

Response to referees [essd-2018-144]

We would like to thank the four referees for their positive and relevant comments, which helped us improve the manuscript. We have followed most of the suggestions as explained below. All changes were marked in bold in the annotated version of the manuscript.

Simon Gascoin on behalf of all coauthors.

Response to Gaia Piazzì (referee 1)

GENERAL COMMENTS

The authors introduce a freely-available collection of snow products derived from Sentinel-2 and Landsat-8 data through a robust retrieval algorithm, which allows overcoming well-known limitations affecting the snow detection. The newly-developed algorithm is described in detail and its potential is tested and discussed through several analyses. The satellite-based dataset presented in this study is of critical importance and great usefulness not only for the scientific research community, but also for users interested in operational applications. I have identified some issues which deserve only minor revisions of the manuscript before its publication.

We appreciate this positive feedback and all the relevant suggestions below.

SPECIFIC COMMENTS

Here are listed my main comments:

1. The manuscript contains several self-explanatory figures, which make it easier to read and understand the text. However, further references to figures need to be introduced, since for some of them the reference is even missing.

This was corrected (see specific comments below).

2. The manuscript comprehensively describes the newly-developed algorithm and the accuracy analysis of the resulting satellite products. However, a description of the satellite datasets used in this study (i.e., Sentinel-2 and Landsat-8) is missing. It would be of key interest to introduce the spectral bands of these satellite imagery to support the description of LIS algorithm.

We have added Tab. 2 to indicate the spatial resolution and central wavelength of the spectral bands used by LIS for each compatible sensor.

3. A further section introducing the evaluation metrics should be added to better enable the assessment of the results.

We have added Sect. 4.1.3 to describe the evaluation metrics.

4. Although the title mentions both Sentinel-2 and Landsat-8 data, references to Landsat-8 dataset are generally missing in the manuscript. Indeed, both methodology and results are actually referred to Sentinel-2 imagery.

todo

5. The final section 6 includes information of critical importance, which should be reported and

discussed in the text. Please consider my comments on Section 6.

[We have responded in the specific comment on Sect. 6 below.](#)

Please consider replacing sen2cor with Sen2Cor throughout the whole text.

Done

Section 1 (Introduction)

Pag. 1, l. 20: Please report the reference within brackets.

Done

Pag. 1, l. 22: Please replace 12h with 12 hours.

Done

Pag. 2, l. 5: Please consider reporting the full name of the plan.

[The full reference is given in the reference list](#)

Pag. 2, l. 10: Please report the reference within brackets.

Done

Pag. 2, l. 14: Please consider replacing 5 day with 5-days.

Done (*"of 5 days"*)

Pag. 2, l. 19, 21: Please explicitly define and explain what level-1C and level-2A products are (i.e., TOA, BOA).

Done

"(...) level-1C tiled product (a monodate ortho-rectified image expressed in top-of-atmosphere reflectance)

"(...) level 2A (L2A) products (monodate ortho-rectified images expressed in surface reflectance, including a cloud mask)"

Pag. 2, l. 21: Please specify the starting sensing date of the processed L2A product provided by ESA.

It was written "Since March 2018 ..."

Pag. 2, l. 24: What are the main limitations of ESA's L2A algorithm for snow detection?

[As explained below, we prefer to keep this for the Discussion. We have added the following paragraph in Sect. 5:](#)

"Both Sen2Cor and LIS use the NDSI to detect the snow cover, however the algorithms differ in many aspects. The better performances of LIS in Sect. 4.2.2 can be related to the mono-date approach for atmospheric correction and cloud detection in Sen2Cor, which can cause snow/cloud confusion in the L2A product, while LIS uses L2A products from the more accurate multi-temporal MAJA processor (Baetens et al. 2019). In addition, a unique feature of the LIS algorithm is the use of the snow line elevation concept to improve the robustness of the snow detection in mountain regions. In addition, LIS was designed to retrieve snow below thin clouds and to reclassify snow pixels which are frequently found near the snow cover edges in L2A products. Apart from the MAJA processing, however, the effect of these algorithmic differences could be less evident in flat areas."

Pag. 2, l. 28: Please consider explicitly specifying here that Theia collection relies on a combination of these two satellite datasets. Currently this information is more clearly reported only in the Conclusions (see main remark n°5 in General comments).

We agree that this was not clearly presented. We have added the specific comment in the introduction:

“The Theia collection was recently upgraded to also provide snow cover maps at 30 m resolution from Landsat-8 data using the same algorithm.”

Pag. 2, l. 28: Please consider rephrasing the sentence as “... at 20-m and 30-m resolution, respectively...”.

The sentence was split in two parts to clarify.

Pag. 2, l. 30: Please report here the reference for the Normalized Difference Snow Index.

Done

Pag. 2, l. 31: Please close the bracket.

Done

Section 2 (Algorithm)

Please add in the text references to Figures 1 and 2, which need to be commented (see main remark n°1 in General comments).

We have added a reference to Fig. 1 and 2 in Section 2.5:

“An example of the evolution of the snow line elevation over the tile 31TCH in the Pyrenees is shown in Fig. 1 and the corresponding snow maps are shown in Fig. 2.”

Please refer to Table 1 whenever an algorithm parameter is introduced.

We have added several references to Tab. 1. We also added a comment in the end of Sect. 2.4.1: *“The values of the parameters are provided in Tab. 1.”* Note that the text often refers to Fig. 3 (flow chart), whose caption also refers to Tab. 1 (parameters).

Generally, in Section 2 the authors do not explain how the parameters values have been defined and optimized. Please add this missing information in the algorithm description.

We added the following sentence:

“The parameter values were set based on previous studies with Landsat (Hagolle et al., 2010; Zhu et al., 2015; Gascoin et al., 2015) and by visually checking many snow maps and snow fraction histograms. From this set of a priori parameter values, only r_2 was adjusted based on the analysis of a first batch of products to enhance the snow detection in shaded slopes.”

Section 2.1 (Scope)

Currently some limitations of the ESA's processor are discussed within the description of LIS algorithm, for instance the critical issues affecting the cloud mask (i.e., thin clouds, edges of snow cover). In order to improve the readability of the manuscript and to properly highlight the potential of LIS algorithm, the authors could clearly identify and explain in this section what the main limitations of the ESA's processor are, before describing their algorithm.

We have started to design LIS before Sen2Cor was available. We agree that it is important to discuss the algorithmic differences between LIS and Sen2Cor to understand the differences in the results, however we prefer to do it as part of the Discussion (a paragraph was added).

Section 2.2 (Input)

In this section the input data are listed, but the data processing to derive input L2A product is not discussed. This section should also briefly explain how the correction of atmospheric and terrain slope effects is performed.

For the sake of simplicity we prefer to refer to MAJA ATBD where all the details are given (reference added in Sect. 2.2).

Pag. 3, l. 13: Please consider replacing remote sensing images with remotely-sensed images.

Done (remotely sensed)

Section 2.3 (Pre-processing)

Pag. 3, l. 22: Please specify the original spatial resolution of the red and green bands. It would be useful to introduce a table reporting the spectral bands and their central wavelengths of the original satellite data (see main remark n°2 in General comments).

Done (new Tab. 2)

Pag. 3, l. 23: Please spell out SRTM.

Done

Section 2.4.1 (Snow detection)

Are the cloud-contaminated pixels preliminary flagged and neglected in this first procedure? If so, please specify this information.

We have indicated "a cloud-free pixel"

Section 2.5 (Snowline elevation)

Please consider renumbering this Section as 2.4.2.

Done. Thank you, this was a mistake.

Pag. 4, l. 15: How are the false snow detections defined? What is the reference dataset to identify false snow detections?

At this stage there is no reference data, it is a guideline we adopt to have as much as possible true snow pixels.

Pag. 4, l. 18: Sentinel-2 tile is 100 km 2 .

No it is 110 km by 110 km (109.8 km), about 12E3 km².

Pag. 4, l. 20: Please specify how snow fraction is evaluated (fraction evaluated over the total number of snow and no-snow pixels?).

We have added "the total snow fraction in the image".

Furthermore, how is defined the threshold of snow fraction f_t ?

Pag. 4, l. 23-24: How are defined d_z and f_s ?

We have added references to Tab. 1.

$f_t = 0.1\%$ was defined based on the fact that it corresponds to 30'000 pixels, which we considered sufficient to have a statistically robust estimation of the snow line in the image. $d_z=100$ m and $f_s=10\%$ are indicated in Tab. 1. These parameters were defined in the early development phase by visually checking the snow maps and the snow fraction histogram (see response to a similar comment above). Here we have separated in the text the algorithm description to the parameters values because LIS is easily configurable to change these parameters. We did not test the sensitivity of these parameters using reference data but this could be the focus of further work.

Section 2.6 (Cloud mask processing)

Please consider renumbering this Section as 2.4.3.

Done

Pag. 5, l. 15: Please either specify how factor r_f is defined or neglect this information by just reporting the corresponding reference.

We prefer to keep it because r_f is a parameter like the others (it must be defined in the LIS configuration file), and we need to introduce it in the algorithm description.

Pag. 6, l. 2, 6: Please specify how r_D and r_B are defined.

We have added references to Tab. 1.

Figure 3: Please introduce MUSCATE in text before this figure.

We have added this sentence in the Figure caption: "*MUSCATE is the scheduler which manages the L2A production for Theia.*"

Section 2.7 (Implementation)

Please consider renumbering this Section as 2.4.4.

Done

Section 3 (Data description)

Pag. 9, l. 3: The number of selected tiles is already reported in pag. 8, l. 15.

We would rather let it here.

Pag. 12, l. 12: Please specify how the pixels classified as no-snow are defined. Are all snow- and cloud-free pixels classified as no-snow ones? Does the no-snow class include water pixels? Here it should be also mentioned the uncertainty in classifying vegetated pixels as snow or no-snow ones.

Indeed no-snow pixels can be anything except no-data, cloud or snow. Regarding the uncertainties in forest regions we do not think that it should be mentioned here since this section is devoted to the data description. We have highlighted this aspect in Sect. 6.

Pag. 12, l. 13: Please specify what pixels are classified as no-data.

These are the pixels outside of the acquisition segment.

We have added : "*Note that no-snow pixels can be any surface, including water surface. No data pixels are the pixels outside of the acquisition segment.*"

Please switch Figure 5 and 6 since they are mentioned reversely in the text.

Actually Fig. 5 is mentioned first, but in the caption of the Fig.1 and 2.

Section 4.1 (Method)

Since in this section the authors refer only to Sentinel-2 (e.g., pag. 13, l. 28; Table 3), they should explicitly specify that the evaluation is based only on Sentinel-2 data, as clearly stated in Section 6 (see main remark n°5 in General comments).

