Response to Anonymous Referee #2

[egusphere-2024-3505] Trends in the annual snow melt-out day over the French Alps and the Pyrenees from 38 years of high resolution satellite data (1986–2023)

We thank the referee for the helpful and relevant comments. We explain below how we aim to take them into account in a revised manuscript.

This study uses multi-source datasets to reveal snow melting day over the long time series. Employing various methods, it improves the spatial-temporal resolution of snow melting day to support the study results. Authors have made a comparison between SWH and landsat, but I am also worried about the data consistency between different landsat sensors, and between landsat and sentinel-2, especially under the case with a long study period and the mixture of all available data. So my suggestion is to add some data validation before using satellite datasets to calculate snow melting day.

We agree that data consistency is important. First of all we note that all sensors are multispectral imagers at 20-30m resolution, i.e. we do not blend high and low resolution imagery (like MODIS, AVHRR), or radar (like ERS or Sentinel-1). Secondly, we have evaluated the annual SMOD products that results from our multi-sensor time series and assessed their uncertainty in the paper. In the end, this is what matters for us since our objective is the recent evolution of the SMOD. We did not detect any trend in the mean SMOD error, but a larger spread in the beginning of the study period, that we attribute mainly to the lower revisit. In this paper we focused on the comparison between Landsat-DLR and SPOT SWH products because these products result from different retrieval algorithms (NDSI-based unsupervised classification vs. deep learning). Landsat products before 2018 (Landsat-DLR) were already evaluated by Hu et al. (2019a, 2019b)[1]. The performances of SPOT4 products were also previously evaluated by Barrou Dumont et al. (2024)[2], and the spectral characteristics of the SPOT products are essentially near-identical outside of the addition of the SWIR band to SPOT4 and 5. However, in response to the reviewer we have performed an additional analysis to evaluate the agreement between Landsat and Sentinel-2 products after 2016 as suggested (see below).

---

[1] Hu, Z., Dietz, A., & Kuenzer, C. (2019a). The potential of retrieving snow line dynamics from Landsat during the end of the ablation seasons between 1982 and 2017 in European mountains. International Journal of Applied Earth Observation and Geoinformation, 78, 138-148.
Hu, Z., Dietz, A. J., & Kuenzer, C. (2019b). Deriving regional snow line dynamics during the ablation seasons 1984–2018 in European Mountains. *Remote Sensing*, *11*(8), 933.

[2] Barrou Dumont, Z, S. Gascoin and J. Inglada, "Snow and Cloud Classification in Historical SPOT Images: An Image Emulation Approach for Training a Deep Learning Model Without Reference Data," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 5541-5552, 2024, https://doi.org/10.1109/JSTARS.2024.3361838

Line 110, when you combined the landsat data from different sensors over the long time series, how did you consider the sensor correction or data consistency between products from various sensors? In addition, can you give a brief description regarding the algorithm to produce level 2B product?

We retrieved the snow cover area from Level-2 Landsat Collection 2 Level-2 that are processed by USGS to provide consistent surface reflectance across different Landsat sensors and thus enable time series analysis (Köhler et al. 2022[3]). The USGS applies radiometric calibration and atmospheric correction algorithms to Level-1 Landsat data. We can give a brief description of the Level-2B algorithm (snow detection) in a revised manuscript.

Line 115, before mixing sentinel-2 and landsat-8 data from 2016 to 2023, I miss the consistency check between both datasets or sensor correction because they are from different sensors. Only if the data from different sensors are proven to be consistent, the trends calculated later will be meaningful.

We performed a comparison between eight pairs of snow products from Landsat 8 (L8) and Sentinel-2 (S2). These image pairs were acquired on the same day between November 2017 and June 2018, covering areas in the Alps and the Pyrenees.

Products were filtered using the TCD mask, the water mask, and the glacier mask, to exclude pixels that are not used in the SMOD trend calculation. Additionally, pixels containing clouds detected in either S2 and/or L8 were excluded from the analysis since images were not acquired at the same time of the day.

The analysis, conducted on a dataset of 77×10^6 pixels, yielded the following results:

34% of the pixels were classified as snow by S2.
31% of the pixels were classified as snow by L8.
S2 detected 96% of the snow pixels identified by L8.
L8 detected 89% of the snow pixels identified by S2.
The F1 score between the two products was 0.92.

We will include these results in a Supplement of the revised manuscript.

Line 183, what is the basis to set the threshold of tree cover 50%?

We aimed to exclude dense forests where the snow detection is very uncertain with multispectral optical sensors such as Landsat and Sentinel-2 - without excluding sparse forests where the snow detection is still accurate (Muhuri et al. 2021[4], Barrou Dumont et al.

[3] Koehler, J.; Bauer, A.; Dietz, A.J.; Kuenzer, C. Towards Forecasting Future Snow Cover Dynamics in the European Alps—The Potential of Long Optical Remote-Sensing Time Series. Remote Sens. 2022, 14, 4461. https://doi.org/10.3390/rs14184461
[4] Muhuri A., Gascoin, S., Menzel, L., Kostadinov, T., Harpold, A., Sanmiguel-Vallelado, A., López-Moreno, J. I., (2021) Performance Assessment of Optical Satellite Based Operational Snow

2021[5], Figure 3). We note that the TCD distribution across our study domain is bimodal with a mode near 0% and a mode near 100%. Therefore we expect the sensitivity to this value to be low.

Line 199, what is NOBS?

This variable refers to the number of observation and was indeed not defined in the main text, we apologize for this omission that we will correct in the revised article.

Check the figure 5. The left sub-figure is landsat only or landsat+spot?

There was indeed a mistake in the caption. The left panel is SPOT+Landsat.

Fig. 9 and Fig. 10, how about mapping pixel-wise? One of advantages of this study is high spatial resolution.

We took advantage of the spatial resolution by stratifying the analysis in 300 m elevation bands up to 3600 m high. The high resolution allowed us to obtain a large number of pixels even at high elevations, which would not be possible with coarser resolution sensors. In addition, the aggregation by elevation and massif contributes to reduce the noise at the pixel level (random error) and thus to show more robust results. A map at 20 m resolution would have to be resampled to a much coarser resolution to fit in the article format (without resampling at 72 dpi, a 300 km wide domain such as the Pyrenees measures approximately 5 m wide). This is why we choose to show only a small region at full resolution (Figure 5). Finally, the products are available at full resolution in a public repository.

Line 322, I am not sure how tree cover density correlates to elevation? Maybe the changes in bias with elevation could be partly explained by tree cover density. How about dividing tree cover density into several classes? Then analyze something, such as the bias and trend.

Following this suggestion we have computed the correlation of TCD and elevation from all the pixels corresponding to elevations > 1200m and TCD > 0 and TCD < 50. We obtained a correlation of -0.01 between TCD and elevation (therefore no correlation between the two). We will add this point to the Discussion.

Cover Monitoring Algorithms in Forested Landscapes, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, https://doi.org/10.1109/JSTARS.2021.3089655.

[5] Barrou Dumont, Z., Gascoin, S., Hagolle, O., Ablain, M., Jugier, R., Salgues, G., Marti, F., Dupuis, A., Dumont, M., and Morin, S. (2021) Brief communication: Evaluation of the snow cover detection in the Copernicus High Resolution Snow & Ice Monitoring Service, The Cryosphere, 15, 4975–4980, https://doi.org/10.5194/tc-15-4975-2021