Reply to comments from Anonymous Referee #1

General Comments

The manuscript investigates the effect of the uncertainty of the initial conditions in the context of soil water movement described by the Richards equation. First the necessary warm-up times for different soils and climates are determined and then the effects of different methods to describe the initial condition on a subsequent data assimilation are compared. The comparison is additionally shown on a real-world case. I think the manuscript is interesting and shows the effects of the uncertain initial conditions nicely. I have few comments that may require some additional investigation or discussion. However, the manuscript is sometimes difficult to read and could be clearer. Therefore, many comments ask for some clarification.

[Response]

Thanks for your positive comments. We have improved the manuscript according to your suggestions.

Major comments:

1. The required computational power varies between the different initial conditions. The most expensive ones (with warm-up period) seem to give the best results in the subsequent data assimilation. I would find it very interesting whether this finding holds if for each method a similar total computation time (computation time for initial condition + computation time for data assimilation) is available. This means that e.g. IC-ObsInt or IC-Flux could afford more ensemble members than IC-WUP. The question is then, if e.g a higher number of ensemble members (in combination with a larger uncertainty in the representation of the initial condition) of IC-ObsInt or IC-Flux could lead to similar, or even better results.

[Response]

Thank you for your valuable suggestion. First of all, we must apologize that we have used mistaken model input folders when analyzing the data from WUE and WUP, and two identical curves for WUE and WUP were generated. Thanks for raising this question, leading to the discovery of this mistake. Nevertheless, the general conclusions from this figure still hold. Several minor modifications are given below (please see our response to your last comment) and will be reflected in the revised manuscript.

Then, in order to investigate whether the results for each method holds if with a similar computation time, we have added the cases for lnK_s estimations utilizing the initial condition IC-HfSatu and IC-ObsInt with the ensemble size of 500 (hereafter referred as IC-HfSatu-500 and IC-ObsInt-500) in the manuscript. The computational costs of them are similar with that of IC-WUP, IC-WUE and IC-Flux. According to the results, we added a new figure and corresponding discussion in the manuscript about the effects of ensemble size on parameter estimation.

[Changes in the manuscript]

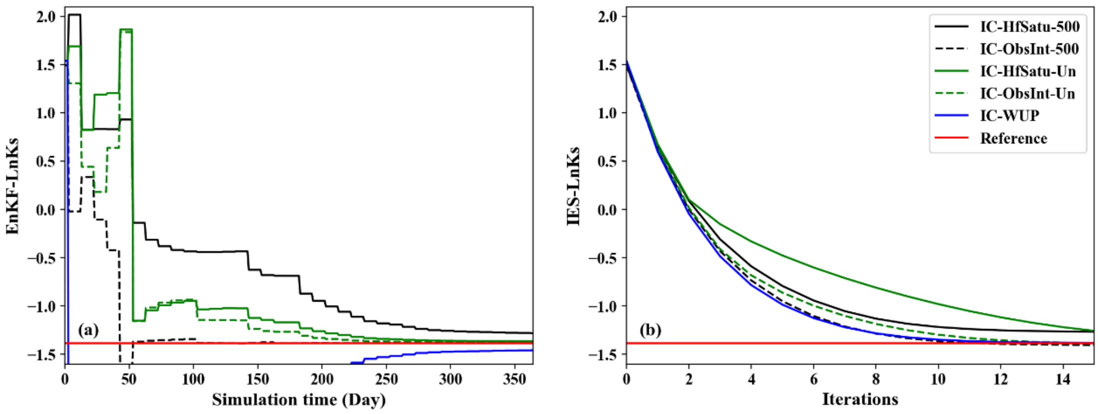

Figure a1. The impacts of ensemble size and the uncertainty of initial ensemble on the results of lnK_s estimations within EnKF and IES.

The effects of increasing ensemble size are totally different for EnKF and IES. In Figure a1(a), when the ensemble size grows to 500, the assimilation results of IC-HfSatu-500 are worse than that of IC-ObsInt-500, implying that the initialization methods still plays an important role in parameter revisions even with a larger number of ensemble size. Besides, compared with the IC-ObsInt and IC-HfSatu with the ensemble size of 300 (Fig. 6(a)), IC-ObsInt-500 and IC-HfSatu-500 both show better data assimilation results, indicating that the results of parameter estimations can be improved by increasing ensemble size for EnKF. The results are reasonable since the cross-correlation between model parameters and states can be better calculated with a large number of realizations. With the similar computational cost, the UIC of IC-ObsInt-500 stabilize the update steps, making the results better than IC-WUP. On the contrary, the impacts of improving ensemble size are slight for IES. As plotted in Figure a1(b), the data assimilation results of IC-HfSatu-500 and IC-ObsInt-500 are similar with that of 300. Since IES is a kind of iterative history-matching algorithm, the nonlinear relationship can be well calculated during the iterations, while the UIC is existed in the whole simulation. Therefore, warm-up methods show better data assimilation results within IES.

