Anonymous Referee #2

Review "Meltwater Storage in the firn of Kaskawulsh Glacier, Yukon Territory, Canada" by N. Ochwat.

The authors study the density profile of two firn cores drilled in spring 2018 in the accumulation zone of Kaskawulsh Glacier (Yukon, Canada). These cores are used to calculate local firn density and the impact of meltwater retention and refreezing on surface lowering that must be accounted for to correct geodetic mass balance estimates. The authors obtain an average firn density of 670 ± 2 kg m^{-3} in the 36 m deep core, and estimate an average surface lowering of 10 ± 0.8 cm per year over the period 2005-2018. The authors also identify a perennial firn aquifer below ~35 m depth. The paper suffers from major issues including the robustness of the methodology, results and uncertainty estimates, making the conclusions difficult to trust. In addition, some terms used are unclear; the authors sometimes expect a priori knowledge from the readers (e.g. Section 3.3). The reviewer also noted that results reported in the main text and tables are often not matching, and that the conclusions lack of novelty. The paper is mostly descriptive and does not provide novel insight on geodetic mass balance uncertainties compared to previous studies. Therefore, the reviewer deems that the manuscript should be rejected in its current form. Below, the authors can find the reviewer's major concerns, listed as General and Point comments.

There is quite a bit to unpack here and on some points we provide more detail with the specific responses below.

We will revise the manuscript to clarify the aims, methodology, results, and uncertainties. We apologize for confusion in the values reported – we were inconsistent and unclear in some places, and this will be fixed. We will also provide more supporting evidence indicating how the density of the firn has changed over time, using a range of studies completed since the 1960s, primarily from the Icefield Ranges Research Reports, as well as from previously unpublished field data collected by colleagues. We will revise Section 3.3 in order to provide more a priori knowledge on the stable isotopes, and how these help to inform our inference of annual accumulation rates. We will also be clearer in the uncertainty analysis and in the discussion of the various density values being reported in order to make the relationship between the table and the text more clear.

The main objective of the original manuscript was to characterize the firn of the upper Kaskawulsh Glacier: a significant ice mass within a major icefield where little or no published data is currently available on firn density or densification rates, meltwater retention, or liquid meltwater storage. The revised paper will be expanded and restructured. We will reiterate and clarify the three main messages: 1) firn density and ice content, 2) changes in densification rate, and 3) the new firn aquifer in this region. The results and discussion will focus around these three points. Number (3) is admittedly a bit of an aside, but it is of great interest and is relevant to meltwater retention and mass balance studies, as well as affecting the glacier thermal and hydrological behavior.

General comments

1. Results are based on "subjective" approximations that may alter the conclusions. For instance, the completeness of the two firn cores section is assessed based on "visual inspection" by three persons. How do the resulting "random" and "human" errors impact the firn density calculated in Eq. 1? In L120, the authors provide a 10-20% uncertainty in estimating the factor f in Eq. 1 (L125-126)? This would lead to a ~100 kg m^{-3} uncertainty in firn density (assuming the 670 kg m^{-3} value reported here), in line with 110 kg m^{-3} estimated in Foy et al. (2011; see L287). However, the authors report uncertainties ranging from 2 to 6 kg m^{-3}. Please elaborate. See also Point comment in L137-140.

We will double check our calculations, but believe this is just confusion regarding the uncertainty in point samples vs. average values for the core. We use standard error analysis in these calculations, which we are happy to walk through in supplementary material if the reviewers would like to see it. To summarize here, our point samples (10-cm core sections) have significant measurement uncertainty, as the reviewer notes. There are numerous sources of measurements error, including the subjective assessment of the completeness of each core section. We will refer to these various sources of uncertainty as "random error" in the revised manuscript, as they are all believed to be random (vs. systematic) – sometimes we will measure or estimate a section as being too long or complete, sometimes we will underestimate it. The uncertainty factor of ~15% applies to point samples (10-cm core sections): say, for instance, 500 ± 75 kg m^{-3} for a given sample. Conceptually, for the case of random errors, averaging of the 10-cm density values for the whole core leads to marked reductions in uncertainty because random errors cancel out. Based on standard error analysis, the standard error in the average follows $s_e = \square/\sqrt{N}$, where $\square$ is the uncertainty in each data point and N is the number of samples. Taking the example of a 30-m core with 300 10-cm density values, $N = 300$. Taking $\square = 75$ kg m^{-3} as an example, $s_e = 4$ kg m^{-3}. We will go through our uncertainty calculations again to ensure that these are accurate, but the main point here is the difference between sample uncertainties and the standard error in the mean.

2. Across the manuscript, the authors report results that are not matching between the main text and tables, making the conclusions hard to trust. For instance, in L18 the authors report an average surface lowering of 10 ± 0.8 cm yr^{-1} between 2005-2018. In L356, the authors report 10 ± 8 cm yr^{-1} for the same period. In L322, this annual rate is cumulated over the period 2005-2018 to obtain 1.3 ± 0.8 m in ~13 years. What uncertainty was used here (0.8 or 8 cm)? Please elaborate. Similar issues can be found across the whole manuscript and are listed in the Point comments.

