

Referee 3

[Answer]: We thank referee 3 for the valuable comments. Please find our detailed answers below.

Buchmann et al. are presenting an in-depth analysis of a large set of closely located station pairs with parallel snow measurement in Switzerland. Most of the sites have several decades of parallel measurements which gives valuable insight in possible differences, caused by the impact of the station environment in contrast to the climate/region the stations may represent.

The achieved results of the presented statistical analysis give an estimated magnitude of the added uncertainties which can be expected due to (the often unknown) variety in the less-than-ideal siting of the station. This information is valuable for any analysis of long-term snow measurement, especially in case of single located stations, as metadata describing the location for such kind of measurements are often lacking the necessary accuracy or may even be missing entirely.

The paper is written in an understandable language and the analysis method is described adequately. I do recommend the publication of the study, but do have some concerns and suggestions that should be addressed before publication in TC:

- 1. The authors introduce a lot of abbreviations for their measured and calculated parameters, station names, time periods (seasons) and analysis methods. While some of those are well-known and their use well-established and justified, are others solely defined in this paper. The sheer amount is severely hampering the reading process, as the reader is required to either memorize the abbreviations or searching the meaning of the abbreviations somewhere in the text. This comment is of technical nature and should be easy to address. Please refer to the attached file with specific comments for possible solutions.*

[Answer]: We agree that too many abbreviations hamper the reading process. We inserted a table (Table 1) in the Data section and named the variables in full in the Abstract and Conclusion. Single months are also named in full throughout the manuscript.

The description of the results sometimes lacks clarity and accuracy. Obvious exceptions are just very briefly and sometimes not at all mentioned, neither are possible reasons for those exceptions discussed. Please find some concrete examples in the attached file with specific comments.

[Answer]: We rephrased several sentences and replaced vague words with more accurate terminology.

The general problem with exceptions and outliers is that they are extremely hard to explain. We tried to argue the case by inserting a new section 4.6 where we discuss an example and the limitations of the metadata in general. See our answers to the specific comments below.

- 2. The authors were not able to conclude on the exact causes for the observed differences due to the lack of metadata for their sites. That fact is well described and creating awareness of this kind of problem that may likely exist for other long-term snow measurements is one of the main messages of the paper. I do think, however,*

that it would be appropriate to recommend a list of necessary metadata to record for ongoing snow measurements and if possible, also give advice on which sites to avoid entirely.

[Answer]: list of necessary metadata is one thing, having (accurate) metadata at all is key. How best to describe the influence of a growing tree in the vicinity of a site? Influence of a structure (house, wall, fence et cetera) on a site? It's basically always a question of resources and will, but at least the exact coordinates should be known, ideally accompanied with pictures of the surrounding area and frequent visits to the sites.

We discussed that topic in the new section 4.6.

3. *Further, as also one of the other reviewers mentions, it is necessary to clearly state that also for the case of parallel measurements of two ideal and neighbored sites, it is likely that significant differences in the described variables may still be observed. I can imagine that some of the still existing neighbored sites (such that metadata could be achieved and described for the more recent measurements) in this study could be used to illustrate, discuss, and possibly quantify these differences and how they compare to the differences found for the long-term analysis presented in this study.*

[Answer]: We agree and created a new section 4.6, where we discuss the example of the station pair Klosters and the use and limitations of metadata as well as possible explanations and remaining uncertainties.

Please find more minor and specific comments and suggestions in the attached file.

R: Specific comments to manuscript TC-2021-125 by Buchmann et al.:

Page 1, lines 22-24: I don't understand the meaning of that sentence, consider rewriting.

[Answer]: We rephrased the sentence.

Page 2, line 41: 1-km scale

[Answer]: done

Page 3, lines 78-79: *Please explain. Measurement routines for new snow are not familiar for all readers and it may remain unclear, why solely taking the difference of to subsequent snow depth measurements is regarded as wrong.*

[Answer]: *We added a sentence with references in L76-77.*

Measurements are taken every morning at 6:00 UTC at least between November and April (for details refer to Haberkorn (2019) and Buchmann et al. (2021)).

