Referee comment on "Coupling saturated and unsaturated flow: comparing the iterative and the non-iterative approach" by Natascha Brandhorst et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2021-15-RC2, 2021

The manuscript presents two methodologies of coupling for dimensionally heterogeneous modelling of subsurface flow in the unsaturated zone and in the saturated zone. The interest of such an approach is that important diminutions of computation times may be obtained compared to fully 3D modeling approaches. The main drawback is that accuracies of the simulations are dommaged, still in comparison with 3D modelling, and more or less along the considered cases and coupling methodologies. Given the very large computation times that may be encountered in fully mechanistic hydrological modeling at the wtareshed scale, this problem is of great interest for the community of hydrological modeling. The manuscript contains an important material in terms of numerical results and provides relevant hints to compare the two coupling strategies under concern.
Nevertheless the presentation of the considered theories and numerical experiments lack of rigor, and the writting of the manuscript is not clear enough. In some places additional computations may even be needed. Thus I think it should be thoroughfully reworks prior to publication. I recommend to reject the paper in its present form, and to encourage the authors to resubmit after having completing and improving it.

We thank the reviewer for the effort and time to revise our manuscript. He provided many helpful comments and remarks that need to be reworked by us. However, we do not understand his suggestion to reject the manuscript based on his comments which mainly require further explanations and only a few additional computations. In the following, we will answer to all comments in detail and hope that we can thus clarify all open points, but also underline our point of view where we do not agree with the reviewer.

General comments:

- The considered equations should be defined more rigourously and rewritten. For instance, the double time derivative term in equation (2) is a non-standard formulation of Richards equation (see for instance Gottardi and Venutelli, 1993). I guess that considerations related to the order on magnitudes of those two time derivatives may be used to justify the adopted formulation, but it should be explicited. Moreover, the use of the same notation S_y for the specific yield in equation (1), which has classically a clear and well identified physical meaning (drainage porosity), and for the iteratively computed, time variable fitting parameter used in the iterative method to handle recharge fluxes from the unsaturated zone is confusing. I think that the latter one should be expressed as the sum of the true specific yield S_y and a new additional term used for the purpose of the coupling between the saturated zone and the non-saturated zone. This would not imply new computation, but simply to rewrite some equations and rescale some results. I think that the added value of such more accurate notations in terms of clarity and of ease of physical interpretation would be important.

We will answer this comment separately for equation (1) and (2). The comment implies that there is a general issue concerning more equations, but as no other equations are mentioned explicitly and we do not see any problems with our formulations we will limit our reply to these two:

Equation (1): This formulation was also noted by the third reviewer, so we will give the same explanation here:

The left member of Eq. 2 is derived from the time derivative in the volume balance (we assume incompressibility of water):

$$\frac{\partial V_w}{\partial t} = \frac{\partial\left(S(h_p)\phi V_t\right)}{\partial t} = \phi V_t\,\frac{\partial S(h_p)}{\partial t} + S(h_p) V_t\,\frac{\partial \phi}{\partial t} + S(h_p)\phi\,\frac{\partial V_t}{\partial t}$$

$$= V_t \cdot \left(\phi \frac{\partial S(h_p)}{\partial t} + S(h_p)\frac{\partial \phi}{\partial t} + \frac{1}{V_t} S(h_p)\phi \frac{\partial V_t}{\partial t}\right)$$
$$= V_t \cdot \left(\frac{\partial (S(h_p)\phi)}{\partial t} + S(h_p)\frac{\phi}{V_t}\frac{\partial V_t}{\partial h_p}\frac{\partial h_p}{\partial t}\right)$$
$$= V_t \cdot \left(\frac{\partial (S(h_p)\phi)}{\partial t} + S(h_p) S_s \frac{\partial h_p}{\partial t}\right)$$

with subscripts *w* and *t* denoting *water* and *total.*

The specific storage is in general defined as $S_s = \frac{1}{V_t}\frac{\partial V_p}{\partial h_p} = \frac{\partial \phi}{\partial h_p} + \frac{\phi}{V_t}\frac{\partial V_t}{\partial h_p}$ with $V_p = \phi V_t$ being the pore volume. This definition is used when the change of porosity is written in dependence of the change in water pressure head (so $\frac{\partial \phi}{\partial t} = \frac{\partial \phi}{\partial h_p}\frac{\partial h_p}{\partial t}$). One then also gets a slightly different formulation of Eq. 2: $\frac{\partial V_w}{\partial t} = V_t \cdot \left(\phi \frac{\partial S(h_p)}{\partial t} + S(h_p) S_s \frac{\partial h_p}{\partial t}\right)$.

