approach or a state augmentation approach were tested for (3). EnKF with state updating only was tested for (2).

(b) Type of model. Both VIC-3L and CLM were studied for (1), (2) and (3).

100 ensemble members or particles (hereinafter: ensemble members) were used in the data assimilation experiments. Precipitation was perturbed were perturbed by multiplicative error ~ N(1,0.1) to represent the uncertainty of measured precipitation at the site. Soil parameters were perturbed as in Table 1. The soil moisture observation error is assumed to be normally distributed with mean equal to 0 and standard deviation equal to 0.02m3/m3, for both VIC-3L and CLM. The model error was set to zero assuming that uncertainty was captured by uncertain parameters and model forcings. Parameter inflation according to Whitaker and Hamill (2012) was applied (Eq. (33)) forcing the ensemble of parameters to have a spread equal to the prior ensemble standard deviations for the parameters.

4.3 Results

Two criteria are used to evaluate the performance of different scenarios: the Nash-Sutcliffe model efficiency (NSE) coefficient and the Root Mean Square Error (RMSE):

$$\text{NSE} = 1 - \frac{\sum_{t=1}^{T}\left(\theta_t^{sim} - \theta_t^{obs}\right)^2}{\sum_{t=1}^{T}\left(\theta_t^{obs} - \frac{1}{T}\sum_{t=1}^{T}\theta_t^{obs}\right)^2} \tag{40}$$

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\theta_t^{sim} - \theta_t^{obs}\right)^2} \tag{41}$$

where θ_t^{sim} is the ensemble mean soil moisture content at time step t, θ_t^{obs} the soil moisture observation at time step t and T is the number of time steps. The NSE and RMSE values were calculated only for soil moisture content as no reliable information was available on the true values for the soil hydraulic properties. These performance measures were evaluated separately for the verification and assimilation period. A NSE value equal

10   to 1 and RMSE equal to 0 imply a perfect prediction.

4.3a Results for VIC-3L

Figure 3 shows the soil moisture time series for the three VIC-3L model layers during the data assimilation period. The figure compares time series for the four scenarios with parameter estimation. The soil moisture time series for the first model layer are characterized by sharper fluctuations related to rainfall. This is especially the

15   case for summer and related to some intensive rainfall events combined with faster drying due to higher evapotranspiration. As expected, the second and third layer show a slower response to rainfall, with flatter soil moisture time series. Soil moisture content for the third layer shows a slow and steady increase. Data assimilation is able to adjust soil moisture values towards the observed ones. However, RRPF does not reproduce measured soil moisture content at 50cm depth well for the period from March to June. From July

20   onwards simulated soil moisture content with RRPF is close to the observations again. As a consequence, the NSE value of RRPF for the third layer is below zero. Also MCMCPF shows a reduced performance for the third layer with a NSE equal to 0.1279, which is related to a dry bias. EnKF results in better simulation results for the third layer (both for state augmentation and dual estimation) compared to RRPF and MCMCPF.

Figure 4 shows the NSE and RMSE values of soil moisture content for the assimilation period and all scenarios.

25   The open loop deviates most from the measured values, but if states are updated RMSE values are reduced by 68%, 82% and 95% for the three layers, compared to the open loop run. This means EnKF without parameter estimation works very well during the assimilation period even though only states are updated. The two EnKF-scenarios show a similar performance during the assimilation period with similar NSE and RMSE values. RMSE-reductions compared to the open loop run are for the augmentation approach 42% for the first layer, and

30   88% both for the second and third layer. The two particle filter algorithms (RRPF and MCMCPF) give for the first and second layer results comparable to the two EnKF-algorithms. Overall, during the assimilation period, EnKF without parameter estimation (noParamUpdate) outperforms DA with parameter estimation, and the EnKF-algorithms give better results than the PF-algorithms, related to the performance for the third model layer. MCMCPF gives better results than RRPF.

35   Figure 5 shows the parameter evolution for the four parameter estimation scenarios during the assimilation (and parameter calibration) period. In general, parameters show similar tendencies during the calibration period for

these scenarios. The parameters estimated by MCMCPF show much larger temporal fluctuations than for the other three methods. This is inherent to the MCMCPF methodology. MCMC allows for relatively large moves with jumps large enough to cover the complete posterior distribution of states and parameters. Even although the soil moisture time series for the state augmentation and dual estimation method are very similar, the temporal evolution of their parameter values are different. Nevertheless, the updating of the AUG and DUAL parameters still follow the same general tendency. The temporal evolution of parameter values for the first layer shows more fluctuations than for the second and third layer. This is related to rainfall events as soil moisture content in the first layer is sensitive to rainfall, which affects also the parameter characterization. Subfigure (h) in Fig. 5 shows the maximum baseflow velocity Dm in the third layer, which is a key parameter to calculate the baseflow. The time series of Dm for the EnKF-algorithms show a fast decrease in the first month and a stable tendency afterwards, whereas the Dm time series for RRPF decreases continuously until the last month. This slower convergence might also explain the worse performance of RRPF for a substantial part of the assimilation period.

Figure 6 shows the temporal evolution of the parameters log10Ks and β for the second layer and the four data assimilation algorithms. The mean of the ensemble members tends to be stable for the four data assimilation algorithms. A too narrow spread of ensemble members would lead to filter divergence. For the state augmentation (AUG) and dual estimation (DUAL), the spread of the ensemble members is kept large enough during the whole assimilation period as the ensemble inflation method helped to keep adequate ensemble spread. RRPF and MCMCPF also have enough ensemble spread because of parameter perturbation and MCMCPF resampling. Parameters change largely from late April onwards, which is related to intensive precipitation events from late April onwards (see also Fig. 2a).

Figure 7 displays soil moisture time series for the verification period, for all three model layers and for all four data assimilation algorithms. Soil moisture content shows stronger fluctuations over the first three months (August, September, and October) related to intensive rainfall events and the higher evapotranspiration during these months. The performance of the data assimilation algorithms shows differences over this period. In the first three months, RRPF shows the worst and DUAL the best performance in terms of reproducing the measurement data. In the last months of the verification period, the opposite behavior can be observed. All the four data assimilation algorithms do not perform well for the third model layer. This might be related to the fact that aquifers are not included in VIC and because of the simple baseflow parameterization.

The NSE and RMSE values for soil moisture characterization in the verification period and the three soil layers are plotted in Fig. 8. Generally, the overall RMSE values for the verification period are high compared to the assimilation period. In the verification period, the RMSE values of the scenario noParamUpdate are close to the RMSE values of the open loop run. If soil parameters were updated during the assimilation period, the RMSE values for soil moisture characterization were reduced. More specifically, state augmentation (AUG) shows a RMSE improvement of 68% and 36% for the second and third model layer (compared with the open loop run), a result very similar to the dual estimation with 67% respectively 36% RMSE-reduction. Results are also not very different for the two particle filter algorithms with 67% and 18% RMSE-reduction for MCMCPF, and 69% and 39% RMSE-reduction for RRPF. The NSE values for the third model layer are negative, indicating the bad performance of the algorithms for this layer.

4.3b Results for CLM

Figure 9 shows the CLM soil moisture time series for the assimilation period as obtained by application of the four different data assimilation algorithms. The performance of the data assimilation algorithms varies more than for the VIC-simulations. State augmentation (AUG) and dual estimation (DUAL) perform slightly better than

5 RRPF and MCMCPF for all the three layers. The soil moisture fluctuations at 5cm depth could not always be reproduced well by data assimilation. RRPF shows the worst performance, especially at 50cm depth. Figure 10 shows the NSE and RMSE values for soil moisture characterization during the assimilation period for all scenarios. In general, the performance is very good if only states are updated. State augmentation (AUG) and dual estimation (DUAL) show a similar performance with a RMSE-reduction (compared to the open loop run) of

10 63% (66%) for layer 1, 80% (82%) for layer 2 and 86% (87%) for layer 3 for the augmentation (dual estimation) method. RMSE-reductions are smaller for MCMCPF (between 47% and 75%) and especially for RRPF (between 30% and 60%).

Figure 11 displays the ensemble of the temporal evolutions of log10ks and the soil hydraulic parameter B at 50cm depth during the assimilation (calibration) period. Overall, changes in parameter values are small and

15 towards the end of the calibration period the behavior has become quite stable. The figure shows that the inflation method is able to keep the ensemble spread except for RRPF with an ensemble spread which is clearly too low. The poorer performance of RRPF compared to the other data assimilation algorithms is likely related to the reduced ensemble spread.

Figure 12 shows time series of CLM-calculated soil moisture content for the three layers for the verification

20 period. The temporal evolution of soil moisture content at shallow depths (5cm and 20cm) for state augmentation (AUG), dual estimation (DUAL) and MCMCPF is characterized by a very similar consistency with the observations. At 50cm depth the differences between the data assimilation algorithms are larger. Figure 13 shows the NSE and RMSE values for soil moisture content characterization in the verification period for the different data assimilation scenarios. The RMSE values for the verification period are higher than for the

25 assimilation period. If parameters were not updated (scenario noParamUpdate) in the assimilation period, soil moisture characterization is close to the open loop run, and even slightly worse than the open loop run at 5cm depth. State augmentation (AUG), dual estimation and MCMCPF show all very similar RMSE-reductions (compared to the open loop run) of 18-23% for 5cm depth, 26%-30% for 20cm depth and 66%-70% for 50cm depth. The performance of RRPF is slightly worse for the second and third layer, compared to the other data

30 assimilation algorithms.

5. Discussion

This study evaluated four sequential data assimilation algorithms in combination with two land surface models for joint state-parameter estimation with measured data at the Rollesbroich site in western Germany. The important novel aspect of this work is that this kind of evaluation and comparison study is done for real-world

35 data.

