Response to Reviewer 1

General Comments

Here the authors explore different model complexities and configurations to highlight the need for coupling upland-wetland interactions in land surface models to better capture downstream hydrologic fluxes. In general, the inclusion of wetlands in LSMs is an important consideration for many types of landscapes. Overall, the manuscript was clear and the approach was sound. Below are some points that I think the authors should consider to help clarify some points for the reader.

Dear Reviewer, we would like to thank you for your constructive comments on our manuscript. In this file we respond to each comment to address your concerns. The response for each point is written in blue.

One broad comment is that there is a bit of a disconnect in how the manuscript was framed (the title, abstract, introduction had a heavier emphasis on wetlands in general) vs the actual analysis and discussion (much larger emphasis on fens, which is the study site). The conclusions were logical based on the data from the fen, but it would be helpful to a) emphasize that fens by definition receive significant amounts of groundwater inputs and b) how other types of wetlands with different levels of GW connectivity could change your conclusions.

We agree that the framing of the manuscript can be improved. We focus mainly on groundwater dominated fen wetlands, since our case study is a fen, but we also conduct numerical experiments that look at non-groundwater dominated wetlands (section 5.4.2). We will edit the abstract, introduction and objectives to better explain this.

Specific Comments

- It would be helpful to reference more existing literature on the need for including proper interaction between wetlands and the upland to help strengthen the case for this particular study. While there may not be as many LSM studies that look at this directly, drawing parallels to watershed scale studies, which has quite a few studies in the recent years.

We agree that we can include more literature on watershed scale applications of LSS and will seek to do this in the revised manuscript.

- Minor suggestions for Figure 1
  - Expand part d to be larger to match the other panels for clarity;
  - add the location of POJP piezometer for clearer connection to panel c);
  - if possible, indicate the extent of the fen in part d) just to give some context for the reader. I understand that this can be variable throughout the fen. If the fen goes beyond the transect, ignore this comment!

We will include these points in the revised manuscript.

- L123: The statement regarding the amount of surface water vs groundwater into the fen seems something unique to this system, and isn't necessarily a feature of the V2 configuration –

consider putting this elsewhere. As a side note, because I don't have much context to the wetland:upland ratio or the water balance, the reader might be surprised by this statement. Might be worth indicating the relative areas for the wetland + upland and/or some estimated water balance in the site description

We agree that as the relative magnitude of the surface water flux is not a general feature of the V2 model configuration, so we will move this text from this part. Regarding the wetland-upland ratio, the total study domain is about 3300 m and the wetland width represent about 150 m of that length. We will focus on that point more in the study description section to make it clearer to the reader.

- L137: does the MESH-CLASS model have saturation-excess and infiltration-excess runoff components to determine runoff vs recharge? Similarly, how does it calculate the R flux? I would not expect a full description of the model, but since this is a major connection to the GW component, it would be helpful to briefly describe it. Also, is snow accumulation and melt modelled in the upland? I assume so, but since it's explicitly mentioned in the fen model, but not here, it can create confusion

The MESH/CLASS model uses an infiltration excess method (based on Green and Ampt) to calculate infiltration. Excess ponded water at the top of soil column is used to calculate overland runoff. We will mention this in the model description.

Regarding, snow accumulation and melt at the upland, they are simulated by the MESH model, and we will mention this in the revised manuscript. However, note that these fluxes are not used directly in our upland-wetland model – rather we use runoff and drainage generated from snowmelt by the MESH/CLASS model. For the fen the snowmelt fluxes are used directly in the fen water balance model.

- L148: I may be mistaken, but I don't think Figure 1 show the groundwater divide/no-flow boundary condition (this is in Fig 2?)

We assume the groundwater divide is the basin boundaries shown in Figure 1-b as black dashed line. We will further clarify that for more clearness in the revised manuscript.

- L151 and paragraph: would be helpful to mention how dx is determined within the Darcy's Law calculation

On Line 153 we added the following: "Equation 1 is solved numerically using a block-centered finite difference solution on a regularly spaced grid with dx = 0.01*L (L=Hillslope length).

- L165: Units are not consistent across equation. If Ro is upland runoff (m3/d) and L is the hillslope length (m), the units are [L4/T]; similarly the (Rg+M-Ef)wf component have units of [L2/T], while the Q terms are [L3/T].

Thanks for pointing out this mistake. The units of RO should be m/d and for Q_G and $Qout$ its m^2/d, knowing that we assume a unit area (1m) in y direction. That will be corrected in the revised manuscript.

- L168: Are Cspill, hspill, n calibrated parameters, and is the model sensitive to them? I believe they are not referenced again later on but one might assume that they have major roles in changing Q.