This is a good suggestion. We have moved this sentence from the conclusion to the beginning of the Method section: "*The evaluation was based only on Sentinel-2 products, because the production for Landsat-8 had just started at the time of writing and thus the Landsat-8 products represented a very minor fraction of the Theia Snow collection.*"

As this manuscript is focused on the mapping of snow cover, it would be more proper to limit the analysis over the snow period, not up to the end of summer season. For instance, Table 3 reports a selected pair of snow products in August, which is not of high interest for the main purpose of this research. Furthermore, it is noteworthy that, when assessing the contingency matrix, the analysis of these snow-free scenes could affect the representativeness of the results, as shown by the significantly higher number of correct negatives in Table 4.

We think it is important to assess the snow product in different seasons. In this product we find 174603 snow covered pixels mainly in the Parc National des Ecrins. It is only 0.6% of the image area, but still 69.8 km² (see Fig. R1 below). Regarding the representativeness of the results we agree that the accuracy may not be an adequate performance measure when the number of no-snow pixels is much greater than the number of snow pixels, however we have also computed the kappa coefficient and F1-score to account for this.

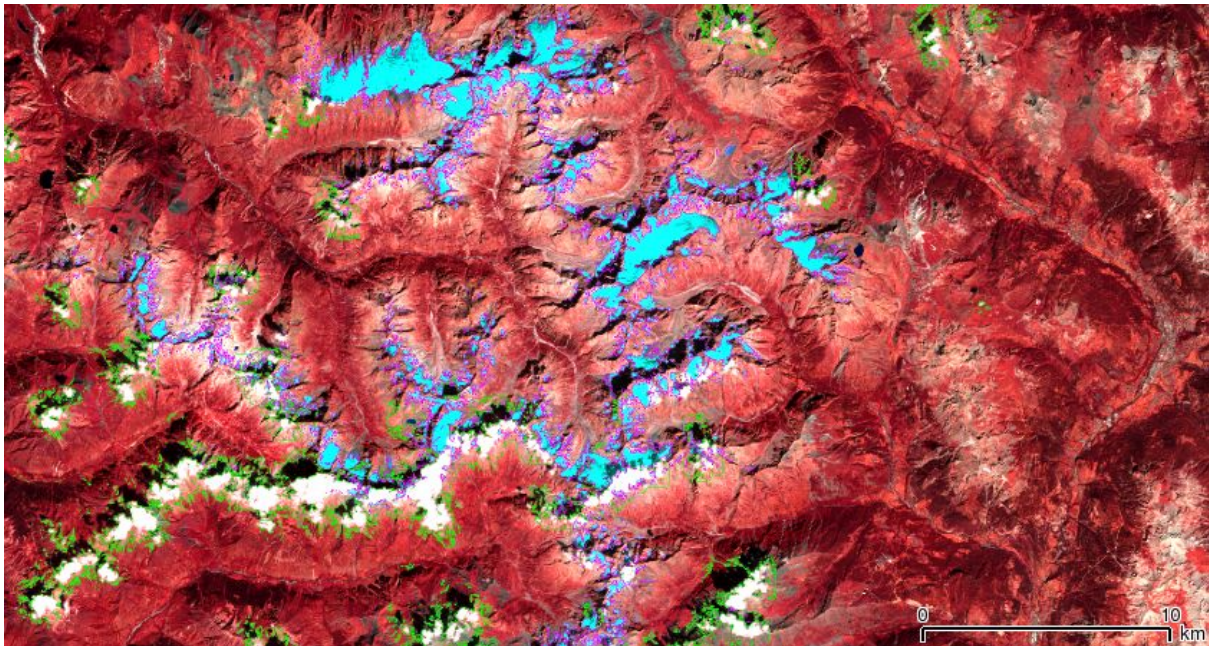


Fig R1: zoom over the Ecrins and Oisans on 2016-Aug-13 (snow covered outline: magenta, clouds: green).

Section 4.1.1 (Comparison with in situ snow depth measurements)

Please add in the text references to Figures 7 and 8, which need to be commented (see main remark n°1 in General comments).

We thank the referee for noting this. Now Fig. 7 is referred to in Sect. 3 and Fig. 8 in Sect. 4.1.1.

Pag. 15, l. 2: How is defined the first tile? Are the Sentinel-2 data of the second tile neglected even when sensed in a different day with respect to the selected first tile? If so, it would entail a loss of useful remotely-sensed snow data.

Given the overlap of between Sentinel-2 tiles (Fig. 8), it can happen that a station is located in two adjacent tiles. We did not assess the differences between selecting one or another tile, however we do not expect significant differences since the Sentinel-2 L1C data are the same. By visual inspection we did not detect discontinuities in the snow cover between tiles, which suggests that the LIS algorithm is robust to the tiling scheme. We aim to study this more extensively in the future by comparing the snow pixels in the overlapping areas independently of the station dataset.

Section 4.1.2 (Comparison with snow maps of higher spatial resolution)

Pag. 16, l. 11: Please remove the.

Done

3Pag. 16, l. 11: Could the authors mention the classification algorithms that have been tested and provide their references?

We tested different algorithms in the Orfeo Toolbox: k-means, decision tree, random forest, support vector machine and gradient boosted trees (boost). We do not think that it is useful to specify this given that we use only the results of the boost method here. However the interested reader can check the details of this work in the master thesis report of Bouchet (2017) that is cited in the manuscript (and available online).

Section 4.2.1 (Comparison with in situ snow depth measurements)

To improve the assessment of the results, please report how the evaluation metrics are defined (see main remark n°3 in General comments).

[We have added a new section 4.1.3 for this purpose.](#)

Section 4.2.2 (Comparison with snow maps of higher spatial resolution)

While in Section 4.1.2 the authors explain that “The Theia and sen2cor snow maps were compared ...”, here Table 5 shows only the results related to Sentinel-2 dataset. Once again, please specify that the evaluations are based on Sentinel-2 data, as currently reported in Section 6 (see main remark n°5 in General comments).

[This is now clearly specified at the beginning of Sect. 4.](#)

Table 5: Please provide the definition of F1, FPR and FNR, how they have been evaluated and their optimal values (see main remark n°3 in General comments).

[Done \(Sect 4.1.3\).](#)

Pag. 19, l. 4-5: Please consider moving the definition of “false positive” and “false negative” pixels in Section 4.2.1, where these terms are introduced for the first time.

[We believe that it could help the reader to interpret the results to let it here as well.](#)

Section 5 (Discussion)

Pag. 21, l. 10: Please provide a reference, if available.

[Done \(Gascoin et al. 2015\)](#)

Pag. 21, l. 25: Please consider removing of preceding the snow cover.

[Done](#)

Pag. 21, l. 26: According to my previous comment on Section 2, it would be useful to explain how the algorithm parameters have been defined.

[This is now explained in Sect. 2.4.4 \(see response to similar comments above\).](#)

Section 6 (Conclusion)

In this final section the authors provide useful information of key importance, which should be provided throughout the manuscript:

1. In the Introduction it should be explicitly specified that most snow products are derived from

Sentinel-2 and only recently from an integration of Sentinel-2 and Landsat-8 data.

[Done](#)

As well, at the beginning of Section 4 the authors should explicitly state that the evaluation was based only on Sentinel-2 products.

[Done](#)

2. In the section describing the methodology it should be properly explained that the analyses have been performed by considering the snow cover detected over the satellite field of view, and not at ground level (no correction of vegetation masking effect).

[This is an important comment and we are thinking of adjusting the snow maps based on the tree cover density for future products. However since this journal focuses on the presentation of data, we deliberately gathered all the limitations of the product in the conclusion to provide a synthetic caveat to the user.](#)

3. The condition of clear-sky should be explicitly mentioned as one criteria for the selection of the scenes to be analysed.

[If the referee refers to the Theia vs. SPOT snow maps, this was specified in Sect. 4.1.2 “We searched in the catalogue of the Kalideos database for available SPOT images that could match a cloud-free \(or nearly cloud-free\) Theia snow map”.](#)

Response to Javier Herrero (referee 2)

General comments

The authors present a high quality snow product derived from satellite information that is of good use to a multitude of studies and applications related to snow. The fact that it is especially oriented to mountainous areas, where it is more difficult for this type of products to provide a high quality standard, is a good indicator of the validity of this product and the algorithms used.

The improvements introduced in the snow detection algorithm and the availability and characteristics of the snow product seem very relevant and worthy of publication. The manuscript is well and clearly written, with a careful editing, ordering and presentation of the concepts.

We are grateful to the referee for this positive appreciation of our work.

Specific comments

My main specific comment refers to the presentation of the Theia collection as a mixed product of Sentinel-2 and Landsat-8. It is true that it will be in the future, but currently, and in what is presented in this manuscript, for all intents and purposes, it is only Sentinel-2. Therefore, I would consider withdrawing references to Landsat-8, especially from the title.

We agree that our manuscript largely focuses on Sentinel-2. That is the main novelty of our work and indeed our main motivation. In addition, the operational production began with Sentinel-2 products, therefore most of the available products at the time of writing were generated from Sentinel-2. However, we have developed the LIS processor to be compatible with Landsat-8 and an increasing number of Landsat-8 products are being made available in Theia. Therefore we would like to keep Landsat-8 in the manuscript. We have modified the abstract (*"We present the algorithm to generate the snow products and provide an evaluation of Sentinel-2 snow products accuracy"*) and conclusion to clarify this point:

*The Theia Snow collection is a free collection of snow products which indicate the presence or absence of snow on the land based on Sentinel-2 (20~m) and Landsat-8 (30~m) data. Most of the snow products are derived from Sentinel-2. **However, in late September 2018, Landsat-8 snow products started being routinely added to the Theia Snow collection. Previous Landsat-8 data were also reprocessed back to March 2017. At the time of writing (15 March 2019), for example in tile T31TCH in the Pyrenees, 20.5% of all available products were derived from Landsat-8 data (63 among 244). In addition, the processing is done to facilitate the combination of Landsat-8 and Sentinel-2 snow products since the data format and the tiling scheme are the same (only the pixel size differs). Although the co-registration of Sentinel-2 and Landsat-8 data can be still problematic (Storey et al. 2017), the combination of "analysis ready" Landsat-8 and Sentinel-2 is important to increase the number of cloud-free observations.***

(...) it remains a preliminary assessment that should be extended in the future. For example, the evaluation was based only on Sentinel-2 products, **since the integration of Landsat-8 data started during the writing of this manuscript and should be extended to Landsat-8 products. However, the processing is done with the same algorithm, which is based on the calculation of the NDSI from slope-corrected surface reflectance (level 2A data).**

Another comment arises on the fact that we do not understand that the validation with SPOT imagery was done for cloud-free data until page 16 (line 3). It is an important detail, which perhaps is not given due importance previously, because one of the presented strengths of the new algorithm is the treatment of the clouds. The validation for clouds discrimination lays on the visual verification, which is not as consistent as other methods. May be this can be addressed in future works, but it would be interesting at least that this issue became clearer throughout the manuscript.