Page 3, lines 70-90: *Here you introduce a lot of abbreviations for your measurands and calculated variables. I am not sure, if all those abbreviations are necessary. Their unfamiliarity is severely hampering the reading flow later in the manuscript. At least, consider a table which you often refer to, to help the reader finding easily back to your definitions throughout the text or (and) simply try to use less abbreviations. I personally think that you also could just continue using "average snow depth" instead of "HSavg" during the*

entire text. Off course you probably still want to use abbreviations in figures with limited space.

[Answer]: We agree and inserted Table 1 in the Data section. Further, we used full names throughout the Abstract and Conclusion. And single months are also written out in full. Table 1 is mentioned in the figure captions where appropriate.

Page 6, line 122: why do you refer to equation 2 here?

[Answer]: That was a mistake. We removed the reference and rephrased the sentence. See next comment.

Page 6, line 123: this equation is probably not necessary. I suggest writing: Absolute changes defined as the difference between the fitted value at the end and the fitted value at the beginning of the time series.

[Answer]: We rephrased the sentence accordingly.

Page 6, line 127: Also here I do suggest to just explain in writing, I don't see any increased value in this equation, it may actually confuse a bit (I pondered sometime about the meaning, as I could not find the time period of 60 days you describe in the sentence before).

[Answer]: We agree and removed the equation.

Figure 1: Figure captions should be more stand-alone. Either refer to a table which explains the abbreviations your simply restate their meaning in the figure caption here.

[Answer]: We referred to section 2 with the definitions in the figure caption.

Page 8-9, Section 4.2:

- I do find this entire section difficult to read because of the (already mentioned) extensive use of abbreviations. As mentioned before, add an easy to find and cite table or reconsider the use of some of the abbreviations.

[Answer]: We agree that too many abbreviations hamper the reading process. Instead of inserting a table which can be confusing as well, we decided to write out the variable names in full on first occurrence in each section. Single months are also named in full. We didn't want to overload the manuscript with yet another table and think this is a suitable compromise.

- I suggest the use of "panel" or "panels" instead of "plot/s" whenever you refer figure parts

[Answer]: done

- line 184: consider writing: "...analysis due to median winter seasons shorter than 60 days" – I assume that is the threshold you have set?

[Answer]: done

- Line 192: Did you find this "empirical value" during this study or earlier? Or is it in use elsewhere, please elaborate or cite.

[Answer]: We tried 40 days as well, but didn't see any differences. We agreed on 60 as it refers to two months.

•*Line 194: please add some concrete numbers to explain "smaller" –also some of the median differences for the ablation period are very high compared to the majority of the lines, I can't find this mentioned anywhere in the text.*

[Answer]: As the differences between the various stations are large, we only want to show that overall, they are smaller during accumulation than ablation, which can be easily confirmed by eye. We rephrased the sentence so that it now contains information about variability as well:

Here we found that the differences and variabilities observed in the first period seem to be smaller than in the last period (Figure 3).

Figure 2: Please use A,B,C,D in the figure caption, maybe add also top and bottom panels, left and right panels where appropriate.

[Answer]: We changed left to panels A and B and right to panels C and D.

Page 12, Line 211ff: Beside the mentioned exception PIO, I also see that GSS, RIE and ZER (the latter two to a smaller extend) show larger variations in spring than autumn. Also, the non-typical large difference of 18 days at 5KK remains uncommented. I'll suggest commenting those in the text.

[Answer]: The lack of reliable metadata only leaves room for speculation for ZER, GSS, and RIE; we have no explanation. Nevertheless, we completed the list of exceptions. We highlighted the issue in section 4.6 and touch upon the large difference for 5KK. But in our opinion, 5KK shows typical behaviour.

Page 12, line 215-216: Here you write the stations are aligned from shortest to longest periods in Figure A2, while you state an alignment according to elevation in the figure caption of Figure A2, please correct

[Answer]: rectified in text

The pairs are aligned from low (left) to high (right) elevation.

Page 12, lines 217ff: Even if belonging to the minority; I do think it is worth mentioning the non-typical behavior of BEV and MVE. Are there any reasons why those stations are so different?

[Answer]: In our opinion, BEV and MVE show 'typical' behaviour, see figure 4. BEV is located in the Engadin in Eastern Switzerland, whereas MVE is in the Vallais. A possible influence could be attributed to the different local climates, but would go beyond the scope of this study to verify, hence not mentioned in the manuscript.

Page 12, line 225: Could you describe the different definitions of snow onset and disappearance?

