

We thank the reviewers for their constructive comments and suggestions. Our answers are the same as the ones given in the interactive discussion. The revised version is the same as the one already shown in the supplement there. We provide the marked-up manuscript in the following.

5 **Answer to Review by Anonymous Referee #1**

Summary: *the manuscript presents a study on parameter estimation for an unsaturated zone model. The authors use real soil water content data to estimate parameters and forcings for a 4 layer model using the EnKF. A full iteration of the filter, divided into different steps is introduced to improve the estimation process. During a rain event it is observed that the estimation gets flawed by the models inability to model preferential flow and the observations corresponding to this rain event are excluded from the parameter estimation. This is called a closed-eye period and it is observed that this improves the stability of the estimated parameters.*

General comments: *The manuscript is interesting and is filling a missing space in the literatur as being a reasonably controlled test case using real data. However, the manuscript is sometimes difficult to overview and the proposed method feels very ad hoc and leaving a number of open questions. Suggestion is major revision.*

10 (throughout the text references to the manuscript are given as: P9,L78 = page 9 Line 78)

Reply: We thank the reviewer for the constructive comments and suggestions. We have revised our manuscript taking them into account. For consistency with your references, we will still refer to the unrevised manuscript as well.

Major concerns

20 **1. Data assimilation vs calibration**

a. The manuscript introduces the concept of a fully iterative filter, which is an approach that I have never seen in the literature before (and obviously the authors neither since there are no citations on this part). Iterative filters are not uncommon, but they are usually local and not global, or restart versions where the model is restarted and ran until next observations. Some relevant references and comparisons to other iterative approaches would be needed as well as a good motivation to the current approach. The authors approach also shows clear similarities to the precalibration exercise by Hubert et al 2010, to the iterative Kalman Ensemble generator by Nowak 2009 (see bottom for references) and to the SODA by Vrugt et al 2005 (already cited, but not compared to the suggested approach).

Reply: Yes, we have not seen the fully iterative approach in the literature as well. We improved the introduction of the approach (page 9, line 12-15) to: “We iterate the whole EnKF scheme and start the next iteration with the final estimation of soil parameters, Miller scaling factors and upper boundary condition of the previous iteration. These iterations differ from the typically applied iterative EnKFs like the Restart EnKF or Confirming EnKF (e.g. Song et al., 2014) and rather resemble the iterations in the Kalman ensemble generator by Nowak (2009), who used a modified EnKF to estimate parameters only. The full iterations applied are required to estimate constant augmented state components even with a small damping factor and a rather short time period of data. This is especially important for the upper boundary condition, where the time periods of constant values are short.”

We have not been aware of the works by Huber et al. (2011) and Nowak (2009). Thank you, for pointing them out to us. We have incorporated Nowak (2009) in the text shown above (page 9, line 12-15) and Huber et al. (2011) in the description of the prior estimation (page 10, line 14): “Highly uncertain properties can exacerbate the performance of the EnKF. For example Huber et al. (2011) showed that the state estimation with an EnKF is superior in case of properly precalibrated model parameters. In our case, there is no prior knowledge about small scale heterogeneity, which can lead to difficulties in the estimation process:”

The SODA approach by Vrugt et al. (2005) is comparable to the Standard EnKF approach in our manuscript, since he as well estimated parameters, while continuously updating the states with an EnKF. But, the estimation with the augmented state instead of the shuffled complex evolution metropolis algorithm is more computational efficient. We added this in the description of the Standard EnKF (page 11, line 6).

b. Further, the suggested approach seems to me a lot more similar to a batch calibration than a filtering, as the same data is used over and over again to calibrate the model towards one final parameter estimate. As it to me is unclear what effect the top boundary estimation has (see below), and as the truly dynamic parts of the data are removed, I have a problem seeing why

a filter and not a proper calibration would provide a good solution, given that the filters after all are after all suboptimal for the unsaturated zone?

Reply: Yes, our approach is not a typical filter application for state estimation. Similar to a batch calibration we focus on the parameter estimation, but we argue that our approach goes beyond. Unlike the batch calibration, the EnKF continuously updates the state and could incorporate model errors. Due to the continuous parameter estimation we were able to detect the times when the model assumptions do not hold (the closed-eye period). Due to the state estimation we can remove these times (which in this case correspond to the dynamic parts of the data). The state estimation enables us to guide the state through this time with a erroneous model and allows us to pick up the estimation afterwards again.

The estimation of the boundary condition is required to prevent the introduction to biases into the parameters. This again is required for a proper detection of the closed-eye period (see below).

2. Top boundary and its update

a. The top boundary leaves a few questions unanswered. It is claimed throughout the manuscript that the top boundary is updated and that this is largely also the reason for the use of a data assimilation method rather than a calibration. However, the authors neither show nor discuss the result of these updates, leaving a reader wondering about the necessity of this inclusion.

b. Further, I do not understand what boundary was used to drive the model. On P8,L10 it is written that the temporal resolution is daily unless a change of mode occurs. On the other hand, all the figures show finer resolved top boundary, and also the scenario descriptions are dealing with temporal resolutions of 10 minutes when describing the rain intensity. If these events are smoothed across a full day (or if the event is shorter than a day, across a full event), it is a considerable smoothing, and it would be much more preferred if the authors also plotted the forcings they are using. Further, wouldn't this risk effecting the parameter estimation during the rain event, if reality has a much stronger peak than the model? Or is the first observation so far down that this has no effect? Please clarify!

Reply: We agree, that we did not discuss the estimation of the upper boundary in sufficient detail in the manuscript. An estimation of the upper boundary condition is not mandatory. A stochastic representation of the uncertainties could be sufficient. But we argue, that a bias in the upper boundary condition could induce a bias in the estimated parameters. Especially if this bias mainly occurs during the rain event, it could cause a parameter shift during the rain event, similar to the one observed during the closed-eye period. To prevent this, we do estimate the boundary condition. This requires the decreased temporal resolution, which could lead to short-term parameter variations within the rain event. We think that this downside is subordinate to the removal of the bias, especially since it is strongly alleviated by the dissipative soil water movement.

We improved the description of the boundary condition in the manuscript (page 8, line 5-11): “The expansion to an augmented state changes the propagation in time. Each component needs an individual forward propagation. We assume the soil hydraulic parameters and Miller scaling factors to be constant in time. This is not possible for the upper boundary condition where the forward equation is unknown from a soils perspective. However, measurements are available to estimate the evaporation and precipitation. Hence, we assume the forward model constant until a new estimation is available. Then, we switch to the estimated boundary condition. To base the improvement of the upper boundary condition on several measurements, we reduce the temporal resolution of precipitation measurements to change daily and at transitions between precipitation and evaporation. This means, that the upper boundary condition is treated like the parameters, except that the value can change in the forward propagation. The original temporal resolution of the precipitation data (10 minutes) is not required due to the dissipative nature of the Richards equation, which smooths the infiltration front before it reaches the first TDR sensor in a depth of 8.5 cm. The estimation of the averaged boundary condition will ensure that there is no global bias on the parameter estimation during the rain event, but could lead to small short-time parameter drifts within a rain event.” and (page 11, line 34): “As we added the boundary condition into our estimation, we can exclude effects due to a bias in the precipitation to cause this parameter shift. The reduced temporal resolution in the boundary condition could only cause short-term parameter changes within the rain event.”

We also show the boundary condition with the applied reduced temporal resolution in Fig. 8 now.

3. Purpose of model?

A nice and stable model is calibrated, that can predict water movement as long as nothing really happens. I find the result very interesting, as it properly questions the use of the standard Richards equation for modelling the unsaturated zone, given that the

rainfalls that caused the problems here are quite moderate. The authors should take their space to consider the implications of what they show; can we use the Richards equation in the field? What purpose does a model have that is seemingly data driven (P11,L15) and that cannot predict during rainfall?

5 **Reply:** Yes, we agree that the Richards equation cannot describe preferential flow, which limits its applicability. Quite to the opposite preferential flow is mostly relevant for the top layers only and less for the lower layers. Consequently the Richards equation is valid there and can be applied.

In our manuscript we are interested in the actual processes. We separated the times with and without preferential flow and show that the Richards equation can really describe the times without preferential flow.

10 Furthermore, if one is not interested in the processes, but an optimal prediction a batch calibration can yield better results, because there effects of preferential flow are incorporated into the parameters, which then in part compensate the non-represented processes.

4. Memory

15 *Since the model used is disconnected from the groundwater, all information within the model has to travel from the top to the bottom in the 1D column. As the lower layers are more dependent on a longer memory to properly assess their behavior (e.g. it takes longer before we see an effect of a rain event in the second layer than in the first), there is a good risk that we smooth this information when continuously updating the first layer.*

a. *Wouldn't a continuous update of the top layer risk always smoothing the model such that the estimation in the bottom layers becomes difficult?*

20 b. *Could this be helped by the closed-eye period (as here, the rain event that is too strong for the top layer, may give valuable insight in the lower ones and if the model is not updated during this period, the front reaches down)?*

Reply: For the estimation of the second layer, the state of the first layer at the border to the second layer has to be represented as well as possible. We can picture the first layer as a boundary condition for the second layer. This means, that the state estimation of the first layer will improve the parameter estimation of the second layer. Especially during the rain event this state estimation becomes very important. Without it, the infiltration front would reach the second layer at the wrong time, leading to biased parameters there.

25 The closed-eye period in the first layer does not improve the parameter estimation in the second layer, because the parameters in the first layer are worse at describing the rain event This requires stronger corrections in the state, which actually might make the parameter estimation in the second layer more difficult.

30 But, we agree, that in case of a proper representation of the preferential flow a better estimation of the second layer might be possible with a longer memory.

5. Validation

35 *Using data that has been used 10-20 times to calibrate a model with as a validation set is not a particularly strong case. More attention could be given to scenario D, where it is nicely shown that the CE-model has a much better performance than the Standard one. Even better would it be if one of the three observations in the top layer could be taken out of the calibration and used for validating the model; this would be a strong case.*

40 **Reply:** Yes, we agree, taking one sensor out of the calibration and using it for validation would be a strong case. Unfortunately this is impossible in the given scenario. Due to the detected small scale heterogeneity, this heterogeneity has to be estimated at each sensor location. Without it, it is impossible to predict the water content at this location. This is a fundamental challenge in a heterogeneous medium and the only way to verify the results is the presented internal consistency.

6. All parameters needed?

45 *The bottom $\frac{3}{4}$ of the model is only briefly discussed in the manuscript. That nothing happens in layer 4 is not so surprising, but how does it look for layers 2 and 3? Do the infiltration fronts reach down here and how is it with the parameter estimation? If nothing happens, do we need to estimate them? How is the effect of the closed-eye period on the second layer?*

Reply: We have updated Fig. 4 and included the layers 2 and 3. In both layers there is basically no dynamics as well. We agree, that the estimation of these parameters is operationally not required in this case, but we wanted to address the challenge to estimate parameters there. We could see, that we can already estimate heterogeneity and reduce uncertainty in soil parameters.

Additionally, we gained information about the robustness of the procedure. The estimated parameters did not reach values, that would be excluded by expert knowledge. Without the representation of heterogeneity the hydraulic conductivity in those layers dropped rapidly to very small values, which could be excluded.

5 **7. Overview**

The authors could consider including a nice block diagram to make their approach clearer (which time period is used for what and where is it iterated and what is feeded to where?). Two example of unclearness: 1) what is used from 2nd iteration onwards as initial condition for period C; the same output from B despite changed parameters? 2) for the top boundary, is it he finally updated value for each model that is also reused in the next iteration?

10 **Reply:** We followed your advice and included a block diagram. To the specific questions: 1) the same initial condition generated once is used for all iterations. 2) the updated mean value and standard deviation for each time interval are used to generate the boundary condition ensemble for the next iteration. Both is mentioned explicitly in the text now.

8. How to select a closed-eye period?

15 *One of the key findings of the manuscript it the improvement using a closed eye period. What is not so clear is how this is to be selected. How can we differentiate between a wanted changes of parameters away from a false prior, from the erroneous updates that the authors show in this work? This feels very ad hoc, and hence also leaves the full paper feeling very problem specific. I think the manuscript would give a more rigorous feeling if this issue was discussed in more depth.*

20 **Reply:** The idea of the closed-eye period is not ad hoc and actually generally applicable. In our case we can detected the closed eye-period due to the iterative approach and thus we can determine, based on the assumption of constant parameters, the closed-eye period due to the change in the direction of the parameter update.

We enhanced the discussion accordingly (page 15, line 14-20): “Due to the iterative approach, we could detect the times when the local-equilibrium assumption is violated: the variations of the parameters are larger than the change from initial to final value during later iterations. The changes in the direction of the parameter update then determine the closed-eye period.

25 Generally, the closed-eye period can be detected, if the operational limits of the model are known. In our case, we base this on the changing parameters, but e.g. a direct detection in the state or forcing could be possible as well.

The Closed-Eye EnKF omits the incorporation of the model structural errors in the parameters and is a generally applicable concept. In this study, it yields better predictions during periods when the underlying assumptions are fulfilled: the drying period after a rain event when there is local equilibrium, showing the strength of the Richards equation there.”

30

Minor concerns

9. Longer development plots of parameters

The parameter plots only show the last iteration, however, I would find it highly interesting to know how the development throughout the iterations also looks. Example: mean alpha is initially sampled at 4.8 and has in Figure 6 a value for the standard filter of maybe 4.7 with a jump of 0.05. Hence, it cannot have looked the same during the other 9 iteration, then the value would be different. Similarly, the closed eye filter has a stable value of around 5, but how did the way there from 4.7 look?

35 **Reply:** We have added the evolution of the parameters during all iterations. Especially during the first iteration parameters can overshoot to compensate other parameters, leading to the observed behaviour. Additionally, the mean value is not perfectly conserved from one iteration to the next, because the ensemble is not kept, but sampled new from a Gaussian distribution with
40 mean and standard deviation of the final parameters of the previous iteration.

10. Performance of org. model

Please include in the water content figures, also the performance of the mean of the original model, so that the reader has the possibility to assess the positive development of the model during the filtering.

45 **Reply:** We included the original model in Fig. 4. It assumes homogeneous soil layers and cannot describe the observed water contents.

11. Why this damping?

As the selection of a damping parameter is anything but obvious, I think it is useful if authors using it gives a one sentence

motivation to why it was chosen this way!

