The topic and the study is interesting and thus worthy publication. But before this significant improvements is necessary.

We want to thank the reviewer for reviewing our manuscript, the reviewer's comments have provided us with valuable insights and suggestions that will improve the quality of this manuscript for future readers. Please note that the line numbers we mention refer to the marked up version of the manuscript.

1. **be strict in the structure. There are elements in the method that is results and in results that are conclusion.**

Here we respond to this general statement by the reviewer together with the specific related suggestions the reviewer makes (i.e., line164-68; 177-179; 189-191).

The reviewer suggest that the text regarding the initial correlation analysis (164-168), the tuning of the model (177-179) and the test between temporal and spatial sub setting when training the model (line 189-191) as something that is better suited for the results section. As authors we find this a difficult distinction as well, as it is right on the edge between the methods and the initial set of results. However all of these instances in question refer to the construction of the machine learning model and we thus think it is better suited for the methods section. This is also very much in line with what other peer-reviewed papers have done and seems to be common practice in the community. For your convenience we provide some examples of other papers that use ML and describe the construction of their models in the methods section.

1- Moradi E, Darabi H, Alamdarloo EH, Karimi M, Kløve B. (2023). Vegetation to stresses in water-scarce areas using machine learning and remote sensing. Ecological Informatics 73:101838
2- Zhang Y, Mao J, Ricciuto DM, Jin M, Yu Y, Shi X, Wullschleger S, Tang R, Liu J. 2023. Global fire modelling and control attributions based on the ensemble machine learning and satellite observations. Science of Remote Sensing 7: 100088.
3- Carranza C, Nolet C, Pezij , Ploeg M. 2021. Root zone soil moisture estimation with Random Forest. Journal of Hydrology 593:125840
4- Farhadi H, Najafzadeh M. 2021. Flood Risk Mapping by Remote Sensing Data and Random Forest Technique. Water 13:3115.
5- Islam KI, Elias E, Carroll KC, Brown C. 2021. Exploring Random Forest Machine Learning and Remote Sensing Data for Streamflow Prediction: An Alternative Approach to a Process-Based Hydrologic Modeling in a Snowmelt-Driven Watershed. Remote Sensing 15: 3999.

Regarding the comment on the results being conclusions. We assume the reviewer meant that there are discussions that are meant to be conclusions, since we only see comments relating to the latter (Line 266, 273-277, 291-292, 294, 299-300). We respond to this issue here in the general sense and expand on the specifics when mentioned below. We agree with the reviewer that these sections are phrased in a way that is better fit for a discussion/conclusion and have rephrased these sections to fit that of a discussion and move some of these points down to the conclusion where suited.

2. Discussion need to reflect on the uncertainties in the results and how uncertainties and methods influence the results. Slope and aspect is probably important for soil wetness, but how can slope on 0.1*0.1 deg make sense or be relevant? Such discussions is not present but relevant.

We agree with the reviewer that the terrain slope is important for soil wetness as it is a major driver of hydrological gradients. When considering only soil wetness in relation to slope the question arises whether it will be relevant at the 0.1˚ resolution. Yet one could argue that in regions with steep topographical gradients the soils dry out faster. On the 0.1˚ resolution even a small difference in slope (Figure Below) would distinguish between flat and hilly terrains. The rationale of including these at 0.1˚ resolution was in the hope that the model could derive some meaningful relationships between slope and vegetation status; for example, the response time of soil moisture during long dry periods or the vegetation composition. We expect there to be a difference in the species composition of plants present in regions that differ in slope (which would in turn affect *evi* responses). And by including this data in the 0.1˚ and 0.01˚ predictions we had hoped that this would result in better predictions, compared to a model where this is not included. But as the reviewer points out such discussion is necessary and we have included this in the discussion section, please see lines 405 – 408. Aspect was not

