Editor Decision: Reconsider after major revisions (15 Aug 2016) by Prof. Dr. Kurt Roth

Comments to the Author:

The paper applies different already reasonably mature data-assimilation (DA) techniques to the estimation of states and parameters of two land surface models (LSM). The problem is important, the methods appropriate, and the paper operationally instructive.

The key finding -- the differences between the LSMs are larger than the differences between the DAs (abstract, lines 23-25) -- corroborates the maturity of the DA schemes, which is an important result. It raises the question, however, if such a study is required at all and if so, if its focus is correct. Given the comparable results of the DAs, my interest is directed towards computational cost, ease of use, and range of applicability beyond these two LSMs. While the authors decided not to address this issue in depth, a paragraph on it is certainly warranted. Looking at the LSMs, the situation is much more difficult since addressing the implications and relevance of their differences, and possible remedies, is clearly beyond the scope of this manuscript. Still, statements like the paragraph on time-varying parameters that emerge when heterogeneity is neglected (reply to an issue raise by reviewer #2) has me wonder what such parameters would be good for. Such approaches are fine for process control, in a chemical production plant for instance, but are they of any use in an environmental context where we eventually want to predict something? Some clarifying statements in the revision appear to be in order.

With the above, the focus of this paper can only be on the methodological aspects of DA, illustrated for the difficult case of LSMs. This must be clearly worked out in the revised version.

Reply: We added in the discussion a paragraph that addresses other aspects of the use of data assimilation in combination with land surface models. This paragraph clarifies the computational cost, ease of use and range of applicability. The following text was added (line 640-654):

"*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 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. We expect that for example with more unknowns (i.e., 2D and 3D-applications) EnKF-based algorithms will perform better than PF, as PF will become extremely CPU-intensive and needs many more particles. For those high-dimensional applications EnKF is expected therefore to be more CPU-efficient than PF. Nevertheless, the small difference in performance between EnKF and PF based algorithms in this study indicates that PF is also an efficient data assimilation algorithm for problems of this size. It can be expected that larger ensemble numbers can improve the performance of EnKF and PF based algorithms. For MCMCPF, multiple MCMC resampling steps can also help improve performance. We expect that both data assimilation methods can relatively easily be used in combination with other land surface models and that the relative performance of the data assimilation methods would also be similar for those models*."

In addition, the discussion on time-varying parameters was also extended (line 620-639):

**"*Generally, parameters are time variant when jointly estimated with state variables as they are updated at each assimilation time step. Time-variant parameters might be dependent on the end of the training sequence, especially for parameters which are very sensitive to model forcings. The fact that we replace heterogeneous soil properties and soil moisture content for a given area by spatially homogeneous values, also introduces temporal variability in the effective parameters that are estimated in this study. In this context, it can be expected that estimated parameters show temporal*

***evolution. Uncertainties and errors in model forcings and model structural errors will introduce additional temporal fluctuation of estimated parameter values. In a batch calibration approach, these temporal parameter variations will be averaged out and parameters are estimated which on average perform better over the period of consideration. The advantage of sequential data assimilation is that parameter estimation is faster whereas temporal parameter variations in some cases are meaningful. Kurtz et al. (2012) were successful in estimating a temporal variable parameter with EnKF, but concluded that the algorithm needs time to adjust to new parameter values. Vrugt et al. (2013) found considerable temporal non-stationarity in parameters estimated by MCMCPF. In our study, this methodology also exhibited non-stationarity. However the other three methodologies in our study (Particle Filter, EnKF with augmentation and EnKF with dual estimation) did not show strong non-stationarity when estimating time-variant parameters. Especially for EnKF, parameters showed asymptotic properties at the end of assimilation period. Shi et al. (2015) also demonstrated the capability of EnKF in parameter estimation. For highly identifiable parameters, parameter uncertainty decreased and parameters converged fast. So we think that joint estimation of states and time-variant parameters with data assimilation still shows great potentials in terms of identifiability of parameters. In our study, we think that most parameters converge in 5 months of assimilation period.*"**

In addition, the numerous issues raised by the reviewers need to be addressed and discussed as confirmed in the authors' replies. Since the focus is on methods, the concepts must be more strict than in an application. What is for instance "fundamental" and what is not (reply to rev #1, pg 16 l 33) or how to measure distance in probability space (reply to rev #1, pg 16 l 1-2). [I do not want to become too philosophical at this point, just focus attention.]

(I should comment that the manuscript would have passed my desk with "minor revisions" if the authors had already included successful results/useful answers of/for their "will do in revision" replies.)

Reply: We repeated some experiments, added some new experiments, and revised the manuscript based on the comments from the two reviewers. See below.

As a very minor point I should like to comment on the authors reply to reviewer #1 concerning the SPADE sensor:

(i) The numbers given for the permittivity refer to the relative value, which is dimensionless (not F/m).

(ii) The key question is: What are the sensor's center frequency and frequency range for the observed range of soil water contents? This is relevant because the relative dielectric permittivity depends on frequency, and temperature for that matter, which may turn this particular setup quite complicated to interpret. (I could not check this aspect myself because I could not get an electronic version of the cited paper. In the context of this paper, this is a minor issue, however.)

Reply: Concerning the first point (i): we modified the unit in the revised manuscript. Concerning the second point (ii): the measurement frequency of the SPADE soil moisture sensor is approximately 150 MHz in water and 340 MHz in air, and there is no clear center frequency.

Reviewer I

The manuscript addresses an important issue in soil hydrology: The application of data assimilation methods to real world data, especially when estimating not only states, but also parameters. The authors state to address three main questions: (1) the performance of the data assimilation methods on Land Surface Models with real world data in general, (2) the differences in performance due to different data assimilation methods and (3) difference in performance due to different Land Surface Models. The study finds small differences due to data assimilation methods and large differences based on different Land Surface Models. These findings can give valuable insight for the applicability of data assimilation methods on Land Surface Models. But, this requires an adequate discussion of the used measurements, data assimilation methods and finally the results.

I have 3 major comments regarding each of these discussions, which fall short off answering the stated main questions enough. Additionally I have one major comment on the quality of the explanations for employed methods and models.

Reply: We thank the reviewer for pointing out the contribution of our work. We revised the manuscript taking into account the comments.

Page 10, Lines 20-23: "… 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." What kind of soil moisture sensors are installed? Please discuss possible uncertainties in the data.

Reply: The SPADE soil water content probes (sceme.de GmbH i.G., Horn-Bad Meinberg, Germany; (*Hübner et al.*, 2009)) were installed at 5 cm, 20 cm and 50 cm depth along a vertical profile. The SPADE probe is a ring oscillator and the frequency of the oscillator is a function of the dielectric permittivity of the surrounding medium, which is strongly dependent on the soil water content because of the high relative permittivity of water (≈80) as compared to mineral soil solids (≈2-9), and air (≈1). The SPADE probe was calibrated according to the procedure outlined in [*Qu et al.*, 2014]. The possible uncertainties in the soil moisture data are related to imperfect contact of the sensors with the soil, imperfection of the model which relates the sensor response and dielectric permittivity and imperfection of the model which relates dielectric permittivity and soil moisture. The measurement error is assumed to be 0.02cm^3/cm^3. We added the text in the revised version of the manuscript to clarify this (line 364-372):

"*The SPADE soil water content probes (sceme.de GmbH i.G., Horn-Bad Meinberg, Germany (Hübner et al., 2009)) were installed at 5 cm, 20 cm and 50 cm depth along a vertical profile. The SPADE probe is a ring oscillator and the frequency of the oscillator is a function of the dielectric permittivity of the surrounding medium, which is strongly dependent on the soil water content because of the high relative permittivity of water (≈80) as compared to mineral soil solids (≈2-9), and air (≈1). The SPADE probe was calibrated according to the procedure outlined in (Qu et al., 2014). The possible uncertainties in the soil moisture data are related to imperfect contact of the sensors with the soil, imperfection of the model which relates the sensor response and dielectric permittivity and imperfection of the model which relates dielectric permittivity and soil moisture.*"

Page 10, Lines 24-35: "In this work, the Rollesbroich site is modeled as a single point and the data of the soil sensor network are averaged to calculate areal averages of soil moisture content at 5cm, 20cm and 50cm depth." Please discuss the importance and implications of this assumption. What are the expected impacts of heterogeneity?

Reply: Data assimilation experiments with land surface models are generally conducted for large scales, especially when remote sensing data are assimilated. Therefore it is important to evaluate the model performance at a larger scale. The Rollesbroich site has an area of 0.27km^2, which is a very small catchment. *Qu et al.* (2014) described the statistics of soil properties for soil samples taken in the Rollesbroich catchment (see Table 1). We can see that soil texture shows a relatively limited variation. In our work only vertical heterogeneity is considered. In this case,

heterogeneity does not seem to be very strong and we do not face a special challenging upscaling case for the land surface model. This was added in the revised version of the manuscript (line 374-380):

"*Data assimilation experiments with land surface models are generally conducted for large scales, especially when remote sensing data are assimilated. Therefore it is important to evaluate the model performance at a larger scale. Qu et al. (2014) described the statistics of soil properties for soil samples taken in the Rollesbroich catchment. Soil texture showed a relatively limited variation. In our work only vertical heterogeneity is considered. In this case, heterogeneity does not seem to be very strong and we do not face a challenging upscaling case for the land surface model.*"

Table 1. Descriptive statistics of soil properties on the basis of 273 soil samples in the Rollesbroich catchment. Table is from *Qu et al.*, (2014).

		Clay %	Sand %	Silt %	Bulk Density (g/cm^3)	Carbon Content (g/kg)	Porosity (cm^3/cm^3)
5cm	mean	18.99	19.90	61.10	0.94	54.47	0.65
	std	2.00	3.82	3.79	0.12	15.82	0.05
20cm	mean	18.03	20.76	61.20	1.28	34.08	0.52
	std	1.99	4.03	3.46	0.15	16.84	0.05
50cm	mean	16.50	22.00	61.50	1.52	11.22	0.43
	std	2.40	5.68	4.53	0.16	6.01	0.06

Page 29, Figure 3: The figure shows higher water contents closer to the surface. Please mention this and discuss reasons and implications.

Reply: Text was added in the revised version of the manuscript (line 474-477):

"*The Rollesbroich catchment is a wet site with a yearly average precipitation around 1200mm. Regular precipitation events cause a wet surface layer. In addition, porosity of the upper soil layer is higher than for the deeper soil layers. This causes that during wet conditions soil moisture content is higher for the upper soil layer than for the deeper layer. It implies that at this site often we have a drainage flux from the top soil towards the aquifer (and drainage channels)*."

Page 11 Lines 34-36: "The soil moisture observation error is assumed to be normally distributed with mean equal to 0 and standard deviation equal to 0.02m^3/m^3, for both VIC-3L and CLM." Please discuss why you assume this uncertainty, especially since it is a mean of 41 values.

Reply: We admit that 0.02m^3/m^3 is a little larger than the uncertainty of the mean soil moisture content averaged over the 41 values. A larger observation error elevates potential problems with filter inbreeding. In addition, it adds flexibility in case of the presence of an observation bias or model structural error. This was added in the revised version of the manuscript (line 448-451):

"*We admit that 0.02m^3/m^3 is a little larger than the uncertainty of the mean soil moisture content averaged over the 41 values. A larger observation error elevates potential problems with filter inbreeding. In addition, it adds flexibility in case of the presence of an observation bias or model structural error.*"

Page 11 Lines 33-34: "Precipitation was perturbed were perturbed by multiplicative error N(1,0.1) to represent the uncertainty of measured precipitation at the site." Please give a reason for this error. What is the assumed error for evaporation?

Reply: In the Rollesbroich catchment, precipitation was measured by a tipping bucket. Therefore only a measurement error was assumed, which is typically around 10% of the measured value [*Hodgkinson et al.*, 2004]. In this work the variables which govern evapotranspiration (incoming shortwave and longwave radiation, air temperature, relative

humidity, wind speed), were not perturbed. This was added in the revised version of the manuscript (line 430-433):

"*In the Rollesbroich catchment, precipitation was measured by a tipping bucket. Therefore only a measurement error was assumed, which is typically around 10% of the measured value (Hodgkinson et al., 2004). In this work the variables which govern evapotranspiration (incoming shortwave and longwave radiation, air temperature, relative humidity, wind speed), were not perturbed*."

Page 2, Lines 13-15: "This approach allows for joint estimation of the states and parameters while taking into explicit consideration model structural error and forcing data errors (Liu and Gupta, 2007)." This is correct, but it is not to the point, since the authors later set the model error to zero (see Page 11, Lines 36-37) and hence do not consider model structural errors. Please discuss this.

Reply: Yes, model structural error is not considered in our work, but parameter uncertainties and forcing uncertainties are considered and we assume that these capture in this case the model uncertainty. However, we agree that it can be expected that we have other model structural errors, for example in relation to the representation of photosynthesis. We added the text in our manuscript (line 451-454):

"*The model error was set to zero assuming that uncertainty was captured by uncertain parameters and model forcings. However, we agree that it can be expected that we have other model structural errors, for example in relation to the representation of photosynthesis*."

Page 6, Line 25 (Eq. 25): Please mention that the way R (and y_t^i) is described, you assume uncorrelated measurement errors.

Reply: This information was added in the revised manuscript as suggested by the reviewer (line 219):

"*R is assumed to be uncorrelated.*"

Page 7, Line 25 (Eq. 33): Please describe the implications of employing this method. How does the performance of the filter depend on the choice of initial uncertainty?

Reply: Filter inbreeding is a problem associated with EnKF. The ensemble spread may narrow down in the course of parameter estimation so that most of the ensemble members would become very close to the ensemble mean value, which is called filter inbreeding and which might cause filter divergence [*Franssen and Kinzelbach*, 2008; *Han et al.*, 2014; *Whitaker and Hamill*, 2012]. The approach according Eq. 33 has been proven to be an efficient method to avoid filter inbreeding [*Han et al.*, 2014; *Whitaker and Hamill*, 2012]. In the revised manuscript, we used VIC model to test the effect of initial uncertainties on the performance of EnKF by increasing the forcing error from 10% to 20%. The additional simulation results were explained in the manuscript (line 549-552):

"*The effect of initial uncertainties on the performance of EnKF with the ensemble inflation method is also tested with the VIC-3L model. The forcing error was increased from 10% to 20%. Table 6 shows the RMSE values for soil moisture content characterization in the assimilation and verification periods. The difference between the results for 10% or 20% perturbation of the forcings is very limited, for both variants of the EnKF-method.*"

Page 9, Line 13. How did you choose the tuning parameter s? How does it influence the performance? Please include the choice of s for the PF and initial uncertainties of the EnKF when comparing different methods.

Reply: The optimal tuning parameter *s* is hardly known in applications [*Yan et al.*, 2015]. It was set to 0.01 in some applications [*DeChant and Moradkhani*, 2012; *Plaza et al.*, 2012]. In our work, to keep particle spread, *s* was set to 0.1. We tested other values for parameter *s*, like 0.01 and 0.5,

to see how it influences the performance. The additional simulation results were explained in the manuscript (line 553-571):

"*The optimal tuning parameter s in RRPF is hardly known in applications (Yan et al., 2015). It was set to 0.01 in some applications (DeChant and Moradkhani, 2012; Plaza et al., 2012). In our work, to keep particle spread, s was set to 0.1. We also tested other values for parameter s, like 0.01 and 0.5, to see how it influences the performance. Table 7 shows the RMSE values for soil moisture content characterization in the assimilation and verification periods for RRPF. In the assimilation period, PF_0.01 performs the worst and PF_0.5 performs the best, especially for the third layer. This result is expected. Larger spread of parameter values results in a larger spread of state values, and larger spread of state values is more likely to cover the true value. From Tab. 4, we can see that the open loop run deviates strongly from the measurement values for the third model layer. If all model simulations are far away from the observation, measurements cannot correct the simulations towards the measured values. Figure 8 shows the temporal evolution of the parameters log10Ks and β for the third layer and the three RRPF scenarios during the assimilation period. Figure 9 shows the corresponding temporal evolution of soil moisture content. Severe particle degeneration happens in PF_0.01 which results in its bad performance in the third layer. In PF_0.1 particle degeneration also happens from March to June and explains its bad performance from March to June in Fig. 3. The spread of parameter members in PF_0.5 is very large but this may also be a disadvantage for parameter convergence. Table 7 illustrates that in the verification period the difference among the three simulation variants with different perturbation factors is limited for the first and second model layers. For the third layer, PF_0.01 still performs the worst, and PF_0.1 and PF_0.5 perform similarly. So neither too small nor too large parameter perturbation is desirable, and therefore s was set to 0.1 in our work.*"

Page 15 Line 32: You state: "It is not surprising that the EnKF is more efficient and effective than the PF." I would follow this statement in case of strictly Gaussian distributions and linear measurement operators. In those cases the EnKF is expected to outperform the PF. Nevertheless, when dealing with non-linear processes that challenge the Gaussian assumption, the better performance is not clear at all.

<u>Reply:</u> We added an explanation already in the manuscript why we think EnKF outperforms PF, even although the PF is in theory more suited for non-linear processes and non-Gaussian statistics. We re-evaluated this part and tried to formulate more precisely (line 655-674):

**"*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 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 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 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 data assimilation schemes if*

***one adopts the update rule of the EnKF during particle resampling with MCMC [Vrugt et al., 2013].*"**

Page 16 Line 33: You state that EnKF and PF "differ fundamentally in their analysis step". I would disagree and argue that the analysis step is similar. The only difference is that the EnKF updates it's posterior based on the Gaussian assumption of the distributions, while the PF drops this assumption.

Reply: We agree that there is of course also a similarity in the updating step, but nevertheless we still think that EnKF and PF differ fundamentally in their analysis step. In EnKF, the difference of the forecasted output variable(s) and corresponding observed value(s) (the part (y-Hx)) is used directly to update states and parameters. However, in PF, the likelihood is calculated and the individual particles are not corrected towards the states, but only weighted differently in correspondence with the likelihood.

Page 25, Table 1: The EnKF assumes Gaussian distributions. What is the reason that you sample for all but one parameter from an initial uniform distribution? What are the implications for the chosen inflation method?

Reply: We want to compare EnKF and PF starting from the same prior distribution in order to make a more meaningful comparison. It is right that EnKF assumes a Gaussian distribution, but the PF not. We believe that assuming an initial uniform distribution is a neutral assumption good for comparing EnKF and PF. This was clarified in the revised version of the manuscript (line 434-437):

"*We want to compare EnKF and PF starting from the same prior distribution in order to make a more meaningful comparison. EnKF assumes a Gaussian distribution, but the PF not. We believe that assuming an initial uniform distribution is a neutral assumption good for comparing EnKF and PF*."

Page 16, Lines 1-2: You state: "The value of the likelihood does generally not say anything about how close the forecasted variables are to their measured counterparts." I disagree, the likelihood yields information about the distance.

Reply: The likelihood yields information about the relative, but not the absolute distance between forecasted variables and measurements. We can say that particles with higher weights are closer to the measurements but we cannot tell how close they are to the measurements.

Discussion of results: Since the results show a wealth of information. I would appreciate a detailed discussion. Especially consider the following points in detail and incorporate them into your conclusion:

Page 29, Figure 3: The large deviations in the Particle Filter might hint at filter inbreeding in the states. You observe the deviation but do not discuss it and actually exclude the possibility of inbreeding by investigating the parameters: "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." (Page 13 Lines 15-19). Please address the question of adequate ensemble spread in the states by actually showing the ensemble of states there.

Reply: We showed the evolution of the state ensemble of PF in our revised manuscript and addressed this question (line 563-567):

"*Figure 9 shows the corresponding temporal evolution of soil moisture content. Severe particle degeneration happens in PF_0.01 which results in its bad performance in the third layer. In PF_0.1 particle degeneration also happens from March to June and explains its bad performance from*

***March to June in Fig. 3. The spread of parameter members in PF_0.5 is very large but this may also be a disadvantage for parameter convergence.*"**

Page 30 Figure 4: During the calibration period the filter without parameter estimation performs better. Please discuss possible reasons.

Reply: When only states (soil moisture content) are assimilated, states are updated directly by observations. However, when states and parameters are updated jointly, the nonlinear relation between states and parameters is considered, which may introduce inconsistency. We further discuss these results in the paper (line 493-495):

"*When only states (soil moisture content) are assimilated, states are updated directly by observations. However, when states and parameters are updated jointly, the nonlinear relation between states and parameters is relevant, which may introduce inconsistencies*."

Page 31 Figure 5: Parameter b estimated by MCMC shows a large difference to the other methods. But MCMC does perform approximately as well as the other filters (Figure 7, Page 33). Why is there no difference? Please discuss.