In our case, we maintain the time derivative of porosity (or rather water content, as $S(h_p)\phi = \theta$) and therefore the specific storage formulation reduces to $S_s = \frac{\phi}{V_t}\frac{\partial V_t}{\partial h_p}$, which is equivalent to assuming that $\frac{\partial \phi}{\partial h_p} \cong 0$.

This formulation of Richards' equation can be found in e.g., Kavetski et al. (2001), Kollet and Maxwell (2006); Fahs et al. (2009). We understand the reviewer's comment that there are more common formulations of this equation and will therefore clarify the definition of the specific storage in the revised manuscript.

Equation (2): Here, we disagree, but we understand the comment since the specific yield in equation (2) is often set equal to drainage porosity. The problem is that this definition neglects the soil water stored above the groundwater table and thus overestimates the free capacity for groundwater storage. However, using drainage porosity is kind of an "established" way since it can be accurate enough in some applications and there is simply no better way of estimating this parameter without modelling or, in case of field studies, monitoring the water content in the unsaturated zone. When deriving the 2D groundwater flow equation for an unconfined aquifer from the volume balance (we assume incompressibility of water), the specific yield turns out to be equal to

$$S_y = \frac{\partial V(h)}{\partial h} = \frac{\partial}{\partial h}\left(\int_{z\_bottom}^{z\_surface} \theta(h,z)dz\right)$$

to ensure mass conservation. h is the water pressure head. This can be close to drainage porosity under specific conditions, but in general it is not and can be even far off, depending on the soil moisture profile. This dependency is no new finding and has been described by others (see the related literature cited in the introductory part of our manuscript) who also object to defining the specific yield simply as drainage porosity. We see that depending on the application this definition might be sufficient and not cause any larger problems. However, in our case (where we have information about the soil moisture profile) a definition as drainage porosity would mean violating the conservation of mass. By the adaptive determination of this value during the iterations, we use this information to get a better estimate of this "true" specific yield. Unfortunately, we cannot get there entirely due to the influence of the lateral fluxes on the recharge as we explain in l. 222-228. It is important to us that it is understood that this is not a mere coupling parameter for handling the recharge fluxes in our model but is as close as we can get to the value of this.
Concerning the notation, we could indeed write the specific yield as the sum of two terms, the drainage porosity and the part that takes the soil moisture into account. However, we would prefer not to do this in order to be consistent with other papers that are relevant for

our work (e.g., Pikul et al., 1974; Crosbie et al., 2005). In these papers, S_y as specific yield is also not used as drainage porosity.

- Convergence studies for mesh refinement and time stepping strategy are not evocated as it should be the case in any study producing Computational Fluid Dynamics results. In some places it may imper the possibility to understand the comparative behaviours of the proposed test cases. For instance if we consider the comparison of accuracies of test 2 and test 3, in the present form of the manuscript it is impossible to say what comes from the differences of meshes and what comes from the different physics under concern (e.g.: homogeneous versus heterogeneous soil).

This is a fair comment and we agree that grid convergence should be addressed. We performed convergence studies testing the general convergence behavior of the model but did not perform any on the presented test cases. One reason for this (regarding the first two test cases) is that we wanted to use the same grid as in Beegum et al. (2018) to keep the results comparable. The other reason is that we had concerns that a convergence analysis would make the manuscript too long. Furthermore, we need to keep the comparability among the coupling approaches and the reference model which we might lose if a convergence study tells us to use different grid sizes or time steps for the different models. Regarding the comparability between the test cases (especially the 2D and the 3D test case), we agree with the reviewer that a convergence study should be performed to exclude possible effects of using different grid sizes. Therefore, our suggestion is to perform a convergence study with the iteratively coupled model for the three applied test cases and then choose the grid and time step size for the other models accordingly. Thus, we could guarantee the comparability of results among the models and the test cases.