included as one of the predictors given its poor correlation with *evi*. This is present in line 200-201.
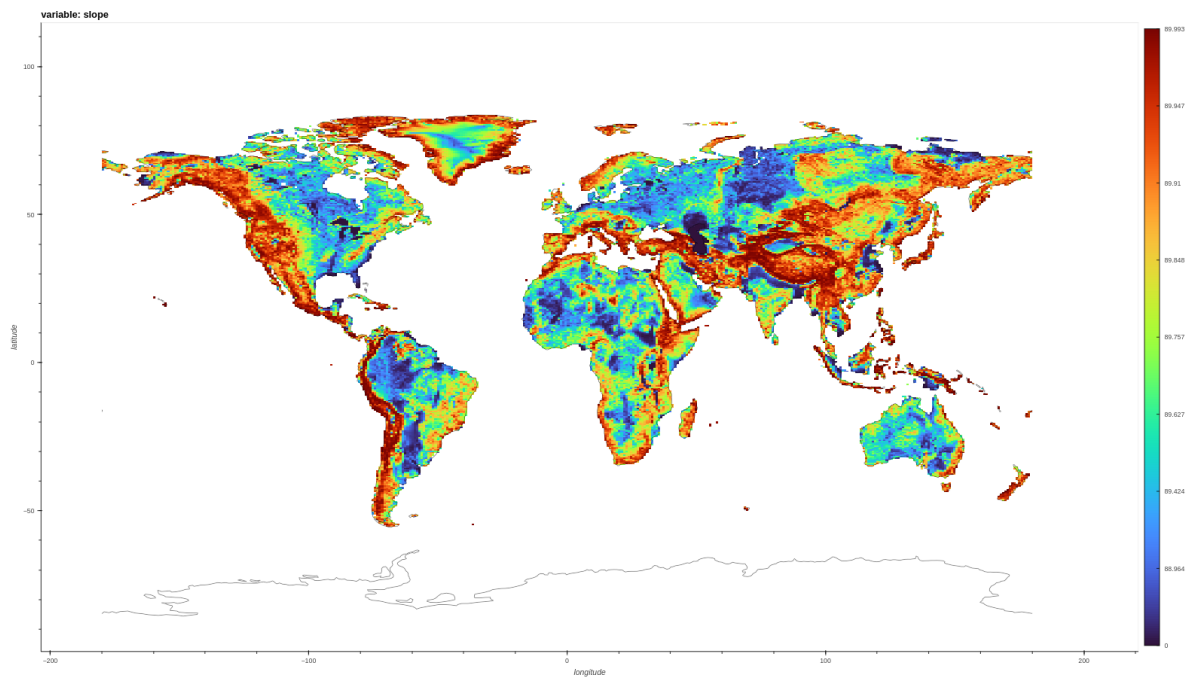
Figure 1: Slope at the 0.1 degree resolution.

3. Uncertainties in met input was mentioned in prior feedback but only reflected on in one sentence. Temp is by far the most important and both in the results section it could be elaborated more on. How does this influence the geographical distribution. And then the comment about uncertainty becomes more important. But what about the other parameters uncertainty?

We have relied on the original evaluation in the ERA5-land publication which validated land surface temperature to further expand our discussion on how uncertainty in the input date might affect our results. In addition, we have sourced an additional study (Lal et al. 2022) that allows us to compare errors in ERA5-Land soil moisture estimates to the geographical distribution of the errors presented in our model. Using these two variables we have been able to match regions of high error in our predictions with regions of high error in the ERA5-Land data. We have expanded on this in the discussion (line 409-416).

Lal P, Singh G, Das NN, Colliander A and Entekhabi D. 2022. Assessment of ERA5-Land Volumetric Soil Water Layer Product Using In Situ and SMAP Soil Moisture Observation. *IEEE Geoscience and Remote Sensing Letters*. 19:1-5.

4. Disaggragation and aggragation of data is not well explained. How is for instance temperature and precipitation scaled to 0.01 deg from Era 5 data? And also how is the GEEMAP functions scale the topographical parameter?

Here we respond in the general sense and include our responses to the specific mentions regarding Line 94-95, 121-122, 144-145, line 266. The current phrasing suggests that the data were downloaded and then independently scaled by the geemap package to the desired resolution. However, what we meant to say is that the data were downloaded using the geemap package, and then downloaded two different sets of data with different scales. The scaling happens at the google earth engine back end and directly relates to the geemap package. More specifically, the scaling relies on google earths engine image pyramiding policy (https://developers.google.com/earth-engine/guides/scale#:~:text=Image%20assets%20in%20Earth%20Engine,at%20the%20next%20lower%20level.). For *evi*, *hnd*, *aspect*, *elevation*, *aspect*, and *slope* the mean is used to aggregate higher resolution data to lower resolution data. Whereas the mode is used for landcover (https://developers.google.com/earth-engine/reference/rest/Shared.Types/PyramidingPolicy). This information has been added and is present on lines 177-183, and lines 140-143.