Unfortunately very-high resolution images are generally not acquired when the sky is cloudy to save production cost. We agree that this is an important limitation of the evaluation and we have specified “very-high resolution **clear-sky** snow maps” in the introduction of the new version of manuscript. Note however that there are still some clouds in the SPOT scenes (but a very minor fraction of the image). There are some clouds too in the Sentinel-2 products, which were correctly identified (e.g. green polygons in Fig R1 below). Note that in the meantime, the MAJA cloud mask has been thoroughly validated using an active learning approach ([Baetens et al. 2019](#)), showing the high accuracy of MAJA cloud mask, which outperforms the one from Sen2Cor. This was added in Sect. 5.

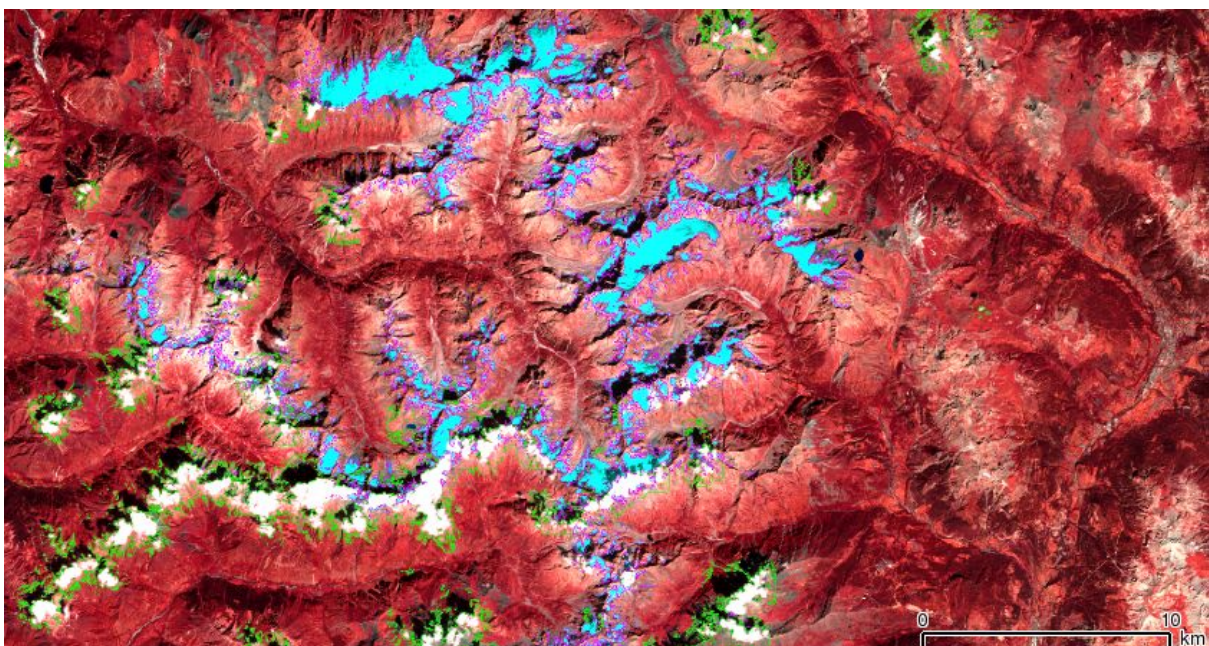


Fig R1: zoom over the Ecrins and Oisans on 2016-Aug-13 (snow covered outline: magenta, clouds: green).

The download of the Snow product through Theia portal (theia.cnes.fr) is a little bit messy in regard to the selection of tiles and dates (compared, for example, to Sentinel Hub portal).

Nevertheless, I was able to download a particular image with a delay of only 1 day, and the accuracy of the snow mask was impressive (for a clear day with patchy snow over Sierra Nevada). Very few shaded or partially-covered snow pixels are lost (false negatives), while no false positives are apparently seen. The product is really high quality.

[We thank the referee for this feedback.](#)

And a suggestion referred to the hard classification as snow/no snow pixels: obviously, the current resolution of Sentinel is superb. But maybe it would be worthwhile to generate some subpixel information, i.e. fractional snow cover. Probably with current technology it is possible to generate this product that could be important in semi-arid areas where patchy snow prevails over large areas during snowmelt periods.

[We take note of this suggestion, which echoes a recent call by the European Environment Agency to produce near-real time Fractional Snow Cover \(FSC\) at 20 m based on Sentinel-2 NDSI. This is indeed a perspective of our work since it would be straightforward to implement a simple relationship \$FSC=f\(NDSI\)\$ in LIS in order to provide the fractional snow cover area instead of a binary value \(snow/no-snow\). However, there is no current consensus on the shape and stability of such a relationship at the resolution of Sentinel-2 and further work shall be conducted to establish a robust algorithm.](#)

Technical corrections

- Page 4. Line 11. n_i and r_i are actually four parameters, not two.

[This is right, we changed to “parameter pairs”.](#)

- Page 4. Line 20. f_t is introduced without any definition or reference to Table 1. Perhaps a reference to the table would be pertinent in this paragraph, where most of the parameters that appear in it are presented.

[We have added references to the parameter Table 1 every time that a new parameter is introduced. This was also suggested by Referee 1.](#)

- Page 7. Figure 3. Caption. Where is Table 2.6?

[Corrected, it was Tab 1.](#)

- Page 8. Table 1. r_1 appears twice. The last one should be r_2 , I guess.

[Corrected, thank you.](#)

- Page 16. Table 3. Caption. Why the reference to Fig. 3?

[Corrected, it was Fig 9...](#)

- Page 16. Table 3. What is the criterion followed to order the rows?

[We sorted the rows by date in this revision.](#)

- Page 18. Line 22. Where is Fig. B4?

[It was Fig. A1 in Appendix. We apologize for all these confusions in the Figure labels.](#)

I want to add two technical corrections to the previous post, regarding the format of the data downloaded from the Theia portal:

- Page 13. Line 6. It seems that the productID_EXS_R2.tif file is not provided in the zipped file right now. The zip file contains then 8 files in total.

[Thank you for checking. Actually it depends on the product version. This file was stored in the root of the zip file in v 1.0, but in more recent versions it is placed in the MASKS folder](#)

within the zip file. We have modified the text and also added this item which was missing in the product description:

“- DATA/productID_HIS_R2.txt (since product version 1.4): a text file indicating the cloud, snow and no-snow fraction area by elevation bins (Sect. 2).”

See the product version details here:
(http://www.cesbio.ups-tlse.fr/multitemp/?page_id=13180).

- Page 13. Line 14. I think that you meant a quicklook picture of the productID_SNW_R2.tif image

We corrected this mistake, thank you!

Response to Jeff Dozier (referee 3)

This paper appears to be quite intriguing. However, in order to review it, I want to access some of the results, but I am not able to create an account at <https://theia.cnes.fr/>. At this point, my comments are rather sparse, but I would be willing to review the paper again if this problem with creating an account can be rectified.

We are sorry that the referee could not rectify this problem during the review process.

For the world's mountains, the "cirrus" band — Landsat 8 OLI (band 9), MODIS (band 26), or Sentinel-2 (band 10) — produces problematic output. The idea is that the wavelengths in this band are in a region of strong absorption by atmospheric water vapor, so that anything with high reflectance is likely a cirrus cloud. However, much of the snow of interest is at high elevation! In some areas of the world, for example, the Himalaya, almost all the snow is misclassified as cirrus in the Landsat 8 BQA file.

The cirrus detection with the 1.38 μm band is handled by the MAJA processor. The detection is done using a threshold which increases with altitude (see Hagolle et al., 2017 MAJA Algorithm Theoretical Basis Document <http://doi.org/10.5281/zenodo.1209633>)

One of the other referees has stated that the Theia archive mostly consisted of Sentinel-2A/B products, with little representation of Landsat 8 data. This is a shame, as the combination of the two systems provides imagery with a high temporal resolution.

The integration of Landsat-8 began in October 2018 and therefore the number of Landsat-8 products was not significant during the preparation of this manuscript. Since then the Landsat-8 data have been routinely produced and reprocessed back to March 2017. Therefore, we have modified the manuscript to highlight that a significant fraction (about 20% depending on the tile) of the products in the Theia Snow collection now consists of Landsat-8 products.

In this context, there is no discussion of the CFMask for Landsat 8. References to consider are: Foga, S., Scaramuzza, P. L., Guo, S., Zhu, Z., Dilley Jr, R. D., Beckmann, T., Schmidt, G. L., Dwyer, J. L., Hughes, M. J., and Laue, B.: Cloud detection algorithm comparison and validation for operational Landsat data products, *Remote Sensing of Environment*, 194, 379-390, doi 10.1016/j.rse.2017.03.026, 2017. Zhu, Z., Wang, S., and Woodcock, C. E.: Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images, *Remote Sensing of Environment*, 159, 269-277, doi 10.1016/j.rse.2014.12.014, 2015. In our experience, some clouds have Landsat 8 OLI signatures that can be similar to some snow, for example thin cirrus clouds compared to fractional snow.

We checked outputs of the Fmask algorithm (Zhu et al. 2015) during the early phase of the project and we also found issues in mountain regions in particular confusions between shaded slopes and water surface, snow and cloud confusion (Fig. R2 below). This also motivated us to develop our own algorithm using slope-corrected surface reflectances.