It was shown that soil properties and model parameters (i.e., hydraulic conductivity, soil texture, and VIC model parameter Dm) estimated with variants of EnKF or PF, resulted in improved model predictions during a verification period (without data assimilation) where the estimated parameters were used as model input. The

improvement (compared to open loop runs) was considerable, especially for deeper soil layers, the land surface model CLM and the EnKF-based algorithms. However, this improvement does not necessarily imply that the estimated parameters are also closer to the real-world values. Updated parameters might compensate for model structural errors and biases. If model structural errors and biases have a strong correlation over time (i.e., are

5 very persistent), estimated parameters which compensate for model bias still give an improved model prediction in the verification period. Whereas in synthetic studies it could be confirmed that parameter estimates indeed approach the true parameter values, this cannot be confirmed for the real-world study.

The performance of the four data assimilation algorithms does not differ very much in this study. However, the EnKF-based algorithms slightly outperform the particle filter based data assimilation algorithms if 100 ensemble

10 members/particles are used. The difference between the data assimilation algorithms is larger for CLM, which is probably related to the fact that indirectly more parameters are affected by the calibration (by the pedotransfer functions) than for VIC. It can be expected that in case a large number of unknown parameters has to be estimated it will be more difficult for PF to find those parameters than it is for EnKF. Nevertheless, the small difference in performance between EnKF and PF based algorithms indicates that PF is also an efficient data

15 assimilation algorithm for problems of this size. The results of this study are obtained for the point scale, a relatively small ensemble size of 100 (which is nevertheless larger than typically used for data assimilation in combination with land surface models) and for relatively short parameter estimation and verification periods. Given the CPU-intensity of the calculations a larger comparison was beyond the scope of this work.

DeChant and Moradkhani (2012) used a range of performance measures, like Nash-Sutcliffe efficiency (NSE),

20 Reliability (α), and Normalized root-mean-square error ratio (NRR), to evaluate EnKF and PF in state-parameter estimation. They also concluded that EnKF and PF showed similar performance. EnKF was more effective in the verification period but its ensemble members had a too low spread. The PF characterized more accurately the tails of the posterior distribution. Dumedah and Coulibaly (2013) found that PF performed better than EnKF when forecasting for longer lead time periods. They observed that model simulations were stronger adjusted

25 towards the observations in case EnKF was used for data assimilation, whereas for PF this was less the case. In our study, an inflation algorithm was applied to the ensemble of parameter values to ameliorate filter divergence, which might have helped EnKF to better characterize the posterior distribution of parameters and states. The disadvantage of this – commonly applied – procedure is that the uncertainty of the estimated parameters is not characterized well, as the parameter uncertainty is kept at the same level as the prior uncertainty. If possible, it is

30 therefore better to use very large ensemble sizes to avoid filter inbreeding and have also a good uncertainty characterization.

It is not surprising that the EnKF is more efficient and effective than the PF. Both approaches use an ensemble of realizations to approximate the forecast distribution, yet they differ fundamentally in their analysis step. The EnKF updates the simulated state variables of each ensemble member using the difference of their forecasted

35 output variable(s) (could be one or more of the simulated states) and corresponding observed value(s). This difference is then transformed into the state space using the measurement operator and determines the analysis values of the state variables. The measured values of the output variable(s) are thus used directly in the analysis step. In the PF on the contrary, not the measured values are used to determine the state update in the analysis step but rather the likelihood of each trajectory. This likelihood measures in probabilistic terms the agreement

40 between the forecasted output variable(s) and their measured values, yet constitutes only a proxy of their

distance. The value of the likelihood does generally not say anything about how close the forecasted variables are to their measured counterparts. What is more, the value of the likelihood is the same for a given distance of the forecasted variables to their measured values, whether they are overestimating or underestimating the data. This makes it much harder to determine an adequate size and direction (up or down) of the state update with

5   MCMC resampling. This explains why PF-MCMC methods cannot be as efficient and effective as EnKF-based data assimilation schemes. Multiple MCMC resampling steps can increase significantly the particle ensemble by allowing each particle trajectory to improve its likelihood. Yet, this deteriorates significantly the efficiency of implementation as each new particle that is generated during resampling requires a separate model evaluation to determine the likelihood of the proposed trajectory. One can improve significantly the efficiency of PF-based

10   data assimilation schemes if one adopts the update rule of the EnKF during particle resampling with MCMC [Vrugt et al., 2013].

Differences between land surface models were larger than differences between data assimilation algorithms in this study. CLM performed better than VIC, especially for the deepest model layer. Although it is important not to over interpret this result, as this is only a study for one site, the worse performance of VIC could be related to

15   the missing groundwater/subsurface component in this model. In CLM, the interaction between the unsaturated zone and groundwater is included. The change of water table depth is calculated and included as boundary condition for solving flow in the unsaturated zone.

6. Conclusion

Different sequential data assimilation algorithms were tested in combination with the Variable Infiltration

20   Capacity Model (VIC) and the Community Land Model (CLM). In total four sequential data assimilation algorithms were evaluated for joint state-parameter estimation: two variants of the Ensemble Kalman Filter (EnKF) (augmentation method and dual estimation), and two variants of the Particle Filter (Residual Resampling Particle Filter (RRPF) and Markov Chain Monte Carlo Particle Filter (MCMCPF)). The performance of the four sequential data assimilation methods in combination with two land surface models was evaluated for the

25   TERENO-observation site Rollesbroich in the western part of Germany. The highly equipped site allows to gain more insight in the performance of data assimilation algorithms for joint state-parameter estimation for land surface models. Measured soil moisture data at 5cm, 20cm and 50cm depth from different wireless sensor network were averaged over the area and used for assimilation. The assimilation period (including parameter estimation) was from March 2012- July 2012. The parameter estimates for the four data assimilation algorithms

30   were evaluated for the period of August 2012- December 2012. The performance of the four different joint state and parameter estimation methods in the verification period was not very different, with a slightly better performance of the augmentation method and dual estimation method and a slightly worse performance of RRPF and MCMCPF. The difference in performance between VIC and CLM was larger than the difference in performance between the four data assimilation methods. CLM performed better than VIC especially for the

35   deep soil layers. This is probably related to the poor representation of groundwater subsurface flow in VIC. The control of groundwater as lower boundary condition and its impact on the vadose zone in the form of moisture supply is neglected in VIC. It results here in an underestimation of soil moisture content for the deeper soil layer.

Appendix A: Parametrization of the VIC Model

The water balance for a given time step is given by:

$$\frac{\partial S}{\partial t} = P - E - Q \tag{A1}$$

where $\frac{\partial S}{\partial t}$ [LT-1] is the change of water storage, P [LT-1] is precipitation, E [LT-1] is evapotranspiration and Q

5 [LT-1] is runoff. E is composed of soil evaporation, transpiration by vegetation and evaporation from intercepted water. Bare soil evaporation is calculated by the equation of Francini and Pacciani (Franchini and Pacciani, 1991). Evaporation from intercepted water is calculated based on canopy potential evapotranspiration which is calculated by the Penman-Monteith equation (Shuttleworth, 2007). Maximum amount of water intercepted by the canopy is 0.2 times LAI (Dickinson, 1984). Vegetation transpiration is estimated using Blondin (1991) and

10 Ducoudre et al. (1993), where canopy resistance is calculated by minimum canopy resistance, LAI, photosynthetically active radiation flux factor, temperature factor, vapor pressure deficit factor, and soil moisture factor. The four factors are available through Wigmosta et al. (1994). Q includes direct runoff Q_d [LT-1] and baseflow Q_b [LT-1]. The VIC model assumes there is no lateral flow in the top two soil layers. Therefore the movement of moisture can be characterized by (Liang et al., 1996):

15 $$\frac{\partial \theta_i}{\partial t} z_i = P - Q_d - Q_{i,i+1} - E \qquad (i=1, 2) \tag{A2}$$

$$\frac{\partial \theta_3}{\partial t} z_3 = Q_{2,3} - E - Q_b \tag{A3}$$

where θ [L3L-3] is volumetric soil moisture content, z_i [L] is soil depth for layer i (i =1,2), $Q_{i,i+1}$ [LT-1] is the vertical drainage between layer i and i+1, Q_d [LT-1] is calculated for layer 1 and layer 2. Evapotranspiration E [LT-1] can occur from soil moisture stored in the three layers. In case of bare soil evaporation only, E is equal to

20 zero in Eq. (A3) because there is no evaporation from layer 3. If plant roots are present in layer 3, E also takes place from layer 3. Base flow Q_b [LT-1] is only generated from the third layer.

Assuming that the drainage is driven by gravity, the Brooks and Corey (1964) relation is used to estimate unsaturated hydraulic conductivity, and the vertical drainage between layer i and i+1 is expressed as (Liang et al., 1994):

25 $$Q_{i,i+1} = k_{s,i} \left(\frac{\theta_i - \theta_{r,i}}{\theta_i^{max} - \theta_{r,i}} \right)^{\beta_i} \qquad (i=1,2) \tag{A4}$$

where $k_{s,i}$ [LT-1] is the saturated hydraulic conductivity for layer i, $\theta_{r,i}$ [L3L-3] is the residual soil moisture content, exponent β_i [-] is a model parameter and θ_i^{max} [L3L-3] is the maximum soil moisture content of layer i:

$$\theta_i^{max} = \phi_i \qquad (i=1,2) \tag{A5}$$

where ϕ_i [-] is the porosity of the soil layer i. Exponent β_i [-] is a function of the pore size distribution index B_p