Reply: We based our choice on values seen in the literature. We enhanced the text on page 8, line 20-21: “Hendricks Franssen and Kinzelbach (2008) investigated damping factors between 0.1 and 1.0, while Erdal et al. (2014) employed a damping factor of 0.3 for the parameters. The smaller the value, the better can a non-linearity be handled, but the slower will the estimated value approach the optimal value. We choose the values $d = [1.0, 0.1, 0.1, 0.1]$. This means, that based on the water content measurements, there is no damping for the update of the water content state, but a strong damping for soil hydraulic parameters, Miller scaling factors and upper boundary condition.”

12. Resolution of model

10 *The resolution is uniformed 1cm (P5,L20), which for a model with strong boundary fluxes may not be so small. Has the authors checked that the grid size is not also causing issues? Why is “the effect” minimized by 1 cm and not at 1mm?*

Reply: We improved our wording to: “We choose a resolution of 1 cm. For the represented soil and boundary conditions this resolution is sufficient to resolve the occurring fronts smoothly, so that we can neglect effects caused by the numerical solver.”

15 **13. Why 100 members?**

The model has 140 unknowns and is 1D, it cannot be a particular difficult to also run a larger ensemble size and reduce the ad hoc tunings, so why is this setup with a small quite small ensemble chosen?

Reply: We did not investigate the ensemble size, as 100 members proofed to be sufficient. By increasing the ensemble size we only see the localization as a possibility to reduce the applied improvements. We do not see the localization as an ad hoc tuning, though. A limited information range is physically based.

The damping factor is included to attenuate the effect of linearizing non-linear relations, the covariance inflation is added to prevent filter inbreeding, which can be caused non-Gaussian distributions and the prior estimation of Miller scaling factors is added due to otherwise occurring wrong correlations between hydraulic conductivity and unrepresented heterogeneity.

25 **14. True parameters or heuristic model (was Wollschläger wrong?)**

a. I’m a bit confused about this discussion. On P12L24 it is stated that the estimated parameters of the closed eye filter better resembles the believed true parameters and that the standard filter ones are more/only heuristic. However, on P14,L1 it is clearly stated that the estimated parameters are all heuristic and only valid in their estimated range, which hence suggests that there are no such things as true parameters.

30 *b. Also an elaboration on the different result presented here to that of Wollschläger 2009; you get quite different result for a plot quite close by. Are the result so local or where the previously published results not so reliable?*

Reply: We agree, there are no true parameters, but there are, what we call ‘believed true material properties’, which are only valid in the observed range. We have added our definition (page 12, line 24): “This means, if the parameters are constant in time, we believe, that the parameters can represent reality in the observed water content range during times, when the underlying assumptions hold.”

Wollschläger et al. (2009) were not wrong. The used data comprised only one TDR sensor in each layer. The obtained results are based on the assumption of homogeneous layers and are optimal in the sense of an inversion when a true model is assumed. As shown in the manuscript the representation of heterogeneity is important, though.

40 **Technical stuff**

1. Figure 6: negative alpha values?

Reply: The sign of the α values is connected to the choice how the potential is defined. If bound water content states are chosen to have a negative potential, the α values have to be negative.

45 *2. Language: should be checked carefully. E.g. what does the sentence “The forcing or embedding in space is the initial condition” mean?*

Reply: We corrected the sentence and improved formulations throughout the manuscript.

3. Consider splitting the “Conclusion” section into a “Summary and discussion” (where some of the discussion points taken up in this review would fit) and a short “Conclusion” section which only contains the actual conclusions.

Reply: We kept the existing sections, but improved the Conclusion.

Answer to Review by Anonymous Referee #2

In the paper by Bauser et al., the authors perform data assimilation experiments for a 1D real-world case (Grenzhaus site, Germany) for a time period of two months. An iterative EnKF is used to correct states (soil moisture), parameters (layer-based hydraulic conductivity, Miller scaling factors, porosity, van Genuchten parameters) and forcings (upper boundary condition) with measured soil moisture profiles. The authors introduce a so called ‘closed-eye’ EnKF scheme, in which parameters are not estimated during an intensive rainfall event. This is done in order to prevent the incorporation of model structural errors into the estimated parameters.

We thank the reviewer for the constructive comments and suggestions. We have revised our manuscript taking them into account. For consistency with your references, we will still refer to the unrevised manuscript as well.

General comments:

1) The introduction would benefit from a more detailed description of previous applications of the EnKF in soil hydrology. The authors should highlight the novelties of their approach compared to previous studies. Additionally, one of the aspects of this paper is a proper description of uncertainties for the utilized test case, so relevant literature on uncertainty description in hydrologic data assimilation should also be cited (e.g. Jafarpour & Tarrahi (2011), Zhang et al. (2015)).

Reply: We have extended our introduction. We had not been aware of the works by Jafarpour and Tarrahi (2011) and Zhang et al. (2015) and cite them now.

We also added our introduced approach into the introduction (starting page 2, line 10):

“Synthetic cases offer the advantage of a direct control and knowledge about all uncertainties, which is a challenge in real world cases. The characterization of uncertainties is critical for the success of a data assimilation scheme (Liu et al., 2012). Assuming the wrong uncertainty sources can lead to poorer performance (e.g. Crow and Van Loon, 2006; Zhang et al., 2015). In case the correct uncertainty is unknown it can be better to overestimate uncertainties than to disregard or underestimate them (e.g. Jafarpour and Tarrahi, 2011; Zhang et al., 2015).

In this study, we focus on a real world case to address the challenge of consistent aggregation of the information. We exercise this aggregation with the EnKF on a small situation: a 1D soil profile equipped with time domain reflectometry (TDR) probes measuring water content, during a time period of less than 2 months. We assess all uncertainties in the representation of this particular situation qualitatively and design a three stage approach to reduce the largest uncertainties or to consider them appropriately: First, we improve the prior knowledge (initial condition and small scale heterogeneity) to facilitate the subsequent estimation. Second, we perform an assimilation with a standard EnKF approach with an augmented state to directly reduce uncertainties in states, soil hydraulic parameters, small scale heterogeneity and upper boundary condition. We introduce iterations over the complete EnKF scheme to cope with the short data set, and to determine times when the underlying local equilibrium assumptions by the Richards equation are violated. We define this specific time as the closed-eye period, because in the third stage, we only estimates states, but keep parameters constant during this time and thus prevent the incorporation of the uncertainties in the dynamics into the parameters. The estimation of soil hydraulic parameters is only performed before and after this closed-eye period.

The remainder of the manuscript is organized as follows: Section 2 describes our understanding of (i) a representation of a physical system in general, (ii) the hydrological system we want to represent, (iii) the specific representation and its uncertainties in detail, (iv) the concept of knowledge fusion and (v) the EnKF. Section 3 presents the three stages: (i) Improving the prior, (ii) Standard EnKF and (iii) Closed-Eye EnKF and compares the prediction capabilities of Standard and Closed-Eye EnKF. In section 4 we present our conclusion and their implications.”

2) More details on the iterative EnKF are needed because this method is not commonly used in hydrologic data assimilation: Are the initial conditions the same for each EnKF iteration? What is the criterion for choosing 50 or 10 EnKF iterations? How do the updated parameters evolve during the different EnKF iterations? In order to judge the EnKF performance better, it would be beneficial if the authors could also provide simulation results with the initial guess of parameters (no data assimilation) as a benchmark.

Reply: We improved the introduction of the iterative approach (page 9, line 12-15):“We iterate the whole EnKF scheme and

start the next iteration with the final estimation of soil parameters, Miller scaling factors and upper boundary condition of the previous iteration. These iterations differ from the typically applied iterative EnKFs like the Restart EnKF or Confirming EnKF (e.g. Song et al., 2014) and rather resemble the iterations in the Kalman ensemble generator by Nowak (2009), who used a modified EnKF to estimate parameters only. The full iterations applied are required to estimate constant augmented state components even with a small damping factor and a rather short time period of data. This is especially important for the upper boundary condition, where the time periods of constant values are short.”

The initial conditions are kept the same for each iteration. We added this information to the manuscript. We followed the advice of Referee #1 and include a block diagram to better illustrate the three stage approach. We added the evolution of the parameters during all iterations and show the results of the initial guess parameters together with the results of the Standard EnKF.

3) The assimilation time period is rather short and already split into different parts for the estimation of initial conditions/ Miller scaling factors (time periods A/B) and the estimation of the full state-parameter vector (time period C). In between time periods A/B and C there is an inconsistency in terms of ensemble size (40/40 versus 100), the number of EnKF iterations (50/10) and an additional perturbation of parameters. I suggest the authors to be more consistent in the data assimilation set-up by using the same meta-parameters (ensemble size/ EnKF iterations) for all time periods. Especially for the change of the ensemble size it is unclear how the states and parameters from the 40 ensemble members are resampled to 100 ensemble members (see specific comments below). This data assimilation scheme also does not allow for an equilibration of model states towards the perturbed/ updated parameters. Therefore, I would suggest to add a certain spin-up phase for the ensemble at the beginning of the assimilation phase and between the EnKF iterations. In my opinion, it would also strengthen the paper, if the comparison between standard and ‘closed-eye’ EnKF could be made for more than one single rainfall event in order to provide more evidence for the effect of model structural errors during rainfall events. This could be achieved, e.g. by applying parameter estimation with EnKF during both rain events (time period B+C) or by extending the simulation period for additional rain events. It would also be advantageous to use an independent verification period (including one or more rain events, not only the recession phase in time period D) to compare the performance of the updated parameter values from the two assimilation schemes.

Reply: The number of iterations between time period A and C was chosen differently, due to the different assimilation time (2 days and 5 days). The ensemble size was increased from time period A to C, because the number of estimated parameters was increased as well. The ensemble size in time period B was chosen as 100 as well, so that we can directly use the states as initial condition, which we now clarify in the text. The information about parameters is passed as mean value and standard deviation, which allows the change in the ensemble size.

We agree that a spin-up phase is conceptually correct. We see the state estimation time period B as an initial spin-up phase to achieve a good estimate of the initial condition. A spin-up phase without state estimation is not suitable due to the unrepresented structural model errors, which could lead to a biased initial condition.

We investigated the benefit of incorporating this initial spin-up into the iterations. The spin-up lead to stronger correlations between states and parameters in the beginning. Without the spin-up the initial correlations are small, but develop rapidly. We see this as an internal spin-up phase. This might require more iterations, but compared to the additional computational time by the spin-up phase (time period B is 3 times longer than time period C, which then quadruples the computation time), these costs are small. Because of that we decided not to include a spin-up phase into the iterations.

We agree, that the concept of the closed-eye period should be investigated on further rain events and verification periods. This will be part of our future work. In this manuscript we focus on putting forward the concept.

4) It is quite unclear from the manuscript how the upper boundary condition is updated. As a filtering method, the EnKF only adapts the state-parameter vector for the current model time step. How is the updated upper boundary condition then incorporated in the subsequent model integration? Additionally, the results for the update of the upper boundary condition are not really discussed in detail. So it is quite unclear, how this update affects the model simulations. The state vector includes almost all components of the water balance for the 1D soil profile. These components are all adapted with the same measurements and it is unclear whether the soil moisture data contain enough information to update all water balance components at once. As the main focus of the paper is on parameter estimation, I suggest the authors to leave out the estimation of the upper boundary

condition and only use perturbed forcing data instead.

Reply: The upper boundary condition is part of the augmented state and is treated like a parameter, except that the value can change in the forward propagation. An estimation of the upper boundary condition is not mandatory. A stochastic representation of the uncertainties could be sufficient. But we argue, that a bias in the upper boundary condition could induce a bias in the estimated parameters. Especially if this bias mainly occurs during the rain event, it could cause a parameter shift during the rain event, similar to the one observed during the closed-eye period. To prevent this, we do estimate the boundary condition. This requires the decreased temporal resolution, which could lead to short-term parameter variations within the rain event. We think that this downside is subordinate to the removal of the bias, especially since it is strongly alleviated by the dissipative soil water movement.

10 The information about the boundary condition is in the covariances between measurements and the boundary condition, which is highest for the topmost TDR.

We improved the description of the boundary condition in the manuscript (page 8, line 5-11): “The expansion to an augmented state changes the propagation in time. Each component needs an individual forward propagation. We assume the soil hydraulic parameters and Miller scaling factors to be constant in time. This is not possible for the upper boundary condition where the forward equation is unknown from a soils perspective. However, measurements are available to estimate the evaporation and precipitation. Hence, we assume the forward model constant until a new estimation is available. Then, we switch to the estimated boundary condition. To base the improvement of the upper boundary condition on several measurements, we reduce the temporal resolution of precipitation measurements to change daily and at transitions between precipitation and evaporation. This means, that the upper boundary condition is treated like the parameters, except that the value can change in the forward propagation. The original temporal resolution of the precipitation data (10 minutes) is not required due to the dissipative nature of the Richards equation, which smooths the infiltration front before it reaches the first TDR sensor in a depth of 8.5 cm. The estimation of the averaged boundary condition will ensure that there is no global bias on the parameter estimation during the rain event, but could lead to small short-time parameter drifts within a rain event.” and (page 11, line 34): “As we added the boundary condition into our estimation, we can exclude effects due to a bias in the precipitation to cause this parameter shift.

25 The reduced temporal resolution in the boundary condition could only cause short-term parameter changes within the rain event”

5) A discussion is missing on how a ‘closed-eye’ EnKF period can be defined in other practical applications. In the paper, a quite heuristic criterion (maximal parameter change within a defined time period) is used which relies on a prior application of the standard EnKF assimilation scheme. Can such a criterion be generalized? The EnKF is usually applied as a pure sequential method without iterations (often in the context of real-time simulations), so the question is whether the ‘closed-eye’ period can also be defined properly for such cases, or if the iterative EnKF is a prerequisite for the ‘closed-eye’ EnKF. An additional question is how such a ‘closed-eye’ period would be defined for a 3D model where precipitation is spatially distributed and the computational demand is much larger than for a 1D soil column.

Reply: In our case the determination of the closed-eye period requires the previous iterative Standard EnKF, but the close-eye period could be possibly determined directly from the current state or the forcing. Such a direct identification then would also enable the application to computationally more expensive models.

We enhanced the discussion accordingly (page 15, line 14-20): “Due to the iterative approach, we could detect the times when the local-equilibrium assumption is violated: the variations of the parameters are larger than the change from initial to final value during later iteration. The changes in the direction of the parameter update then determine the closed-eye period.

