Reply to Anonymous Referee #1

We thank the anonymous referee for his/her positive, thorough and constructive review. We provide an answer to each comment below.

Comment 1:

Study objectives are not clearly and consistently stated

According to the abstract, introduction and conclusion, the study has two main objectives. The first isto propose a top-down approach to include vegetation change into hydrological models via the root-zone storage capacity (l. 4-5, 575, 581-583). The second is the quantification of the sensitivity of modelled hydrology to changes in root-zone storage capacity under climate change and related to that, the testing of the hypothesis that changes will be more pronounced when considering an adapted root-zone storage (l. 93 ff).

Although these two objectives are clearly connected, they are never stated together. The first objective (proof-of-concept and methodological aspect) of the study is stressed in the discussion and conclusion, whereas the introduction highlights only the second objective (application and sensitivity analysis). The objectives of the study should be more clearly stated in the introduction and the discussion and conclusion should build on these objectives.

Reply 1:

We agree that these are two main aspects of the manuscript. However, in the revised version of the manuscript, we will more clearly state that our objective is to evaluate the sensitivity of hydrological model predictions to ecosystem adaptation in response to climate and potential land use change. To reach this objective, we introduce an approach, subject to assumptions, to estimate future evaporation and associated changes in the root-zone storage capacity. As the future is unknown, we cannot evaluate our results against observations. Therefore, we would rather not qualify the introduced methodology as an objective as the underlying hypothesis cannot be tested. In the revised version of the manuscript, we will not use the terminology "proof-of-concept". However, we fully agree with the suggestions of the reviewer to more thoroughly discuss the limitations and opportunities of the proposed methodology in the discussion (see our reply to the next comment).

Comment 2:

Discussion and conclusions leave open questions

The discussion could be more thorough and consistent regarding both, the modelling results and the methodological approach.

The discussion is structured into two separate parts: *Implications* (l. 500- 539) and *Limitations and knowledge gaps* (541-577).

However, two paragraphs from the first section (Implications) are better suited for the second section (limitations): l. 512-519 on possible further exploration of the space-for-time concept and l. 535-539 on the limitations of the simulated climate time series used in the study.

Also, given that one major objective of the study is to propose an approach to include vegetation change into hydrological models, I feel that the model results are not thoroughly discussed as to whether the proof-of concept of the method was successful.

The following questions/issues remain unaddressed:

a) The approach showed that the root-zone storage capacity parameter has a potentially large effect on future water flows. How realistic are the values for root-zone storage capacity that were calculated for the different scenarios? Is there any evidence from literature regarding the extent to which plants adapt their root system to changing climate? Does this adaptationdepend on vegetation type (e.g. crop/grass vs. tree) or species?

b) Are the results regarding the water flow under future conditions realistic? Is this what couldbe expected under climate change?
c) In which situations can this method be applied? Which hydrological models? Which ecosystems?
d) What are the limitations and chances of this approach?

__Reply 2:__

In the revised version of the manuscript, we will more thoroughly and consistently discuss the limitations and opportunities of the methodological approach and the modelling results.

We agree that the two paragraphs in the Implication section mentioned by the reviewer are more related to "Outlooks" of possible future work. Initially we had treated this as implications of our work for future work, but we agree that it may better fit in a section called: "Limitations and outlook". We will adapt this in the revised version of the manuscript.

As the future is unknown, we cannot evaluate our results against observations. Therefore, "proof-of-concept" may not be the right terminology, we will adapt this in the revised version of the manuscript. Our study should really be considered as a sensitivity analysis to test the sensitivity of hydrological model predictions to non-stationary systems through plausible assumptions of ecosystem adaptation. To emphasize this more strongly, we propose to adapt the title of the manuscript to: "The sensitivity of hydrological model predictions to ecosystem adaptation in response to climate change."

In the discussion of the revised manuscript, we will address the following discussion points:

a) The estimated values of the root-zone storage capacity for the different scenarios have median values below 250 mm for a return period of 20 years, which is within the range of global root-zone storage capacity values estimated by Wang-Erlandsson et al. (2016). We will mention this in the discussion of the revised manuscript.

There is increasing evidence that vegetation efficiently adapts to its (changing) environment (Gentine et al., 2012, Troch et al., 2013, Hrachowitz et al., 2020). Guswa (2008) shows that the active root zone tends to be larger in water-limited ecosystems in comparison to wet environments. A distinction should be made between individual plant adaptions of roots and the adaptation of the root system of the collective of plants at the ecosystem scale. The study of Brunner et al. (2015) describes several strategies of tree root to cope with drought, which include root biomass adjustments, anatomical alterations and physiological acclimations. Individual plants that have not adapted to meet their water and light requirements will disappear and be replaced by better adapted plants. Therefore, the root system at the ecosystem scale and associated root-zone storage capacity continuously adapt to changing environmental conditions in a state of dynamic equilibrium (Hrachowitz et al., 2020). While the adaptation of individual plants depends on vegetation type and species, here, we determine effective values of the root-zone storage capacity at the catchment scale to reflect the adaptation of the whole ecosystem.

