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 will revise 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 permittivity of water (≈80F/m) as compared to mineral soil solids (≈2-9F/m), and air (≈1F/m). 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 will modify the text in the revised version of the manuscript to clarify this.

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 will be clarified in the revised version of the manuscript.

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: 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). This will be clarified in the revised version of the manuscript.

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. We will add an explanation in the paper.

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. An explanation will be added in this paper.

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 will revise this part in our manuscript.

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

Reply: We will add this information as suggested by the reviewer.

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]. We will test the effect of initial uncertainties on the performance of EnKF, e.g. increasing the forcing error from 10% to 20%. The additional simulation results will be explained in the manuscript. If larger changes would be observed, a more detailed discussion will be included.

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 will test other values for parameter *s*, like 0.01, to see how it influences the performance. The additional simulation results will be explained in the manuscript. If larger changes would be observed, a more detailed discussion will be included.

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.

Reply: 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 will re-evaluate this part and try to formulate more precisely.

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 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, not the difference but 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. We will add an explanation in the paper.

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 will show the evolution of the state ensemble in our manuscript and address this question.

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 will further discuss these results in the paper.

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 in 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] have a similar conclusion. In our work, 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. An explanation will be added in the paper.

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: 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 will show the evolution of the parameter ensemble from time step 0 onwards in the revised version of the manuscript. 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 0.053m^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. We will extend the discussion of these results.

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 will provide additional discussion of the results and discuss reasons for the larger deviations in the fit.

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 will evaluate the impact of filter settings on the results of this paper. However, it will be difficult to evaluate the significance of the difference between the DA-algorithms on the basis of this single study.

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.

Reply: We will revise this part in our manuscript.

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 will be added in the revised version of the manuscript.

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 will be 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 will be corrected in the revised version.

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. We will revise this part in our manuscript. 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.

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: The unmodified Richards equation is:

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

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

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

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

***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:**

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

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

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

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

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

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 would not be presented in the revised manuscript and would be referred 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. Soil layer definition where soil moisture is calculated 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) to 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 t 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.

When 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.

When 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:

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

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:

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

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

$c_i = 0$

$r_i = q^t_{i\text{-}1}$

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

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

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

Reply: We admit that 4D-Var is also commonly used. The sentence will be adapted to account for this, and we will specify shortly in which contexts the different methods are normally 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 will revise and extend this part in our manuscript.

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 will clarify 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 will reformulate the text to make this clearer.

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 includes a state vector and a parameter vector. 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?

<u>Reply:</u> 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.

<u>Reply:</u> 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 will modify the sentence to accommodate this result.

All Technical corrections are applied in our revised manuscript.

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.