Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2018-21-RC1, 2018

Open Access Earth System Science Data Discussions

Interactive comment on "The SUMup Dataset: Compiled measurements of surface mass balance components over ice sheets and sea ice with preliminary analysis over Greenland" *by* Lynn Montgomery et al.

Anonymous Referee #1

Received and published: 31 March 2018

Summary

The Paper by Montgomery and co-authors introduces the newest release of the SUMup dataset and explains the data structure. The authors provide an overview on spatial and temporal distribution of the data and thereby highlight biases in the representativeness of the data. Eventually, the authors use the data to analyse dependency of Greenland snow density and air temperatures as well as to compare the density distribution of the top 1 m layer of the ice sheet to model results.

Printer-friendly version

Discussion paper

General Remarks

Montgomery and colleagues have made a very valuable contribution to cryospheric sciences by compiling the very large SUMup database, containing snow/firn density snow accumulation on land ice and snow depth on sea ice. These data are very important in evaluating and calibrating, e.g., model output. While the database compiled and presented is very valuable, I am not entirely convinced by the quality of the manuscript. I find the structure sometimes confusing and I wonder whether the relatively simple application of the database is really valuable. In its current shape, the comparison to MAR output appears somewhat appended and major components of a comparison between measurements and model output are missing or addressed superficially. Maybe this comparison could be removed and instead the manuscript fully focussed on a compelling presentation of the data itself? The current presentation of the data does not fully convince me. In particular, I find the use of terms confusing. For instance, it remains rather unclear what is meant under the term “measurement” and a number of similar terms is used without defining them. I am also confused to what is finally shown in the figures illustrating spatial and temporal distribution of the data: It appears that ∼95% of the data points are excluded from the analysis because their spatial resolution is too high? As outlined below, describing your data set through a number of clearly defined parameters could make excluding the majority of your data obsolete. Finally, the authors need to improve the manuscript by addressing numerous inaccuracies (e.g. units in the figures) and wording.

Detailed Remarks

Line 20: I would remove this statement, as the database is probably not the largest set of “field measurements”. Even when specifying that it is the largest data set of ’glaciological field measurements’, I would still remove the statement, as the sheer number of data points says nothing about quality or value of the data.

Lines 48-51: There have been efforts to (i) compile glaciological datasets in a consis-

tent form, tailored for the use in models. (ii) Not only accumulation but also ablation datasets have been compiled. Recent examples that addressed both (i) and (ii) are, e.g. Ahlstrøm et al. (2008) and Machguth et al. (2016). I recommend that the authors provide a brief overview of similar work that has been done in the past, also highlighting the achievements rather than implying that previous efforts were either not publicly available or not formatted consistently.

Line 70: Stating here that 40 measurements have been excluded is puzzling. As a reader, I ask myself why these data have been excluded. I would rather move this statement and maybe also the comment on co-location to the figure caption (indeed, both statements are there already). Since this is the only time that excluded data are mentioned, I also ask myself whether any quality testing has been performed or not? Table 1 and 2 show that there is a column reserved for uncertainties, but how is that column used? I believe it is a viable approach to include all available data, but it needs to be stated whether quality test were performed or if the user should do such tests.

Line 103: The title refers to structure and metadata, however, the following text has a strong focus on the data itself. Maybe consider changing the title or revising the structure? Furthermore, I think the paper would benefit from adding a figure that provides a graphical overview of the entire database, showing for instance its three subdatasets and, e.g., the number of measurements per subdataset together with a listing of the most typical types of measurement that contribute to the subdatasets.

Lines 111 – 112: I do not fully understand why these ten measurements are mentioned here. I would rather include a section to state your criteria for including or excluding data, e.g. “Snow density data that exceeded a physically plausible range from >0 kg m-3 and <1000 kg m-3 were rejected.” Such criteria for data compilation are currently missing from the manuscript but I believe they are an integral part of presenting a database similar to yours.

Line 134: Here you write of “data points”, above (line 116) of “point measurements at

different depths", on line 137 of "records" and elsewhere more generally of "measurements". While I do not have a clear solution for the issue, I think there needs to be at least a definition of what is considered a "measurement" in the context of this paper, followed by consistent use of the terms. It might be confusing that a 2 m snow pit can be one measurement and a density measured over a depth range of a few millimetres (using, e.g., a neutron probe) is also a measurement (together with all the thousands of other density values derived from just one bore hole).