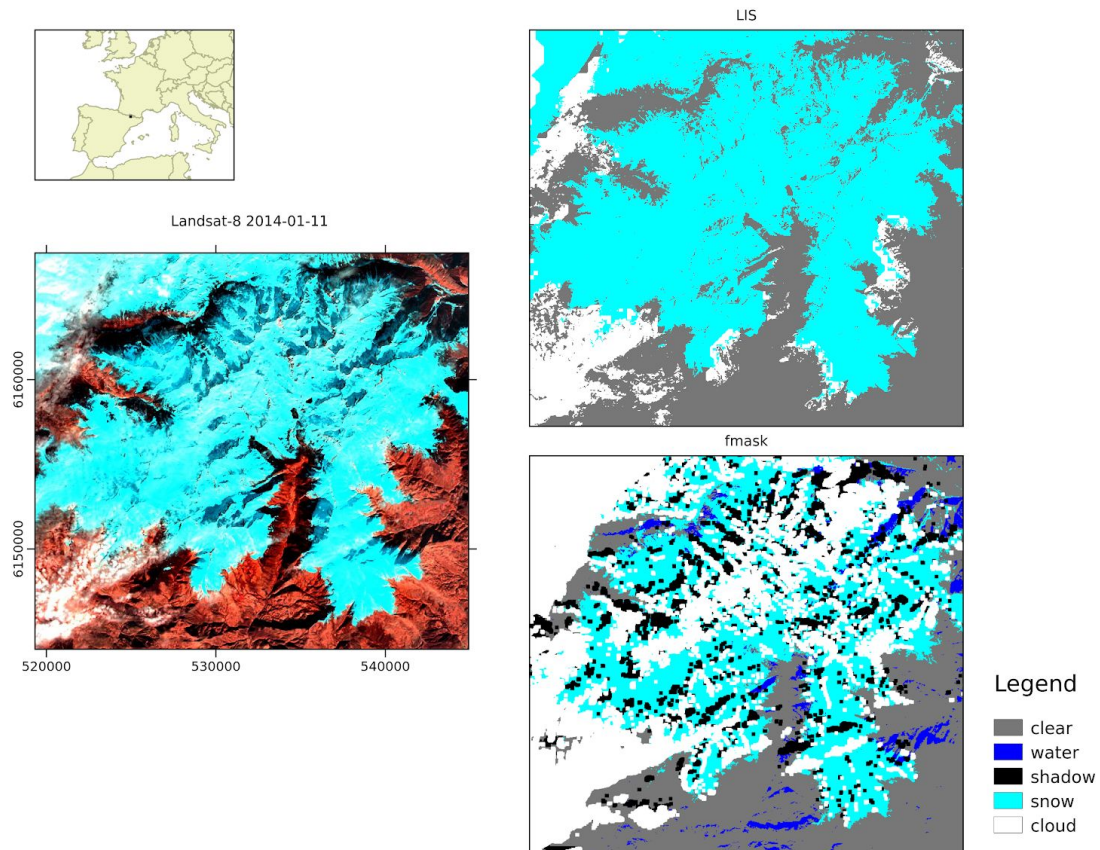


Figure R2: comparison of LIS and Fmask outputs for a Landsat-8 scene in the Central Pyrenees (2014-Jan-11).

Moreover, Selkowitz et al. have shown that in mountainous regions, the majority of Landsat snow pixels are not fully snow covered at 30 m resolution. Reference to consider is: Selkowitz, D. J., Forster, R. R., and Caldwell, M. K.: Prevalence of pure versus mixed snow cover pixels across spatial resolutions in alpine environments, *Remote Sensing*, 6, 12478-12508, doi 10.3390/rs61212478, 2014.

This is a very interesting study and we thank the referee for bringing it to our attention. The results suggest that in the study area with maritime climate the fraction of mixed pixels at the Sentinel-2 resolution is below 40%, whereas it can be 80% in a more continental mountain area. We also started looking at this issue recently by analyzing the fractional snow cover area from 6 m resolution SPOT-6 snow maps in the French Alps. The SPOT-6 sensor has a lower resolution than the one of WorldView used by Selkowitz et al. (2014) but covers larger areas (swath width at nadir: 60 km vs. 14.6 km). Our results (Fig. R4 below) suggest that the fraction of mixed pixels is lower. This might be due to the lower resolution of the original snow map (6 m) and the fact that Selkowitz et al. (2014) defined mixed pixels as pixels with FSC between 0.02 and 0.98, while in Fig. R4 we adopted a less strict representation by

choosing FSC bins of 10%. However, current state-of-the-art FSC algorithm do not reach an accuracy as low as 0.02. In another study, Selkowitz et al. 2017 report a root mean square error in FSC of 0.2. (“The USGS Landsat Snow Covered Area Products: Methods and Preliminary Validation.” [Available online](#)). Therefore the mixel pixel definition in the pioneer study by Selkowitz et al. (2014) seems rather strict for an assessment of the error that would be done in operational context by using binary snow maps instead of FSC.

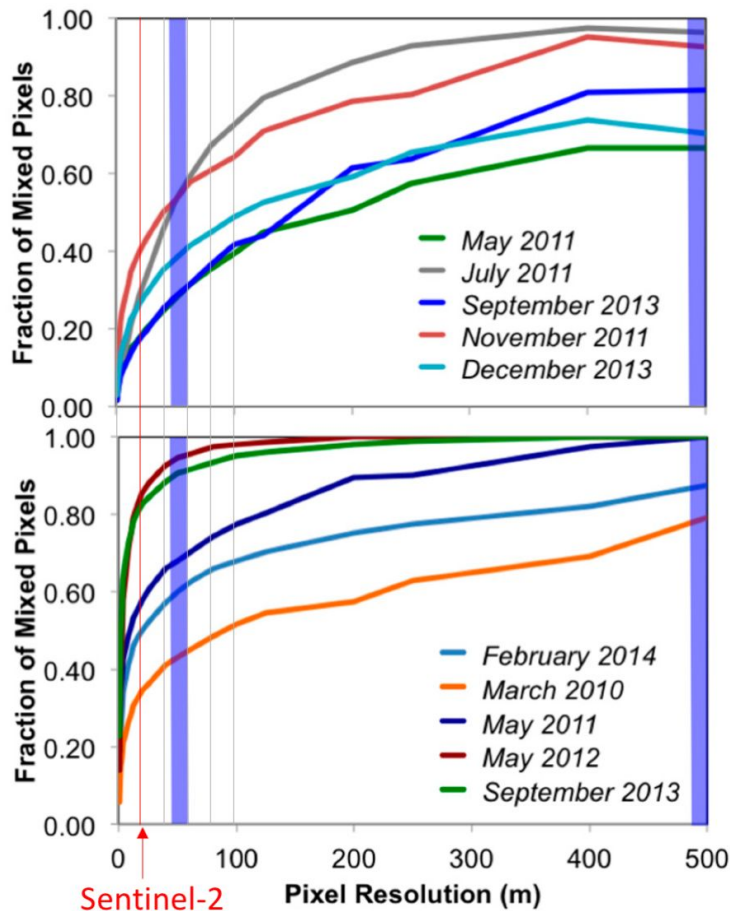


Figure R3: Fraction of partially snow-covered pixels in snow covered pixels, for pixel resolutions between 1 m and 500 m, for **(top)** a study area in the Oregon Cascades, and **(bottom)** a study area in the Rocky Mountain NP. Vertical red line indicates 20 m spatial resolution (Sentinel-2 Theia snow products). Vertical blue lines indicate 40 m spatial resolution (slightly above the nominal spatial resolution for Landsat) and 500 m spatial resolution (the nominal spatial resolution for MODIS) (source: modified from Selkowitz et al. *Remote Sens.* **2014**, 6(12), 12478-12508; doi:[10.3390/rs61212478](https://doi.org/10.3390/rs61212478))

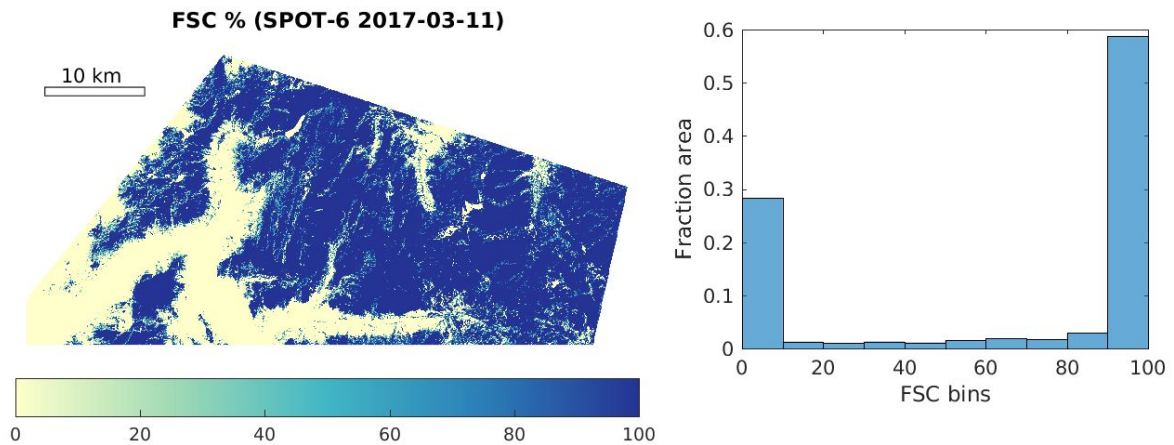


Figure R4: Fraction snow cover (FSC in %) from a SPOT-6 image at 6 m resolution in the French Alps (Oisans). The SPOT-6 FSC was computed as follows: (1) supervised classification of the image in snow and no-snow classes from the multispectral and panchromatic data at 1.5 m and 6 m resolutions (2) resampling of the binary snow cover map to a FSC map at the same resolution as Sentinel-2 SWIR band (20 m) by spatial averaging.

As we replied to Referee 2 above, we are actually considering to distribute near-real time FSC at 20 m based on Sentinel-2. It would be straightforward to implement a simple relationship $FSC=f(NDSI)$ in the LIS code. However, there is no current consensus on the shape and stability of such a relationship at the resolution of Sentinel-2 and further work shall be conducted to evaluate if a robust algorithm can be obtained from this function or if a more computer-intensive approach based on spectral unmixing should be considered (within the constraints of operational production).

As mentioned previously, I can take a closer look when I can get access to Theia.

We are sorry that the referee could not access the data during the period of open discussion. However, we are grateful to the referee for his insightful comments and we remain available for further discussion or assistance in accessing the data even after the open discussion of this paper.

Response to Referee 4

General comments:

The authors present the algorithm for the derivation of high-resolution operational snow cover maps from Sentinel-2 images. They further present an evaluation using station measurements, fine-scale satellite-derived snow maps as well as visual control. Overall, they found good performances though some tendencies were revealed. On cloudless days, the Theia snow product more likely underestimated snow coverage when compared to station measurements and fine-scale satellite-derived snow maps. When including cloudy days, the visual control with the false color composites revealed that the Theia snow product more likely overestimated snow coverage due to confusion with cloud coverage. This manuscript presents a complete package consisting of the algorithm to derive a satellite-derived snow product, a thorough evaluation as well as the access point to freely download the snow products via the Theia portal. It is therefore interesting for a broad range of scientists and practitioners. I already have an account at Theia and previously downloaded snow cover maps via the command line tools on Linux provided on github (https://github.com/olivierhagolle/theia_download). The manuscript is well written and I have only minor issues, which should be addressed before publication.

We appreciate the positive comments and we are pleased that the referee downloaded some products. We also thank the referee for the suggestions below.