Reply: *Demaria et al.* (2007) evaluated the sensitivity and identifiability of ten parameters which control surface and subsurface runoff in the VIC model for four U.S. watersheds along a hydroclimatic gradient. They found that parameter b is crucial in a dry environment, while its impact on model performance is not significant for wet sites. They concluded that parameter b plays a key role in partitioning rainfall into soil moisture and surface runoff in dry environments. [*Liang and Guo*, 2003] and [*Atkinson et al.*, 2002] reached a similar conclusion. In our work, as the Rollesbroich catchment is very wet, even though parameter b estimated by MCMC shows a large difference with other methods, it shows small impact on the soil moisture content for layer 1 and layer 2. In the revised manuscript, all experiments of VIC-3L were done again. Evolution of parameter b estimated by MCMC was more reasonable (figure 4).

Page 32 Figure 6: Initial parameter uncertainties are the same for PF and EnKF but at time 0 the ensemble spreads are different. Please explain. You only show the two parameters with the least change over time. Give a reason or show the one with the smallest and the one with the largest changes.

Reply: In original manuscript figure 6 shows the evolution of the parameter ensemble from time step 1 but not time step 0. At time step 0, the ensemble spreads are the same, but at time step 1, the parameter ensemble is updated by PF or EnKF, and the ensemble spreads between EnKF and PF differ. We showed the evolution of the parameter ensemble from time step 0 onwards in the revised version of the manuscript (figure 5, figure 8 and figure 11). We think that the saturated hydraulic conductivity $\log_{10}k_s$ and the model parameter β are the two most important parameters for the VIC Model for the Rollesbroich catchment, so they are shown in the figure.

Page 34, Figure 8: There is basically no difference for the prediction of the water content whether there are parameters estimated or not. Please discuss why this is the case.

Reply: Predictions of soil moisture content for layer 2 and layer 3 (in the verification period) improved significantly for the case of parameter estimation. Concerning the soil moisture content of layer 1, the RMSE value of the open loop run is $0.053 m^3/m^3$, which is already quite close to the observed values. In addition, the soil moisture content for the upper layer is strongly driven by single precipitation events. We extended the discussion of these results (line 533-536):

"*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, the four methods show a RMSE improvement of about 54% and 42% for the second and third model layer (compared with the open loop run).*"

Page 37 Figure 11: You do not show the parameters from the augmented state. Instead you show parameters derived from those in a non-linear way. Please discuss this.

Reply: We showed on purpose the soil hydraulic parameters k_s and B, which are in CLM calculated from soil texture. We believe that displaying these parameters is more meaningful than soil texture.

Figures 7, 8, 12, 13: Especially water content of the top layer can almost not be represented by either model, although improved with state of the art data assimilation methods. Please discuss this including the representation of the physics in the models and implications of the perfect model assumption.

Reply: We provided additional discussion of the results and discuss reasons for the larger deviations in the fit (line 675-682):

"*In the verification period soil moisture of the top layer cannot be represented perfectly by the two LSM´s, in spite of parameter updating with state of the art data assimilation methods. Table 5 and table 9 illustrate that the RMSE values of the four joint state and parameter assimilation methods are similar for both models, which means that both models have larger errors for the top layer. There is a number of reasons for the larger soil moisture mismatches for the upper layer: (i) the memory effect from initial conditions, very well identified at the beginning of the verification period is smaller for the upper soil layer, as this layer is more affected by precipitation events and evaporation; (ii) these soil moisture changes make that it is also more affected by model structural errors, for example concerning evaporation processes.*"

The difference in the assimilation methods is small. Please discuss if this difference is significant. Do you expect the same results for other applications? What is the influence of specific filter settings on the performance?

Reply: The performance of the four data assimilation algorithms in which states and parameters are jointly estimated does not differ very much in our study. Nevertheless, the small difference in performance between EnKF and PF based algorithms indicates that PF is also an efficient data assimilation algorithm for problems of this size. It can be expected that larger ensemble numbers can improve the performance of EnKF and PF based algorithms. For MCMCPF, multiple MCMC resampling steps can also help improve performance. Given the CPU-intensity of the calculations a larger comparison is beyond the scope of this work. In our work, all algorithms are evaluated with real-world data, so we think our results are meaningful for other applications. We expect that for example with more unknowns (i.e., 2D and 3D-applications) EnKF-based algorithms will perform better than PF, as PF will become extremely CPU-intensive and needs many more particles. We evaluated the impact of filter settings on the results of this paper, like the impact of initial conditions on the inflation method for EnKF, the effect of tuning parameter *s* on PF and how the ending date of assimilation affects the verified results. However, it will be difficult to evaluate the significance of the difference between the DA-algorithms on the basis of this single study.

We extent the discussion (line 640-654):

"***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 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. We expect that for example with more unknowns (i.e., 2D and 3D-applications) EnKF-based algorithms will perform better than PF, as PF will become extremely CPU-intensive and needs many more particles. For those high-dimensional applications EnKF is expected therefore to be more CPU-efficient than PF. Nevertheless, the small difference in performance between EnKF and PF based algorithms in this study indicates that PF is also an efficient data assimilation algorithm for problems of this size.***"

Please improve the explanations on the Particle Filter and the two LSMs. Page 7 Line 32 - Page 8 Line 11: I do understand Particle Filters, but the given explanation is not clear. Especially clarify your description of transition and proposal densities.

<u>Reply:</u> We revised this part in our manuscript (line 261-285):

"***The particle filter was first suggested in the research area of object recognition, robotics and target tracking (Gordon et al., 1993). 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.***

Based on the recursive Bayes Law, the posterior PDF of state variables at time t given the observations y_t is:

$$p(x_t|y_{1:t})=\frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{\int p(y_t|x_t)p(x_t|y_{1:t-1})dx_t} \quad (34)$$

where $p(y_t|x_t)$ is the likelihood function for time step t and $p(x_t|y_{1:t-1})$ is the prior PDF. The prior PDF is in a data assimilation framework typically obtained from the predicted model states (probably including parameters), before data assimilation.

The likelihood $p(y_t|x_t^i)$ is considered to be Gaussian:

$$p(y_t|x_t^i)=\frac{exp(-\frac{1}{2}(y_t-Hx_t^i)^T R^{-1}(y_t-Hx_t^i))}{(2\pi)^{m/2}|R|^{1/2}} \quad (35)$$

where R is the measurement error covariance matrix, $|R|$ is the determinant of matrix R and m is the dimension of vector y.

The posterior PDF is approximated by the PF according to:

$$p(x_t|y_{1:t}) \approx \sum_{i=1}^{N} w_t^i \delta(x_t-x_t^i) \quad (36)$$

where x_t^i is assumed to be the i^{th} state sample (in our case soil moisture) drawn from the posterior PDF $p(x_t|y_{1:t})$ with weight w_t^i and δ is the Dirac delta function. For the sequential updating case, the recursive weight update equation is defined:

$$w_t^i = w_{t-1}^i p(y_t|x_t^i) \quad (37)$$

The normalized weights for the particles are given by:

$$\widetilde{w}_t^i = \frac{w_t^i}{\sum_{i=1}^N w_t^i} \quad (38)$$

The state estimated from the N particles is given by:

$$x_t = \sum_{i=1}^N \widetilde{w}_t^i x_t^i \quad (39)"$$

Page 8 Lines 21-30: You describe SIR and RR. What is the reason to choose RR?

Reply: RR is developed from SIR by [*Liu and Chen*, 1998]. RR is one of the most popular methods for PF to reduce particle degeneration. For RR, the variance of particles is smaller than the one given by the SIR scheme. Moreover, RR is computationally cheaper than SIR. This explanation was added in the revised version of the manuscript (line 298-299):

"*The particles have more similar weights than the particles in SIR (Weerts and Serafy, 2006). Moreover, RR is computationally cheaper than SIR*."

Page 7 Lines 31-32: "The particle filter was first suggested in the research area of object recognition, robotics and target tracking (Arulampalam et al., 2002)." The PF was actually mentioned earlier (Gordon et al., 1993).

Reply: This was corrected in the revised version.

For this study the understanding of the different LSM models is important, since the main difference in performance is attributed to different models. Because of that a good description is necessary. Although I am not an expert on LSMs, I noticed the following: Page 17, Line 15 (Eq. A2): The equation describes the soil water movement in the top two layers. To me it was not clear if this description is valid for both layers individually. If so, why is the precipitation P and evaporation E the same?

Reply: This was corrected in the revised version (line 723-730):

"

$$\frac{\partial \theta_1}{\partial t} z_1 = P - Q_d - Q_{1,2} - E_1 \quad (A2)$$

$$\frac{\partial \theta_2}{\partial t} z_2 = Q_{1,2} - Q_{2,3} - E_2 \quad (A3)$$

$$\frac{\partial \theta_3}{\partial t} z_3 = Q_{2,3} - E_3 - Q_b \quad (A4)$$

where θ [L^3L^{-3}] is volumetric soil moisture content, z_i [L] is soil depth for layer i (i =1,2,3), $Q_{i,i+1}$ [LT^{-1}] is the vertical drainage between layer i and i+1, Q_d [LT^{-1}] is calculated for layer 1. 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 zero in Eq. (A3 and A4) because there is no evaporation from

***layer 2 and 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.* "**

Page 18, Line 1 (Eq. A7): The dimensions are inconsistent: You add [LT-1], [] and [L]. i_m should be I_m (or is not introduced). Please also explain the distinction of the cases $P+I < I_m$ and $P+I > I_m$.

Reply: i_m is I_m. When $P+I > I_m$, the upper soil layers (layer 1 and layer 2) will be saturated. When $P+I < I_m$, the upper soil layers are assumed unsaturated, and infiltration capacity is variable, as clarified in Eq. (A9). We revised this part in our manuscript (line 741-749):

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

$$Q_d = \begin{cases} P - \left(\theta_1^{max} - z_1\theta_1\right) - \left(\theta_2^{max} - z_2\theta_2\right) + \left(\theta_1^{max} + \theta_2^{max}\right)\left(1 - \frac{Iv+P}{I_m}\right)^{1+b}, & P+Iv \le I_m \\ P - \left(\theta_1^{max} - z_1\,\theta_1\right) - \left(\theta_2^{max} - z_2\,\theta_2\right), & P+Iv > I_m \end{cases} \quad (A8)$$

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. $z_1\phi_1$). When $P+Iv > I_m$, the upper soil layers will be saturated and when $P+Iv \le I_m$, the upper soil layers are assumed unsaturated, and infiltration capacity Iv [L] is variable which is a function of the maximum filtration capacity I_m [L] [Zhao, 1992]:

$$Iv = I_m\left(1 - (1 - A)^{\frac{1}{b}}\right) \text{ with } I_m = (1 + b)\left(\theta_1^{max} + \theta_2^{max}\right) \quad (A9)$$"

Page 19, Lines 18-22: the Richards equation is formulated for a continuum. Please explain the modifications and the implications of applying it to layers. Please give the extent of these layers.

Reply: in the revised manuscript, table 3 was added to show the extent of the 10 layers, and text about the Richards equation was revised (line 789-801):

"***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:***

$$\frac{\partial\theta_i}{\partial t} = \frac{\partial}{\partial z}\left[k_i\left(\frac{\partial(\psi_i + z_i - C)}{\partial z}\right)\right] - E = \frac{\partial}{\partial z}\left[k_i\left(\frac{\partial(\psi_i - \psi_{E,i})}{\partial z}\right)\right] - E \quad with \quad C = \psi_{E,i} + z_i \quad (B7)$$

$$\psi_{E,i} = \psi_{sat,i}\left(\frac{\theta_E(z_i)}{\theta_{sat,i}}\right)^{-B_i} \quad with \quad \theta_E(z_i) = \theta_{sat,i}\left(\frac{\psi_{sat,i} + z_\nabla - z_i}{\psi_{sat,i}}\right)^{-\frac{1}{B_i}} \quad (B8)$$

where ψ_E [L] is the equilibrium soil matric potential, z_∇ is water table depth [L] 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. Further details about Eq. (B7) can be found in the CLM-manual (Oleson et al., 2013). General Richards equation used a "θ" -based solution which cannot account for the variation of ψ below water table because "θ" is constant (at saturated value) while ψ varies temporally and spatially, which leads to the failure to maintain the hydrostatic equilibrium soil moisture distribution. However, the modified Richards equation in which a constant hydraulic potential C is explicitly subtracted at each time step can fix this deficiency. Details about the implementation of the modified method are given in (Zeng and Decker, 2009)."

More explanation related to this question:

The unmodified Richards equation is:

The soil water flux *q* can be described by Darcy's law:

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

For one-dimensional vertical water flow in soils, the conservation of mass is stated as:

$$\frac{\partial\theta}{\partial t} = -\frac{\partial q}{\partial z} - E = \frac{\partial}{\partial z}\left[k\left(\frac{\partial(\psi+z)}{\partial z}\right)\right] - E \quad (2)$$

***E* is the ET loss. *Zeng and Decker* (2009) note that this equation cannot maintain the hydrostatic equilibrium soil moisture distribution because of the truncation errors of the finite-difference numerical scheme. They show that this deficiency can be overcome by subtracting the equilibrium state as:**

$$\frac{\partial\theta}{\partial t} = \frac{\partial}{\partial z}\left[k\left(\frac{\partial(\psi+z-C)}{\partial z}\right)\right] - E \quad (3)$$

Where C is a constant hydraulic potential above the water table z_∇:

$$C=\psi_E + z \quad (4)$$

Substitution of equation (4) into equation (3) yields the modified Richards equation (B7):

$$q=-k\left(\frac{\partial(\psi-\psi_E)}{\partial z}\right) \quad \text{and} \quad \frac{\partial\theta}{\partial t} = -\frac{\partial q}{\partial z} - E = \frac{\partial}{\partial z}\left[k\left(\frac{\partial(\psi-\psi_E)}{\partial z}\right)\right] - E \quad (5)$$

Implication of this modified method: Richards equation (2) used a θ-based solution which cannot account for the variation of ψ below water table because θ is constant (at saturated value) while ψ varies temporally and spatially, which leads to the failure to maintain the hydrostatic equilibrium soil moisture distribution. However, the modified Richards equation in which constant hydraulic potential C is explicitly subtracted at each time step can fix this deficiency. Details about the implementation of the modified method can be seen in [*Zeng and Decker*, 2009].

Below we explain the application of the modifications to layers. These details are not presented in the revised manuscript and we refer to the CLM-manual (Oleson et al., 2013) where more details can be found. Table 2 shows the soil layer definition in CLM for N_{levsoi} =10 layers.

Table 2. Division of a soil column in layers in CLM. Layer node depth(z), thickness(Δz), and depth at layer interface(z_h) for 10 soil layers. Unit is meter.

Layer *i*	*z*	Δz	z_h
1 (top)	0.0071	0.0175	0.0175
2	0.0279	0.0276	0.0451
3	0.0623	0.0455	0.0906
4	0.1189	0.0750	0.1655
5	0.2122	0.1236	0.2891
6	0.3661	0.2038	0.4929
7	0.6198	0.3360	0.8289
8	1.0380	0.5539	1.3828

9	**1.7276**	**0.9133**	**2.2961**
10	**2.8646**	**1.5058**	**3.8019**

Numerical solution of equation (5) for soil layer *i*:

$$\Delta z_i \frac{\Delta\theta_i}{\Delta t} = -q_{i-1}^{t+1} + q_i^{t+1} - e_i \qquad (6)$$

where Δt is the time step, Δz_i is soil layer thickness, $\Delta\theta_i = \theta_i^{t+1} - \theta_i^t$, q_i is the outgoing flux of water from layer *i* to layer *i*+1, q_{i-1} is the incoming flux of water from layer *i*-1 to layer *i* and e_i is layer-averaged ET loss. The water fluxes (q_{i-1}^{t+1} and q_i^{t+1}) in equation (6) are linearized about θ using Taylor series expansion which results in a general tridiagonal equation set of the form [*Oleson et al.*, 2013]:

$$r_i = a_i \Delta\theta_{i-1} + b_i \Delta\theta_i + c_i \Delta\theta_{i+1} \qquad (7)$$

a_i, b_i, c_i and r_i will be discussed under different conditions below. This tridiagonal equation set is solved over *i*=1,…, N_{levsoi}+1 where *i*= N_{levsoi}+1 is a virtual layer representing the aquifer.

For *i*=1, the boundary condition is the infiltration rate, and

$$a_i = 0$$

$$b_i = \frac{\partial q_i}{\partial \theta_i} - \frac{\Delta z_i}{\Delta t}$$

$$c_i = \frac{\partial q_i}{\partial \theta_{i+1}}$$

$$r_i = q_{infl}^{t+1} - q_i^t + e_i$$

q_{infl}^{t+1} is the infiltration into the soil which is partitioned between water input flux (sum of precipitation reaching the ground and melt water from snow), surface runoff, and surface water storage.

For *i*=2, …, N_{levsoi} - 1,

$$a_i = -\frac{\partial q_{i-1}}{\partial \theta_{i-1}}$$

$$b_i = \frac{\partial q_i}{\partial \theta_i} - \frac{\partial q_{i-1}}{\partial \theta_i} - \frac{\Delta z_i}{\Delta t}$$

$$c_i = \frac{\partial q_i}{\partial \theta_{i+1}}$$

$$r_i = q_{i-1}^t - q_i^t + e_i$$

For the lowest soil layer (*i* = N_{levsoi}), the bottom boundary condition depends on the depth of the water table. If the water table is within the soil column, a zero flux boundary condition is applied ($q_i^t = 0$) and

$$a_i = -\frac{\partial q_{i-1}}{\partial \theta_{i-1}}$$

$b_i = -\frac{\partial q_{i-1}}{\partial \theta_i} - \frac{\Delta z_i}{\Delta t}$

$c_i = 0$

$r_i = q_{i-1}^t + e_i$

And for the aquifer layer $i = N_{levsoi} + 1$:

$a_i = 0 \quad b_i = -\frac{\Delta z_i}{\Delta t} \quad c_i = 0 \quad r_i = 0$

If water table is below the soil column, for $i = N_{levsoi}$:

$a_i = -\frac{\partial q_{i-1}}{\partial \theta_{i-1}}$

$b_i = \frac{\partial q_i}{\partial \theta_i} - \frac{\partial q_{i-1}}{\partial \theta_i} - \frac{\Delta z_i}{\Delta t}$

$c_i = \frac{\partial q_i}{\partial \theta_{i+1}}$

$r_i = q_{i-1}^t - q_i^t + e_i$

And for the aquifer layer $i = N_{levsoi} + 1$:

$a_i = -\frac{\partial q_{i-1}}{\partial \theta_{i-1}}$

$b_i = -\frac{\partial q_{i-1}}{\partial \theta_i} - \frac{\Delta z_i}{\Delta t}$

$c_i = 0$

$r_i = q_{i-1}^t$

Upon solution of the tridiagonal equation set, soil water content is updated as:

$$\theta_i^{t+1} = \theta_i^t + \Delta\theta_i \Delta z_i \qquad (8)$$

Specific comments: Page 5, Line 36: "Commonly used data assimilation algorithms are EnKF, PF and variants of them." 4D-Var is also commonly used.

<u>Reply:</u> We admit that 4D-Var is also commonly used. The sentence was adapted to account for this, and we specified shortly in which contexts the different methods are normally used (line 192-195):

"*Commonly used data assimilation algorithms are four-dimensional variational method (4D-Var), EnKF, PF and variants of them. All these algorithms are successfully applied for the atmospheric, oceanic, biogeochemistry and hydrologic assimilation systems. In hydrology the EnKF, PF and their variants are most frequently used*."

Page 6, Line 19: "H […] is the identity matrix if y refers to in-situ ground measurements available at all grid cells." This is only the case if the same quantity as the state is observed. Otherwise the quantity has to be transferred. Additionally mention that H has to be linear for the EnKF.

Reply: We revised this part in our manuscript (line 213-215):

"*H is an observation operator that connects measurements and model states, it should be linear for EnKF and it is the identity matrix if y refers to in-situ ground measurements available at all grid cells and if the same variable as the state are observed*."

Page 6-7: You do not mention the use of a damping factor (Hendricks Franssen and Kinzelbach, 2008) for the parameters. Is there a specific reason you do not employ it?

Reply: The filter inbreeding problem could be reduced with help of a damping factor, which limits the intensity of the perturbation of the parameters [*Franssen and Kinzelbach*, 2008]. In our work, the inflation algorithm proposed by *Whitaker and Hamill* (2012) was applied to the ensemble of parameters to reduce filter inbreeding.

Page 25, Table 1: It is not clear which parameters are estimated for each layer individually and which are estimated for the entire profile.

Reply: We clarified this in table 1.

Page 11 Lines 11-12: "The parameters of the other layers were updated with help of the calculated spatial covariance in case of EnKF." This statement implies that the parameters were not estimated with the PF? Please clarify.

Reply: Parameters are also estimated in PF. In PF, each particle includes a state vector and a parameter vector. But the weight vector is calculated only by the state vector. Both the state vector and parameter vector are resampled with help of the weight vector. We added the text to make this clearer (line 404-406):

"*In PF, parameters are resampled with help of the weight vector which is calculated for each particle, and therefore linked to both states and parameters associated to the particle*."

