We would like to thank the reviewer for their thorough response to our manuscript. Their comments will be very helpful to improve our manuscript. We are glad to take the opportunity of this discussion format to address the points they raise.

My main concern about the research is, this study area is not an ideal watershed for doing the experiments. Since we are talking about coupling a forest model with a hydrological model, we better have a catchment that is covered with forest. The five subcatchments of Navizence, have 30% forested area in Chippis (but no streamflow data) and ~15% forested area in Vissoie. The other three catchments have negligible fraction of forested area. That might be part of the reason that we don't see much hydrograph difference between uncoupled and coupled runs even in the most forested catchment Vissoie. In fact the uncoupled-coupled streamflow differences are smaller in the other less forested catchments

The relative fractions of currently forested, potentially forested and never-forested area are shown on Fig. 2. The fractions of forested area are higher than the numbers stated in the review: more than half of the Chippis subcatchment is currently forested (> 25 km^2 out of ~50 km^2), and for Vissoie, this number is about 25% (~15 out of ~60 km^2).

The abandonment of pastures and replacement with forests is an ongoing process in this region, and is expected to continue given current socio-economic trends (see Price et al. 2016). Therefore, the potentially forested area is also relevant, and the behavior of the model for this situation is also tested in this study. For the five subcatchments of this study, the fractions of currently and potentially forested area are ~ 90% (Chippis), 66% (Torrent du Moulin), 70% (Vissoie), 25% (Mottec), 33% (Moiry/Lona).

Considering also the other reasons for selecting this study region, listed in Section 2.2 (p.14f) (strong ecological and hydro-climatic gradient due to elevation differences), we argue that this is a suitable study region for testing the model. Another study is planned in which we apply FORHYCS to several catchments representing different hydro-climatic regions of Switzerland.

It would be great if authors could acquire downscaled meteorological data from latest climate modeling scenarios and test the FORHYC with those time series data. It's not that the delta method (e.g. ±degrees and/or precipitation) isn't scientifically sound, but the climate modeling data provides more variation and insight to the change of climate.

In the framework of an ongoing study, we have acquired the CH2018 scenarios (NCCS 2018), based on the EURO-CORDEX simulations. Given the scope of this study, it does not seem necessary to run FORHYCS with all 39 model chains. Instead, we suggest providing results for the three chains selected by Brunner et al. (2019) as representative for dry, medium and wet conditions (in addition to the delta runs).

P1L12-13: give names of the two new metrics.

The names for these metrics (dH95 and 1-ABC) are introduced in Section 2.2.3.

P2L22: what aspects of mountainous regions are sensitive to what type of global change?

We will be more specific on this in a revised version and mention the following aspects:

- Strong influence of snow storage on hydrology, which may change with warming.
- Warmer and drier conditions may alter species composition in temperature-limited ecosystems.
- Upwards shift of treeline.

P8L7: explain Al, is it the stand leaf area?

Yes – we will include this in the text.

P9L3: define EDI, is it Evaporative Demand Index?

This is a remainder of an old version of the text, thank you for catching this. This refers to the drought index used in TreeMig, and described in the rest of the paragraph. The symbol "EDI" will be removed.

P10L5: again, define fDS.

Vitality reduction function due to drought stress. Based on the comments of Reviewer #1, this passage will need to be clarified. We will also consider this remark.

P14L7: forest minimum stomatal resistance at 180 s m-1 seems to be at the lower end of what has been reported. The stomatal/canopy resistance could be one of the most influential parameters when it comes to estimating transpiration (not sure about this PREVAH model). Why is a single number resistance superior to the previous "minimum canopy resistance for each land cover class"? In any case, the number needs to be justified according to the region and species being applied in this particular study.

This concerns only uncoupled FORHYCS, i.e. a simultaneous run of the hydrological and forest parts of the model without any transfer of information between them. For coupled FORHYCS, minimum stomatal resistance is parameterized according to species and tree height, as described in Section 1.3.2 of the Supplement (which we forgot to reference from the main text – this will be fixed in a new version).

Uncoupled FORHYCS is not meant to represent an improvement from stand-alone PREVAH. Rather, its purpose here is to provide a comparable baseline against which to assess the effect of the coupling. This was not done directly against stand-alone PREVAH for the reason stated in Section 2.1.7 (PREVAH uses a different dataset for soil water storage capacity which is incompatible with the new water balance module).

In another study (Speich et al. 2018a), we found that RSMIN was a rather influential vegetation property for simulated annual water balance (though not as influential as soil water storage capacity or LAI). This was one of the motivations for coupling the models. For uncoupled FORHYCS, the value of 180 s/m was selected as a standard value from the literature (see the cited reference of Guan and Wilson 2009, who base their choice on the review by Körner (1994)). For comparison, in another study in a region nearby

(Peters et al. 2019), stomatal resistances of 128 s/m and 285 s/m were found for *Larix decidua* and *Picea abies*, respectively.