We will correct the typos noted in the text, thank you for pointing those out. We also now realize that reporting so many different densities for the calculated background firn and actual firn may have led to some confusion. We report the upper 10 m of firn density to allow comparability to other literature on firn density. We also chose to compare the partial core densities because the length of Core 1 and Core 2 differed – this permits a direct comparison. We will remove some of these values in order to make the relationship between the values in the table and text clearer, and will be more precise and consistent with the wording we use in the revised manuscript.

3. The 13-year period (2005-2008) is estimated using calculated total water content of 23.22 m w.e. at the drilling site and assuming an average accumulation rate of 1.76 m w.e. yr^{-1} (1960s). The authors do not assess the robustness of this estimate given the uncertainty in firn density. Please elaborate.

We agree with the reviewer that we did not adequately address the uncertainty in the estimate of average accumulation (hence age) of the core. We will add this uncertainty analysis to the manuscript. We have obtained a dataset from 2003-2013 of snow accumulation and density data from the Divide site (12 km from our drill site, similar elevation) on the upper Kaskawulsh Glacier. We will include this data in our estimate of annual accumulation rates in order to provide more supporting evidence for the accumulation rate chosen here. The interannual variability in this data could be used as an estimate of uncertainty. We have three additional lines of evidence for the annual accumulation rate: (i) preserved peaks in the oxygen isotope record, (ii) our own winter accumulation measurements from spring 2016, and (iii) the (much) earlier published data from the Icefield Ranges Research Reports. These will all be taken into account to provide an uncertainty in the annual accumulation rate, which can then be propagated through to the uncertainty in the age of the core.

4. The term "melt-affected firn" is often used in the manuscript but not explained. Is this firn affected by the presence of refrozen meltwater in pore space? What are the associated visual features as stated in L204-205? Perhaps a photo of the cores would help the interpretation. The same holds for "Ice content" in L134, that is sometimes defined as the cumulative thickness of ice layers in the core expressed in m, or as a fraction after being normalized by the length of the firn core (see e.g. L192 and Table 1).

A definition of "melt-affected firn" will be added to increase clarity in the methods section. "Melt-affected firn" is any firn that displays physical characteristics indicating that there was the presence of liquid water at some point. This can result in ice layers, ice lenses, or can be indicated by the lack of grain boundaries, the presence of air bubbles, texture, and opacity. "Melt-affected firn" can also be identified using stable isotopes and the cation/anions, however, this was not done in the field.

The use of the term "ice content" will be used with more precise wording in the text in order to clarify as to whether or not it is describing cumulative ice layers or the normalized fraction. We will refer to the former as the "total ice content" and the latter as "ice fraction".

5. The authors sometimes expect "a priori" knowledge from the reader. Section 3.3 on stable isotopes is a good example: how to interpret the summer peaks at -22‰ in Fig. 4? This section is not necessary and the results are not further discussed in the

text, except in L244-246 that relates low ion concentrations to active meltwater percolation/motion in firn.

We agree that more information needs to be included in section 3.3. The following will be added into the manuscript in the paragraph that contains L244-246:
"δ^{18}O records in ice cores are a proxy for paleo-temperature, and are thus often utilized to assist in deriving age scales for ice cores. This method relies on the strong high-latitude temperature modulation of the isotopic composition of precipitation [*Jouzel et al.*, 1997; *Schneider et al.*, 2005]"

Although we realize the amount of melt and percolation rendered the major ion data useless, and the stable isotope data largely washed out, we wish to keep it within the manuscript as these results support our other findings of active meltwater percolation. Additionally, we discuss the sections of the stable isotope record that are preserved, and thus useful in determining an age-depth relationship.

6. The conclusions lack of novelty compared to previous studies that also estimated surface lowering in the region (see L334-339). The paper does not provide a convincing estimate of (local) surface lowering uncertainty for geodetic mass balance measurements, nor estimate the regional mass change accounting for density correction. In L328-330, the authors claim that density estimated at the two cores are representative of a larger region, which cannot be proved using only two cores as stated in L371-376. The authors should consider combining their core measurements with firn modeling to obtain spatially continuous density profiles and estimate regional mass balance uncertainty due to firn processes.

There are several good points here, mostly related to points addressed above. Within our objective to characterize the firn density, densification rates, and meltwater retention on the upper Kaskawulsh Glacier, we did not initially set out to quantify changes in densification rates over time – only the annual densification associated with meltwater refreezing, and the resulting firn density profile. We understand that we need to clarify and expand on this to study to include changes in the densification rate as well, as to be more relevant to geodetic mass balance studies. As discussed above, we have reached out to colleagues that have worked in this region for many decades, including previous ice-core work (C. Zdanowicz, K. Kreutz, S. Campbell), and will use this unpublished data, as well as other published data, to quantify how the firn ice content and densification rate has changed over time. We can also develop models of this process, using climate reanalyses as forcing data.

