Reviewer 1: Kristian Förster

P1L17: Here, I would recommend to provide the annual runoff in mm per year too, since it helps to compare the values with other studies.

Response: This is a good suggestion, the values in m/yr have been added to the abstract and the conclusions in the final draft.

P1L23: Please provide an explanation for the abbreviation CTD. In the current version of the manuscript, it becomes only clear on page 5.

Response: A correction for this abbreviation has been added to the final draft.

P3L9: What do you mean by "large uplift rates"? Please be more specific.

Response: We are referring to the regional uplift rates for Glacier Bay and southeast Alaska from isostatic rebound caused by glacial wastage since the little ice age. See Larsen et al. (2005) for more information. For clarity, in Section 1, ¶6 we changed large to *rapid* and added *from isostatic rebound* to the final draft.

Larsen, C.F., Motyka, R.J., Freymueller, J.T., Echelmeyer, K.A. and Ivins, E.R., 2005. Rapid viscoelastic uplift in southeast Alaska caused by post-Little Ice Age glacial retreat. *Earth and Planetary Science Letters*, *237*(3-4), pp.548-560.

P3L16: Here, you explain that the model output is available as daily output. What is the internal time step of the (energy balance) model? P4L25pp.: Here you could provide some more details on the time step of the model. Since it is an energy balance model, I would expect sub-daily time steps (even though the output is daily).

Response: The internal timestep of the model is 3 hrly and the results have been aggregated to daily and monthly for the climatological analysis. A sentence about the timestep information has been added to the model Section 3.1, ¶3.

P5L15p.: ET is computed by SnowModel? Was there any attempt to compare the results with the MODIS data – at least for reasons of plausibility, given that there is a mismatch in scales between MODIS and SnowModel?

Response: ET of the land surface is not calculated by SnowModel. When snow is present in the grid cell, sublimation of the snowpack is calculated by the energy balance sub-model (EnBal) and sublimation of blowing snow is calculated by the snow transport sub-model (SnowTran-3D), see Liston et al. (2006) for a review of all of the model physics and subroutines. It should be noted that the snow transport model is not recommended for model resolutions above 100m, and since our model resolution is 250m, SnowTran3-D was not utilized for the simulations.

However, obviously ET does make up a portion of the water balance and we therefore use the MODIS ET dataset to supplement the results of our SnowModel simulations. Hill et al. (2015) estimated the ET component of annual runoff for the entire Gulf of Alaska region to be ~17%. Beamer et al. (2016) estimate ET for the Gulf of Alaska region to be 10-15% less than Hill et al. (2015) results. Since much of the GBNPP domain is glaciated or covered in snow for many months of the year, the authors decided to simply estimate ET values from the MODIS ET dataset for the historic simulation time period and spatially subset by watershed or grouped watershed. See section 3.2.3 for more information, as well as the section below in our responses regarding another ET question by Reviewer1. We find that the MODIS based ET values range from 5%-13% of annual runoff in the GBNPP watersheds and we've added a table with this information below. The authors decided to add this table (new Table 4) to the manuscript to clarify the ET process and results.

Watershed Name	Historic MODIS ET (m/yr)	Percentage of Annual Precipitation (%)	Adjusted Annual Runoff (m/yr)
GBNPP	0.3	9	3.1
North	0.3	9	2.8
West	0.2	5	4.1
West-Arm	0.2	5	3.2
East-Arm	0.3	9	3.3
Tarr	0.2	3	2.7
Carroll	0.2	8	2.7
Dundas	0.4	9	2.7

Beamer, J.P., Hill, D.F., Arendt, A. and Liston, G.E., 2016. High-resolution modeling of coastal freshwater discharge and glacier mass balance in the Gulf of Alaska watershed. *Water Resources Research*, *52*(5), pp.3888-3909.

Hill, D.F., Bruhis, N., Calos, S.E., Arendt, A. and Beamer, J., 2015. Spatial and temporal variability of freshwater discharge into the Gulf of Alaska. *Journal of Geophysical Research: Oceans*, *120*(2), pp.634-646.