- To the knowledge of the reviewer, an important example of hydrological model that couples dimensionnaly heterogeneous descriptions of flow in the saturated zone and in the unsaturated zone is MIKE-SHE (e.g.: Graham and Butts, 2005), which is for instance included in recent international benchmarking efforts for physically based hydrological modeling (e.g.: Kollet *et al.*, 2016). The fact that works related to MIKE-SHE do not appear in the references of the manuscript make me think that the bibliographical survey on which the presentation of the background of the study is done should be consolidated.

We are fully aware of the MIKE-SHE model and will include references to MIKE-SHE in our introduction. Its coupling strategy does not differ substantially from the models we already referred to (no overlap of the two compartments, compartments solved separately, iterative procedure with step-wise adjustment of water table to improve mass balance), but we agree that it is a widely-used model for coupling unsaturated and saturated flow and should be therefore mentioned as an example in our literature survey.

Specific comments:

- l 136-137: "a Neumann boundary representing net flux from precipitation and evapotranspiration" : with the source/sink term of the equation (2), it is possible to represent actual evapotranspiration distributed in time and space according to water avalaibility in the soil (see for instance Orgogozo et al., 2019) ; please discuss the limitation associated with an a priori estimation of the actual evapotranspiration directly embedded in the Neumann boundary.

We agree that the source/sink term can be used to represent actual evapotranspiration in the root zone. Here, we do it in a simplified way merging the net fluxes into the Neumann boundary at the top because we have no information on the root zone distribution or vegetation. We could assume one, but the good match of the simplified approach and the results presented in Beegum et al. (2018) show that it is not needed. In the other two test cases there is only precipitation and evaporation does not play a role. Also, our focus is on the groundwater table. The implementation of evaporation would be more important if the focus would be on land surface processes. We will change the sentence to "The upper boundary condition of the unsaturated zone models is a Neumann boundary representing the flux across the land surface. In this work, the flux will be either precipitation or the net flux from precipitation and potential evapotranspiration depending on the test case." to make the boundary condition clearer.

- l 151 : "collect the computed recharge (i.e. flux leaving over the bottom boundary) and interpolate the 2D map of groundwater recharge." : You mean collect all the computed regarges for all time steps of the 1D Richards model since the previous time step of coupling ? Should be clarified.

Indeed, that is what we mean. This will be made more explicit in the revised manuscript.

- l 157 : "Add (or subtract) a ratio r of this water to the recharge computed in the next time step." you mean the next time step of coupling ? Should be explicited.

Indeed, that is what we mean. This will be made more explicit in the revised manuscript.

- l 159 – 171 : The proposed way of chosing the ratio r is difficult to accept. In case of water table elevation, the ratio r could be fitted to keep unchanged across the mesh resizing process the total amount of water contained in the part of the domain that stays unsaturated, while in case of water table lowering a 'field capacity' water saturation could be prescribed to the cells that experienced desaturation in order to compute a total water amount to be distributed in the new 1D mesh, with an associated proper r ratio? Here the formulae proposed for the computation of the ratio r seems somewhat arbitrary. For instance the point (1) "the groundwater table rise or fall is also effected by lateral flows" is already taken into account in the 2D groundwater model. Besides, "the unsaturated zone is really compacted by a rise of groundwater levels" does not sound physical at all.

In principle we do not disagree with the reviewer on this point. The non-iterative method consists to a fair degree of ad-hoc solutions that originate from the aim to keep the model as simple as possible while maintaining acceptable results. In this, we found the suggested r-ratio to give a reasonable result. A lot of testing, similar to those suggested by the reviewer have been done on the model, but the sparse representation of the unsaturated zone often causes practical problems with refined methods. Nonetheless, it should clearly be noted here that the non-iterative model, which is taken from a previous publication on speeding up model spin-up (Erdal et al 2019), in this work serves mainly as a comparison to the iterative one. This is the new method for this publication and the better one, and the one we recommend using. Hence, we will, upon revision, make sure to discuss the reviewer's suggestions, but also make it clearer what the purpose of the non-iterative model is.

As for the remarks on the formulations, we will address those upon revision.

- l 183 : "(v) the iteration counter" : with which loop is related this iteration counter is unclear at this point (it could be for instance with the time stepping of the 1D Richards equation or with the coupling time steping)? Although it becomes clear afterward, it should be explicited here, at the first occurence of (v).

We agree that this can be misleading at this point. We will add the information by changing the sentence to "… and the superscript (v) the iteration counter for the coupling time step loop."

