

An extensive set of snow depth measurements from the European Alps is evaluated to explore snow climatology regions and temporal trends. Raw data is collected from various sources and subsequently harmonized, quality controlled and temporal gaps were filled. Based on the snow depth time series, five regions with different snow climates are derived using Principal Component Analysis and k-means clustering. Linear long-term trends and short-term trend variability are computed from the snow depth time series for the five derived regions. Finally, snow depth observations are compared with gridded air temperature and precipitation data, which is either reanalysis based or inferred from spatially interpolating observational data. The authors find decreasing trends of snow depth for the majority of the stations and substantial differences in trends between regions.

This study addressed a very relevant topic: the evolution of snow depth in the European Alps during the last half century. The authors quantify these changes based on an extensive compilation of in-situ time series of snow depth from different regions. Additionally, the authors provide the unified dataset as an online resource, which can be very useful for further applications. The manuscript is well written and applied data and methods are sufficiently explained. My comments concern therefore primarily smaller ambiguities and (potential) errors.

### General/major comments

- Some sentences are extremely long and should be shortened to increase readability.
- **Selection of different time periods**  
I'm confused about the different time periods that were used for this study: The PCA and k-means clustering was performed for 1981 – 2010, long-term trends were computed for 1971 – 2019 and short-term trends for 1961 – 2019. Could you elaborate in more detail why you use (these) three periods?
- **Which data (raw, gap-filled) was used for which analysis/plot?**  
In the manuscript, I was sometimes a bit confused which data was used for which application. E.g. in section 2.4 you use the raw data without any gap filling, right? But for the subsequent analyses you always use the gap-filled data (as explained in appendix section A.3)? This point should be more explicitly stated in my opinion.
- **Gap filling method**  
I struggled to follow the explanation of the gap filling method – particularly from line 641 to 654:
  - I do not understand what “crossing a calendar day window with a year window” means.
  - Shouldn't the “window of 31 days” be “window of gap length + 30 days”?
  - “mean of the daily values” → does this refer to the daily climatology?
  - “and the weights were based on the vertical distance between candidate and reference station.” → why are horizontal distances not considered?
  - Finally, I wondered if reconstructed values have a smaller temporal variability (on smaller scales) because you apply climatological values in your method. But this is not the case, right? Because you only compute the ratios from climatological values (daily means)?

### Point comments

**L150-151:** I don't understand this sentence – could you rephrase it?

**L157-158:** this sentence is a bit odd: why do three different climate forcing zones create four main climate regions?

**L179-189:** I struggled to read this sentence (because it is so long). To increase readability, the providers from Italy could e.g. be listed with bullet points.

**L195-196:** “with a few expectations of monthly/seasonal data from the HZB and SMI.” → was the monthly/seasonal data also used in this study?

**L196-199:** I'm confused by this sentence: automatic measurements are used both from France and the Aosta Valley, right? Then I would change “only for France or” to “only for France and”. Anyway, I think the sentence should be rephrased to improve comprehensibility.

**L205:** “(see also Fig. 2b)” should be “(see also Fig. 2c)”, right?

**L245:** why is the data for Austria only available until 2016?

**L253:** “The frequency by elevation (50 m bins were used to calculate to proportion)...” → I don't understand this part.

**L257-258:** why is this criterion applied? Wouldn't e.g. a threshold of 50% make more sense?

**L273:** did you consider to use the “elbow method” to find the optimal number of clusters?

**L273-274:** what do you mean by “as well as clustering directly on the daily observations.”?

**L287-288:** remove “using assessed”

**L337:** “and is highly correlated to elevation up to 1000-1500 m, and mostly constant above.” → I don't understand this part.

**L348:** the clustering was performed on the 5-dimensional PCA loadings, right?

**L367-368:** please provide a reference here for the HISTALP subregions  
**L370:** what do you mean by “estimated data-driven”?  
**L377:** what do you mean by “unique stations”?  
**L386:** I would not write “matches” here because there are 4 vs. 5 regions. I would rather write that the obtained regions are similar.  
**L404:** does this statement refer to a specific month? Or the entire winter?  
**L408-409:** “mean North & high Alpine” is odd. Do you mean: “While in December, the mean negative trend was stronger in North & high Alpine”?  
**L422:** “Points with lines indicate” → “Lines indicate” (or are there points with no lines?)  
**L467-468:** “Moreover, we assume that most of this seasonal imbalance is because there is no or no significant snow cover in that month” → I don’t understand this part  
**L493:** I guess “~100m” should be “~100 mm”  
**L515-516:** “because this implies less chances that precipitation falls at the “right” time.” → I don’t understand this part  
**L532-533:** I’m not sure if I understand this sentence correctly. Do you mean homogenization is not so important because such a large number of stations is used?  
**L589:** how do you define a network? A country?  
**L612:** how is this surrounding band defined in terms of horizontal distance?  
**L673-674:** I don’t understand this sentence.  
**L686:** what is meant by “ablation scheme of the different stations”?  
**L701-702:** “gap filling snow depth series using simulations of the Crocus snow model for the French Alps” → I’m a bit confused by this part. Does it state that gap filling was performed by running the Crocus snow model with meteorological forcing?  
**L721:** what is meant by “original observations”? Available observations?

**Figure 1b:** this panel is hard to read. Could you enlarge it? To increase readability, station density could be plotted instead (i.e. the number of stations per a certain area).

**Figure 2b:** how was the polygon for the DEM generated? With a convex hull?

**Table 2:** there are typos in the first row (e.g. “(0,1000] m”).

**Table 3:** the spacing between the columns should be improved (→ it is currently confusing that the columns “DJF #” and “MAM mean (min, max)” are so closed together)

**Figure A1:** I guess a subset of stations was used to produce this figure, right?

**Figure A2:** which statistical quantities (percentiles, outliers, etc.) do the points, lines and box edges represent?

**Figure B4:** how is the numerical quantity “silhouette width” computed?

**Figure C4:** For completeness, the table for MESCAN-SURFEX (March to May) should also be shown

### Stylistic comments and typos

**L120:** “1960–1990. (Lejeune et al., 2019).” → “1960–1990 (Lejeune et al., 2019).”

**L144:** “while Section 4 concludes.” → “while Section 4 convers conclusions.”

**L147:** “with their typical arc-shaped” → “with their arc-shaped”

**L231-232:** change to “Station numbers are shown for fresh snow (HN) and snow depth (HS) time series.”

**L279-280:** I would remove the line break here.

**L312:** “significantly to the” → “significantly from the”

**L348:** “There were” → “This yielded”

**L353-354:** “South of the main ridge, there were two regions:” → “Two regions emerge south of the main ridge:”

**L362:** “as has in the north” → “analogue to the north”

**L376:** “the station in common, and the same common stations” → “the stations, and the same stations”

**L400:** I would rather use present tense here (and in the following lines).

**L469:** remove “supposed to be”

**L561:** there seems to be a space in the word “scientific”

**L661:** “were useful” → “are justified”

**L664:** “has not been yet used” → “has not yet been used”