We appreciate the reviewer's positive evaluation. Thank you for the fast review with useful and detailed comments, which will improve the manuscript. To facilitate some discussion, we will respond below to the general comments and provide a detailed reply to the specific/technical comments with the actual revision.

- Abstract: the reviewer suggests to add the observation periods for the different regions in the abstract. We agree that this is a good idea. However, the different catchments have all different observation periods. For the Alps, streamflow and meteorological data overlap for the period 1961-2015/2016. Still, for Norway and Canada, individual catchments have a different observation period (but all at least 50 years), and therefore it is not easy to summarize that in the abstract. Nonetheless, we could indicate the range of years that were analyzed (1945-2016).

- Introduction: When revising the manuscript, we will try to shorten the introduction - thank you for pointing that out. We will also carefully check again the use of the terms buffering and compensating. In our opinion, these concepts mean something different but are in the literature often used as synonyms when describing glaciers and droughts. We addressed this in the discussion but suggest that we will clarify this difference already in the introduction in the revised version. In our view, buffering means 'something that helps protect from harm', i.e. it is usually a passive process but could also be an active process. In our context, glaciers act as a buffer to (meteorological) droughts, because they still provide water during periods when rainfall input is low, thereby preventing a severe impact on water availability, especially in relative terms. In this respect, glaciers are always a buffer, regardless of the situation, because their water supply (on the short-term) is dependent on temperature and radiation rather than rainfall. 'To compensate', in our view, means 'to provide something to *reduce* the effect of something that has been lost or damaged', i.e. it is an active process. In our context, it means that excess (more than average) glacier melt can reduce the effect of a rainfall deficit on streamflow.
  In the case of buffering, the effects of a rainfall deficit on streamflow can still be present, but streamflow will not be affected as much as without a buffer present, because part of the streamflow comes from groundwater/snowmelt/glacier melt. In case of compensation, excess glacier melt can compensate for the rainfall deficit, and streamflow can be close to average or above normal levels. Since glaciers do not only buffer but also compensate during warm and dry events, we argue (in the discussion) that such a distinction should be made and that this approach helps to quantify this specific role of glaciers. We would appreciate to hear from the reviewer whether we here were able to clarify our argumentation.

- The reviewer has some concerns about how we used measured mass balance data in our analyses. One concern relates to the way we averaged different mass balance observations to have country-wide mass balance time series, and the other concern relates to the use of this country-wide mass balance instead of having mass balance time series for each individual catchment.

  The aim of this analysis was to investigate how event-scale level of compensation relates to variations in annual, winter and summer balances, keeping in mind the different time scales analyzed (one event versus integrated response over the year, or season). Since only a few mass balance measurements are available, and only a few of them have long-term observations, we decided to combine them in one country-wide mass balance time series. This time series was then used to test correlations with *C* of each catchment in that respective country. We do acknowledge the reviewers suggestion on using area-weighted average mass balance on country scale and suggest that we will recalculate region-specific mass balance time series using the same approach. We tested the effect of using area-weighted average mass balances (for Austria and for Switzerland) for the correlations between *C* and mass balances in the European Alps (see Figure below), and observed some small differences. We suggest that we will use the area-weighted average approach in the revised manuscript.

Regarding the use of these regional/country averages mass balance time series, we agree that this is not the best approach. Still, we are rather limited because of limited observations. The ideal option would be to have long-term mass balance time series available for all glaciers in each catchment. In reality we are far from that situation. The reviewer suggests to not average mass balance time series per country because of significant regional climatic conditions. Using geodetic mass balance data, or mass balance data calculated from a model, as done in the studies referred to by the reviewer, is not really an option here, as we need annual variations and would like to base this study solely on observations rather than mixing observations and simulations. We would thus need to find a way to extrapolate the few observations that exist, to all the other glaciers in the studied catchments. The study of Huss (2012) nicely illustrates different options to do that, i.e. arithmetic average, using glacier hypsometry or multiple regression. Although Huss did such an analysis for the European Alps, doing something similar for southwestern Norway and western Canada is beyond the scope of this study. The only option left would be to find out which catchments have a measured glacier in their boundaries that has long-enough time series to do a correlation analyses with the level of compensation. Area-weighted averages will not need to be used then, as it is unlikely that more than one glacier in the same catchment is measured. This will reduce the number of catchments analysed in this part of the study significantly, but may represent better the relation between mass balances and level of compensation and will be worth testing.

- Another general comment of the reviewer relates to the clarification of section 4.2, especially the header and Figure 3. The reviewer is right that section 4.2 only presents WD events; the other events are presented in section 4.3. As the reviewer suggests, we will add this in the title of section 4.2. The catchments in Figure 3 were chosen so that they represent the streamflow responses for catchments with different glacier cover, ranging from low to high. The example years were selected because in those years a warm and dry event occurred in multiple catchments, and in a specific month. We wanted to show an example for an event in September (European Alps 1985 in this case), one in August (Canada in 1981 in this case) and one in July, but found that in 2006 in Norway, events in multiple months occurred, therefore nicely illustrating responses, trends and *C* values for different months and in a catchment with different glacier cover. Figure 3 is meant as an example to illustrate what we analyze in this study and to show the range of responses, regarding streamflow trends during the events, compensation levels, and dependence on glacier cover. We will add the explanation of this selection of example catchments and years to the revised manuscript.