Page 11 Line 30: "EnKF with state updating only was tested for (2)." Why did you not test state updating with the PF?

Reply: In our work, parameters are perturbed to generate particles for PF, which means each particle is associated with given states and parameters. After calculating the weight vector for the state vector, the resampling is (automatically) applied for both state vector and parameter vector. In this context, it is not possible to only update states. If we would have deterministic parameters, it would be possible to only update states with PF, but then not a comparison with EnKF could be made.

Page 12 Lines 2-10: Why do you need 2 different characterizations of the uncertainty. What is the additional gain by showing NSE?

Reply: Because RMSE values are usually affected by both a mean bias and random variations, NSE is added as another measure [*Han et al.*, 2012]. NSE values represent the correlation between the estimation and the observation.

Page 12 Line 9-10: "A NSE value equal to 1 and RMSE equal to 0 imply a perfect prediction." This is wrong. A RMSE equal to the measurement uncertainty is perfect.

Reply: Thanks, we admit that for a prediction in the verification phase we cannot expect a better result than a RMSE equal to the measurement uncertainty but still an RMSE equal to 0 would be the best result. We modified the sentence to accommodate this result (line 464-465):

"*A larger NSE value and smaller RMSE value imply a better prediction.*"

All Technical corrections are applied in our revised manuscript.

Page 18, Line 25: "soil matric potential ψ_i [L]" is actually the matric head. In soil physics the matric head is typically described with h_m, while potentials are described with ψ.

Reply: Here we mean soil matric potential not the matric head.

Reference

Atkinson, S. E., R. A. Woods, and M. Sivapalan (2002), Climate and landscape controls on water balance model complexity over changing timescales, *Water Resources Research, 38*(12), doi:10.1029/2002wr001487.

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.

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.

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.

Han, X., X. Li, H. J. H. Franssen, H. Vereecken, and C. Montzka (2012), Spatial horizontal correlation characteristics in the land data assimilation of soil moisture, *Hydrol. Earth Syst. Sci., 16*(5), 1349-1363, doi:10.5194/hess-16-1349-2012.

Hodgkinson R. A., T. J. Pepper, and D. W. Wilson (2004), Evaluation of tipping bucket rain gauge performance and data quality, Environment Agency, Rio House, Waterside Drive, Aztec West, Almondsbury, Bristol, BS32 4UD, ISBN:1844323242.

Hübner, C., R. Cardell-Oliver, R. Becker, K. Spohrer, K. Jotter, and T. Wagenknecht (2009), Wireless soil moisture sensor networks for environmental monitoring and vineyard irrigation: Helsinki University of Technology, 1, 408-415.

Liang, X., and J. Z. Guo (2003), Intercomparison of land-surface parameterization schemes: sensitivity of surface energy and water fluxes to model parameters, *Journal of Hydrology, 279*(1-4), 182-209, doi:10.1016/s0022-1694(03)00168-9.

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.

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.

Plaza, D. A., R. De Keyser, G. J. M. De 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.

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.

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.

Yan, H., C. M. DeChant, and H. Moradkhani (2015), Improving Soil Moisture Profile Prediction With the Particle Filter-Markov Chain Monte Carlo Method, *Ieee Transactions on Geoscience and Remote Sensing*, *53*(11), 6134-6147, doi:10.1109/tgrs.2015.2432067.

Zeng, X., and M. Decker (2009), Improving the numerical solution of soil moisture-based Richards equation for land models with a deep or shallow water table, *Journal of Hydrometeorology*, *10*, 308-319, doi:10.1175/2008JHM1011.1.

Reviewer II

The manuscript presents different data assimilation methods for a joint estimation of soil moisture states and model parameters for the VIC hydraulic model and the Community Land Model. The models were tuned and evaluated at a single site and the main objectives include the advantages of the joint state and parameter estimation incorporating real data from the field, performance of the DA methods as well as the different land surface schemes.

The topic is interesting for the scientific community and the paper is clearly structured and well written. I agree for the most part with referee #1, who emphasizes shortcomings with respect to the presentation and discussion of the given objectives, to which I will add only a few more comments.

Reply: We thank the reviewer for pointing out the contribution of our work. We will revise the manuscript taking into account the comments.

Furthermore, I have a comment regarding the usefulness of the presented data assimilation techniques for land surface modeller. In my opinion, the merits of a joint state and parameter estimation should not only be discussed with respect to DA schemes updating states only, but also with respect to alternative methods like conventional Bayesian interference, e.g. (Yang et al., 2008), as well as the issue of optimizing time-invariant parameter vs. time-variant parameter, which has been intensely studied in the group of the authors (Vrugt et al., 2005, 2013). Therefore a discussion of the following (which might be too obvious for the authors to mention) can help to increase the significance for a broader community: What are the advantages of the joint parameter estimation versus optimizing time-invariant parameter? There seem to be shortcomings as time-variant parameter may be highly dependent on the end of the training sequence, especially when it ended shortly after a large precipitation event, like in this study. Will the parameter converge in the given training data set of 5 months? Vrugt et al. (2013) show that time-variant parameter can exhibit considerable non-stationarity, which is caused by changing sensitivity of the target variable on the parameters. Is there a difference/advantage of the joint estimation with time-variant parameter in terms of equifinality/identifiability of the parameter?

Reply: Yes, we agree that there are other relevant methods for parameter estimation/calibration of hydrologic models, for example Bayesian recursive estimation [Thiemann et al., 2001], particle swarm optimization [Scheerlinck et al., 2009] and differential evolution adaptive metropolis [Vrugt and Ter Braak, 2011]. However these methods require in general a large number of model evolutions, which is often prohibitive for large scale land surface models. We refer therefore in the revised version of the manuscript shortly to alternative methods and point to the limitations of those methods. We modified the text in the revised version of the manuscript to clarify this (line 48-53):

"*Many methods for parameter estimation/calibration of hydrologic models are proposed, for example Bayesian recursive estimation (Thiemann et al., 2001), particle swarm optimization (Scheerlinck et al., 2009) and differential evolution adaptive metropolis (Vrugt and Ter Braak, 2011). However these methods require in general a large number of model evolutions, which is often prohibitive for large scale land surface models, and other uncertainty sources like forcings (e.g. precipitation and air temperature) are not considered*."

Generally, parameters are time variant when jointly estimated with state variables as they are updated at each assimilation time step. It is true that time-variant parameters may be dependent on the end of the training sequence, especially for the parameters which are very sensitive to model forcings. The fact that we replace heterogeneous soil properties and soil moisture content for a given area by spatially homogeneous values, also introduces temporal variability in the effective parameters that are estimated in this study. In this context, it can be expected that estimated parameters show temporal evolution. Uncertainties and errors in model forcings and model structural errors will introduce additional temporal fluctuation of estimated parameter values. In a batch calibration approach, these temporal parameter variations will be averaged

out and parameters are estimated which on average perform better over the period of consideration. The advantage of sequential data assimilation is that parameter estimation is faster whereas temporal parameter variations in some cases are meaningful. Kurtz et al. (2012) were successful in estimating a temporal variable parameter with EnKF, but concluded that the algorithm needs time to adjust to new parameter values. Vrugt et al. (2013) found considerable temporal non-stationarity in parameters estimated by MCMCPF. This was discussed in the revised version of the manuscript (line 620-639):

"***Generally, parameters are time variant when jointly estimated with state variables as they are updated at each assimilation time step. Time-variant parameters might be dependent on the end of the training sequence, especially for parameters which are very sensitive to model forcings. The fact that we replace heterogeneous soil properties and soil moisture content for a given area by spatially homogeneous values, also introduces temporal variability in the effective parameters that are estimated in this study. In this context, it can be expected that estimated parameters show temporal evolution. Uncertainties and errors in model forcings and model structural errors will introduce additional temporal fluctuation of estimated parameter values. In a batch calibration approach, these temporal parameter variations will be averaged out and parameters are estimated which on average perform better over the period of consideration. The advantage of sequential data assimilation is that parameter estimation is faster whereas temporal parameter variations in some cases are meaningful. Kurtz et al. (2012) were successful in estimating a temporal variable parameter with EnKF, but concluded that the algorithm needs time to adjust to new parameter values. Vrugt et al. (2013) found considerable temporal non-stationarity in parameters estimated by MCMCPF. In our study, this methodology also exhibited non-stationarity. However the other three methodologies in our study (Particle Filter, EnKF with augmentation and EnKF with dual estimation) did not show strong non-stationarity when estimating time-variant parameters. Especially for EnKF, parameters showed asymptotic properties at the end of assimilation period. Shi et al., (2015) also demonstrated the capability of EnKF in parameter estimation. For highly identifiable parameters, parameter uncertainty decreased and parameters converged fast. So we think that joint estimation of states and time-variant parameters with data assimilation still shows great potentials in terms of identifiability of parameters. In our study, we think that most parameters converge in 5 months of assimilation period.***"

P.10, ll.28ff: For me there seems to be no need to show the spin-up time series (Figure 2). Precipitation and temperature of the assimilation and verification period seem to be enough.

<u>Reply:</u> We revised this figure in our revised manuscript.

P.11: What is the reason of choosing July 31 as the date to switch from assimilation to verification period? This choice seems to be critical for me, as the parameters of the final time step are chosen for the verification period. What would be the impact, if e.g. July 20 would have been chosen, as Figure 5 suggests, that some parameter for the MCMCPF method were significantly different?

<u>Reply:</u> For 2013, there are issues with a large number of sensors in the area and the mean soil moisture content would have to be estimated from fewer (and different) sensors. We started the assimilation in March 2012 as in the winter before soil moisture content readings were affected by soil freezing and therefore unreliable (at least in February). We tested the impact of the choice of the last assimilation day on the parameter estimation with the MCMCPF method (line 542-548):

"***To address this issue, we tested different ending dates of the assimilation period for MCMCPF: June 11 2012, June 30 2012, July 20 2012, and July 31 2012, which are indicated by MCMC_0611, MCMC_0630, MCMC_0720 and MCMC_0731 respectively in Fig. 7. The only difference between the assimilation scenarios is the assimilation ending date. Figure 7 shows the soil moisture time series from August 1 2012 (verification period) for the 4 scenarios. We can see that MCMC_0611 differs strongly from the other scenarios whereas the differences among MCMC_0630, MCMC_0720 and MCMC_0731 are limited, although parameters showed temporal varability.***"

This issue is also described in the discussion in the revised manuscript (line 620-639):

"*Generally, parameters are time variant when jointly estimated with state variables as they are updated at each assimilation time step. Time-variant parameters might be dependent on the end of the training sequence, especially for parameters which are very sensitive to model forcings. The fact that we replace heterogeneous soil properties and soil moisture content for a given area by spatially homogeneous values, also introduces temporal variability in the effective parameters that are estimated in this study. In this context, it can be expected that estimated parameters show temporal evolution. Uncertainties and errors in model forcings and model structural errors will introduce additional temporal fluctuation of estimated parameter values. In a batch calibration approach, these temporal parameter variations will be averaged out and parameters are estimated which on average perform better over the period of consideration. The advantage of sequential data assimilation is that parameter estimation is faster whereas temporal parameter variations in some cases are meaningful. Kurtz et al. (2012) were successful in estimating a temporal variable parameter with EnKF, but concluded that the algorithm needs time to adjust to new parameter values. Vrugt et al. (2013) found considerable temporal non-stationarity in parameters estimated by MCMCPF. In our study, this methodology also exhibited non-stationarity. However the other three methodologies in our study (Particle Filter, EnKF with augmentation and EnKF with dual estimation) did not show strong non-stationarity when estimating time-variant parameters. Especially for EnKF, parameters showed asymptotic properties at the end of assimilation period. Shi et al., (2015) also demonstrated the capability of EnKF in parameter estimation. For highly identifiable parameters, parameter uncertainty decreased and parameters converged fast. So we think that joint estimation of states and time-variant parameters with data assimilation still shows great potentials in terms of identifiability of parameters. In our study, we think that most parameters converge in 5 months of assimilation period.*"

P.11: state updating only: How does the model then learn for the verification period? How are the parameter chosen in this case? Please describe this more clearly.

Reply: when only the state is updated in the assimilation period, the model gets more accurate initial state conditions in the verification period. We would indeed expect that an improved characterization of initial states has some positive impact during the first weeks, but vanishes over time. We address this issue in the revised manuscript (line 418-420):

"*When only the state is updated in the assimilation period, the model gets more accurate initial state conditions in the verification period. We would expect that an improved characterization of initial states has some positive impact during the first weeks, but vanishes over time*."

P.11, ll.34ff, Table 1: soil moisture observation errors and parameter perturbations are given by normal and uniform distributions and corresponding ranges, means and standard deviations are given with numbers without further reasoning. As a comprehensive set of soil moisture measurements and soil core data is available, I would assume, the range of perturbation is related to the measured distributions, but I did not see a hint in the text. Referee #1 already addressed this issue related to measurement uncertainty and spatial heterogeneity, and the authors gave detailed reply, but I still miss, how the prior distributions and measurement uncertainties are related to the measured pdfs. It is surprising for me, that the uncertainty of the soil moisture measurements related to spatial heterogeneity is smaller than the given instrument uncertainty of 0.02m^3/m^3.

Reply: We addressed this issue in the revised manuscript (line 437-447):

**"*For the CLM model parameters, the parameter perturbations were taken from Han et al. (2014), and for the model parameter perturbations for VIC, we refer to Demaria et al. (2007) and Troy et al. (2008). Also measurements were available at the Rollesbroich site to estimate parameter uncertainty like soil texture measurements. If we calculate the uncertainty of the mean soil texture based on those data, we get very small uncertainties. The range of parameter perturbations should be large enough to create enough spread among the ensemble members, which helps for better assimilation performance. In this case, the uncertainty has to be increased in order to fit the data. This is related*

***to the fact that ultimately soil hydraulic parameters, and not soil texture, are important for calculating water and energy fluxes in the soil. The pedotransfer functions which are used to relate soil texture and soil hydraulic parameters are also subject to uncertainty. We therefore did not directly use the uncertainty of the soil texture estimated from the measurements, but increased it*.”**

Qu et al. (2014) calculated the root mean square error (RMSE) associated with soil water content estimation, which is 0.026 m^3/m^3, after the two-step calibration procedure for this catchment. The uncertainty of the mean soil moisture content, assuming a Gaussian distribution, from 41 measurements in this catchment was $\frac{0.026}{\sqrt{41}}$ m^3/m^3. However, in our study we used 0.02m^3/m^3 as observation error, because a larger observation error elevates problems with filter inbreeding. We found also in this case that the small measurement error estimated from the data was too small for our purposes. We address this issue in the revised manuscript (line 448-451):

“***We admit that 0.02m3/m3 is a little larger than the uncertainty of the mean soil moisture content averaged over the 41 values. A larger observation error elevates potential problems with filter inbreeding. In addition, it adds flexibility in case of the presence of an observation bias or model structural error*.”**

P 13. ll3-5: You state: "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". This hints at the issue of equifinality and identifiability of the parameters with respect to the time series to be predicted. Please discuss this problem.

Reply: Equifinality is handled by both methods because not a single best solution is calculated but an ensemble of different solutions, which are all compatible with the measurement data. The ensemble mean values are plotted. The updating of the parameters follows for both methods the same general tendency. However, as the reviewer stresses, the ensemble mean values also differ for the two assimilation methods. We believe that in this case differences are related to the assimilation methods. The land surface model is ran twice in EnKF in case of dual estimation but only once for the augmentation approach. Model structure errors and biases “contribute” to different extents to parameter updating by these two data assimilation methods. Therefore the temporal evolution of parameter values is different. We addressed this issue in the revised manuscript (line 501-507):

“***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 is different. Nevertheless, the updating of the AUG and DUAL parameters still follows the same general tendency. We believe that in this case differences are related to the assimilation methods. The land surface model is ran twice for EnKF with dual estimation but only once for the augmentation approach. Model structural errors and biases “contribute” to different extents to parameter updating by these two data assimilation methods. Therefore the temporal evolution of parameter values is different*.”**

Figures 3,7,9,12: Legend: "OBS" were coded with 2 dots. Please make use of different line types for a better discrimination between the displayed series. Especially red and green will be indistinguishable for many readers

Reply: As the lines are so close in these figures that different line types are indistinguishable. We change the line color to avoid red and green. We plot the figure by python which shows the legend of “OBS” with 2 dots automatically.

Reference

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.

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.

Kurtz, W., H. J. H. Franssen, and H. Vereecken (2012), Identification of time-variant river bed properties with the ensemble Kalman filter, Water Resource Research, 48, W10534, doi:10.1029/2011WR011743.

Qu, W., H. R. Bogena, J. A. Huisman, G. Martinez, Y. A. Pachepsky, and H. Vereecken (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.

Scheerlinck, K., V. R. N. Pauwels, H. Vernieuwe, and B. De Baets (2009), Calibration of a water and energy balance model: Recursive parameter estimation versus particle swarm optimization, Water Resource Research, 45, W10422, doi:10.1029/2009WR008051.

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.

Thiemann, M., M. Trosset, H. Gupta, and S. Sorooshian (2001), Bayesian recursive parameter estimation for hydrologic models, Water Resource Research, 37(10), 2521–2535.

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.

Vrugt, J. A., and C. J. F. Ter Braak (2011), DREAM(D): An adaptive Markov Chain Monte Carlo simulation algorithm to solve discrete, noncontinuous, and combinatorial posterior parameter estimation problems, Hydrology Earth System Sciences, 15(12), 3701–3713.

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.

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~~these model 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~~is 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 functions. (Vereecken et al., 2016). 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 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). Many methods for parameter estimation/calibration of hydrologic models are proposed, for example Bayesian recursive estimation (Thiemann et al., 2001), particle swarm optimization (Scheerlinck et al., 2009) and differential evolution adaptive metropolis (Vrugt and Ter Braak, 2011). However these methods require in general a large number of model evolutions, which is often prohibitive for large scale land surface models, and other uncertainty sources like forcings (e.g. precipitation and air temperature) are not considered. 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 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; Lee, et al., 2014), 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. 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 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), 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 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 (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 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 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 algorithms which ~~are~~were 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~~were 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 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., 2013). A Metropolis ratio ~~is~~was 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.

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.

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 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 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.

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.

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 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 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

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 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 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 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 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.

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 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

Data assimilation algorithms combine observations and model predictions together and update model states and parameters. Commonly used data assimilation algorithms are four-dimensional variational method (4D-Var),

EnKF, PF and variants of them. All these algorithms are successfully applied for the atmospheric, oceanic, biogeochemistry and hydrologic assimilation systems. In hydrology the EnKF, PF and their variants are most frequently used.

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):

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

where i refers to the i^{th} 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}~~ v_t 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:

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

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:

$$y_t^i = y_t + w_t^i \quad (23)$$

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:

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

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

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

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. R is assumed to be uncorrelated. Finally, the Kalman gain K is calculated by:

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

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:

$$x^i = \begin{bmatrix} s^i \\ p^i \end{bmatrix} \quad (27)$$

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

$$H^* = [H_s, H_p] \quad (29)$$

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:

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

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

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

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 according to:

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

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.

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 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:

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

where $\overline{p_t}$ is the ensemble mean for a parameter p_t 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.

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~~Gordon et al., ~~2002~~1993). 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.

T~~he~~ Based on the recursive Bayes Law, the posterior PDF of state variables at time t given the observations y_t is:

$$p(x_t|y_{1:t}) = \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{\int p(y_t|x_t)p(x_t|y_{1:t-1})dx_t} \quad (34)$$

where $p(y_t|x_t)$ is the likelihood function for time step t and $p(x_t|y_{1:t-1})$ is the prior PDF. The prior PDF is in a data assimilation framework typically obtained from the predicted model states (probably including parameters), before data assimilation.

The likelihood $p(y_t|x_t^i)$ is considered to be Gaussian:

$$p(y_t|x_t^i) = \frac{\exp(-\frac{1}{2}(y_t - Hx_t^i)^T R^{-1}(y_t - Hx_t^i))}{(2\pi)^{m/2}|R|^{1/2}} \quad (35)$$

where **R** is the measurement error covariance matrix, |**R**| is the determinant of matrix **R** and m is the dimension of vector ***y***.