b) Given the changes in temperature and precipitation, the future predicted hydrological response does not seem unrealistic, although, of course, this cannot be tested against observations. Common practice in hydrological studies on the impact of climate change is to assume a stationary system (Benchmark Scenario $2K_A$ in our analysis). In addition to this *scenario, we suggest a possible approach to consider ecosystem adaptation in response to climate change and test the sensitivity in the resulting hydrological response. Our approach is subject to considerable uncertainties in the estimation of the future transpiration (required to estimate the root-zone storage capacity) as we are using the Budyko framework for future conditions (Berghuijs et al., 2020; Reaver et al. 2021). Besides, we do not explicitly consider that vegetation can adapt to drier conditions by regulating their stomata and hence reducing transpiration (which is the topic of your comment 3). Moreover, the increased CO_2 concentration may, on the one hand, increase water use efficiency, while on the other hand increase green foliage due to fertilization effects (Donohue et al., 2013; Frank et al., 2015; Yang et al., 2019). Hence, we cannot predict what will exactly happen, but we can at least test the sensitivity of the hydrological response to changes in the system representation.*

c) *Root-zone storage capacity estimates derived from the water-balance approach are applicable in various hydrological and land surface models, provided that they include a root-zone parameterization, which is the case for most models (Nijzink et al. 2016, van Oorschot et al., 2021). The water-balance approach to estimate the root-zone storage capacity has successfully been applied in a variety of climate zones and across various ecosystems (New-Zealand in de Boer et al. 2016; Australia in Donohue et al., 2012, United States in Gentine et al. 2012 and Gao et al., 2014; and at the global scale in Wang-Erlandsson et al., 2016). The method was also applied along rainforest-savanna transitions to reveal drought-coping strategies (Singh et al., 2020). However, the method is not suitable in areas where the water table is very close to the surface and where vegetation directly can tap from the available groundwater instead of creating a buffer capacity (e.g. Fan et al. 2017). Another limitation of the water-balance approach relates to equation 6, in which we scale the daily transpiration estimates with a constant factor to the patterns of potential evaporation minus interception evaporation, implying that vegetation can extract water for transpiration from dry soils as easily as from wet soils.*

d) *The proposed methodology to estimate future root-zone storage capacities relies on the underlying assumption that past empirical relations between aridity index and evaporative index (i.e. the Budyko framework) still apply in the future. The Budyko framework reflects the long-term hydrological partitioning under dynamic equilibrium conditions. Therefore, when using the Budyko framework to estimate the future rate of transpiration, we assume that the future vegetation has adapted to the future climatic conditions and that it is in a state of dynamic equilibrium. This is a considerable uncertainty of our methodology because it implies that vegetation has had the time to adapt to the rapidly changing environmental conditions. There is no doubt that vegetation eventually will adapt, otherwise we would not see the hydrological partitioning of catchments around the world broadly plotting along the Budyko curve. However, unanswered questions are how long it will take for vegetation to adapt and how it will adapt. While the Budyko framework is a well-established concept, the recent study by Reaver et al. (2021) shows that it should be cautiously applied in changing systems which are not in equilibrium. We will include this discussion in the revised version of the manuscript.*

Despite these uncertainties, there are also strong aspects of our methodology. Current practice in most climate change assessment studies assumes constant system properties in the future, thereby neglecting adaption of vegetation to local climate conditions. Our analysis is a first step in evaluating what may happen if we consider ecosystem adaptation in response to climate change in hydrological model predictions. Our method is based on readily available data and is therefore easily applicable. Furthermore, if we assume space and time symmetry, i.e. the exchange of spatial knowledge with temporal knowledge, we may be able to transfer root-zone storage capacity estimates from a location X with a current climate similar to the future climate of a location Y.

Comment 3:

Methods: no limitation of the root-zone storage capacity

The methodological approach assumes a limitless adaptation of the root-zone storage capacity to changing aridity index (compare l. 243). I was wondering whether this is realistic. The adaptability of the root-zone depends on the vegetation's capability to change the root system following a change inclimate/water demand. This capability probably depends on the vegetation type (crop, grass, or tree)but also on the species. Also, adapting the root-zone storage capacity is not the only way that plants/vegetation might adapt to a change in aridity index. Plants can adapt to drier conditions by closing their stomata and reducing gas exchange with the atmosphere and hence transpiration. Also,overall vegetation cover could decrease if the water supply is not sufficient to support the same cover.Although I think it is not necessary to consider this limitation in this proof-of concept study, it is nevertheless an important point to discuss in the discussion section.

Reply 3:

This is a very good point, we briefly mention it in the discussion when we refer to the study of Zhang et al. (2020). However, we fully agree that the different strategies of vegetation to cope with changing environmental conditions need to be discussed in more detail in the revised version of the manuscript.

Comment 4:

Links to ecohydrological modelling or dynamic global vegetation models (DGVMs) missing Although this study is about hydrological modelling, I think that the advances and contributions ofecohydrological models and DGVMs to studying the feedbacks between vegetation and the watercycle should be mentioned and discussed in the introduction and, if applicable, also in the discussionsection of the manuscript. Please find some hints on where to start in the following:

One prominent model is e.g. the DGVM LPJmL which dynamically models carbon, nitrogen and waterflows. The model has been applied to various question among them also questions related to water flows under climate and land-use change.

You can e.g. have a look at the following publication:
Rost et al. (2008), *Water Resources Research*.

https://doi.org/10.1029/2007WR006331Here you can find a list of some key

publications of the model:

https://www.pik-potsdam.de/en/institute/departments/activities/biosphere-water- modelling/lpjml/key-publications

In the field of ecohydrological modelling, you could have a look at the works of Ignacio Rodriguez-Iturbe and Amilcare Porporato. An ecohydrological study to look at might be Tietjen et al. (2017), *Global Change Biology* https://doi.org/10.1111/gcb.13598. The study looks at feedbacks betweensoil water availability, vegetation change and climate change and they disentangle the effects of climate change alone and climate change in combination with vegetation change.