- l 203 eq (10) (see the first general comment): According to the basic derivation of the diffusivity equation for unconfined aquifer, the specific yield is equal to the drainage porosity of the considered porous medium – although it seems that it might be different for more elaborated derivations, according to the literature cited by the authors. What is the physical interpretation of the variations the specific yield computed by eq (10)? Is there a theoritical reason why the iterating on the values of the specific yield field in the aquifer should lead to convergence? In case this is a purely empirical methodology, are there cases for which divergence may occure? Other questions : the value of the 'physical' S_y parameters that appears in the equation (1) is only the seed of the iterative process at the first time step of simulation, and do not appear directly anymore in the course of the simulation for the evaluation of S_y $_v$, right? That is what I understand from table 2 for instance. It should be clarified here.

The first part of this comment is already answered in the general part so we will focus here on the remaining questions.

(a) Regarding the convergence of the specific yield: There can only be a consistent solution for the coupled model if the specific yield converges . As in any iterative process the success of this iteration depends on the chosen method, starting value and closure criterion. Divergence can occur if the model error is too large. This could mean here: too coarse grid, too large (coupling) time steps, too few unsaturated zone models, large differences in the hydraulic conductivity in the saturated and unsaturated model (due to the averaging). We will add a paragraph where we discuss this.

(b) Regarding the value of S_y given in table 2: This is indeed only the seed. This is stated in the footnote of table 2 and shown in Figure 3.

- l 221 – 224 : "The source/sink terms $q_{lat,i}$ have an effect on the recharge ('R(Q lat = 0) ≠ R(Q lat ≠ 0)'), which due to the nonlinearity of Richards equation cannot be quantified." However at step (4) (l 216 – 217), an updated R_v is computed that takes into account the $q_{lat,I}$ $_v$? I don't understand.

We agree that this is confusing. We cannot quantify the effect of the lateral fluxes on the recharge, so we keep the water table fluctuation caused by the fluxes in the unsaturated zone during one coupling time step ΔH_{UZ}^1 constant. The recharge R is only changed because the specific yield is updated during the iterations (see equation (7)). We will add the sentence "Changes of the recharge R^v during the iteration are then only caused by changes of the specific yield." at the end of line 225.

- l 229 – part 2.3 Activity score: Difficult to follow. Lack of explanations and of references. There is also a problem of structure: since 'The parameters and the model output f are defined in Sections 3.4 and 4.4, respectively.', this part 2.3 is not possible to understand by itself at this point of the reading. This part should be reworked so that the reader may understand why it is interesting to use the activity scores for the sensitivity study, and on what motivated the choices of the parameters x and the output f(x). By the way, in table 3 part 3.4, the variable K_{GW} and K_{UZ} seems not to be defined in the manuscript ? And why chosing S_y as a parameter of the sensitivity study while it is subjected to iterative evolution of its value along computation in the iterative method (l 218 – table 2)?

In a previous version of the manuscript this part was more detailed, but we considered it too long and taking too much space, so we shortened. We see that we overdid the shortening and will put more explanation back in. We will also add a list of the used parameters and model output already in this section to make it easier to understand. We will also introduce the parameters K_{GW} and K_{UZ} properly.
The aim of the sensitivity analysis is to reveal the most influential parameters. As we claim that the specific yield has a large impact on the model output, we need to include it in the sensitivity analysis to confirm this statement. The goal of the sensitivity analysis is also to check if the not fully coupled models do not generate unphysical behavior in the sense that parameters influence the properties they should influence. We will make this clearer in the revised manuscript.

- l 267 : '3.1 Test case 1: 1D flow' lack of a figure that presents the geometry, the boundary conditions and the meshes for each models.

Agreed. We will provide such a figure.

- Figure 4 : Wrong title for y-axis (Precipitation-PET, not just PET)

The title is correct. We are aware that the abbreviation PET is reserved for potential evapotranspiration in some communities, but we chose to define it otherwise (precipitation evapotranspiration, see l.256). To avoid confusion, we will change the term to P-ET.

- l 276 : Precise which 1D model (pure Richards I suppose ?)

Yes, this is a pure Richard's model. We stated this is in l. 261-262, but we will mention it here again to avoid any misunderstanding.

- l 281 : "The groundwater domain is divided into a 2 × 2 grid. Each groundwater cell is assigned a 1D model." Then the groundwater model is 2D with only 2 cells in each direction ? I don't understand.