The posterior PDF is approximated by the PF according to:

$$p(x_t|y_{1:t}) \approx \sum_{i=1}^{N} w_t^i \delta(x_t - x_t^i) \quad (36)$$

~~where~~ x_t^i is assumed to be the i^{th} state sample (in our case soil moisture) drawn from the posterior ~~pdf~~PDF $p(x_t|y_{1:t})$ 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:~~

~~$$w_t^i \propto \frac{p(x_t^i|y_t)}{q(x_t^i|y_t)} \quad (35)$$~~

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

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

~~$= w_{t-1}^i p(y_t|x_t^i)$~~The state estimated from the N particles is given by:

~~$$x_t = \sum_{i=1}^{N} w_t^i x_t^i \quad (37)$$~~

The normalized weights for the particles are given by:

$$\widetilde{w}_t^i = \frac{w_t^i}{\sum_{i=1}^{N} w_t^i} \quad (38)$$

The state estimated from the N particles is given by:

$$x_t = \sum_{i=1}^{N} \tilde{w}_t^i \, x_t^i \qquad (39)$$

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:

$$N_{eff} = \frac{1}{\sum_{i=1}^{N} (\tilde{w}_t^i)^2} \qquad (38 \; (40)$$

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) $\hat{N}_i=[Nw_t^i]$, 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 $w_t^j=(Nw_t^i-\hat{N}_i)/N_j$. The ~~variance of the~~ particles have more similar weights than the particles ~~is smaller than the variance given by~~in SIR (Weerts and Serafy, 2006). Moreover, RR is computationally cheaper than SIR. 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 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

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

- Assign weights

  $w_t^i = 1/N$

END IF

Where s is a small tuning parameter and σ^2_{prior} is the prior variance for parameter p. ~~s was 0.1 in our work.~~ The optimal tuning parameter s is hardly known in applications [Yan et al., 2015]. In our work, to keep particle spread, s was set to 0.1. We also test other values for parameter s with VIC-3L model, like 0.01 and 0.5, to see how it influences the performance.

3.3 Markov Chain Monte Carlo PF (MCMCPF)

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. 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.

$$\alpha = \min\left(1, \frac{p\left(x^{pro}_{i,t-1}\middle|x^{i}_{t-2}\right)p\left(y^{i}_{t-1}\middle|x^{pro}_{i,t-1}, p^{pro}_{i,t}\right)p\left(y^{i}_{t}\middle|x^{pro}_{i,t}, p^{pro}_{i,t}\right)}{p\left(x^{i}_{t-1}\middle|x^{i}_{t-2}\right)p\left(y^{i}_{t-1}\middle|x^{i}_{t-1}, p^{i}_{t}\right)p\left(y^{i}_{t}\middle|x^{i}_{t}, p^{i}_{t}\right)}\right) \quad (39 \text{ (42)}$$

where $x^{pro}_{i,t}$ is the i^{th} proposed state sampled from the proposal state distribution at time step t, $p^{pro}_{i,t}$ is the i^{th} proposed parameter sampled from the proposal parameter distribution at time step t, and y^{i}_{t} represents the i^{th} observation at time step t. The proposed state-parameter combination is accepted if ($\alpha > U(0,1)$) 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 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^{pro}_{t-1} based on x_{t-1}

  Create proposal p^{pro}_{t} based on $\hat{p}_t$

  Simulate proposal x^{pro}_{t} based on proposal x^{pro}_{t-1} and proposal p^{pro}_{t} using model

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

  Calculate the accept index vector j

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

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

- Assign weights

  $w^{i}_{t} = 1/N$

END IF

4. Case study

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 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.~~The SPADE soil water content probes (sceme.de GmbH i.G., Horn-Bad Meinberg, Germany (Hübner et al., 2009)) were installed at 5 cm, 20 cm and 50 cm depth along a vertical profile. The SPADE probe is a ring oscillator and the frequency of the oscillator is a function of the dielectric permittivity of the surrounding medium, which is strongly dependent on the soil water content because of the high relative permittivity of water (≈80) as compared to mineral soil solids (≈2-9), and air (≈1). The SPADE probe was calibrated according to the procedure outlined in (Qu et al., 2014). The possible uncertainties in the soil moisture data are related to imperfect contact of the sensors with the soil, imperfection of the model which relates the sensor response and dielectric permittivity and imperfection of the model which relates dielectric permittivity and soil moisture. 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 averaged to calculate areal averages of soil moisture content at 5cm, 20cm and 50cm depth. Data assimilation experiments with land surface models are generally conducted for large scales, especially when remote sensing data are assimilated. Therefore it is important to evaluate the model performance at a larger scale. Qu et al. (2014) described the statistics of soil properties for soil samples taken in the Rollesbroich catchment. Soil texture showed a relatively limited variation. In our work only vertical heterogeneity is considered. In this case, heterogeneity does not seem to be very strong and we do not face a challenging upscaling case for the land surface model. 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~~during the ~~years 2011 and 2012~~simulation period. Soil texture was determined for the area based on 273 soil samples, 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. We started the assimilation in March 2012 as in the winter before soil moisture content readings were affected by soil freezing and therefore unreliable (at least in February). For 2013, there were issues with a large number of sensors in the area and the mean soil moisture content would have to be estimated from fewer (and different) sensors. So our experiments ended in December 2012.

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. In PF, parameters are resampled with help of the weight vector which is calculated for each particle, and therefore linked to both states and parameters associated to the particle.

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. When only the state is updated in the assimilation period, the model gets more accurate initial state conditions in the verification period. We would expect that an improved characterization of initial states has some positive impact during the first weeks, but vanishes over time.

(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.~~In the Rollesbroich catchment, precipitation was measured by a tipping bucket. Therefore only a measurement error was assumed, which is typically around 10% of the measured value (Hodgkinson et al., 2004). In this work the variables which govern evapotranspiration (incoming shortwave and longwave radiation, air temperature, relative humidity, wind speed), were not perturbed. Soil parameters were perturbed as in Table 1. Most

parameters were sampled from an initial uniform distribution. We want to compare EnKF and PF starting from the same prior distribution in order to make a more meaningful comparison. EnKF assumes a Gaussian distribution, but the PF not. We believe that assuming an initial uniform distribution is a neutral assumption good for comparing EnKF and PF. For the CLM model parameters, the parameter perturbations were taken from Han et al. (2014), and for the model parameter perturbations for VIC, we refer to Demaria et al. (2007) and Troy et al. (2008). Also measurements were available at the Rollesbroich site to estimate parameter uncertainty like soil texture measurements. If we calculate the uncertainty of the mean soil texture based on those data, we get very small uncertainties. The range of parameter perturbations should be large enough to create enough spread among the ensemble members, which helps for better assimilation performance. In this case, the uncertainty has to be increased in order to fit the data. This is related to the fact that ultimately soil hydraulic parameters, and not soil texture, are important for calculating water and energy fluxes in the soil. The pedotransfer functions which are used to relate soil texture and soil hydraulic parameters are also subject to uncertainty. We therefore did not directly use the uncertainty of the soil texture estimated from the measurements, but increased it. The soil moisture observation error is assumed to be normally distributed with mean equal to 0 and standard deviation equal to $0.02 m^3/m^3$, for both VIC-3L and CLM. We admit that $0.02 m^3/m^3$ is a little larger than the uncertainty of the mean soil moisture content averaged over the 41 values. A larger observation error elevates potential problems with filter inbreeding. In addition, it adds flexibility in case of the presence of an observation bias or model structural error. The model error was set to zero assuming that uncertainty was captured by uncertain parameters and model forcings. However, we agree that it can be expected that we have other model structural errors, for example in relation to the representation of photosynthesis. 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~~were used to evaluate the performance of different scenarios: the Nash-Sutcliffe model efficiency (NSE) coefficient and the Root Mean Square Error (RMSE):

$$NSE = 1 - \frac{\sum_{t=1}^{T} (\theta_t^{sim} - \theta_t^{obs})^2}{\sum_{t=1}^{T} (\theta_t^{obs} - \frac{1}{T}\sum_{t=1}^{T} \theta_t^{obs})^2} \quad (~~40~~43)$$

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T} (\theta_t^{sim} - \theta_t^{obs})^2} \quad (~~41~~ (44)$$

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 larger NSE value and smaller RMSE value imply a better 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

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~~layers 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. The figure shows higher soil moisture closer to the surface. The Rollesbroich catchment is a wet site with a yearly average precipitation around 1200mm. Regular precipitation events cause a wet surface layer. In addition, porosity of the upper soil layer is higher than for the deeper soil layers. This causes that during wet conditions soil moisture content is higher for the upper soil layer than for the deeper layer. It implies that at this site often we have a drainage flux from the top soil towards the aquifer (and drainage channels). 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 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.~~The large deviations in the RRPF might hint at filter inbreeding in the states (see also Fig. 9(b)). More details will be discussed below. EnKF results in better simulation results for the third layer (both for state augmentation and dual estimation) compared to the RRPF and MCMCPF.

FigureTable 4 shows the NSE and RMSE values of soil moisture content for the assimilation period and all scenarios. The open loop deviates most from the measured values, but if states are updated RMSE values are reduced by ~~68%, 82~~58%, 76% and ~~95~~94% 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~~two approaches 44% for the first layer, and ~~88% both~~81% and 89% for the second and third layer respectively. 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~~. When only states (soil moisture content) are assimilated, states are updated directly by observations. However, when states and parameters are updated jointly, the nonlinear relation between states and parameters is relevant, which may introduce inconsistencies. 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.

Figure ~~5~~4 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~~is different. Nevertheless, the updating of the AUG and DUAL parameters still ~~follow~~follows the same general tendency. We believe that in this case differences are related to the assimilation methods. The land surface model is ran twice for EnKF with dual estimation but only once for the augmentation approach. Model structural errors and biases “contribute” to different extents to parameter updating by these two data assimilation methods. Therefore the temporal evolution of parameter values is different. 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~~4 shows the maximum baseflow velocity D_m in the third layer, which is a key parameter to calculate the baseflow. The time series of D_m for the EnKF-algorithms show a fast decrease in the first month and a stable tendency afterwards, whereas the D_m 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~~5 shows the temporal evolution of the parameters $\log_{10}K_s$ 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~~2).

Figure ~~7~~6 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.~~limited differences over this period. 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. Results of the scenario noParamUpdate are close to that of the open loop run after several days which are not shown in the figure.

The NSE and RMSE values for soil moisture characterization in the verification period and the three soil layers are plotted in ~~Fig. 8~~Tab. 5. 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~~the four methods show a RMSE improvement of ~~68~~about 54% and ~~36~~42% 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.

In Fig. 4 MCMCPF shows much larger temporal fluctuations than the other three methods because of its jumping mechanism. However the parameters become more stable towards the end of the assimilation period. One could argue that the choice of the last assimilation day on the parameter estimation with the MCMCPF method may have an impact on the results in the verification period. To address this issue, we tested different ending dates of the assimilation period for MCMCPF: June 11 2012, June 30 2012, July 20 2012, and July 31 2012, which are indicated by MCMC_0611, MCMC_0630, MCMC_0720 and MCMC_0731 respectively in Fig.

7. The only difference between the assimilation scenarios is the assimilation ending date. Figure 7 shows the soil moisture time series from August 1 2012 (verification period) for the 4 scenarios. We can see that MCMC_0611 differs strongly from the other scenarios whereas the differences among MCMC_0630, MCMC_0720 and MCMC_0731 are limited, although parameters showed temporal varability.

The effect of initial uncertainties on the performance of EnKF with the ensemble inflation method is also tested with the VIC-3L model. The forcing error was increased from 10% to 20%. Table 6 shows the RMSE values for soil moisture content characterization in the assimilation and verification periods. The difference between the results for 10% or 20% perturbation of the forcings is very limited, for both variants of the EnKF-method.

The optimal tuning parameter s in RRPF is hardly known in applications (Yan et al., 2015). It was set to 0.01 in some applications (DeChant and Moradkhani, 2012; Plaza et al., 2012). In our work, to keep particle spread, s was set to 0.1. We also tested other values for parameter s, like 0.01 and 0.5, to see how it influences the performance. Table 7 shows the RMSE values for soil moisture content characterization in the assimilation and verification periods for RRPF. In the assimilation period, PF_0.01 performs the worst and PF_0.5 performs the best, especially for the third layer. This result is expected. Larger spread of parameter values results in a larger spread of state values, and larger spread of state values is more likely to cover the true value. From Tab. 4, we can see that the open loop run deviates strongly from the measurement values for the third model layer. If all model simulations are far away from the observation, measurements cannot correct the simulations towards the measured values. Figure 8 shows the temporal evolution of the parameters $\log_{10}$Ks and β for the third layer and the three RRPF scenarios during the assimilation period. Figure 9 shows the corresponding temporal evolution of soil moisture content. Severe particle degeneration happens in PF_0.01 which results in its bad performance in the third layer. In PF_0.1 particle degeneration also happens from March to June and explains its bad performance from March to June in Fig. 3. The spread of parameter members in PF_0.5 is very large but this may also be a disadvantage for parameter convergence. Table 7 illustrates that in the verification period the difference among the three simulation variants with different perturbation factors is limited for the first and second model layers. For the third layer, PF_0.01 still performs the worst, and PF_0.1 and PF_0.5 perform similarly. So neither too small nor too large parameter perturbation is desirable, and therefore s was set to 0.1 in our work.

4.3b Results for CLM

Figure ~~9~~10 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 RRPF and MCMCPF for all the three model 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~~Table 8 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 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%). Concerning the soil moisture content of layer 1, the RMSE value of the open loop run is

0.053m^3/m^3, which is already quite close to the observed values. In addition, the soil moisture content for the upper layer is strongly driven by single precipitation events.

Figure 11 displays the ensemble of the temporal evolutions of ~~log10ks~~$\log_{10}$Ks and the soil hydraulic parameter B at 50cm depth during the assimilation (calibration) period. Overall, changes in parameter values are small and 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 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~~Results of the scenario noParamUpdate are close to that of the open loop run after several days which are not shown in the figure. Table 9 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 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 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 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 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.

Generally, parameters are time variant when jointly estimated with state variables as they are updated at each assimilation time step. Time-variant parameters might be dependent on the end of the training sequence,

especially for parameters which are very sensitive to model forcings. The fact that we replace heterogeneous soil properties and soil moisture content for a given area by spatially homogeneous values, also introduces temporal variability in the effective parameters that are estimated in this study. In this context, it can be expected that estimated parameters show temporal evolution. Uncertainties and errors in model forcings and model structural errors will introduce additional temporal fluctuation of estimated parameter values. In a batch calibration approach, these temporal parameter variations will be averaged out and parameters are estimated which on average perform better over the period of consideration. The advantage of sequential data assimilation is that parameter estimation is faster whereas temporal parameter variations in some cases are meaningful. Kurtz et al. (2012) were successful in estimating a temporal variable parameter with EnKF, but concluded that the algorithm needs time to adjust to new parameter values. Vrugt et al. (2013) found considerable temporal non-stationarity in parameters estimated by MCMCPF. In our study, this methodology also exhibited non-stationarity. However the other three methodologies in our study (Particle Filter, EnKF with augmentation and EnKF with dual estimation) did not show strong non-stationarity when estimating time-variant parameters. Especially for EnKF, parameters showed asymptotic properties at the end of assimilation period. Shi et al. (2015) also demonstrated the capability of EnKF in parameter estimation. For highly identifiable parameters, parameter uncertainty decreased and parameters converged fast. So we think that joint estimation of states and time-variant parameters with data assimilation still shows great potentials in terms of identifiability of parameters. In our study, we think that most parameters converge in the 5 months assimilation period.

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 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. We expect that for example with more unknowns (i.e., 2D and 3D-applications) EnKF-based algorithms will perform better than PF, as PF will become extremely CPU-intensive and needs many more particles. For those high-dimensional applications EnKF is expected therefore to be more CPU-efficient than PF. Nevertheless, the small difference in performance between EnKF and PF based algorithms in this study indicates that PF is also an efficient data assimilation algorithm for problems of this size. ~~The results of this study are obtained for the point scale, a relatively small~~ It can be expected that larger ensemble ~~size of 100 (which is nevertheless larger than typically used for~~ numbers can improve the performance of EnKF and PF based algorithms. For MCMCPF, multiple MCMC resampling steps can also help improve performance. We expect that both data assimilation methods can relatively easily be used in combination with other land surface models~~) and for relatively short parameter estimation and verification periods. Given~~ and that the ~~CPU intensity~~relative performance of the ~~calculations a larger comparison was beyond the scope of this work.~~ data assimilation methods would also be similar for those models.

~~DeChant and Moradkhani (2012) used a range of performance measures, like Nash Sutcliffe efficiency (NSE), 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 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 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 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 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 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 data assimilation schemes if one adopts the update rule of the EnKF during particle resampling with MCMC [Vrugt et al., 2013].

In the verification period soil moisture of the top layer cannot be represented perfectly by the two LSM´s, in spite of parameter updating with state of the art data assimilation methods. Table 5 and table 9 illustrate that the RMSE values of the four joint state and parameter assimilation methods are similar for both models, which means that both models have larger errors for the top layer. There is a number of reasons for the larger soil moisture mismatches for the upper layer: (i) the memory effect from initial conditions, very well identified at the beginning of the verification period, is smaller for the upper soil layer, as this layer is more affected by precipitation events and evaporation; (ii) these soil moisture changes make that it is also more affected by model structural errors, for example concerning evaporation processes.

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~~overinterpret this result, as this is only a study for one site, the worse performance of VIC could be related to 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 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 TERENO-observation site Rollesbroich in the western part of Germany. The highly equipped site allows ~~to gain~~gaining more insight in the performance of data assimilation algorithms for joint state-parameter estimation for land surface models. Measured soil moisture ~~data~~contents 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 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 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:

$$\frac{\partial S}{\partial t} = P - E - Q \qquad \text{(A1)}$$

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 [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 Leaf Area Index (LAI) (Dickinson, 1984). Vegetation transpiration is estimated using Blondin (1991) and 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):

$$\frac{\partial \theta_1}{\partial t} z_1 = P - Q_d - Q_{1,2} - E_1 \qquad \text{(A2)}$$

$$\frac{\partial \theta_2}{\partial t} z_2 = Q_{1,2} - Q_{2,3} - E_2 \qquad \text{(A3)}$$

$$\frac{\partial \theta_3}{\partial t} z_3 = Q_{2,3} - E_3 - Q_b \qquad \text{(A4)}$$

where θ [L^3L^{-3}] is volumetric soil moisture content, z_i [L] is soil depth for layer i (i =1,2,3), $Q_{i,i+1}$ [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 zero in Eq. (A3 and A4) because there is no evaporation from layer 2 and 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):

$$Q_{i,i+1} = k_{s,i} \left(\frac{\theta_i - \theta_{r,i}}{\theta_i^{max} - \theta_{r,i}} \right)^{\beta_i} \qquad (i=1,2) \qquad \text{(A5)}$$

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

$$\theta_i^{max} = \phi_i \qquad (i=1,2) \qquad \text{(A6)}$$

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 [-]:

$$\beta_i = \frac{2}{B_p} + 3 \qquad \text{(A7)}$$

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

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

~~(A8)~~ $\begin{cases} P - (\theta_1^{max} - z_1\theta_1) - (\theta_2^{max} - z_2\theta_2) + (\theta_1^{max} + \theta_2^{max}) \left(1 - \frac{Iv+P}{I_m}\right)^{1+b}, & P+Iv \le I_m \\ P - (\theta_1^{max} - z_1\theta_1) - (\theta_2^{max} - z_2\theta_2), & P+Iv > I_m \end{cases}$ (A8)

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~~When P+Iv $>I_m$, the upper soil layers will be saturated and when P+Iv $\le I_m$, the upper soil layers are assumed unsaturated, and infiltration capacity Iv [L] is variable which is a function of the maximum filtration capacity I_m [L] [Zhao, 1992]:

$$IIv = I_m(1 - (1 - A)^{\frac{1}{b}}) \text{ with } I_m = (1 + b)(\theta_1^{max} + \theta_2^{max}) \quad \text{(A9)}$$

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

$$A = 1.0 - (1.0 - \frac{z_1\theta_1 + z_2\theta_2}{\theta_1^{max} + \theta_2^{max}})^{\frac{b}{1+b}} \quad \text{(A10)}$$

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

$$Q_b = \begin{cases} \frac{D_S D_m}{W_S \theta_3^{max}} \theta_3 z_3, & 0 \le \theta_3 z_3 \le W_S \theta_3^{max} \\ \frac{D_S D_m}{W_S \theta_3^{max}} \theta_3 z_3 + \left(D_m - \frac{D_S D_m}{W_S}\right)\left(\frac{\theta_3 z_3 - W_S \theta_3^{max}}{\theta_3^{max} - W_S \theta_3^{max}}\right)^2, & \theta_3 z_3 > W_S \theta_3^{max} \end{cases} \quad \text{(A11)}$$

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

Table 3 shows the soil layer definition where soil moisture is calculated in CLM. The hydraulic conductivity k_i [LT^{-1}], soil matric potential ψ_i [L] and soil thermal conductivity λ_i [$WL^{-1}K^{-1}$] 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:

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

where

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

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

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

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

where θ_i [L^3L^{-3}] is soil moisture content for layer i, $\theta_{sat,i}$ [L^3L^{-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 [$ML^{-2}T^{-1}$] 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):

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

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~~if ground evaporation is calculated, or the saturated specific humidity within the canopy ~~when~~if 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:

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

$$\left[k_i\left(\frac{\partial(\psi_i + z_i - C)}{\partial z}\right)\right] - E = \frac{\partial}{\partial z}\left[k_i\left(\frac{\partial(\psi_i - \psi_{E,i})}{\partial z}\right)\right] - E \quad \text{with} \quad C = \psi_{E,i} + z_i \qquad (B7)$$

$$\psi_{E,i} = \psi_{sat,i}\left(\frac{\theta_E(z_i)}{\theta_{sat,i}}\right)^{-B_i} \quad \text{with} \quad \theta_E(z_i) = \theta_{sat,i}\left(\frac{\psi_{sat,i} + z_\nabla - z_i}{\psi_{sat,i}}\right)^{\frac{1}{B_i}} \qquad (B8)$$

where ~~z_∇~~ψ_E [L] is ~~groundwater~~the equilibrium soil matric potential, z_∇ is water table depth~~,~~ [L] 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. Details about modifications in Eq. (B7) can be found in the CLM-manual (Oleson et al., 2013). General Richards equation used a "θ" -based solution which cannot account for the variation of ψ below water table because "θ" is constant (at saturated value) while ψ varies temporally and spatially, which leads to the failure to maintain the hydrostatic equilibrium soil moisture distribution. However, the modified Richards equation in which a constant hydraulic potential C is explicitly subtracted at each time step can fix this deficiency. Details about the implementation of the modified method are given in (Zeng and Decker, 2009).