2. Line 222-226: When the initial condition ensembles are generated for IC-HfSatu, IC-ObsInt and IC-Flux, is uncertainty added? How exactly? The uncertainty in the initial water content must be represented in the initial ensemble. If no uncertainty is added, this could explain partly the inferior result compared to IC-WUP. Please clarify and discuss.

[Response]

Thank you for your careful reading. We are sorry that we did not explain the problem clearly.

(1) The initial conditions of IC-HfSatu and IC-ObsInt were assumed to be deterministic without uncertainty in the last manuscript. In contrast, IC-Flux was conducted by warming up the model for a period (length = “warm-up” time t_{wu}) with a constant infiltration flux until a steady-state soil profile can be obtained. Thus, the uncertainty of parameter is introduced to IC-Flux, IC-WUP and IC-WUE during the construction of initial ensembles.

(2) The IC-HfSatu and IC-ObsInt were assumed to be deterministic without uncertainty. They are the most common and convenient methods to initialize the soil water/hydrological model, while most applications of these two methods do not consider including the parameter uncertainty (much larger than the magnitude of observation error) during the construction of initial conditions. We want to know how

exactly they affect the data assimilation results and whether we can utilize these simple methods to initialize the model within data assimilation framework.

(3) In order to explore the effects of the uncertainty of the initial ensemble on the parameter estimations, the standard deviation of initial ensemble from IC-WUP is counted (0.017). Then, a Gaussian noise (with a standard deviation of 0.017) is added to both IC-HfSatu and IC-ObsInt (hereafter referred as IC-HfSatu-Un and IC-ObsInt-Un). We compared the results of parameter estimations between them and the other initial conditions, as presented in the updated Figure a1.

[Changes in the manuscript]

(1) We will add the content "It should be noted that IC-HfSatu and IC-ObsInt are assumed to be deterministic without uncertainty, while for the IC-Flux, the steady state is constructed by warming up model for a period of time which introduces the uncertainty of parameters"

(2) We will add Figure a1 in the manuscript. As plotted in Figure a1, the data assimilation results of IC-HfSatu-Un and IC-ObsInt-Un are better than those of IC-HfSatu and IC-ObsInt (Fig. 6(b)), indicating that the results of parameter estimations can be improved by increasing the variance of initial ensemble for EnKF. That may because the covariance between parameters and states can be better calculated with a large uncertainty of initial ensemble. However, with respect to IES, the assimilation results of IC-HfSatu-Un and IC-ObsInt-Un are worse than those of IC-HfSatu and IC-ObsInt, implying that a large variance of initial ensemble may deteriorate the data assimilation results within IES, since this uncertainty is difficult to diminish during the iterations. Hence, the optimal initialization method for EnKF may be the ObsInt approach with a large ensemble size and suitable uncertainty, while for IES, warm-up method can be a better choice, since the uncertainty of initial condition and parameter can both decrease during the iterations.

3. Line 247: The spatial resolution of 5cm is rather low for a 1 dimensional case. Is this a computational limitation? Otherwise, I would suggest to reduce the grid size to e.g.

2cm. This is especially relevant for sandy soils where sharp infiltration fronts can develop and require such high resolutions. Could this impact the results?

[Response]

Thank you for your suggestion. To understand the effects of grid size on variably saturated flow, we compared the results of temporal change of the soil moisture profile for the sandy soil with the grid size of 2 cm and 5 cm, as presented in Figure below.

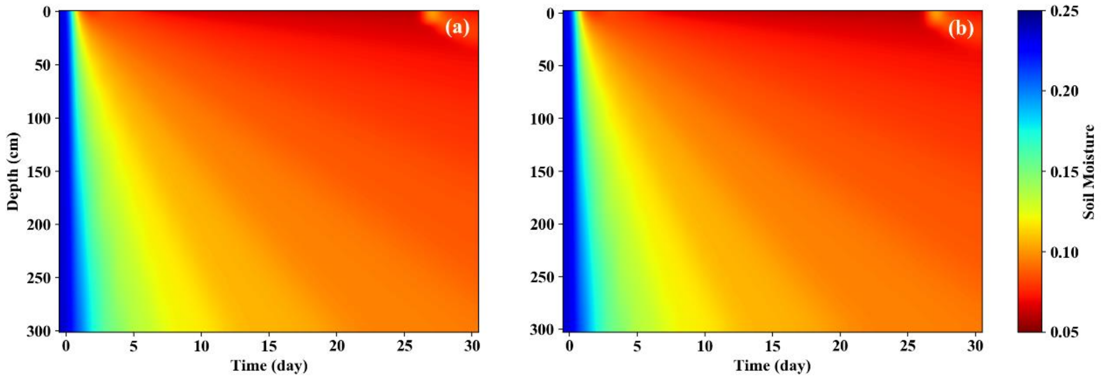

Figure 2. The temporal change of soil moisture profile for the sandy soil with the grid size of 2 cm (a) and 5 cm (b).

