Reply to Anonymous Referee #2 comments on "Comparison of manual snow water equivalent measurements: questioning the reference for the true SWE value" by Maxime Beaudoin-Galaise and Sylvain Jutras, The Cryosphere Discuss., https://doi.org/10.5194/tc-2021-354-RC2, 10 Jan 2022

On behalf of Sylvain Jutras, co-author, and myself, I thank the anonymous referee #2 for his comments on the submitted manuscript. Following the reading of his general and specific comments, responses to each comment have been formulated point-by-point. Reading these comments allowed us to identify concrete modifications to the manuscript, as mentioned in our replies. At the request of the publisher, it will be possible to provide the revised version of the manuscript.

Comments from the reviewer in blue
Answer in black

1. Manuscript presents comparison between three different snow samplers and a density cutter which are used for SWE measurements. Uncertainties and errors are presented and evaluated for the methods. The study results that the large samplers should be used as reference for SWE measurements.

Indeed, our manuscript aims to present these elements. However, we believe that this comment includes only a fraction of what we want our results and conclusions to present. The variability of the density measurements per snow layer with the density cutter are presented alone and not in comparison with the snow sampler. Since there were not several snow pits made per day, it was not possible to calculate a CV for the snow pit methods, like the snow samplers. The CV of the density cutter is representative only of a fraction of the uncertainty of the SWE estimation by a snow pit. The uncertainty due to instruments and the measurement error results allow a comparison between the snow pit and snow samplers methods used in our study.

2. Topic of the study is scientifically relevant, and results are supported by other existing studies. Data set is novel, comprehensive and collected with stable manner ensuring quality. The results are interesting for groups starting SWE measurements and choosing instruments for that, in addition to groups using already one of the instruments or similar ones, or groups, such as modelers, using data from the instruments. However, I would like to see more novelty and progress beyond current understanding in the study.

This is indeed the intended scope of our manuscript. However, we believe that this article brings an interesting level of novelty. We answer this in more details in replies to Anonymous Referee #1 (comments 1) and for comment 7 of this reply submitted to Anonymous Referee #2.

3. Snowpack structure and snow properties should be presented to be able to evaluate for which snow conditions the presented results could be applicable. The reference for the true SWE is highly related to experienced snow conditions. For example, large and short samplers are not suitable for very deep and hard snow. It could be stated already in the title that which type of the snowpack is the comparison made for.

This is a comment similar to some observations raised by Anonymous Referee #1 (comments 6 and 7). Additional information has been added to the manuscript to better describe the typical snow conditions encountered at our study site. Histograms of the distribution of average snow depths and snow densities encountered during the 4 years of measurements have been added to the results section. A new subsection (3.1 snow measurements distribution) has been created in order to add the new distribution figure as well as the values of the coefficient of variation for each sampler according to snow depth, density and SWE. These results focused on the description of the study site clarify in which type of environment the following results are applicable.

This aspect was neglected during writing because we thought that our results impact more widely than the snow conditions specific to our study site. For example, the results for uncertainty due to instruments are not site-specific, but method-specific. The calculation of this type of uncertainty could be done only from a protocol and the instruments that will be used (with their precision) in order to estimate a minimum uncertainty associated with each SWE estimation method. For us, this is a type of results never before presented in the literature for the snow pit and snow sampler methods used in our study, but that would be relevant for any type of SWE estimation method.

4. As stated in lines 365-367, it is possible to study methods in several snow conditions with the data set. I would like to see these results, how samplers are working in accumulation period and melting period, and what could create additional errors in those conditions, such as increasing amount of ice layers in melting period or crust layers during melt-freeze events in accumulation period.