In CLM, water table depth z_∇ is calculated according to Niu ~~(Niu~~ et al~~.,~~. (2007). An unconfined aquifer is assumed to ~~lie~~be 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:

$$q_{recharge} = -k_{wt}\frac{(-\psi_{wt})}{(z_\nabla - z_{wt})} \qquad (\text{~~B8~~}B9)$$

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 $q_{drainage}$ [$ML^{-2}T^{-1}$] is calculated by a simple TOPMODEL-based (SIMTOP) scheme (Niu et al., 2005)

$$q_{drainage} = 10\sin(\varepsilon)\exp(-2.5z_\nabla) \qquad (\text{~~B9~~}B10)$$

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

$$\Delta z_{\nabla} = \frac{\Delta W}{S_y} \quad \text{with } \Delta W = (q_{recharge} - q_{drainage})\Delta t \qquad (\text{~~B10~~}B11)$$

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~~The first author of this paper was funded 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.

Hodgkinson R. A., T. J. Pepper, and D. W. Wilson (2004), Evaluation of tipping bucket rain gauge performance and data quality, Environment Agency, Rio House, Waterside Drive, Aztec West, Almondsbury, Bristol, BS32 4UD, ISBN:1844323242.

Hübner, C., R. Cardell-Oliver, R. Becker, K. Spohrer, K. Jotter, and T. Wagenknecht (2009), Wireless soil moisture sensor networks for environmental monitoring and vineyard irrigation: Helsinki University of Technology, 1, 408-415.

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, and H. Vereecken (2012), Identification of time-variant river bed properties with the ensemble Kalman filter, Water Resources Research, 48, W10534, doi:10.1029/2011WR011743.

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.

Lee, J. H. (2014), Spatial-Scale prediction of the SVAT soil hydraulic variables characterizing stratified soils on the Tibetan Plateau from an EnKF analysis of SAR soil moisture, Vadose Zone Journal, 13(11), doi: 10.2136/vzj2014.06.0060.

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., M. Camporese, and M. Putti (2012), Ensemble Kalman filter versus particle filter for a physically-based coupled surface-subsurface model, Advances in Water Resources, 47, 1-13, doi:10.1016/j.advwatres.2012.06.009.

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.W., H. R., H.. Bogena, J. A., M.. Huisman, G., P.. Martinez, Y. A., . Pachepsky, and V. H. Vereecken (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.

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.

Scheerlinck, K., V. R. N. Pauwels, H. Vernieuwe, and B. De Baets (2009), Calibration of a water and energy balance model: Recursive parameter estimation versus particle swarm optimization, Water Resources Research, 45, W10422, doi:10.1029/2009WR008051.

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 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.

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 exchange fluxes: The role of spatial patterns of riverbed hydraulic conductivities, Journal of Hydrology, doi:10.1016/j.jhydrol.2015.08.019.

Thiemann, M., M. Trosset, H. Gupta, and S. Sorooshian (2001), Bayesian recursive parameter estimation for hydrologic models, Water Resources Research, 37(10), 2521–2535.

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 value of soil moisture measurements in vadose zone hydrology: A review, Water Resources Research, 44, 21, doi:10.1029/2008wr006829.

Vereecken, H., et al. (2013), Modeling Soil Processes: Review, Key Challenges, and New Perspectives, Vadose Zone Journal, 15, 5, doi: 10.2136/vzj2015.09.0131.

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.

Vrugt, J. A., ~~C~~and C. J. F. Ter Braak (2011), DREAM(D): An adaptive Markov Chain Monte Carlo simulation algorithm to solve discrete, noncontinuous, and combinatorial posterior parameter estimation problems, Hydrology Earth System Science, 15(12), 3701–3713.

Vrugt, J. A., C. J. F. Ter. 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), 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.

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.

Yan, H., C. M. DeChant, and H. Moradkhani (2015), Improving Soil Moisture Profile Prediction With the Particle Filter-Markov Chain Monte Carlo Method, Ieee Transactions on Geoscience and Remote Sensing, 53(11), 6134-6147, doi:10.1109/tgrs.2015.2432067.

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

Zeng, X., and M. Decker (2009), Improving the numerical solution of soil moisture-based Richards equation for land models with a deep or shallow water table, Journal of Hydrometeorology, 10, 308-319, doi:10.1175/2008JHM1011.1.

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

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

Table 3 Soil layer definition where soil moisture is calculated in CLM. Layer node depth (z), thickness(Δz), and depth at layer interface(zh) for 10 soil layers. Unit is meter.

Layer i	z	Δz	zh
1 (top)	0.0071	0.0175	0.0175
2	0.0279	0.0276	0.0451
3	0.0623	0.0455	0.0906
4	0.1189	0.0750	0.1655
5	0.2122	0.1236	0.2891
6	0.3661	0.2038	0.4929
7	0.6198	0.3360	0.8289
8	1.0380	0.5539	1.3828
9	1.7276	0.9133	2.2961
10	2.8646	1.5058	3.8019

Table 4 NSE and RMSE values for soil moisture content characterization at 5cm, 20cm and 50cm for different scenarios in the assimilation period with the VIC-3L model.

Criteria	**Soil depth**	**MCMC**	**PF**	**AUG**	**DUAL**	**noParamUpdate**	**openloop**
NSE (-)	5cm	0.82	0.73	0.80	0.82	0.89	0.33
	20cm	0.80	0.84	0.92	0.91	0.86	-1.16
	50cm	0.27	-11.77	0.69	0.58	0.91	-26.65
RMSE (m^3/m^3)	5cm	0.019	0.023	0.020	0.019	0.015	0.036
	20cm	0.011	0.010	0.007	0.007	0.009	0.037
	50cm	0.021	0.088	0.014	0.016	0.008	0.129

Table 5 NSE and RMSE values for soil moisture content characterization at 5cm, 20cm and 50cm in the verification period with the VIC-3L model.

Criteria	**Soil depth**	**MCMC**	**PF**	**AUG**	**DUAL**	**noParamUpdate**	**openloop**
NSE (-)	5cm	0.39	0.39	0.39	0.39	0.35	0.36
	20cm	0.38	0.47	0.40	0.39	-1.75	-1.87
	50cm	-10.33	-8.41	-10.54	-11.33	-26.83	-32.96
RMSE (m^3/m^3)	5cm	0.052	0.052	0.052	0.052	0.054	0.053
	20cm	0.026	0.024	0.026	0.026	0.055	0.056
	50cm	0.076	0.069	0.077	0.079	0.119	0.132

Table 6 RMSE values for soil moisture content characterization at 5cm 20cm and 50cm in the assimilation and verification periods for EnKF with the VIC-3L model. AUG_10%, AUG_20%, DUAL_10% and DUAL_20% represent forcing errors of 10%, 20%, 10% and 20% respectively.

Period	**Soil depth**	**AUG_10%**	**AUG_20%**	**DUAL_10%**	**DUAL_20%**
Assimilation	5cm	0.020	0.019	0.019	0.019
	20cm	0.007	0.007	0.007	0.007
	50cm	0.014	0.014	0.016	0.014
Verification	5cm	0.052	0.052	0.052	0.052
	20cm	0.026	0.025	0.026	0.025
	50cm	0.077	0.077	0.079	0.079

Table 7 RMSE values for soil moisture content characterization at 5cm, 20cm and 50cm in the assimilation and verification periods for 3 scenarios of RRPF with the VIC-3L model. PF_0.01 represents the scenario in which s=0.01, PF_0.1 represents s=0.1 and PF_0.5 represents s=0.5.

Period	**Soil depth**	**PF_0.01**	**PF_0.1**	**PF_0.5**
Assimilation	5cm	0.025	0.023	0.015
	20cm	0.012	0.010	0.007
	50cm	0.113	0.088	0.037
Verification	5cm	0.053	0.052	0.056
	20cm	0.025	0.024	0.020
	50cm	0.119	0.069	0.071

Table 8 NSE and RMSE values for soil moisture content characterization at 5cm, 20cm and 50cm in the assimilation period with the CLM model.

Criteria	**Soil depth**	**MCMC**	**PF**	**AUG**	**DUAL**	**noParamUpdate**	**openloop**
NSE (-)	5cm	0.63	0.63	0.82	0.85	0.72	-0.31
	20cm	0.73	0.23	0.94	0.95	0.98	-0.57
	50cm	0.50	-0.26	0.85	0.86	0.47	-6.90
RMSE (m^3/m^3)	5cm	0.027	0.027	0.019	0.017	0.024	0.051
	20cm	0.013	0.022	0.006	0.006	0.004	0.031
	50cm	0.017	0.028	0.009	0.009	0.018	0.069

Table 9 NSE and RMSE values for soil moisture content characterization at 5cm, 20cm and 50cm in the verification period with the CLM model.

Criteria	**Soil depth**	**MCMC**	**PF**	**AUG**	**DUAL**	**noParamUpdate**	**openloop**
NSE (-)	5cm	0.26	0.23	0.32	0.33	-0.19	-0.14
	20cm	0.39	0.21	0.44	0.46	0.24	-0.11
	50cm	0.35	-0.23	0.51	0.42	-3.87	-4.58
RMSE (m^3/m^3)	5cm	0.057	0.058	0.055	0.054	0.072	0.071
	20cm	0.026	0.029	0.025	0.024	0.031	0.035
	50cm	0.018	0.025	0.016	0.017	0.050	0.053

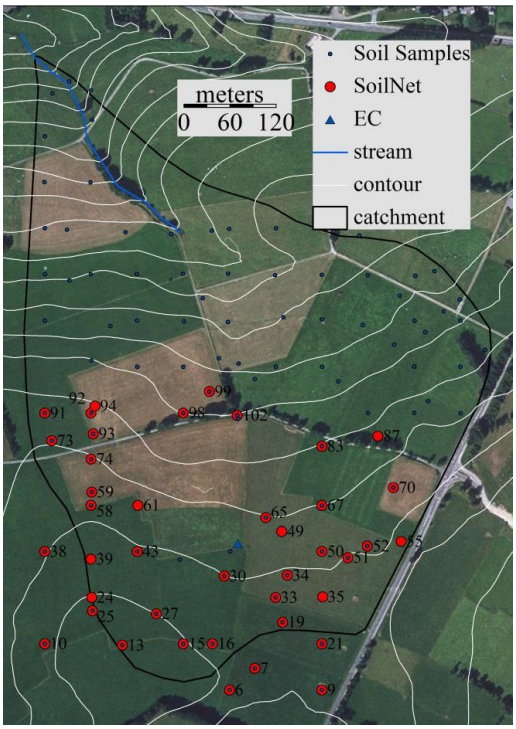
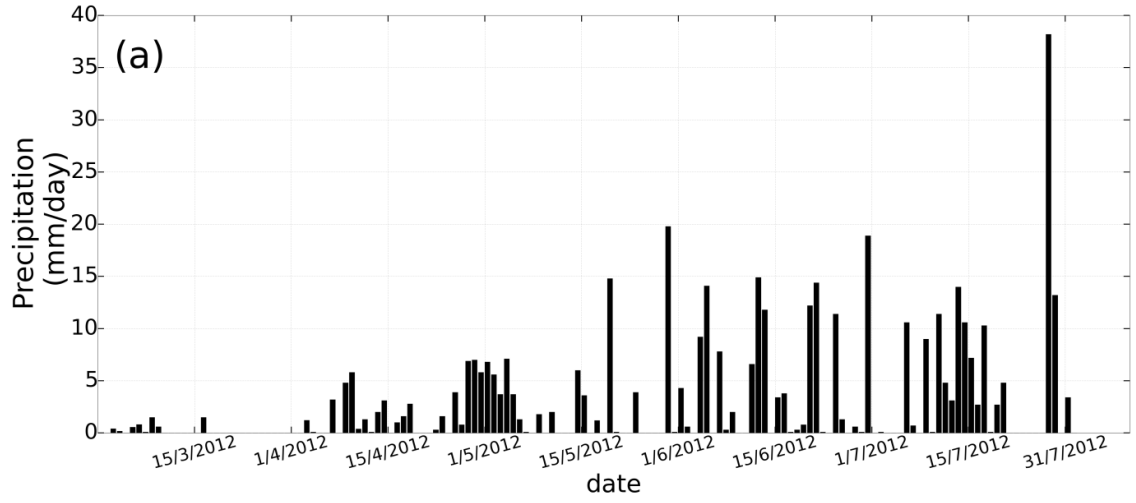
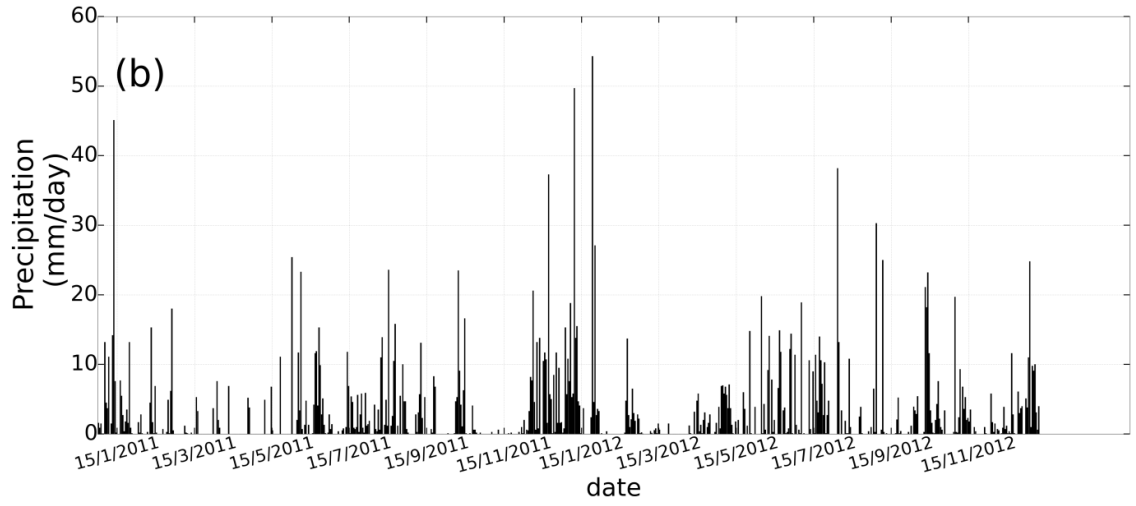
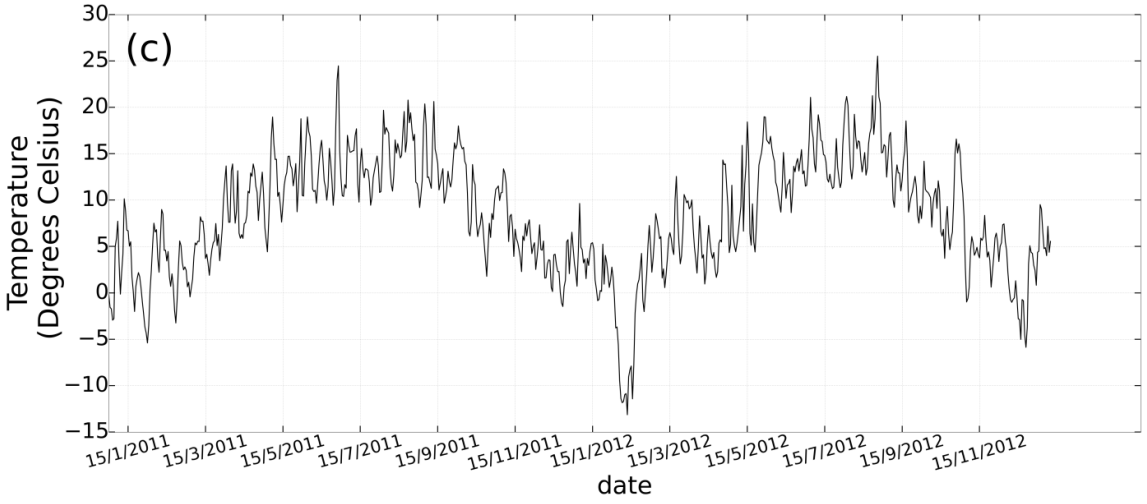


Figure 1 Overview of measurement devices in the Rollesbroich catchment. The blue dots are soil sample locations, red dots are soil network locations (soil moisture content and soil temperature are measured here), and the blue triangular indicates the eddy covariance tower. The Figure is taken from *Qu et al*. (2014).


(a)
Precipitation (mm/day)
0
5
10
15
20
25
30
35
40
15/3/2012
1/4/2012
15/4/2012
1/5/2012
15/5/2012
1/6/2012
15/6/2012
1/7/2012
15/7/2012
31/7/2012
date


(b)
Precipitation (mm/day)
0
10
20
30
40
50
60
15/1/2011
15/3/2011
15/5/2011
15/7/2011
15/9/2011
15/11/2011
15/1/2012
15/3/2012
15/5/2012
15/7/2012
15/9/2012
15/11/2012
date


(c)
Temperature (Degrees Celsius)
−15
−10
−5
0
5
10
15
20
25
30
15/1/2011
15/3/2011
15/5/2011
15/7/2011
15/9/2011
15/11/2011
15/1/2012
15/3/2012
15/5/2012
15/7/2012
15/9/2012
15/11/2012
date