40 Generally, the closed-eye period can be detected, if the operational limits of the model are known. In our case, we base this on the changing parameters, but e.g. a direct detection in the state or forcing could be possible as well.

The Closed-Eye EnKF omits the incorporation of the model structural errors in the parameters and is a generally applicable concept. In this study, it yields better predictions during periods when the underlying assumptions are fulfilled: the drying period after a rain event when there is local equilibrium, showing the strength of the Richards equation there.”

45

6) To my knowledge, the covariance inflation method proposed in Anderson (2009) has not been applied in hydrologic data assimilation so far. It would be interesting to see, how this inflation scheme affects model results, e.g. by providing some information on the calculated inflation factors or by comparison with a standard EnKF assimilation scheme.

Reply: We show the inflation factor during the last iteration of the Standard EnKF and Closed-Eye EnKF in an additional figure now.

Specific comments:

5 **page 1 line 6:** *I think the definition of the ‘closed-eye EnKF’ should already been given here or at least in the introduction.*

Reply: We incorporated it in the introduction (see answer to general comment 1).

page 2 line 15-16: *This not fully representative for the study. You should also add your experiments with the standard/ ‘closed-eye’ EnKF here.*

10 **Reply:** We extended the introduction (see answer to general comment 1).

page 3 line 22-25: *Please clarify what you define as ‘non-equilibrium conditions’ in this context. Do you mean preferential flow?*

15 **Reply:** Preferential flow is one example for a local non-equilibrium condition. Generally, it means that the state does not follow the retention curve. This could also be the case if there is a rapid ground water lift. We added the preferential flow example in the manuscript.

20 **page 4 line 20-27:** *The definition of Miller scaling requires more details. How exactly are equations 5+6 incorporated into your model simulations? I guess you define a scaling factor for each model grid cell and use the parameters of the respective soil layer as the “reference material property”. Additionally, why does a scaling factor of 1 introduce large uncertainties? This is not obvious from equations 5+6.*

Reply: You are correct. We have improved the explanation in the manuscript to:

“A possibility to describe the texture inside soil layers in a simplified way is Miller scaling (Miller and Miller, 1956). They assume geometrical similarity and scale the material properties at each location (grid cell) according to:

25
$$K(\theta) = K^*(\theta) \cdot \xi^2, \tag{5}$$

$$h_m(\theta) = h_m^*(\theta) \cdot \frac{1}{\xi}, \tag{6}$$

30 where * denotes the reference material properties of the corresponding layer and ξ is the Miller scaling parameter at this location. This means, that we have to define a Miller scaling factor for each grid cell. The assumption of geometric similarity does not necessarily hold and leads to some uncertainties. The scaling factor themselves are a priori unknown. We assume no heterogeneity (scaling factors of 1), which is most likely wrong and introduces large uncertainties.”

35 **page 5 line 5-9:** *The TDR measurements for 111 and 136 cm in Figure 4b suggest that there is some dynamics in the lower part of the soil profile. On the contrary, the simulated soil moisture contents are almost constant. This is an indication that there might be some problems with the assigned lower boundary condition. The constant simulated soil moisture values are probably an effect of the sand layer that was introduced as layer 5 which acts as a hydraulic barrier to the upper soil profile. You could try to use the soil hydraulic parameters from layer 4 also for layer 5 to see, if the simulated soil moisture dynamics can be improved in the lower part of the soil profile. Additionally, it is not clear why you have chosen a water table depth of 4m. Is this value based on water level data from surrounding wells?*

40 **Reply:** Considering the noise in the sensor, we do not see dynamics in the TDR measurements at 111 and 136 cm depth. Additionally, the measurements are in an hourly interval, but we interpolated linearly when there was missing data leading to the linear stretches. We have added this information to the Figure.

We agree, that in case of dynamics, we should use the soil parameters of layer 4 for layer 5 as well. Without dynamics, we chose the sand instead to decouple the from the ground water table. Because of this decoupling, the position of the groundwater table doesn’t matter. We do not have measurements of the ground water table and chose the depth of 4m to be consistent with

Wollschläger et al. (2009).

page 7 line 25: *If the soil moisture measurements exceed the calibrated saturated water, why don't you use the data from the volumetric soil samples (page 5 line 24) as an initial guess?*

- 5 **Reply:** The data from volumetric samples is recorded for each TDR location individually, but we still assume a homogeneous reference material for the complete layer with only one saturated water content value. Furthermore, changing the saturated water content affects the material properties directly. We did not want to alter the material properties estimated by Wollschläger et al. (2009) strongly and thus, chose to estimate the saturated water content instead, to keep the change as small as possible. As long as we do not actually reach saturation, the value is not important.

10

page 8 line 9-11: *Does this mean that you only use daily averages of precipitation? In the data assimilation experiments rather short time periods are investigated and the simulations are compared to hourly soil moisture measurements. So it would make much more sense to also use atmospheric forcing data with a similar time resolution which should be available from the weather station (page 3 line 3-4).*

- 15 **Reply:** In this case the precipitation is averaged over roughly 5 hours during day 17 and 8 hours during day 18. The reduced temporal resolution is necessary to estimate the boundary condition (see answer to general comment 4). Additionally, for the evaporation the estimation of daily values has a smaller uncertainty than the estimation of hourly values (Foken, 2006).

page 8 line 30-31: *Which value for the cut-off radius did you use in the 5th order polynomial for soil water content?*

- 20 **Reply:** We empirically chose a very large value of 153 cm for the cut-off radius (corresponds to $2c$ in the 5th order piecewise function). We added this information in the manuscript.

page 9 line 2-4: *Why is the localization different for hydraulic conductivity and Miller scaling factors? Miller scaling factors are strongly related to hydraulic conductivities and should therefore also have a similar localization.*

- 25 **Reply:** The Miller scaling factors have a local influence at the corresponding position of the TDRs in each layer, while K_0 is defined for the complete layer. Because of that we applied a stricter localization to the scaling factors.

page 9 line 9-11: *Did you also use localization in the derivation of the inflation factors? This is necessary because otherwise uncertainties are increased in model areas that are not updated by the EnKF.*

- 30 **Reply:** Yes.

page 10 line 20-23: *How are initial Miller scaling factors perturbed?*

Reply: The uncertainty of the natural logarithm was set to 0.4. We added this information in the manuscript.

- 35 **page 11 line 7-11:** *It is unclear how the initial conditions for time period C were exactly created. How were the final (updated) states from time period B re-sampled from 40 to 100 ensemble members? Did you perturb the ensemble mean of Miller scaling factors from time period A or each of the 40 ensemble members individually? Was this perturbation constant in space or was the perturbation on the grid cell/ soil horizon level? Please also take into account that this perturbation creates inconsistencies between states and parameters which could be alleviated by a spin-up period.*

- 40 **Reply:** We used 100 ensemble members for time period B (we added this in the manuscript). Thus, we could use the exact states as initial conditions. For the Miller scaling factors we only used the mean value. The block diagram now shows these details. As we added the Miller scaling factors at measurement locations to the augmented state, we perturbed the values there. We are aware that the perturbations destroy previously acquired correlations, but decided not to use a spin-up phase (see answer to general comment 3).

45

page 11 line 32 - page 12 line 4: *It is not obvious from Figure 5 whether hydraulic conductivity reaches its initial value by the end of time period C. Please be more quantitative here: How large is the change within the 'closed-eye' period and how large is the difference between initial and final values? It would also be good to show how the update of hydraulic conductivity evolves over the different iterations of the EnKF. Is there a continuous increase of this parameter during the iterations? Additionally,*

the increase of hydraulic conductivity during the closed-eye period could also be related to the adaptive covariance inflation. The calculation of the inflation factors is influenced by the mismatch between simulations and measurements. As this mismatch increases during the infiltration event, this could also lead to higher inflation factors which increase the Kalman gain and thus the parameter update for this period. Did you check how the inflation factors evolve over this time period? Maybe it's also worthwhile to repeat these simulations without covariance inflation in order to exclude possible artefacts from the inflation scheme.

Reply: We added the temporal evolution during the 10 iterations to the Figure. The observed changes in the parameters cannot be caused by the covariance inflation. We agree, it is possible that the inflation could cause parameters to change stronger during the infiltration, but this does not explain the changing directions, on which we base the closed-eye period. The maximal inflation factor λ during the last iteration of the Standard EnKF was 1.013. The values are larger during the Closed-Eye EnKF when the mismatch is larger during the infiltration. We added a new Figure with the temporal evolution during the last iterations in the manuscript.

page 12 line 14-17: The 'closed-eye' EnKF experiments are added on top of the standard EnKF experiments with an additional parameter perturbation. This makes it quite difficult to compare these two experiments. I suggest to use exactly the same model set-up for both experiments, i.e. to repeat the 'closed-eye' EnKF experiment with the same initial conditions as the experiment with standard EnKF. Otherwise, a direct comparison of both methods is not possible. An additional question is why the 'closed-eye' EnKF and the parameter perturbation is only performed for the first soil layer.

Reply: We did not see changing parameters in the second layer. Thus, we only applied the closed-eye period on the first layer. We do not want to change the work flow in the three stages. Instead, we now compare the results from the Closed-Eye EnKF to a Standard EnKF with 20 iterations to achieve a fair comparison.

page 12 line 23-24: What are the 'believed true material properties'? Given the fact that this is a real-world experiments, where the true material properties are unknown, this is quite speculative.

Reply: We agree, there are no true material properties, but there are, what we call 'believed true material properties', which are only valid in the observed range. We have added our definition (page 12, line 24): "This means, if the parameters are constant in time, we believe, that the parameters can represent reality in the observed water content range during times, when the underlying assumptions hold."

page 13 line 13-22: Please be more quantitative here by providing performance measures such as root mean square error or Nash-Sutcliffe efficiency.

Reply: We deliberately did not include a quantitative measure here. We wanted to focus on the qualitative result, whether the model can describe the data or not. We think this result can be seen directly. Especially, since the quantitative result of these measures can be easily misinterpreted. The root mean square error is only meaningful, if the model can represent the data with statistically independent errors, which is not the case here.

page 13 line 26-30: In Figure 10b the simulated ensemble mean for each layer is always above the measured soil moisture values at the end of time period C. Why does this offset not appear in the beginning of time period D in Figure 10d? Shouldn't period D start with the simulation results from period C?

Reply: We start with the last states from the last iteration from the Standard EnKF or Closed-Eye EnKF respectively, as they represent the best possible estimation for the initial condition for time period D. The results from the free run in the time period B and C already have a bias and consequently should not be used. For time period B and C we performed an EnKF with state estimation to achieve a good initial condition. We added the information in the manuscript.

page 14 line 28-31: You should also discuss how such a 'closed-eye' period can be defined in practical applications.

Reply: (See answer to general comment 5.)

Figure 2: Why is there no evaporation flux in time period A and at the beginning of time period B? If it is a dry period, there should also be evaporation.

Reply: Due to the previous dry period of over one month, the evaporation will be far below the evaporation during wet conditions and almost 0.

Technical corrections:

- 5 **Reply:** We made the corrections.

EnKF with closed-eye period - towards a consistent aggregation of information in soil hydrology

Hannes H. Bauser^{1,2}, Stefan Jaumann^{1,2}, Daniel Berg^{1,2}, and Kurt Roth^{1,3}

¹Institute of Environmental Physics (IUP), Heidelberg University, Heidelberg

²HGS MathComp, Heidelberg University, Heidelberg

³Interdisciplinary Center for Scientific Computing (IWR), Heidelberg University, Heidelberg

Correspondence to: Hannes H. Bauser (hannes.bauser@iup.uni-heidelberg.de)

Abstract. The representation of soil water movement exposes uncertainties in all model components. We assess the key uncertainties for the specific hydraulic situation of a 1D soil profile with TDR measured water contents. Addressed uncertainties are initial condition, soil hydraulic parameters, small scale heterogeneity, upper boundary condition, and the local equilibrium assumption by the Richards equation. We employ an ensemble Kalman filter (EnKF) with an augmented state to represent and estimate all key uncertainties, except for the intermittent violation of the local equilibrium assumption. For the latter, we introduce a Closed-Eye EnKF to bridge the gap. Due to an iterative approach, the EnKF was capable to estimate soil parameters, Miller scaling factors, and upper boundary condition based on TDR measurements during a single rain event. The introduced closed-eye period ensured constant parameters suggesting that they resemble the believed true material properties. This closed-eye period improves predictions during periods when the local equilibrium assumption is met, but requires a description of the dynamics during local non-equilibrium phases to be able to predict them. Such a description remains an open challenge. Finally, for the given representation our results show the necessity to include small scale heterogeneity. A simplified representation with Miller scaling already yielded a satisfactory description.

1 Introduction

The description of soil water flow in the vadose zone with a mathematical model requires knowledge about material properties (typically characterized by soil hydraulic parameters), initial conditions, and boundary conditions. Especially the material properties are difficult to determine, since they can neither be measured directly nor transferred directly from the laboratory to the field.

Soil hydraulic parameters have been estimated inversely based on measurements of the temporal development of the hydraulic state with reviews e.g. by Hopmans et al. (2002) and Vrugt et al. (2008). Inversions are based on a perfect model assumption and incorporate all errors and uncertainties into the estimated parameters, leading to a suboptimal estimation.

In contrast, data assimilation methods are capable to combine information from measurements and models into an optimal estimate of the geophysical field of interest, but depend on the correct description of corresponding uncertainties (Reichle, 2008). Data assimilation methods include variational methods, particle filters, and ensemble Kalman filters (EnKF).

The EnKF, introduced by Evensen (1994), became popular in hydrology and was used to estimate e.g. water content states

from satellite data (e.g. Reichle et al., 2002) or local water content measurements (e.g. De Lannoy et al., 2009).

Liu and Gupta (2007) reminded that in ~~hydrologic~~hydrologic data assimilation the focus on state estimation is not sufficient and called for the incorporation of the errors in model structures and parameters.

Moradkhani et al. (2005) used a dual EnKF approach, whereas Vrugt et al. (2005) combined an EnKF and the shuffled complex evolution metropolis algorithm in an iterative way, to estimate parameters and states in a rainfall-runoff model. Chen and Zhang (2006) used an EnKF with augmented state to simultaneously estimate pressure and conductivity field for saturated flow.