In addition, we do believe that the methods on how the meteorological data were regridded to the 0.01 degree resolution were already present in the two previous manuscripts. In its current form this information can be found on line 251 – 253.

Line 94 -95: rewrite this sentence…when data have different resolution I assume some aggregation or disaggregation is needed. Is this the case here… and if so how is then the disaggregated or values calculated? Sum? Average? Median? or? i.e. precip from Era5 is at 0.1 deg and how do you this to 0.01 deg… trend, average

This has been removed but we refer the reviewer to the response to general comment 4. We do believe that the methods on how the meteorological data were regridded to the 0.01 degree resolution were already present in the two previous manuscripts. In its current form this information can be found on line 251 – 253.

Line 94 -95: error here..

This sentence has been removed and we refer to the previous comment for the explanation.

Line 99-100: repeating previous more or less… try to be less repetitive

We removed this sentence according to the reviewers suggestion.

Table 1: What about cross correlation between parameters. I cannot see this mentioned. Does that have any influence? TP and Soil wetness must be correlated.

Inherently there are many variables in the hydrological and atmospheric system that are cross-correlated. As such the reviewer is correct to say that these two are cross correlated. However, we have retained this in the predictor set to account for localities where they are not cross-correlated and otherwise relevant information would be lost.

Like wetlands, humid forests and groundwater dependent ecosystems. We have included this in the manuscript, line 205 -208.

Table 1: I think you can add resolution to all lines to avoid misunderstanding

We thank the reviewer for this suggestion, which we implemented and can be found in the new version of this table in the updated manuscript.

For the readability I would suggest to include Description in the table. And what is level 1 meaning?

An additional column with a description of every variable was included in the new version of the table, including the information regarding 'level 1' which represents the depth of the soil layer represented.

Line 118: Layer level 1? What does that mean?

We have now added the depth in cm it refers too.

Line 118: What resolution is this data 0.1x0.1 deg I presume....https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=overview or?

Please see response to general comment 4, where we elaborate more on the data and spatial scales. Further information regarding the spatial scales used for the different variables is also present in Table 1.

Line 121-122: But how was they spatially disaggregated? From 0.1 deg to 0.01? How do you handle temperature?

We kindly refer the review to response to general comment 4, where this has been addressed in more detail.

Line 144-145: And how is GEEMAP doing this... averaging or... this can shortly be explained and then linked t reference.

We kindly refer the review to response to general comment 4, where this has been addressed in more detail and where we added the corresponding links for better understanding.

Line 164-168: this is a result and not method!

We kindly refer the reviewer to our response to the general comment number 1.

Line 177 – 179: this is also a result

In line with the previous comment, we kindly refer the reviewer to our response to the general comment number 1.

Line 189-191 and Figure 2: result

Here as well, we kindly refer the reviewer to our response to the general comment number 1, as this one falls in the same line.

Line 217: Is Ai,j in eq1 ACC value? You need to explain this if not? or use ACC in eq 1

We improved the description of equation 1 for more clarity.

Figure 3: There are no reference to this figure before the figure. This need to be! A RMSE map could be a third map to include here and that would give more information. Combine fig 3 and 5

We thank the reviewer for this suggestion but we think adding another RMSE map would take away the attention from the main results and overview as the figures would be too small to see any detail. The past and currents reviewer' s comments have helped we realized throughout the review process that the four figures are a bit redundant and therefore we have chosen to include only three of them and keep the RMSE results (Figure 5) separate. At the end of the day our goal is to show the broad global patterns here. We have also made sure that reference to figures precede the figures in the manuscript.

Line 230: Reference to figures should be before the figure
We have adapted the figure reference and figure position accordingly.

Figure 4: To increase readability the full explanation can be give on the figure i.e t2m - temp. at 2 m, swvl1 - soil water, lc - land cover.... you have the space - why not use it

We thank the reviewer for this suggestion and agree with it. The figure was adapted accordingly and we hope this helps the reader..