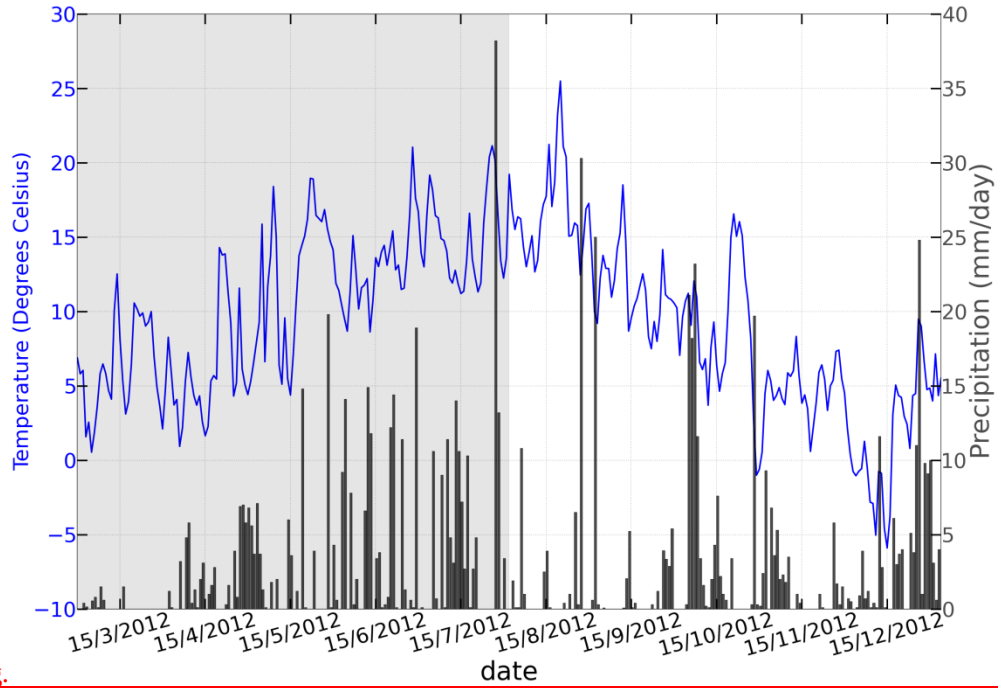
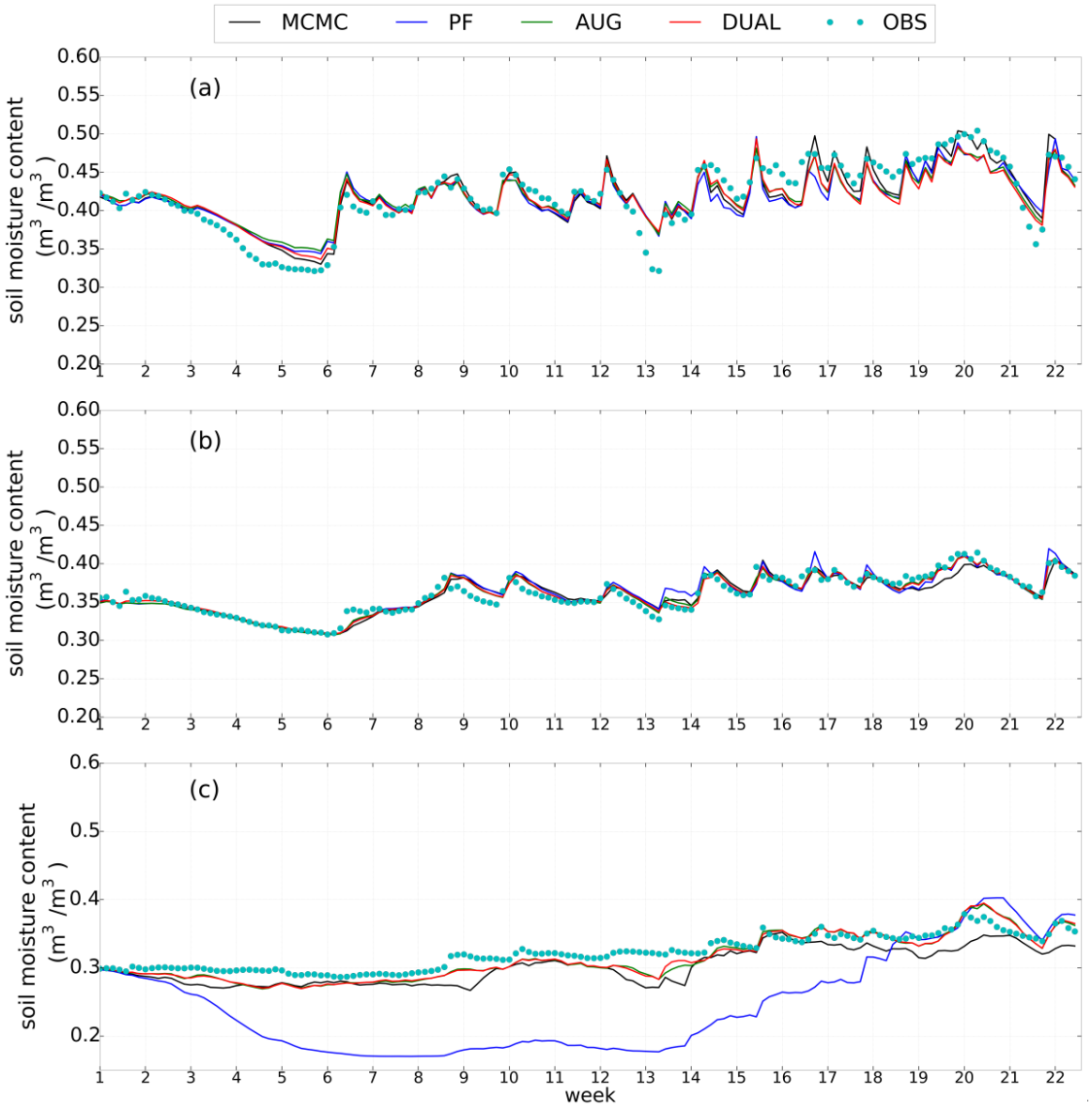


Fig.

~~Fig~~ure 2 ~~(a)~~daily mean air temperature (blue curve) and daily precipitation (black bars) from March 1, 2012 to ~~July~~December 31, 2012 ~~(~~, gray background area indicates parameter estimation period~~), week 1 is 01-03-2012 and week 22 coincides with 26-07-2012, (b) daily precipitation for the years 2011~~ (from March 1, 2012 to July 31, 2012) and white background area verification period (from August 1, 2012~~, and (c) daily mean air temperature for the years 2011 and~~ to December 31, 2012~~,~~), all measured at the Rollesbroich site.


MCMC
PF
AUG
DUAL
OBS
(a)
(b)
(c)
soil moisture content
(m3 /m3)
week

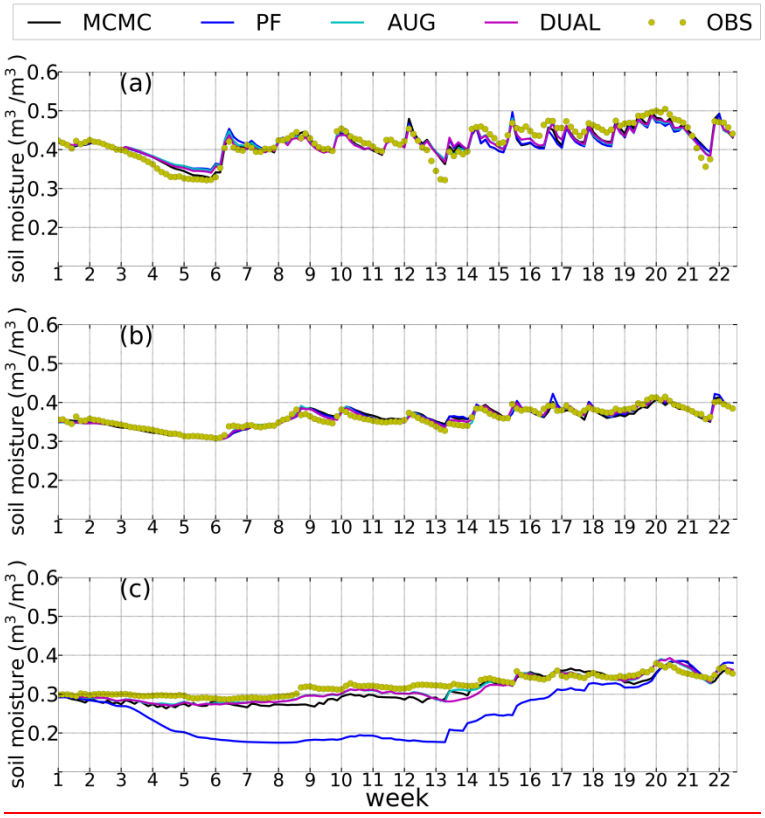
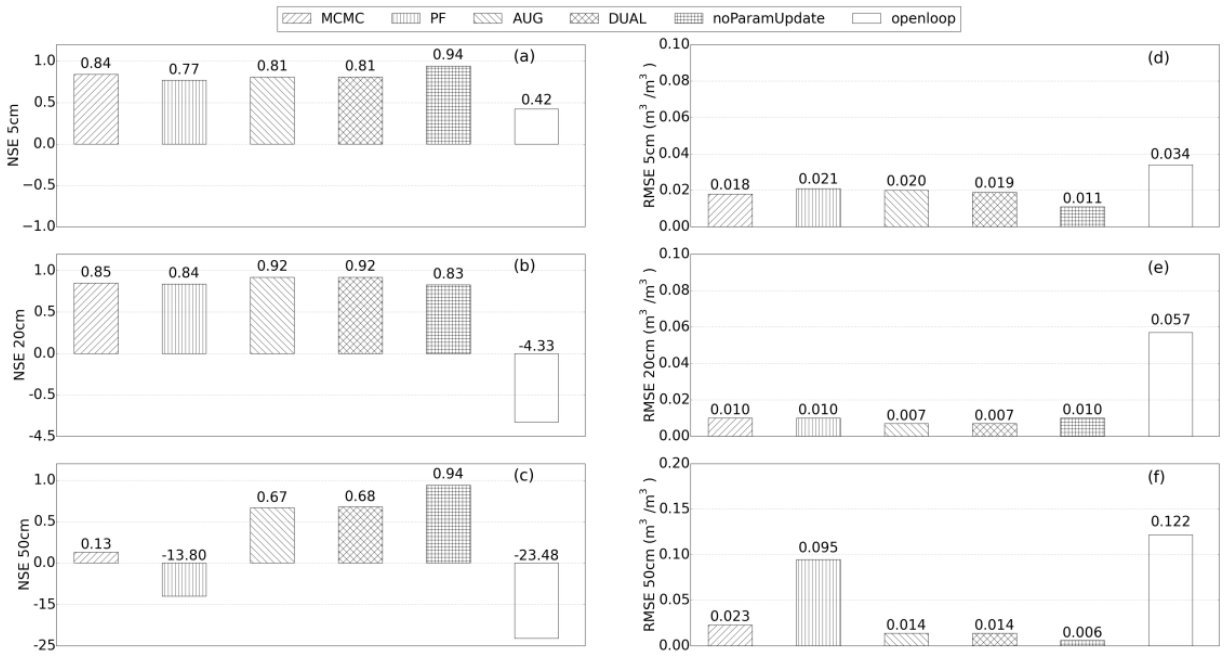
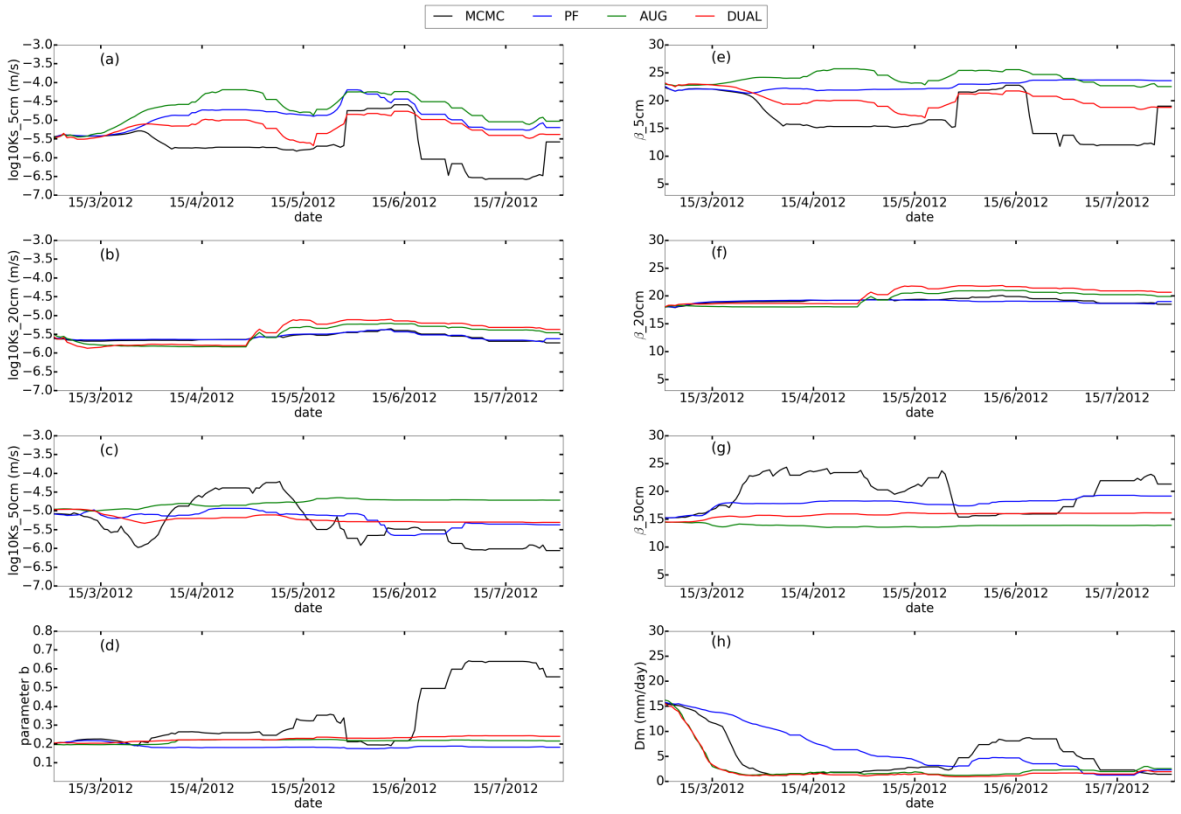


Figure 3 Time series of soil moisture content for different assimilation scenarios during the assimilation period (March-July 2012) for the VIC-3L model: (a) 5 cm depth, (b) 20 cm depth and (c) 50 cm depth. Week 1 starts at 01-03-2012 and week 22 at 26-07-2012.



~~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.~~



F~~igure 5~~

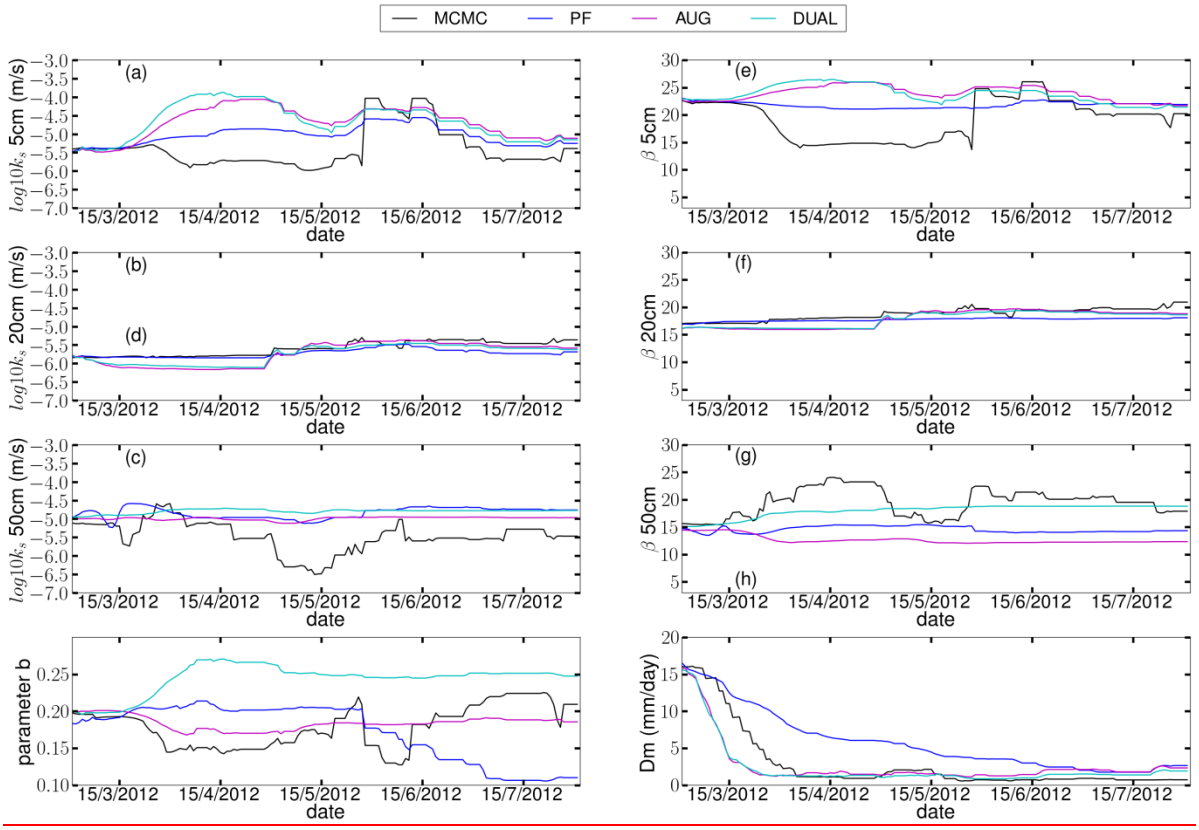
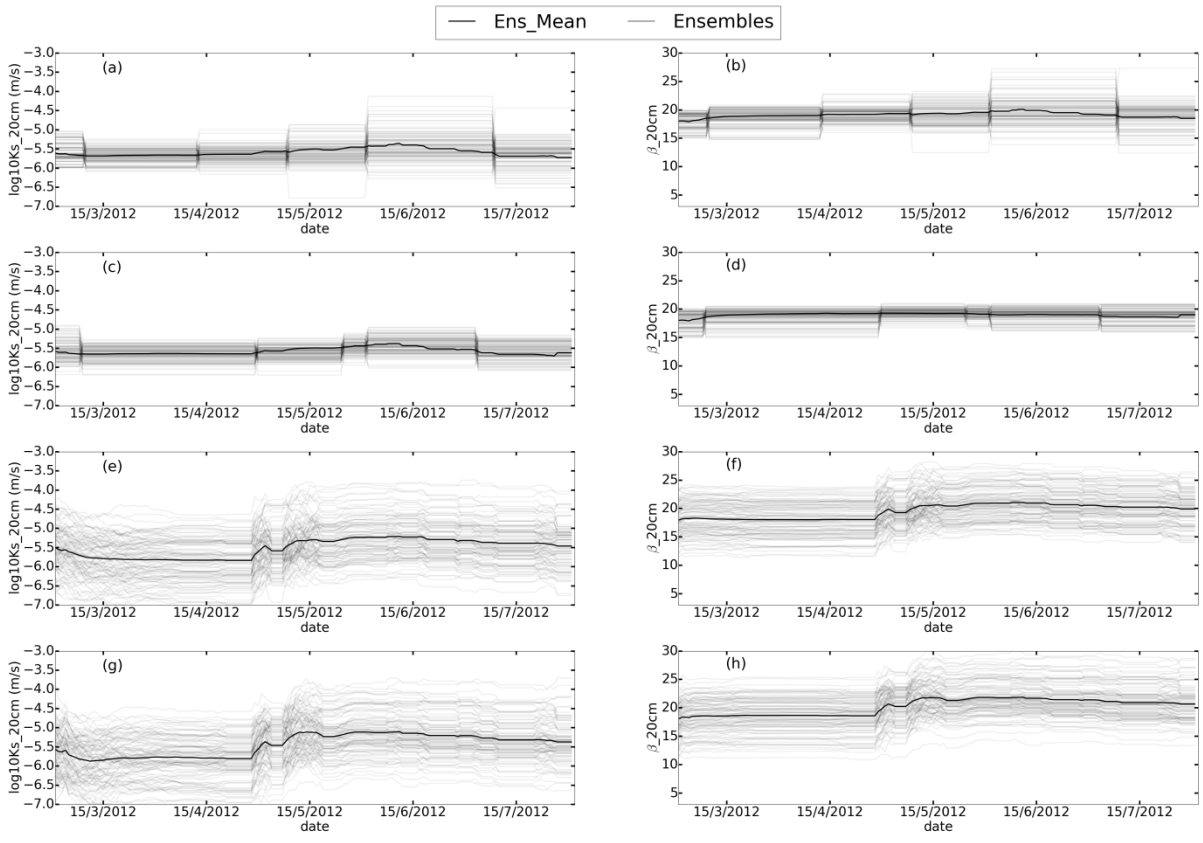


Figure 4 Temporal evolution of parameter values in the parameter estimation period (March 2012-July 2012), for the four data assimilation scenarios and the model VIC-3L. (a) Saturated hydraulic conductivity $\log_{10}k_s$ (m/s) at 5cm depth, (b) 20cm depth and (c) 50cm depth, (d) model parameter b, (e) model parameter β at 5cm depth, (f) 20cm depth and (g) 50cm depth, and (h) maximum velocity of baseflow D_m (mm/day).