Reply 4:

We thank the reviewer for providing the references of these relevant studies on ecohydrological modelling. It is interesting to read that in the study of Tietjen et al. (2017), the future vegetation cover is determined based on empirical relations relating the fraction of each plant functional type to mean *annual temperature and precipitation. The rooting depth for each plant functional type is a fixed estimate derived from a re-analysis of a global root dataset. Instead in our approach, we do not impose a fixed rooting depth, but it is estimated from the future climate data and our estimate of future transpiration. In the LPJmL4 model (Schaphoff et al., 2018), transpiration depends on the water accessible for plants, which is computed from the relative water content at field capacity and the root distribution within each soil layer. These root distribution estimates are also fixed parameters for each plant functional type considered in the model. Accounting for this climate control on root development and root-zone parameterization in ecohydrological model could potentially also be very interesting (van Oorschot et al., 2021). We will discuss the links with ecohydrological and vegetation models in the revised version of the manuscript.*

Comment 5:

l. 58: optimality principles: is this an established term? If not specify what is optimized in this approach(probably it's vegetation growth or something similar)

Reply 5:

We will clarify in the revised version of the manuscript that the optimality principles indeed refer to vegetation growth through optimal allocation of aboveground and belowground resources. This implies that ecosystems have developed root systems to ensure access to sufficient (but not more) water to overcome dry periods (Guswa 2008; Schymanski et al., 2008).

Comment 6:

l. 95: "land-use change under future conditions": The manuscript does not tackle land-use under future conditions. The authors test what happens if land-use is the same in the whole catchment basedon what is already there. But it is never discussed which land-use types are realistic for the future or whether there is a trend in land-use towards any of the present land-use types. Rephrase to make clear that this is just a theoretical assessment of the sensitivity towards different types of land-use instead of a projection into the future. Also, the statement "we exchange space-for-time" (l. 96) suggests, that there is a known land-use trend for the future.

Reply 6:

We agree that we perform a sensitivity assessment of potential/theoretical land-use change and not necessarily projected land-use change and we will rephrase this to "land-use change under potential future conditions". However, the potential changes applied are based on a space-for-time exchange, using characteristics from the Budyko framework of a set of existing catchments to simulate potential changes in a set of different catchments. We believe that the statement "exchange space-for-time" can also be used in case the land-use change for the future is only theoretical.

Comment 7:

The method description is generally a bit confusing. I feel that generally it could be a bit shorter (e.g. the scenario description and the description of the 4 different root-zone storage capacities) are repetitive at some points. It might also help to provide a supportive figure of the study's workflow that clearly separates between different sources of input data, generation of scenarios and model application (instead of Fig. 3 which would fit better in the Supplemental material). Please revise the method section for more clarity and structure. The specific comments below hopefully help to do that.

Reply 7:

We thank the reviewer for his detailed comments to improve the clarity of the method section. In the revised version of the manuscript, we will try to improve Figure 4 to clarify the workflow, the scenarios and the data used for each scenario. We agree that the current structure is sometimes repetitive, but we think it has the advantage of clearly distinguishing the four different scenarios. Nevertheless, in the revised version of the manuscript, we will try to restructure the Method section in such a way that repetitions are reduced while keeping the distinction in modeling results for each of the scenarios. We agree with the suggestion of the reviewer to move Figure 3 to the Supplement.

Comment 8:

l. 109: "divided into three main zones": It would be nice to see these three main zones in the Figureas well. In the figure, it is unclear which part of the catchment represents which of these three zones.

Reply 8:

Yes, you are right, we will indicate the three zones on the map.

Comment 9:

l. 120: reference is missing for the meteorological variables

Reply 9:

Indeed. The numbers are based on the E-OBS data (Section 3.1) and the historical streamflow data (Section 3.3), we will add these references in the text.

Comment 10:

l. 122: always refer to the specific label of the figure if possible (here it's Fig. 1c and not Fig. 1)

Reply 10:

Agree, we will be more specific.

Comment 11:

l. 147-161: A figure or some numbers comparing the simulated historical and 2K climate scenarios could be a nice addition. From the description, it remains unclear what a "globally 2K warmer world" (l. 158) will translate to in this regional data set. Does this 2K warmer world lead to a mean 2K warmer regional climate? What's the difference in mean annual temperature and mean annual precipitation in 2K vs. historical climate?

Reply 11:

We agree that a Table summarizing mean annual temperature, potential evaporation and precipitation for the different data sources is a useful addition. Differences in mean annual potential evaporation and precipitation between the simulated 2K and historical climate are now shortly described in the result section 5.1.3 (L396). We will elaborate this further in the revised version of the manuscript.

Comment 12:

l. 164: It would be helpful to add Borgharen to the catchment map in Fig. 1

Reply 12:

Good point, we will add Borgharen on the map of Figure 1.

Comment 13:

Methods: The decision to divide the land-use types into broadleaved forest on the one hand, and coniferous forest/agriculture on the other hand needs better explaining. Why is a tree-dominated (coniferous) vegetation grouped with crops? I would expect that crops and trees are very different with regard to their effect on the water cycle and concerning their root-storage capacity.

Reply 13:

We understand that it may sound confusing. However, the division of both groups was made according to the percentage of broadleaved forest, as we found that omega values tended to be lower for areas with relatively more broadleaved forests (25-38%) in comparison to catchments with relatively low fractions of broadleaved forests (1-12%), as also shown in 5b. We then related this finding to the fact that in the Walloon part of the catchment, most of the old broadleaved forest has been converted to coniferous plantations and agricultural areas, whereas the broadleaved forest mostly remained in the French part of the catchment. In the manuscript, when we refer to "broadleaved forest" versus "coniferous and agriculture", we implicitly mean catchments with relatively high or relatively low percentages of broadleaved forest. However, it is easy to overlook the words "high" and "low" when reading these descriptions, which is why we refer to "broadleaved" and "coniferous and agriculture". We will add a note on this in the revised version of the manuscript.