It is difficult to know whether our two cores are representative of the larger accumulation area of this icefield or others in the St. Elias region, but they still provide information where little other recent work on firn properties has been undertaken. We will revise our discussion and not over-extend our claims. Firn modelling is difficult when there is a large amount of meltwater percolation and refreezing – firn densification models do not do well under this condition – which is part of the need for the kind of data that we present.

Point comments L92-94: Are the measurements from the snow pit discussed somewhere in the manuscript or shown in Fig. 2? Please clarify.

The snowpit measurements are presented in Figure 2 and are included in the density data and inference about annual snow accumulation. We will edit the caption of the figure in order to clarify that the measurements are part of the density figure.

L135-136: What does "melt percent" mean? How is this calculated?

We will explain this more clearly in the manuscript. "Melt percent" has been used in the literature (e.g., Koerner, 1977) to refer to the percent of annual snow accumulation that melts (and refreezes), in the accumulation area of polar environments. At our site we don't use this concept but use "ice content" to refer to the fraction of a core sample that is made up of refrozen meltwater. For clarity, we will refer to this as "ice fraction" in revisions.

L137-140: This is unclear, why should the thickness of ice lenses be divided by a factor two?

Ice lenses were partial ice layers, where the ice did not extend horizontally through the whole core section. We assume that, on average, the ice lens occupied 50% of the core; therefore the measured thickness was divided by two. We will clarify this in the text.

L161-164: The authors should provide some references on the methods used to study isotopes.

Please kindly refer to our response to comment #5 above.

L184-190: This paragraph includes numerous errors in reporting results. In L186, "571 ± 3 kg m^{-3}" is reported in the text while Table 1 lists 518 kg m^{-3} at core 2 between 4-14 m depth. In L187, "608 ± 2 kg m^{-3}" is reported while Table 1 lists 618 kg m^{-3} between 4-21 m depth. The authors report an extremely small density uncertainty of 2-3 kg m^{-3} while Figs. 2a and b show much larger uncertainties. In L229- 230, the authors state that densities larger than 917 kg m^{-3} are eliminated. However, Fig. 2a shows values of ~1000 kg m^{-3} or larger at 6 and 10 m depth. To the reviewer, it is hard to judge whether these errors are due to negligence or calculation errors. Please elaborate.

The uncertainties of point samples versus average values for the cores are discussed above (please see the response to point #1). Figure 2 shows the point data (10-cm sections). The outliers were removed from the calculations to determine density and background firn density, but we left the outliers in the figure in order to allow the readers to see all of the data used and because the outliers were still within the uncertainty associated with each point. We can eliminate the outliers from the graphs in the revised manuscript if this is a source of confusion. These are part of the uncertainty in the measurements, but of course these values are not possible so they can be set to the maximum density.

L185, 187, 188: For clarity, the authors should better write: "between 4 and 14 m depth" instead of "in the upper 10 m"; "between 4-21 m depth" instead of "in the upper 17 m"; and "between 4-36 m depth" instead of "representing ~32 m". The same holds for L284-286.

Thank you for the suggestion, we will edit the manuscript accordingly. We were attempting to be brief, but it is more important to be clear and therefore use longer wording.

L193: 660 kg m^{-3} is actually 1.5% smaller than the firn density of 670 kg m^{-3} reported in L189.

Thank you for noticing this error, we have recalculated and will make sure that the correct percentage is used in the revised paper.

L276: What do the authors mean by "summer melt extent"? Do they mean meltwater production in mm w.e. yr^{-1} as listed in Table 2? Please clarify.

This sentence will be reworded so that instead of saying "summer melt extent", it will be clarified as meltwater production in mm w.e./yr, as listed in Table 2.

L278: It is hard to assess the robustness of the results in this paragraph. In L278, the authors state that summer 2015 was the warmest in the period 2014-2018, whereas Table 2 shows that it was actually summer 2016 (-1.0ºC in 2016 vs. -1.8ºC in 2015). The same goes for annual mean temperature in 2015-2016 (-9.0ºC in 2016 vs. -9.6ºC in 2015). How to interpret the larger PDD and melt rates in 2015 then? Please clarify.

Thank you for pointing out the unclear text. We will rewrite this. It is that PDD are not the same as average temperatures, and they don't always correlate. PDD refer to the cumulative temperature above 0°C, which can deviate from the average temperature. For instance, days or overnights of −10°C will bring down the average temperature vs. −1°C, but 0 PDD accumulate in each case. Melt is assumed to scale with PDD rather than average temperature, as a proxy for available melt energy. We will revise to say that summer 2015 likely experienced the most melt in recent years, based on higher PDD totals.

L284: Again 608 kg m^{-3} is reported in the text whereas 618 kg m^{-3} is listed in Table 1.

Thank you for pointing out this error, it will be fixed.

L315: What do the authors mean by “certain amount”? Ice layer thickness?

“Certain amount” will be removed, as it is ambiguous. This will be reworded to include ice layer thickness, spatial extent, and pore space availability, and the reference Machguth et al., (2016) will be added.

Table 3: What does “1.5-2g” mean in the personal communication of Sass and O’Neel?

Thank you for noticing the “g” that should not have been there and will be removed.