Response to referee comment Massimiliano Zappa

We would like to thank Massimiliano Zappa for reviewing our manuscript and the feedback and the useful suggestions for improvement. We will reply to the comments below.

The reviewer's comments are in **bold**, our response in *italic.*

Dear authors,

I like the manuscript but I have a couple of concerns:

a) Glacier dynamics: C D: You are comparing a state-of-the-art approach (dynamical glaciers), with an approach which is acknowledged to be inadequate in transient mode (constant glaciers). I think you should have a look at the option without glacier change, but tax it as inadequate early enough in the paper (Figure 6) and continue with the dynamical model only. Up to Figure 6 you could also work with a version where you remove the glaciers from the beginning.

>> We agree that constant glaciers are not realistic if used for prediction and discuss this in the manuscript. As the reviewer notes, it is interesting to have a look at this option, since not all hydrological models (at all scales) do yet use a dynamical glacier approach and we want to show the effect of this glacier modelling choice on streamflow drought analysis. In the revised manuscript, we will make clearer from the beginning that constant glaciers are not realistic in transient mode. However, we do think this setting is useful and we therefore do not agree that we should only continue with dynamical glacier modelling from Figure 6 onwards. From Figure 6 onwards, especially in Figure 8, we show that modelling with constant glaciers is actually an interesting benchmark simulation. We can use it to isolate the effect of short term anomalies in precipitation and temperature from the effect of long term glacier changes on streamflow droughts and it thus gives more insight in the processes causing streamflow droughts in glacierised catchments (see also comment to reviewer #1). In relation to this, a version where we would remove the glaciers from the beginning would be a very interesting option, because it would give more insight to what extent streamflow droughts are caused by precipitation anomalies and/or snow melt anomalies (no glacier version), glacier melt anomalies (constant glacier version) and glacier dynamics (dynamical glacier option). However, as mentioned in the reply to reviewer #1, model parameters are calibrated to glacierised catchments and therefore reflect the typical sensitivities and relations among fluxes of glacierised catchments and cannot directly be used to simulate catchments without a glacier.

b) HVT TVT: Here I fear that regime shift (shown by HVT) and actual drought-analysis (shown by TVT) are mixed up. In the discussion part this is presented as finding, but I think this should be stated from the beginning.

>> Interestingly, reviewer #1 had the opposite suggestion, of using only a HVT approach for future drought analysis. We mention in the introduction that using a HVT leads to severe 'droughts' in case of regime shifts (Van Huijgevoort et al., 2014). Despite this, HVT is used in studies to analyse future streamflow droughts at the global scale, which we also mention in the introduction. In the threshold level method droughts are defined as discharges below the threshold, whether HVT or TVT, so purely looking at the definition both are droughts. However, we do discuss this issue of regime shifts and whether regime-shift 'droughts' should also be interpreted as droughts. Given the confusion of both reviewers and the lack of consideration of this drought definition issue in most global future drought studies we infer that it is useful to discuss the effects of certain methods (HVT vs. TVT). We conclude (with the reviewer) that some are better suitable for the drought analysis than others, depending on the focus of a study/what the research interest is. We will clarify this in the revised version (in the discussion).

c) For calibration and validation I would use a seasonally varying discharge value instead of the average discharge as benchmark (see detailed comment in the commented manuscript)

>> This is a good point. We agree that Nash Sutcliffe criterion is not the best objective function to use in areas with strong seasonal discharge. It would indeed be better to use the monthly mean of the observations. However, such an objective function is not available (yet) within the calibration tool of the HBV-light model. We will discuss this issue in the discussion. Also, the objective function that we used is not based on the whole discharge time series, but for 40% on glacier mass balances, 40% on seasonal discharge and 20% on the peak flows. We therefore do not calibrate on the whole seasonal cycle and avoid this problem to a certain extent. But still, using the monthly mean of the observations in the seasonal discharge calibration part would be beneficial.

Please clarify these issues and the other points in the commented PDF.

>> Thank you for the comments in the PDF. We respond to the more substantial comments here below in the online reply. The minor technical/editorial suggestions such as literature we missed are much appreciated and will all be considered and addressed during the revision phase. Thanks also for the suggestions to improve the visualization details, which we will also address in the revision.

- **Sample of two catchments**

>> The two catchments are indeed not particularly dry areas when one uses the definition of arid. However since drought is a relative term, droughts can occur and have impacts in wet regions as well. In this initial study, we focused on these two catchments because of the high data availability, in particular glacier data that we needed to constrain the model simulations. The two catchments are illustrative of the effects of glacier modelling strategy and threshold level method on future streamflow droughts in glacierised regions. We aim to apply the results of this study in further research to areas around the world that are more sensitive to anomalies in glacier melt.

- **Glacier routine**

>> Thanks for commenting on the understandability of the glacier routine together with the Seibert et al. (2017) paper. We hope that this can help settle the original concern by the Editor whether the description was sufficient.