Using data from snow pits, we also checked the impact of the number of ice layers or the cumulative thickness of the ice layers present in the snowpack. These tests did not show any trend or significant relationship between these variables and the uncertainty or measurement error on the SWE estimate by the manual methods used. For the coefficient of variation of each sampler, the coefficient of determination (R^2) obtained by linear regression shows values ranging from 0.0009 to 0.003 depending on the cumulative thickness of the ice layers and values ranging from $3 * 10^{-5}$ to 0.04 depending on the number of ice layers. The same statistical analysis was done for the mean bias error (%). The R^2 value range from 0.013 to 0.017 depending on the cumulative thickness of the ice layers and values ranging from $4 * 10^{-4}$ to 0.004 depending on the number of ice layers.

Thank you for your comment, it is a relevant suggestion to improve our analysis. In order to test this, we divided our coefficient of variation results according to whether the measurements were taken during periods of snow accumulation or melting. Relevant results were found and added to the analysis. For SFS (p-value = 0.41) and HQS (p-value = 0.15), there is no significant difference, but the ULS shows a significantly lower CV during the accumulation period (p-value = $9.8 * 10^{-3}$). During the snowmelt period, the three samplers did not show any significant difference between them with CVs of 4.2%, 4.2% and 4.5% for SFS, HQS and ULS respectively. During snow

accumulation, CV values are of 5.9%, 3.8% and 2.9% for SFS, HQS and ULS respectively. During snowmelt, while there is no significant difference between SFS and HQS (p-value = 0.087) as well as HQS and ULS (p-value = 0.38), the ULS shows a significantly lower CV than SFS (p-value = 0.0031)

These results have been illustrated by modifying Figure 2 (line 317). The figure has been divided into two panels, showing the same result but with a distinction of the period of snow accumulation or snowmelt.

5. It would be interesting addition to density sampler comparison if results for different layer properties (e.g. hardness and grain shape classes) would be presented, similarly as for different density ranges. However, a problem with wedge cutter is that the sample includes more snow from the bottom of the sample location than top of it due to shape of the cutter. When samples are taken inside a layer, related error should be small, but layers are still naturally changing gradually. Related uncertainty could be checked. For thin layers, problem with using density from the same grain type is that it may also vary depending on snowpack structure. At least, analysis and description on how densities of the same grain types vary should be added. You could consider also average density from the closest measurements above and below instead of averaging the whole snowpack.

This would indeed be a relevant analysis to make. Still, we think a deeper analysis of the snow properties deviates from the main purpose of our article. We did not want to place too much importance on the analysis of the estimation of the snow density per layer. Although the snow pit is useful for gathering more information about snow layer properties, we want our article to focus on the various SWE measurements methods, including the snow pit and snow samplers, as tools to estimate the SWE of the snow cover. As mentioned in the introduction (lines 60-64), the snow pit, according to different protocols, is often used in different contexts to serve as references to true SWE. In order to stay close to this objective, we have written our article considering the snow pit, like the snow samplers, a manual SWE measurement method of the snow cover.

As mentioned in comment 1 of this reply, the SWE estimation of a single snow layer with a density cutter cannot be compared to the SWE estimation by a snow sampler. Although the CVs obtained for measurements with the density cutter are similar to those obtained from three snow samplers, it must be considered that the density cutter uncertainty is associated with a single snow layer, whereas the sampler is for the entire snow cover. The results of the uncertainty due to instruments (table 3 – line 325) shows this difference. This type of uncertainty was calculated in order to follow our main objective, and thus to be able to compare the manual SWE measurements by a snow pit and the snow samplers. Even if the uncertainty is small for the SWE estimation for a single snow layer, it does not mean that the uncertainty of the SWE obtained for a snow pit is. According to the principle of propagation of uncertainties, it is necessary to cumulate all the uncertainties of each snow layer to evaluate the whole snow pit. By this method, the snow pit shows a higher uncertainty than the larger samplers in estimating the SWE of a snow cover. However, we understand that there are limits to the measurement error obtained for the snow pit according to our methodology and we explain this in the discussion (lines 448-464).

6. I think that novelty and impact of the study are not strong enough for publishing in The Cryosphere at present. However, I recommend improving the manuscript and publishing in another journal. In the case of significant improvements on the manuscript, resubmission to The Cryosphere could be considered.

