I have been recruited late as a reviewer, after initial reviewers were unable to come through, so in the interests of an expedient return I will offer a relatively quick and high-level assessment. I note that R1 has been incredibly thorough in her/his response, and I agree with many concerns expressed in that review. My comments will build off of that discussion.

Overall assessment:

The manuscript makes the most out of a limited but valuable data record, bias-adjusting NARR climate reanalyses from 1979-2016 to develop a point-scale meteorological forcing relevant to the AWS location adjacent to Saskatchewan Glacier. This point 'data' is then distributed or extrapolated over the glacier to force a surface energy balance model for glacier melt. Combined with NARR-based precipitation estimates, mapped onto the glacier using similar methods (bias-adjusted, then distributed over the glacier using a lapse rate), this provides an estimate of glacier mass balance from 1979-2016. Perturbation methods are used to assess energy and mass balance sensitivity to increases in temperature and precipitation, representative of future climate change scenarios.

There is good potential in this manuscript and I found the methods to be mostly well-explained, although I agree with R1 that inconsistencies in terminology, notation, and units need to be improved. Many previous studies have used bias-adjusted climate reanalyses or climate model output to force distributed surface energy and mass balance models on a glacier, so this is not particularly novel or innovative. However, there are several judicious choices that seem to me well-judged, such as the assignment of glacier albedo and its variability (uncertainty), based on satellite images, the inclusion of available precipitation data to inform the NARR bias adjustments, and the introduction of diurnal temperature lapse rates on the glacier based on the vertical transect of HOBO temperature sensors. Available in situ data is limited but is well-leveraged. The numerical experiments are well-designed and nicely visualized, and the authors reach significant and (mostly) well-supported conclusions. This is a significant and interesting outlet glacier of Columbia Icefield, an important glacier system in the Canadian Rocky Mountains which nonetheless lacks in situ surface mass balance data. I therefore think this study has potential value, and would be interested to see a revised manuscript that addresses several of the major concerns of R1. I have several additional points that also need to be
addressed. Within these, it may just be that I did not understand or follow the authors' methods. In some places though, I am not confident in the results (see points 5 to 8) and would encourage a careful re-examination. In particular, the reported AWS wind speeds are implausibly high and the modelled turbulent fluxes also do not seem correct; sensible heat fluxes are extremely high and the finding of positive latent heat flux is possible but unusual. The values reported here are not typical of mid-latitude glaciers in a continental environment (point 8).

Major Concerns

1. I don't think it is unreasonable to use the single NARR grid cell that coincides with this glacier, given the 32-km scale of NARR grid cells and the good correlation with available ground T, RH, and shortwave radiation data. I do wonder why NARR incoming longwave radiation data was not used in this study though, for consistency. Was this considered, or compared with the longwave radiation parameterization that is used? The parameterization requires cloud cover fraction that is taken from NARR, so why not just use incoming LW from NARR, which appropriately considers the 3-hourly vertical temperature, humidity, and cloud structure, vs. a parameterization that only uses near-surface values?

2. The authors note that ERA5 was not available at the time of the study, but it has been out for more than one year now - about 1.5 years I believe, and offers 1/4 degree resolution with hourly data, which can avoid some of the complexities and assumptions in mapping the 3-hour NARR data onto hourly estimates. ERA-land is even higher resolution. I would not insist on this, as it is a lot of additional work, but I think it would be valuable and would strengthen the manuscript a lot to explore model results with ERA5 and/or ERA-land as well - this could very much help to test the robustness of the results and conclusions, and lower the reliance on NARR's relatively unproven veracity in the complex mountain terrain. This could be follow-up work, but I also have the nagging sense that this study risks being already out of date. The argument that ERA5 precipitation is questionable is not really valid, as all of the reanalysis output is bias-adjusted and NARR is weak in this respect.

3. I agree with R1 on the confusion regarding terminology in the methods. As I read it, the NARR output is not downscaled to the AWS; this study uses bias-adjusted rather than downscaled NARR output, as I understand it. Bias-adjusted NARR fields are then distributed over the glacier from a reference site (the AWS) using locally-relevant lapse rates and sophisticated methods for the incoming SW radiation. As I understand it, the reference site for the precipitation differs (e.g., Parker Ridge), but it is a similar approach. I am happy to leave it to the authors to decide how they would like to refer to this process (extrapolated, downscaled, lapsed, or distributed over the glacier), as long as it is clearly defined and consistent in the manuscript. It is not what I think of or what is commonly referred to as downsampling though.

4. I admire the use of the regional network of permanent weather stations to develop temperature and precipitation lapse rates, but I worry about the relevance of these values to the glacier itself. These are all off-glacier sites with a maximum elevation of 2025 m, while Saskatchewan Glacier extends from about 1800 to 3300 m. Glacier near-surface temperatures (and the surface energy balance that influences these) are very specific, as is the snow accumulation regime on glaciers and in the unsampled elevation band from 2025-3300 m. I don't have great confidence in the applicability of the lapse rates as determined by the off-glacier climate station network. For temperature, why not use the average daily or monthly lapse rates as determined by the HOBO temperature transect? I realize that these are summer-only, and the data are limited, but this is what is used for the diurnal lapse rates so this would seem relevant and consistent. Winter temperature lapse rates are not important to the glacier melt, so could be assigned an average or May
value. For the precipitation lapse rates over the glacier, is there a way to use available winter mass balance data (in situ and/or LIDAR-inferred) to look at this? The current precipitation lapse rate may be appropriate, but it would be helpful to constrain and evaluate this, as well as the assumption of a sustained (and strong) linear increase in precipitation across the icefield plateau from 2800 to 3300 m.