- **Snow routine and snow towers**

>> Snow redistribution and snow towers are not accounted for in the model. The snow redistribution routine, which is included in the model version of Seibert et al. (2017) was not yet available in our model version. Since snow is not redistributed from the higher elevation zones, snow towers are present in the model simulations of both the historical and future period. Snow towers can influence the glacier retreat and also the snowmelt contribution to streamflow. In case snow is redistributed on the glacier, the glacier will melt slower than in our case where there is no additional supply of snow to the glacier. However, the timing of the glacier retreat is also influenced by the various other simplifications in the modelling of the glacier retreat. The storage of snow in snow towers could result in less snowmelt contribution at the end of summer. However, we checked the elevation zones where snow towers occur and compared the amount of SWE stored in the snow towers with the total discharge and found that the influence of snow towers on the streamflow simulation is small (negligible to a few percent). Moreover, for our drought analysis we used a threshold which is based

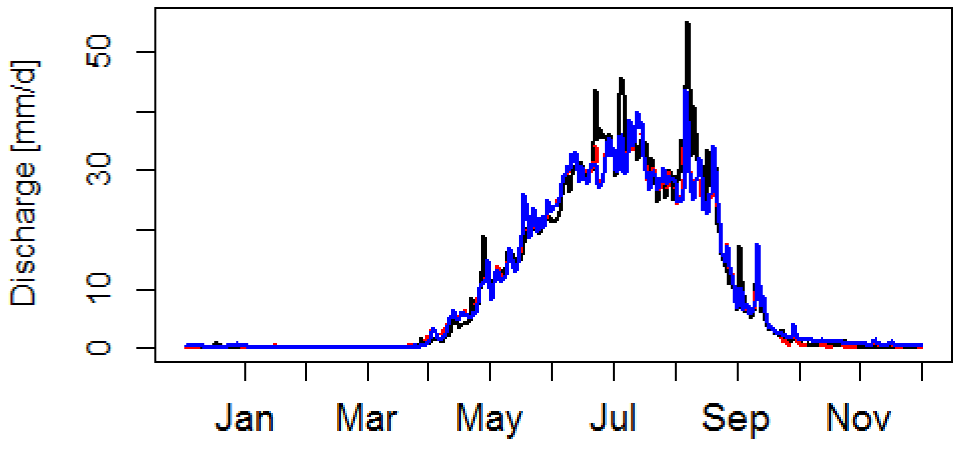
on the simulated streamflow and therefore the small effect of the snow towers is present in both the streamflow and the threshold and it won't affect our drought analysis. When snow redistribution would be taken into account, both streamflow and threshold would have slightly other values (possibly resulting in slightly other drought characteristics), but we expect the same main processes to take place. When comparing the historical and future period (with the HVT) a slight mismatch in streamflow regimes caused by the snow towers could occur, because snow towers are built in the historical period run, but in the future period simulation snow towers also melt in some elevation zones or have disappeared in the far future (2071-2100). However, we think that this effect of the snow towers on the drought analysis is negligible because of the small ratio between SWE and total discharge. We will discuss the snow towers and the possible implications in the discussion in the revised version.

- **Calibration period**

>> For both catchments we wanted to have at least a 10 years calibration period. For Wolverine this period was only available in the period 2005-2014. For Nigardsbreen a longer time series was available and we decided to use the first part of the time series for calibration and the last 10 years for validation. The two study regions are very far apart from each other and have different climate developments. We hence think the benefit of comparable periods is low compared to the benefit of optimal use of available data.

- **Fig. 3 Wolverine**

>> In Figure 3, the observed and simulated regime do indeed not agree very well. However, in this Figure 3 the regime based on observations is only calculated based on 3 years for Wolverine (due to data availability in the historical period), while the simulated regimes are calculated based on 30 years of data. The observed regime is therefore more sensible to extreme years or measurement errors. We can show the agreement between the regimes of observations and Qsim$_o$ for the calibration period (not Qsim$_{cm}$, since they are not available for the calibration period of Wolverine) (Figure below) as inset in Figure 3 in the revised version and explain Figure 3 better.



- **Thresholds in Switzerland**

>> Yes we are aware of studies using HVT to quantify drought in climate impacts studies, see e.g.:

- *Lehner et al., 2006 - Estimating the impact of global change on flood and drought risks in Europe, A continental, integrated analysis*

- *Arnell, 1999, The effect of climate change on hydrological regimes in Europe: a continental perspective*
- *Prudhomme et al., 2014 - Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment*

Thanks for the information on Switzerland using the last 10 years as a threshold/index baseline and thus essentially a moving threshold. We will extend the discussion on this.

<u>References:</u>

Arnell, N. W. (1999). The effect of climate change on hydrological regimes in Europe: a continental perspective. *Global environmental change*, *9*(1), 5-23.

Lehner, B., Döll, P., Alcamo, J., Henrichs, T., & Kaspar, F. (2006). Estimating the impact of global change on flood and drought risks in Europe: a continental, integrated analysis. *Climatic Change*, *75*(3), 273-299.

Prudhomme, C., Giuntoli, I., Robinson, E. L., Clark, D. B., Arnell, N. W., Dankers, R., ... & Hagemann, S. (2014). Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. *Proceedings of the National Academy of Sciences*, *111*(9), 3262-3267.