P5L27: A new subsection 3.3.1 is introduced in section 3.3 but there is not any other subsection (e.g., 3.3.2 etc.). I was wondering if it is worth to merge the sections 3.3 and 3.3.1?

Response: This is a good idea, and we merged these two sections in the final draft.

P6L19: The term "forecast" is used throughout the manuscript to describe the scenario data and the corresponding results. I am not sure if this term is correct in this context. I would suggest using "projection" instead since this term acknowledges additional uncertainty involved in climate scenarios which arise from uncertain greenhouse gas emissions (i.e. external forcing that is not exactly known). For instance, in a recent paper we also used the term projection to highlight this type of forcing (Hanzer et al., 2018). In contrast, according to Kirtman et al. (2013), the term forecast refers to initialized climate model runs (e.g., seasonal to decadal predictions, see their Box 1.1, or http://glossary.ametsoc.org/wiki/Climate_prediction).

Response: The difference in these terms is important, and the authors agree that the term 'projection', as defined by Reviewer1 above, is more in line with the intention of the use of 'forecast' in the original manuscript. We are not attempting to make a climate prediction (as defined in the provided weblink) of what will happen in the future in Glacier Bay. We are modeling one of the potential scenarios that may occur in the hydrology of the region that would accompany the RCP8.5 emissions scenario. For these reasons, the authors have chosen to change the term 'forecast scenario' to 'projection scenario', in every instance, throughout the manuscript. These changes are added to the final draft.

P6L30: Here, I would suggest to add some thoughts why you have selected RCP8.5 only. It is clear that running impact models for numerous RCPs is expensive in terms of computational costs. However, you could argue that you are interested in a worst-case.

Response: See Section 3.5.1, ¶2 for text added as an explanation for why we chose the RCP8.5 scenario. The modified paragraph is below, added text is italic.

"Although future climate simulations *from SNAP* exist for numerous RCP (representative concentration pathway) scenarios, in this study we restrict ourselves to the RCP 8.5 scenario and to the 5-model mean. *The other RCP scenarios (RCP 2.5, RCP 4.5, RCP 6.0) represent concentrations of greenhouse gases (GHGs) in the atmosphere that peak earlier in the 21st Century or at lower levels of GHGs than the RCP 8.5 scenario. Keep in mind that the choice of the RCP 8.5 scenario is not an attempt to evaluate the likelihood of the future GHG concentrations. Rather, we use the RCP 8.5 scenario for the projection scenario because we are interested in the hydrologic changes that might occur in the worst-case scenario.*"

P7L14: Does it mean that you did not apply SnowModel to future periods, e.g. by forcing the model with modified MERRA data (scaling of meteorological forcing)? From your explanations, you compared the historic run from SnowModel with the future simulation of Beamer et al. (2017) in terms of long-term averages on runoff. If I understood this correctly, this would suggest a simple approach that contradicts the first line of your abstract (". . . is

used to estimate current and future runoff into Glacier Bay.”). I would encourage the authors to provide more details on the setup of future scenarios.

Response: In section 3.5.3 Future Climatologies, we discuss creating the projection scenario climatologies. First, a review of the process. Beamer et al. (2017) conducted a SnowModel historic and forecast simulation for the larger Gulf of Alaska study area. The authors subset the GBNPP study area results from the more spatially extensive Gulf of Alaska simulations. These SnowModel simulations were forced with CFSR for the historic and projection scenarios, for the entire model space (at the lower 1km resolution) and complete model timeframe (3hrly). Climatologies were created by temporal averaging and spatial aggregation into the GBNPP watersheds or grouped watersheds.

