

Referee #2

I think this is interested study that is worthy of publication. However, a lot of improvements have to be made for its current form. I agree with that ground water dynamics may be a good proxy of surface hydrological processes in some places. However, this may not be the truth in some other areas. So this proposed approach may have its limits. This have to be clarified in the introduction and discussion.

We thank the reviewer for their comments and feedback. We have made significant changes to the manuscript. Please see below for our detailed responses to the comments. The reviewer's original comments are in black, our responses are in blue italic fonts, and modifications made to the original manuscript in blue plain text.

We have clarified the limitations of using groundwater dynamics as they may not be suitable in other regions with perched systems for example. We have added the following statement to the introduction "Nevertheless, we acknowledge that groundwater dynamics in some regions such as arid areas could be disconnected to land surface processes and less dependent on many key physical features of the hillslope, which may impede the ability of the proposed classification in these regions."

There are a lot of indices and methods used in this study were not presented in the Methodology section. I also feel that the methods section did not clearly present how the authors process the data and generate the results.

The revised manuscript has a "methods" section with the definition of all the indices as well as all the methods and data processing methods we used.

In addition, and especially, discussion of the findings of this study has to be strengthen. currently, the discussion is weak, maybe due to the reason that the results and discussion were combined. References are needed for the interpretations. Explanation of the results and comparions with other published studies have to be improved.

The revised manuscript has a discussion section in which we discussed and compared the findings of this study to previous works.

1. the abstract lacks quantified description

We have added statistics to the abstract. We, specifically, added these lines:

"The ΔP classification performs very well in identifying hillslopes with 6 out of the 9 characteristics studied. The variability among similar hillslopes as quantified by the coefficient of variation is less (0.2) in the ΔP and clustering classification than in the others (>0.3 for TWI, elevation, and land cover)."

2. line 41-44, references are needed to support this statement.

We have added the following references: (McDonnell & Woods, 2004).

3. line 106, maybe give some examples of such models

We have named these models, the sentence is now “These models (e.g., HydroGeoSphere (Brunner and Simmons, 2012), ParFlow (Maxwell & Miller, 2005), Advanced Terrestrial Simulator, (Coon et al., 2016)) that can be constrained with ground observations and measurements at ultra-high resolutions through aerial or remote sensing (i.e., drones, planes, or satellites) account for the two-way interactions between groundwater and land surface processes”.

4. line 138, maybe starting with a sentence to tell the reads the pupose or the reason of using ParFlow-CLM in this study

We have added the reason for using ParFlow-CLM in this study. Please refer to the sentence below: “We use the integrated hydrologic model, ParFlow, which has the advantages of simulating the water and energy balance from the bedrock to the lower atmosphere and therefore connect groundwater dynamics with land surface processes.”

5. line 164, provide the examples of the application of ParFlow-CLM

We have now added examples of ParFlow-CLM applications in the following sentence.

“ParFlow-CLM has been used in many studies to understand the interactions between groundwater dynamics and lower atmosphere (Maina et al., 2022; Maina and Siirila-Woodburn, 2020) at different scales from the watershed (Foster and Maxwell, 2019; Maina et al., 2020) to the continental scale (Maxwell and Condon, 2016).”

6. line 239-240, what were those thresholds tested, specify

We have tested these thresholds, we have now clarified it in the revised manuscript by adding the following sentence: “For the ΔP_1 , elevation, TWI, and AI classifications, we define the thresholds of each zone by analyzing the distributions of the hillslope values of these indices.”

7. line 238-243, what were the thredholds of drainage area you finally used?

We used a threshold for drainage area equal to 810,000 m². We tested different thresholds to select this one, more details could be found in Wainwright et al (2022). We have added this information to the revised manuscript.

*Wainwright, H. M., Uhlemann, S., Franklin, M., Falco, N., Bouskill, N. J., Newcomer, M. E., ... & Hubbard, S. S. (2022). Watershed zonation through hillslope clustering for tractably quantifying above- and below-ground watershed heterogeneity and functions. *Hydrology and Earth System Sciences*, 26(2), 429-444.*

8. line 316, the clustering approaches have to be introduced in the methods section

We have now described the clustering approaches in the methods sections.

9. line 369-379, most of the part would be better to move to the methods section

We have moved this paragraph to the methods section

10. section 3.2.3, why were surface runoff not considered? I thinks it might be one of the most important hydrological processes.

We have not added surface runoff because not all hillslopes have a quantifiable surface runoff.