Line 233: Is it a goal to use as little of the data on training as possible? Explain why you say "only 6% of the data" and why you do not use more. Does it not improve when using more? Did the performance drop when you use less? Or?

This relates to the spatial vs temporal splitting of the data, we have added a reference to the relevant section in the text and expanded on this idea here, please see line 284-285.

Line 239: How can you distinguish between landcovers in fig 5. I agree it is a nice figure, but what does it tell? Not landcover influence without any other linked to

it. This need improvement... a figure should have a purpose beyond being nice. I think it have information but which?
The three representative regions were chosen for the different landcover types, this has been made more clear in the revised manuscript by mentioning it in the figure legend. The current figure layout and idea was also result of previous review rounds, where more detailed results were asked for to also see the difference in spatial resolution and the performance of the model.

Figure 5: Suggest to combine fig 3 and 5.

Please see previous comment directed at figure 3 above.

Line 245: the RMSE map give more info than visual comparison
We thank the reviewer for catching this type, which has been corrected in the updated manuscript.

Figure 6: I assume that the observed is the same in both a and b. Then print a and b in one graph to get the differences clearer presented.

We have updated the figure according to the reviewer's suggestions to make it more clear and easier to interpret.

Line 254: At least? a more precise formulation is needed when you compare to another number.

Has been replaced "more than". Please see line 311-312.

Lined 245: What is concidered a good ACC value... this is not described.
We have labeled all ACC values that are above 0.0 are able to capture anomalies that go beyond the seasonal cycle. We would argue that a ACC above 0.2 is deemed good given the strong seasonal cycle that is present making it hard to get really high ACC values. This has been added in the manuscript please see lines 263-265.

Line 255: How can "this" indicate... what is the basis for this statement.

We adapted this line for more clarification. See line 305 - 3011.

Line 259: This introduction to how the discussion is structured is not needed and not common to include.

We agree with the reviewer and this has been removed.

Line 266: This paragraph is not discussion but more a conclusion. A discussion should focus on how good and/or uncertain the findings are, what makes them relevant or reduce relevance etc and the findings should be compared to what other studies have found and why these are different, new or similar etc

Please see response to comment 2. This paragraph has been removed and we have made sure this information is present in the conclusion, please see line 440 – 449.

Line 267: Not completely true… Landcover was equally important as soilwater and tp was less important than hnd… It is correct to say temperature was the most important… but lc was more important than most of the rest.

This has been removed, and repositioned in the discussion. We have made it clear that it is landcover, elevation, energy and water. See lines 442 – 444.

Line 273-277: Highlighted red.

We kindly refer to general comment number 1. This has been deleted and the information has been included in the conclusion, see 455-463.

Line 278 – 281: Do this study confirm something already done or is it something new or what… or does it bring in something new? Be more clear and discuss

We have added clarifications which can be found in lines 335 – 347.

Line 291-292: This a conclusion and recommendation for further work

We kindly refer to general comment number 1. This sentence has been removed.

Line 294-295: This is more like a discussion should be focused. I have marked in red what is not discussion

As for the previous comment, we kindly refer to to general comment number 1 as we believe that falls in the same line.

Line 294: If you include thought about why then this is a part of the discussion… otherwise this is a conclusion.

Please see response to general comment number 1. We have related this to one of the known limitations of the RF model, which is an inability to reproduce extremes and keep it as a discussion. Please see lines 362-364.

Line 299 – 300: Highlighted red

Please see response to general comment number 1. This sentence has been removed.

Line 309: What do you mean with "natural weather processes"... this is not a scientific precise description. More interesting here is how the weather influence the result. This is not much discussed. And it should be! Temperature is the most important. How well do Era 5 reflect this and how does this ucertainty influence the results. Tp has weak influence and the spi and spei ... this seems strange as soil water has strong influence... these must be cross correlated but why they do not give similar influence is strange and need discussion.

We kindly refer to the response to general comment 3. We have also elaborated on what we meant by natural weather processes, which can be found in line 379.

Line 331: What about the other met parameters?

As with the comment before, we kindly refer to the response to general comment 3, where a more detailed answer regarding these parameters can be found.