Specific comments:

Since Landsat-8 data was not part of the snow product evaluation I would also suggest removing it everywhere (also in the title). However, I would keep mentioning it in the conclusions such that the reader is aware of the Landsat-8 snow maps.

As commented to Referee 2 we would like to keep Landsat-8 in the title because these products now represent a significant fraction of the Theia Snow collection (about 20%). We have modified the abstract to make it clear that the evaluation was done only with Sentinel-2 products. The conclusion was also modified to highlight this point.

It is only in the conclusion that the user learns he has to apply a forest mask to exclude forested areas from Theia snow maps. I think this should be mentioned earlier, e.g. as a bullet in the 2.1 Scope?

We agree with the referee. We have added this in Section 2.1: ***“An algorithm was designed to determine the snow presence or absence from Sentinel-2 and Landsat-8 observations outside areas of dense forest (...)”***

In section 2.2 you mention the correction for atmospheric and terrain slope effects in the Sentinel images. Can you cite something, since it is not explained or referenced later on?

We have added a reference to MAJA Algorithm Theoretical Basis Document.

In section 4.1.2 the authors present the evaluation with Spot images having a higher spatial resolution than Sentinel. I might not have fully understood the description of the Spot evaluation with Sentinel data. Did the size of each of the fifteen Spot polygons match one Sentinel pixel, such that fifteen pixels were evaluated for each of the six dates (or per tile)? If this is correct, how many of each were homogeneously snow-covered respectively snow free?

These polygons were not used for Sentinel-2. The fifteen polygons refer to the training data in the SPOT images that were used to perform a supervised classification of the snow cover area in the SPOT images. We hope that this clarifies this section.

Two of the selected six dates were in August and in October, i.e. in late summer/early fall. How much snow cover was there on 2016-08-13 in tile 131TGK and on 2016-10-12 in tile 32TLS? I guess the snow coverage could have an impact on the performances?

As responded to Referee 1, we find 174'603 snow covered pixels in the August product, mainly in the high elevation glacierized areas of Parc National des Ecrins, i.e. 0.6% of the image area (69.8 km²). In the October product we find 7.1% (856 km²) of the image pixels are classified as snow. We accounted for this asymmetry in the number of snow/no-snow by using the kappa coefficient and F1-score. Note that it is equally important to have a correct classification of snow and no-snow pixels. Therefore, an image with a small snow covered area has the same value as an image with a small snow-free area (e.g. winter image). With a summer image we validate that the snow product does not contain false snow detection (e.g. turbid lakes, bright urban areas, etc.)

Technical comments:

1) Please check all figure and table numbers, the references to them in the text as well as the order when they are introduced.

We apologize for the mistakes in the figures and table references. We have revised this thoroughly, also thanks to the comments of the other referees.

2) Page 19, line 5: Should this be: "False negative pixels (i.e. pixels falsely classified as snow-free) ..." ?

Thank you for noting this error, now corrected.

3) Page 21, line 1: "...an excellent accuracy.."

Corrected.

4) Page 21, line 6: "In addition, in situ measurements may not be representative of the snow conditions..."

Corrected.

Theia Snow collection: high resolution operational snow cover maps from Sentinel-2 and Landsat-8 data

Simon Gascoin¹, Manuel Grizonnet², Marine Bouchet¹, Germain Salgues^{2,3}, and Olivier Hagolle^{1,2}

¹CESBIO, Université de Toulouse, CNRS/CNES/IRD/INRA/UPS, Toulouse, France

²CNES, Toulouse, France

³Magellium, Toulouse, France

Correspondence: Simon Gascoin (simon.gascoin@cesbio.cnes.fr)

Abstract. The Theia Snow collection routinely provides high resolution maps of the snow cover area from Sentinel-2 and Landsat-8 observations. The collection covers selected areas worldwide including the main mountain regions in Western Europe (e.g. Alps, Pyrenees) and the High Atlas in Morocco. Each product of the Snow collection contains four classes: snow, no-snow, cloud and no-data. We present the algorithm to generate the snow products and provide an evaluation of **Sentinel-2** snow products accuracy using in situ snow depth measurements, higher resolution snow maps, and visual control. The results suggest that the snow is accurately detected in the Theia snow collection, and that the snow detection is more accurate than the **Sen2Cor** outputs (ESA level 2 product). An issue that should be addressed in a future release is the occurrence of false snow detection in some large clouds. The snow maps are currently produced and freely distributed in average 5 days after the image acquisition as raster and vector files via the Theia portal (DOI:10.24400/329360/F7Q52MKN).

10 1 Introduction

The snow cover is an important driver of many ecological, climatic and hydrological processes in mountain regions and in high latitude areas. In these regions, in situ observations are generally insufficient to characterize the high spatial variability of the snowpack properties. The snow cover was included as one of the 50 “Essential Climate Variables” (ECVs) to be monitored by satellite remote sensing by the Global Observing System for Climate (GCOS) in accordance with the Committee on Earth Observation Satellites (CEOS) agencies. ECVs are intended to support the work of the United Nations Framework Convention on Climate Change and the Intergovernmental Panel on Climate Change.

The snow cover is not a variable *sensu stricto*, but an object which can be characterized through many variables, including: snow cover area (SCA), fractional area (fSCA), albedo, liquid water content, snow depth and snow water equivalent (Frei et al., 2012). An international survey by the Cryoland consortium¹ shed light on the user requirements for snow services based on satellite remote sensing. SCA/fSCA products were ranked as the most important products by the respondents (Malnes et al., 2015). The survey also indicated the need for operational products provided on a daily basis, with latency time shorter than 12 hours. High resolution data down to 50 m resolution were sought by road and avalanches authorities (Malnes et al., 2015).

¹<http://cryoland.enveo.at/consortium>

The respondents requested regional products e.g. at the scale of entire mountain range like the Alps or even Pan-European and preferred the UTM projection (Malnes et al., 2015). At the national level in France there is also a need for operational high resolution snow cover area maps as revealed by the recent roadmap for satellite applications issued by the French Government (*Plan d'applications satellitaires 2018*, Commissariat général au développement durable, 2018). Based on a wide panel of end-users, this plan selected the monitoring of the snow cover area in French National Parks as a one of the key actions which should be achieved using Earth Observation satellites in the near future (2018-2022).

Operational SCA maps have been generated from satellite observations since the 1960's (Ramsay, 1998). Current SCA and fSCA products are mostly derived from MODIS data (Hall et al., 2002; Sirguey et al., 2008; Painter et al., 2009; Metsämäki et al., 2015), but their spatial resolution (1 km to 250 m) can be too coarse for various applications, especially in mountain regions where the snow cover properties vary at scales of 10 m to 100 m (Blöschl, 1999). High resolution (30 m) snow cover maps can be generated from Landsat images but the low temporal revisit of the Landsat mission (16 days) is an important limitation for snow cover monitoring, especially considering that the cloud probability can exceed 50% in temperate mountains (Parajka and Blöschl, 2008; Gascoin et al., 2015). Since 2017 with the launch of Sentinel-2B, the Copernicus Sentinel-2 mission offers the unique opportunity to map the snow cover extent at 20 m resolution with a revisit time of **5 days** (cloud-permitting) (Drusch et al., 2012). The combination of Sentinel-2 and Landsat-8 data provides the opportunity for even more frequent observations of the snow cover with a global median average revisit interval of 2.9 days (Li and Roy, 2017).

The principles of snow detection from multispectral optical imagery is well established since the pioneering work of Dozier (1989) with the Landsat Thematic Mapper. Today, the challenge for scientists and end-users is rather to cope with the large amount of data that is generated by a mission like Sentinel-2. The generation of a single snow map from a Sentinel-2 level-1C tiled product (**monodate ortho-rectified image expressed in top-of-atmosphere reflectance**) involves the download of a 700+ Mo zip file (once uncompressed, the product contains 12 folders and 108 files, including 15 raster files and 13 XML metadata files). Since March 2018, ESA began the operational processing of level 2A (L2A) products (**monodate ortho-rectified images expressed in surface reflectance, including a cloud mask**). Each ESA L2A product also includes a snow cover mask. However, the size of a single L2A product before unzipping can exceed 1 Go (15 folders, 137 files). In addition, the quality of the ESA L2A snow mask can be improved for two reasons (i) ESA's L2A algorithm is a general purpose algorithm, which was not optimized for snow detection (ii) Because ESA's L2A processor treats each image independently (i.e. monodate approach, Louis et al., 2016), the output cloud mask has a lower accuracy than a cloud mask generated by multi-temporal algorithm (Hagolle et al., 2010).

In this article we introduce the Theia Snow collection, a new collection of snow cover maps, which are derived from Sentinel-2 at 20 m resolution in an operational context. **The Theia Snow collection was recently upgraded to also provide snow cover maps at 30 m resolution from Landsat-8 using the same algorithm.** The data are currently being produced and freely distributed via the Theia land data center ². The snow retrieval algorithm is based on the Normalized Difference Snow Index (Dozier, 1989), but also uses a digital elevation model to better constrain the snow detection. We first describe the algorithm

²Theia is a French inter-agency organization designed to foster the use of Earth observation for land surface monitoring for academics and public policy actors.

(Sect. 2) and its implementation in the let-it-snow processor (LIS, Sect. 2.4.4). Then, we provide a detailed description of the product characteristics (coverage, period, format, etc., Sect. 3). In Section 4, we present an evaluation of Theia snow products using in situ snow depth measurements, very-high resolution **clear-sky** snow maps and visual control. We finally discuss the main limitations and potential applications of the Theia Snow collection.

5 2 Algorithm

2.1 Scope

An algorithm was designed to determine the snow presence or absence from Sentinel-2 and Landsat-8 observations **outside areas of dense forest** with the following requirements:

- It should be scalable, i.e. allow the processing of large areas (10^4 km²) with a reasonable computation cost (typically less than 1 hour for a single product).
- It should be robust to seasonal and spatial variability of the snow cover and land surface properties.
- It should maximize the number of pixels that are classified as snow or no-snow.
- It is preferable to falsely classify a pixel as cloud than falsely classify a pixel as snow or no-snow.

2.2 Input

15 The algorithm works with multispectral **remotely sensed** images, which include at least a channel in the visible part of the spectrum and a channel near 1.5 μ m (referred to as shortwave-infrared or “SWIR”). It takes as input:

- a L2A product, including:
 - the cloud and cloud shadow mask (referred to as “L2A cloud mask” in the following),
 - the green, red and SWIR bands from the flat surface reflectance product. These images are corrected for atmospheric and terrain slope effects (**Hagolle et al., 2017**). The slope correction is important in mountain regions since it enables to use the same detection thresholds whatever the sun-slope geometry.
- A digital elevation model (DEM).