Yes, the groundwater model is 2D (we used the same grid as in Beegum et al. (2018) for comparison) but due to the soil structure and boundary conditions, flow is only vertical. This is similar to test case 2, which is 3D but only has 2D flow.

- l 285 : "Since there is no variability along the 8000 m side, flow is effectively 2D in this test case." Then it is useless and misleading to present it as a 3D computation ; the figures and the discussions should be reshape for presenting directly the test case as a 2D one. The comparison of computation times is also questionnable : to deal with a 2D case in 3D increase tremendously (and artificially) the computation time with a fully mechanistic 3D model. Here some additionnal simulations (dealing with the 2D problem in 2D) are needed for making the comparison of computation times.

We disagree to this comment. Firstly, we do not claim that this is a 3D computation, the title of this subsection says clearly that it is a 2D problem. Secondly, we solve it in 3D, so giving 2D figures would be wrong. We agree that there is no real point in using a 3D model for this, but again (as in the previous comment) we would like to state that we did not design this test case (this was done by Morway et al., 2013 and Beegum et al., 2018) but only apply it to compare with already published results. Besides, the comparison of computation times is still valid as we use the same 3D setup for all models (the two coupled ones and Parflow). We want to stress here too that we do not intend to determine exact values for the computation time, which will anyhow strongly depend on the computational resources, but are only interested in the comparison between the models.

- l 290 : "[...] assigning a minimal initial pressure head of −1.25 m" ; you mean that -1.25 m is the pressure head at the top of the domain ? Please clarify.

We mean that the pressure head cannot have a value lower than -1.25m. Our formulation is not clear here. We will change it to "applying hydrostatic equilibrium and assigning a minimal initial pressure head of -1.25m at locations where the pressure head is below this value."

l 291 : "Monthly varying rainfall (Fig. 6) is used as Neumann boundary condition for the land surface". More precision about these data would be useful – e.g.: are they synthetic ? Of which type of climate are they representative ?

The data is taken from Morway et al. (2013) and information about the origin of the data is not given in their publication. We assume that it is purely synthetic, but cannot be sure and therefore do not want to give any information here that may be wrong in the end.

- l 292 : Table 2 is not timely introduced ; since it contains information for the 3 test cases, it should be placed either in the beginning or at the end of the presentation of the considered test cases, but not at the middle.

We agree, and will change this.

- l 293 – 294 : "grid size $\Delta x = \Delta y = 100$ m and $\Delta z = 0.1$ m.". It makes a form factor of 10^3 ... Any convergence study done for the mesh refinement?

Not yet, but as stated in our reply to the general comments, we will perform one. In our opinion the form factor is not so unrealistic, as the 1D models need a much finer grid for convergence, but the convergence study will clarify this point.

- l 296 : "With the flow problem being 2D this means that the entire domain is actually covered by 1D models." Nevertheless as far as I understood the proposed methodologies it would be exactly the same if the case was a 3D one? And I don't understand to which extent a 1D approximation for a 2D problem would be essentially more "acute" than a 1D approximation for a 3D problem?

For a 3D case, we would need 80x40 1D models to cover the entire domain, as conditions would vary along the 8000m side. The word "acute" is a typo and should be "actually", we will correct that.

- l 306-307 : "three different soil units are distributed throughout the domain as depicted in Fig. 7." More information is needed here. Is this distribution synthetic? How has it been acquired / built ? Of which type of soil (sand, loam, clay ...) each unit is representative ?

The distribution is synthetic. The units 1 to 3 are representative for loam, loamy sand and sandy clay. We will add this information.

- l 308 : "averaged arithmetically" Any tests for the use of harmonic or geometric mean instead of arithmetic mean?

No, from our point of view such tests are not needed. As we only have lateral flow in the groundwater domain, we can consider this problem as a horizontal flow through a vertically layered (stratified) medium. For such a setting, the arithmetic average is the adequate choice for all effective parameters.

- l 309 : " In the vertical direction a non-uniform grid is used with smaller grid sizes close to the surface and a total of 50 cells." Please provide the minimum and maximum sizes.

The grid size in the vertical direction ranges between and 0.003m and 1.2m. We will add this information.

- l 315 : "The 1D models are placed at the center of each zone." How are laterally averaged the porous medium properties in each 1D models covering 10*8 cells laterally?

They are not averaged but taken at the specific location. We did not mention this and will add this information.