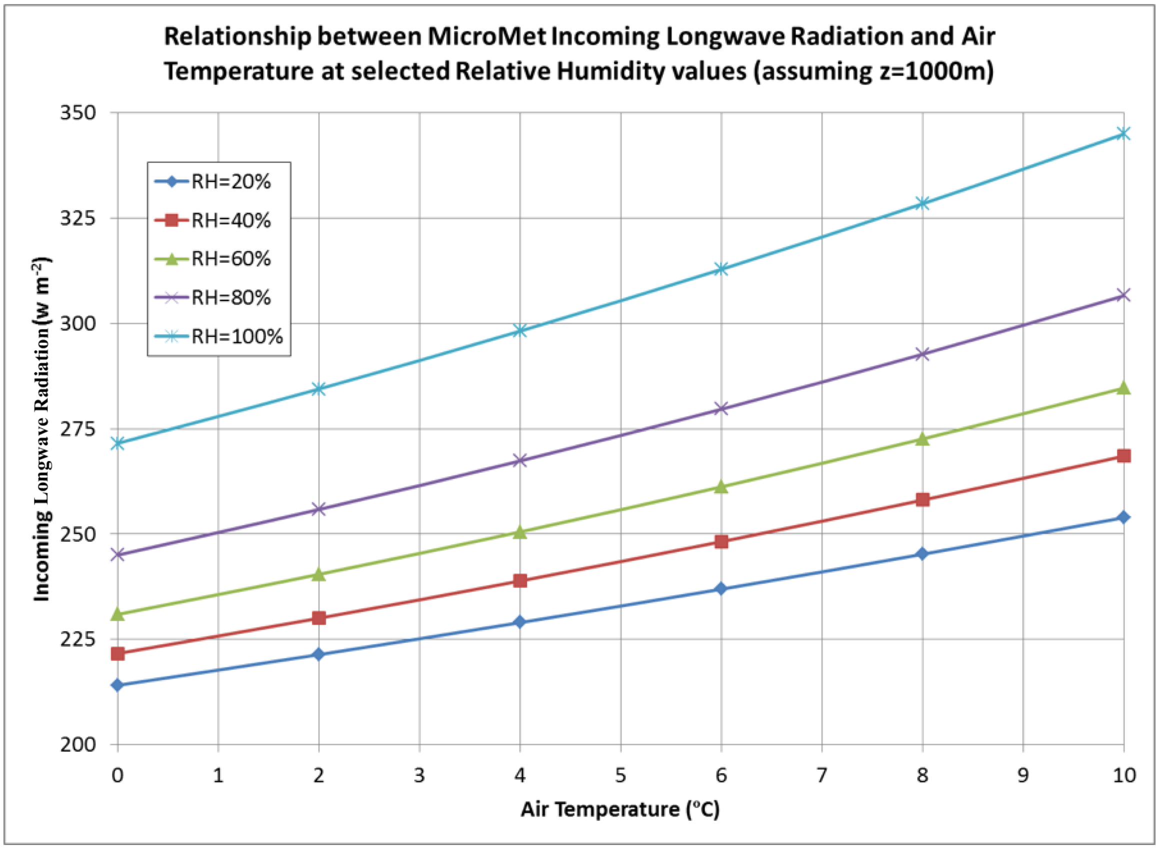
Next, a full SnowModel historic simulation for the GBNPP study area was conducted at the higher spatial resolution of 250m, forced by the MERRA reanalysis product. Climatologies were then created by temporal averaging and spatial aggregation by watershed or grouped watershed. At this point, we have spatially and temporally averaged climatologies for each watershed or grouped watershed (1. CFSR-based historical climatologies, 2. CFSR-based projection scenario climatologies, 3. MERRA-based historical climatologies). We created a scaling factor from the CFSR historic and projection scenarios to apply to the MERRA historic climatologies. After the application of these scaling factors we have the MERRA-based projection scenario climatologies by watershed or grouped watershed.

This study presents projection scenario results as 30-year climatologies. We do not present results related to the frequency characteristics of the runoff. As such, we are not presenting results on frequency distributions, or peak flows, etc. For long-term characteristics of runoff, however, we believe that our approach is appropriate because changes in runoff are driven by long-term changes in precipitation and temperature, which vary relatively slowly in space, and these changes are preserved in the scaling from CFSR-based historical climatologies to CFSR-based projection scenario climatologies.