In general, the soil moisture values at various depths all drop significantly in a short time, since the sandy soil has a great ability of drainage. Besides, owing to the effects of meteorological condition, the soil moisture at the surface decreases more quickly than that at the deep layer. However, compared with the model outputs using the grid size of 2 cm (Figure 2(a)), the results with the grid size of 5 cm (Figure 2(b)) are extremely similar, showing a slight impact of grid size.

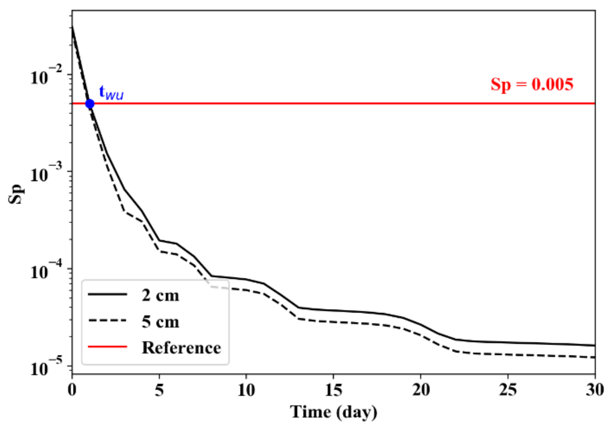

Figure 3. The spread index value for the sandy soil over time with the grid size of 2 cm and 5 cm.

In order to explore the effects of grid size on warm-up time, the spread value over time for the sandy soil is plotted in Figure 3. Generally speaking, the spread value with the grid size of 5 cm shows a similar trend with that under the grid size of 2 cm. Besides, although there is a deviation between two spread values after 2 days' simulation, the deviation is too small that can be ignored. Furthermore, the t_{wu} value of both cases are the same. Hence, the grid size has little effects on the soil moisture simulation and t_{wu} for the sandy soil in our case.

[Changes in the manuscript]

(*previous manuscript*) The flow domain is discretized into 60 grids with a grid size of 5 cm.

(*revised manuscript*) The flow domain is discretized into 60 grids with a grid size of 5 cm which has been proved to be sufficient for one-dimensional soil water flow modeling in our case (results not shown).

4. Lines 379-383 and Figure 7: The biases of Ks, α, and n as well as their uncertainties differ. Therefore, their RMSEs should not be compared directly. I think it is not a meaningful result that α, which has the largest initial bias and uncertainty, also has the larger RMSE and that n, which has the smallest initial bias and uncertainty, has the smaller RMSE. Their relative improvement might be a better measure.

[Response]

Thank you for your suggestion. As the reviewer pointed out, *RMSE* index may be not suitable to evaluate the data assimilation results of different parameters, since their uncertainties are different. To identify the improvement of data assimilation results, we will add the relative error index (*RE*) into the manuscript, which can be calculated as follow,

$$RE = \frac{RMSE_e}{RMSE_p} \tag{1}$$

where $RMSE_e$ and $RMSE_p$ represent the *RMSE* of the estimated and prior parameters. *RE* varies from 0 to positive infinity. As *RE* approaches to 0, the analysis results are considered to be identical, but a large value of *RE* (more than 1) indicates that the results of parameter estimation are inferior to the prior parameters. Compared with the *RMSE*, this index can better present the impacts of different types of parameters on the data assimilation results. Hence, we modified Fig. 7 and Fig. 8 in the manuscript.

[Changes in the manuscript]

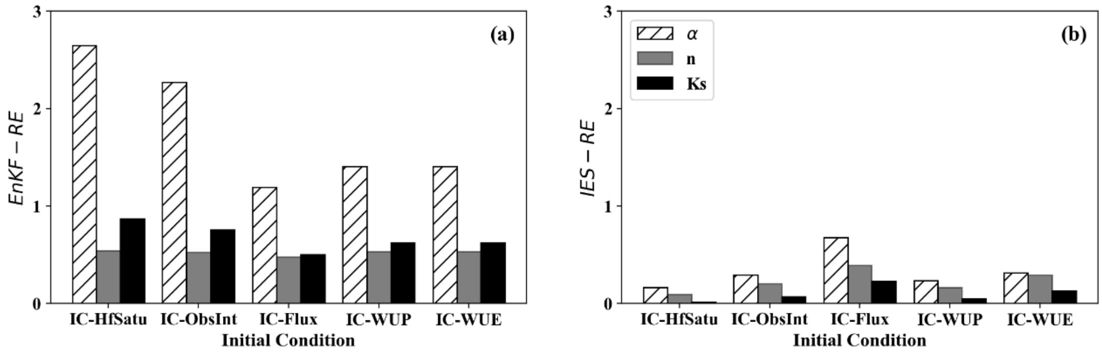

Fig. 7. The *RE* results of parameter estimations (α, n and K_s) under five initialization methods with (a) EnKF and (b) IES.