- l 327 : "The residual saturation $S_r = \theta_r / \theta_s$ and the specific storage S_s are excluded from the analysis and set to 0.01 and 0.0015, respectively." Why have they been excluded ? To be justified, or at least discussed.

We will complement this sentence by "…, as a previous smaller sensitivity analysis had shown no impact of these parameters (not shown)."

- Table 3 : The parameters K_{GW} and K_{UZ} are appearing in the manuscript for the first time in this table. The notations used in table 3 and those used in the equations (especially (1) and (2)) should be the same, of at least explicitly related.

We agree. As we already mentioned in our reply to a previous comment, we will introduce these parameters properly at their first appearance.

- l 329 -330 : "The horizontal spatial resolution is again $\Delta x = \Delta y = 10$ m, whereas the vertical resolution is $\Delta z = 0.1$ m as in the 2D flow case." Once again a convergence study must have been done to justify the use of this mesh with a form factor of 10_2.

We do not see a problem with the form factor, as the 2d (x-y) groundwater flow and the 1D (z) vadose zone flow are decoupled. However, we will perform a convergence study (see reply to comment on l. 293-294).

- l 332 : "The time step sizes are the same as in the previously described test cases." Any convergence study for justifying the use of the proposed time stepping policy?

Not yet, but we will perform a convergence study.

- l 345 : "A visual comparison indicates that the coupling applied by Beegum et al. (2018) yields a comparable accuracy." Why not plotting the results of Beegum in Figure 8?

We agree, but we did not plot the data because we do not have them. We could put their figure next to ours if this helps, but we would have to clarify copyright issues for this.

- l 356 Table 4: This table contains information for all test cases and then it is not at the right place, being presented in a part specific to test case 1. Besides, since in test case 1 there is no lateral flux and thus no iteration in the iterative methods, I wonder why the iterative method has a wall time twice time more long than non-iterative method, while this later one include an addtionel step of remeshing? To be discussed.

Regarding the placing of the table, we will move the table to an appropriate place. The reason why the iterative model needs more time is because the grid is larger (including the saturated zone). We will add the sentence "This is due to the larger grid of the iterative model which also includes the saturated part." at the end of this paragraph for explanation.

- l 363-364 : "The results by Beegum et al. (2018) have a similar accuracy and shape as the results of the iterative coupling approach." Why not plotting them in Figure 9?

See reply to the previous comment: Because we do not have the data. We could put their figure next to ours if this helps, but first would need to clarify the copyright issues.

- l 365 : "When considering the non-iterative model, it is notable that initial time steps are an issue [...]" Any convergence study on time step ? What happens if smaller time steps are used?

We have extensively tested the case with smaller time steps, but the general behavior still remains, and does not seem to be an issue with time step size. This information is clearly missing in the original submission and will be added to the revised manuscript.

l 366-368 : "Both of these issues may be related to the reference model essentially acting as a bucket without any plausible steady state solution (i.e. steady state for the groundwater model would have groundwater tables far above the top of the domain)." Then why not chosing lower values of precipitation , so that a steady state may be reached?

We use this test case to compare to the results of Beegum et al. (2018), so it would not make sense to change the precipitation.

- l 374 : "All values [of $S_{y\,0}$] are smaller than the proposed value of 0.28, although the difference is less than 0.03." How the proposed value of 0.28 has been choosen? Are there correlations between the S_y values and the state of the groundwaters (e.g.: water table altitude, lateral fluxes intensity)?

The value of 0.28 was taken from Beegum et al. (2018), who do not justify their choice. We assume that they fitted it beforehand. And yes, there are correlations. There is a linear relationship between the water table position and the Sy values in this case (so Sy values increase with increasing water table position). We will add the figure, but we also want to point out that this relationship is probably case specific and should not be overrated.

- l 379 : "Both coupling schemes show a good agreement with the fully integrated 3D model." It is hard to understand why the matching between the fully 3D computation and the 2.5D ones is better here for this 3D heterogeneous test case than in the 2D homogeneous test case 2. I noted that in test case 3 a finer mesh is used than in test case 2. May be that convergence issues are at stake?

This is explained in l. 452-459 (It is caused by the boundary condition). We will refer to those lines here, so that the reader knows an explanation will come later in the discussion part. We will also do the convergence analysis to exclude possible effects of the grid resolution.

