Answer to Reviewer #2:

As interactive comment on: „A satellite-based snow cover climatology (1985–2011) for the European Alps derived from AVHRR data“

by F. Hüsler et al.

We also would like to thank the second reviewer for his valuable comments, which contributed to enhance the quality of the manuscript. Each concern is addressed separately below:

→ **Bold: Comment of the reviewer**
→ Regular: Answer of the authors
→ *Italics: Changes to the initial manuscript*

Specific comments:

1. The validation of cloud removal methodology is done only using modeled monthly mean SCA. However, in order to assess the accuracy of applied approaches for cloud removal and to assure generation of accurate daily data on which further analysis (e.g. SCD, SCOD and SCMD) depend, it would be better to validate them against daily point or spatially distributed data. I assume authors have daily point or gridded station dataset which could be used to test specific cloud covered days. Moreover using MODIS snow cover product could also be useful to validate the methodology for cloud removal on individual scenes given that clear sky condition from MODIS is available due to different capturing time and possible clear sky condition during MODIS observation. The latter could be better due to the fact that MODIS observes snow cover independent from effects such as south/north facing mountain ridges which I assume is not considered in the interpolation methodology of gridded station data. Alternatively, the point data from stations could be used to test the accuracy of cloud removal methods.

Thank you for this important comment and also for the ideas for validation. Indeed, the validation is a crucial step in order to assess the product’s accuracy. In response to the comments on this subject from all three reviewers, we extended the validation of the gap-filling product to assess the seasonal performance and validate longer time periods. As suggested by the reviewer, we used the comparatively dense ground-station network in Switzerland and the available ECA&D snow measuring stations within the Alpine Region as a reference. Even though many other problems arise from ground station validation (such as biased positioning, mixed pixels in complex terrain, area vs. point measurement etc.) we decided to use *in-situ* observations as a reference instead of MODIS images. This decision was based on the fact, that snow retrieval and differentiation from clouds in MODIS imagery suffers from similar problems as does the AVHRR snow retrieval. Hence it would be considered a comparison rather than a validation due to same uncertainties in the “reference” data as in the data to be validated.

Finally, the following descriptions of the station validation and results (and figures) were included in the manuscript:

Station locations were included in Figure 1:

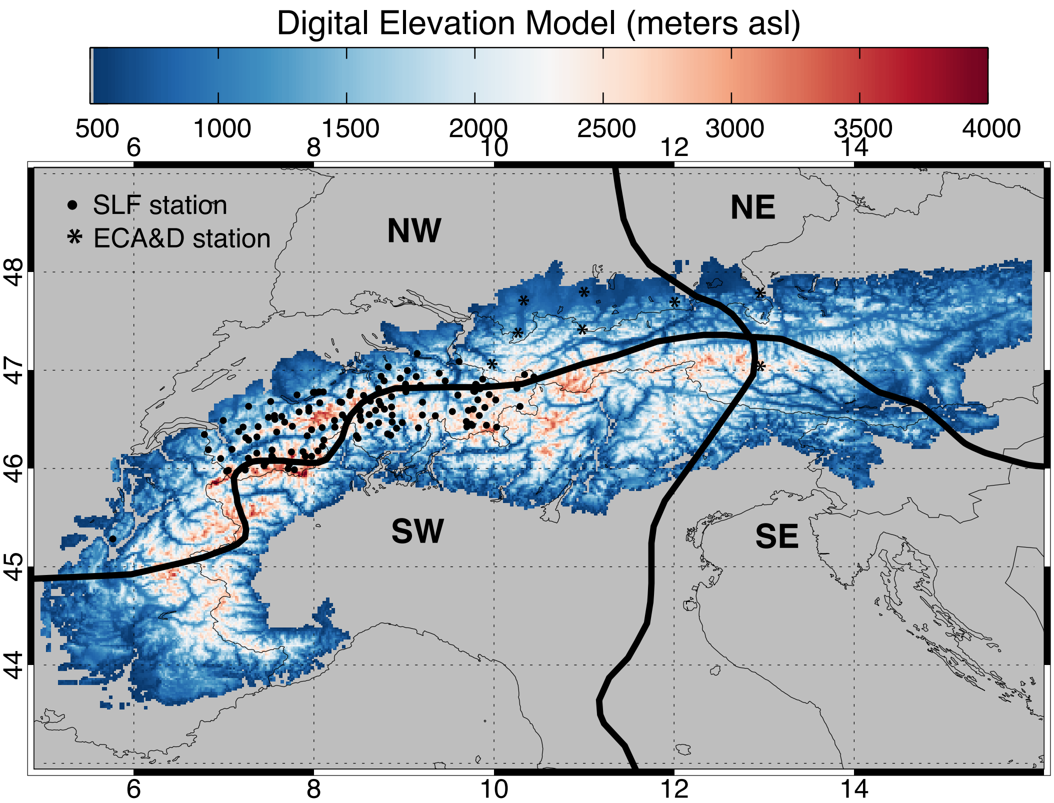


Figure 1: Digital elevation model of the Alpine region. Black dots (SLF station) and stars (ECA&D station) represent station locations used for the validation of the product. Bold lines indicate the subdivision into the four subregions (north-west, north-east, south-east and south-west) according to Auer et al. (2007). Data for the extra-alpine areas is not displayed.

A description of the dataset was added to the data section:

3.2 Gridded station dataset and snow-depth measurements

(…)

In addition, in situ snow depth information from 48 automatic stations over the time period 2003 - 2010 were used to validate the gap-filling procedure to include longer time periods, to assess the seasonal performance, and to ensure that the gap-filling does not significantly lower the initial accuracy. The stations are distributed all over Switzerland covering different elevation ranges and the data are provided by the WSL Institute for Snow and Avalanche Research (SLF). Snow-depth data is recorded automatically and available daily. To expand the validation beyond Switzerland, the 8 European Climate Assessment&Data stations (Klok 2005) lying within the Alpine Region (mainly Germany and Austria) were additionally included in the reference data set. The locations of all stations are illustrated in Fig. 1 and they are positioned *at elevations between 480 and 3100 meters asl.*

A new section was added to the Methods section:

4.2 Snow gap-filling accuracy assessment

To assess the accuracy and consistency of the gap-filled product, two types of reference data sets were used: gridded snow-depth data and in-situ snow-depth data.

(…)

For the point-wise validation the automated snow stations were considered as ground truth and were compared to the pixel closest to each station. As for the SLF stations, the snow depth is measured from above with an ultrasonic snow depth sensor (Lehning 1999). During the snow-free season, the same sensor measures vegetation height, which requested a pre-processing of the station data. For this purpose, a robust automated correction algorithm was implemented to find the first and the last occurrence of a location specific minimum value to determine the snow melt-out and onset timing. Values in between were set to zero as these are assumed to correspond to vegetation height rather than to snow depth. Quality checking indicated a good performance of the correction although short-term snow events after complete melt-out and before permanent snow onset might not be represented in the reference data.

Methodologically, the validation was carried out by means of investigating the accuracy index (ACC), defined as the sum of correctly classified pixels divided by the total number of compared (i.e. cloudfree) pixels. As above, the station was classified to be snow covered when the measured snow-depth was equal or higher to 5 cm. Furthermore, the missclassified pixels were evaluated in terms of snow over- (SO) and underestimation (SU), both relating the sum of incorrectly classified pixels to the total number of compared (i.e. cloudfree) pixels as suggested by Parajka et al. (2010). For this type of validation it must be kept in mind, that, in mountainous regions, steep elevation gradients are common and one single AVHRR pixel may include both valleys and mountain peaks with elevation differences up to 1200 m at alpine sites. As a consequence, differences may be observed when a grid cell value is compared to an in-situ point observation in complex terrain.

The results section ("Gap-filled product") was expanded with the results from the station validation:

The results of the station validation to assess the seasonal and filtered accuracy of the snow product are shown in Fig. 4 and consist of three boxplots. The topmost plot (a) illustrates the ACC for the four products (unfiltered, spatially filtered, forward and backward filtered) for all months while (b) and (c) show the SU and SO errors, respectively. The mean annual accuracy for the whole year ranges from 90% (for unfiltered and spatially filtered) up to 91% (for the temporally filtered products). As expected, the ACC is high for the full snow season (Dec–Mar) and the summer season with values constantly over 90%, whereas it decreases to 82%–89% in the transition season. The mixed pixels including patchy snow cover, different land cover types such *as forest, combined with strong elevation differences are suggested to be the reason for this. Also the IQR increases for the transition season, however the general pattern is quite consistent throughout all the years expressed in the comparatively small IQR values. Except for the month of June and October, the temporally filtered product does not decrease in ACC through the filtering process. On the contrary, it even slightly increases the performance of the snow mapping. This is assumed to be caused by the reduction of miss-classified cloudy as snowy pixels (or vice versa). Concerning the errors, they show constantly very low values close to 0%. The reduction of ACC in the transition season is attributed to a slight underestimation of the snow cover in spring (SU error of 6% (May) and 7.5% (June) for the unfiltered product). As stated above, this is most probably caused by mixed pixels (i.e.snow lying under trees or snow patches and therefore not being detectable by means of remote sensing). The temporally filtered products show a comparatively higher SU error (up to 12%–15% in June). As daily values are compared, this kind of validation is partic- ularly sensitive to the uncertainty of ± 10 days inherent in the composite product. During the snow onset season (Oct, Nov), however, there is a tendency towards snow overestimation by the temporally filtered products. This is more pronounced for the backward filtered product (SO error of 11%) and is probably explained by the filling of pixels with clear-sky information from a couple of days (maximum 10) ahead of the actual day. Especially at lower elevations, the snow-onset process is characterized by several snow-onset events interrupted by melting events. Likewise, it needs to be considered that the correction of the station data described in Section 4.2 also leads to some uncertainties in the snow melt-out and onset timing as derived from the station. Particularly short-term spring or summer snowfall events after complete melt- out and minor snowfall events before permanent snow onset are not represented in the station measurements.*

A new figure (Fig. 4) was added to the “Gap-filled product”-Section in the results:

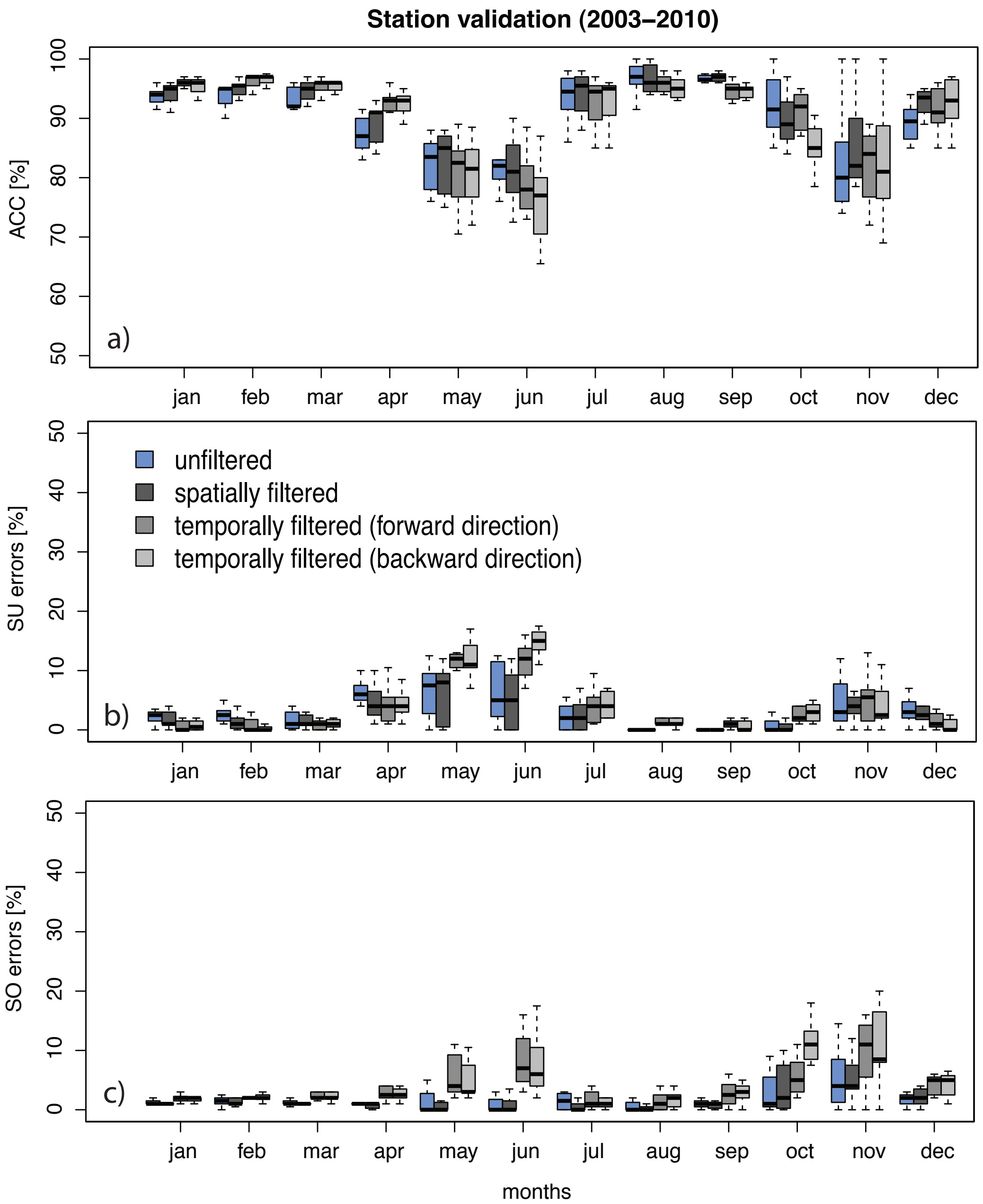


Figure 4: Monthly median and Inter Quartile Range (IQR) of (a) ACC, (b) SU error, and (c) SO error as from station validation for the period 2003-2010. Different colors refer to different types of filtering procedures (see legend).

2. I suggest to separate “Study Area” description from “Data” section unless you name the section “Study area and data”. Study Area itself can be one section and this would be a standard publication structure.

Thank you for this suggestion. We changed the “study area” to a separate section.

3. Figure 7: Please extend the explanation in text. Please give reasons why e.g. SCD at lower elevation should / could have positive departure from long-term value whereas SCD at higher elevation have at the same time negative departure from long term?

Indeed, this is a very interesting case. To give potential reasons, we added some sentences for more detailed explanation, which is based on the Winterbericht 2005/2006 (Stucki et al. 2011) annually compiled by the WSL Institute for Snow and Avalanche Research: *"The pattern of comparatively (short) long SCD at (higher) lower altitudes in 1996 is assumed to be caused by a late snow-onset (at all altitudes) and repeated snowfall events in February and March (Stucki et al. 2011). These led to record snow amounts also at lower altitudes which lasted longer and delayed the melt-out day in spring 1996."*

4. Page 3008, lines 3-8: Please explain how the monthly mean values are computed for different data availability periods (1985-1989, 1989-2002 and 2003-2011)?

Unfortunately, page 3008, lines 3-8 do not correspond to the content of this comment and we could not figure out the correct lines which this comment refers However, the detailed explanations about the computation of the montly mean values and anomalies can now be found section 4.3. To clarify, the monthly mean SCA value was calculated as the mean of all daily snowmaps available for the respective month. To fill the missing data in the timeseries, the long-term mean value was used in order not to influence the timeseries and its analysis.

5. Page 3012, lines 10-15 and Figure 3a and 3b: It is difficult to recognize the difference (improvement) of SNOWL method on cloud removal from figures 3a and 3b visually. I would suggest to also including a table with % values of cloud removal after each step to see the performance of each approach. This is also important to see how much cloud cover remains at the end.