[Answer]: As pointed out by another referee, this whole paragraph is somewhat redundant. We decided to leave it out entirely and only mention that our values are "similar" to values

obtained by Klein et al. (2016), even though they did not analyse the same time periods and applied another definition compared to our study. We added the references for the other definitions in the method section.

Page 12, line 227: I disagree that the absolute changes for especially Dstop by this study and by Klein et al (2016) are “very similar” in all cases. While some numbers are differing just by a decimal day, differ some by a factor two or in one case (BOS) even by more than a factor 3. Even if it just “days”, I do think this justifies a more accurate description of the size of the differences.

[Answer]: As mentioned above, we removed that paragraph in its entirety as it (1) is not strictly necessary for our analysis and (2) as you pointed out not quite accurate. Explaining the differences would go beyond the scope of this study. However, the main message is to put our values into context by comparing them to values obtained by Klein et al. (2016) for the few stations that had been analysed by both studies.

Figure A2: mismatch between description of alignment in figure caption and describing text. Please correct.

[Answer]: done, changed in the text to:

L225: The pairs are aligned from low (left) to high (right) elevation.

Page 14, line 243: This sentence sounds strange as it reads that 64 hours amount to 1 hour a day, Please consider rewriting.

[Answer]: We rephrased the sentence to:

The largest difference in a station pair (64 hours during MA) amounts to approximately one hour per day.

Page 14, line 248. Which abbreviation do you use for the station pair Klosters? Would you have any material to illustrate the “no difference on the first glance” and the differences you nevertheless describe? It is difficult to understand that it is not possible to see the difference between stations that close to building constructions. Please give an example on the metadata you do have and the findings you obviously have made on-site.

[Answer]: The no difference at first glance was meant to concern the metadata: same village, same elevation, but if using a proper map AND exact coordinates as well as local knowledge and metadata, differences arise.

Page 16, line 262: I think the referral to figure 5, can already be added to first sentence in section 4.5.

[Answer]: done

Section 5, conclusions:

In case you are keeping the abbreviations in the rest of your paper (see comments above), I would really recommend the use of complete variable descriptions here in the conclusions, to accommodate readers which start their reading with the conclusions.

[Answer]: We agree and use full variable names in the conclusion.

Page 17, line 275: which engineers are you referring to here?

[Answer]: civil or environmental engineers. Sentence rephrased to:
[...] which may be useful for applications where there is normally only a single time series available.

Page 18, lines 301-302: Is that speculation or do you have any research on that?

[Answer]: It is more an educated guess than speculation. But varying instructions combined with potentially varying executions are a likely source; but impossible to verify. We mentioned that in 4.6.

Page 18, lines 304-307: while this is an important factor to consider for the analysis of past timeseries, it is nowadays possible to get a better set of metadata. Why don't you recommend a set of metadata which ought to be recorded for each station and what kind of locations should be avoided.

[Answer]: WMO guidelines already give recommendations, but as we are dealing with stations that are well established, there is always balance between keeping the long-term (but maybe not perfect) measurements and begin from scratch, especially if the environment of a station is changing – which in Switzerland with its very limited space is always a danger.

Page 18, line 306-307: What exactly do you mean with your last sentence? Is that a wish for sites in Switzerland, worldwide? Only parallel sites? Is it enough with one of the sites in case of parallel measurements?

[Answer]: That sentence was intended as a wish for sites in Switzerland, especially in view of current efforts to optimise resources and save costs. But it obviously applies on a global scale as well, as parallel measurements are invaluable for many applications. However, as we only analysed stations in Switzerland, we added the locality and rephrased the sentence: This means that the current number of available long-term snow measurement sites in Switzerland should at least be maintained.

The numbering of the tables and figures are somewhat odd, but it seems that some of the figures/tables are supposed to be in the annex despite they appear mixed with the "normal" figures/tables of the paper. Some of the figures numbered with A are not less used than these "normal" figures, please consider to add those figures in the main document. Also, even if not stated on each figure/table, please insert all necessary information for understanding the figures/tables directly in the caption, independent from the describing text in the manuscript.*

[Answer]: We agree and improved the figure/table captions accordingly. Further, we included Figure A2 (now Figure 5) and Figure A3 (now Figure 6) in the manuscript and removed Table A2 altogether.