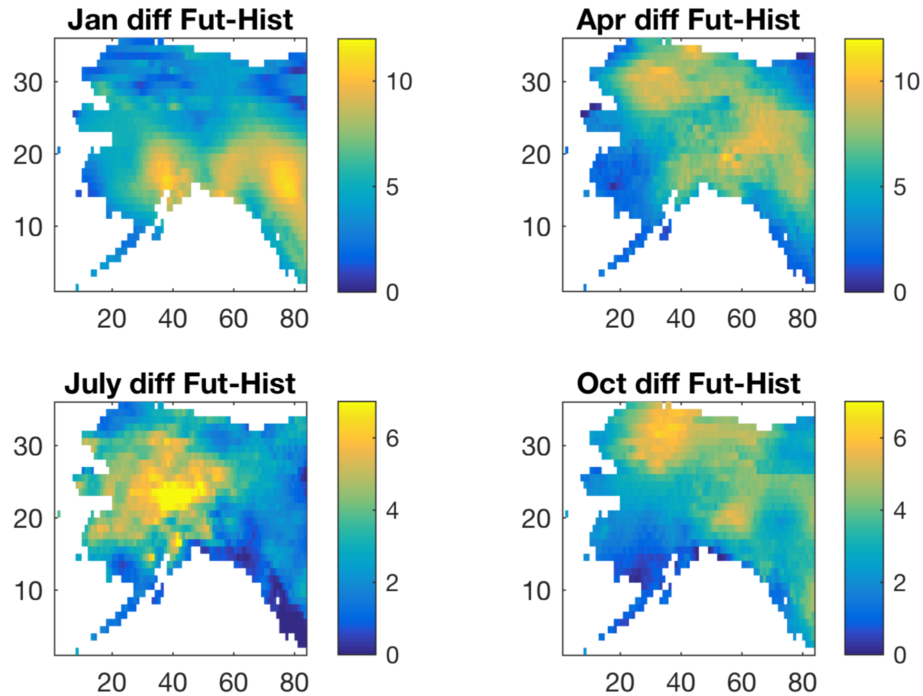
P12L32p.: I was wondering why only temperature and precipitation have been considered, given that SnowModel requires additional meteorological quantities?

Response: This is a decision based in part on the fact that only temperature and precipitation anomalies from SNAP for the RCP scenarios are available for the AK region. While MicroMet inputs include relative humidity, wind speed, wind direction, shortwave and longwave radiation, and surface pressure, those have not been modified in the projected scenario. In order to investigate the impacts of leaving out any potential changes in other variables, such as humidity, the authors considered the changes in relative humidity for their Gulf of Alaska results and the results from this additional analysis are below.

First, it's important to understand that MicroMet will modify the radiation balance through temperature increases. See the figure below, where its obvious that longwave radiation changes substantially with both Temperature and Relative Humidity. Since we already have a good idea of the expected increases in Temperature from SNAP (from the ΔTemp analysis discussed in our Results Section 4), the question becomes how different might relative humidity values be in the projection scenario?



To answer this question, we turned to the VEMAP project (https://daac.ornl.gov/VEMAP/guides/VEMAP_Alaska.html) which provides low resolution downscaled monthly grids of many variables including relative humidity. These grids come from two climate models, and we looked at the results from one (HadCM2). The results provided were for the GHG+A1 scenario, sometimes referred to IS92a. This scenario is quite comparable to the SRESA2 scenario which, in turn, falls between the RCP6.0 and RCP8.5 projection scenarios. We computed climatologies of relative humidity for the periods 1966-1996 and 2070-2100 for four months of the year (Jan, Apr, Jul, Oct). The figure below shows the differences in relative humidity between the two climatological periods (Future – Historical).