Based on the Richards equation, Li and Ren (2011) first employed the simultaneous estimation of states and soil hydraulic parameters with the EnKF, which has been studied in more detail for purely synthetic cases (e.g. Wu and Margulis, 2011; Song et al., 2014; Erdal et al., 2015; Shi et al., 2015; Man et al., 2016) or with focus on synthetic cases (e.g. Wu and Margulis, 2013;

10 Erdal et al., 2014).~~The advantage of synthetic cases comprises the~~

Synthetic cases offer the advantage of a direct control and knowledge about all uncertainties, which is a challenge in real world cases. The characterization of uncertainties is critical for the success of a data assimilation scheme (Liu et al., 2012).

Assuming the wrong uncertainty sources can lead to poorer performance (e.g. Crow and Van Loon, 2006; Zhang et al., 2015).

In case the correct uncertainty is unknown it can be better to overestimate uncertainties than to disregard or underestimate them (e.g. Jafarpour and Tarrahi, 2011; Zhang et al., 2015).

15 In this study, we focus on a real world case to address the challenge of consistent aggregation of the information,~~which requires a proper characterization of uncertainties in all model components.~~ We exercise this aggregation with the EnKF on a small situation: a 1D soil profile equipped with time domain reflectometry (TDR) probes measuring water content, during a time period of less than 2 months.

20 Our goals are (i) to We assess all uncertainties in the representation of this particular situation qualitatively and (ii) design a three stage approach to reduce the largest uncertainties ~~with the EnKF~~ or to consider them appropriately.

2 Background

~~This section:~~ First, we improve the prior knowledge (initial condition and small scale heterogeneity) to facilitate the subsequent estimation. Second, we perform an assimilation with a standard EnKF approach with an augmented state to directly reduce uncertainties in states, soil hydraulic parameters, small scale heterogeneity and upper boundary condition. We introduce iterations over the complete EnKF scheme to cope with the short data set, and to determine times when the underlying local equilibrium assumptions by the Richards equation are violated. We define this specific time as the closed-eye period, because in the third stage, we only estimates states, but keep parameters constant during this time and thus prevent the incorporation of the uncertainties in the dynamics into the parameters. The estimation of soil hydraulic parameters is only performed before and after this closed-eye period.

30 The remainder of the manuscript is organized as follows: Section 2 describes our understanding of (i) a representation of a physical system in general, (ii) the hydrological system we want to represent, (iii) the specific representation ~~of this system and its uncertainties~~ in detail, (iv) the concept of knowledge fusion and (ivv) the EnKF. Section 3 presents the three stages: (i)

Improving the prior, (ii) Standard EnKF and (iii) Closed-Eye EnKF and compares the prediction capabilities of Standard and Closed-Eye EnKF. In section 4 we present our conclusion and their implications.

2 Background

2.1 ~~Respresentation~~Representation

5 In this work, we call the mathematical description of a physical system a representation, which comprises in the most general sense the four components: dynamics, forcing, subscale physics and states.

Dynamics: The dynamics specifies how the state of the system is propagated in space and time at the scale of interest.

Forcing: The forcing or coupling of the system to the superscale physics in space (initial state) and time (boundary condition).

10 **Subscale ~~pyhsies~~physics:** The subscale physics not explicitly described by the dynamics, but with a parameterization comprising one or more parameters and their spatial and temporal distribution.

State: Distribution of the variable propagated by the dynamics, at a specific time. Notice that this excludes parameters.

According to this definition a representation is specific for each system. Hence, we first introduce the system, and the details of the representation afterwards.

15 2.2 Testsite

We aim to represent the water movement in a soil profile at the Grenzhof test site close to Heidelberg, Germany. Since 2003, experiments are conducted at the test site. In 2004, a weather station was built, which measures precipitation and further atmospheric data (wind, temperature, incoming and outgoing long- and shortwave radiation, relative humidity and air temperature) in 10 minute intervals. A detailed description of the test site can be found in Wollschläger et al. (2009).

20 In 2009, a soil profile was equipped with 11 TDR probes measuring water content hourly. The soil profile itself, described explicitly by Schenk (2011), is depicted in Fig. 1. It consists of four different soil layers up to the depth of 1.4 m. The top three layers are equipped with three TDR probes each, and ~~two TDR probes in~~ the bottom layer with two TDR probes. The profile is kept free from vegetation at the surface.

25 The complete time period considered comprises 60 days from October 1, 2011 (day 1) until November 29, 2011 (day 60). The boundary condition along with the water content measured by the topmost TDR are shown in Fig. 2.

2.3 Representation of soil water flow

For this specific soil profile we can now formulate the representation consisting of the four components described in section 2.1. For each we give the a priori description, as well as the corresponding assumptions and sources of uncertainty.

Dynamics: For a continuous porous medium, the Richards equation (Richards, 1931) describes the change of volumetric water content θ

$$\frac{\partial \theta}{\partial t} - \nabla \cdot [K(\theta) [\nabla h_m(\theta) - 1]] = 0, \quad (1)$$

with the isotropic conductivity K and matric head h_m . The Richards equation itself is process-based and is expected to describe the actual physics adequately as long as its underlying assumptions are fulfilled: It assumes the soil water movement to be a single-phase process. Hence, the Richards equation is expected to fail in very dry or wet conditions. Both are not observed at the Grenzhof test site during the time period considered. The corresponding uncertainties are small. The Richards equation further assumes local equilibrium. This is assumed to be violated during strong infiltration events, e.g. [during heavy rainfall leading to preferential flow](#). We expect local non-equilibrium at the Grenzhof test site, but it is a priori not clear, how strong the forcing has to be or how to describe these cases. This creates uncertainty in the dynamics during rain events.

We additionally assume, that horizontal flow is negligible and that we can describe the system one dimensional without additional sources or sinks. The flat terrain and horizontal layering at the Grenzhof test site supports this assumption. Still, heterogeneity, that is not represented, might introduce two-dimensional flow.

Material properties: The subscale physics not explicitly represented in Eq. 1, are the dependencies of K and h_m on θ . Several parameterizations are available. We follow Wollschläger et al. (2009) and choose the Mualem-van Genuchten parameterization (Mualem, 1976; Van Genuchten, 1980):

$$K(\Theta) = K_0 \Theta^\tau \left[1 - \left[1 - \Theta^{n/[n-1]} \right]^{1-1/n} \right]^2, \quad (2)$$

$$h_m(\Theta) = \frac{1}{\alpha} \left[\Theta^{-n/[n-1]} - 1 \right]^{1/n}, \quad (3)$$

with the saturation

$$\Theta := \frac{\theta - \theta_r}{\theta_s - \theta_r}. \quad (4)$$

The parameterization reduces the description of the subscale physics to a set of six parameters: θ_s , θ_r , α , n , K_0 and τ . We do not expect significant uncertainties due to assuming this particular parameterization, because we only observe a rather small water content range.

The parameters θ_s and θ_r are the saturated and residual water content. The parameters α and n can be associated with an inverse air entry value and the width of the pore size distribution respectively, although there is no direct physical relation. K_0 is the saturated conductivity and τ is a fudge factor.

The spatial distribution of these material properties has to be defined on the scale of the dynamics. The large-scale structure for this soil profile are 4 soil layers, where we assign an individual parameter set to each layer.

Although these parameters cannot be measured directly and are often assigned with the largest uncertainties in soil hydrology, we can in this case use the soil parameters determined by Wollschläger et al. (2009) through an inversion for a soil profile close

by. They are expected to differ from the soil hydraulic parameters at the represented soil profile due to spatial heterogeneity.

We do not include a description of hysteresis, leading to additional uncertainties connected with the parameterization.

We assume the parameters to be constant in time, which appears justified, since we only investigate a short time period of less than 2 months.

- 5 The assumption of homogeneous soil layers is presumably wrong. A possibility to describe the texture inside soil layers in a simplified way is Miller scaling (Miller and Miller, 1956). They assume geometrical similarity and scale the material properties at each location (grid cell) according to:

$$K(\theta) = K^*(\theta) \cdot \xi^2, \quad (5)$$

$$h_m(\theta) = h_m^*(\theta) \cdot \frac{1}{\xi}, \quad (6)$$

- 10 where * denotes the reference material properties for each of the corresponding layer and ξ is the Miller scaling parameter -at this location. This means, that we have to define a Miller scaling factor for each grid cell. The assumption of geometric similarity does not necessarily hold and leads to some uncertainties. The scaling factor themselves are a priori unknown. ~~They are assumed to be~~ We assume no heterogeneity (scaling factors of 1, which-), which is most likely wrong and introduces large uncertainties.

15

Forcing: The forcing in time is the upper and lower boundary condition. For the upper boundary condition, we can measure rain and use the meteorological data from the on-site weather station to calculate the evaporative water flux using the reference FAO Penman-Monteith equation following Allen et al. (2006). Here, especially the estimation of the evaporation is highly uncertain. The daily mean value can differ 10-20% (Foken, 2006). In dry conditions the evaporation is limited by the soil, which leads to additional uncertainties.

20

At the Grenzhof test site the rain gauge had not been calibrated during the period of consideration. Because of that we have to assume an uncertainty of about 20% for the precipitation as well.

For the lower boundary condition there are no measurements at the site. However, constant water content measurements in the lowest soil layer during the observed time period indicate, that the dynamics is decoupled from the groundwater table.

- 25 Therefore, we follow Wollschläger et al. (2009) and add a 5th layer with soil parameters for sand from Carsel and Parrish (1988) up to a depth of 4 m in our representation. The sand layer decouples the dynamics from the groundwater table, that we keep constant at the depth of 4 m. This leads to uncertainties, but they are expected to be negligible.

The forcing ~~or embedding~~ in space is the initial condition. The initial condition is difficult to estimate. A simple approach is to use the measured water contents and interpolate them linearly with constant extrapolation to layer boundaries. This leads to

- 30 large uncertainties in between.

States: With the three components dynamics, material properties and forcing we can predict the temporal development of the states, in this case the water content. These states will have large uncertainties due to the uncertainties in the other compo-

nents.

An additional source of uncertainty is the simulation of the dynamics with a numerical solver, here with MuPhi (Ippisch et al., 2006). The assumption of local equilibrium in the Richards equation has to be granted on the grid size of the numerical model as well. ~~Since we have a very small test case, we~~ We choose a resolution of 1 cm ~~to minimize the effect.~~ For the represented soil and boundary conditions this resolution is sufficient to resolve the occurring fronts smoothly, so that we can neglect effects caused by the numerical solver.

Water content determined with TDR probes will be used to improve the representation. The water content values are calculated from the temperature corrected electric permittivity with the complex refraction index model (CRIM) following Roth et al. (1990). The porosities required there, were obtained from volumetric soil samples.

We assume an unbiased uncertainty of 0.01 for the water contents. This is the largest noise observed with the TDR probes. The assumptions made for the representation of the soil water movement at the Grenzhof test site are summarized in Table 1. Along we give our qualitative assessment of the uncertainties: small, no need to represent the uncertainty explicitly; intermediate, it might be necessary to represent these uncertainties, we decide not to but keep them in mind; large, uncertainties must be represented with the goal to reduce them. Our findings depend on this uncertainty assessment and might be affected if ~~is-it~~ does not hold.

2.4 Knowledge Fusion

We define knowledge fusion as the consistent aggregation of all information pertinent to some observed reality. In the presented situation at the Grenzhof test site this would require the quantitatively correct description of all uncertainties in the representation and a subsequent optimal reduction of all these uncertainties based on all additionally available information. For the measurement part these are primarily the water content data from the 11 TDR probes, but any other information, even expert knowledge, should be incorporated as well. So far, this goal is only partly feasible, though.

With an inversion all information and uncertainties would be included into the parameters, which is not a consistent aggregation of the information, since the structural errors in the other components are not represented.

Data assimilation methods are capable to represent all uncertainties and have been expanded to not only estimate states, but also parameters. For an aggregation of all information we have to reduce uncertainties in all representation components. In this study we aim to describe and reduce the uncertainties classified as large in Table 1 and decided to neglect the ones classified as small or intermediate with the EnKF.

30 2.5 Ensemble Kalman Filter

The EnKF is a data assimilation method that uses a Monte Carlo approach for an optimal state estimation, based on the assumption of unbiased Gaussian error distributions. The EnKF incorporates the measurements sequentially by alternating between a forecast step (superscript f), which propagates the state in time, and an analysis step (superscript a), which incorporates the

information from the measurements at this time to improve the state.

These two steps are now explained in more detail - specifically for the given representation. A general description can be found e.g in Burgers et al. (1998).

Forecast: First an ensemble of water content states θ^i is propagated forward in time from k-1 to k:

$$5 \quad \theta_k^{f,i} = M(\theta_{k-1}^{a,i}, \phi_k^i, \xi_k^i, \mu_k^i) + \beta_k^i \quad (7)$$

where M denotes the propagation with the Richards equation and ϕ^i , ξ^i and μ^i are soil hydraulic parameters, Miller scaling parameters, and the upper boundary condition of ensemble member i , respectively. In this way, all the uncertainties from θ , ϕ , ξ and μ are explicitly represented and propagated forward in time with the non-linear dynamics leading to the forecast ensemble θ^f , which represents these uncertainties. Uncertainties not explicitly represented are incorporated into the process noise β , which is difficult to describe, however. We set the process noise to 0. In this way, we represent all the uncertainties classified as ‘large’ in Table 1 directly, except for uncertainties caused by the local equilibrium assumption, which is expected to only hold outside of strong rain events. We decide to not represent ‘intermediate’ and ‘small’ uncertainties.

The uncertainty of the forecast ensemble ~~can be is~~ characterized based on the Gaussian assumption with the state error covariance matrix \mathbf{P} :

$$15 \quad \mathbf{P}_k^f = \frac{1}{N-1} \sum_{i=1}^N \left(\theta_k^{f,i} - \bar{\theta}_k^f \right) \left(\theta_k^{f,i} - \bar{\theta}_k^f \right)^T, \quad (8)$$

with the number of ensemble members N and the ensemble mean $\bar{\theta}$.

Analysis: To incorporate measurements, we have to quantify their uncertainties by defining the measurement error covariance matrix \mathbf{R} :

$$\mathbf{R}_k = \frac{1}{N-1} \sum_{i=1}^N \left(\epsilon_k^i \right) \left(\epsilon_k^i \right)^T, \quad (9)$$

20 with the measurement errors ϵ . We assume the measurement errors to be uncorrelated and drawn from a normal distribution with the assumed TDR standard deviation of 0.01.