Comment 14:

l. 233: Why is Imax taken as 2mm?

Reply 14:

We estimate the interception storage capacity (Imax) at 2 mm based on analyses performed in previous studies which report a low sensitivity of the root-zone storage capacity to the value of Imax *(de Boer-Euser et al., 2016, Bouaziz et al. 2020). In Bouaziz et al. (2020), we tested the sensitivity of applying interception storage capacities of 0.5, 1.0, 2.0 and 3.0 mm and found a relatively limited impact on the root-zone storage capacity. To reduce the complexity of our analyses, and because of this low sensitivity and our interest in the effect of stationarity versus non stationarity of the root-zone storage capacity in the four scenarios, we decided to use a single value for the interception storage capacity. A single value was also used in van Oorschot et al. (2021). We will include these references in the revised version of the manuscript to explain our choice.*

Comment 15:

l. 262: Why are E-OBS data taken from 1980-2018 while streamflow data is only from 2005-2017? Would the results have been different if E-OBS data from 2005-2017 were used instead?

Reply 15:

Thank you for pointing this out. When calculating the root-zone storage capacities, we actually used the period 2005-2017 for both the streamflow data and the meteorological data. This was then not correctly reported in the text, we will make sure to correct this in the revised version of the manuscript. Using the period 2005-2017 or 1980-2017 for the meteorological data in the estimation of $S_{R,max}$ leads to relatively similar ranges of root-zone storage capacities across the scenarios, as shown in Figure 1.

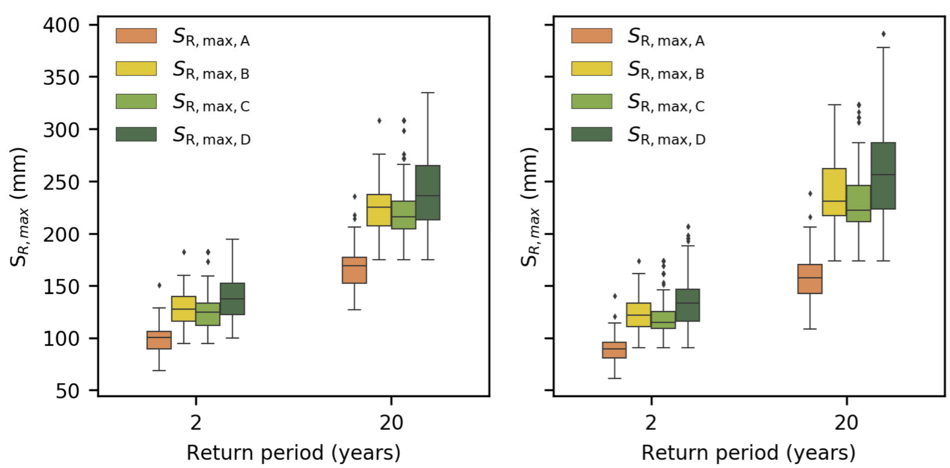

Figure 1 Left: Root-zone storage capacities for the 35 catchments of the Meuse basin for the four scenarios derived using meteorological data between 2005-2017 (sane as Figure 5c of the manuscript). Right: Root-zone storage capacities derived using meteorological data between 1980-2017.

Comment 16:

l. 289 ff: How were the ω values sampled?

Reply 16:

When we estimated the root-zone storage capacities for the land-use change scenarios C and D, we estimated the long-term actual evaporation from the Budyko curve through a horizontal shift along the parametric Budyko curve to account for a change in aridity index, and a vertical shift towards a *different parametric Budyko curve to account for a change in land-use. For each catchment under change, we assigned an omega value randomly sampled from the set of catchments with current characteristics representing the future characteristics of the catchments under change. We repeated this random sampling seven times, which resulted in seven parameter combinations of $S_{R,max}$ for scenario C and seven parameter combinations for scenario D. We will clarify this in the revised manuscript.*

Comment 17:

l. 306: hillslopes are associated with forest and plateau with agriculture. But which type of forest do you mean here? Broadleaved or coniferous?

Reply 17:

The three hydrological response units defined in the hydrological model are determined from topographical data (based on thresholds for Height Above the Nearest Drain and slope) and land-use data (where broadleaved and coniferous forests were both included in the hillslope class, while agricultural land was included in the plateau class). The three classes have slightly different parameterization to reflect different dominant hydrological processes. In the land-use scenarios, we did not change the percentages of each HRU in our model representation. We agree that this is a limitation of our approach, which we will add in the Discussion. However, the data to determine how the link between land-use and HRU may change in the future is not known at this detailed level. Additionally, we expect a limited impact of adapting the fractions of HRU on the hydrological response and we therefore consider this to be an acceptable limitation of our study.

Comment 18:

l. 331ff: "the performance … for the ensemble of retained parameter sets": From the 10000 calibration runs: how many parameter sets were obtained for the model runs? From the supplemental material it looks like the prior is almost the same as the posterior parameter distribution.

Reply 18:

We retained 124 parameter sets based on the defined criteria for model performance. To deal with the relatively long computational costs of running the model, we applied a preliminary first calibration to pre-scan the range of prior distributions. The real calibration was performed with these reduced parameter ranges as prior, which explains the limited difference between prior and posterior distributions.

Comment 19:

l. 334-337: This section can be removed as it is a repetition of what was already mentioned above in lines 272-274.

Reply 19:

We agree and will remove the repetition.