Lines 169-171: Maybe it would be worth briefly explaining what kind of ablation process took place at Summit Station (I assume wind erosion?).

Line 177: remove "an" (maybe should read "one" instead?)

Lines 187 to 200: I find this rather confusing. It is hard to follow which measurements were finally included in the analysis (there is no statement on the exact number of measurements). Furthermore, I do not understand why more than 90 % of the data are simply excluded from the analysis. By excluding these data, you also change the distribution of measurements over time and space. Would it be a solution to bring the data, for the analysis, to a uniform depth resolution and then perform the analysis?

Figures 2 and 3: Your study would benefit strongly from providing clear definitions of the different parameters defining your "measurements". These could be "location", "transect", "depth", "data point" and more. Such definitions could be given early in the manuscript. For instance, in Figure 2 mainly the location matters and I would recommend visualizing that parameter. Because all high resolution density data taken in one bore hole share the same location, they could be easily visualized in the figure.

Line 232: Maybe "contributes" rather than "contains"?

Line 233: What kind of spatial resolution do you mean? I assume horizontally, along track?

Lines 249/250: This is unclear to me, maybe consider rewording?

Section 3.4: I generally found this section confusing and I think that using a consistent and clearly defined terminology (see comment to Figs. 2 and 3) would improve clarity. E.g. on line 293 you write of “collection bias” and on the following line of “sampling bias”. I think it is mandatory to use a consistent terminology, even more when focussing on the presentation of a scientific data set in a journal like ESSD.

Section 3.4.1: What kind of density is meant in the title? The following section suggests you analyse both accumulation and density.

Line 340: “>3000 m”

Line 354: no comma in “...a low density snow crystal, formation ...”

Line 357: what is meant with “1 m”?

Section 5: unless this is a requirement from the journal, I would move this important information somewhere to the top of the manuscript.

Line 378: What do you mean with “properly” described?

Line 381: As mentioned in the introduction, I do not think that such a statement is helpful in any kind.

Line 388: Please reword the sentence that starts with “in the future...”, something is wrong there.

Line 405: Not sure here, but would “scientists working in the field of remote sensing” be better than “remote sensors”?

Figure 2: Maybe use the same time periods for all three plots.

Figure 3: Unclear why the bins are so extremely different among the three plots.

Figure 4: Using the same time frame for all three plots would substantially improve the comparison of the temporal distribution of measurements at the three locations. Why is this the case in Figure 5 but not in Figure 4?

Figure 8: You mention the grid cells but if I am not mistaken, the spatial resolution of the MAR output is never mentioned in the study. I assume it is 20 km? But even if you used a higher resolution model run, the question of how to compare very coarse gridded model output to point measurements is important and, to my understanding, not discussed in the current paper. For this and other reasons, I am not convinced that the comparison to MAR contributes substantially to your study.

Figure 9: Unit cannot be g m-3 if the maximum value is 0.6. Throughout the entire manuscript, please check for consistent use of units.

Figure 10: same as Figure 10.

Figure 12: Here you use kg cm-3 but rather mean kg m-3.

Citations

Ahlstrøm, A., P. Gravesen, S. Andersen, D. van As, M. Citterio, R. Fausto, S. Nielsen, H. Jepsen, S. Kristensen, E. Christensen, L. Stenseng, R. Forsberg, S. Hanson and D. Petersen (2008): A new programme for monitoring the mass loss of the Greenland ice sheet Geological Survey of Denmark and Greenland Bulletin, Review of survey activities, 15, 61-64.

Machguth, H., H.H. Thomsen, A. Weidick, J. Abermann, A.P. Ahlstrøm, M. L. Andersen, S.B. Andersen, A.A. Bjørk, J.E. Box, R.J. Braithwaite, C.E. Bøggild, M. Citterio, P. Clement, W. Colgan, R.S. Fausto, K. Gleie, B. Hasholt, B. Hynek, N.T. Knudsen, S.H. Larsen, S. Mernild, J. Oerlemans, H. Oerter, O.B. Olesen, C.J.P.P. Smeets, K. Steffen, M. Stober, S. Sugiyama D. van As, M.R. van den Broeke and R.S. van de Wal (2016): Greenland surface mass balance observations from the ice sheet ablation area and local glaciers, Journal of Glaciology, 62(235), 861-887.

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2018-21, 2018.