The Kalman gain \mathbf{K} weights the forecast state error covariance and the measurement error covariance

$$\mathbf{K}_k = \mathbf{P}_k^f \mathbf{H}_k^T \left[\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k \right]^{-1}, \quad (10)$$

25 with the measurement operator \mathbf{H} mapping from the state space to the measurement space, and the Kalman gain \mathbf{K} mapping from the measurement space back to the state space - weighted with the uncertainties of measurements and states. Based on the Kalman gain the weighted difference of forecast and measurements is added to the forecast states to obtain the analysis states

$$\theta_k^{a,i} = \theta_k^{f,i} + \mathbf{K}_k \left[z_k - \mathbf{H}_k \theta_k^{f,i} + \epsilon_k^i \right], \quad (11)$$

where z are the measurements. The measurement errors ϵ have to be added to the difference so that the ensemble describes the analysis error covariance correctly (Burgers et al., 1998). The descriptive explanation is that the measurements z already

include measurement errors, while the measurements of the forecast state $\mathbf{H}\theta^f$ do not. Hence, the error has to be added to compare equivalents.

By alternating between forecast and analysis, the information of all measurements is incorporated to achieve an improved estimation of the states at each time step.

5

By replacing the water content state θ with an augmented state ψ , not only the water content states, but all components of the augmented state can be estimated with the EnKF. We choose to include all the directly represented uncertain components in the augmented state $\psi = [\theta, \phi, \xi, \mu]$ consisting of water content, soil hydraulic parameters, Miller scaling parameters and the upper boundary condition.

10 Adding components to the augmented state increases its dimension, which in turn requires a larger number of ensemble members, leading to a higher computational effort. To minimize this effect, we keep the added components as small as possible.

As parameters to be incorporated, we choose α , n and K_0 for each of the four layers as well as the saturated water content θ_s for layers 3 and 4. For the upper two layers, the measured water contents were far away from θ_s and θ_r estimated by Wollschläger et al. (2009). Thus, they are irrelevant for the estimation, as changes in θ_s and θ_r can be compensated by the other parameters.

15 For the bottom two layers measurements exceeded θ_s . Thus, it has to be added in the augmented state. We do not include the fudge factor τ , since the observed water content ranges are rather small. For other situations, incorporation of τ might improve the results, however.

Since the EnKF assumes Gaussian distributions and linear correlations, we do not directly use K_0 but the logarithm (to which the water content response behaves more linearly).

20 The dimensions of the Miller scaling factors are reduced as well. Only the scaling factors at measurement locations are added. The whole Miller scaling field is determined by linear interpolation between measurement locations inside a layer and a constant extrapolation towards layer boundaries. As the measurements yield only little information about the small-scale architecture away from their location, we expect that this assumption has only little influence on the results. Again, the logarithm of the parameter is used in the augmented state.

25 The upper boundary condition is already a scalar and is added as the flux at the surface to the augmented state.

In this way, soil parameters, Miller scaling factors and the upper boundary condition are updated besides the states in the analysis.

The expansion to an augmented state changes the propagation in time. Each component needs an individual forward propagation. We assume the soil hydraulic parameters and Miller scaling factors to be constant in time. This is not possible for the upper boundary condition where the forward equation is unknown from a soils perspective. However, measurements are available to estimate the evaporation and precipitation. Hence, we assume the forward model constant until a new estimation is available. ~~To be able to improve the boundary condition with~~ Then, we switch to the estimated boundary condition. To base

30 the improvement of the upper boundary condition on several measurements, we reduce the temporal resolution of precipitation measurements to ~~only~~ change daily and at transitions between precipitation and evaporation. This ~~is possible~~ means, that the
35 upper boundary condition is treated like the parameters, except that the value can change in the forward propagation. The

original temporal resolution of the precipitation data (10 minutes) is not required due to the dissipative nature of the Richards equation-, which smooths the infiltration front before it reaches the first TDR sensor in a depth of 8.5 cm. The estimation of the averaged boundary condition will ensure that there is no global bias on the parameter estimation during the rain event, but could lead to small short-time parameter drifts within a rain event.

5 The expanded forward propagation of the augmented state is:

$$\boldsymbol{\theta}_k^f = M_\theta(\boldsymbol{\theta}_{k-1}^a, \phi_{k-1}^a, \boldsymbol{\xi}_{k-1}^a, \mu_{k-1}^a), \quad (12)$$

$$\phi_k^f = \phi_{k-1}^a, \quad (13)$$

$$\boldsymbol{\xi}_k^f = \boldsymbol{\xi}_{k-1}^a, \quad (14)$$

$$\mu_k^f = \begin{cases} \mu_{\text{meas}} & \text{if a measurement is available,} \\ \mu_{k-1}^a & \text{else.} \end{cases} \quad (15)$$

10 The analysis information for the additional augmented state components lies in the covariance between states at measurement locations and the components. The covariance linearizes this relation that can be highly non-linear.

To reduce the impact of the linearization of non-linear dependencies, we employ damping factors (Hendricks Franssen and Kinzelbach, 2008), different for each augmented state component (Wu and Margulis, 2011) $\mathbf{d} = [\mathbf{d}_\theta, \mathbf{d}_\phi, \mathbf{d}_\xi, \mathbf{d}_\mu]$. The correction vector in Eq. 11 is multiplied by the corresponding factors. ~~We choose the values $\mathbf{d} = [1.0, 0.1, 0.1, 0.1]$~~ Hendricks Franssen and Kinzelbach

15 damping factors between 0.1 and 1.0, while Erdal et al. (2014) employed a damping factor of 0.3 for the parameters. The smaller the value, the better can a non-linearity be handled, but the slower will the estimated value approach the optimal value. We choose the values $\mathbf{d} = [1.0, 0.1, 0.1, 0.1]$. This means, that based on the water content measurements, there is no damping for the update of the water content state, but a strong damping for soil hydraulic parameters, Miller scaling factors and upper boundary condition.

20

Due to a limited ensemble size, the EnKF will show spurious covariances between uncorrelated state components. These can be reduced by increasing the ensemble size, or - more computationally efficient - by introducing a localization (Houtekamer and Mitchell, 2001), where external information about uncorrelated states is introduced. This can be achieved by multiplying the error covariance matrix by the correlation matrix ρ (with entries from 0 to 1) with ~~the Schur-an entrywise~~ product. This

25 alters the calculation of the Kalman gain to

$$\mathbf{K}_t = \left(\rho \circ \mathbf{P}_t^f \right) \mathbf{H}_t^T \left[\mathbf{H}_t \left(\rho \circ \mathbf{P}_t^f \right) \mathbf{H}_t^T + \mathbf{R}_t \right]^{-1}, \quad (16)$$

where the ~~Schur-entrywise~~ product is indicated by \circ .

For the water content state, the covariances are reduced with increasing physical distance. We use the 5th order piecewise rational function defined by Gaspari and Cohn (1999), which is similar to a Gaussian but with local support, which introduces a

30 cut off length. We chose a very large empirical value of 153 cm (which corresponds to $2c$ in the 5th order piecewise function).

For the soil parameters, we can localize even stronger by only allowing covariances (entries of 1) between parameters and

measurement locations in the respective soil layer and the first measurement locations in the neighboring layers. For the Miller scaling factors, only covariances to the corresponding measurement locations are used. All other entries are set to 0.

Spurious correlations and non-Gaussian distributions can lead to filter inbreeding and ultimately filter divergence (e.g. Houtekamer and Mitchell, 1998; Hamill et al., 2001). To prevent this, the ensemble can be inflated by a multiplicative factor λ (Anderson and Anderson, 1999):

$$\psi^i = \sqrt{\lambda} (\psi^i - \overline{\psi^i}) + \overline{\psi^i}. \quad (17)$$

We choose the inflation as proposed by Anderson (2009), where the inflation factor λ is adapted temporally and spatially for the states. We expand this for the other augmented state components by simply summing over the inflation factors at the measurement locations weighted with the correlations.

10 ~~Additionally, we~~ We iterate the whole EnKF scheme ~~. As we assume constant forward models for the soil parameters and and~~ start the next iteration with the final estimation of soil parameters, Miller scaling factors ~~, we can use the final distributions after~~ one iteration as the initial for the next. This enables us and upper boundary condition of the previous iteration. These iterations ~~differ from the typically applied iterative EnKFs like the Restart EnKF or Confirming EnKF (e.g. Song et al., 2014) and rather~~ resemble the iterations in the Kalman ensemble generator by Nowak (2009), who used a modified EnKF to estimate parame-
15 ters only. The full iterations applied are required to estimate constant augmented state components even with a small damping factor and a rather short time period of data. This is especially important for the upper boundary condition, where the time periods of constant values are short.

The operational assumptions of the method lead to a sub-optimal estimation of the state in each time step. Furthermore, due
20 to the non-linear dynamics, the assumption of Gaussian distributions does not hold. It is not clear how this ~~effects~~ affects the EnKF performance in detail. The Gaussian assumption leads to a linearized state update in the analysis step. This induces erroneous updates of those state components with dominant non-linear relation between states and measurements. These errors are alleviated by employing the damping factor, which reduces the update but as a consequence also reduces the incorporation of measurement information. Furthermore, we use the largest observed measurement noise to characterize the TDR uncertainties.

25 This leads to possibly too large measurement uncertainties, which have a similar effect as an additional damping factor.

We do not expect a strong influence from this sub-optimal state estimation on the mean value of the results for the soil hydraulic parameters, Miller scaling factors, and upper boundary condition. However, the final value will be approached slower than in an ideal estimation. This effect ~~can be easily circumvented by~~ is handled by the iterative approach.

On the downside, incorporating the same measurement information several times will lead, together with the other limitations,
30 to incorrect quantitative uncertainties. However, the spatially and temporally adaptive covariance inflation by Anderson (2009) ensures that measurements and states agree with each other within their uncertainties. This is only a partial solution, however, as the information is only propagated through the correlations to the measurement locations. Therefore, we will not interpret the uncertainties of the ensemble in a quantitative way. Nevertheless, we do expect them to hold qualitative information that

can be interpreted.

3 Results

The complete time period considered comprises 60 days and ranges from October 1 (day 1) until November 29 (day 60). The boundary condition along with the water content measured by the topmost TDR are shown in Fig. 2. The whole time period is separated into 4 sections: A (day 1-2), B (day 3- 17), C (18-22) and D (23-60).

Since soil parameters can only be estimated within the observed water content range and will not be valid outside of this range, a rather large rain event is desirable. On the other hand we do not represent uncertainties associated with the assumption of local equilibrium by the Richards equation, which is violated during strong rain events. Time period C (October 18-22, 2011) combines both: a rather large total rain amount (18.2 mm) and a small maximal intensity (0.7 mm in 10 minutes).

We designed a three stage approach (Fig. 3) to improve the representation by incorporating the information from the water content measurements of the 11 TDR probes: (1) improving the prior information during time periods A and B, (2) a standard EnKF approach, and (3) a ~~closed-eye-closed-eye~~ EnKF approach, both during time period C with the considered rain event. The three stages are discussed in the following, with results analyzed through forecasts in time periods B, C and D.

3.1 Stage 1: Improving the prior

~~In~~ Highly uncertain properties can exacerbate the performance of the EnKF. For example Huber et al. (2011) showed that the state estimation with an EnKF is superior in case of properly precalibrated model parameters. In our case, there is no prior knowledge about small scale heterogeneity, which can lead to difficulties in the estimation process: In a uniform layer, the water content is typically higher in deeper locations. Heterogeneity can reverse this and can lead to higher water contents above lower water contents inside a layer. Until the infiltration front of the rain event reaches these positions, the EnKF cannot distinguish between heterogeneity and an infiltration front propagating at very low speed. This can easily lead to too small saturated conductivities especially in lower soil layers that are not reached by the infiltration front. Even upper layers can require a tuned relation of the prior uncertainties of Miller scaling factors and saturated conductivity.

To avoid too small saturated conductivities, we improve the prior of the heterogeneity using the measurements from time interval A. Prior to A, there had been one month without rainfall leading to little dynamics during A. We start with the soil parameters estimated by Wollschläger et al. (2009) and linearly interpolated initial condition (with constant extrapolation to layer boundaries). Parameters and boundary condition are kept constant, but states and Miller scaling factors are updated with an EnKF of 40 ensemble members, where the initial uncertainty of the natural logarithm of the scaling factors was chosen to be 0.4.

As time interval A is only a 2 days period, the filter won't be able to reach constant Miller scaling values during this short time. Therefore, we iterate 50 times over this period, which then does lead to constant scaling factors. The resulting scaling field of the improved prior ~~is depicted in~~ (Fig. 4-~~It~~) shows that there is heterogeneity within each layer that requires a representation.

Especially layer 4 exhibits a strong heterogeneity, which is already observable in Fig. 1.

Additionally, a good initial state can improve the estimation. Therefore, we guide the state with ~~the~~ an EnKF (only state estimation) with 100 ensemble members through time period B and achieve a better representation for the initial state for time
5 period C, than interpolating between the measurements there.

3.2 Stage 2: Standard EnKF

~~Here~~ During time period C, we improve the representation with the following uncertain components: improved initial condition, soil hydraulic parameters estimated by Wollschläger et al. (2009), improved Miller scaling factors, and upper boundary condition from precipitation measurements and estimated with the reference FAO Penman-Monteith equation. We ~~iterate 10 times~~
10 ~~over time period C and~~ improve soil parameters, Miller scaling factors and the upper boundary condition along with the states. The estimated mean values and standard deviations of soil parameters, Miller scaling factors and upper boundary condition are used to generate the ensemble for the next iterations. The initial condition is kept the same. We iterate 10 times over the time period C. This estimation is performed with an ensemble size of 100 members and will be referred to as the Standard EnKF. This approach is comparable to the SODA approach by Vrugt et al. (2005) but is computationally more efficient.

15 The initial ensembles of initial condition and Miller scaling are determined as described in stage 1. For the Miller scaling factors the uncertainty is increased again compared to the estimations in the prior, since they were estimated under the assumption of fixed soil parameters. Now, they have to be able to adapt to changing parameters. The uncertainty of the natural logarithm of the scaling factors is chosen to be 0.1. For the boundary condition shown in Fig. 2, an uncertainty of 20% is assumed. The initial soil hydraulic parameters and their uncertainties are summarized in Table 2.

20 In order to check the improved results, we do not show the states from the last iteration but actually run another iteration, however, this time, without incorporating the measurement information. This corresponds to an ensemble of forward runs during time period C. It is a much more strict test for the quality and objectivity of the assimilation, because the states now cannot be adjusted. Actually, allowing incorporation of measurements leads to a much better agreement. The results for ~~the top and bottom layer~~ all four layers are shown in Fig. 5. The graphs show that the measurements can be represented well by the
25 ensemble, but deviations occur, mainly during the infiltration.