30 [-]:

$$\beta_i = \frac{2}{B_p} + 3 \tag{A6}$$

Q_d is calculated for layer 1 and layer 2 as follows (Liang et al., 1996):

$$Q_d = \begin{cases} P - (\theta_1^{max} - z_1\theta_1) - (\theta_2^{max} - z_2\theta_2) + (\theta_1^{max} + \theta_2^{max})\left(1 - \frac{I+P}{i_m}\right)^{1+b}, & P+I \leq I_m \\ P - (\theta_1^{max} - z_1\,\theta_1) - (\theta_2^{max} - z_2\,\theta_2), & P+I > I_m \end{cases} \quad (A7)$$

where the parameter b [-] is the infiltration shape parameter which is a measure of the spatial variability of the infiltration capacity. Because of the lack of hydrologic information at site, it is usually determined by calibration. The reason for calculating Q_d for the entire upper soil (layer 1 and layer 2) is that the top layer has a very small water holding capacity (i.e. $z_1\phi_1$). The variable filtration capacity I [L] of the upper soil is a function of the maximum filtration capacity I_m [L] [Zhao, 1992]:

$$I = I_m(1 - (1 - A)^{\frac{1}{b}}) \text{ with } I_m = (1 + b)(\theta_1^{max} + \theta_2^{max}) \quad (A8)$$

where A [-] is the fraction of area where infiltration capacity is less than I_m:

$$A = 1.0 - (1.0 - \frac{z_1\theta_1 + z_2\theta_2}{\theta_1^{max} + \theta_2^{max}})^{\frac{b}{1+b}} \quad (A9)$$

Q_b is formulated according the Arno model equation (Franchini and Pacciani, 1991):

$$Q_b = \begin{cases} \frac{D_S\,D_m}{W_S\theta_3^{max}}\theta_3 z_3, & 0 \leq \theta_3 z_3 \leq W_S\theta_3^{max} \\ \frac{D_S D_m}{W_S\theta_3^{max}}\theta_3 z_3 + \left(D_m - \frac{D_S D_m}{W_S}\right)\left(\frac{\theta_3 z_3 - W_S\theta_3^{max}}{\theta_3^{max} - W_S\theta_3^{max}}\right)^2, & \theta_3 z_3 > W_S\theta_3^{max} \end{cases} \quad (A10)$$

where D_m [LT-1] is the maximum baseflow velocity, D_S [-] is the fraction of D_m where nonlinear baseflow begins and W_S [-] is the fraction of maximum soil moisture (θ_3^{max}). In VIC-3L, there is no distinction between unsaturated and saturated zones in the lower layer. In other words, the unsaturated and saturated zones are treated in a lumped sense. Therefore Q_b includes both drainage from the unsaturated part and baseflow from groundwater (Liang et al., 1996; Liang et al., 2003). Liang et al. (2003) developed a new parameterization into the VIC-3L model (called VIC-ground) to represent the interaction between surface water and groundwater. Their results showed that soil moisture content for the lower VIC-ground layer was in general higher than for VIC-3L.

Appendix B: Parametrization of the CLM Model

The hydraulic conductivity k_i [LT-1], soil matric potential ψ_i [L] and soil thermal conductivity λ_i [WL-1K-1] for layer i are determined by sand and clay content (Clapp and Hornberger, 1978; Cosby et al., 1984) and organic matter density (Lawrence and Slater, 2008). The calculation of the hydraulic conductivity k_i at the interface of two adjacent layers i and i + 1 is described in detail in (Oleson et al., 2013; Han et al., 2014).

The soil matric potential ψ_i [L] is given by:

$$\psi_i = \psi_{sat,i}\left(\frac{\theta_i}{\theta_{sat,i}}\right)^{-B_i} \quad (B1)$$

where

$$\psi_{sat,i} = -10 \cdot 10^{1.88 - 0.0131f_{s,i}}(1 - f_{om,i}) - 10.3f_{om,i} \quad (B2)$$

$$B_i = (1 - f_{om,i})(2.91 + 0.159f_{c,i}) + 2.7f_{om,i} \quad (B3)$$

$$\theta_{sat,i} = \left(1 - f_{om,i}\right)\theta_{sat,min,i} + 0.9 f_{om,i} \tag{B4}$$

$$\theta_{sat,min,i} = 0.489 - 0.00126(f_{s,i}) \tag{B5}$$

where θ_i [L3L-3] is soil moisture content for layer i, $\theta_{sat,i}$ [L3L-3] is saturated soil moisture content, $\psi_{sat,i}$ [L] is the saturated soil matric potential, B_i [-] is the Clapp-Hornberger exponent, $f_{s,i}$ [-] is sand fraction, $f_{c,i}$ [-] is clay fraction and $f_{om,i}$ [-] is organic matter fraction.

The water balance is given by Eq. (A1). ΔS includes the changes in canopy water, surface water, snow water, soil water, soil ice and water stored in the unconfined aquifer. In addition to surface and subsurface runoff, Q also includes runoff from glaciers, wetlands and lakes. Latent heat fluxes E [ML-2T-1] include ground evaporation, canopy evaporation and transpiration. The basic processes can be described by the fundamental expression (Schwinger et al., 2010; Oleson et al., 2013):

$$E = \frac{\rho}{r}(q-q_a) \tag{B6}$$

Where ρ is the density of air [ML-3], r is aerodynamic resistance [TL-1], q [MM-1] is the specific humidity of soil pore space (or canopy space) or saturated specific humidity of snow or surface water and q_a [MM-1] is specific humidity at the atmospheric level when ground evaporation is calculated, or the saturated specific humidity within the canopy when canopy evapotranspiration is calculated. r, q and q_a are based on Monin-Obukhov similarity theory (Schwinger et al., 2010; Oleson et al., 2013).

The one-dimensional vertical flow in the unsaturated zone is influenced by infiltration, surface and subsurface runoff, canopy transpiration, and interactions with groundwater. A modified Richards equation is used to predict vertical soil water flow:

$$\frac{\partial \theta_i}{\partial t} = \frac{\partial}{\partial z}\left[k_i\left(\frac{\partial(\psi_i - (\psi_{sat,i} + z_\nabla - z_i))}{\partial z}\right)\right] - E \tag{B7}$$

where z_∇ [L] is groundwater table depth, and E [LT-1] is evapotranspiration loss. This equation has different boundary conditions depending on the presence of a water table in the soil column.

In CLM, water table depth z_∇ is calculated according to Niu (Niu et al., 2007). An unconfined aquifer is assumed to lie below the soil column. If the water table is within the soil column, water storage in the unconfined aquifer is assumed to be constant as the soil column is saturated with water below the water table and a zero-flux bottom boundary condition is applied. The recharge to the unconfined aquifer is calculated by:

$$q_{recharge} = -k_{wt}\frac{(-\psi_{wt})}{(z_\nabla - z_{wt})} \tag{B8}$$

where k_{wt} [LT-1] is the hydraulic conductivity of the layer containing the groundwater table, ψ_{wt} [L] the soil matric potential of that layer, z_{wt} [L] the depth of that layer and z_∇ [L] the water table depth. Drainage $q_{drainage}$ [ML-2T-1] is calculated by a simple TOPMODEL-based (SIMTOP) scheme (Niu et al., 2005)

$$q_{drainage} = 10\sin(\varepsilon)\exp(-2.5z_\nabla) \tag{B9}$$

where ε [Rad] is the mean topographic slope in the grid cell. The change in the water table depth is then given by:

$$\Delta z_{\overline{v}} = \frac{\Delta W}{S_y} \quad \text{with } \Delta W = (q_{recharge} - q_{drainage})\Delta t \tag{B10}$$

where S_y [-] is the specific yield depending on the soil properties.

Acknowledgment

We thank Terrestrial Environmental Observatories (TERENO) for providing the measurement data. We also would like to acknowledge the support by the supercomputing facilities of Forschungszentrum Julich (JUROPA). This work was financed by a stipend from the China government.

Reference

Anderson, J. L. (2007), An adaptive covariance inflation error correction algorithm for ensemble filters, Tellus Series a-Dynamic Meteorology and Oceanography, 59(2), 210-224, doi:10.1111/j.1600-0870.2006.00216.x.

Arulampalam, M. S., S. Maskell, N. Gordon, and T. Clapp (2002), A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, Ieee Transactions on Signal Processing, 50(2), 174-188, doi:10.1109/78.978374.

Bailey, R. T. and D. Bau (2012), Estimating geostatistical parameters and spatially-variable hydraulic conductivity within a catchment system using an ensemble smoother, Hydrology Earth System Sciences, 16, 287–304.

Bateni, S. M., and D. Entekhabi (2012), Surface heat flux estimation with the ensemble Kalman smoother: Joint estimation of state and parameters, Water Resources Research, 48, doi:10.1029/2011wr011542.

Bertoldi G. (2004), The water and energy balance at basin scale: a distributed modeling approach, University of Trento, Monograph of the School of Doctoral Studies in Environmental Engineering, 202 pp., ISBN 88-8443-069-0.

Blasone, R. S., Vrugt, J. A., Madsen, H., Rosbjerg, D., Robinson, B. A., and Zyvoloski, G. A. (2008), Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. Advances in Water Resources, 31(4), 630-648.

Blondin, C. (1991), Parameterization of land-surface processes in numerical weather prediction, in Land Surface Evaporation: Measurements and Parameterization, edited by T. J. Schmugge and J. C. Andre, Springer-Verlag, New York, 31-54.

Bogena, H. R., M. Herbst, J. A. Huisman, U. Rosenbaum, A. Weuthen, and H. Vereecken (2010), Potential of Wireless Sensor Networks for Measuring Soil Water Content Variability, Vadose Zone Journal, 9(4), 1002-1013, doi:10.2136/vzj2009.0173.

Brooks, R. H., and A. T. Corey (1964), Hydraulic Properties of Porous Media, Colorado State University, 3.

Burgers, G., P. J. van Leeuwen, and G. Evensen (1998), Analysis scheme in the ensemble Kalman filter, Monthly Weather Review, 126(6), 1719-1724, doi:10.1175/1520-0493(1998)126<1719:asitek>2.0.co;2.

Carpenter, J., P. Clifford, and P. Fearnhead (1999), Improved particle filter for nonlinear problems, Radar, Sonar and Navigation, IEE Proceedings - , 146(1), 2-7, doi:10.1049/ip-rsn:19990255.

Chen, Y., and D. Zhang (2006), Data assimilation for transient flow in geologic formations via ensemble Kalman filter, Advances in Water Resources, 29, 1107–1122.