In the figure above, we see that Relative Humidity is projected to increase throughout the Alaska region, including the Glacier Bay region in the lower, right portion of the figure. In July and October, the increases are typically less than five percent. In January, the increases throughout Alaska are in the range of 5-10 percent. The first figure above, which shows longwave radiation as a function of temperature and RH, suggests that the ~4 deg C temperature increases that are predicted by RCP 8.5 produce a change in longwave of about 25 watts per square meter. An increase in RH of about 5-10 percent appears to produce a smaller (~10 watts per square meter) change in longwave. So, while it would be conceptually preferable to adjust all weather forcing variables, it appears that the expected changes in temperature dominate. This fact is what led us to adopt the future SNAP climatologies for Temperature and Precipitation as a good 'leading order' study of the effects of changing climate with the RCP8.5 projection scenario.

Additionally, the focus of this manuscript and scientific questions is not changes in climate variability, weather extremes, or high runoff events, but rather longer-term, climatological averages. The authors think it's not within the scope of this project to create our own statistically downscaled, projection scenario weather variables (RH, shortwave and longwave radiation). Since we are primarily focused on the climatological averages of the model output, we find the choice of perturbing the temperature and precipitation inputs for the projection scenario to be adequate, and aligned with the methods summarized in Beamer et al. (2017).

P12L36: The validation could be done in a quantitative way too. The only linkage between your results and oceanographic data is provided on page 12, lines 7 to 8 (by comparing Figure 7a with Figure 9). Since the model calibration is done for another region (indeed, in which your region is included), I would expect a closer look on this dataset, since it is the only dataset available for assessing model accuracy.

Response: The authors decided to use the oceanographic dataset primarily in a qualitative way because the of the complex, understudied open boundary of the bay system, where water (fresh and salt) moves freely in and out of the boundary into Icy Strait, the Cross Sound, and eventually the Pacific Ocean. Critically, freshwater fluxes are not measured or analyzed at the mouth of Glacier Bay, and to the best of the authors' current knowledge, a dataset that includes these fluxes entering and exiting the system does not exist. Thus, we are not able to explicitly determine what component of FWV is sourced by freshwater runoff as opposed to fluxes of highly stratified water at the bay's inlet. Therefore, we've chosen to focus on the change in freshwater volumes from month to month within the bay, instead of quantifying total bay freshwater volume at any given time. Given the sharp temporal gradients in freshwater runoff, the oceanographic dataset is presented to qualitatively assess whether the FWV signal shows a strong temporal gradient that is synchronous with freshwater runoff predicted by the model.

P22L9 (Figure 7): Why do you plot ET derived by MODIS only, given that your model accounts for ET too? If ET computations are available for the model too, you could plot ET for the future scenarios as well. In my opinion, analyzing changes in ET would be an interesting asset to describe the hydrological change.

Response: In Figure 7 and Appendix A we plot monthly ET values derived from MODIS because SnowModel calculates sublimation of the snowpack when solving the energy balance equations but does not calculate ET from the land surface when no snowpack is present. See previous answer on ET for more details. This is why Beamer et al. (2017) added the SoilBal sub-model to their analysis of the Gulf of Alaska SnowModel simulations. Many other previous studies using SnowModel from the Arctic, Patagonia, Greenland, and Alaska do not calculate ET using an additional sub-model, and this manuscript is no different. (Mernild et al., 2012; 2013; 2014; 2017a)

Our intention in plotting the MODIS derived ET values on the runoff plots (Figure 7 and Appendix A) is to give the reader an estimation of how much the monthly runoff would be altered if ET were included. See the explanation in section 3.2.3 for more details. We temporally average (14-year), and spatially aggregate the MODIS ET values for each watershed/grouped watershed. We do not subtract these monthly values from the partitioned climatologies (snowmelt, glacier melt, rain runoff) because the model does not resolve which of these sources would be the appropriate origin of the ET-based water. Even more importantly, in the projection scenario, we have no land cover evolution beyond a simple estimation of glacier area change. If glaciers continue to recede, as they have over the last several hundred years and are projected to change in the future in Glacier Bay, these changes will continue to alter the landscape, in both landcover species and landcover type designations. These landcover changes would inevitably cause changes to ET in the future. We admittedly make no attempt to quantify or characterize these types of landcover changes, nor the subsequent ET changes, because it is outside the scope of our current project and analysis.