2.3 Pre-processing

20 The red and green bands are resampled with the cubic method to a pixel size of 20 m by 20 m to match the resolution of the SWIR band. The DEM is generated from the **Shuttle Radar Topography Mission** seamless DEM (Jarvis et al., 2008) by cubic spline resampling to the same 20 m resolution grid.

2.4 Algorithm description

2.4.1 Snow detection

The snow detection is based on the Normalized Difference Snow Index (NDSI, Dozier, 1989) and the reflectance in the red band. The NDSI is defined as:

$$5 \quad \text{NDSI} = \frac{\rho_{\text{green}} - \rho_{\text{SWIR}}}{\rho_{\text{green}} + \rho_{\text{SWIR}}} \quad (1)$$

where ρ_{green} (resp. ρ_{SWIR}) is the slope-corrected surface reflectance in the green band (resp. SWIR at $1.6 \mu\text{m}$). The NDSI expresses the fact that only snow surfaces are very bright in the visible and very dark in the shortwave infrared. Turbid water surfaces like some lakes or rivers may also have a high NDSI value, hence a additional criterion on the red reflectance is used to avoid false snow detection in these areas. A **cloud-free** pixel is classified as “snow” if the following condition is true:

$$10 \quad (\text{NDSI} > n_i) \text{ and } (\rho_{\text{red}} > r_i) \quad (2)$$

where n_i and r_i are two **parameter pairs** (i.e. $i \in \{1, 2\}$) as explained below. Otherwise the pixel is marked as “no-snow”.

The values of the parameters are provided in Tab. 1.

2.4.2 Snowline elevation

The algorithm works in two passes. The snow detection (Sect. 2.4.1) is performed a first time using parameters n_1 and r_1 ,
15 which are set in such way as to minimize the number of false snow detections (**Tab. 1**). This first pass enables to estimate the minimum elevation of the snow cover z_s , above which a second pass will be performed with less conservative parameters values n_2 and r_2 (**Tab. 1**). **An example of the evolution of the snow line elevation over the tile 31TCH in the Pyrenees is shown in Fig. 1 and the corresponding snow maps are shown in Fig. 1.** Based on regional analyses of the snowline elevation variability in mountain ranges, (e.g. Gascoin et al., 2015; Krajčič et al., 2016), we assume that large altitudinal variations of
20 the snowline elevation are not likely at the scale of a tile of 110 km by 110 km. Therefore, the minimum snow elevation z_s is considered uniform at the scale of the processed image.

Before proceeding to pass 2, the total snow fraction **in the image** after pass 1 is computed. If this snow fraction is below f_t (**Tab. 1**), then pass 2 is skipped as the sample of snow pixels is not considered as statistically significant to determine the snow line elevation. Otherwise, the pass 2 is activated. For that purpose, the DEM is used to segment the image in elevation bands
25 with a fixed height of d_z (**Tab. 1**). Then, the fraction of the cloud-free area that is covered by snow in each band (after pass 1) is computed. The algorithm finds the lowest elevation band b at which the snow cover fraction is greater than a given fraction f_s (**Tab. 1**). The value of z_s is the lower edge of the elevation band that is two elevation bands below band b .

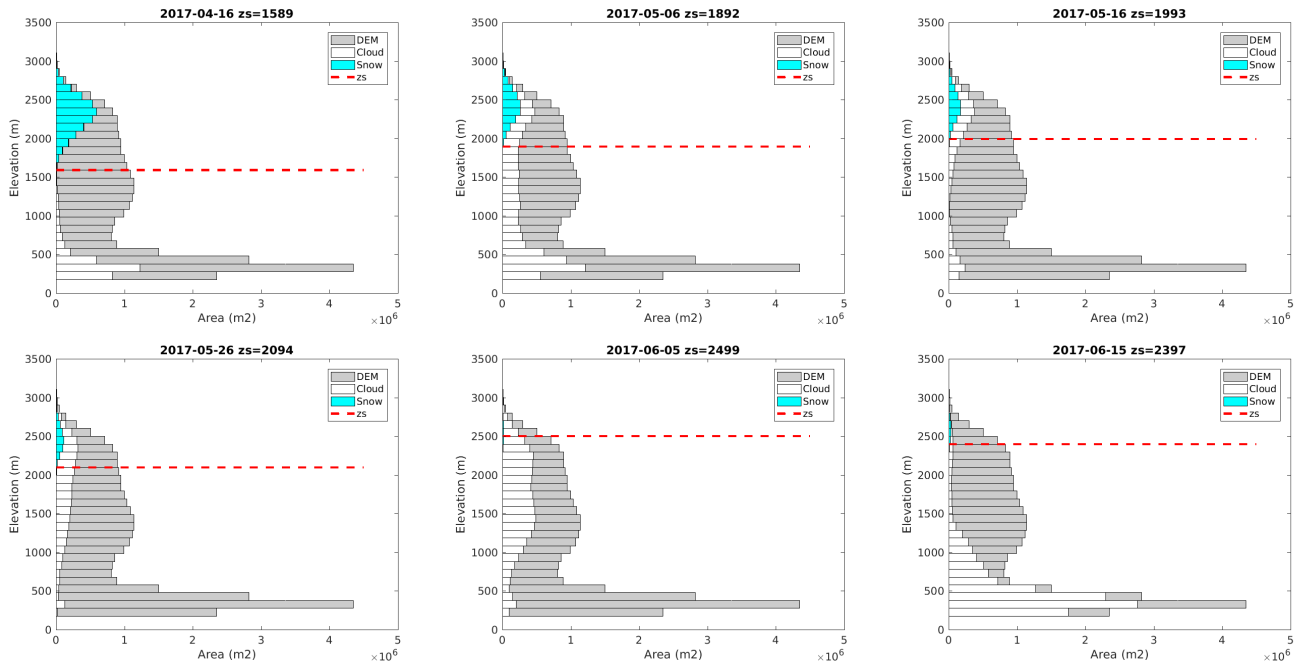


Figure 1. Time series of snow and cloud cover area distribution by elevation band over tile 31TCH (tile location in Fig. 5). The red dashed line indicates the position of z_s as determined by the LIS algorithm. The corresponding snow maps are shown in Fig. 2.

2.4.3 Cloud mask processing

The cloud mask in the input L2A product is conservative because (i) it is computed at a coarser resolution and (ii) it was developed to remove surface reflectance variations due to cloud contamination. However, the scattering of some thin clouds is low in the SWIR, green and red bands which are used for snow detection (Sect. 2.4.1). Hence, the human eye can “see” the snow (or the absence thereof) through these semi-transparent clouds in a color composite. In addition, the L2A cloud mask tends to falsely classify the edges of the snow cover as cloud. Therefore, the algorithm includes some additional steps to recover those pixels from the L2A cloud mask and reclassify them as snow or no-snow. This step is important because it substantially increases the number of observations as specified in Sect. 2.1.

A pixel from the L2A cloud mask cannot be reclassified as snow or no-snow if any of these conditions is satisfied:"

- it is coded as “cloud shadow” in the L2A cloud mask;
- it is coded as “high altitude cloud” (or “cirrus”) in the L2A cloud mask;
- it is not a “dark cloud” (see below).

The cloud shadows are excluded because the signal-to-noise ratio can be very low in these areas. The “high clouds” are excluded because they can have a similar spectral signature as the snow cover, i.e. a high reflectance in the visible and a low

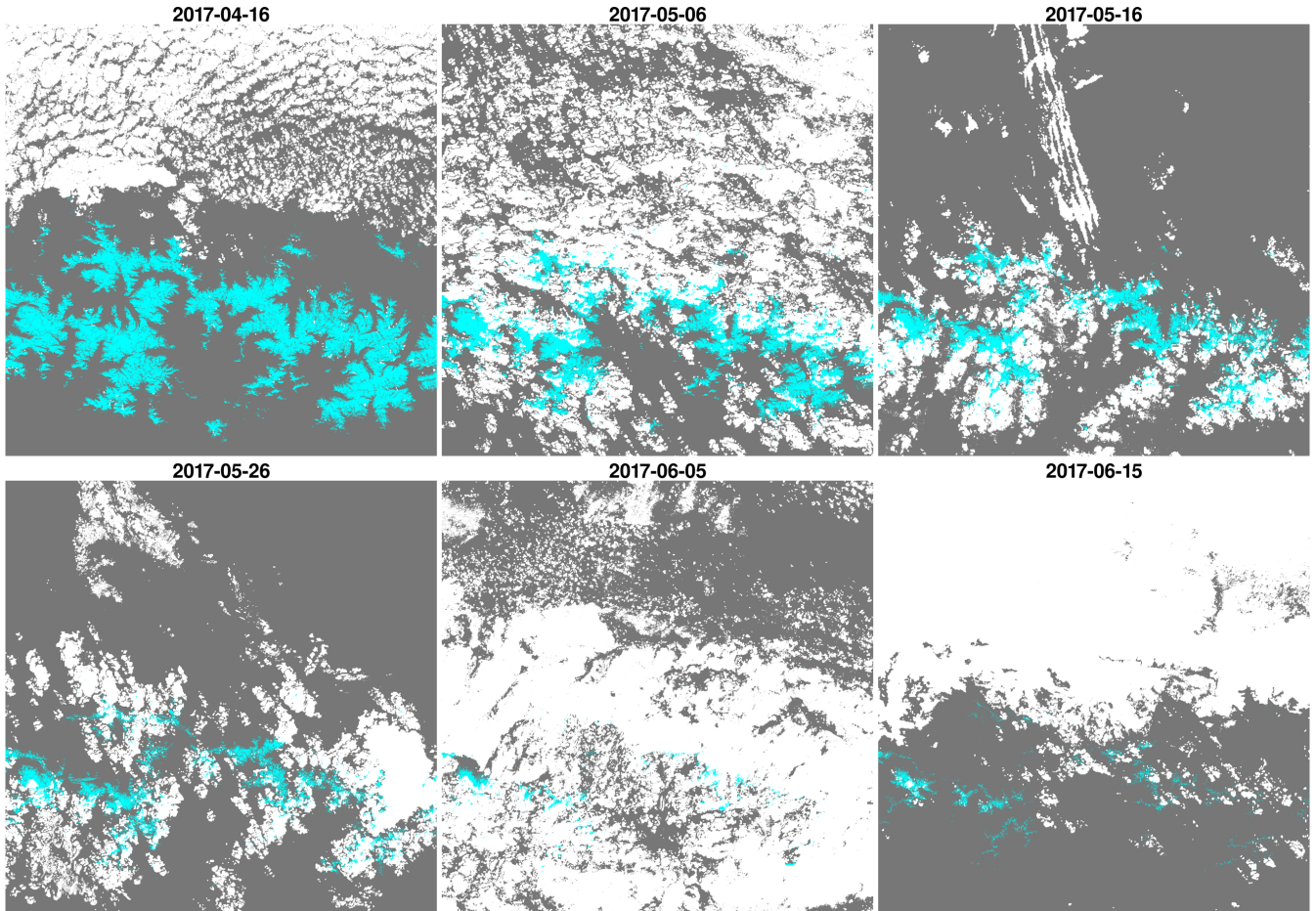


Figure 2. Time series of snow maps over tile 31TCH (tile location in Fig. 5). The corresponding charts of snow and cloud cover area distribution by elevation band are shown in Fig. 1.

reflectance in the SWIR. This type of cloud are detected in Sentinel-2 and Landsat-8 L2A products based on the spectral band centered on the $1.38 \mu\text{m}$ wavelength (Gao et al., 1993).