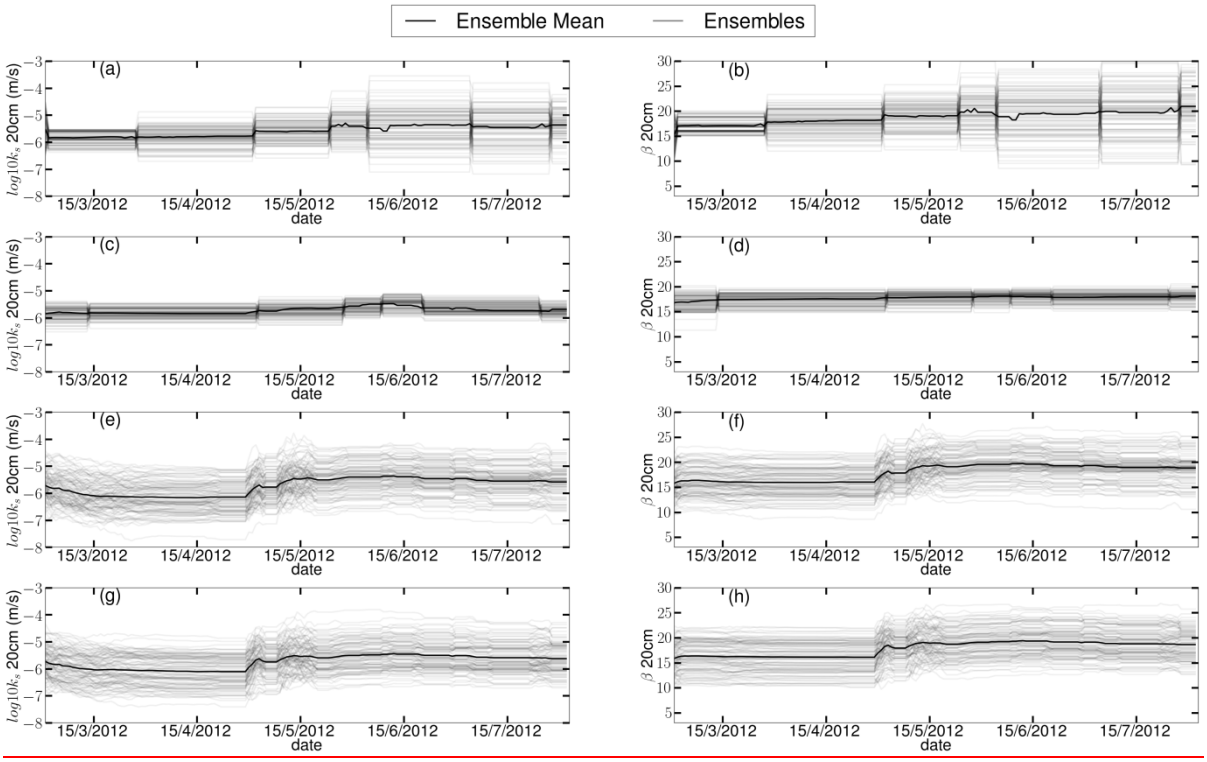
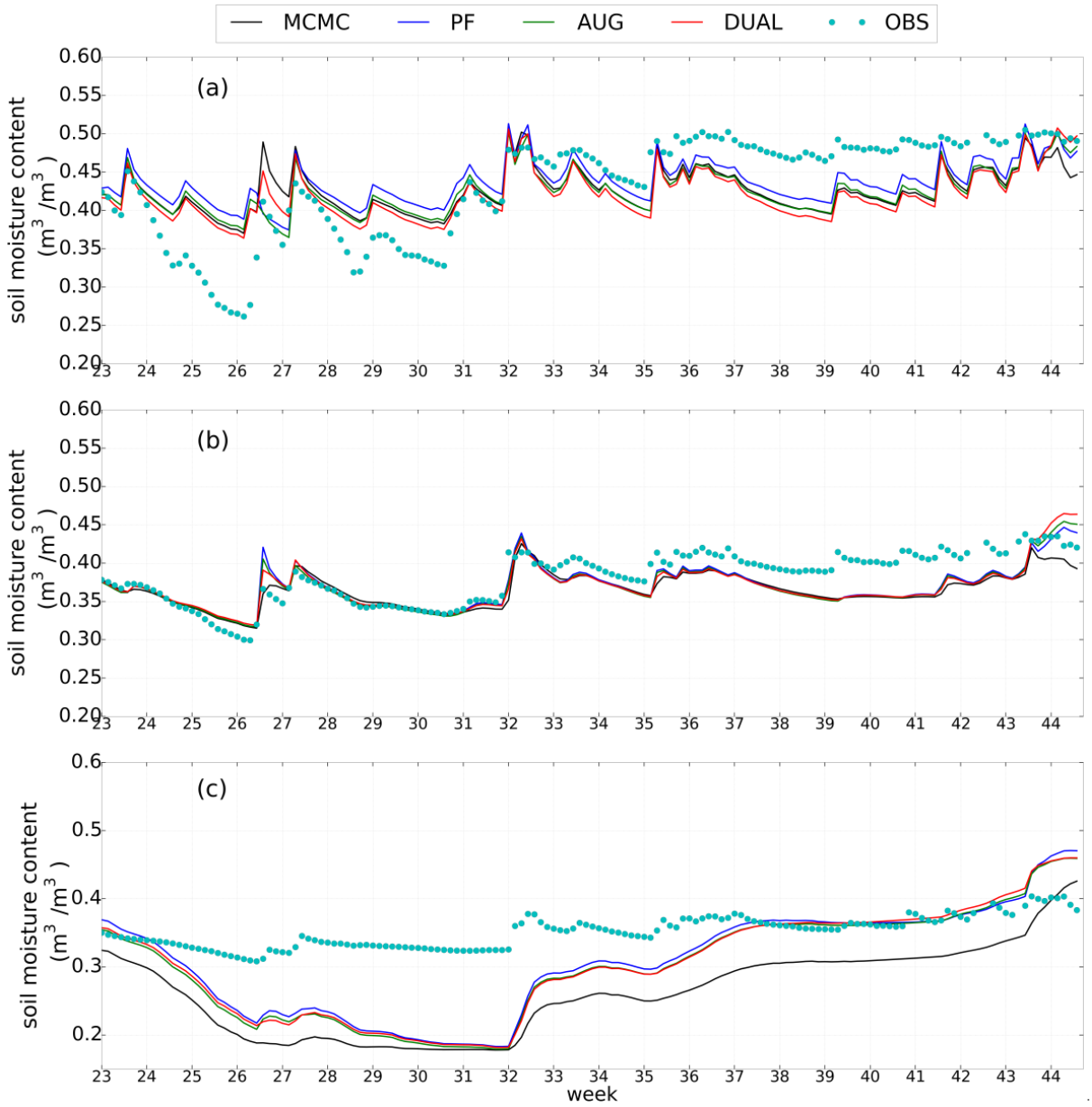


Figure ~~6~~5 Temporal evolution of parameter values for 100 ensemble members in the parameter estimation period, for the four data assimilation scenarios and the model VIC-3L and the second model layer. Saturated hydraulic conductivity $\log_{10}k_s$ (m/s) at 20cm depth is displayed for the four methods: (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.


MCMC
PF
AUG
DUAL
OBS
(a)
(b)
(c)
soil moisture content
(m3 /m3)
week

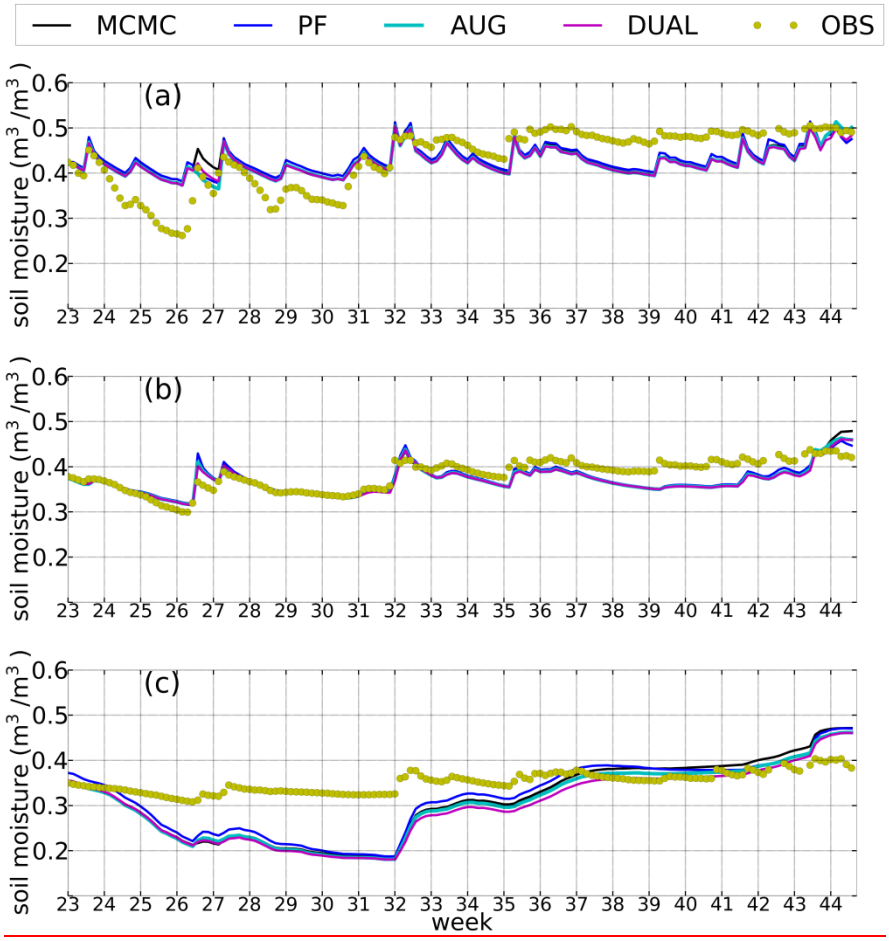


Figure ~~7~~6 Time series of soil moisture content for different assimilation scenarios during the verification period and for the VIC-3L model: (a) 5 cm depth, (b) 20 cm depth and (c) 50 cm depth. Week 23 starts with 01-08-2012 and week 44 with 26-12-2012.

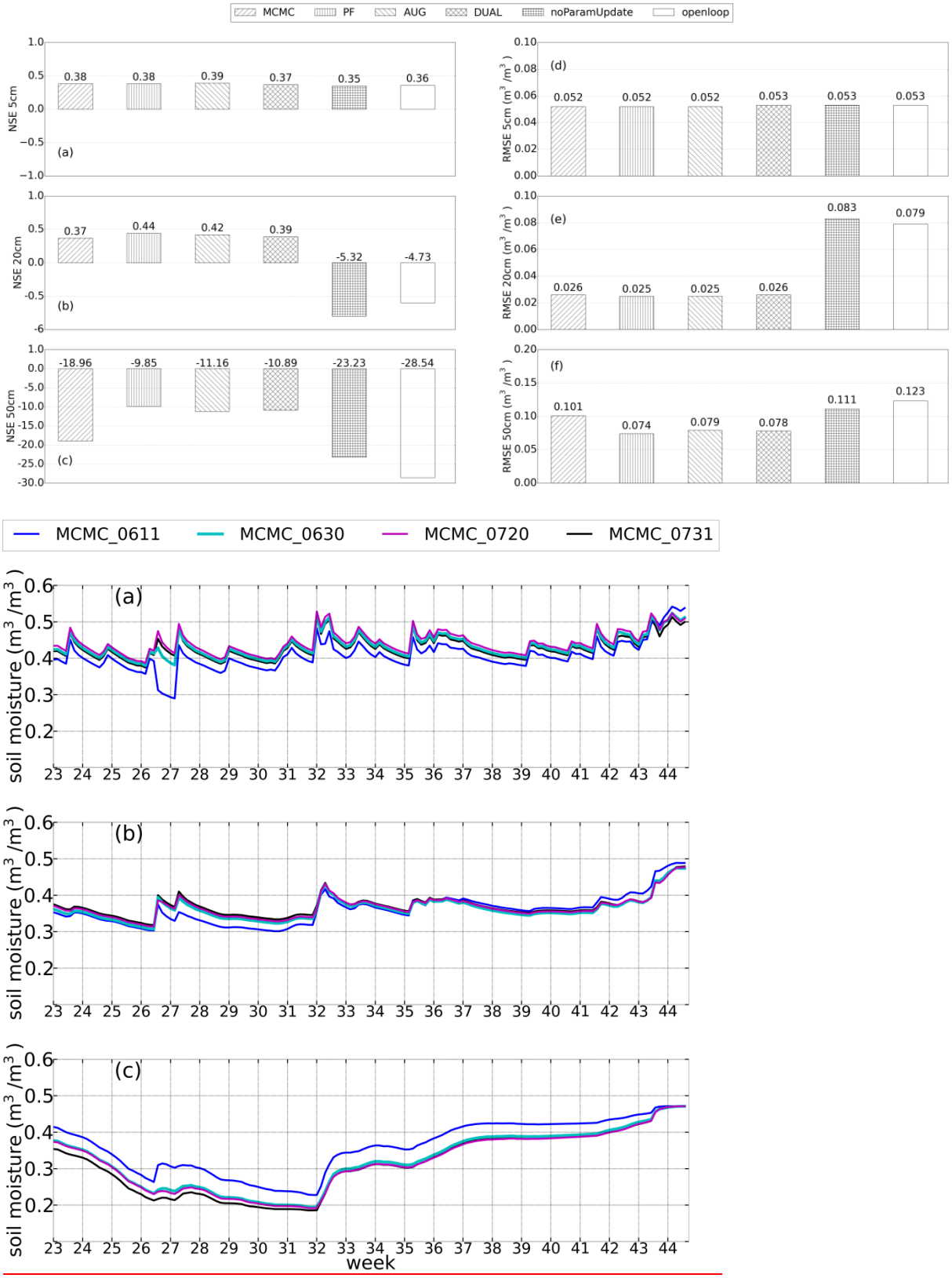


Figure ~~8 NSE and RMSE values for~~7 Time series of soil moisture content ~~characterization in the~~ for the four MCMCPF assimilation scenarios with different ending date (0611, 0630, 0720 and 0731) of assimilation period during the verification period ~~with~~and for the VIC-3L model: (a) ~~NSE values for soil moisture content at 5cm~~5 cm depth, (b) ~~NSE values at 20cm,~~20 cm depth and (c) 50 cm depth. Week 23 starts with 01-08-2012 and week 44 with 26-12-2012.

NSE values

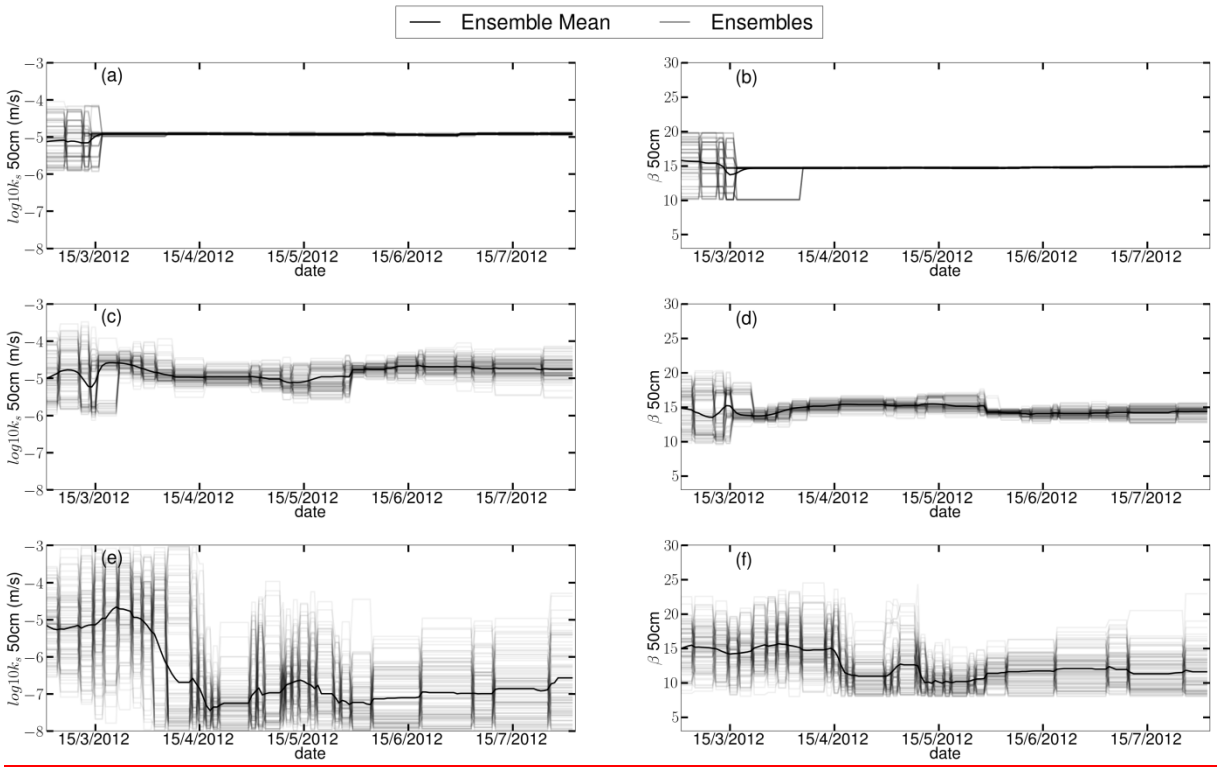


Figure 8 Temporal evolution of parameter values for 100 ensemble members in the parameter estimation period, for RRPF and the model VIC-3L and the third model layer. Saturated hydraulic conductivity $\log_{10}k_s$ (m/s) at 50cm, (d) RMSE values at 5cm, depth is displayed for the three methods: (a) PF_0.01, (c) PF_0.1, and (e) RMSE values at 20cmPF_0.5. Model parameter β at 50cm depth for the three methods: (b) PF_0.01, (d) PF_0.1, and (f) RMSE values at 50cmPF_0.5.

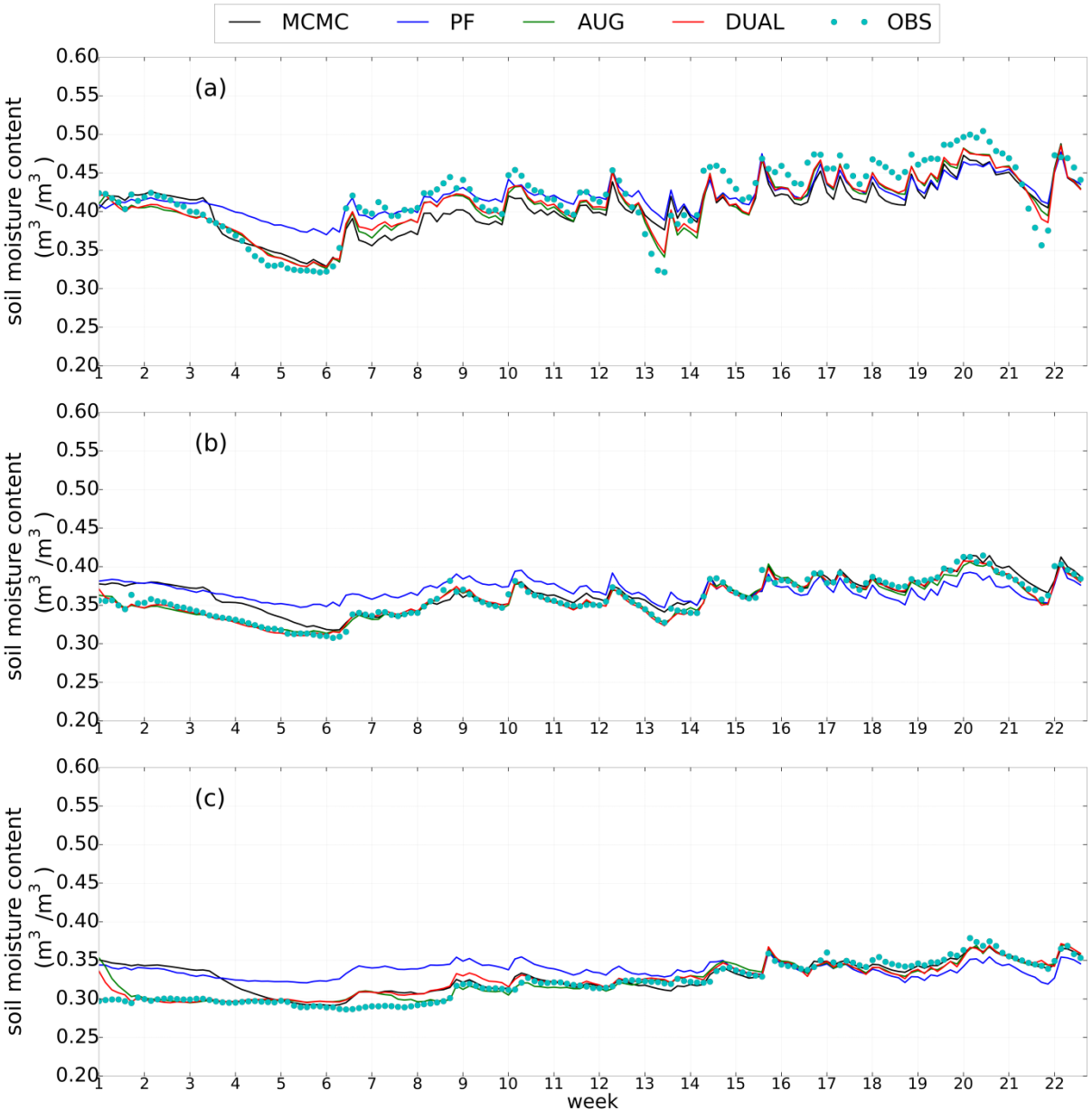

MCMC
PF
AUG
DUAL
OBS
(a)
(b)
(c)
soil moisture content
(m3 /m3)
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
week

Figure 9

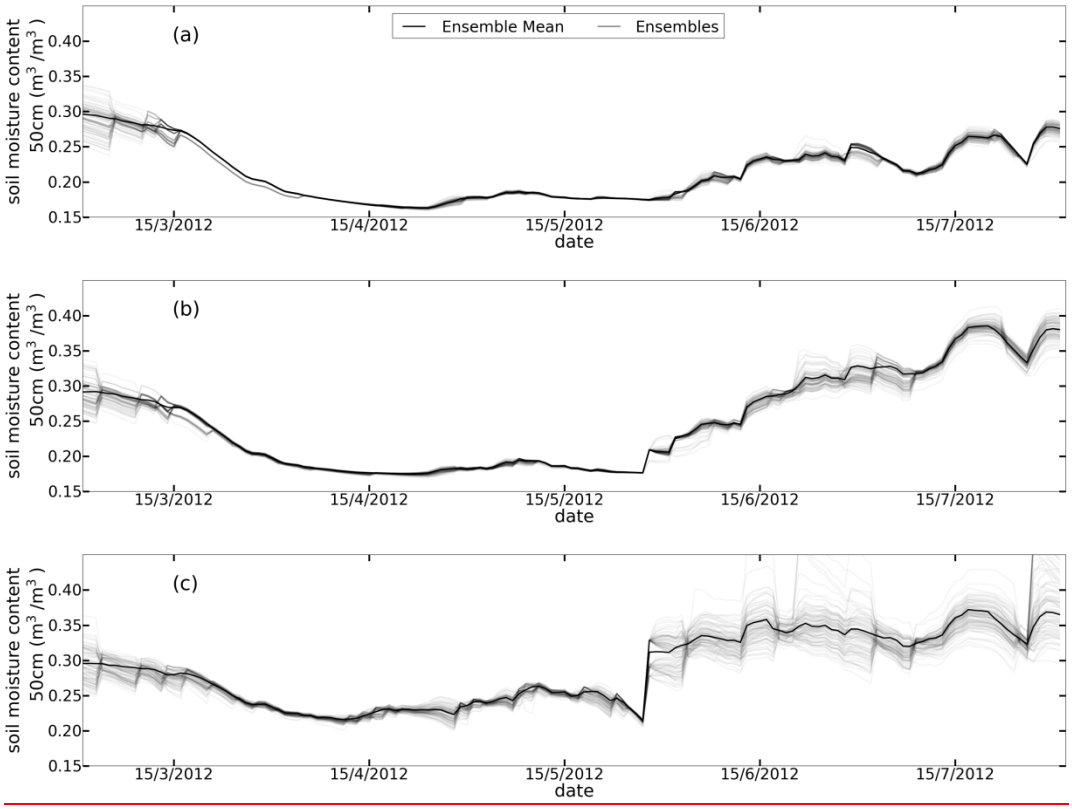


Figure 9 Temporal evolution of soil moisture content for 100 ensemble members during the assimilation period, for RRPF and the model VIC-3L and the third model layer: (a) PF_0.01, (b) PF_0.1 and (c) PF_0.5.