We agree we should provide the parameter values used, and explain how these values were selected, and will do this in the revised manuscript. These three parameters control the timing and magnitude of discharge from the fen into the river channel. We did not have streamflow data to calibrate these parameters, and therefore our approach was to use sensible fixed parameter values that enable us to perform controlled numerical experiments. Our objective is not to simulate streamflow precisely, but rather to explore the sensitive in simulations to changing coupling between the wetland and the upland.

h_spill represents the spilling threshold of the fen, so that when the fen water level is below h_spill, there is not discharge from the fen into the river. We set this value equal to the elevation of the fixed head boundary from the uncoupled groundwater model (Figure 2b/c).

The values of c_{spill} and n (c_{spill} = 0.1 and n = 1.5) were arbitrarily chosen within the reasonable ranges. The ranges were defined based on the recommended ranges in (Razavi and Gupta, 2019), which they used in modelling the fast reservoir with non-linear response in the HBV-SASK model.

We will make this clear in the revised manuscript.

- L181: Is there any downstream gage to calibrate? I would be hesitant to say that calibrating the GWT in the upland represents the performance of the collective fen-upland-GW models - especially since three of them do not have the backwards interaction in the model structure

We agree that calibrating the GWT is not sufficient to conclude that the model accurately represents our specific field site – however, the main objective in this study is to perform a comparison between the alternative modeling approaches with using the one reasonable parameter set (as we explained in the previous response).

- L185: were these all generated from uniform distributions?

Yes, the 15,000 realizations were generated using a uniform distribution. That will be added to the text.

- L186: maybe use "(threshold is chosen arbitrarily based on…)" instead of "chosen rather arbitrary"

This threshold value was selected based on identifiability analysis and considering the behavioural realizations, which is not included in the manuscript since it is out of scope.

- L189: I would consider putting in the best parameter set in Table 1 to give context for readers rather than just embedded in the text later on

Agreed - the table will be modified in the revised manuscript to include the calibrated parameter set.

- L194: are these L values corresponding to likely wetland-upland areas?

These values correspond to the length of the upland hillslope. At this case we explore the effect of changing the area of upland with having the same area for the fen.

- Figure 3: Should write in caption the simulation number (V1?)

Yes, these results are for model version V1. That will be added.

- Figure 6d: missing y axis label to be consistent

That should be the same label as in Figure 6c. However, that will be added for clarity.

- L291: It is slightly hard to understand why at low L values, the groundwater table does not fluctuate in the chained model – it would be worth discussing why. I would assume there's still stochastic inputs to the groundwater from recharge/precipitation, and it's not that the groundwater table has reached the lower boundary/bedrock

The water table does not fluctuate very much in the chained model when the hillslope length is small because of the fixed head boundary condition. In the coupled model, there is no fixed head boundary, so the water table fluctuates more.

- L306: Land is capitalized mid-sentence

That will be modified.

- L306: Because the reader does not know how the forest and grass (is it solely a runoff-coefficient difference, or does ET get affected too?) affects the model fluxes, it's hard to attribute the changes in the model to solely the soil properties, which is the focus of this section.

In our model, changing the vegetation impacts ET and runoff, while changing the soil properties changes the runoff vs infiltration, recharge, and ET. The first model configuration in this numerical experiment was our original setup (OJP in the southern boreal forest). The second configuration (grass over finer soil) was designed to maximize surface runoff, to provide a contrasting case with the groundwater dominated case. The configuration was based on a model configuration for a grassland site (St Denis in Ireson et al., 2022 - MESH point scale paper, in the prairies south of the forest). We did this because we have credible parameters for each configuration – they are not entirely hypothetical. We will explain this rationale more clearly in the manuscript and provide the citation.

- L313: I think that for this instance, it is true that the chained approach is adequate to illustrate the coarse grained soil texture. But I think it's worth commenting that in areas of smaller hillslopes/contributing areas, that may not be the case (as proven in your previous experiment)

We agree and will mention this at this point in the revised paper.

L346: I would not necessarily include wetlands in 'fen/wetlands' as fens by definition have a lot of GW inputs. Having wetlands here can cause readers to assume that wetlands that either have more bi-directional interactions with the upland via groundwater, or don't receive groundwater, should be treated the same way. While the authors wouldn't run more simulations to capture other wetland types, I think it's a valuable discussion point

We agree with you, this should be clearly discussed in the conclusions to prevent any confusion. We would say that this is an investigation and fens are taken as an example of wetlands that have more dependency on two-way groundwater exchange.

References

Ireson, A. M., Sanchez-Rodriguez, I., Basnet, S., Brauner, H., Bobenic, T., Brannen, R., Elrashidy, M., Braaten, M., Amankwah, S. K., and Barr, A.: Using observed soil moisture to constrain the uncertainty of simulated hydrological fluxes, Hydrological Processes, 36, e14465, https://doi.org/10.1002/hyp.14465, 2022.

Razavi, S. and Gupta, H. V.: A multi-method Generalized Global Sensitivity Matrix approach to accounting for the dynamical nature of earth and environmental systems models, Environmental Modelling & Software, 114, 1–11, https://doi.org/10.1016/j.envsoft.2018.12.002, 2019.