~~Already from the TDR water contents~~ There is heterogeneity inside the layers. For example, in the first layer, ~~it becomes clear that there is heterogeneity inside the layer.~~ The ~~the~~ middle TDR shows the highest water content. This effect is ~~even stronger~~ strongest in layer 4. Here, the water content of the TDR about 25 cm above the next has a water content almost 0.1 larger. The results calculated with soil parameters by Wollschläger et al. (2009) with the assumption of homogeneous soil layers are not
30 able to describe the measured water content.

The results for the estimated Miller scaling factors are shown in Fig. 4. The Miller scaling factors were adjusted along with the hydraulic parameters during this Standard EnKF approach. This is necessary, because changing parameters require adjusted Miller scaling factors as well. The largest changes occur in layer 1, where information is introduced by the water content changes during the infiltration.

The estimated soil parameters including their uncertainties are summarized in Table 3. Although the uncertainties are quantitatively incorrect, they still indicate how much information about each parameter is in the measurements. This can be seen especially well for K_0 . The estimation of this parameters requires dynamics. Therefore, the smallest uncertainty is observed in layer 1 and increases with the layers. In layer 4 the uncertainty of $\log_{10}(K_0)$ was reduced least, but still from 1.0 to 0.4.

5 This shows the difficulties of the EnKF. As we did not observe a change in water content, it is unlikely that any flux was introduced during the rain event. Nevertheless, the uncertainty was reduced, which can lead to difficulties when there actually is information.

The development of the parameters K_0 , α and n of the first layer ~~during the last iteration~~ are shown in the Fig. ~~??-??~~. All 6. Towards later iterations, all three show the same feature: The value increases during the rain event and decreases before and after, but does not change significantly between initial and final values. ~~We interpret this~~ As we added the boundary condition into our estimation, we can exclude effects due to a bias in the precipitation to cause this parameter shift. The reduced temporal resolution in the boundary condition could only cause short-term parameter changes within the rain event.

We interpret the parameter shift during the rain event as follows: The assumption of local equilibrium during the rain event is wrong, which leads to preferential flow. The infiltration is thus too fast for the actual parameters (Fig. 5). The preferential flow 15 leads to an apparent increase in the saturated hydraulic conductivity during the rain event, which is reduced again afterwards. This also shows that our initial assumption of zero process noise is wrong. The uncertainty introduced by the dynamics is not represented.

3.3 Stage 3: Closed-Eye EnKF

We enhance the EnKF with a closed-eye period during times when the local equilibrium assumption of the Richards equation 20 does not hold. During this time, soil hydraulic parameters and Miller scaling factors are not adjusted but kept constant. In contrast, the water content state is continuously updated. In this way, the state is guided on the basis of measurements through times with uncertain dynamics without incorporating the dynamics uncertainties into the parameter estimation. The estimation of soil hydraulic parameters is only performed before and after this closed-eye period.

We use an identified non-constant parameter to define the closed-eye period. We choose K_0 , because it shows the changes most 25 prominently and we can directly relate it to observed physics. The closed-eye period is defined as the time period between the minimum and the maximum value of K_0 during the last iteration with the Standard EnKF (Fig. 7).

To perform the Closed-Eye EnKF estimation, we add another 10 iterations to the previous Standard EnKF iterations, but with the closed-eye period for the first layer. We continue with the previously estimated mean values, but increase the uncertainty for the first layer again (to half the initial uncertainty) to allow that the new values can possibly deviate from the Standard 30 EnKF results.

The additional iterations result in new soil parameters and Miller scaling factors. The scaling factors changed only marginally compared to their regular estimates (Fig. 4). Changes in the parameters of the first layer are larger (Fig. ~~??-??~~6). We recognize two things: First, the values differ from each other. For n and α the changes are rather small, but for K_0 the difference is significant. The closed eye estimation yields approximately half the hydraulic conductivity as before. Second, the parameters

now are constant over time and do not show fluctuations any more.

This indicates that we could extract times when the Richards equation is actually valid and we are able to determine soil hydraulic parameters that resemble the believed true material properties. This means, if the parameters are constant in time, we believe, that the parameters can represent reality in the observed water content range during times, when the underlying

5 assumptions hold. There is also an apparent downside, though. We cannot use the measurements during the closed-eye period to estimate the parameters. This leads to a smaller observed water content range, limiting the parameter estimation possibilities. We call this ‘apparent’, because the corresponding interval does not contain valid information about the modeled system in the first place.

The forecast during time period C (analogous to the Standard EnKF) is shown in Fig. 8. The performance is acceptable, but 10 worse than the Standard EnKF (Fig. 5). This is exactly what is expected. In the Standard EnKF we incorporated errors in the dynamics into the parameters e.g. by estimating a too large hydraulic conductivity, which can represent preferential flow up to a point. By introducing the closed-eye period, we do not incorporate these errors into the parameters any more. The parameters are thus believed to be closer to the believed true material properties. Obviously, however, these parameters fail to describe rain events that lead to hydraulic non-equilibrium.

15 Guiding the states through the closed-eye period is a challenge. A representation of the dynamics’ uncertainty would be required to estimate optimal states. We did not do this. The adaptive covariance inflation reduces the issue by increasing the ensemble spread when measurements cannot be explained by combined measurement and ensemble uncertainty ~~=(Fig. 9)~~. Still, it cannot solve the problem completely, since it was not designed to follow rapid changes of the inflation factor ~~and it~~ and it is merely based on the correlations in the ensemble. These correlations are not necessarily valid any more, when a major, spatially 20 structured uncertainty is not represented. In our case this did not lead to complications, because the deviations are still small and could be adjusted towards the end of the closed-eye period. In cases with stronger rain (and stronger non-equilibrium) and larger deviations of representation and measurements, we expect that a description of the dynamics uncertainties becomes mandatory.

3.4 Prediction capabilities

25 We investigate the predictive capabilities of the Standard EnKF and Closed-Eye EnKF during the combined time period of B and C as well as time period D. As the Closed-Eye EnKF added another 10 iterations for the estimation, we compare it to Standard EnKF results with a total of 20 iterations as well.

During time period B the total amount of rain was 23.2 mm with a maximal flux of 2.5 mm in 10 minutes (compared to a total of 18.2 mm with a maximal flux of 0.7 mm in 10 minutes for time period C). As we already assume local non-equilibrium 30 during the rain event in time period C, we expect an even stronger effect during time period B.

The forecast results for time period B and C for soil parameters and Miller scaling parameters from the Standard EnKF and Closed-Eye EnKF are depicted in Fig. 10a and 10b (there is no improved estimation of the upper boundary condition during time period B and we did not use the improved boundary condition from time period C). We performed a spin-up with an EnKF with state estimation during time period A to achieve a good initial condition. Neither of them is able to predict the water con-

tents during the rain event in time period B in detail. However, the Standard EnKF describes the rain event in time period C better. The Closed-Eye EnKF is not able to predict the water contents during rain events, as for both local non-equilibrium is expected, which is not incorporated into the parameter estimation. For the Standard EnKF only the rain event from time period C is incorporated into the parameters and it predicts this rain event better. However, it fails to predict the stronger rain event in time period B showing that the local non-equilibrium in different rain events affects the dynamics differently. Hence, their effects should not be included into parameters, but rather considered separately. The uncertainties are large for both EnKFs. The reason is the large uncertainty in the boundary conditions.

We emphasize that the Standard EnKF excels if the goal is to guide some partly flawed representation with a stream of data. However, if in addition it is the goal to estimate objective parameters for the representation within its range of applicability, then the Closed-Eye EnKF must be used, at the cost of a lower heuristic performance.

The forecast during the dry time period D is depicted in Fig. 10c and 10d. [Here, we used the final state of the last iteration in the Standard EnKF and Closed-Eye EnKF estimation as the initial condition.](#) In the observed water content range the Closed-Eye EnKF performs better than the Standard EnKF. This is the result that we expected. We were able to estimate the believed true material properties in the observed water content range and can make predictions in this range as long as the basic assumptions of the Richards equation are met.

Both EnKF can describe the topmost TDR better than the lower TDRs. This is expected as well, since the first TDR measures the strongest dynamics and hence dominates the parameter estimation. The small deviations in the other TDR sensors show, that the assumption of Miller similarity, including the linear interpolation in between is capable to improve the representation. It cannot represent heterogeneity completely, however.

Below the observed soil water content we do not expect predictive capabilities, since the material properties are heuristic and hence cannot be applied outside of the calibrated range. Nevertheless we can see, that the Closed-Eye EnKF does describe the topmost TDR acceptably. Again, we see deviations in the two lower TDR sensors. This time even stronger. Again, we attribute this to the limited description of the heterogeneity with Miller scaling.

As a last comment, we did investigate the possibility of improving the results by estimating a multiplicative factor to the uncertain evaporation in analogy to Wollschläger et al. (2009). This factor can improve the estimation of the topmost TDR towards later times, but does not change the qualitative results above.

4 Conclusions

In this study we improved the representation of soil water movement in a 1D soil profile at the Grenzhof test site close to Heidelberg, Germany with an EnKF based on water content measurements by TDR probes.

We assess key uncertainties of this specific representation. These are initial conditions, soil hydraulic parameters, Miller scaling factors (describing small-scale heterogeneity), upper boundary condition, and the local equilibrium assumption by the Richards

equation. These were accounted for in the process. Other components deemed to be of lesser importance and were neglected. The most noteworthy ones of these are hysteresis, violations of the 1D assumption and measurement biases, which might affect our results.

We designed a three stage approach to directly represent and estimate all key uncertainties, except for errors caused by the local hydraulic equilibrium assumption. These intermittent errors are handled by introducing a closed-eye period. Due to an iterative approach we can perform the estimation on a single rain event.

In the first stage, the prior is improved. In our case this is the initial water content distribution and Miller scaling factors. A good prior for the Miller scaling proved particularly beneficial, because during phases without dynamics, water content distributions generated by heterogeneity cannot be distinguished from low hydraulic conductivities.

The second stage is the iterative application of a Standard EnKF with an augmented state to improve the representation. This state consists of soil water contents, soil hydraulic parameters, Miller scaling factors, and the upper boundary condition, to improve the representation. This approach incorporates possible errors in the dynamics into the augmented state components. Since parameters are assumed to be constant in time, we can detect intermittent errors in the dynamics from fluctuating parameters.

The third stage is the application of a Closed-Eye EnKF, that only estimates the full augmented states when and where we assume the dynamics to be correctly represented. Outside of this range we ‘close an eye’ and do not estimate the previously varying parameters. In this way, the state is guided through this ill-represented phase and the full estimation is picked up afterwards again.

The approach showed that homogeneous soil layers are not sufficient to represent heterogeneity. We were forced to also assume small-scale heterogeneity by observations of consistently higher water contents at probes that are located higher up within a soil layer. This emphasizes the role of heterogeneity which must be considered more extensively in studies that rely on local measurements like the used TDR measurements.

We employed a simplified representation of heterogeneity by estimating Miller scaling factors at measurement locations and interpolation between them. This captures the main features. Miller scaling is not capable to represent the heterogeneity completely, however. Predictions for the first layer showed, that measurements from the topmost TDR can be predicted better than the measurements of the second and third TDR probes. We attribute this to limitations of the Miller scaling. The parameter estimation is primarily influenced by the topmost TDR. Describing the other measurements with the same parameter set scaled by the Miller factor cannot fully describe the material properties there.

An iterative Standard EnKF was used to successfully estimate soil parameters, Miller scaling parameters and upper boundary condition. It was capable to predict the rain event that was used for its calibration very well. However, it fails at the prediction of a different rain event. The reason is the violation of the local equilibrium assumption by the Richards equation (Eq. 1) during rain events. ~~This can be seen from variations of the parameters that are larger than the change from initial to final value during later iterations.~~ Non-represented errors are incorporated into the parameters. Since these errors are different for different rain

events (with different strengths and different local non-equilibrium), they cannot be predicted. Due to the iterative approach, we could detect the times when the local-equilibrium assumption is violated: the variations of the parameters are larger than the change from initial to final value during later iterations. The changes in the direction of the parameter update then determine the closed-eye period.

- 5 Generally, the closed-eye period can be detected, if the operational limits of the model are known. In our case, we base this on the changing parameters, but e.g. a direct detection in the state or forcing could be possible as well.

The Closed-Eye EnKF omits ~~these~~ the incorporation of the model structural errors in the parameters and is a generally applicable concept. In this study, it yields better predictions during ~~dry~~ periods when the ~~local equilibrium assumption is fulfilled~~ underlying assumptions are fulfilled: the drying period after a rain event when there is local equilibrium, showing the strength of the Richards equation there. As a consequence, however, its predictions are worse during rain events, when the local equilibrium assumption is violated.

During the closed-eye phase, a description of these uncertainties caused by the non-equilibrium is desirable to be able to optimally guide the states through this phase. In our case, the method still performed well without such a description, as the errors were small enough to be compensated by the adaptive covariance inflation.

Our approach is capable to find parameters closer to the believed true material properties of the soil than a standard EnKF. Predictions during rain events would require an additional representation of the dynamics during the rain event, though. Still, our approach shows a way to limit the incorporation of errors into parameters and is one step towards the concept of Knowledge Fusion - the consistent aggregation of all information pertinent to some observed reality.

20

Author contributions. HHB designed, implemented, performed, and analyzed the presented study. SJ provided computational software. DB provided discussions on the statistical foundations from a particle filter point of view. KR came up with the idea of knowledge fusion and the closed-eye period. All authors participated in continuous discussions. HHB prepared the manuscript with contributions from all authors.

Acknowledgements. This research is funded by Deutsche Forschungsgemeinschaft (DFG) through project RO 1080/12-1.