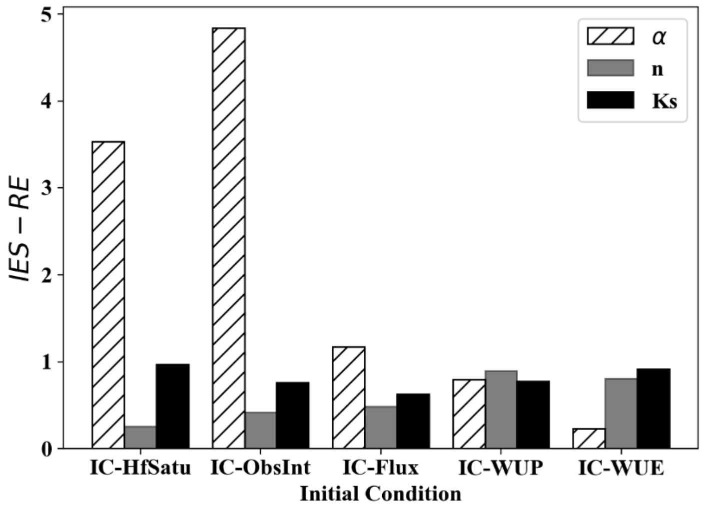

Fig. 8. The *RE* results of parameter estimations under five initialization methods with IES when the simulation time is 60 days.

5. Line 286: You find that t_{wu} is less than one day for sand. I think that this result might be due to the chosen initial condition for the warm-up. It is true, that for the chosen high water contents sand will drain very fast and rapidly approach a similar water content state. However, in case of an initial condition in a very dry state (which should be relevant for the arid climate), the hydraulic conductivity of sand drops to very low values and the initial spread can extend for a very long time, or until a sufficiently large *rain event increases the water content and then leads again to the rapid approach of the similar water content. I think it would be interesting to investigate this by choosing a different (dry) initial condition. At least this should be discussed in the manuscript.*

[Response]

Thank you for your valuable comment. To investigate this problem further, we conducted another case of Monte-Carlo simulations for sandy soil with a wet and dry initial condition (i.e., the mean of soil moisture ensemble is 0.235 and 0.15 respectively with a standard deviation of 3%). The temporal change of the soil moisture ensemble and the corresponding t_{wu} are presented in Figure 3.

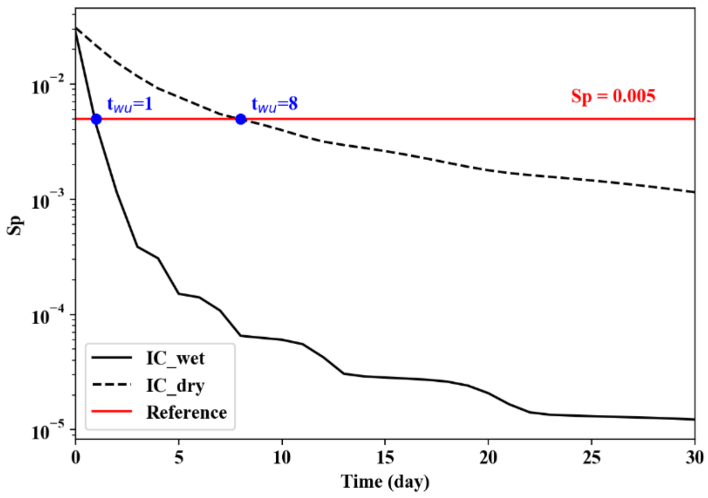

Figure 4. The temporal change of spread for sandy soil with a wet and dry initial condition separately (i.e., the mean of soil moisture ensemble is 0.235 and 0.15 respectively with a standard deviation of 3%).

As the reviewer pointed out, the warm-up time t_{wu} for the wet and dry condition is 1 day and 8 days, implying that a more dry initial condition would lead to the increase of warm-up time. The results are reasonable since with a high soil moisture, the sand would drain fast, making the soil moisture ensemble approach to a similar state. However, regarding the initial condition in a very dry state, the hydraulic conductivity of sand drops to very low values and the initial spread can extend for a very long time.

[Changes in the manuscript]

We will add the discussion about the effects of the mean of the initial soil moisture ensemble on the warm-up time in Section 3.2.2.

6. Line 273-275 and 477-478: Since you recommend the choice 0.5% as a threshold: Please explain why. What is the advantage? Why should I not choose the other mentioned thresholds (e.g. 1% or 0.1%)?

[Response]

Thank you for your careful reading. The threshold of 0.5% is recommended mainly because we want to balance the estimation accuracy and computational cost. An explanation is added in Section 3.2.2.

[Changes in the manuscript]

As shown in the Fig. 3(c), there is still a big diversity of the results between Spin-up and Monte-Carlo methods at index value of 1%, indicating that the UIC still plays a significant role and may deteriorate the model outputs at this stage. Regarding the threshold of 0.1%, the requested t_{wu} is more than 15 months which may introduce a large computation cost.

