We appreciate this detailed review of our manuscript. We are glad that our somewhat unusual style is seen as positive. The reviewer lists three more general remarks and a number of detailed comments. Below we discuss how we will address these in the revised version of the manuscript (reviewer comments in blue italic font)

"Physically based" vs. "conceptual" ?

This is an interesting question. We agree that there is no clear distinction which makes some models physically-based and others not. However, as discussed in the review of Jens Christian Refsgaard, there is a difference in the degree to which models are physically-based. We fully agree that in the end, the crucial question is, which situations/tasks a model is suitable for. We will clarify this in the text.

Number of parameters and desire for optimization

There is a difference between how many parameters are in a model and how many are actually allowed to take different values. As discussed in the review, this (also) depends on how much spatial variability is considered. In the revision, we will follow the reviewer's advice and extend the discussion on when a distributed model is (not) needed.

Need for a classification of hydrological models

As also noted in the review of of Jens Christian Refsgaard, we used the term hydrological models in a rather narrow sense and will clarify that we in this manuscript are focusing on catchment (runoff) models. We agree that also within these models there are different 'families' and will extend the text slightly to clarify this.

- Line 115: "advocation for realism in model development"? What is 'realism' in this regard? And Is this still valid for today?

We will change the word 'realism' to 'best practices'

- Line 130: Instead "represent all hydrological processes", better term it "represent all relevant hydrological processes",

Agreed, will be changed

- Line 285: what do you mean with "re-organising observed flood generating factors'?

We will change 're-organising observed flood generating factors' to 'combining observed flood generating factors such as heavy rainfall, extreme snowmelt situations and wet antecedent conditions'

- Line 300: "risk for overparameterisation"? You may discuss it in the relation with complexity of model and nature? Also nature may have several means to come to one state? What can one do in this regards?

We will add 'i.e., the inability to determine one single best parameter set' for clarification

- Line 315: "groundwater dynamics"? You may elaborate a bit how far the catchments wetness in the HBV model can be related to observed gw-dynamics.

Groundwater dynamics were adressed in some early applications of a modified version of the HBV-model as described by Bergström and Sandberg (1983) and we have used groundwater dynamics as

additional criteria in previous studies (e.g., Seibert, 2000, HESS). We will extend section 4.4 and discuss opportunities and challenges of multi-criteria model evaluations.

Seibert, J., 2000. Multi-criteria calibration of a conceptual rainfall-runoff model using a genetic algorithm. Hydrology and Earth System Sciences, 4(2): 215-224.

Bergström, S. and Sandberg, G. (1983) Simulation of groundwater response by conceptual models - Three case studies. Nordic Hydrology, Vol. 14, No. 2.

- Lines 465-480: about the application for ungauged catchments: You mention the option of ensemble applications (many sets of parameters). You may elaborate a bit more on this. For my understanding this is a very valuable approach.

An early indication of the value of using ensembles for HBV model simulations can be found in Seibert and Beven (2009, HESS) where we found that the ensemble mean performed generally better than the best single parameter set (see Figure 2). This finding was more recently confirmed in studies such as the two we refer to here (Seibert et al., 2018; Pool et al., 2021). We certainly add a bit more text on this approach.

- Figure 6: The value of a model depends on its purpose!! If the purpose is to look like the reality, the left figure is much "better".

We agree, but the question we posed is 'Which modell will fly?' – and trying to let the left plane fly will not end well ☺

Again, thanks for these valuable comments and interesting thoughts that will help us clarify the manuscript.

Best regards,

Jan and Sten