Comment 20:

Scenario description in 4.4:

- It is unclear which values of S_R with regard to the return period are used (2 years or 20 years?)
- How did you decide for the return period in the mixed agricultural/coniferous land-use? Agriculture should be 2 years and forest 20 years (l. 251-253)

Reply 20:

In the distributed model, each cell has a percentage wetland, hillslope and plateau. The root-zone storage capacity parameter for the wetland and plateau hydrological response units were assigned a return period of 2 years, while a return period of 20 years was assigned to hillslope. We refer to the studies of Nijzink et al. (2016) and Gao et al. (2014) where return periods of 20 years are associated with forested areas. Lower return periods of 2 years are better suited for agricultural areas (Wang-Erlandsson et al. 2016). We will clarify this in the revised version of the manuscript.

Comment 21:

l. 357, 362, 369: no need to repeat that $S_{Rmax,a}$ is used as a parameter in the historical run for every scenario. Better to mention it once, when the historical run is explained.

Reply 21:

Agree, we will adapt this in the revised version of the manuscript.

Comment 22:

Results: It is not always clear what the reported numbers represent. Median and standard deviation? Mean and standard error of the mean? E.g. l. 374 & 377, l. 382, l. 390 & 391, l. 402, l. 408.If the reported values are always the same, you could also mention it once and state that all subsequent values represent the same measures.

Reply 22:

Good point, the reported numbers represent the median and standard deviation, we will make sure to mention this once clearly.

Comment 23:

l. 376: should this be ω_{obs} instead of ω?

Reply 23:

Correct, we will adapt this.

Comment 24:

l. 377: should this be transpiration instead of evaporation? This is a general issue: there is no clear distinction between evaporation, transpiration and evapotranspiration in the text.

Reply 24:

Throughout the manuscript, we use the term evaporation to represent all the different evaporation components (interception, transpiration and soil evaporation). It is perhaps a matter of taste, but we like to follow the terminology proposed by Savenije (2004) and Miralles et al. (2020), where evaporation instead of evapotranspiration is used to refer to all evaporative fluxes.

Comment 25:

l. 377-379: The differences of ω between the catchments is mainly attributed to the differences in the main vegetation type (broadleaved vs. coniferous/agriculture, l. 377-379). However, the catchments also differ substantially in other characteristics (French part: thick soils and gentle slopes, thin soils and steep terrain in the Ardennes, porous chalk in Wallonia (l. 109-113)). It should be discussed to what extent the differences in ω might not be dependent on the vegetation cover alone but also on the topography and soil type/thickness. Also, what are the implications of this regarding the method?How sure are you that the differences in hydrology between land-use types are really caused by the vegetation cover and not by the underlying topographical and soil characteristics?

Reply 25:

This is an interesting question which we will include in the discussion of the revised version of the manuscript. The differences in omega-values are most probably related to a combination of biophysical features. However, considering that transpiration is the largest continental water flux (Jasechko, 2018) and that omega values determine the hydrological partitioning, we assume that the variability in omega values is largely controlled by the root-accessible water volume $S_{R,max}$. This root-accessible water volume is independent from the soil type, as root systems will develop in a way to ensure sufficient access to water. In clayey soils, the rooting depth might be shallower than in sandy soils for an identical root-zone storage capacity. In our opinion, geology, soils characteristics and topography are implicitly integrated in other model parameters, e.g. the time scales of the linear reservoirs which represent the subsurface flow resistance in different parts of the system.

Comment 26:

l. 394: Fig. 2b should either be referenced earlier in the text, e.g. when talking about the difference between the historical and the 2K climate time series in the method section or it should be a separateresult figure that comes later in the text.

Reply 26:

Is it perhaps possible that you overlooked the reference to Fig 2b earlier in the text in Section 4.1.2 (L243) to illustrate the water-balance approach to estimate the root-zone storage capacity?

Comment 27:

l. 424: “median values of approximately 0.93”: why approximately?

Reply 27:

You are right, we will remove ‘approximately’

Comment 28:

l. 431: “streamflow during the wettest months”: include which months you mean by “wettest months”

Reply 28:

Good point, we will clarify that here we refer to the months December and January as wettest months.

Comment 29:

l. 500: “shows distinct patterns of change”: more precise language could be used: Which response variables differ and are they larger or smaller compared to the stationary scenario?

Reply 29:

This is a good suggestion, we had not included more details to avoid repetition from the result section. However, we think changes in streamflow and evaporation can briefly be repeated here to be more precise. We will clarify this in the revised manuscript.

Comment 30:

l. 512-519: This section does not fit in the “Implications” section of the discussion. It is more of a limitation of the current study or an outlook of what could be done next. It could e.g. be moved to the“Limitations and Knowledge gaps” section of the discussion.

Reply 30:

We agree that this paragraph contains an outlook of what could be done next. Initially, we had seen this as an implication of our work for future work, but we agree that it would better fit in a "Limitations and outlook" section of the discussion. We will adapt this in the revised version.

Comment 31:

l. 524-256: It is not clear to me, why the results on actual evaporation differences between the scenarios indicate disagreements among model process representations. Please elaborate more on this point. Also, what are the specific "processes that become relevant in the future"?

Reply 31:

What we mean here is that in the future scenario, evaporation demand increases. In scenario 2K$_A$, where the root-zone storage capacity has not adapted to the future climate, we see water stress conditions that do not occur in the other scenarios. The different model representations amongst scenarios lead to different hydrological responses. However, we might consider removing this point in the revised version of the manuscript and add the other relevant points of discussion mentioned earlier in our reply to Comment 2.