7. Lines 306-313 and Figure 5: If I understand correctly, you investigate when the uncertainty for the full profile drops below the 0.5% threshold. In addition (or possibly as replacement) I would find it interesting to see the spatially resolved times for each depth for the deepest profile (20 m).

[Response]

Thank you for your valuable comment. As the reviewer pointed out, the t_{wu} is calculated by taking the uncertainty of the full profile into considerations in Fig. 5. We will add a new subfigure (Fig. 5(b)) in the manuscript, which presents the t_{wu} value for each depth with the profile length of 20 m, as presented below. The result is similar with our previous conclusion, showing that more warm-up time is needed if the soil layer is deeper, and we should be very careful to deal with simulation with a long unsaturated profile.

[Changes in the manuscript]

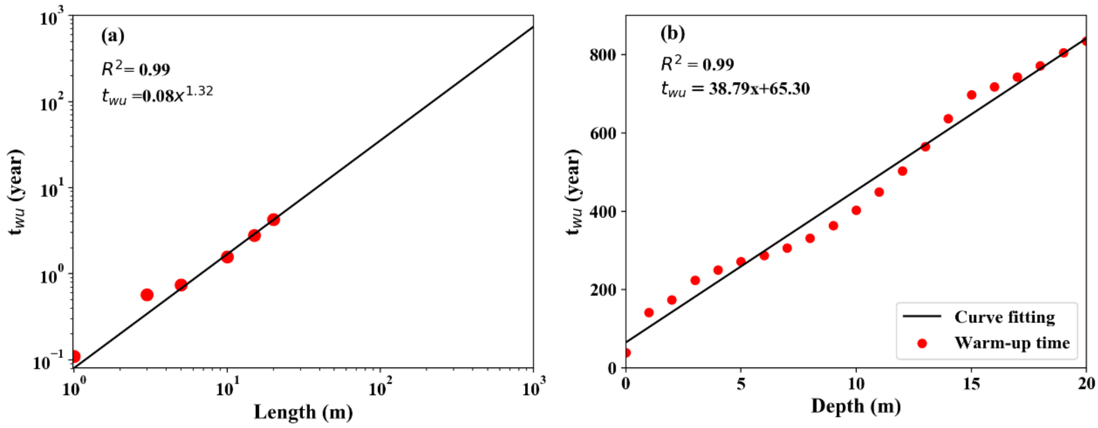

Fig. 5. The value of the warm-up time t_{wu} for (a) various soil profile lengths and (b) each depth with a 20-m soil profile.

8. Line 152 (Equation 8): If the monthly average of the previous year is required, wouldn't that imply that PC is not defined for the first year? However Fig. 3 shows PC starting from time 0. Please clarify.

[Response]

Thank you for pointing out this problem. We are sorry that we did not explain it clearly. PC is an index that reflects the deviation of soil moisture between two adjacent years after a period of warm-up time t_{wu}. Following Ajami et al. (2014), PC at t=t_{wu} month is calculated by comparing the relative difference of soil moistures at t=t_{wu} month and t =t_{wu} + 12 month. As presented in Fig. 3(a) in the manuscript, PC at t=12 month is close to 0.

[Changes in the manuscript]

We will update Equation 8 as

$$PC(t) = 100\left|\frac{M(t) - M(t+12)}{M(t+12)}\right| \tag{8}$$

where $M(t)$ is the monthly mean of soil moisture after model spins up for t months and $M(t$+12) is the monthly average of soil moisture after M(t) one year later.

9. Figure 3: Why does the water content state after 24 months differ between panel (a) and (b)? Since both are initialised with the same parameter values and the UIC has *already decayed, they should show the same soil moisture. Please clarify.*

[Response]

Good eyes! Thank you for pointing out this problem. We made a mistake when calculating the monthly-average soil moisture at the t = 24 month in Fig. 3(b). This error is amended in the updated figure.

[Changes in the manuscript]

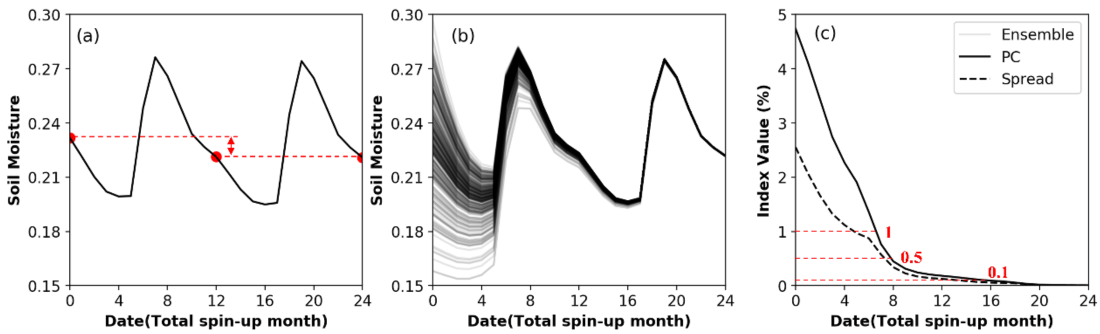

Fig. 3. Comparison of Spin-up and Monte-Carlo methods in determining warm-up time. (a) Spin-up method with monthly-averaged soil moisture versus time by running a simulation recursively for several years, (b) Monte-Carlo method with monthly-averaged soil moisture of different realizations versus time based on various initial conditions, and (c) Comparison of PC and S_p versus time. For the purpose of demonstration, the parameter uncertainty is not considered.