To clarify the ET results, the authors will be adding this table in the final draft of the estimated monthly ET values from MODIS by watershed and we include the percentage of annual runoff and the adjusted annual runoff values. We have also added the adjusted runoff values in the abstract and conclusion paragraphs for clarity. Note: these are historic values because we do not estimate ET for the projection scenario.

Watershed Name	Historic MODIS ET (m/yr)	Percentage of Annual Precipitation (%)	Adjusted Annual Runoff (m/yr)
GBNPP	0.3	9	3.1
North	0.3	9	2.8
West	0.2	5	4.1
West-Arm	0.2	5	3.2
East-Arm	0.3	9	3.3
Tarr	0.2	3	2.7
Carroll	0.2	8	2.7
Dundas	0.4	9	2.7

Mernild, S.H. and Liston, G.E.: Greenland freshwater runoff. Part II: Distribution and trends, 1960-2010, *Journal of Climate*, *25*(17), pp.6015-6035, DOI: 10.1175/JCLI-D-11-00592.1, 2012.

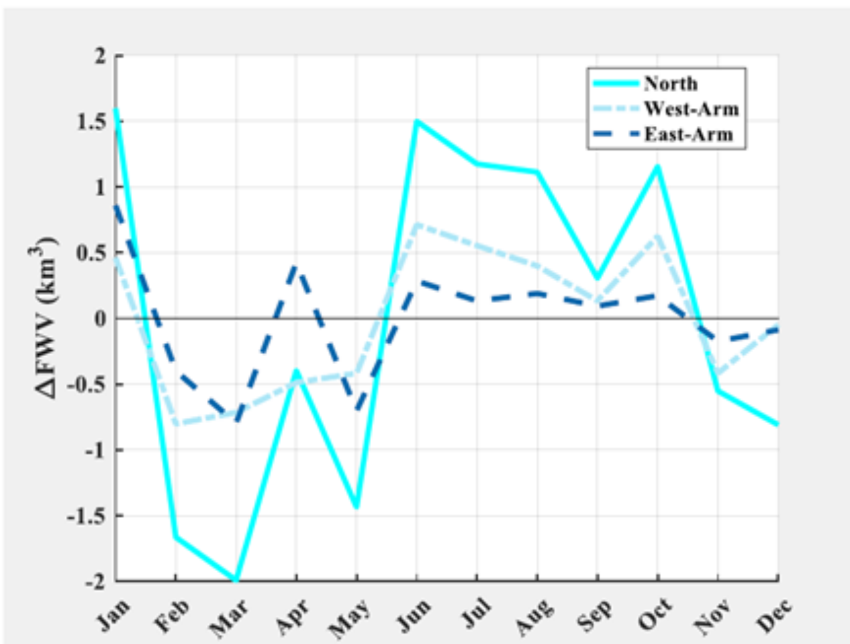
Mernild, S.H., Lipscomb, W.H., Bahr, D.B., Radić, V., Zemp, M.: Global glacier changes: a revised assessment of committed mass losses and sampling uncertainties, The Cryosphere 7, 1565–1577, DOI: 10.5194/tc-7-1565-2013, 2013.

Mernild, S.H., Liston, G.E. and Hiemstra, C.A.: Northern hemisphere glacier and ice cap surface mass balance and contribution to sea level rise, *Journal of Climate*, *27*(15), pp.6051-6073, DOI: 10.1175/JCLI-D-13-00669.1, 2014.