Comment 32:

l. 333-334: The conclusion, that vegetation is important for regulating the water cycle is correct but itis also quite established and not really a specific discussion of your results.

Reply 32:

We agree that this conclusion is already quite established. We will rephrase this statement to emphasize how our study contributes to the quantification of the potential impact of vegetation adaptation in regulating the water cycle.

Comment 33:

l. 535-539: This discussion is also a limitation of your study or an outlook to further work. It should notbe under the "Implications" subheading of your discussion.

Reply 33:

We agree that this part of the discussion is also more an outlook for future research and will move this to the Limitation and outlook section.

Comment 34:

l. 542: “it is unclear how ecosystems will cope with climate change”: A discussion of how useful your approach to include vegetation into hydrological models under climate change in the light of this uncertainty would be interesting. To what extent can we be sure that the root-zone storage capacity can adapt to changing climate? What evidence is there from other studies regarding this issue? How would you proceed with your approach if vegetation changes to a vegetation type for which there is no data from the same region?

Reply 34:

This is a very interesting point. There is increasing evidence that vegetation efficiently adapts its root-zone storage capacity to ensure sufficient access to water (Guswa 2008, Schymanski et al. 2008). However, while we know that the ecosystem will eventually adapt to changing environmental conditions, partly by changing the mix of vegetation species and partly by vegetation adjusting its rooting depth or density, the question is how long it will take for an ecosystem to adapt in relation to the rate of climate change. Also, there are limits to the capacity of an ecosystem to adapt, for instance when is the threshold passed for the adaptability of rainforest to become savannah, or where lies the threshold for savannah to become desert? In this study we assume that adaptation thresholds are not reached. We refer to our reply to comment 2 for further details on this matter.

An interesting next step for our methodology will be to apply it in a climate-matching approach (Fitzpatrick and Dunne, 2019), where the current climate and landscape characteristics of a location X match the future climate or landscape characteristics of a location Y. This climate matching could be applied over distant regions, using datasets which combine landscape and climatological data over large samples of catchments (e.g. the various CAMELS datasets). Despite considerable uncertainties, this may allow us to infer vegetation adaptation and the associated changes in root-zone storage capacity from identifying regions in the world where the current climate resembles the projected future climate in a different region.

Comment 35:

At the end of the discussion, you mention several times that this study should be read as a sensitivityanalysis (l. 571) and a proof-of-concept (l. 575). This should also be made clear in the abstract. Also, athorough discussion of the advantages and disadvantages of the presented method is missing. What are possible applications of it, to what types of regions/questions can it be applied? What are the limitation and what could be improved?

Reply 35:

This is a very good suggestion, in the revised abstract, we will more strongly emphasize that our study should be understood as a sensitivity analysis. As also mentioned in our replies to the main comments (1 and 2), we will not use the terminology “proof-of-concept” anymore as we cannot test our results against future observations. We agree with your suggestion to more thoroughly discuss the advantages and disadvantages of the presented method in the discussion. We refer to our detailed reply to Comment 2 for the specific points that we will address.

<u>Comment 36:</u>

Figure labels should be in the same position for all figures (e.g. top left)

<u>Reply 36:</u>

Agree, we will adapt this in the revised version.

<u>Comment 37:</u>

Figure labels could be bold for better visibility?

<u>Reply 37:</u>

Good suggestion, we will adapt this in the revised version.

<u>Comment 38:</u>

Why are the scenario names (2Ka-d) that are defined in Fig. 4 never used? Instead $S_{rmaxa-d}$ isused in Figs. 5,8,9? If scenario names are given, they should be used consistently.

<u>Reply 38:</u>

Very good point. In Figure 9, we are actually showing values of $S_{R,max}$. However, in Figure 8 and 9, it indeed makes more sense to refer to Scenario 2KA etc in the labels.

<u>Comment 39:</u>

Fig. 1:

- colours of figure b): better use some continuous colour scheme
- Figure labels are inconsistent, b and c not on the same height
- Fig. 1b: what are the black points? Are they the streamflow measurement locations? Mention in the caption
- Fig. 1 does not reflect well many aspects mentioned in the text (2.1 landscape and 2.2 landuse)
    - Which are the three zones mentioned in l. 109? Are they represented in Fig. 1b? If yes you could add this to the caption. It is not clear what is the French, the Ardennesand the Wallonia part mentioned several times in the text
    - Fig. 1b: The numbers don't really match with the text. In Walloon 44% of the broadleaved forest should be there (l. 126), but in the figure the max. percentage is38%.

<u>Reply 39:</u>

- *We will test if an alternative color scheme improves readability.*
- *We will move the labels*
- *The black points are indeed the streamflow measurement locations, we will add this in the*

caption.

- *We will add the location of the three zones*
- *When we refer to 44% in the text, we mean 44% of the 18th century Walloon forests of Belgium that have remained from the original broadleaved forests. The 38% in the figure refer to the fraction of broadleaved forest within a catchment.*

Comment 40:

Fig. 3:

- Maybe this figure fits better in Supplement S3 because it is part of the model description? Idon't find it very helpful in the manuscript without the context of the model formulas

Reply 40:

We agree that Figure 3 can be moved to the Supplement to be connected to the model description. We will modify this in the adapted version.

Comment 41:

Fig. 5:

- Labels are missing
- Figures are a bit small: Could be a made bigger if empty space between panels is reduced
- 5b:
  - ω_{obs} should be on the y-axis not just ω
  - Axis text: No % because it's already in x-axis title
- 5c:
  - Caption last sentence: "A similar but reversed approach is applied …" It is the *same* and not a *similar* approach that was used.