10. Lines 370-377: Based on Figure 6, I disagree with the statement that filter inbreeding is not a significant issue for the EnKF case. In Figure 6, it seems that the final parameter value does not change any more over time and is over 5 standard deviations away from the truth. This means that the uncertainty is too small. Part of the reason could be that the initial uncertainty is chosen way too small. It is over 9 standard deviations away from the true value. This makes it very difficult for the EnKF to find the true value. I would suggest to repeat the simulations with a larger parameter uncertainty.

[Response]

Thank you for your suggestion. In order to explore the effects of parameter uncertainty on the data assimilation results, we compared the parameter estimations of lnK_s with various standard deviations of initial parameters (0.1 and 1 respectively). We agree with you that the data assimilation results of IC-ObsInt, IC-HfSatu and IC-Flux are improved since the cross-correlations between parameters and states are better calculated with a larger parameter uncertainty. However, regarding the warm-up methods, the prior parameters can be updated greatly at the first assimilation step, leading to the filter inbreeding of model, even with a larger parameter uncertainty (Figure 5). The results are reasonable since the UIC is diminished with the warming-up procedure, leading to a strong cross-correlation between estimated parameters and states. Hence, the ensemble of parameters and states converge too quickly. Compared with IES using a Levenberg-Marquardt algorithm, it is a kind of defect of EnKF without a damping coefficient for updating.

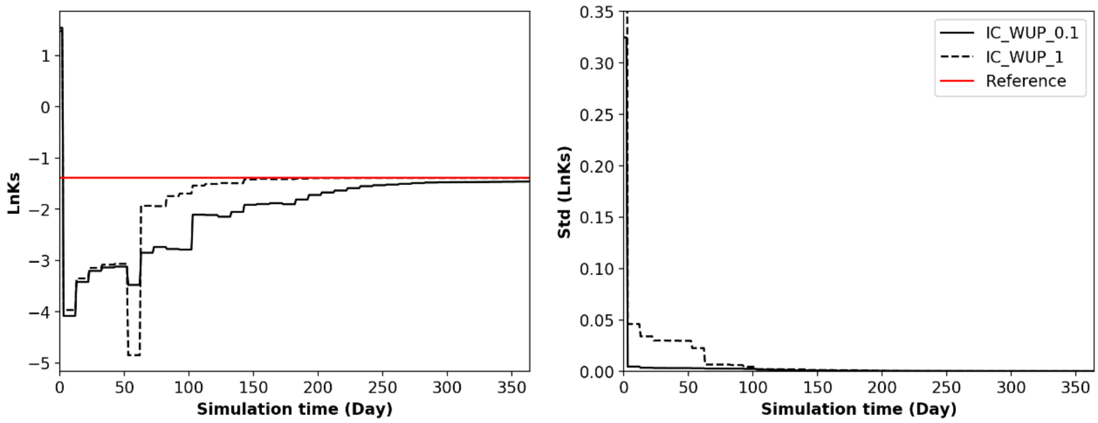

Figure 5. The temporal change of estimated mean and standard deviation of lnK_s with various parameter uncertainties.

[Changes in the manuscript]

We will revise our discussion about filter inbreeding in Section 3.3.3.

11. Line 177-178: ": : :, uk are state variables (i.e., pressure head and soil moisture) : : :". Do you update water content and matric potential of the same node simultaneously in the augmented state? Due to their nonlinear relation, the analysis would lead to inconsistencies between water content and matric potential for the analysis. How is this handled in the forecast? Please clarify.

[Response]

Thank you for your careful reading. In this study, we only update the soil moisture in the simulation and converse the updated soil moisture ensemble into the pressure head to drive the model. The above statement may be confused and we have revised it.

[Changes in the manuscript]

(*Previous manuscript*) $\mathbf{u}_k$ are state variables (i.e., pressure head and soil moisture) at time t_k.

(*Revised manuscript*) $\mathbf{u}_k$ are the state variables (i.e., soil moisture) at time t_k. The updated soil moisture ensemble can be conversed to pressure head to drive the model.

Technical comments:

Lines 228-238: This part describes IC-WUP and IC-WUE. However, this is not a general description. In Section 3.2, when the spin up periods are investigated, a different procedure is used. This confused me when reading the paper the first time. Please, only mention the general settings in 3.1 (i.e. climates, soils and model representation), and not specifics that only apply to 3.2 or 3.3. Therefore I would suggest to move this part to Section 3.3. Additionally, here it is not clear how the parameter and initial condition ensembles are exactly generated. Please clarify.