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

Cherkauer, K. A., and D. P. Lettenmaier (1999), Hydrologic effects of frozen soils in the upper Mississippi River basin, Journal of Geophysical Research-Atmospheres, 104(D16), 19599-19610, doi:10.1029/1999jd900337.

Clapp, R. B., and G. M. Hornberger (1978), Empirical equations for some soil hydraulic properties, Water Resources Research, 14(4), 601-604, doi:10.1029/WR014i004p00601.

Cosby, B. J., G. M. Hornberger, R. B. Clapp, and T. R. Ginn (1984), A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils, Water Resources Research, 20(6), 682-690, doi:10.1029/WR020i006p00682.

DeChant, C. M., and H. Moradkhani (2012), Examining the effectiveness and robustness of sequential data assimilation methods for quantification of uncertainty in hydrologic forecasting, Water Resources Research, 48, doi:10.1029/2011wr011011.

Demaria, E. M., B. Nijssen, and T. Wagener (2007), Monte Carlo sensitivity analysis of land surface parameters using the Variable Infiltration Capacity model, Journal of Geophysical Research-Atmospheres, 112(D11), 15, doi:10.1029/2006jd007534.

Dickinson, R. E. (1984), Modeling evapotranspiration for three-dimentional global climate models, in Climate Processes and Climate Sensitivity, Monogr. Ser., edited by J. E. Hansen and T. Takahashi, 58-72, Washington, D.C.

Douc, R., C. Olivier, and M. Eric (2005), Comparison of resampling schemes for particle filtering, Image and Signal Processing and Analysis, Proceedings of the 4th International Symposium on, 64-69, doi: 10.1109/ISPA.2005.195385.

Ducoudre, N. I., K. Laval, and A. Perrier (1993), SECHIBA, a new set of parameterizations of the hydrologic exchanges at the land-atmosphere interface within the LMD atmospheric general circulation model, Journal of Climate, 6(2), 248-273.

Dumedah, G., and P. Coulibaly (2013), Evaluating forecasting performance for data assimilation methods: The ensemble Kalman filter, the particle filter, and the evolutionary-based assimilation, Advances in Water Resources, 60, 47-63, doi:10.1016/j.advwatres.2013.07.007.

Erdal, D., I. Neuweiler, and U. Wollschläger (2014), Using a bias aware EnKF to account for unresolved structure in an unsaturated zone model, Water Resources Research, 50, 132–147, doi:10.1002/2012WR013443.

Evensen, G. (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model usin Monte-Carlo methods to forecast error statistics, Journal of Geophysical Research-Oceans, 99(C5), 10143-10162, doi:10.1029/94jc00572.

Franchini, M., and M. Pacciani (1991), Comparative-analysis of several conceptual rainfall runoff models, Journal of Hydrology, 122(1-4), 161-219, doi:10.1016/0022-1694(91)90178-k.

Franssen, H. J. H., and W. Kinzelbach (2008), Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem, Water Resources Research, 44(9), doi:10.1029/2007wr006505.

Gao, H., Q. Tang, X. Shi, C. Zhu, B. T. J., F. Su, S. J., P. M., D. P. Lettenmaier, and E. F. Wood (2010), Water Budget Record from Variable Infiltration Capacity (VIC) Model, In Algorithm Theoretical Basis Document for Terrestrial Water Cycle Data Records.

Gharamti, E. M., I. Hoteit, and J. Valstar (2013), Dual states estimation of a subsurface flow-transport coupled model using ensemble Kalman filtering, Advances in Water Resources, 60, 75-88, doi:10.1016/j.advwatres.2013.07.011.

Gordon, N. J., D. J. Salmond, and A. F. M. Smith (1993), Novel-approach to nonlinear non-Gaussian bayesian state estimation, Iee Proceedings-F Radar and Signal Processing, 140(2), 107-113.

Han, X., H. J. H. Franssen, C. Montzka, and H. Vereecken (2014), soil moisture and soil properties estimation in the community Land Model with synthetic brightness temperature, Water Resource Research, doi:10.1002/2013WR014586.

Kitanidis, P. K., and R. L. Bras (1980), Real-time forecasting with a conceptual hydrologic model .1. analysis of uncertainty, Water Resources Research, 16(6), 1025-1033, doi:10.1029/WR016i006p01025.

Kurtz, W., H.-J. Hendricks Franssen, H.-P. Kaiser, and H. Vereecken (2014), Joint assimilation of piezometric heads and groundwater temperatures for improved modeling of river-aquifer interactions, Water Resources Research, 50(2), 1665-1688, doi:10.1002/2013wr014823.

Lawrence, D. M., and A. G. Slater (2008), Incorporating organic soil into a global climate model, Climate Dynamics, 30(2-3), 145-160, doi:10.1007/s00382-007-0278-1.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hyhrologically based model of land-surface water and energy fluxes for general-circulation models, Journal of Geophysical Research-Atmospheres, 99(D7), 14415-14428, doi:10.1029/94jd00483.

Liang, X., E. F. Wood, and D. P. Lettenmaier (1996), Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification, Global and Planetary Change, 13(1-4), 195-206, doi:10.1016/0921-8181(95)00046-1.

Liang, X., Z. H. Xie, and M. Y. Huang (2003), A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model, Journal of Geophysical Research-Atmospheres, 108(D16), 17, doi:10.1029/2002jd003090.

Liu, and R. Chen (1998), Sequential Monte Carlo methods for dynamic systems, Journal of the American Statistical Association, 93(443), 1032-1044, doi:10.2307/2669847.

Liu, Y., and H. V. Gupta (2007), Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, Water Resources Research, 43(7), doi:10.1029/2006wr005756.

Lue, H., Z. Yu, Y. Zhu, S. Drake, Z. Hao, and E. A. Sudicky (2011), Dual state-parameter estimation of root zone soil moisture by optimal parameter estimation and extended Kalman filter data assimilation, Advances in Water Resources, 34(3), 395-406, doi:10.1016/j.advwatres.2010.12.005.

Milly, P. C. D., and A. B. Shmakin (2002), Global modeling of land water and energy balances. Part II: Land-characteristic contributions to spatial variability, Journal of Hydrometeorology, 296–307.

Montzka, C., H. Moradkhani, L. Weihermueller, H. J. H. Franssen, M. Canty, and H. Vereecken (2011), Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter, Journal of Hydrology, 399(3-4), 410-421, doi:10.1016/j.jhydrol.2011.01.020.

Montzka, C., J. P. Grant, H. Moradkhani, H. J. H. Franssen, L. Weihermüller, M. Drusch and H. Vereecken (2013), Estimation of radiative transfer parameters from L-band passive microwave brightness temperatures using advanced data assimilation,. Vadose Zone Journal, Special Section on Remote Sensing of Vadose Zone Hydrology, 12(3), DOI:10.2136/vzj2012.0040.

Moradkhani, S. Sorooshian, H. V. Gupta, and P. R. Houser (2005a), Dual state-parameter estimation of hydrological models using ensemble Kalman filter, Advances in Water Resources, 28(2), 135-147, doi:10.1016/j.advwatres.2004.09.002.

Moradkhani, H., C. M. DeChant, and S. Sorooshian (2012), Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-Markov chain Monte Carlo method, Water Resources Research, 48, doi:10.1029/2012wr012144.

Moradkhani, H., K. L. Hsu, H. Gupta, and S. Sorooshian (2005b), Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, Water Resources Research, 41(5), doi:10.1029/2004wr003604.

Niu, G. Y., Z. L. Yang, R. E. Dickinson, and L. E. Gulden (2005), A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, Journal of Geophysical Research-Atmospheres, 110, D21106, doi:10.1029/2005JD006111.

Niu, G. Y., Z. L. Yang, R. E. Dickinson, L. E. Gulden, and H. Su (2007), Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, Journal of Geophysical Research-Atmospheres, 112(D7), 14, doi:10.1029/2006jd007522.

Oleson, K., et al. (2013), Technical description of version 4.5 of the Community Land Model (CLM), NCAR Technical Note NCAR/TN-503+STR, 422, doi:10.5065/D6RR1W7M.

Pasetto D., et al. (2015), Impact of sensor failure on the observability of flow dynamics at the Biosphere 2 LEO hillslopes, Advances in Water Resources, doi: doi:10.1016/j.advwatres.2015.04.014.

Pauwels, V. R. N., A. Balenzano, G. Satalino, H. Skriver, N. E. C. Verhoest and F. Mattia (2009), Optimization of Soil Hydraulic Model Parameters Using Synthetic Aperture Radar Data: An Integrated Multidisciplinary Approach, Geoscience and Remote Sensing, IEEE Transactions on , 47(2), 455-467, doi: 10.1109/TGRS.2008.2007849.

Plaza, D. A., R. D. Keyser, G. J. M. D. Lannoy, L. Giustarini, P. Matgen, and V. R. N. Pauwels (2012), The importance of parameter resampling for soil moisture data assimilation into hydrologic models using the particle filter, Hydrol. Earth Syst. Sci., 16(2), 375-390, doi:10.5194/hess-16-375-2012.

Qin, J., S. Liang, K. Yang, I. Kaihotsu, R. Liu, and T. Koike (2009), Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal, Journal of Geophysical Research-Atmospheres, 114, doi:10.1029/2008jd011358.

Qu, B. H. R., H. J. A., M. G., P. Y. A., and V. H. (2014), Effects of soil hydraulic properties on the spatial variability of soil water content: evidence from sensor network data and inverse modeling, Vadose Zone Journal, 13(12), doi:10.2136/vzj2014.07.0099.

5   Rasmussen, J., H. Madsen, K. H. Jensen, and J. C. Refsgaard (2015), Data assimilation in integrated hydrological modeling using ensemble Kalman filtering: evaluating the effect of ensemble size and localization on filter performance, Hydrology Earth System Science, 19, 2999-3013, doi:10.5194/hess-19-2999-2015.