P22L9-23: I've got confused by the streamflow simulations. How was the model calibrated to generate these KGE scores? I assume the PREVAH/FORHYCS were "tuned" to their best performance before the series of experiment. It seems three modeling runs generated model efficiencies that vary from catchment to catchment. Would some other combinations of parameterization make the results look differently?

As Reviewer #1 had a similar comment, here is our answer to their question regarding calibration:

Indeed, there are various parameters related to non-vegetation aspects. Some of these parameters are constant for the whole study area; others are spatially variable and have a different value for each grid cell. The spatially variable parameter values were determined in a previous study (Zappa and Bernhard 2012) based on the regionalization method of Viviroli et al. (2009). The spatially constant parameters were also taken from previous studies. As a different dataset for soil water holding capacity was used in this study (see Section 2.2.2), some parameters were manually adjusted to improve the optical fit of the streamflow lines. Due to the proof-of-concept nature of this study, a full calibration was not undertaken.
In a revised version, we will provide the reference for the spatially variable parameter values, as well as the values used for the spatially constant parameters, in the Supplement.

P23-Figure 5: I'd suggest giving PREVAH a solid light grey line to make it easier to read the difference between FORHYCS00 and FORHYCS11 (which matters more).

OK

P26L8-20: labels from different simulations need to be unified. For example, Succ_TM_BEK is BEK in the Figure 8?

Based on comments from Reviewer #1, we will reconsider the naming of the different runs. We will also take this comment into account.

P33L3-15: looks like a large part of the reason that uncoupled and coupled modeling runs were making rather subtle differences is, the catchment areas are not forested enough for the forest model to pass signals back to the hydrology module. If we look at the hydrologic modeling performance for catchments, Vissoie has more vegetated areas (other than Chippis, which has no observations), thus has the lowest KGE score. The other three catchments, none with meaningful forest fraction, perform much better without interplay with the forests.

It is plausible that the degree of forest cover had some influence for the KGE scores. However, I would also see it in the view of the different soil parameterizations used in PREVAH and uncoupled FORHYCS. Those differ in currently and potentially forested areas, so that lower performance is to be expected in those areas.

It is also worth considering that streamflow from the Mottec and Moiry subcatchments stems to a large extent from the glaciers, the water balance of which the model is able to simulate relatively well.

References

Brunner, MI, Björnsen Gurung, A, Zappa M, Zekollari, H, Farinotti D, Stähli, M (2019): Present and future water scarcity in Switzerland: Potential for alleviation through reservoirs and lakes. Science of The Total Environment 666, 2019, 1033-1047, https://doi.org/10.1016/j.scitotenv.2019.02.169

Guan, H. and Wilson, J. L. (2009): A hybrid dual-source model for potential evaporation and transpiration partitioning, Journal of Hydrology, 377, 5 405–416, https://doi.org/10.1016/j.jhydrol.2009.08.037

Körner, C (1994): Leaf diffusive conductances in the major vegetation types of the globe. E.D. Schulze, M.M. Caldwell (Eds.), Ecophysiology of Photosynthesis, Ecological Studies v.100, Springer-Verlag (1994), pp. 463-490

National Centre for Climate Services (NCCS) (2018): CH2018 — Climate Scenarios for Switzerland Technical Report, NCCS, Zurich (2018)

Peters, RL, Speich, M, Pappas, C, et al. (2019) Contrasting stomatal sensitivity to temperature and soil drought in mature alpine conifers. *Plant Cell Environ*. 2019; 42: 1674– 1689. https://doi.org/10.1111/pce.13500

Price, B., Kaim, D., Szwagrzyk, M., Ostapowicz, K., Kolecka, N., Schmatz, D. R., Wypych, A., and Kozak, J. (2016): Legacies, socio-economic and biophysical processes and drivers: the case of future forest cover expansion in the Polish Carpathians and Swiss Alps, Regional Environmental Change, https://doi.org/10.1007/s10113-016-1079-z

Speich, M. J., Zappa, M., and Lischke, H. (2018a): Sensitivity of forest water balance and physiological drought predictions to soil and vegetation parameters - A model-based study, Environmental Modelling & Software, 102, 213–232, https://doi.org/10.1016/j.envsoft.2018.01.016

Viviroli, D, Mittelbach, H, Gurtz, J, Weingartner, R (2009): Continuous simulation for flood estimation in ungauged mesoscale catchments of Switzerland – Part II: Parameter regionalisation and flood estimation results. Journal of Hydrology 377, 1-2, 208-225, https://doi.org/10.1016/j.jhydrol.2009.08.022

Zappa, M, and Bernhard, L (2012): Klimaänderung und natürlicher Wasserhaushalt der Grosseinzugsgebiete der Schweiz. Technical report. Birmensdorf, Eidg. Forschungsanstalt WSL.