References

- Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: FAO Irrigation and Drainage Paper No. 56: Crop Evapotranspiration, FAO Water Resources Development and Management Service, 2006.
- Anderson, J. L.: Spatially and temporally varying adaptive covariance inflation for ensemble filters, *Tellus A*, 61, 72–83, 2009.
- 5 Anderson, J. L. and Anderson, S. L.: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts, *Monthly Weather Review*, 127, 2741–2758, 1999.
- Burgers, G., Jan van Leeuwen, P., and Evensen, G.: Analysis scheme in the ensemble Kalman filter, *Monthly Weather Review*, 126, 1719–1724, 1998.
- Carsel, R. F. and Parrish, R. S.: Developing joint probability distributions of soil water retention characteristics, *Water Resources Research*, 10 24, 755–769, 1988.
- Chen, Y. and Zhang, D.: Data assimilation for transient flow in geologic formations via ensemble Kalman filter, *Advances in Water Resources*, 29, 1107–1122, 2006.
- Crow, W. T. and Van Loon, E.: Impact of incorrect model error assumptions on the sequential assimilation of remotely sensed surface soil moisture, *Journal of Hydrometeorology*, 7, 421–432, 2006.
- 15 De Lannoy, G. J., Houser, P. R., Verhoest, N. E., and Pauwels, V. R.: Adaptive soil moisture profile filtering for horizontal information propagation in the independent column-based CLM2.0, *Journal of Hydrometeorology*, 10, 766–779, 2009.
- Erdal, D., Neuweiler, I., and Wollschläger, U.: Using a bias aware EnKF to account for unresolved structure in an unsaturated zone model, *Water Resources Research*, 2014.
- Erdal, D., Rahman, M., and Neuweiler, I.: The importance of state transformations when using the ensemble Kalman filter for unsaturated 20 flow modeling: Dealing with strong nonlinearities, *Advances in Water Resources*, 86, 354–365, 2015.
- Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *Journal of Geophysical Research: Oceans (1978–2012)*, 99, 10 143–10 162, 1994.
- Foken, T.: *Angewandte Meteorologie*, Springer Berlin Heidelberg, 2006.
- Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, *Quarterly Journal of the Royal Meteorological 25 Society*, 125, 723–757, 1999.
- Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter, *Monthly Weather Review*, 129, 2776–2790, 2001.
- Hendricks Franssen, H.-J. and Kinzelbach, W.: 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, 2008.
- 30 Hopmans, J., Šimůnek, J., Romano, N., and Durner, W.: 3.6.2. Inverse Methods, *Methods of Soil Analysis: Part 4 Physical Methods*, pp. 963–1008, 2002.
- Houtekamer, P. L. and Mitchell, H. L.: Data assimilation using an ensemble Kalman filter technique, *Monthly Weather Review*, 126, 796–811, 1998.
- Houtekamer, P. L. and Mitchell, H. L.: A sequential ensemble Kalman filter for atmospheric data assimilation, *Monthly Weather Review*, 35 129, 123–137, 2001.
- Huber, E., Hendricks-Franssen, H., Kaiser, H., and Stauffer, F.: The role of prior model calibration on predictions with ensemble Kalman filter, *Ground water*, 49, 845–858, 2011.

- Ippisch, O., Vogel, H.-J., and Bastian, P.: Validity limits for the van Genuchten–Mualem model and implications for parameter estimation and numerical simulation, *Advances in Water Resources*, 29, 1780–1789, 2006.
- Jafarpour, B. and Tarrahi, M.: Assessing the performance of the ensemble Kalman filter for subsurface flow data integration under variogram uncertainty, *Water Resources Research*, 47, 2011.
- 5 Li, C. and Ren, L.: Estimation of unsaturated soil hydraulic parameters using the ensemble Kalman filter, *Vadose Zone Journal*, 10, 1205–1227, 2011.
- Liu, Y. and Gupta, H. V.: Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, *Water Resources Research*, 43, 2007.
- Liu, Y., Weerts, A., Clark, M., Hendricks Franssen, H., Kumar, S., Moradkhani, H., Seo, D., Schwanenberg, D., Smith, P., Van Dijk, A.,
 10 et al.: Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities, *Hydrology and Earth System Sciences*, 16 (10), 2012, 2012.
- Man, J., Li, W., Zeng, L., and Wu, L.: Data assimilation for unsaturated flow models with restart adaptive probabilistic collocation based Kalman filter, *Advances in Water Resources*, 92, 258–270, 2016.
- Miller, E. and Miller, R.: Physical theory for capillary flow phenomena, *Journal of Applied Physics*, 27, 324–332, 1956.
- 15 Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P. R.: Dual state–parameter estimation of hydrological models using ensemble Kalman filter, *Advances in Water Resources*, 28, 135–147, 2005.
- Mualem, Y.: A new model for predicting the hydraulic conductivity of unsaturated porous media, *Water Resources Research*, 12, 513–522, 1976.
- Nowak, W.: Best unbiased ensemble linearization and the quasi-linear Kalman ensemble generator, *Water Resources Research*, 45, 2009.
- 20 Reichle, R. H.: Data assimilation methods in the Earth sciences, *Advances in Water Resources*, 31, 1411–1418, 2008.
- Reichle, R. H., McLaughlin, D. B., and Entekhabi, D.: Hydrologic data assimilation with the ensemble Kalman filter, *Monthly Weather Review*, 130, 2002.
- Richards, L. A.: Capillary conduction of liquids in porous mediums, *Physics*, 1, 318–333, 1931.
- Roth, K., Schulin, R., Flühler, H., and Attinger, W.: Calibration of time domain reflectometry for water content measurement using a com-
 25 posite dielectric approach, *Water Resources Research*, 26, 2267–2273, 1990.
- Schenk, G. U.: Soil-Atmosphere coupling: Looking from within the soil towards the surface, Ph.D. thesis, Ruperto-Carola University Heidelberg, Germany, <http://www.ub.uni-heidelberg.de/archiv/12840>, 2011.
- Shi, L., Song, X., Tong, J., Zhu, Y., and Zhang, Q.: Impacts of different types of measurements on estimating unsaturated flow parameters, *Journal of Hydrology*, 524, 549–561, 2015.
- 30 Song, X., Shi, L., Ye, M., Yang, J., and Navon, I. M.: Numerical comparison of iterative ensemble Kalman filters for unsaturated flow inverse modeling, *Vadose Zone Journal*, 13, 2014.
- Van Genuchten, M. T.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Science Society of America Journal*, 44, 892–898, 1980.
- Vrugt, J. A., Diks, C. G., Gupta, H. V., Bouten, W., and Verstraten, J. M.: Improved treatment of uncertainty in hydrologic modeling:
 35 Combining the strengths of global optimization and data assimilation, *Water Resources Research*, 41, 2005.
- Vrugt, J. A., Stauffer, P. H., Wöhling, T., Robinson, B. A., and Vesselinov, V. V.: Inverse modeling of subsurface flow and transport properties: A review with new developments, *Vadose Zone Journal*, 7, 843–864, 2008.

- Wollschläger, U., Pfaff, T., and Roth, K.: Field-scale apparent hydraulic parameterisation obtained from TDR time series and inverse modelling., *Hydrology & Earth System Sciences*, 13, 2009.
- Wu, C.-C. and Margulis, S. A.: Feasibility of real-time soil state and flux characterization for wastewater reuse using an embedded sensor network data assimilation approach, *Journal of Hydrology*, 399, 313–325, 2011.
- 5 Wu, C.-C. and Margulis, S. A.: Real-time soil moisture and salinity profile estimation using assimilation of embedded sensor datastreams, *Vadose Zone Journal*, 12, 2013.
- Zhang, D., Madsen, H., Ridler, M. E., Refsgaard, J. C., and Jensen, K. H.: Impact of uncertainty description on assimilating hydraulic head in the MIKE SHE distributed hydrological model, *Advances in Water Resources*, 86, 400–413, 2015.

Table 1. Prior uncertainties of the representation of soil water movement at the Grenzhof test site during October and November 2011. We assess the uncertainty caused by assumptions in a qualitative way: small, no need to represent the uncertainty explicitly; intermediate, it might be necessary to represent these uncertainties, we decide not to but keep them in mind; large, uncertainties must be represented with the goal to reduce them.

Assumption	Uncertainty
Dynamics:	
Single phase description (Richards equation)	small
Local equilibrium (Richards equation)	small- large
1D, no sources and sinks	intermediate
Material Properties:	
Mualem-van Genuchten parameterization for material properties	small
Mualem-van Genuchten parameters from Wollschläger et al. (2009)	large
Mualem-van Genuchten parameters constant in time	small
No hysteresis	intermediate
Large-scale architecture described with layers	small
Small-scale architecture described with Miller scaling	small-intermediate
Miller scaling factors are 1	large
Forcing:	
Upper boundary condition from rain gauge and FAO Penman-Monteith eq.	large
Lower boundary condition constant groundwater table decoupled with soil layer	small
Linear interpolation for initial condition	large
Numerical representation	small
Water content measurements without biases and uncertainty of 0.01	intermediate

Table 2. Soil hydraulic parameters as estimated by Wollschläger et al. (2009) and our initial uncertainties for the Standard EnKF

Layer	α [1/m]	$\log_{10}(K_0)$ [m/s]	n	θ_s	θ_r	τ
1	4.8 ± 2.0	-4.6 ± 1	1.61 ± 0.20	0.39	0.020	0.5
2	11.0 ± 2.0	-5.3 ± 1	1.80 ± 0.20	0.40	0.070	0.5
3	7.4 ± 2.0	-6.0 ± 1	1.43 ± 0.20	4.27 <u>0.43</u> ± 0.02	0.090	0.5
4	2.9 ± 2.0	-5.4 ± 1	1.36 ± 0.20	4.33 <u>0.43</u> ± 0.02	0.100	0.5
5	14.5	-4.1	2.68	0.43	0.045	0.5

Parameters without uncertainty are not included in the augmented state and are not estimated. The saturated water content θ_s of the layers 3 and 4 were already estimated during the prior estimation in time period A, because measurement values exceeded the value of 0.36 determined by Wollschläger et al. (2009).

Table 3. Results for the soil hydraulic parameters estimated with the Standard EnKF.

Layer	α [1/m]	$\log_{10}(K_0)$ [m/s]	n	θ_s	θ_r	τ
1	4.7 ± 0.3	-5.23 ± 0.04	1.43 ± 0.01	0.39	0.020	0.5
2	10.1 ± 0.7	-4.08 ± 0.25	1.66 ± 0.02	0.40	0.070	0.5
3	8.2 ± 1.0	-6.04 ± 0.31	1.33 ± 0.03	0.42 ± 0.01	0.090	0.5
4	2.3 ± 0.3	-5.71 ± 0.40	1.31 ± 0.01	0.43 ± 0.01	0.100	0.5

Parameters without uncertainty are not included in the augmented state and are consequently not estimated.

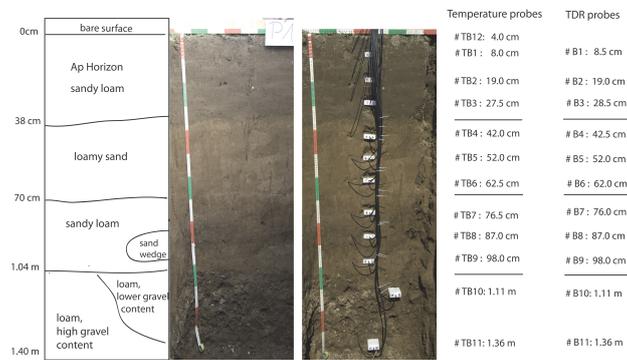


Figure 1. Soil profile at the Grenzhof test site close to Heidelberg, Germany (Schenk, 2011) . It consists of four different soil layers and is equipped with twelve temperature probes and eleven TDR probes. The surface is kept free from vegetation.

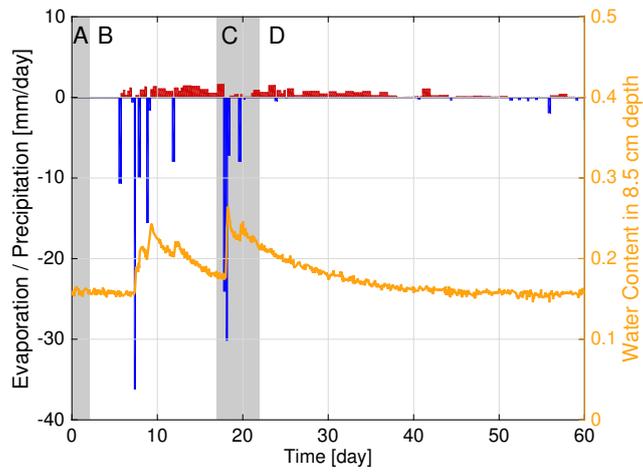


Figure 2. Boundary condition and topmost TDR from Oktober 1 (day 1) until November 29 (day 60) at the Grenzhof test site. We distinguish 4 sections: A (day 1-2), B (day 3-17), C (day 18-22) and D (day 23-60). The evaporation is calculated using the reference FAO Penman-Monteith equation. Before the first rain the evaporation is set to 0, due to a previous dry period of over a month. The precipitation is displayed with a 6 hour resolution.

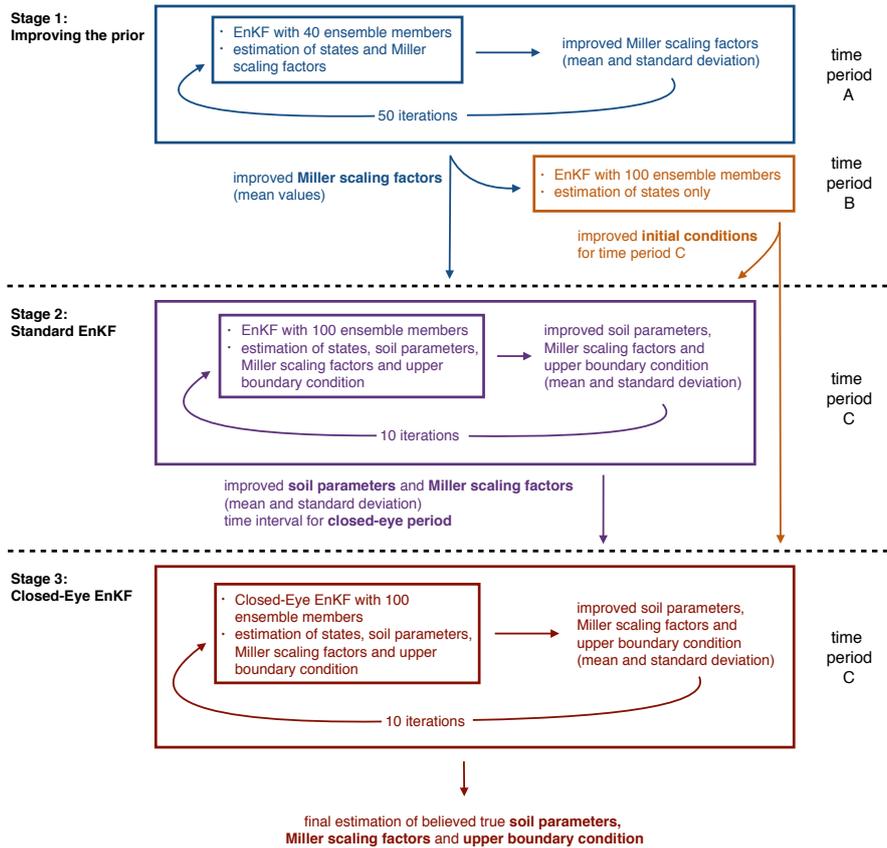


Figure 3. Block diagram illustrating the three stage approach. The different time periods are shown in Fig. 2.