- l 389-391 : "Areas with larger differences appear at similar locations for both coupling schemes showing the largest deviations of up to ΔH GW = 1.5 m along the y = 800 m boundary". Why are they such discrepancies, and why there? These points should be discussed here.

We thank the reviewer for pointing us to this part. Doing the analysis, we noticed that we made a mistake during the postprocessing of the data. The larger discrepancies occur at the lower Dirichlet boundary (so along x=400m). These differences are then again due to the low groundwater table and strong gradient of the water tables along that boundary (as in the second test case). We will of course correct the figures and the corresponding discussion part.

- l 394-396 : "Overall, the specific yield values are decreasing when the groundwater table is rising and increasing when the groundwater table is falling (roughly between t = 1100 d and t = 2200 d)." Once more, a careful discussion of the physical meaning of S_y and its variation is needed.

The question was raised also by the reviewer 3 and we copy here the answer to his / her comment:
One cannot really say at this point what is expected because the heterogeneities have a strong impact which is hard to evaluate. We see a different behavior for the homogeneous second test case. More tests, maybe starting with a less complex layered soil structure, would be needed to be able to make a solid statement on the specific yield values.

- l 414 : "Note that the specific yield in the iteratively coupled model is not the value used for the non-iteratively coupled model defined in Table 3 but the value calculated by the model during the simulation." I don't understand how it is possible to make a sensitivity analysis on a parameter that is not constant and specified prior to computation, but timevariable, calculated along computation?

The activity scores are calculated individually for each time step, therefore a time variable parameter as input is no problem. We hope that this will become clearer if we explain the method of activity scores better in the second section.

- l 417 : "Acitivity"

Yes, this is a typo, and we will correct it.

- l 421-422 : "When looking at Eqs. 1 and 7, one sees that S_y can be eliminated which explains why there is no influence of S_y under these conditions." You mean that dh/dt = 0 at extremas ? To be clarified.

We will change the sentence to "When looking at Eqs. 1 and 7, one sees that S_y can be eliminated when lateral fluxes are negligible compared to the recharge fluxes, which explains why there is no influence of S_y under these conditions." for clarification.

- l 431-432 : "The average S_y value shows some smaller fluctuations, but overall it converges to a value around $S_y = 0.17$, which is a plausible value.". This is a too short discussion of the value of this key parameter that controls the exchanges between the saturated zone and the unsaturated zone in the iterative method. It should be interpreted physically. It seems to potentially encompasses a non clearly identified list of physical phenomena.

We do not really understand the need for a discussion at this point. It is already interpreted physically and a further discussion of the value is not possible.

- l 436-437 : "This means that the specific yield is mainly depending on the unsaturated zone parameters. This is reasonable as its intention is to represent the missing unsaturated zone in the groundwater model." Somewhat strange. According to the basic derivation of the diffusivity equation for unconfined aquifers, the specific yield should be a property of the saturated zone (drainage porosity). So may be that if this parameters depends mainly on the properties of the unsaturated zone, it means that it is not, or not only, a specific yield (see the first general comment)?

As outlined in our reply to the general comment, we disagree to this interpretation of the specific yield. We would argue that the specific yield is not a property of the saturated zone but a property of the unsaturated zone which is used in the saturated zone model. We refer to the reply to the general comment.

- l 442 : "in the case of the iterative model even consistent." I am not sure of what you want to say, please be more specific.

We will change the sentence to "and in the iterative model the model compartments are even consistent."

- l 448 : "On the contrary, using more models could help decreasing the discrepancies in the less accurate areas close to the no-flow boundary at y = 800 m which are most likely caused by the soil heterogeneities and the simplified recharge and specific yield pattern due to the zonation." These discrepancies are important (~1,5m), and their causes must be carefully assessed. Additional numerical experiment with lower and stronger soil heterogeneities or various zonation startegies could help to ensure that the proposed diagnostic is correct. From my point of view stating that "As this is a general issue for these kind of models and does not relate to the presented coupling strategies themselves, we do not investigate it further." is not sufficient, at least without any bibliographical references as it is at present.

Since we now found that the discrepancies do not occur at that location, but at the lower Dirichlet boundary (which is explained at the comment on l. 389-391), this discussion part is not needed any more and we will simply remove it. Otherwise, we would have agreed that further investigation would be necessary here.

