We thank you for the insights and comments which will help us to improve the manuscript. We will for now focus on the major points brought up in the review.

One issue raised in several parts of the review is related to the characteristics of the selected catchments, which we agree to have neglected slightly. We will provide more information on the catchments, also including soil type or texture, and relate this info to our results and the discussion accordingly. Data on soil properties, however, are not widely available, or are limited in accuracy and representativeness, especially at the catchment scale. We also want to emphasize that the point of our study was not to explain the as-is state of the catchments (or their calibrated parametric representations), but to use them in a virtual experiment approach with 16 catchments and realistic boundary conditions, in order to explore the role and sensitivity of soil storage characteristics in deviations from the Budyko curve. We found rather consistent sensitivity patterns - most of the systems reached the Budyko curve through variation of the soil characteristics in reasonable ranges. But we agree that it will be interesting to compare the individual deviations with actual soil data.

A second major point concerns the coevolution of catchments, which we adopted from Troch (2015) for the discussion part. We agree with the reviewer that human activities have altered catchments and their soils, this influence being however comparably recent (more extensive agriculture maybe over the past 200-300 years), whereas soils developed with the climate at least since the last ice age which ended around 10.000 years ago. Coevolution of catchments or landscape elements happens at a range of time scales (Troch 2015), from geological ones to climatic and ecosystem scales. The rate of "hydrologic aging", however, is hard to assess and depends on the drivers as nicely described by Troch.

In our study, we only tried to limit direct water balance-impacting measures like abstractions, to be able to estimate at least the present hydrologic partitioning/water balance. We did not aim to assess the potential alteration of catchment elements due to anthropogenic activity. We are also not really capable of modeling coevolution of form and functioning in hydrological systems. However, the Budyko curve yields a widely accepted framework to discriminate which combinations of catchment properties might have co-evolved and which not (as done e.g. by Schaefli 2011).
With our interpretation and visualization of parameter spaces regarding soil storage and its relationship to the Budyko curve, we aimed at understanding potential trajectories, analyze patterns found for climate groups, and tried to discuss deviations in the pattern for seasonal catchments in terms of the hydrologic age concept. We want to point out again that the focus of our study is not on best explaining the current state of the catchment but to quantify how even small changes in total storage or in the field capacity of the root zone control and explain frequently observed offsets from the Budyko curve. As catchments and soil properties are not static but changing, this naturally involves a co-evolutionary aspect.

The study by Hartmann et al. 2020 shows such a soil development at scales of centuries and millennia, both in terms of soil storage and retention characteristics. It is an exemplary study, that is correct. However, it shows for two different parent materials (on proglacial moraines), a similar qualitative development at different rates. That was supposed to corroborate our assumed weathering processes to discuss our variation scheme, but this aspect can of course not be generalized to all pedogenetic environments. We will discuss that point more critically in a revised version of the paper, and also include that other mechanisms including human activity can influence soil storage characteristics.

The last major concern was on model parameters and parameter interaction. We agree that varying only certain model parameters, while others remain unaltered, is a simplification of how real-world systems would develop or change. Any parameter variation or sensitivity study needs to take into account if the underlying model is representative. We are confident that conceptual hydrological models such as the HBV approach can represent hydrologic systems sufficiently well for describing catchment dynamics at the long-term scale, especially as we are considering that the virtual catchments we are exploring are not changing their properties during the simulation. We will clarify these points in the discussion.

We have tried to visualize parameter interactions in the 2d parameter spaces (Fig. 8). We focused on the parameters $S_{\text{max}}$ and $FC_{\text{frac}}$, which can be related to two forms of soil storage volumes. They are important in the two stages of the evaporation process: at first, when the atmospheric demand is met, and secondly, when the capillary soil transport/supply is impeded by retention. The other two parameters were not included in this analysis. The $k_{\text{res}}$ parameter is insensitive for the long-term annual water balance, since it does not affect the hydrologic partitioning but the dynamics of stream flow. The beta parameter is lumping several catchment characteristics and affects water partitioning and runoff dynamics, and thus cannot be interpreted in a straightforward manner. It originally was a shape parameter that is supposed to describe the growth of contributing areas with increasing relative saturation, which can be related to steepness, to macropore structures, bedrock permeability etc. Including beta in the analysis and Fig. 8 would in our opinion not add to their interpretability and understanding of the points we try to explore.