Comment on essd-2021-160

Charles Fierz (Referee)

Referee comment on "Canadian historical Snow Water Equivalent dataset (CanSWE, 1928–2020)" by Vincent Vionnet et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2021-160-RC2, 2021

This paper presents an update of the 2019 Canadian Historical Snow Survey dataset (CHSSD). Errors in metadata as well as a large amount of duplicate data led to this overhaul. Furthermore, the dataset was renamed 'Canadian historical SWE dataset (CanSWE)' to reflect the inclusion of automated data and the weight put on water equivalent of snow cover (SWE) data. The paper is well written and well structured, describing clearly the processing steps as well as the cleaning procedures used. Great work!

The paper is timely as it emphasizes the need for standardized formats and procedures, "*Python routines specific to each agency and corresponding data format were written to process the data and metadata and arrange them in a consistent NetCDF format*". While I very much welcome this effort towards standardization I may have wished an even larger convergence towards WMO standards, for example regarding terminology (see comments below). However, as the authors conclude that "*Regular updates are required to make such datasets useful for the community.*", I would strongly suggest following WMO standards even more closely in future. For example, I anticipate a standardized WMO NetCDF format will emerge soon to which the *CanSWE* format may comply too.

In summary I very much appreciate the effort put in homogenising the Canadian snow survey data sets and therefore recommend accepting the paper after the authors addressed the issues below and do some minor revisions as suggested in the annotated manuscript.

Specific comments

- General introduction about snow courses and snow surveys, lines 23-60:

I miss a clear statement whether you only consider snow courses as relevant measurements of water equivalent of snow cover. While on lines 39-40 you state that, "SWE observation networks using different measurements methods have been deployed at a national scale in various countries to provide valuable in situ information.", the following lines only include countries performing snow surveys, which are most typical in North America indeed, but not everywhere. I would suggest including a short paragraph on other methods, referring as you do to the 'European Snow Booklet' but also to WMO-No.8 (WMO, 2018) and also to Lópezâ□□Moreno et al. (2020).

- Section 3, Quality control of the final dataset :
  The homogenization of data quality flags shows how important standardized report practices are.
  Do I understand correctly that a H or W flag automatically results in a D flag? In other words, a quality flag set at one step influences the final number of flagged values?

- Section 4, Data availability and Table 6, lines 328-336:
  From my perspective, except snw, snd, den, most of the entries in this table are observational metadata in the sense defined by WMO (WMO-No. 49, Volume I and WMO-No. 1160). It may be worth considering whether the WMO standard could be implemented in a future version of CanSWE, both regarding the name of metadata and data as well as within the NetCDF file. A note on this in the text would be appreciated.

Comments on terminology

- *'snow water equivalent'* vs **"*water equivalent of snow cover*",** lines 1 (title), 10, 23, 336 (table 6), 338:
  Consider switching to new terminology introduced in WIGOS Metadata Standard (WMO-No. 1160) and WMO-No. 8, keeping SWE as abbreviation.

- '*the depth of water that would be produced if all the snow melted*', line 10:
  Both the WIGOS Metadata Standard & WMO-No. 8, Vol 2, p. 13 state: "Water equivalent of snow cover (SWE) is the vertical depth of water that would be obtained if the snow cover melted completely, which equates to the snow-cover mass per unit area."

- '*[derived] bulk snow density':*
  I very much welcome you using '*bulk snow density*'!
  - Consider defining it on line 211 as ratio of SWE to SD – and also in the abstract, line 14 – Consider adding '*derived*' only in tables (4, 5, and 6) but not in the text (line 339)
  - Use it throughout the text (see lines 245, 246, 250, 278, and 344) and not on line 255 as it is referring to a range.

References

Lópezâ□□Moreno, J. I., Leppänen, L., Luks, B., Holko, L., Picard, G., Sanmiguelâ□□Vallelado, A., Alonsoâ□□González, E., Finger, D. C., Arslan, A. N., Gillemot, K., Sensoy, A., Sorman, A., ErtaÅ□, M. C., Fassnacht, S. R., Fierz, C., and Marty, C.: Intercomparison of measurements of bulk snow density and water equivalent of snow cover with snow core samplers: Instrumental bias and variability induced by observers, 34, 3120–3133, https://doi.org/10.1002/hyp.13785, 2020.

WMO: Guide to instruments and methods of observation: Volume II - Measurement of Cryospheric Variables, 2018th ed., edited by: WMO, World Meteorological Organization, Geneva, Switzerland, 52 pp., 2018.
https://library.wmo.int/doc_num.php?explnum_id=9870

Please also note the supplement to this comment:
https://essd.copernicus.org/preprints/essd-2021-160/essd-2021-160-RC2-supplement.pdf