- l 458-459 : "Therefore the results of the coupled model are on average more accurate even though this test case is more complex than the 2D flow case." Meshes also are different, and without proper convergence studies the impact of this point may not be assessed. The convergence studies must be done, and used for consolidating the discussions.

We will do the convergence analysis to exclude that these differences are due to differences in the mesh.

- l 464 : "is constantly $h_{UZ} = -1.25$ m at ≥ 1.25 m above the groundwater table" This should be made clear sooner (see teh comment on l 290).

Yes, we agree and refer to our reply to the comment on l 290, too.

- l 467 : "Which parameter is dominating depends on the current flow conditions." This should be discussed in more details.

This is discussed in detail in l. 417-427. We will change this sentence to "Which parameter is dominating depends on the current flow conditions as described in Section 4.4."

- l 470 : "comformably" is not specific/quantitative enough.

We will change to "consistently".

- l 478-480 : "This is not the case in this model as we cannot calculate this effect properly and we therefore keep ΔH due to recharge fixed (see Eq. 7)." However in equations (7), (8) and (10), it is clear that there is an iterative procedure that involves $\Delta H_{UZ\ \nu}$ and $\Delta H_{GW\ \nu}$ that evolve at each iteration ν? I don't understand.

This comment is answered by our reply to the comment on l. 221-224. ΔH_{UZ} and ΔH_{GW} are changed during the iteration, but ΔH_{UZ} consists basically of two components: A change caused by lateral fluxes and a change caused by the recharge. The latter is what we refer to here and this is fixed. The fluctuations due to lateral fluxes change during the iteration and thus also the total water table fluctuation. We will try to explain this better.

- l 481 : "In the end, the specific yield is not a physical quantity but a model parameter.". This statement seems too general ; while it is clearly the case in the proposed modeling approach, it is not the case in all formulation of the diffusivity equation in unconfined aquifers. Overal all this paragraph should be rewritten to better discuss the meaning of the concepts that are specific to the proposed methodology with a wording that should not rise ambiguities between these concepts and previously existing concepts. For instance:

We agree that the distinction between general statements and those that only concern our model is not done well here. The part was also commented on by reviewer 1. We will reformulate this paragraph and change it to:

"In the end, the specific yield is not a physical quantity here but a model parameter as it needs to compensate the influence of the lateral fluxes on the recharge, which cannot be quantified properly in this model, as well. Therefore, it should be treated as such and fitted for each application. This is a general problem that does not only concern this model, as there is no good way of determining this parameter properly in advance. Calculating it within the iteration substitutes its calibration and is therefore an advantage over other methods. The sensitivity analysis for the parameters' influence on the calculated specific yield shows that it behaves reasonably. It depends on the parameters of the unsaturated zone models, especially on porosity $\boldsymbol{\phi}$ and the saturated hydraulic conductivity K_{UZ}. The aim of the specific yield is to represent the missing unsaturated zone in the groundwater model, therefore a strong dependency on the unsaturated zone model's parameters is plausible."

- l 485-486 : “The aim of the specific yield is to represent the missing unsaturated zone in the groundwater model” You are talking about what you called a specific yield in your model. I think that it should have another name that ‘specific yield’, this latter word designing a concept that do have physical meaning and that is related to the properties (drainage porosity) within the saturated zone in the basic form of the diffusivity equation for unconfined aquifers (see the first general comment).

We disagree to this interpretation and would like to refer to the reply to first general comment.

REFERENCES

Crosbie, R. S., Binning, P., and Kalma, J. D.: A time series approach to inferring groundwater recharge using the water table fluctuation method, Water Resources Research, 41, 2005

Fahs, M., Younes, A., and Lehmann, F.: An easy and efficient combination of the Mixed Finite Element Method and the Method of Lines for the resolution of Richards’ Equation. Environmental Modelling & Software 24.9: 1122-1126, 2009

Kavetski, D., Binning, P., and Sloan, S. W.: Adaptive time stepping and error control in a mass conservative numerical solution of the mixed form of Richards equation. Advances in Water Resources 24.6: 595-605, 2001

Kollet, S. J., and Maxwell, R.M.: Integrated surface–groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. Advances in Water Resources 29.7: 945-958, 2006

Pikul, M. F., Street, R. L., and Remson, I.: A numerical model based on coupled one-dimensional Richards and Boussinesq equations, Water Resources Research, 10, 295–302, 1974