- The reviewer indicates that it is unclear how we calculated correlations between changes in *C* and glacier changes over time, referring to Figure 10. The phrasing here might have caused some confusion, and we will clarify that in the revised manuscript; we do not correlate *C* with glacier changes, but analyze trends of *C* over time. Maybe the relative glacier cover on the x-axis is confusing here. Time trends were calculated for each catchment, so glacier cover, in this case, refers to the trends for different catchments, not to trends in glacier area or glacier cover of the catchment.

  To calculate trends, correlations can be used. For each catchment and each month, we checked if there were 8 or more events occurring in the time series. If so, the correlation between *C* and the year of the event was calculated. A negative value indicates decreasing C over time, and a positive correlation means an increasing level of compensation (*C*) over time. A decreasing trend may be attributed to retreating glaciers.

We agree with the reviewer that in some regions, glaciers were not only retreating but also had intermittent phases of advance (we looked at glacier length data from WGMS and literature). However, since only a few (extreme) events occur in the time series, limiting the time series only to phases of glacier retreat, will result in too few events to do a trend analysis. On page 22 L438-449 (discussion, 5.5), we will add a discussion on the phases of glacier advance and their possible effects on trends in C. These effects of phases of glacier advance on C may, however, not be straightforward. Also, an event can occur in this advancing phase, or not because these periods are generally characterized by colder temperatures, possibly influencing the effect of such phases on the overall trend.

- The last general comment relates to the use of ‘Norway’ and ‘Canada’, while actually only specific regions within these countries are analyzed, namely southwestern Norway and western Canada. We agree with the reviewer and will change ‘Norway’ and Canada’ into ‘southwestern Norway’ and ‘western Canada’ in the revised manuscript.

Specific comments:

We thank the reviewer for all detailed comments and technical corrections which will greatly help to improve the paper and its clarity. We will consider them to revise the manuscript carefully.

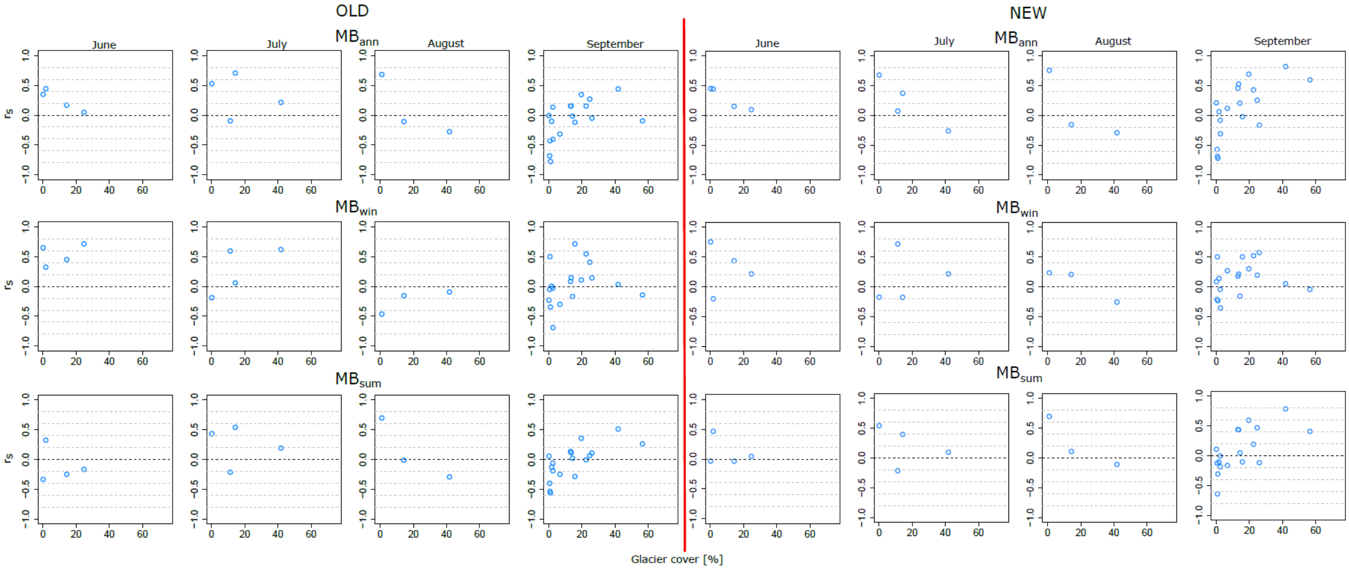

Figure 1 Comparision of the correlation between country-wide mass balances and level of compensation (C), left: using the median to average mass balance observations and create a country-wide mass balance time series (Figure 9 in the manuscript), right: using area-weighted average to calculate a country-wide mass balance time series. Results are shown for the European Alps.