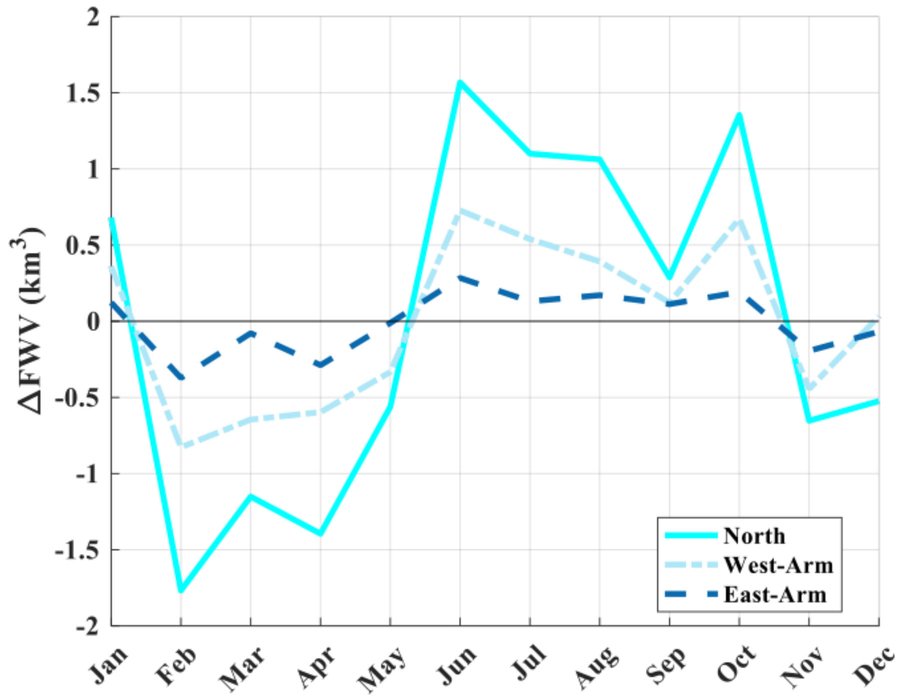
Mernild, S.H., Liston, G.E., Hiemstra, C.A., Malmros, J.K., Yde, J.C. and McPhee, J.: The Andes Cordillera. Part I: snow distribution, properties, and trends (1979–2014), *International Journal of Climatology*, *37*(4), pp.1680-1698, DOI: 10.1002/joc.4804, 2017a.

P23L10 (Figure 9): Why do we see a maximum in delta FWV in January? I would at least expect a brief discussion on that maximum in the text.

Response: This is a good question, and we decided to look at these specific months (Jan and Dec) in our analysis below. For clarity, the January value ΔFWV_jan in Figure 9 is equal to FWV_jan – FWV_dec. In assessing the certainty in this signal, it is important to consider: (1) Winter is vastly under-sampled as compared to other seasons, and (2) there can be great variability between monthly FWC from year to year. The authors realized that because of these two factors, some months were excluded in the FWV extrapolation for all months. After incorporating these excluded months into the analysis, we made changes to Figure 9 that is more representative of all months in the entire dataset. See the old and new versions of Figure 9 below. Some of undersampled winter monthly FWV values have changed slightly, the more highly-sampled summer monthly FWV values remain largely intact, and the ΔFWV_jan value is dampened in the final Figure 9. In the revised figure, it is evident that the strongest signals in monthly changes in freshwater occur in the summer (May to October). Although the figure does suggest that there is an increase in FWV from December to January, it must be weighed in the context of the uncertainties described herein.



Old Figure 9



New Figure 9

P27 (Table 3): It would be helpful for the readers to have a separate column for each existing column which provides the runoff in mm too. In your text, you already highlight the benefit of using specific runoff for reasons of comparison.

Response: A new column has been added to Table 3 in the final draft with the specific runoff in meters for the historic and projection scenarios.

P7L17: There is no Sect. 3.4.1.

Response: Section 3.4.1 has been removed in the final draft.

P10L: Figure 9 does not show any trends. Why do you plot the delta in FWV instead of FWV?

Response: The word *trends* in Section 4.6, ¶1 has been changed to *seasonal timing of changes in freshwater*, which is a more precise wording of the sentence.

P11L12: Please correct the reference to the figure (there is no Fig. 9a).

Response: This mistake has been removed in the final draft.