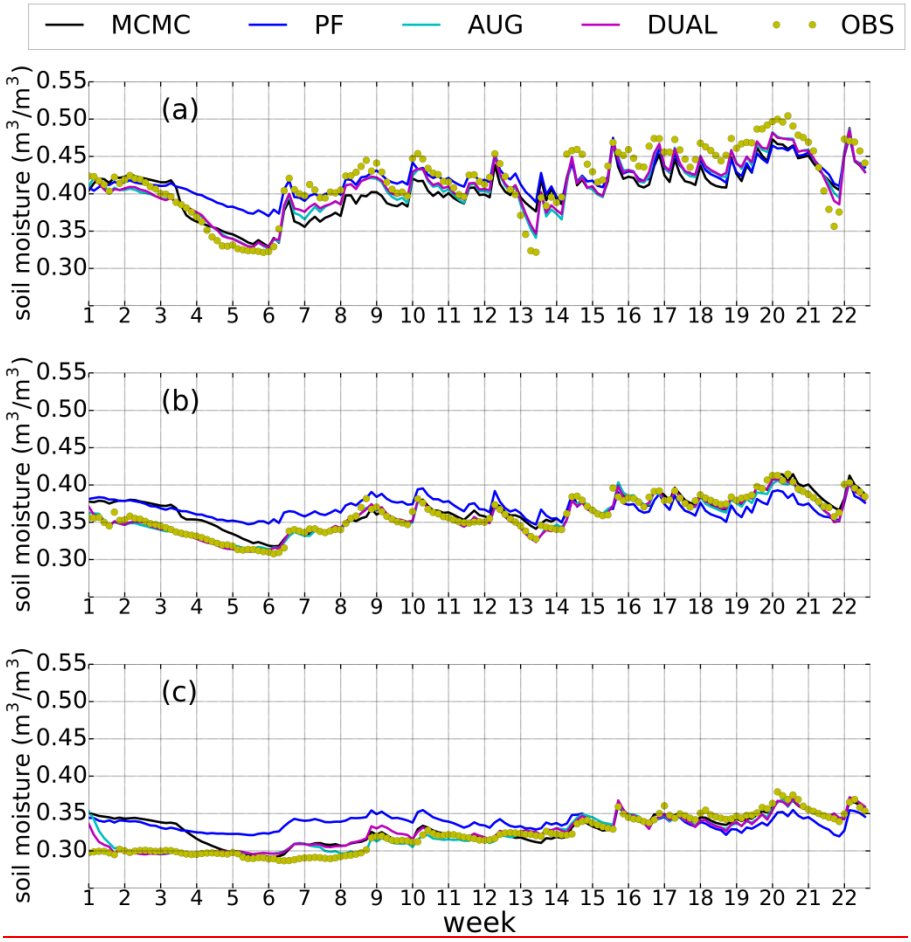
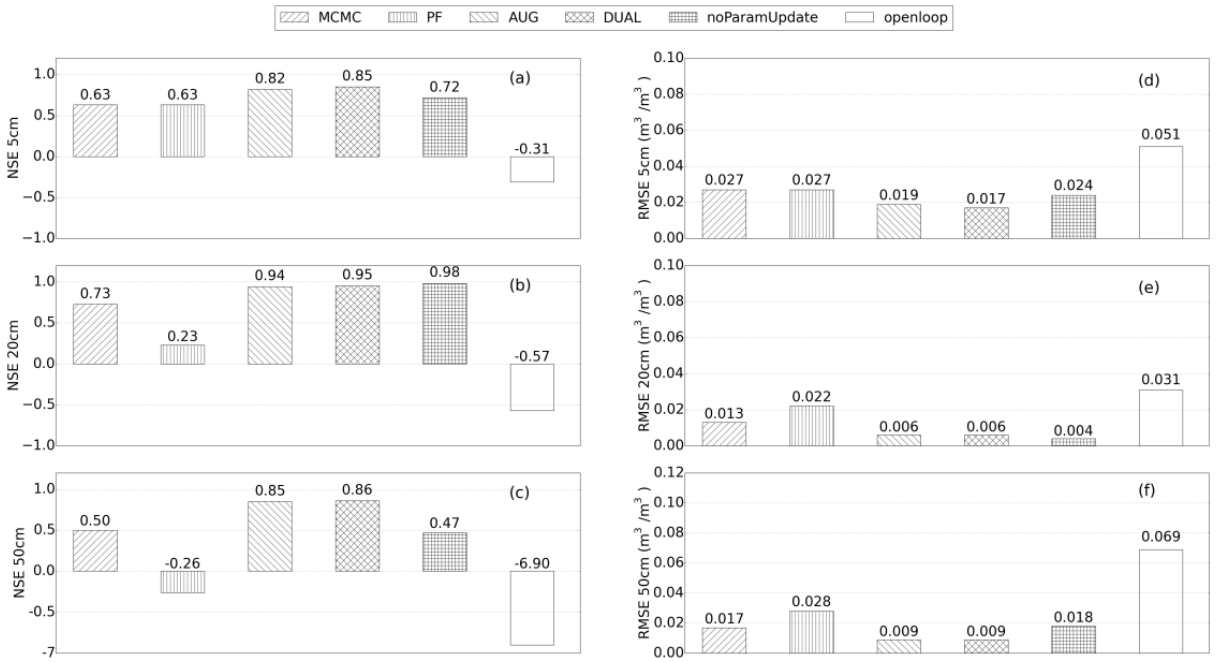


Figure 10 Time series of soil moisture content for different data assimilation scenarios during the assimilation period and for the CLM model: (a) 5 cm depth, (b) 20 cm depth and (c) 50 cm depth. Week 1 starts with 01-03-2012 and week 22 with 26-07-2012.



~~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.~~

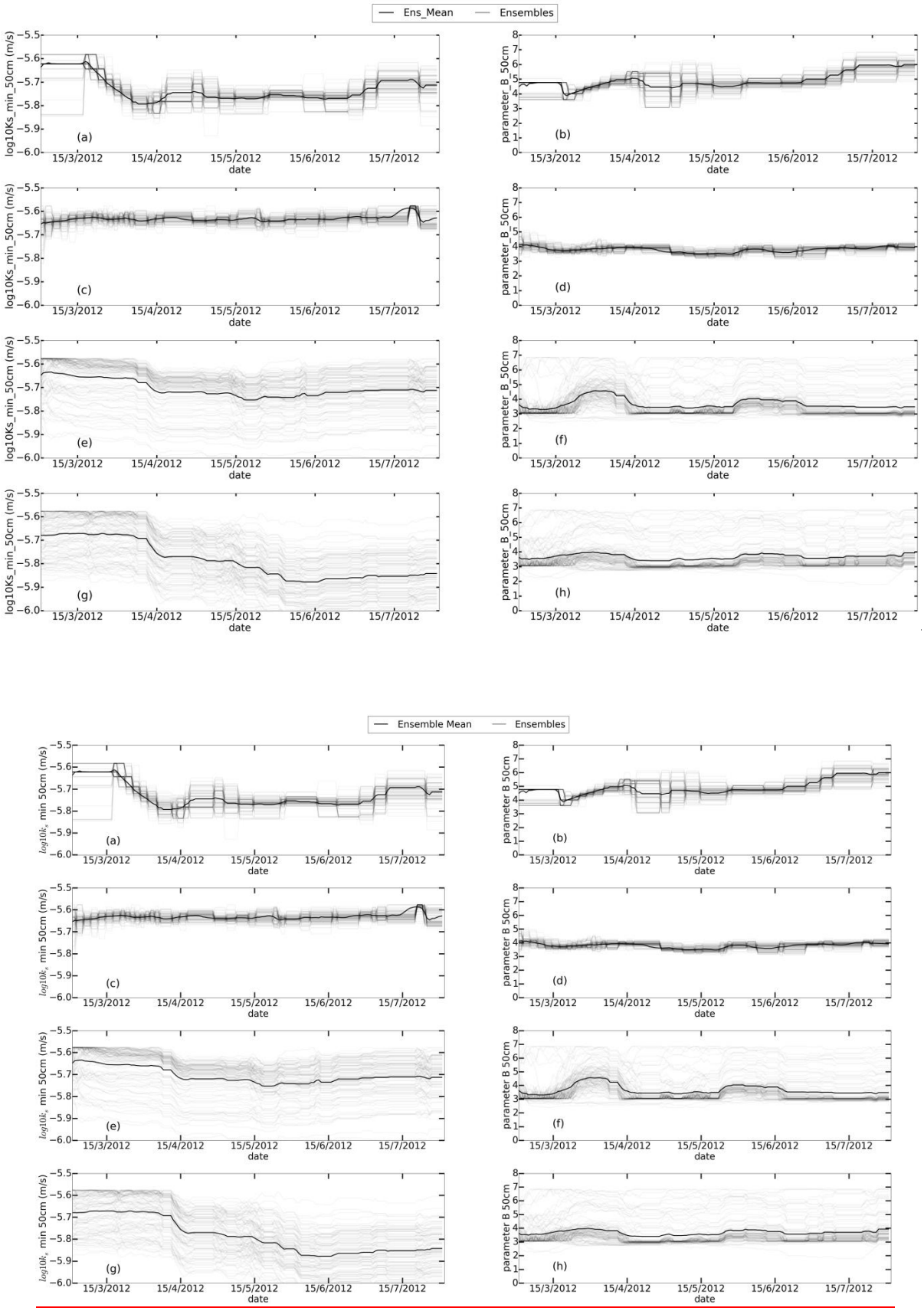
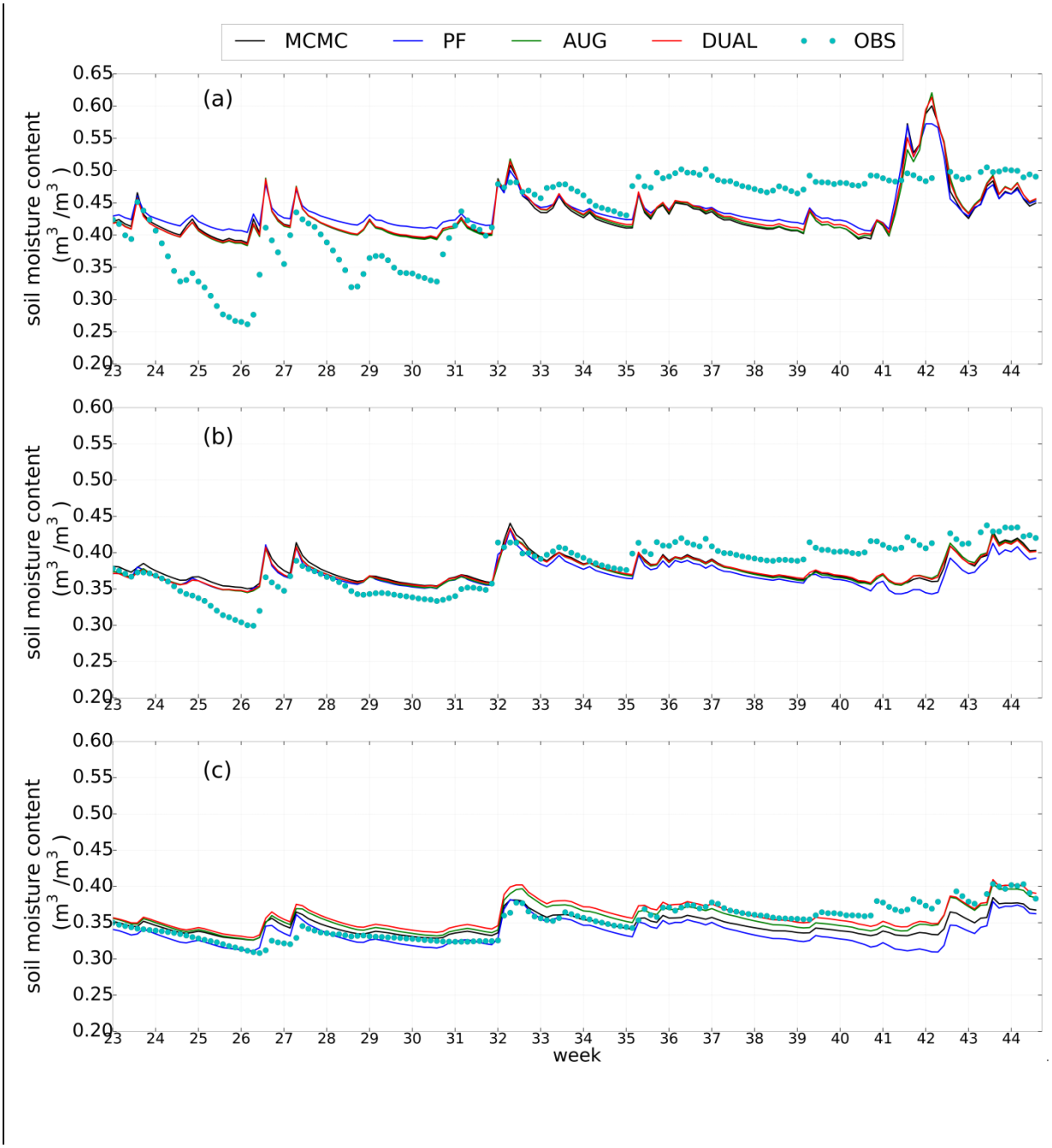


Figure 11 Temporal evolution of parameter values in the assimilation and parameter estimation period, for the four data assimilation scenarios and the CLM-model. Saturated hydraulic conductivity $\log_{10}K_s$ (m/s) at 50cm depth is displayed for the four methods: (a) MCMC, (c) PF, (e) AUG, and (g) DUAL. 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.



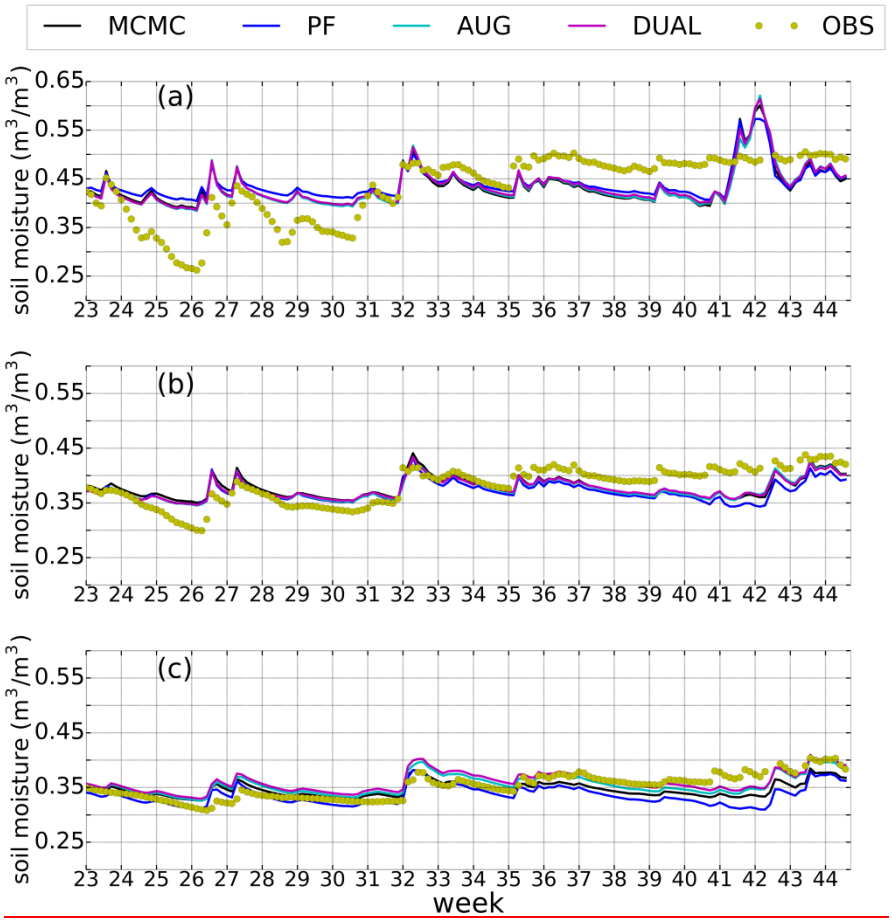
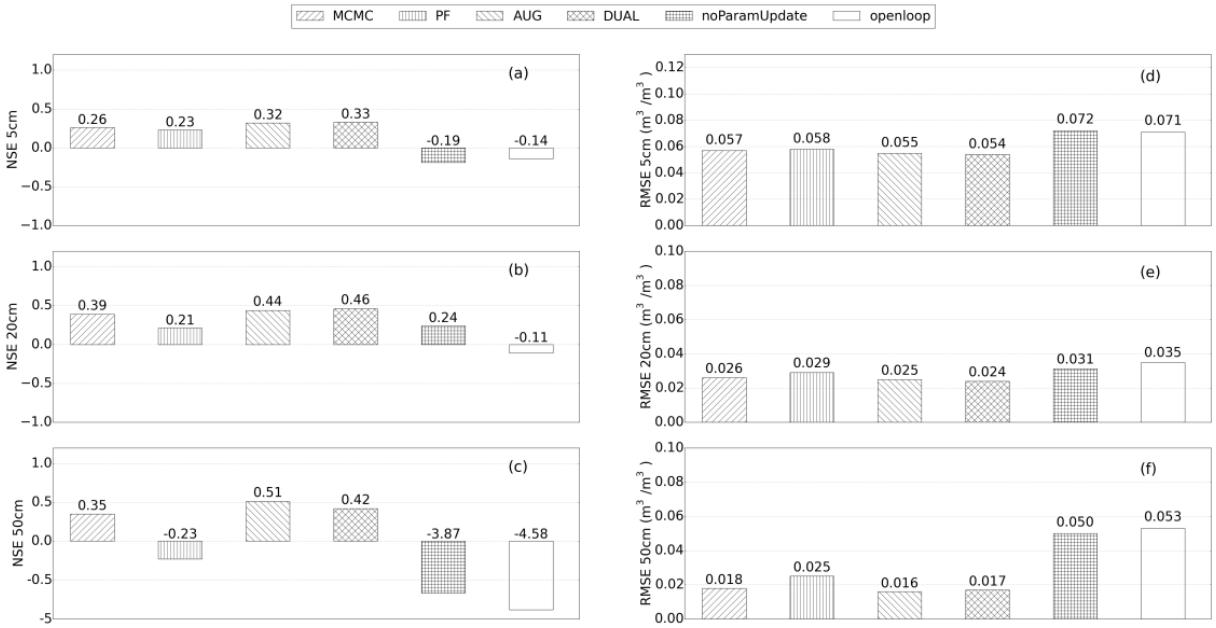


Figure 12 Time series of soil moisture content for ~~different assimilation scenarios with~~ CLM during the verification period: (a) 5 cm depth, (b) 20 cm depth and (c) 50 cm depth. Week 23 starts with 01-08-2012 and week 44 with 26-12-2012.



~~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.~~

Week 23 starts with 01-08-2012 and week 44 with 26-12-2012.