5. The precipitation lapse rate that is used is based on the reference climate station data from November to March (l.206). This does not coincide with the accumulation season on the glacier, which is more like September to May. Is this same precipitation lapse rate used for April to October, and is there objective support for that? This needs to be discussed and addressed, perhaps with an examination of the primary data or perhaps by bringing in the winter mass balance data from the glacier, if there is some from the 2014-2016 study. November to March is relevant for the lower-elevation snow season, but not that of the glacier, where autumn and spring often bring a lot of snow.

6. Wind speed results on ll.393-395. These are extremely high average wind speeds, an annual average of 16 m/s and up to 23 m/s in February. I appreciate this is likely a windy site, and there are katabatic winds here, but those are typically stronger in summer. Are the authors confident that these units are correct - is this perhaps km/hr, or are these maximum (vs. mean) wind speeds that are reported here and plotted in Figure 2? An average monthly wind speed of 23 m/s equates to 83 km/hr, which is not plausible. Values reported and used later in the manuscript (e.g., from NARR, means closer to 5 m/s) are more reasonable. I would also add that I have spent some time on this glacier, and there is a steady and reliable down-glacier wind, but not of the knock-you-over variety.

7. I am not sure what 'homogenized' means here in the context of the observational precipitation records that are spliced. Homogenized has a very specific meaning for meteorological data sets, involving corrections for discontinuities associated with station moves or changing conditions/instruments/methods at an observation site. The precipitation data also seem to have a lot of gaps, which makes me worry about the time series of mean annual values. It seems best from about 1972 to 1994, not for the full period plotted in Figure 2. What methods were used to gap-fill this data for missing months? Apologies if I missed this. My sense is that it would be best to use these data for long-term mean monthly values from 1979-1994, using all available monthly data over this period. This can then inform a bias-adjustment of NARR mean monthly values for the same period, 1979-1994. Then go with bias-adjusted monthly NARR (or ERAS) precipitation for the study. Just my surficial thoughts on inspection of the observational data in Figure 2.

8. Perhaps my most significant concern: the sensible heat fluxes seem far too high for a mid-latitude continental glacier, and compared with other data from the region (Peyto, Haig Glaciers). Also, it is surprising and unusual that latent heat fluxes are positive. I don't trust either of these results. Are the erroneously high winds speeds (point 6) the reason for this? This could explain the high values of sensible heat flux, though it is still odd that latent heat flux is positive. What is the basis for determining the snow/ice surface temperature in these calculations? This is critical to the turbulent heat flux calculations, and I did not see a discussion of this in the paper - apologies if I missed it. Is a melting glacier surface assumed in the summer? What is assumed through the rest of the year? I wonder too if the snow roughness value is appropriate for winter conditions - 6 mm is high, perhaps more reflective of sun cups than the smooth winter and spring snow surface. Snow roughness values closer to 1 mm are commonly adopted in glacier modelling. The sensitivity to this variable could be more thoroughly explored, perhaps considering order of magnitude rather than \pm 1 mm variations.

My sense of what is 'normal' is based on mid-latitude glacier turbulent fluxes reported in
many prior studies, e.g.:


In the Rockies:


9. AWS snow accumulation is reported on l.485. How is this recorded? Isn't this just a tipping bucket rain gauge at this site? Or is this based on measurements from site visits? It would be helpful, per the comment above, to report winter snow accumulations on the glacier and use this data to help evaluate the precipitation modelling.

10. Agree with R1 that it will be really helpful to use mm or m w.e. throughout for mass balance, rather than a mix of mm, cm and m.

11. l.494, the balance gradients. Please give in units of m or mm w.e. per m. This is an interesting result, though I worry that the unusually high value (steep gradient) on the upper glacier is in part due to the unconstrained precipitation/accumulation gradient on the glacier. Looking at the available data in Figure 5 from 2015 and 2016 (2014 data are not sufficient), it would be hard to justify a bi-linear vs. linear relationship for ba. This is purely a model result then, as I understand it - can it be explained via mass balance processes here, since the authors describe the balance ratio as an unusual result? Is it reflected in the geodetic mass balance profiles? This is a significant point, as the balance gradients (values, linear vs. bi-linear) are potentially significant for regional-scale mass balance modelling - I can imagine other authors using the values that are reported in this study. Perhaps it is early to think about this too much, as the results may change upon revisiting of the wind speeds, modelled surface temperatures, and modelled turbulent fluxes.

A few minor points, not comprehensive

l.106, 1970s

l.108, I think this should be negative for the net mass balance
1.189, GSC not defined I think

1.305, I always think of the terrain radiation as coming largely from valley walls, not the glacier.

1.485, AWS snow accumulation - is this in w.e.?