[Response]

Thanks for your valuable comments. We are sorry that we did not make the description clear.

IC-HfSatu is a uniform soil moisture profile with the 50% relative saturation of soil (e.g., a uniform soil moisture value of 0.254 over the profile for loam); IC-ObsInt is a linear interpolation between observations at the beginning of simulation. The depths of the initial observations are 10 cm, 80 cm, 150 cm, 220 cm and 290 cm; IC-Flux is a steady-state soil moisture profile by warming up the model with a constant infiltration flux (1 mm/d). Besides, the initial conditions of two warm-up methods are given by running the model prior to the beginning of simulation period with available meteorological data (as shown in Fig. 1). If the meteorological data before the simulation period is available, it is used in the warm-up method to obtain the initial condition (IC-WUP); otherwise, we use the meteorological data at the simulation period

(IC-WUE) as a surrogate. The length of warm-up time for IC-Flux, IC-WUP and IC-WUE is equal to t_{wu} (242 days) according to the results in Section 3.2.2(a). In addition, IC-HfSatu and IC-ObsInt are assumed to be deterministic without uncertainty, while for the IC-Flux, IC-WUP and IC-WUE, the uncertainty of parameters are introduced by warming up the model.

The initial realizations of soil hydraulic parameters K_s, α and n are generated following logarithm normal distributions, with mean values of 4.7 md^{-1}, 8.6 m^{-1} and 1.8, and variances (log-transformed) of 0.1, 0.3 and 0.006. The saturated soil moisture θ_s and residual soil moisture θ_r are assumed to be deterministic with the value of 0.43 and 0.078.

[Changes in the manuscript]

We will add the explaination in the manuscript and remove this part to Section 3.3 as your suggested.

Line 222-226: I think this part should be moved to Section 3.3 as well.

Thanks. This has been rewritten.

Line 243: "Fig. 1" should be "Fig. 2".

Thank you. The error has been corrected.

Line 254 and Fig. 3: The text mentions a simulation length of 10 years, the figure shows only 2 years. I would suggest to mention that you only show the first 2 years.

Thanks. This has been revised.

Lines 338-342: How many observations are there? In what depths are the observations? What is the assimilation frequency? Or is only a single observation in the depth of 10 cm assimilated every 10 days? If that is the case this has to be clarified.

[Response]

Thank you for your suggestion. Indeed, the observations are only collected at the depth of 10 cm and assimilated every 10 days. Unless otherwise specified, the total numbers of the observations are 37 (3rd, 13th, 23th,…, 363th days).

[Changes in the manuscript]

(*Previous manuscript*) In addition, the observation at 10 cm is assimilated into model every 10 days.

(*Revised manuscript*) In addition, a total number of 37 observations are assimilated into the model. The depth of the observation is a single point at 10 cm and the assimilation frequency is 10 days.

Lines 343-350. I think this part should be moved to methods in Section 2.

Thanks. Revised.

Line 352-353 and Figure 6: I would mention that this is case 1 and case 2.

Thanks. This has been modified.

Line 399: "Field" instead of "Filed".

Thank you. Revised.

Figure 4: Since essentially the times for sand for all climates as well as silt and clay loam for the M-Ac and the M-SC climate can not be properly displayed: Maybe the logarithm of the time could be more meaningful (like in Fig. 5).

[Response]

Thank you. Fig. 4 has been revised in the manuscript according to the suggestions from you and another reviewer. We added a case to investigate the t_{wu} value in a multiple-layers soil.

[Changes in the manuscript]

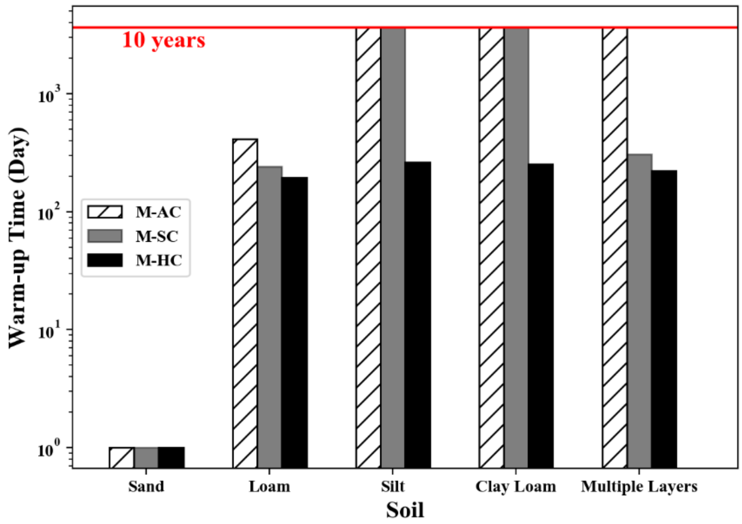

Fig. 4. The length of warm-up time t_{wu} with various soils and meteorological conditions. Note that t_{wu} of Silt and Clay loam with M-AC and M-SC exceed 10 years as well as t_{wu} of multiple layers with M-AC. The consistent layers of heterogeneous soil are the loam (0-75 cm), clay loam (75-150 cm), silt (150-225 cm), and (225-300 cm).