Schaake, J. C., V. I. Koren, Q. Y. Duan, K. Mitchell, and F. Chen (1996), Simple water balance model for estimating runoff at different spatial and temporal scales, Journal of Geophysics Research, 101(D3), 7461–7475, doi:10.1029/95JD02892.

10   Schwinger, J., S. Kollet, C. Hoppe, and H. E (2010), Sensitivity of latent heat fluxes to initial values and parameters of a Land-Surface Model, Vadose Zone Journal, 9(4), 984-1001.

Shi L., Q. Zhang, X. Song, and X. Fang (2015), Application of groundwater level data to data assimilation for unsaturated flow, Advances in Water Science, 26(3), 404-412.

Shi, Y., K. J. Davis, F. Zhang, C. J. Duffy, and X. Yu (2014), Parameter estimation of a physically based land
15   surface hydrologic model using the ensemble Kalman filter : A synthetic experiment, Water Resources Research, 50, 706–724, doi:10.1002/2013WR014070.

Shi, Y., K. J. Davis, F. Zhang and C. J. Duffy, and X. Yu (2015), Parameter estimation of a physically-based land surface hydrologic model using an ensemble Kalman filter: A multivariate real-data experiment, Advances in Water Resources, doi: 10.1016/j.advwatres.2015.06.009.

20   Shuttleworth, W. J. (2007), Putting the 'vap' into evaporation, Hydrology Earth System Science, 11(1), 210-244.

Song, X., L. Shi, M. Ye, J. Yang, and N. I. Michael (2014), Numerical Comparison of Iterative Ensemble Kalman Filters for Unsaturated Flow Inverse Modeling, Vadose Zone Journal, 13(2), doi: 10.2136/vzj2013.05.0083.

Tang Q., W. Kurtz, P. Brunner, H. Vereecken, and H. J. H. Franssen (2015), Characterisation of river–aquifer
25   exchange fluxes: The role of spatial patterns of riverbed hydraulic conductivities, Journal of Hydrology, doi:10.1016/j.jhydrol.2015.08.019.

Troy, T. J., E. F. Wood, and J. Sheffield (2008), An efficient calibration method for continental-scale land surface modeling, Water Resources Research, 44(9), doi:10.1029/2007wr006513.

Vereecken, H., J. A. Huisman, H. Bogena, J. Vanderborght, J. A. Vrugt, and J. W. Hopmans (2008), On the
30   value of soil moisture measurements in vadose zone hydrology: A review, Water Resources Research, 44, 21, doi:10.1029/2008wr006829.

Vrugt, J. A., C. G. H. Diks, H. V. Gupta, W. Bouten, and J. M. Verstraten (2005), Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, Water Resources Research, 41, W01017, doi:10.1029/2004WR003059.

35   Vrugt, J. A., C. J. F. ter Braak, C. G. H. Diks, and G. Schoups (2013), Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications, Advances in Water Resources, 51, 457-478, doi:10.1016/j.advwatres.2012.04.002.

Weerts, A. H., and G. Y. H. El Serafy (2006), Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models, Water Resources Research, 42(9),
40   doi:10.1029/2005wr004093.

Whitaker, J. S., and T. M. Hamill (2012), Evaluating Methods to Account for System Errors in Ensemble Data Assimilation, Monthly Weather Review, 140(9), 3078-3089, doi:10.1175/mwr-d-11-00276.1.

Wigmosta, M. S., L. W. Vail, and D. P. Lettenmaier (1994), A distributed hydrology-vegetation model for complex terrain, Water Resource Research, 30(6), 1665-1679.

45   Wu, C. C., and S. A. Margulis (2011), Feasibility of real-time soil state and flux characterization for wastewater reuse using an embedded sensor network data assimilation approach, Journal of Hydrology, 399(3-4), 313-325, doi: 10.1016/j.jhydrol.2011.01.011.

Wu, C. C., and S. A. Margulis (2013), Real-Time Soil Moisture and Salinity Profile Estimation Using Assimilation of Embedded Sensor Datastreams, Vadose Zone Journal, 12(1), doi: 10.2136/vzj2011.0176.

Xie, Z., F. Yuan, Q. Duan, J. Zheng, M. Liang, and F. Chen (2007), Regional parameter estimation of the VIC land surface model: methodology and application to river basins in China, Journal of Hydrometeorology, 8(3), 447-468, doi:10.1175/jhm568.1.

Zhao, R. J. (1992), The Xinanjiang model applied in China, Journal of Hydrology, 135(1-4), 371-381.

Table 1 summarizes soil parameters chosen to be updated during the assimilation period for the VIC model and CLM (N is normal distribution and U is uniform distribution).

Models	Variables	Description	Unit	Ranges	Magnitude of Perturbation
VIC	$\log_{10}k_s$	Saturated hydrologic conductivity	m/s	[-7, -3]	$+N(0, 1)$
	β	Exponent of the Brooks-Corey drainage equation	-	[8, 30]	$+U(-5, 5)$
	b	Infiltration shape parameter	-	[0.001, 0.8]	$+U(-0.1, 0.1)$
	D_m	Maximum velocity of baseflow	mm/day	(0, 30]	$+U(-10, 10)$
CLM		Clay fraction	percentage	[1, 100]	$+U(-10, 10)$
		Sand fraction	percentage	[1, 100]	$+U(-10, 10)$
		Organic matter density	kg/m^3	[1, 130]	$+U(-15, 15)$

Table 2 summarizes the scenarios used for CLM and VIC-3L and the introduced abbreviations will be used in tables and figures.