We select only the “dark clouds” as possible reclassification areas, because the NDSI test is robust to the snow/cloud confusion in this case. The “dark clouds” are defined using a threshold in the red band after down-sampling the red band by a factor r_f using the bilinear method. This resampling is applied to smooth locally anomalous pixels, following the MAJA algorithm, which performs the cloud detection at 240 m for L2A products (Hagolle et al., 2017). Therefore, if a (non-shadow, non-high-cloud) cloud pixel has a red reflectance at this coarser resolution that is lower than r_D (**Tab. 1**), then it is temporarily removed from the cloud mask and proceeds to the pass 1 snow detection. The new cloud mask at this stage is the pass 1 cloud mask in Figure 3.

Table 1. LIS algorithm parameters description and default values for Theia’s L2A Sentinel-2 (and Landsat-8) products.

Parameter	Description	Value
r_f	Resize factor to produce the down-sampled red band	12 (8)
r_D	Maximum value of the down-sampled red band reflectance to define a “dark cloud” pixel	0.300
n_1	Minimum value of the NDSI for the pass 1 snow test	0.400
n_2	Minimum value of the NDSI for the pass 2 snow test	0.150
r_1	Minimum value of the red band reflectance the pass 1 snow test	0.200
r_2	Minimum value of the red band reflectance the pass 2 snow test	0.040
d_z	Size of elevation band in the DEM used to define z_s	100
f_s	Minimum snow fraction in an elevation band to define z_s	0.100
f_{c_t}	Minimum clear pixels fraction (snow and no-snow) in an elevation band to define z_s	0.100
f_t	Minimum snow fraction in the image to activate the pass 2 snow test	0.001
r_B	Minimum value of the red band reflectance to return a non-snow pixel to the cloud mask	0.100

After passing the pass 1 and 2, some former cloud pixels (pixels that were originally marked as cloud in the L2A cloud mask) will not be reclassified as snow because they did not fulfill the conditions of Equation 2.4.1. These pixels are flagged as cloud in the final snow product if they have a reflectance in the red that is greater than r_B (**Tab. 1**). Otherwise they are classified as no-snow. Here the red band at 20 m resolution is used, to allow an accurate detection of the snow-free areas near the snow cover edges that the L2A product tend to falsely mark as cloud. The resulting cloud mask is the pass 2 cloud mask in Figure 3.

2.4.4 Implementation

The algorithm was implemented in an open source processor called let-it-snow (LIS). LIS is written in Python 2.7 and C++ and relies on the Orfeo Toolbox and GDAL libraries. GDAL is used for input/output operations, image resampling, and meta-data access (GDAL/OGR contributors, 2018). Orfeo Toolbox enables to make the computations with good performances under memory constraints (i.e. the amount of available memory can be predefined), which is critical for operational production (Grizonnet et al., 2017). LIS takes as input a digital elevation model and a Theia L2A product from Sentinel-2 MSI, Landsat-8 OLI, SPOT-4 HRVIR, or SPOT-5 HRG (Gascoin et al., 2018). **The spatial resolution and central wavelength of the spectral bands used by LIS for each sensor are given in Tab. 2.** It is also compatible with Landsat-8 USGS level 2 products (U.S. Geological Survey Earth Resources Observation And Science Center, 2014) and Sentinel-2 ESA’s **Sen2Cor** level 2A products (Louis et al., 2016).

The parameter values were set based on previous studies with Landsat (Hagolle et al., 2010; Zhu et al., 2015; Gascoin et al., 2015) and by visually checking many snow maps and snow fraction histograms. From this set of a priori parameter values, only r_2 was adjusted based on the analysis of a first batch of products to enhance the snow detection in shaded slopes.

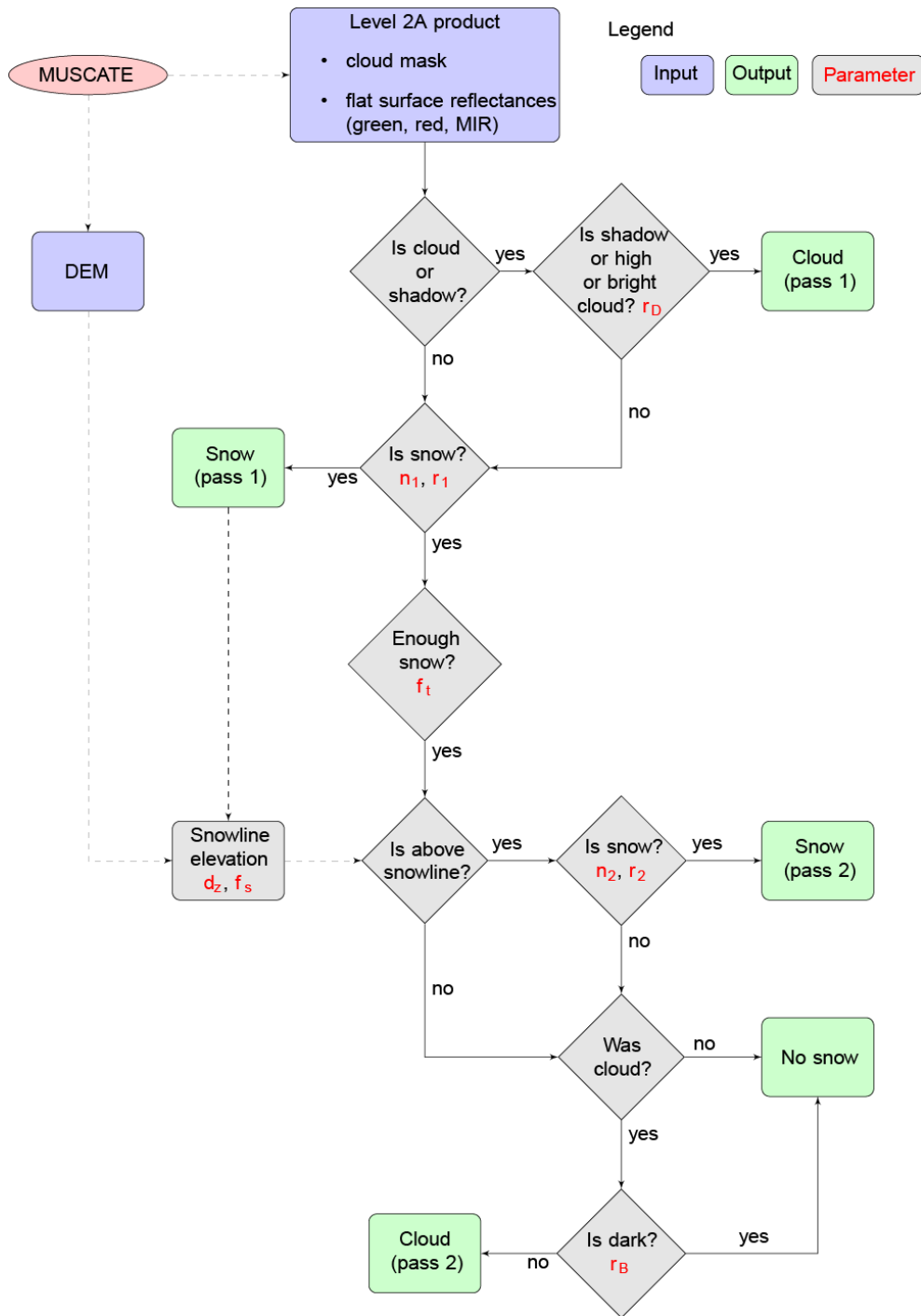


Figure 3. Flowchart of the snow detection algorithm. Table 1 gives the description and value of the algorithm parameters (written in red in this chart). MUSCATE is the scheduler which manages the L2A production for Theia.

Table 2. Spatial resolution and central wavelength of the spectral bands used by LIS for each compatible sensor.

Sensor	Green band	Red band	SWIR band
SPOT-4 HRV	20 m, 0.55 μm	20 m, 0.65 μm	20 m, 1.6 μm
SPOT-5 HRG	10 m, 0.55 μm	10 m, 0.65 μm	10 m, 1.6 μm
Sentinel-2 MSI	10 m, 0.56 μm	10 m, 0.66 μm	20 m, 1.6 μm
Landsat-8 OLI	30 m, 0.56 μm	30 m, 0.65 μm	30 m, 1.6 μm

3 Data description

The Snow Collection data can be freely accessed using a web browser by connecting to <http://theia.cnes.fr> or using this free command-line tool: https://github.com/olivierhagolle/theia_download. The user must first create an account in Theia to have the permission to download the data.

- 5 The Theia Snow collection is organized following the tiling system of Sentinel-2 (Fig.4). Each tile covers a square area of 110 km by 110 km in the Universal Transverse Mercator (UTM) coordinate system. There are currently 127 tiles in the Snow collection, mostly over the mountain ranges of western continental Europe (France, Spain, Switzerland, Italy, western Austria). The snow products are also provided for the southern Quebec in Canada, the Issyk-Kul lake catchment in Kirghistan, the Kerguelen Islands. These extra-European sites were selected to support specific ongoing projects in relation with Theia.
- 10 An opportunistic tile is produced in Central Nevada, USA, because this tile is available as level-2A anyway for calibration purposes. The Theia Snow collection covers a number of mountain ranges with seasonal snow cover (Table 3).

The Theia Snow collection products are available for 127 tiles starting December 2017. At the time of writing there were over 17'000 available products. A set of 1300 products tagged as version 1.0 are available between November 2015 and June 2017 for a subset of 15 tiles (Tab. 3). These products were produced in pre-operational using a different configuration of LIS,

15 which underestimated the snow cover area in shaded slopes. All other products were generated with the same configuration as presented in this article. The products after December 2017 can still have a different version tag because of changing versions in the upstream L2A product. However, the different L2A versions have a low impact on the snow products. It is expected that the cloud mask has improved gradually with time, due to algorithms enhancements in the successive L2A versions, and also because of the increasing revisit frequency of the Sentinel-2 mission (the full 5-day revisit became nominal

20 above all land masses in March 2018). The different versions of the Snow Collection are listed and updated in this page: http://www.cesbio.ups-tlse.fr/multitemp/?page_id=13180.

