Dear authors,
Thanks for taking care of the reviewers comments and providing an improved manuscript.
After reading the revised version, I still have the following comments/questions which motivate minor revisions:

Dear editor,
We thank you for the effort and time to review our manuscript and our reply to the reviewers' comments. In the following we will answer to your comments in detail.

L145: The flux you apply is the real evapotranspiration. Should not be called potential evapotranspiration.

Yes, this is right. We removed the word "potential".

L221-223: Could you provide some guidelines to better qualify what is 'adequate' and 'sufficient' ?

We agree that this formulation is too imprecise. A convergence study on grid resolution and number of zones is needed to determine whether the grid is "adequate" and the number of unsaturated zone models "sufficient". We added "which can both be found by a convergence study" at the end of this sentence. With the convergence study one would refine the grid and the number of unsaturated zone models until results do not change more than a pre-defined criterion. We hope that the meaning of this is clear and therefore do not specify this further.

L268, typo 'Terefore'

Yes, right. We corrected that.

L366-369. Unclear to me. Does it mean that the heterogeneity of the unsaturated zone changes with time i.e. KUZ is replaced by KGW when the water level rises? Please clarify.

Yes, this is how it is done. The saturated cells of the unsaturated zone model are assigned the KGW value. This means that the conductivity changes when a cell becomes saturated or unsaturated. As we wanted to analyze both, the influence of the conductivity of the groundwater and that of the unsaturated zone, we had to choose: Either keep the K values in the cells fixed and get an inconsistency in the saturated (overlapping) part or change the K values in the unsaturated model according to the water table position. We decided to follow the second approach to avoid any possible numerical issues that could arise from large differences of K in the groundwater model and the saturated part of the unsaturated zone model. Besides, since we have (almost) hydrostatic equilibrium in the saturated part of the 1D models, the effect of changing K in those parts should be minor. We added the two sentences: "This means that the hydraulic conductivity of an unsaturated model cell is changed when the cell becomes saturated or unsaturated. Thus, inconsistencies between the two model compartments that may lead to numerical problems are avoided." for clarification.

Sensitivity analyses
You have to convince me that 316 simulations and a second order polynomial are enough to described f(x) (equation 16) for your highly coupled non-linear models... The simplest way to do this is to perform the polynomial fitting without taking care of some model outputs (say about 20 outputs) and then compare the predicted values with the model outputs for these 20 values (cross validation). A nice plot with the differences between both will show the quality of f(x).

We did such a plot for all three observation types (water table position, water table fluctuation and specific yield). We chose the last time step to be sure that no spin-up

effects tamper with the result and used 26 samples for validation. The result is shown in Figure 1. For the water table fluctuations, the fitting is less accurate with $R^2$ at 0.81, but for the other two observation types, the model output and the approximated output fit very well. We added the sentence "The polynomial representation (Eq. 16) was cross validated with 26 randomly chosen parameters and found to be appropriate (result not shown)." at the end of the first paragraph in Section 4.4. We can also add the plot to the manuscript but prefer to keep it out as it is no central point and would make the manuscript even longer.

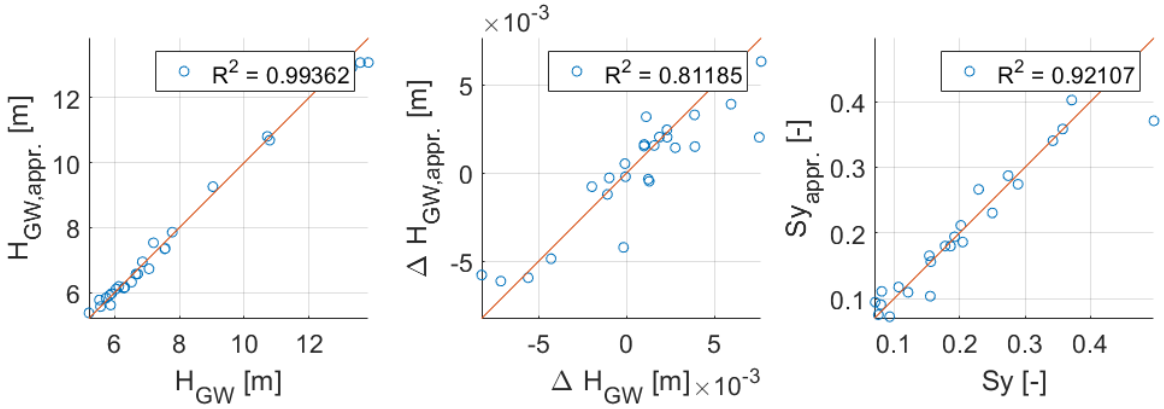


*Figure 1: Observations and observations approximated by f(x).*

You nicely showed that the iterative algorithm is more accurate. What is the interest of performing the sensitivity analyzes with the non-iterative scheme? The physics embed in the model is the same and the model outputs for both scheme very close. Therefore, L529 is obvious. What about removing the non-iterative scheme in the GSA?

This is a good suggestion. We removed the non-iterative coupling scheme from the GSA.