Like we respond to Anonymous Referee #1 (comment 8), we believe that the comments miss the main purpose and novelty of this study. In the context of the boreal biome, which is different from an arctic or alpine environment for example, the snow pit made with a small size density cutter is frequently used to estimate the true SWE of the snowpack. The objective of our study is to compare this method with snow samplers already used in the field (SFS and HQS) and a larger sampler developed for research purposes (ULS). We believe that the article presents novelty by studying the snow pit at the same level as the samplers instead of considering it as a reference. This is a method used to measure the SWE of the snowpack, but where there is no evaluation in the literature of its uncertainty and its measurement error at this level.

Due to the presence of snow cover for part of the year, the boreal biome is part of the cryosphere. Although our conclusions have a scope more focused on stakeholders in a territory similar to ours, we present a study that we consider to be innovative and which, to our knowledge, has not been carried out on other biomes of the cryosphere, such as alpine and arctic biomes. This type of methodology could be extended to other snow conditions in order to assess whether it is the best reference for true SWE in these conditions. In order to avoid confusion, clarifications in this sense has been added in the discussion and conclusion for specifying the results with a more local scope, and those with a more global scope.

Specific comments:

7. It would be nice to have map and figures from the sampling locations.

This is something we thought about in the process of writing. We think it is not necessary because the measurements have always been in the same place on the NEIGE site for the 4 consecutive years of this study. We only mention the coordinates and a description of the site in the section 2.1. Using theses coordinates, readers can easily find numerous maps of the site through free online earth observation portals.

8. Lines 28-29, 32: The first documentations about SWE samplers and snow courses have been published a bit earlier in 1920's in Europe, but those are quite not possible to find since written in German and not available online. I recommend rephrasing such as "On our knowledge, the first documentation in English…"

This is a good observation that I hadn't considered. This sentence has been modified in the manuscript as suggested.

9. Line 29: add 1¾ inch also in cm

The dimension in cm has been added in the revised version. This will make it more consistent with the rest of the text.

10. 2.10 Lines 139 and 240: I would recommend using uniformly unit of mm when writing about SWE instead of cm, then it will not mix with snow depth that easily. Now, both units are used which is confusing.

This is a relevant suggestion to avoid confusion. All the manuscript has been modified in order to have the SWE measurements in mm and the snow depths in cm.

11. Line 204: "...is the total thickness of all snow layers (other than ice layers) (cm)" Otherwise it might look like thickness also includes ice layers, even though it is written in the next sentence.

During writing, we considered the expression "snow layers" did not include the ice layers. This is indeed a mistake on our part and we thank you for pointing it out. This sentence has been modified as suggested.

12. Line 379: Also "under similar snow conditions" would require better description earlier on what kind of snow conditions you had.

As answered to the comments 6 and 7 of Anonymous Referee #1 and comment 3 of this reply, a better description of the snow conditions has been added in results section with a figure of the distribution of snow depth, snow density and SWE.

13. Line 404: "sections. Although"

This error has been corrected in the revised version of the manuscript.

14. Line 408: Chapter 4.2 could be simplified and main points better clarified

A revision of chapter 4.2 has been made so that the interpretation of the results of uncertainty is clarified.

15. Line 455: "drier snow" - newly fallen snow can be also wet (defined by liquid water content). Replace with "This lighter snow".

This is indeed a more appropriate term. The change has been made in a revised manuscript.

16. Line 520: Replace "in the methods section" with "in the Section 2 Material and methods"

This is a very relevant modification, it has been replaced in the manuscript.

17. Lines 560-565: Mention that using large samplers as "true" SWE is also environment related, like in deep snow conditions it is more reasonable to use extendable samplers.

That's a good suggestion. A sentence has been added in a revised version of the manuscript to clarify this detail. If we limit ourselves to our study area in the boreal biome, it is rare to encounter a snowpack with a snow depth greater than 1.5 m. On the other hand, for a broader view of our conclusions, it is an important limitation of large samplers for different environments.