We thank you for reviewing and commenting our paper. We would like to clarify some of the irritations expressed in the review.

The focus of the study is to understand how soil storage characteristics relate to offsets from the Budyko curve. The Budyko curve is therefore an inherent part of the study, and could not simply be left out.

The HBV model evidently does not provide a "perfect picture" of hydrological processes, but it is a well-known and widely accepted conceptual hydrological model. The modeling concept appears suited for our approach, especially to distinguish between free and capillary-bound water storage. The soil parameters are, to a certain degree, interpretable and relatable to physical characteristics of catchments. We agree with you that observation data on this kind of catchment properties is lacking, and we therefore chose virtual experiments as a way of investigation. We cannot see, however, how this approach would suffer from a circular reasoning.

We want to stress that we do not claim that the soil storage parameters can explain any deviation from the Budyko curve, or that it is surprising that soils affect the long-term water balance. There is a practical motivation to use the Budyko curve to constrain hydrological modelling, or estimate water balances in data-scarce or ungauged catchments. This matter was brought forward in several other studies. It was, however, vastly unclear what abundant or deficient soil storage could mean in terms of offsets from this expected behavior, and if a variation of soil parameters alone within plausible ranges would lead to matching the Budyko curve, or not. We found these to be interesting questions.

Our results show that soil storage, but also capillarity characteristics, can lead to considerable deviations from Budyko, and thus should be included in any interpretation of Budyko offsets. We have also discussed how this could be interpreted in terms of soil development.

The focus of the study on soil properties does not mean that other influences are not important. Other studies have already shown other factors can be responsible for deviations as well. Some of our catchments also did not reach the curve through variation
of the soil parameters, while others did. While this is important to discuss, it is beyond the scope of the paper to analyze and compare all possible influences.

We encourage the reviewer to take a second look at our study along the viewpoints sketched above. We will check if our objectives could be stated more clearly in a revised manuscript to avoid possible misunderstanding.

The additional comments are also gratefully acknowledged. Please find specific answers to these below.

Additional comments:

- The impact of observational errors have not been taken into account. It possible that, at least for some of the study basins, the deviation of the observed EVR value from the Budyko curve is due to observational errors.

You are right, observational errors have not been part of the discussion, and can of course be the reason for deviations from the Budyko curve. We would include your point in the discussion in a revised manuscript. However, our study does not aim at judging or explaining the measured catchment water balances, but to explore the sensitivity of Budyko deviations to storage parameters. In that sense, measurement errors in the study catchments' water balances could affect the parametrization of the remaining model parameters that are fixed during the variation process. However, typical observational errors would not change the overall picture we got in terms of sensitivity patterns and the relationship to Budyko. In addition, at least the MOPEX dataset and the one from Germany are supposed to ensure a certain data quality.

- Line 125: The reasoning is not clear. Why do you need to select only the catchments with a closed water balance?

The Budyko framework predicts the mean hydrologic partitioning of rainfall into evapotranspiration and runoff – while storage changes average out at the long-term scale. If the water balance is not closed, for example because water passes a gauge underground and the evaporation ratio is estimated by $ET_a = P - Q$, runoff is mistaken for evapotranspiration, and the evaporation ratio is thus overestimated. Several catchments in the German dataset thus plotted outside the Budyko space. We compared evaporation ratios computed by $P-Q$ with actual evaporation estimates from an agrometeorological model and selected catchments where the two estimates were more or less (5% error) in accordance.

- Figure 9: Is it possible that number of rainy days is working as a proxy for something else, say mean precipitation? Otherwise, please provide a solid reasoning of why number of rainy days should matter.

Climatic variability such as frequency of rainfall events (expressed here by rainy days/year) has already been identified by Milly (1994) to be a sensitive parameter for the mean water balance. In the annex pdf we attached additional correlation plots for the EVR range / the sensitivity of EVR to soil storage volume:

- Figure 1 shows again the plot from the manuscript (Figure 9)
- Figure 2-4 correlates the EVR range to mean rainfall depth of a rainy day, to mean annual precipitation and to the dryness index

Figure 2 and 3 show that no. of rainy days is not working as a proxy for mean
precipitation. Dryness explains 65% of the variance, while no. of rainy days explains 93%.

With a given amount of total annual rainfall, the number of rainfall days relates to the mean depth of rainfall events as well as to the average interstorm period. The time between rainfall events influences the soil depletion (ETa resulting from ETp and rel. soil saturation) and thus the mean antecedent soil moisture state before rainfall events. More storage capacity leads to less saturated soils on average which in turn enhances infiltration and the amount of water available for evapotranspiration. We will include these explanations in the discussion of this point.

Please also note the supplement to this comment: