RC1:

General comments:

This paper presents a dataset in the Upper Susitna Basin in Alaska and contains the data itself, the data collection and data processing of meteorological, glacier mass balance, snow cover and soil measurements. Data collection in this region is hard and this data can be great value for model validation or calibration in this region, however, I think not in the way it is framed now by focussing on implications for the dam (that is not there) and climate change (data is too limited).
We've made a set of changes to address the framing (see below), but the dam proposal was the main motivation for the project (and the reason the project was funded) so we chose to leave this mention in the intro. The wording makes it clear that the paper focuses on the data, this section of the intro is providing the context for why the data was collected. We added phrases about data scarcity and model validation and calibration to bolster the motivations.

I think the dataset presentation can be more thorough and better structured than is presented now. I miss at various places context of statements and the naming of the different stations can be structured better, such that it is clear to the reader which data you are referring to. I specified this in the specific comments below.
 Addressed specifics below

I suggest major comments, predominantly since I think the focus of this paper should be on the data and not too strongly framed to dam-implication work and/or climate change work, since this work does not address that.

Specific comments:
The introduction does not focus on the relevant subjects. You present solely a measurement dataset and you focus in the introduction on climate change and (modelled) river runoff, which you did not do in the paper. Please restructure the introduction, remove this information or at least shift focus to the data. You could include more information about previous field works/data sets in this region?
We believe this larger context will be important to many readers. For people unfamiliar with the area, this gives them enough information that they don't have to search elsewhere for an understanding of why our data might be interesting or useful.

In general: update the captions of figures, such that those are complete. In Figure 1 I miss for example explanation of the subpanels. Also update the figure labels and text inside the figures such that those are readable (mainly Figure1) and resolution is high enough (also for the tables).
Added to caption of figure 1: "The main map focuses on the glacierized portion of the basin, the large inset shows the whole Upper Susitna basin which drains to the proposed dam site, and the small inset shows the basin in the context of the state of Alaska."

P1L16: You did not raise any questions yet. Please rephrase.
Done

P1L16: Your introduction is focused on climate change and river runoff, however I do not think 1 data set is sufficient to solve that. I would focus more on fundamental understanding rather than climate change. Please shift the focus of the importance of the data or even present it only as a dataset.
It is true that this data set won't "solve" climate change or changes to river runoff, but the introduction is meant to connect the paper to the larger topics that it relates to. That's what we're doing here. The wording makes it clear that the paper focuses on the data, this section of the intro is providing the context for why the data was collected. We added a sentence about data scarcity to bolster the motivations.

P2L1-4: in the paper itself you do not make the link to water runoff or the dam so I find this information misleading.
We include this information for context and motivation. On line 4 we clearly state "this paper focuses on the measurements."

P2L5: state 127 glaciers instead of 'more than 120'. Be specific
While we are generally in favor of being specific, here we use 'more than 120' because the number of glaciers in the basin can change e.g. as climate change removes glaciers, or splits one into two. I changed it to a nice round 'more than 100.'

P2L9-10 similar as stated above, you do not make this link to dam operations and environmental resources. Or looked at river discharge.
We include this information for context and motivation. On line 4 we clearly state "this paper focuses on the measurements."

P2L11-14: So you do not include modelling (as stated in P2L3)?
We include the information on modeling for context and motivation. On line 4 we clearly state "this paper focuses on the measurements."

Section 2. Study area: A lot of unnecessary information. Please exclude some information or make more clear why you discuss this information. Connect the information to your paper and rewrite to a smoother story.
Ice sheet
Surging
Permafrost
We believe all the information presented is useful background to introduce the reader to the area we're focusing on. In fact we already connect the information to our paper: "The glacier monitoring work focused on the five largest glaciers..." We include a succinct summary of the characteristics of the glaciers (area, debris cover, area change, surging behavior, and velocities) to give the reader a sense of the glaciers' behavior and previous work that has been completed on the glaciers. The one sentence on ice sheet extent provides context. The paragraph on surging is important because surging can dramatically influence the mass balance of a glacier (one of the main foci of this paper). The two sentences on permafrost and the map (figure 2) provide context for the soil temperature measurements.

P2L19: why Kienholz reference needed?
Because Kienholz and co-authors did the work we're building on here. Their outlines were then incorporated into the RGI.

P2L24: "few"=how many glaciers?
Nine, as mentioned on line 32.

P2L28: change "ninety-three" → 93
Convention states that sentences should not start with Arabic numerals.

P2L28: "127 glaciers": numbers are not congruent ("more than 120 glaciers" in L5, "Most glaciers (in total 127)" in L23). If the area contains 127 glaciers please change in L5 to "more than 120 glaciers" to "127 glaciers", remove in L23 "(in total 127)".
127 in Alaska Range + 9 in Talkeetna Mountains, but see also my reply to P2L5

P2L29: how did you estimate this volume? Explain in text.
The first part of this sentence explains how we estimate volume.

P2L30: why do you use these scaling coefficients? Please explain how you got those or insert reference.
Added:
Radic, V., Hock, R., and Oerlemans, J.: Analysis of scaling methods in deriving future volume evolutions of valley glaciers, Journal of Glaciology, 54, 601–612, https://doi.org/10.3189/002214308786570809, http://dx.doi.org/10.3189/002214308786570809, 2008.

P2L31. Place '.' After m a.s.l..
Done

P2L33: change point into comma or insert 'and'
That is a period at the end of the sentence.

Page 3: Figure1: Combine Figure 1 and 2 by shading/colouring the permafrost areas >50% in Figure 1. For reader it is not clear why permafrost is so important.
Figure 1 is too complicated to add the permafrost classes and keep it readable. I added text to explain significance. "Permafrost affects water runoff, soil temperature, vegetation, and soil carbon fluxes. These factors have complex interactions with climate change."

P3L1: Why is debris cover relevant?
P3L4 explains why. I reordered these sentences.

P3L1-2 please rephrase and combine L1 and 2
No longer needed with reordering.

P3L5: include Östrom reference.
Upon searching, I can't find any relevant reference by that author.

P3L6: Do you exactly study the same area?
Yes, which is why we felt we did not need to say that we study the same area.

P3L10: "History of surging", are they still defined as surging glaciers or long time stable aleardy??
That's an interesting debate, that we choose not to delve into here.

P4L1-8: why is surging relevant? Mass balance not tat different than non-surging glaciers.
Mass balance can in fact be affected by surging (and vice versa). The latest surge of West Fork occurred shortly after the 1981-83 mass balance measurements to which we are comparing our data. While the effect of a surge on the mass balance may not be quantifiable given the data that is available, the surge still needs to be mentioned.

P4L9: this line suggests that glacierized parts have vegetation and human development. Please rephrase.
Text reads "The non-glacierized part..."

P4L9-15: why is permafrost relevant?
Added text to explain significance. "Permafrost affects water runoff, soil temperature, vegetation, and soil carbon fluxes. These factors have complex interactions with climate change."

Page 5 Figure 2: Do not understand the importance of this figure, maybe only the two upper permafrost colors needed? Than combine with Figure 1.
See above

Page 5 Figure 3: put all photos next to each other or make panel a and d bigger such that they are similar height as b+c.
Figure is designed to be one-column wide and be aesthetically pleasing with the mix of vertical and landscape orientations.

P6L4: do you refer to figure three, or do you mean you installed three AWS? Are those the 'station type AWS" in Table A1? Not clear for reader that energy balance weather stations are 'AWS' and that simple AWS is 'HOBO' in table A1.
Added "Figure" to the 3. Edited appendix to remove "AWS" and "HOBO".

P6L6: "floated" is not appropriate here.
Floated is the best descriptive word we have for this installation type. Added quotes around it to make clear that it is not floating in the literal sense.

P6L25: "The station records…surface elevation changes", put earlier in section (for example at the end of P6L4.
We mention this sensor on line 14, in what we think is a logical place to mention it.

Page 7: Figure 4: caption "(i) outgoing longwave radiation" use abbreviation LW as defined before.
Done

Page 8: Table 1: It is not clear to me whether the "energy balance weather stations" are the onces indicated by "AWS" of "HOBO" in Table A1? Please use consistent naming for all stations throughout paper.
Edited appendix to remove "AWS" and "HOBO".

P8L1-2 : what do you want to show with correlation coefficients in Table 2?
Section 3.1.2 describes what we want to show with the correlations.
P8L4: "variable", you indicate only the range of the data in Table 2. Please provide also the standard deviation of the values in Table 2 and refer to the standard deviations in the text.
The standard deviations are not particularly meaningful given the short records we're dealing with, so we chose not to include them. The reader can get a sense of the variability we refer to here by looking at figure 4.

P8L7: "more than 4 degrees Celcius lower": please rephrase to "temperatures by a minimal bias of 4 degrees" or similar.
Replaced "lower" with "cooler"

P8L8: connect the sentences by adding "when" P8L6-10: structure is missing, please rephrase
Reworded.

P8L11-13: please provide numbers instead of only calling it 'higher', 'greater availability', 'less variability' etc.
The table and figure have the numbers. Adding all the numbers here makes the text unreadable.

P9 Table 2: daily correlations? Add standard deviations to the means to indicate the variability. You can remove the range, since that does not add more information than the standard deviation
Standard deviation is not meaningful for wind speed, precipitation, or shortwave radiation since their distributions are so skewed. Range does give important information. The reader can see the variability in the figure.

P10L9: Incoming longwave radiation is also influenced by the surroundings, especially in complex terrain. Insert "mainly"→"which depends mainly on the effective.."
Done

P10L12: high correlations means meteorology is affected by larger scale forcings rather than micrometeorology. You could add discussion about that.
We can't generalize from "LW" to "meteorology." P10L2 already discusses the "relatively homogeneous cloud conditions at those sites"

P10L23: change "Figure 5" to "Figure 5C". and add ABC in Figure 5
Done

P10L25: not necessary melt when T>0 degress, the surface energy balance should be positive.
We are not directly measuring sensible or latent heat fluxes, so we can't calculate the surface energy balance precisely.

P10L30: What about the influence of precipitation.
There was no precipitation on these days, which is why we don't mention it.

P10L31: rephrase "in the ice rather than the snowpack" to "the layer consists of ice instead of snow" or similar
Done

P11 Figure 5: explain the reversed patter of ice temperature (red lines) with depth in the text. At 21 April lines are ordered from light to dark lines with depth, while in June this pattern is reversed. With other words explain why the temperature gradient reverses.
"As air temperatures rose, the subsurface temperatures of the upper layers increased but with a time lag that increased with 25 depth (Figures 5 and 6)."

P11 Figure 5: the lightest lines (Ice2.5m) are not clearly visible and Ice3m not present at all. Please make those lines more clear.
With 15 lines we did our best to have distinct colors that show the short term variability. Perhaps the confusion comes because we do not have data from 4 m depth so there is a visual gap between the data from 3 m and 5 m, particularly at the beginning of the record. The legend clearly indicates the depths.

P12 Figure6: the reversed temperature gradient is here not visible why (not)? Go more in depth in the data (general comment)
The reversed temperature gradient is visible. On 28 April the coldest temperatures are at the surface. On 26 May the coldest temperatures are at -1 m from the initial snow-ice interface, with warmer temperatures at the surface. The pattern continues through the end of the season.

P12L4: "simple weather stations"? Why is station type in Table A1 than indicated by "HOBO glacier"? please make naming consistent throughout the paper.
Edited appendix to remove "AWS" and "HOBO" in favor of "energy balance" "simple glacier" and "simple tundra."

P13L7: which of the two sensors are more trustworthy? And why the comparison? Please explain in text.
The manufacturer's stated accuracies are 0.1 C for Rotronic and 0.2 C for HOBO as shown in Tables 1 and 3. Whether that equates to "trustworthiness" is a matter of opinion. The comparison is done for calibration (as stated in the section header). Added a new first sentence: "To ensure the validity of data collected from many individual sensors scattered across our study area, we set them up side-by-side before deploying them to the field."

P13L9-11: I think these argumentations do not match:
The HOBO sensor is slower than the Campbell, but coefficient of 1, and then conclusion is that there is a lack of consistent pattern. I do not follow this, please explain and rephrase
As stated in the text: 5 minute data show that the HOBO sensors are slower, but hourly averages have a correlation of 1. Figure 7 shows the lack of consistent pattern. Added references to figures in the conclusion sentence: "The lack of a consistent pattern in these comparisons (Figure 7) prevented us from adjusting the HOBO temperature data to match the Campbell data. The high correlation and low temperature offsets (Figure 8) among sensors gave us confidence that using HOBO stations to assess temperature patterns across the basin was valid."

P13L20: add some explanation/conclusion. I miss in this whole section why you do the comparison between the sensors and eventually the physical interpretation or conclusion from your statements.
Added explanation to the top of the section and this physical interpretation to this paragraph: "Again, this gave us confidence that we could use our data to assess humidity variations across the basin."

P14L1-2: did more people had this problem? Is it a random tip that can also occur during dry periods (since this can not be filtered out)? Or is the tip sometimes 'stuck'?
We don't know of others that have run into this problem. It is not a random tip because when the data show rain at one station, there is almost always also rain at other stations. We never observed the bucket getting stuck in our calibration.

P14L4: or conclusion is the HOBO has a sensor problem.

Yes that is the conclusion of the previous paragraph. This paragraph focuses on the Campbell sensor and the conclusion is that if the internal electronics create a double tip, the logger is filtering it out.

P14L16: is this katabatic flow measured or a assumption it develops?
We measured wind speeds and directions at the On-Ice station consistent with a katabatic flow in summer, as the text now notes.

P15 Figure 7: include the colours in the caption
Added "Each colored line represents a HOBO sensor, blue dots are for the reference Rotronic sensor, and red dots are for a Campbell 107 Temperature Probe."

P16-P17: Section 3.3 I do not think this section is a great addition to your purpose of the paper and not supported by any in depth discussion, please remove.
We included this section to help explain differences in glacier mass balance discussed later. For the revisions, I added text to the mass balance section to make the connection more direct. "These mass balance patterns are consistent with the warming temperatures and relatively stable precipitation measured at Talkeetna Airport."

P18 Figure 11: is this the same transect measured every week at same location? What do you mean with "plotted relative to a reference station"? Does this mean steepness in line is varying in time? Please explain in caption. Mention in caption what upper stations in winter are not operating/measurement problems.
The transect is not the same in every week. The stations are fixed in place, and whenever the station has enough data to calculate a weekly average, it gets plotted. Changed caption to: "Weekly air temperature profiles show the winter inversions and summer differences between glacier and tundra temperature. For each week, the reference station (Windy Creek Lower, 940 m a.s.l., triangles and black dots) was plotted on the horizontal axis according to the date. The other stations (triangles or circles) were plotted to the left or right of the reference station according to their temperature difference."

P19 figure 12: Add coloured lines for each of the dots to show whether the gradients change in time/how sensitive they are.
Given the inconsistent recovery of data from different sites in different years, we chose not to display best fit lines. It doesn't make statistical sense to try to compare these lines from year to year.

P19 figure 12: Add the resolution of the glacier inventory in the caption.
Vector outlines don't really have a resolution. The reader can see the detail of the outlines in Figure 1.

P19 figure 12: Insert in caption how the mass balance in computed (from the "HOBOglacier" station in Table A1?
Inserted "...measured with the glaciological method"

P19L1-2: or the measurements are not representative for the whole region.
Added "... or the best-fit line might be too sensitive to data from lower elevations."

P19L4-end: the linear interpolation is done with all data such that no division is made between years (1 average value for all glaciers and all times?)? If so this is a very simplified method and I am not convinced in the numbers you present. For example P is highly spatially variable, as you also state in P20L14-15.
Added "... for each year"

P20 Section 4.2 Very limited, please expand or consider removing or merging this section.
It might not take much space, but this section includes some key results. Reworded a few sentences to make it read better.

P20L17: I do not follow this, you did not do any model simulation of snow accumulation, or this is not mentioned.
Reworded to emphasize comparison to our stake network.

P21L2-3 please rephrase. Absolute difference at 2000 and 1000m or did you do some averaging?
This is an approximate estimate based on binned radar data and point measurements for multiple years and multiple glaciers. We're trying to convey the big picture here, and the reader can refer to the figure for the numerical details.

P21L4: how do you know surface roughness is responsible? Please add explanation or supporting material for this statement.
Added: "Over short spatial scales in the ablation zone, surface roughness is responsible for high spatial variability in SWE. The end-of-summer glacier surface is rough due to streams, crevasses, melt ponds, and moraine material. The end-of-winter snow surface tends to be relatively smooth compared to the summer surface, but can also have wind-derived roughness features that contribute to the variability in SWE over small distances."

P21L6-8 I am not convinced
What do you think it is then? We tried our best to adequately account for errors in our methods and this lists what we couldn't account for.

P22L9: again, why the roughness of the ice surface?
See above

P22L16: add explanation why data become noisier. Why does ice give more noise signal?
Added: This was likely caused when the acoustic signal bounced off different elements of the rough ice surface in successive measurements.

P22L20: you assumed constant density? What are the implications of this assumption?
Changed "assumed" to "used" since this is based on measurements. New sentence reads: "To calculate ablation from the observed distance change, we used a density of 350 kg m^{-3}, based on the average of 5 snow density measurements at the site."

P24 Table 6: increase fontsize
Table doesn't fit on the page with a larger font. The copy editors may be able to fudge this somehow.

Section 5.2: please remove, I do not think is Section is of additional value
Readers who are familiar with the prior work or the field site will be interested in this information.

P24L11: "though we did not do a detailed texture analysis", But still you know it matches with the STATSGO soil map? I am not convinced.
Soil texture is relatively easy to observe in the field and the soils we encountered were largely sandy loam. This matches with the STATSGO map. Sites with very little soil development (Windy Creek Upper, Valdez, Two Plate) show as bedrock on the STATSGO map. We also looked into some of the STATSGO attributes such as organic content and drainage characteristics, which aligned broadly with our field observations. As mentioned, we did not do a more detailed texture analysis (e.g. taking samples back to the lab), so what more do you want to convince you?
Changed: "Though we did not do a detailed texture analysis on the soils, the characteristics we observed generally matched up with the State Soil Geographic (STATSGO) soil map (https://datagateway.nrcs.usda.gov/)."
To: "Though we did not do a detailed texture analysis on the soils, our observations of soil texture, organic content, and drainage characteristics generally matched up with the State Soil Geographic (STATSGO) soil map (https://datagateway.nrcs.usda.gov/)."

P25 Section 6.2 Explain the uncertainties and effect by the disturbance of the disturbed soil on the measurements.
Added: "By pushing the sensors into the undisturbed soil on the side of the pit, we hope to minimize any temperature errors due to disturbance. We did not quantify the effects of disturbance on temperature, though we expect them to be small compared to the large temperature changes observed over the record. Deep sensors were deployed at the bottom of a 2.5

cm-wide hole drilled into the ground. Disturbance of the soil around the deep sensors may affect temperature too, but we tried to minimize that by using the smallest drill we could."

P27L11-13: why is this relevant, please remove.
Unclear what the reviewer thinks is not relevant - this whole paragraph is describing the data in figure 18.

P27L25: How do you relate this to the dam? No runoff analysis is done.
Deleted reference to dam here.

P27L27: this is only a minor section and for me not strong part of your paper and now you present the climate change numbers as one of your main conclusions.
I think it is totally fair to present our data as a baseline for future measurements. The point of concluding statements is to highlight how this paper and the data within it are relevant to the broadest audience possible.

P28 Figure 18: air temperature is gray colour?
Yes, as indicated by the legend.

P28L6: not new conclusion, snow amounts are generally higher at high elevations
True, but it has not been shown for this study area and time period, so it is still a valid conclusion.

P28L7: you did not measure the soil and your conclusion is that these match with the mapped soil descriptions. Please remove this statement out of your conclusions and preferably also out of the text
Detail added above to support these conclusions.

Conclusion in general: please do not focus on climate change and dam implications, but give conclusions about the data you found in the field. What did you find and why is it special?
We do give conclusions about our data and we highlight how the data are relevant to the broadest audience possible (people who are interested in climate change and runoff).