scenario description	Abbreviation
model open loop	Openloop
EnKF with updating states only	noParamUpdate
EnKF using the augmentation approach	AUG
EnKF using the dual estimation approach	DUAL
RRPF with parameter perturbation	PF
MCMCPF	MCMC
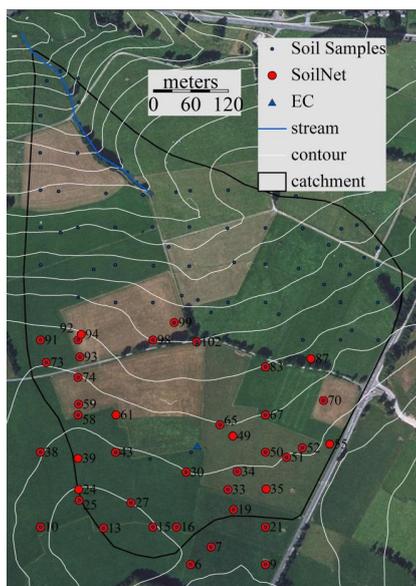
Figure 1 Overview of measurement devices in the Rollesbroich catchment. The blue dots are soil sample locations, red dots are soil network locations (soil moisture content and soil temperature are measured here), and the blue triangular indicates the eddy covariance tower. The Figure is taken from *Qu et al.* (2014).
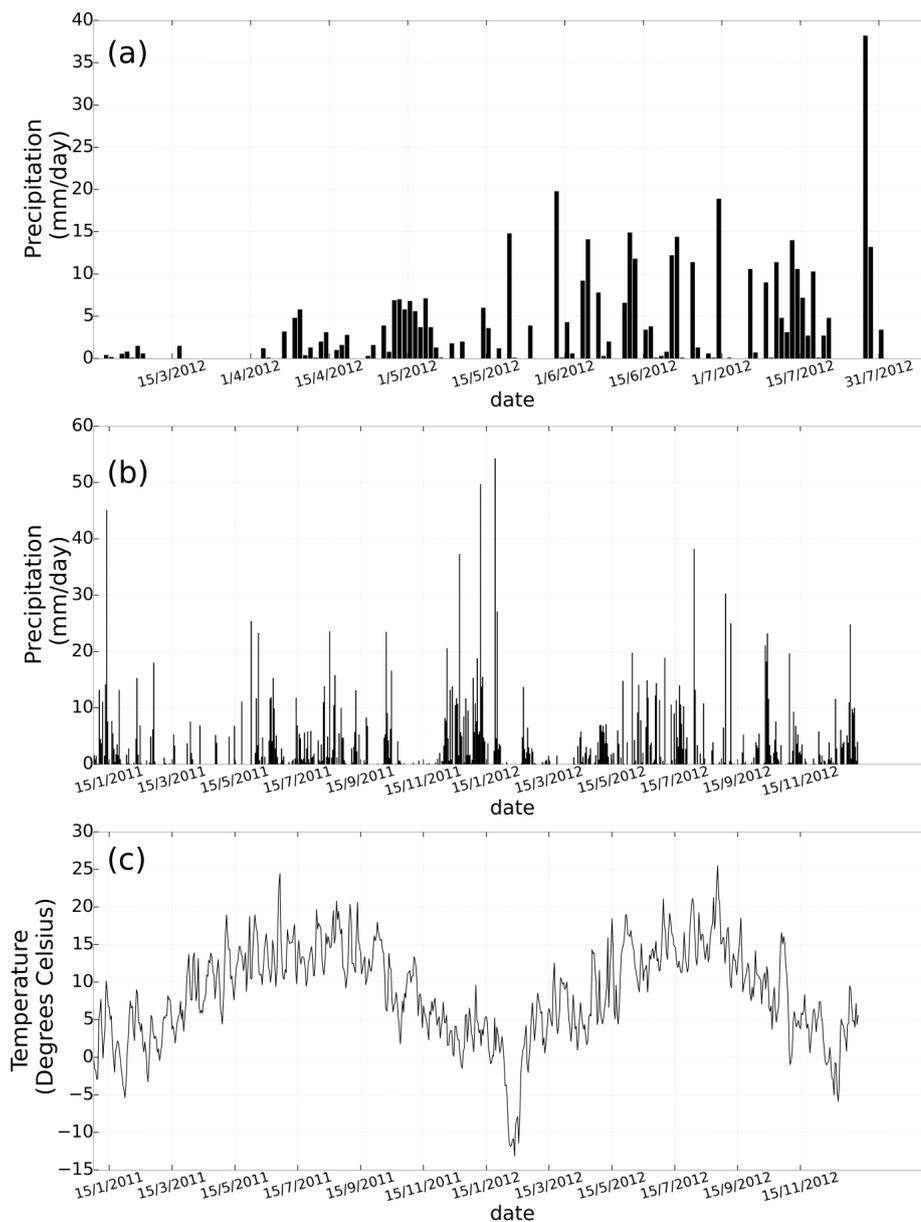
Fig. 2 (a) daily precipitation from March 1, 2012 to July 31, 2012 (parameter estimation period), week 1 is 01-03-2012 and week 22 coincides with 26-07-2012, (b) daily precipitation for the years 2011 and 2012, and (c) daily mean air temperature for the years 2011 and 2012, all measured at the Rollesbroich site.
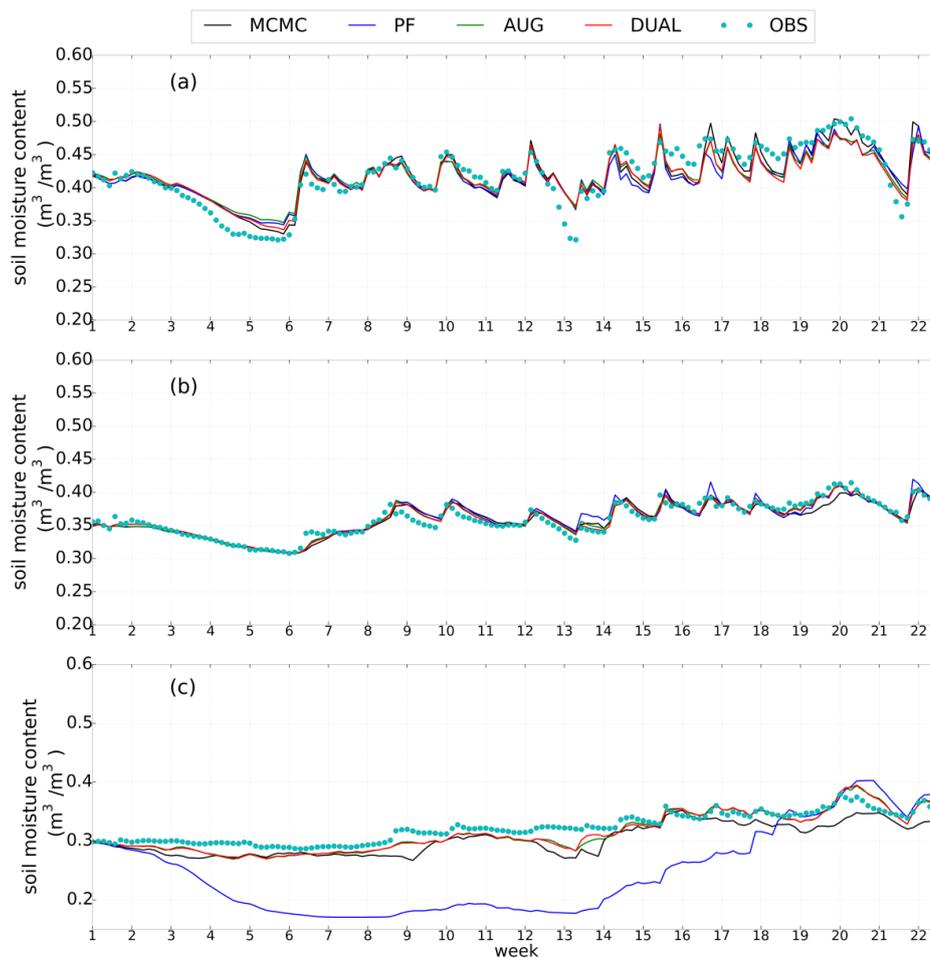
Figure 3 Time series of soil moisture content for different assimilation scenarios during the assimilation period
(March-July 2012) for the VIC-3L model: (a) 5 cm depth, (b) 20 cm depth and (c) 50 cm depth. Week 1 starts at
01-03-2012 and week 22 at 26-07-2012.
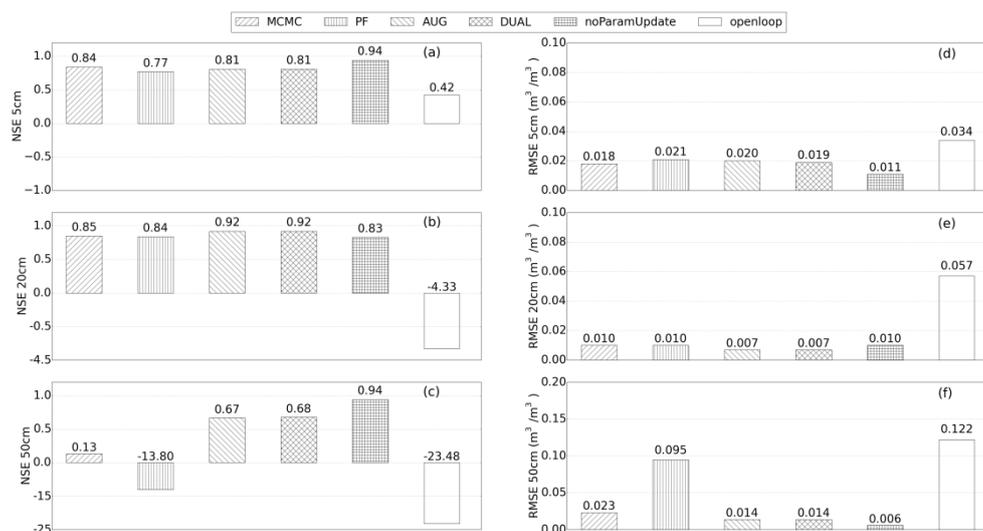
Figure 4 NSE and RMSE values for soil moisture content characterization for different scenarios in the assimilation period with the VIC-3L model: (a) NSE values for soil moisture content at 5cm, (b) NSE values at 20cm, (c) NSE value at 50cm, (d) RMSE values at 5cm, (e) RMSE values at 20cm, and (f) RMSE values at 50cm.
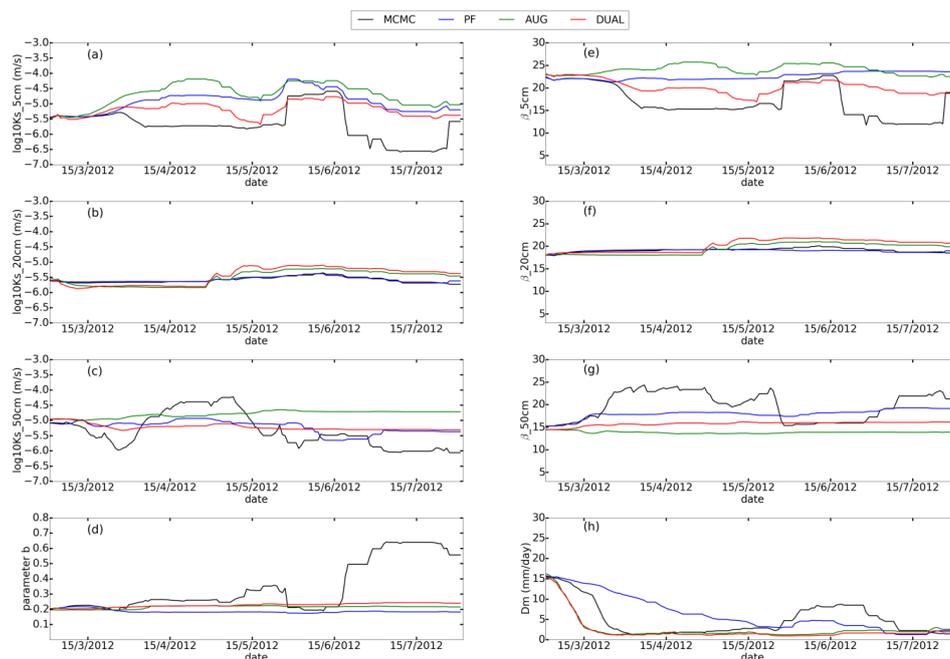