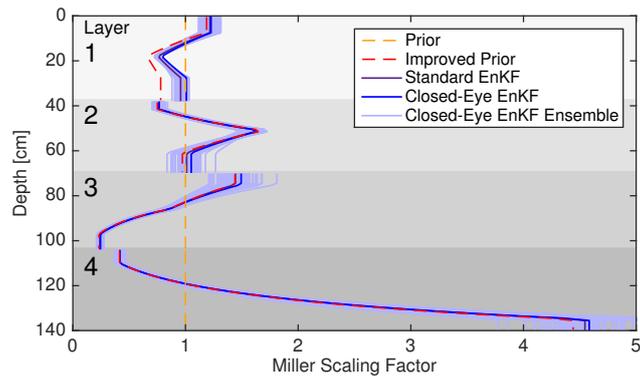


Figure 4. Mean values of the Miller scaling field. Soil layers are indicated by different gray scales. The heterogeneity is a priori unknown and the according prior is set to 1. The natural logarithm of the scaling factors is estimated at the measurement locations and interpolated linearly between (with constant extrapolation to layer boundaries). Already the improved prior can describe the main features. The further estimations with the Standard and Closed-Eye EnKF lead to further small changes mainly in the first layer. For the Closed-Eye EnKF the actual ensemble indicating the uncertainties is additionally shown.

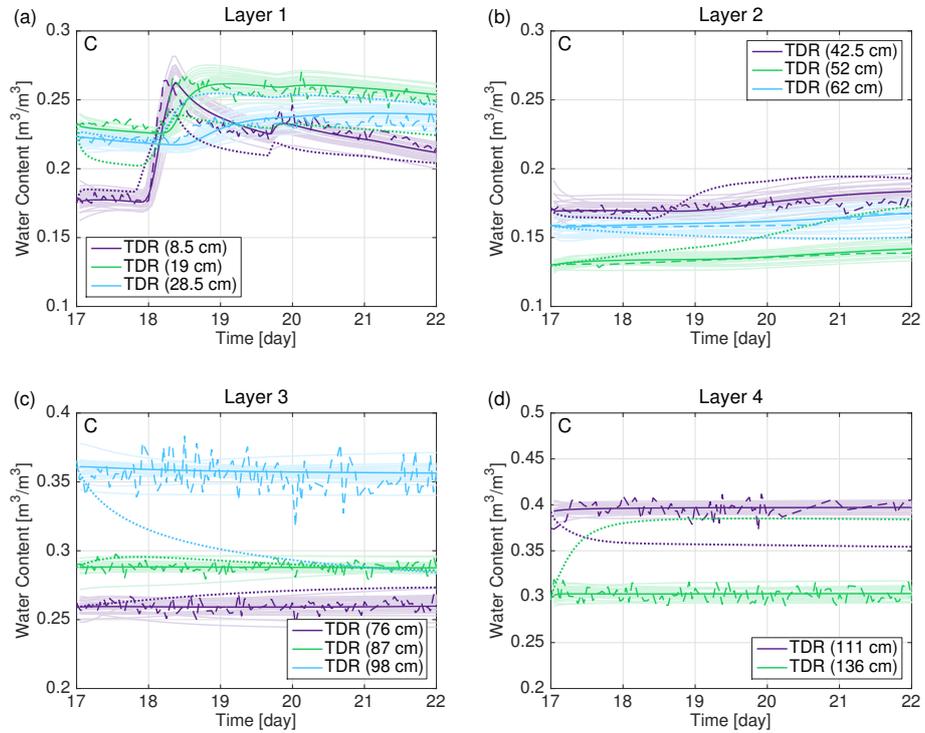


Figure 5. Forecast at the measurement positions of the four different layers during time period C. The solid lines show the mean value of a total of 100 different forecasts. The pale colors show the results from 25 of these, with soil parameters, Miller scaling parameters and boundary condition sampled from the distributions estimated with the Standard EnKF and the initial condition, that was actually used for the estimation itself. The dashed lines show the measurements, the dotted lines results from the original soil parameters by Wollschläger et al. (2009). The measurements were recorded hourly, occurring data gaps were interpolated linearly.

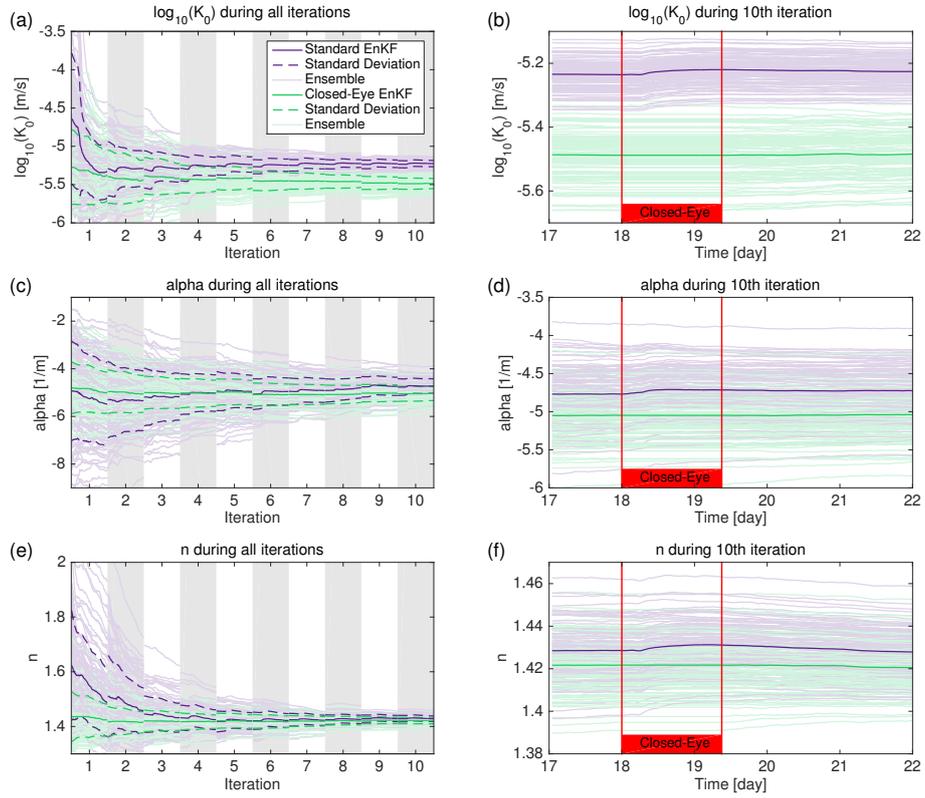


Figure 6. Evolution of the parameters K_0 , α and n of the first layer during the estimation with Standard EnKF (purple) and the Closed-Eye EnKF (green) for all iterations (a), (c), (e) and the last iteration (b), (d), (f).

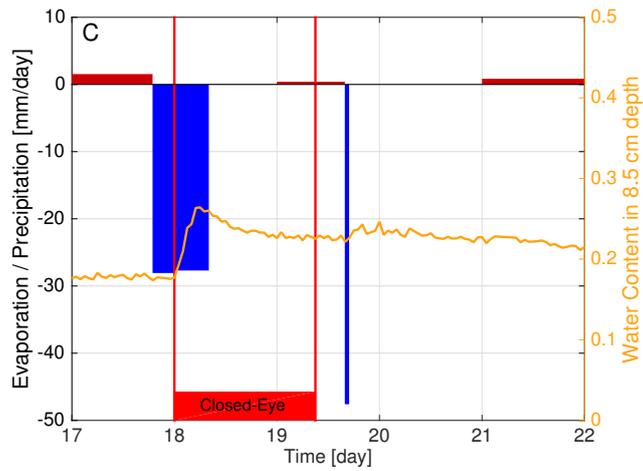


Figure 7. Boundary condition along with the topmost TDR water content measurements during time period C. The closed-eye time period was determined based on the changes in parameter K_0 (Fig. 6). During this time the dynamics cannot describe the water content measurements any more, most likely because of the violation of the local equilibrium assumption. The boundary condition is displayed in the reduced resolution assumed during the estimation.

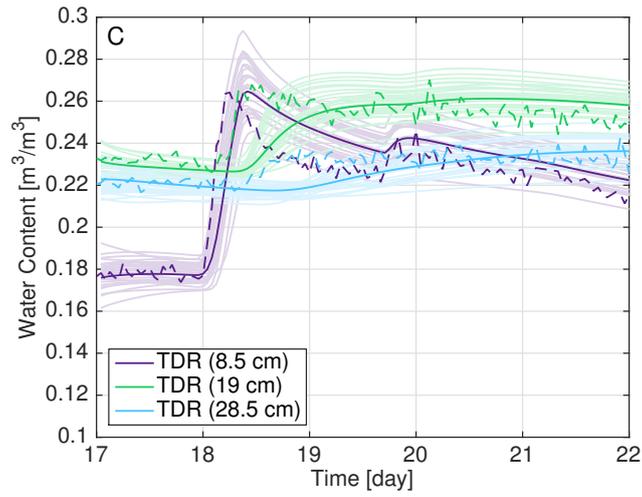


Figure 8. Forecast at the measurement positions of the first layer during time period C with the results from the Closed-Eye EnKF. The solid lines show the mean value of a total of 100 different forecasts. The pale colors show the results from 25 of these, with soil parameters, Miller scaling parameters and boundary condition sampled from the distributions estimated with the Closed-Eye EnKF and the initial condition, that was actually used for the estimation itself. The dashed line shows the measurements.

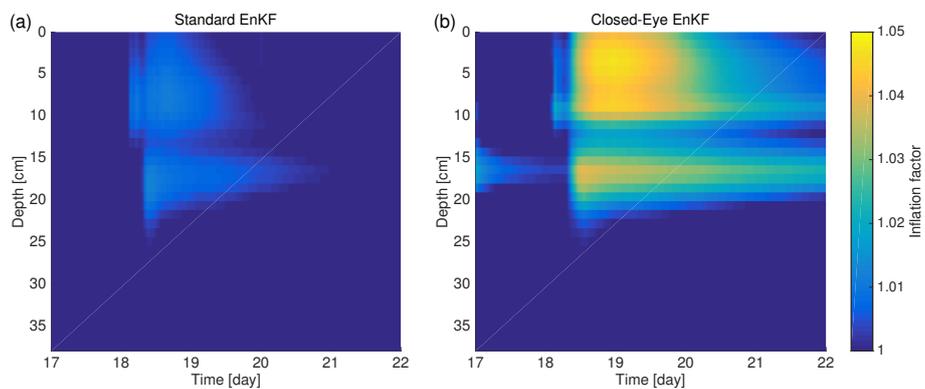


Figure 9. Inflation factor λ in the first layer during the last iteration of the Standard EnKF and Closed-Eye EnKF. The adaptive covariance inflation increases the inflation factor around a measurement location when the deviations due to the infiltration reach the corresponding TDR. The deviations and consequently the inflation are stronger for the Closed-Eye EnKF, as the preferential flow is not incorporated into the parameters any more. The values at the end of an iteration are used as the initial inflation factors for the next iteration.

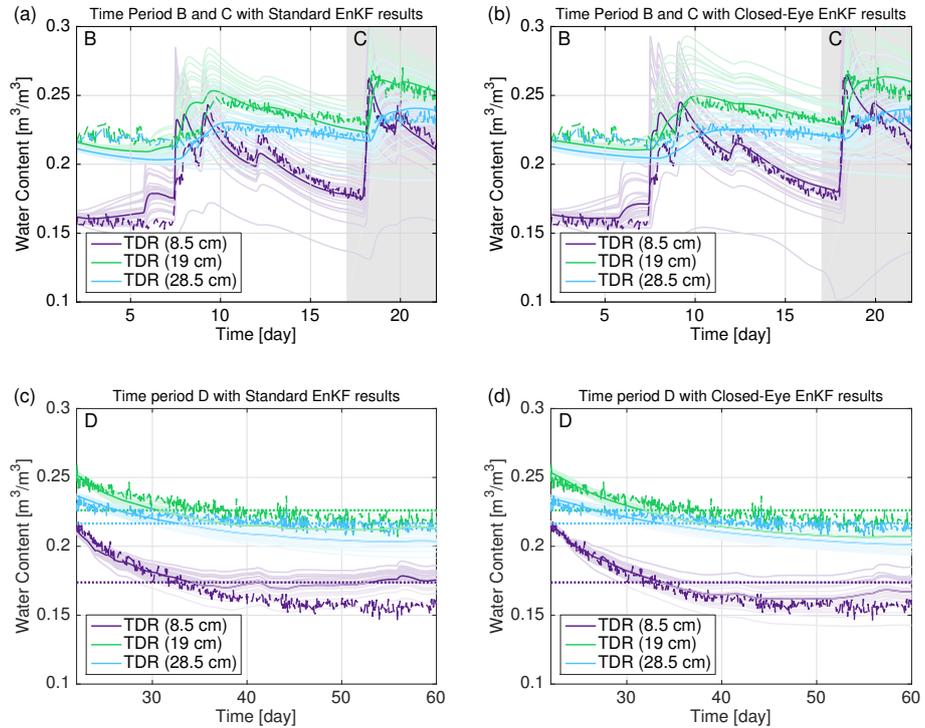


Figure 10. Forecast at the measurement positions of the first layer during time period B and C (a,b) as well as D (c,d). The solid lines show the mean value of a total of 100 different forecasts. The pale colors show the results from 25 of these, with soil parameters and Miller scaling parameters sampled from the distributions estimated with the Standard EnKF (a,c) and Closed-Eye EnKF (b,d). The dashed line shows the measurements, the dotted line the lowest water content measured during time period C, indicating the lower limit of the water content range, for which the parameters are deemed valid.