Reply 41:

- *Indeed, we will add the labels in the revised version.*
- *We will try to decrease the empty space between the panels to increase the panels themselves.*
- *We will replace ω by ω_{obs}*
- *We will remove % from the x-axis title*
- *Indeed, we will replace similar by same.*

Comment 42:

Fig. 6:

- What is the ribbon for the modelled values: range from all realistic parameter sets of thecalibration?

Reply 42:

Indeed, the ribbon represents the ensemble of feasible parameter sets, we will clarify this in the caption.

Comment 43:

Fig. 7:

- Could be larger: box is not visible
- Don't use transparent colours to distinguish the panels. In my opinion they are alreadydistinguished enough by the panel titles and labels in the caption (same for figures in Supplement S3)
- Labelling is not consistent (compare to labelling of Fig. 6)
- Why is there such a big difference between Borgharen and the 34 catchments? Isn'tBorgharen just a summary of all the catchments?

Reply 43:

- *It is more the shape of the violin plots (left and right) which are important here.*
- *We consistently applied a color code throughout the Figures and would like to keep it as we believe it increases the clarity.*
- *We will change the labeling order.*
- *Borgharen is the most downstream outlet point considered. Often, model performance tends to decrease for smaller catchments. Additionally, the calibration was performed at Borgharen.*

Comment 44:

Fig. 8:

- Caption 8e) maybe mention that y-axis is different scale (compare to caption of Fig. 7)

Reply 44:

Yes, we will add this in the revised version.

Comment 45:

Fig. 9:

- What are the ribbons and lines? Median + conf. interval?

Reply 45:

Good point, they indeed show median and range of ensemble retained sets, we will clarify this in the caption.

Comment 46:

S1: Monthly correction factors for E-OBS precipitation data

- First sentence: Citation missing

Reply 46:

Indeed, we will add the missing reference.

Comment 47:

S4: Prior and posterior parameter distributions

- State in table heading, that the last 3 columns are the posterior parameter distributions

Reply 47:

Yes, we will add this in the revised version.

Comment 48:

l. 54: rephrase to: sensitivity of the hydrological response to change in …

Reply 48:

Yes, we will rephrase.

Comment 49:

l. 62: remove “as often referred to”

Reply 49:

Agree.

Comment 50:

l. 79: remove the full stop before the list of references

Reply 50:

Yes.

Comment 51:

l. 191 & 1197: same style for (p1), (p2) and p3 (either with or without brackets)

Reply 51:

Yes, we will make this consistent in the revised version.

Comment 52:

l. 392: replace “return periods of 2 year” with either “2 year return period” or “return period of

2years”. Also check the subsequent text as this mistake happens several times.

Reply 52:

Good point, we will replace.

Comment 53:

l. 410: Vertical space is missing as a new paragraph begins in line 411

Reply 53:

Not sure what is meant here, the spacing looks the same as in the other paragraphs.

Comment 54:

l. 500: “compared to” instead of “with respect to”?

Reply 54:

Ok, we will adapt.

Comment 55:

l. 592: “distinct change of sign”: remove distinct

Reply 55:

Agreed.

Comment 56:

Avoid unspecific adverbs. Either remove them, or state specifically what you mean by them. E.g.

- l.114: “relatively short response time” (how short is relatively short?)
- l. 422: “relatively well reproduced”
- l.423: “slight underestimation” and “relatively similar performance”

Reply 56:

L114, we will be more specific about the response time in the revised version.
L422 and 423, numbers are given later in the sentence, we will clarify this in the revised version.

References

Berghuijs, W. R., Gnann, S. J., & Woods, R. A. (2020). Unanswered questions on the Budyko framework. *Hydrological Processes*, (October), 1–5. https://doi.org/10.1002/hyp.13958

de Boer-Euser, T., McMillan, H. K., Hrachowitz, M., Winsemius, H. C., & Savenije, H. H. G. (2016). Influence of soil and climate on root zone storage capacity. *Water Resources Research*. https://doi.org/10.1002/2015WR018115

Bouaziz, L. J. E., Steele-Dunne, S. C., Schellekens, J., Weerts, A. H., Stam, J., Sprokkereef, E., et al. (2020). Improved understanding of the link between catchment-scale vegetation accessible storage and satellite-derived Soil Water Index. *Water Resources Research*. https://doi.org/10.1029/2019WR026365

Brunner, I., Herzog, C., Dawes, M. A., Arend, M., & Sperisen, C. (2015). How tree roots respond to drought. *Frontiers in Plant Science*, *6*(JULY), 1–16. https://doi.org/10.3389/fpls.2015.00547

Donohue, R. J., Roderick, M. L., & McVicar, T. R. (2012). Roots, storms and soil pores: Incorporating key ecohydrological processes into Budyko's hydrological model. *Journal of Hydrology*, *436–437*, 35–50. https://doi.org/10.1016/j.jhydrol.2012.02.033

Donohue, R. J., Roderick, M. L., McVicar, T. R., & Farquhar, G. D. (2013). Impact of CO2 fertilization on maximum foliage cover across the globe's warm, arid environments. *Geophysical Research Letters*, *40*(12), 3031–3035. https://doi.org/10.1002/grl.50563

Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., & Otero-Casal, C. (2017). Hydrologic regulation of plant rooting depth. *Proceedings of the National Academy of Sciences*, 201712381. https://doi.org/10.1073/pnas.1712381114

Fitzpatrick, M. C., & Dunn, R. R. (2019). Contemporary climatic analogs for 540 North American urban areas in the late 21st century. *Nature Communications*, *10*(1), 1–7. https://doi.org/10.1038/s41467-019-08540-3