Figure 5 Temporal evolution of parameter values in the parameter estimation period (March 2012-July 2012), for the four data assimilation scenarios and the model VIC-3L. (a) Saturated hydraulic conductivity $\log_{10}k_s$ (m/s) at 5cm depth, (b) 20cm depth and (c) 50cm depth, (d) model parameter b, (e) model parameter β at 5cm depth, (f) 20cm depth and (g) 50cm depth, and (h) maximum velocity of baseflow D_m (mm/day).
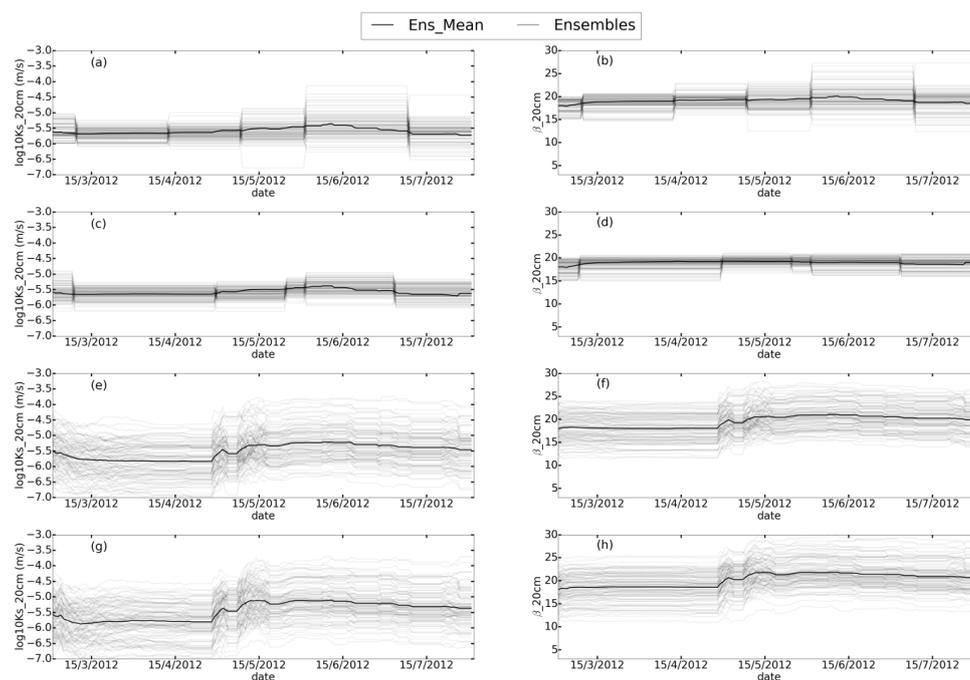
Figure 6 Temporal evolution of parameter values for 100 ensemble members in the parameter estimation period, for the four data assimilation scenarios and the model VIC-3L and the second model layer. Saturated hydraulic conductivity $\log_{10}k_s$ (m/s) at 20cm depth is displayed for the four methods: (a) MCMC, (c) PF, (e) AUG, and (g) DUAL. Model parameter β at 20cm depth for the four methods: (b) MCMC, (d) PF, (f) AUG and (h) DUAL.
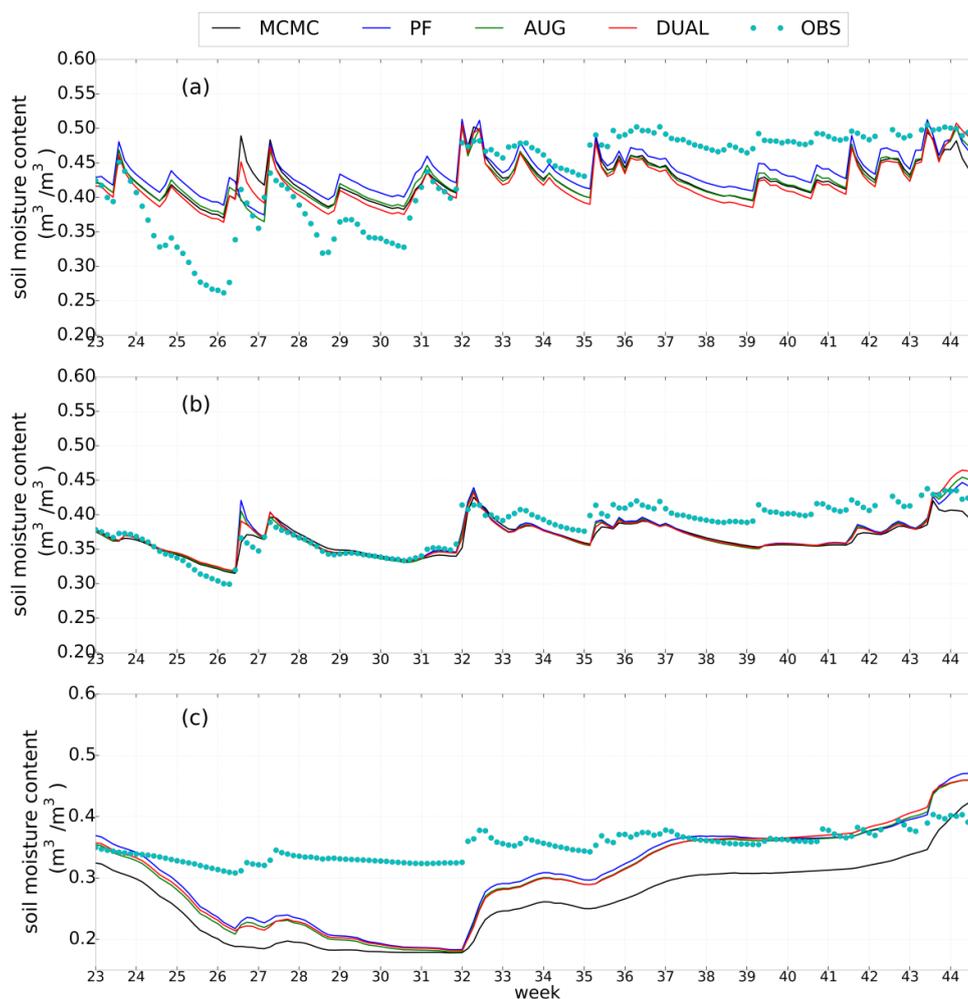
Figure 7 Time series of soil moisture content for different assimilation scenarios during the verification period
and for the VIC-3L model: (a) 5 cm depth, (b) 20 cm depth and (c) 50 cm depth. Week 23 starts with 01-08-2012
and week 44 with 26-12-2012.
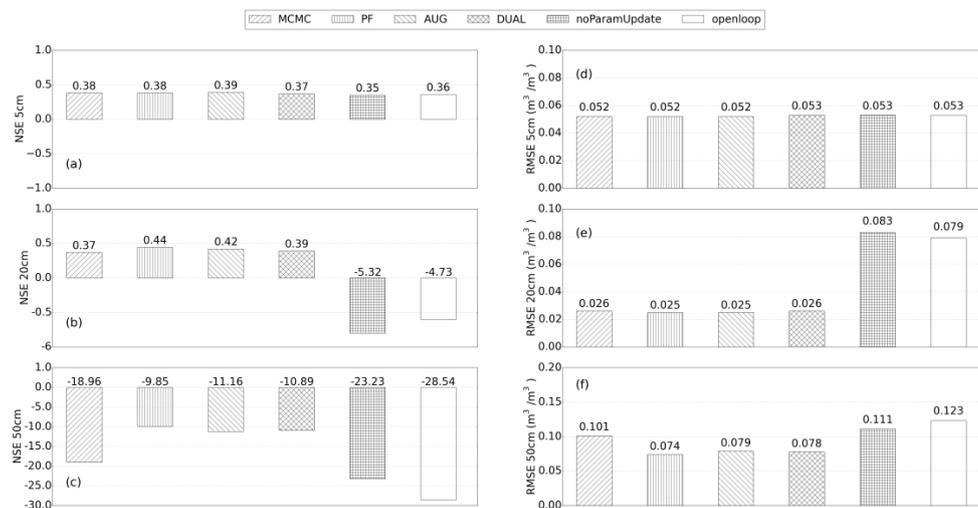
Figure 8 NSE and RMSE values for soil moisture content characterization in the verification period with the VIC-3L model: (a) NSE values for soil moisture content at 5cm, (b) NSE values at 20cm, (c) NSE values at 50cm, (d) RMSE values at 5cm, (e) RMSE values at 20cm, and (f) RMSE values at 50cm.
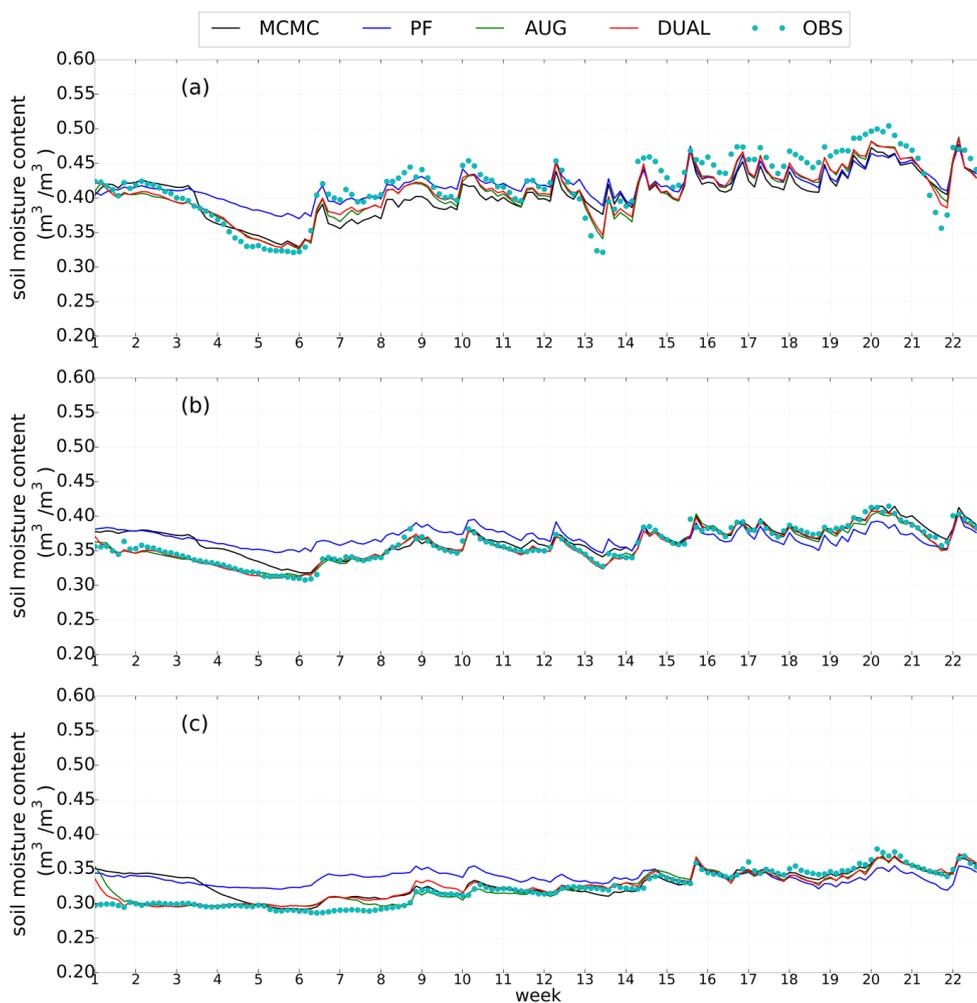
Figure 9 Time series of soil moisture content for different data assimilation scenarios during the assimilation