Figure 6: The line for IC-WUE is essentially not visible. Is it below IC-WUP? At least mention this in the caption.

Thank you for pointing out this problem. We are sorry that we have used mistaken model input folder when analyzing the data from WUE and WUP, and two identical curves for WUE and WUP were generated. Although there is a little difference in Fig. 6 (a) and (c), the general conclusions are consistent with the previous ones.

[Changes in the manuscript]

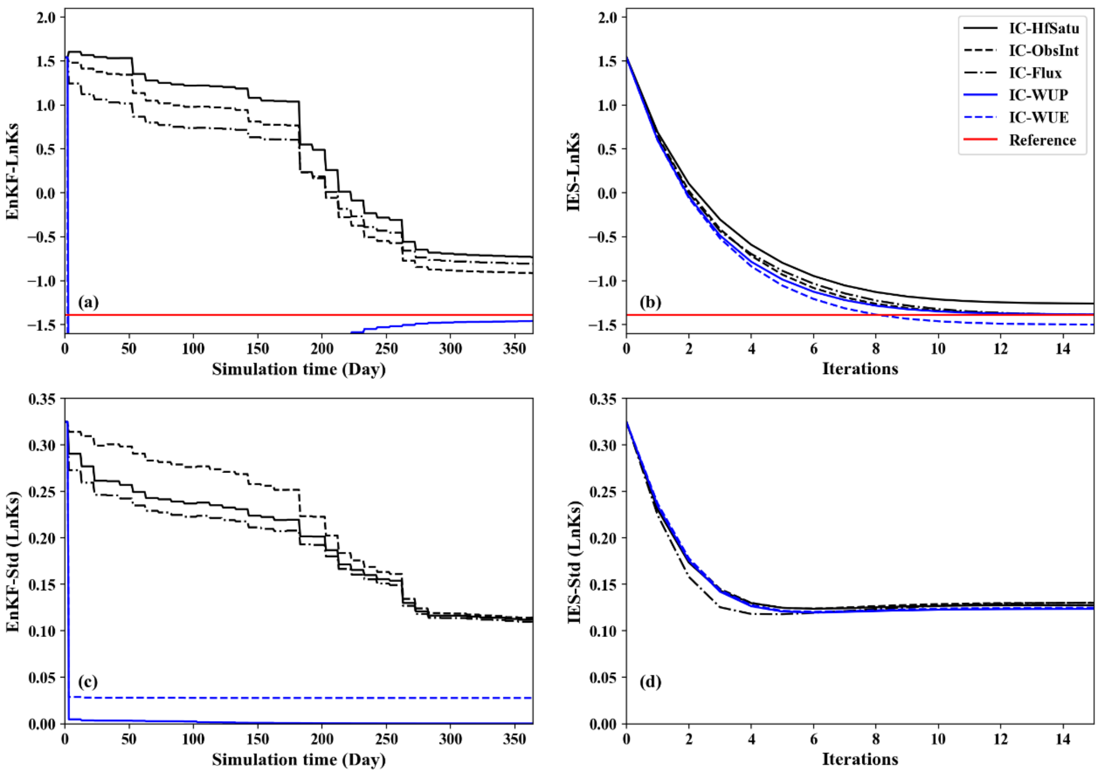

Fig. 6. The results of lnK_s estimations (first row) and their associated standard deviations (second row) within two data assimilation frameworks (left: EnKF; right: IES) under five initialization methods.

In Fig. 6(a), the final parameter estimation results of the IC-WUP are still the best among the five initialization methods within EnKF. Regarding the IC-WUE, the value of estimated parameter approaches to and crosses the true value rapidly (the results of IC-WUE are below the IC-WUP at the beginning of the simulations, but lower than IC-WUP in the late period). The results implies that the update steps of warm-up methods tend to be extremely large during the simulations, because when UIC diminishes, there is a strong cross-correlation between parameter and state, leading to a large update step and filter inbreeding for EnKF. Whereas within IES, there is a stability multiplier at every iterations which promises a robust estimation of parameter for warm-up method (Fig.6 (b)).

Reference

Vereecken, H., Kamai, T., Harter, T., Kasteel, R., Hopmans, J. and Vanderborght, J.: Explaining soil moisture variability as a function of mean soil moisture : A stochastic unsaturated flow perspective, , 34(October), 1–6, doi:10.1029/2007GL031813, 2007.

Ajami, H., McCabe, M. F., Evans, J. P. and Stisen, S.: Assessing the impact ofmodel spin-up on surface water-groundwater interactions using an integrated hydrologic model, Water Resour. Res., 50, 1–21, doi:10.1002/2013WR014258.Received, 2014.