Frank, D. C., Poulter, B., Saurer, M., Esper, J., Huntingford, C., Helle, G., et al. (2015). Water-use efficiency and transpiration across European forests during the Anthropocene. *Nature Climate Change*, *5*(6), 579–583. https://doi.org/10.1038/nclimate2614

Gao, H., Hrachowitz, M., Schymanski, S. J., Fenicia, F., Sriwongsitanon, N., & Savenije, H. H. G. (2014). Climate controls how ecosystems size the root zone storage capacity at catchment scale. *Geophysical Research Letters*, *41*(22), 7916–7923. https://doi.org/10.1002/2014GL061668

Gentine, P., D'Odorico, P., Lintner, B. R., Sivandran, G., & Salvucci, G. (2012). Interdependence of climate, soil, and vegetation as constrained by the Budyko curve. *Geophysical Research Letters*, *39*(19), 2–7. https://doi.org/10.1029/2012GL053492

Guswa, A. J. (2008). The influence of climate on root depth: A carbon cost-benefit analysis. *Water Resources Research*, *44*(2), 1–11. https://doi.org/10.1029/2007WR006384

Hrachowitz, M., Stockinger, M., Coenders-Gerrits, M., van der Ent, R., Bogena, H., Lücke, A., & Stumpp, C. (2020). Deforestation reduces the vegetation-accessible water storage in the unsaturated soil and affects catchment travel time distributions and young water fractions. *Hydrology and Earth System Sciences*, *i*(June), 1–43. https://doi.org/10.5194/hess-2020-293

Jasechko, S. (2018). Plants turn on the tap. Nature Climate Change, 8, 560–563.

Miralles, D. G., Brutsaert, W., Dolman, A. J., & Gash, J. H. (2020). On the use of the term "Evapotranspiration." *Earth and Space Science Open Archive*, 8. https://doi.org/10.1002/essoar.10503229.1

Nijzink, R., Hutton, C., Pechlivanidis, I., Capell, R., Arheimer, B., Freer, J., et al. (2016). The evolution of root-zone moisture capacities after deforestation: A step towards hydrological predictions under change? *Hydrology and Earth System Sciences*, *20*(12), 4775–4799. https://doi.org/10.5194/hess-20-4775-2016

van Oorschot, F., van der Ent, R., Hrachowitz, M., & Alessandri, A. (2021). Climate controlled root zone parameters show potential to improve water flux simulations by land surface models. *Earth System Dynamics Discussions*, 1–26. https://doi.org/10.5194/esd-2021-3

Reaver, N., Kaplan, D., Klammler, H., & Jawitz, J. (2020). Reinterpreting the Budyko Framework. *Hydrology and Earth System Sciences Discussions*, (November), 1–31. https://doi.org/10.5194/hess-2020-584

Savenije, H. H. G. (2004). The importance of interception and why we should delete the term evapotranspiration from our vocabulary. *Hydrological Processes*, *18*(8), 1507–1511. https://doi.org/10.1002/hyp.5563

Schaphoff, S., Von Bloh, W., Rammig, A., Thonicke, K., Biemans, H., Forkel, M., et al. (2018). LPJmL4 - A dynamic global vegetation model with managed land - Part 1: Model description. *Geoscientific Model Development*, *11*(4), 1343–1375. https://doi.org/10.5194/gmd-11-1343-2018

Schymanski, S. J., Sivapalan, M., Roderick, M. L., Beringer, J., & Hutley, L. B. (2008). An optimality-based model of the coupled soil moisture and root dynamics. *Hydrology and Earth System Sciences*, *12*(3), 913–932. https://doi.org/10.5194/hess-12-913-2008

Singh, C., Wang-Erlandsson, L., Fetzer, I., Rockström, J., & van der Ent, R. (2020). Rootzone storage capacity reveals drought coping strategies along rainforest-savanna transitions. *Environmental Research Letters*. https://doi.org/10.1088/1748-9326/abc377

Tietjen, B., Schlaepfer, D. R., Bradford, J. B., Lauenroth, W. K., Hall, S. A., Duniway, M. C., et al. (2017). Climate change-induced vegetation shifts lead to more ecological droughts despite projected rainfall increases in many global temperate drylands. *Global Change Biology*, *23*(7), 2743–2754. https://doi.org/10.1111/gcb.13598

Troch, P. A., Carrillo, G., Sivapalan, M., Wagener, T., & Sawicz, K. (2013). Climate-vegetation-soil interactions and long-term hydrologic partitioning: Signatures of catchment co-evolution. *Hydrology and Earth System Sciences*, *17*(6), 2209–2217. https://doi.org/10.5194/hess-17-2209-2013

Wang-Erlandsson, L., Bastiaanssen, W. G. M., Gao, H., Jägermeyr, J., Senay, G. B., Van Dijk, A. I. J. M., et al. (2016). Global root zone storage capacity from satellite-based evaporation. *Hydrology and Earth System Sciences*, *20*(4), 1459–1481. https://doi.org/10.5194/hess-20-1459-2016

Yang, Y., Roderick, M. L., Zhang, S., McVicar, T. R., & Donohue, R. J. (2019). Hydrologic implications of vegetation response to elevated CO2 in climate projections. *Nature Climate Change*, *9*(1), 44–48. https://doi.org/10.1038/s41558-018-0361-0

Zhang, B., Hautier, Y., Tan, X., You, C., Cadotte, M. W., Chu, C., et al. (2020). Species responses to changing precipitation depend on trait plasticity rather than trait means and intraspecific variation. *Functional Ecology*, (September), 2622–2633. https://doi.org/10.1111/1365-2435.13675