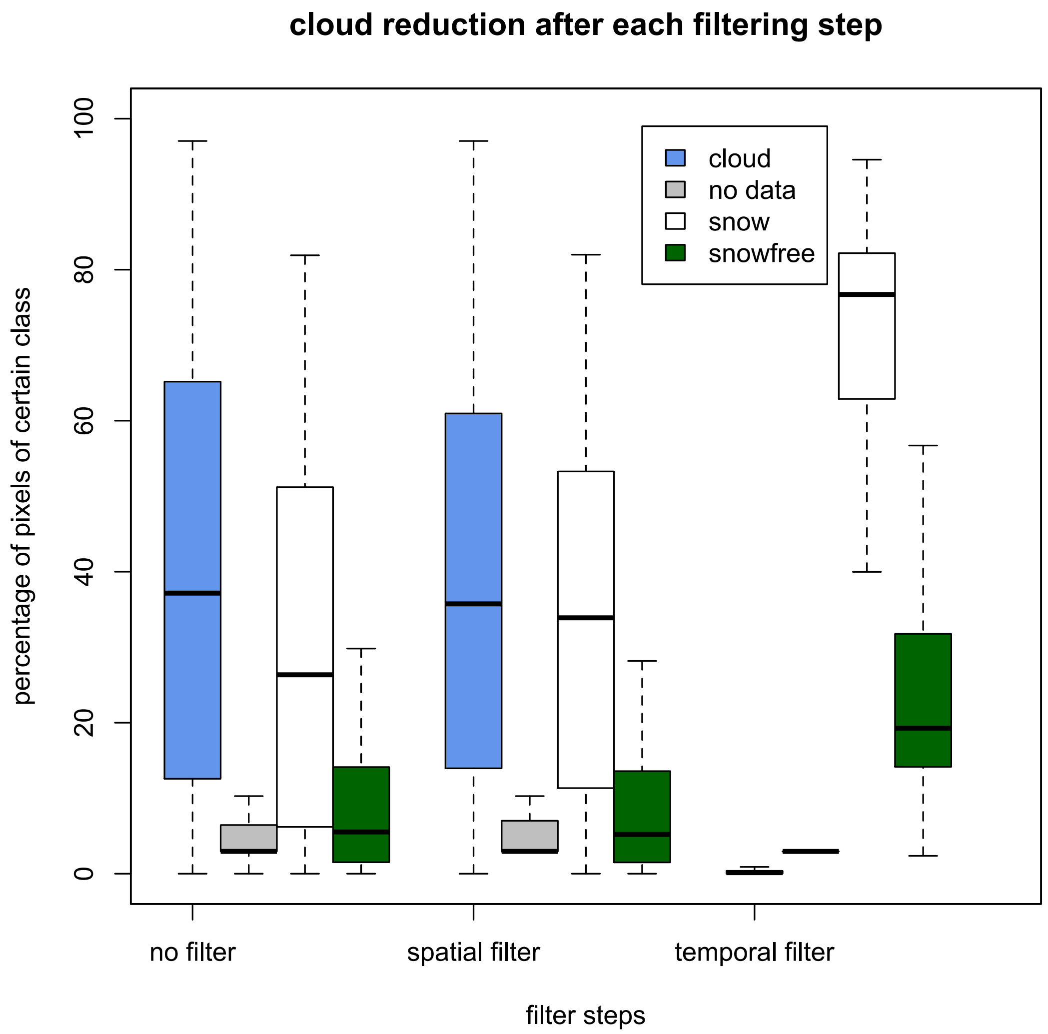
We agree with the reviewer and, as was suggested, included a table with the corresponding values. Table 2 and a short description were added to Section 5.1. The added paragraph reads: *"The quantitative improvement in terms of cloud reduction after each filtering step is shown in Tab. 2. While the change from the unfiltered data to the spatially filtered data is relatively small (change in median from 37.2% to 35.7%), the temporal filtering step strongly impacts on the number of cloudy pixels. The amount of clouds and non-valid pixels is, on average, reduced by 35% (clouds) and 3% (non-valid). The maximum percentage for the remaining clouds after temporal filtering is only 0.9%. Vice versa, the number of clear-sky pixels (snow covered and snow free) is significantly increased. The snow percentage rises from 26.3% up to 76.7% and finally corresponds very well to the same measure derived from the gridded station data set with a median of snow percentage of 73%. Likewise, the number of clear-sky snowfree cases shows an increase of 14%."*

The new table added to the manuscript:

Table 2. Quantitative assessment of classification redistribution after each type of filtering according to Fig. 3.

		min [%]	lower quartile [%]	median [%]	upper quartile [%]	max [%]
cloud	no filter	0.0	12.6	37.2	65.2	97.0
	spatial filter	0.0	14.0	35.7	61.0	97.0
	temporal filter	0.0	0.0	0.1	0.4	0.9
nodata	no filter	0.0	3.0	3.0	6.4	10.3
	spatial filter	2.9	3.0	3.0	7.0	10.3
	temporal filter	2.9	2.9	2.9	3.0	3.0
snow	no filter	0.0	6.2	26.3	51.2	81.9
	spatial filter	0.0	11.3	33.9	53.3	82.0
	temporal filter	40.0	62.9	76.7	82.2	94.6
	snow gridded	53.9	68.2	73.0	78.1	92.6
snowfree	no filter	0.0	1.5	5.5	14.1	29.8
	spatial filter	0.0	1.5	5.2	13.6	28.2
	temporal filter	2.4	14.2	19.3	31.8	56.7

In addition, the figure below graphically illustrates the reduction of clouds by each type of filter. The boxplot shows the median and 25% and 75% percentile of all categories within all scenes over the period Jan-Apr 2006. While the reduction of clouds by the spatial filter is small due to only high-elevation pixels being affected, the effect of the temporal filter is significant. After both filter-steps almost all clouds are successfully eliminated.