5 period and for the CLM model: (a) 5 cm depth, (b) 20 cm depth and (c) 50 cm depth. Week 1 starts with 01-03-

2012 and week 22 with 26-07-2012.
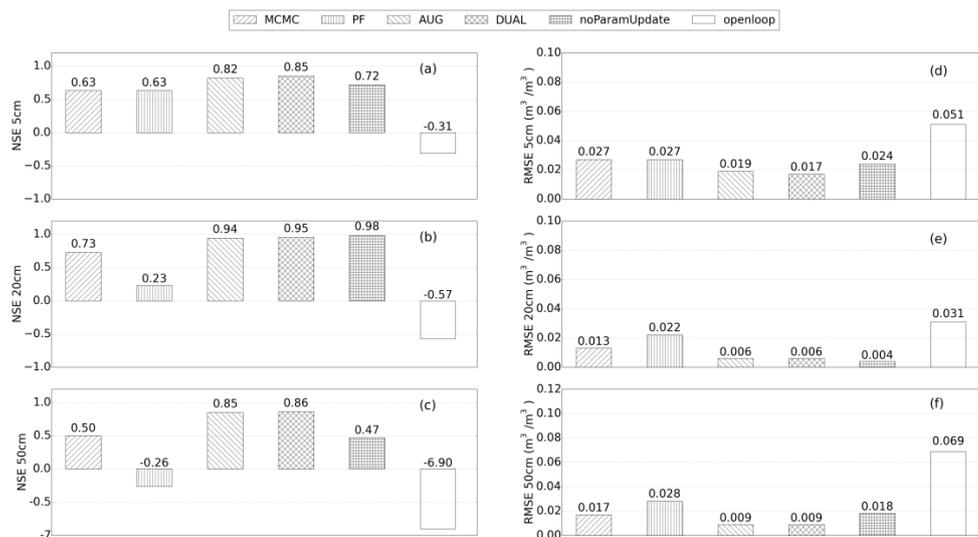
Figure 10 NSE and RMSE values for soil moisture content characterization in the assimilation period with the CLM model: (a) NSE values for soil moisture content at 5cm, (b) NSE values at 20cm, (c) NSE values at 50cm, (d) RMSE values at 5cm, (e) RMSE values at 20cm, and (f) RMSE values at 50cm.
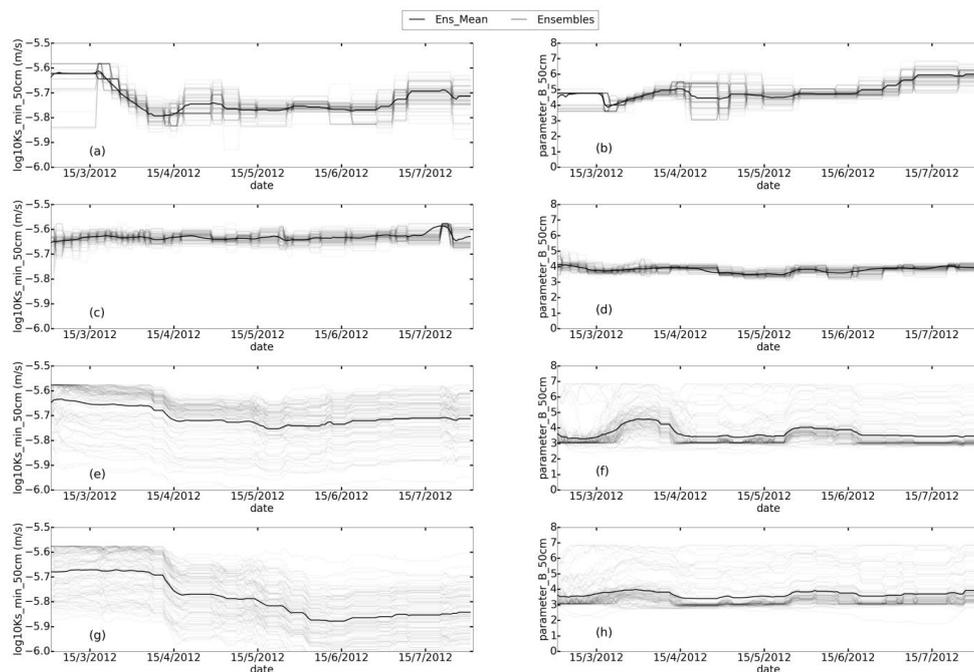
Figure 11 Temporal evolution of parameter values in the assimilation and parameter estimation period, for the four data assimilation scenarios and the CLM-model. Saturated hydraulic conductivity $\log_{10}K_s$ (m/s) at 50cm depth is displayed for the four methods: (a) MCMC, (c) PF, (e) AUG, and (g) DUAL. Soil hydraulic parameter B at 50cm depth for the four methods: (b) MCMC, (d) PF, (f) AUG and (h) DUAL. Displayed are temporal evolutions for 100 ensemble members.
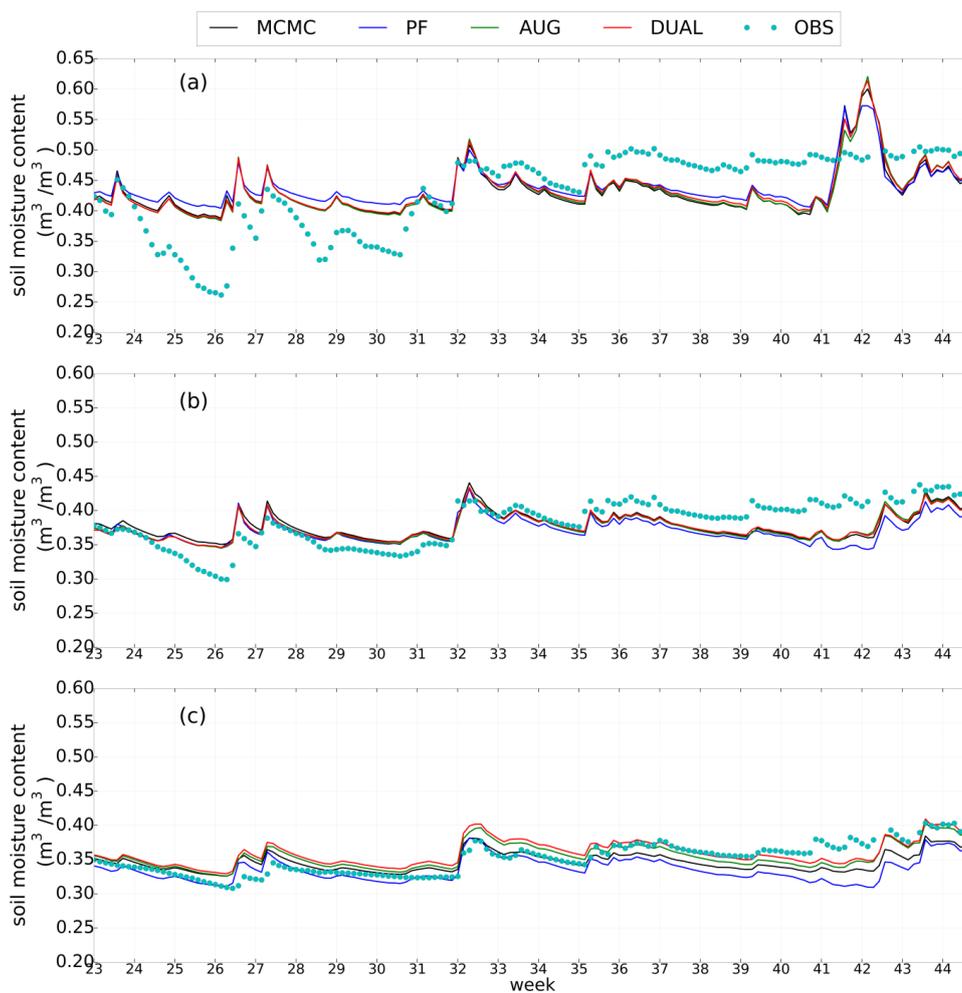
Figure 12 Time series of soil moisture content for different assimilation scenarios with CLM during the verification period: (a) 5 cm depth, (b) 20 cm depth and (c) 50 cm depth. Week 23 starts with 01-08-2012 and week 44 with 26-12-2012.
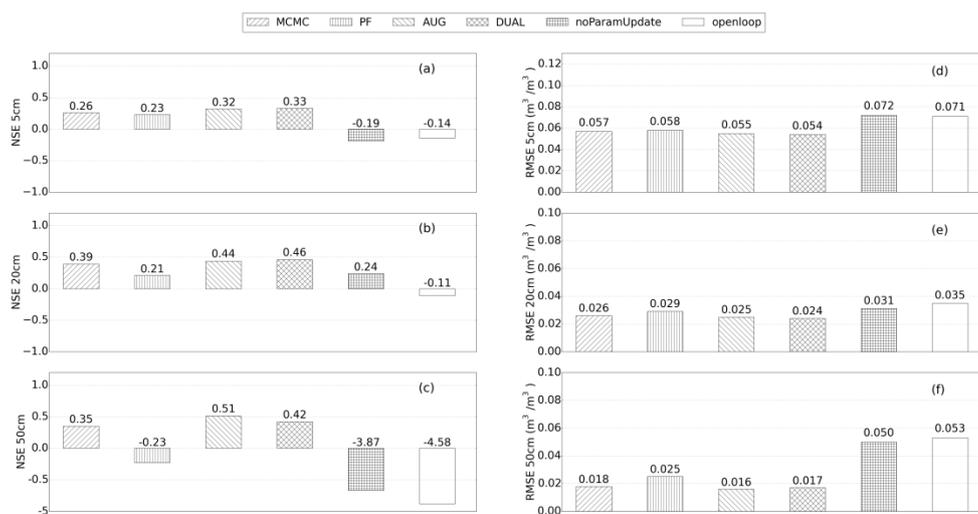
Figure 13 NSE and RMSE values for soil moisture content characterization in the verification period with the CLM model: (a) NSE values for soil moisture content at 5cm, (b) NSE values at 20cm, (c) NSE values at 50cm, (d) RMSE values at 5cm, (e) RMSE values at 20cm, and (f) RMSE values at 50cm.