6. Page 3012, lines 15-17 and Figure 3d. It is not clear from the figure and in the text how "mean" value is computed. It would be better to include some more explanation to this figure as it is not easily understandable.

Thank you for mentioning this. In order to improve the explanations to the figure we not only added Table 2 but also some additional information to the corresponding section (5.1). The sentence "*For the final product, the mean SCA value between the two values (the one from the forward and the one from the backward filtered product for a specific day) was taken.*" was added to Section 4.1 while the sentences "*The average value is calculated as the mean between the SCA value derived from the forward filtering product and the SCA value of the backward filtering product for a specific day*." were adapted in Section 5.1 to better explain the procedure.

Furthermore, the figure caption (Fig. 3) was extended and now reads: "*Stepwise results from the spatial and temporal cloud removal process for the period January--April 2006: Respective area percentage of Switzerland for the classes snow, snow-free, cloud and not valid for (a) unfiltered scenes, (b) after SNOWL spatial gap-filling* *filter, (c) after temporal backward and forward gap-filling (mean value), (d) differences between forward and backward direction and averaged value thereof (calculated as the mean between the SCA value from the forward gap-filled product and the SCA value from the backward gap-filled product for the same day) and gridded station data.*"

Technical comments:

1. Figure 6 Caption: In text you write "long-term mean" whereas in the caption it is "long-term median". Please correct.

Thank you for careful reading. It is the long-term mean and it was changed accordingly in the figure caption.

2. Figure 7: Delete "Connecting lines are displayed for easier readability of the figures only". It is clear anyway.

Other reviewers of the manuscript requested this sentence for clarity. Therefore we think that leaving it is ok.

3. Figure 8: Please use lines for sat1 and sat2 legend instead of dots as they are not dots in the figure but lines.

We have changed the legend accordingly.

4. Figure 8: The dots (model) is unfilled in figure but filled in the legend. Please be consistent in both cases.

We've changed the legend accordingly. The dots are unfilled in both cases now.

5. I would suggest to use AVHRR1 and AVHRR2 instead of sat1 and sat2 in the legend.

Thank you for this reasonable suggestion. We changed it in the manuscript to AVHRR1 and AVHRR2, as you proposed.

6. The sentence "To assess the influence of the number of data available, the dataset was artificially reduced to one satellite at the time" belongs to text and not figure caption. In the figure caption you could write something like: e.g. "AVHRR2 – snow cover data with reduced number of scenes".

The figure caption was adapted according to the reviewers suggestion.

7. References: a. Hosmer & Lemeshow, 2000 is missing, please add it to references b. Remove page numbers (3004, 3003, 3020..) at the end of each reference as this is not standard referencing style of TC.

The requested reference was added to the literature. The page numbers could not be removed as the official standard TC style file was used for the references, which cannot be changed by the author.

8. In some of the figures the legend text is too small. Please enlarge them for better readability (e.g. figures 4, 5, 10 and A1).

Thank you for having a careful look at the graphs. The font size was increased in the respective figures.

9. Figure 3: Please use labels other than circle with outline which does not represent the figure and is difficult to read.

We've changed the legend in Figure 3 according to the reviewers comment.

10. Figure 7: Please use better indication for the years 1995 and 2001 if the statistics were not computed for these years (e.g. gap in connecting line between 1994-1996, 2000-2002). "No point" indication is not recognizable.

Thank you for this important comment. The no-data-years in Figure 7 are now clearly indicated.

References:

Klok, E.J. and Klein Tank A.M.G.: Updated and extended European dataset of daily climate observations. Int. J. Climatol., 29, 1182, doi:10.1002/joc.1779, 2009.

Lehning, M., Bartelt, P., Brown, B., Russi, T., Stöckli, U. and Zimmerli, M.: SNOWPACK model calculations for avalanche warning based upon a new network of weather and snow stations. *Cold Regions Science and Technology*, 30:145-157, 1999.

Parajka, J., Pepe, M., Rampini, A., Rossi, S., and Blöschl, G.: A regional snow-line method for estimating snow cover from MODIS during cloud cover, Journal of Hydrology, 381, 203–212, doi:10.1016/j.jhydrol.2009.11.042, 2010.

Stucki, T., Dürr, L., Etter, H.J.: Schnee und Lawinen in den Schweizer Alpen. Winter 2005/06. Wetter, Schneedecke und Lawinengefahr. Winterbericht SLF. Davos, WSL-Institut für Schnee- und Lawinenforschung SLF, 94 S., 2011.