The snow products are currently routinely generated at CNES using the MUSCATE scheduler, which also manages the L2A production for Theia. MUSCATE performs a series of operations in high performance computing environment (i) download the level 1C product as soon it is available from the French mirror of the Sentinel products (PEPS, Plateforme d'Exploitation

25 des Produits Sentinel), (ii) process and distribute the L2A product (iii) process and distribute the snow product. This workflow

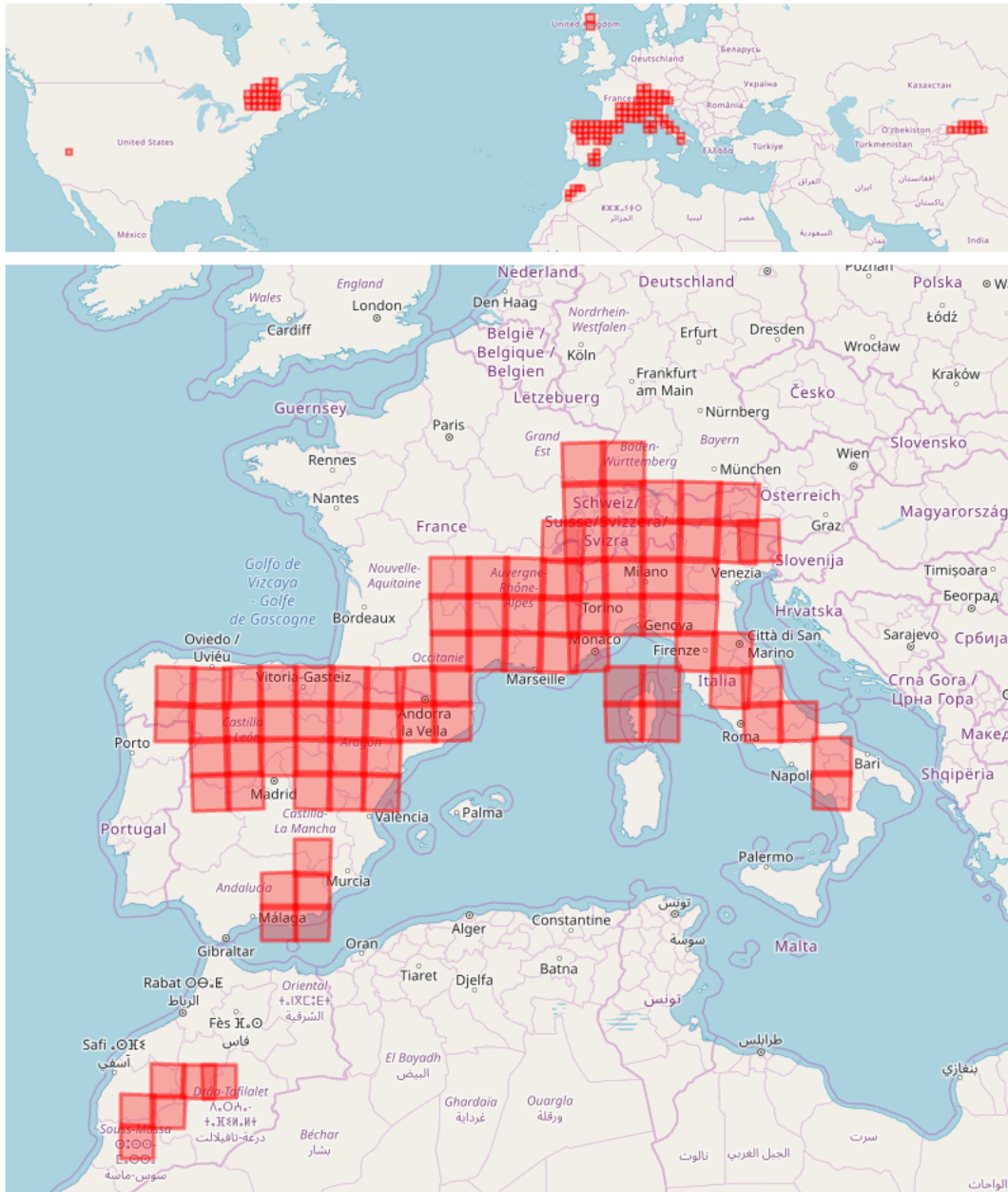






Figure 4. Coverage of the Theia Snow collection tiles



Legend

-  Tile extent
- Snow product
 -  No snow
 -  Snow
 -  Cloud including cloud shadow

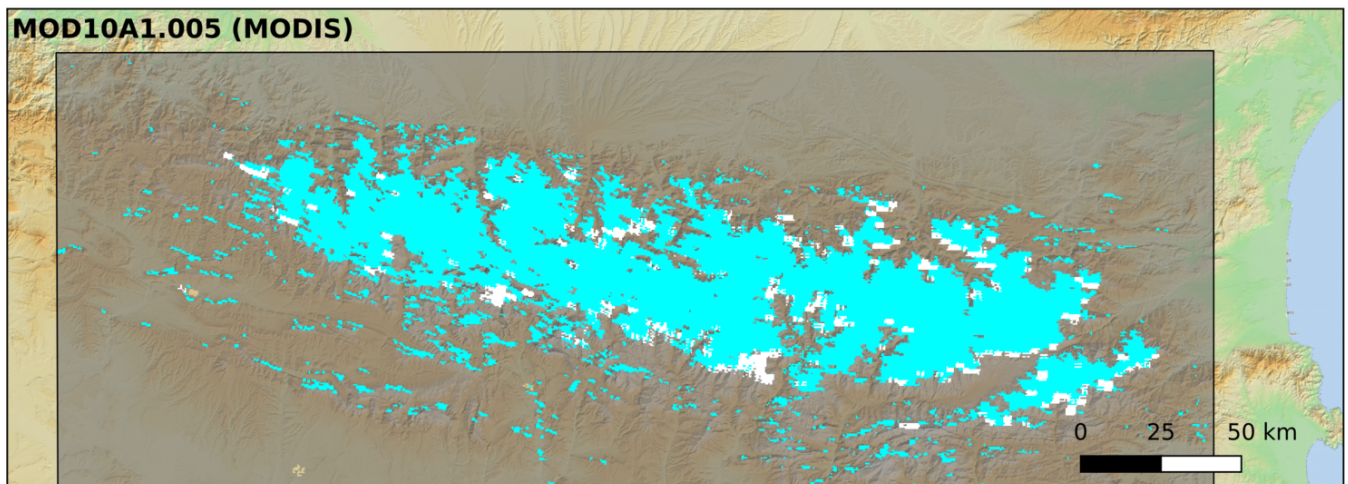
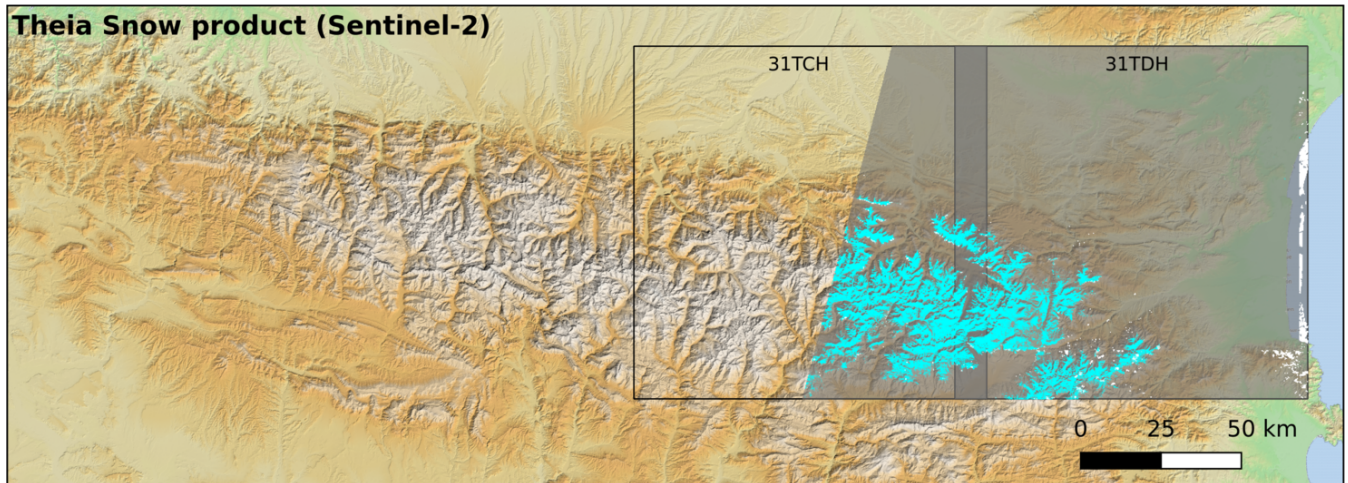


Figure 5. Map of the first two products in the Theia Snow collection and comparison with MOD10A1.005 snow product of the same day. The map background is the colored version of the EU-DEM.

Table 3. Non exhaustive list of mountain ranges covered by the Theia Snow collection and corresponding files. The tiles marked with an asterisk were made available before the start of the operational production (see Sect. 3)

Mountain range	Tiles
Alps	32TMS,32TMR,32TMT,32TNS,32TNR,32TLQ*,32TLP,32TLS*,32TLR*,32TQT,32TQS,33TUM,31TGK*,32TNT,31TGJ,32TPR,31TGL*,32TPT,32TPS
Appenines	33TUH,32TNQ,33TUG,33TVG,33TWF,32TPP,33TWE,32TPQ,32TQP,32TQN
Cantabrian	30TUN,29TPH,29TPG,29TQH
Corsica	32TMM,32TMN,32TNM, 32TNM
Grant Range	11SPC
High Atlas	29RPQ*,29RNQ*,29SQR*,29SPR*,30STA*
Sistema Iberico	30TVM,30TWK,30TWM,30TWL,30TXK,30TXL
Jura	31TGM*,32TLT
Massif Central	31TDL,31TDK,31TEK,31TEL
Pyrenees	30TXN*,30TYM,30TYN*,31TCG,31TDG,31TCH*,31TDH*
Sierra de Cazorla	30SWH
Sierra Nevada	30SVG,30SWG
Sistema Central	30TUL,30TUK,30TVL,30TTK
Vosges	32TLT,32ULU

These tiles can be viewed on an interactive map in this page <https://theia.cnes.fr/atdistrib/rocket/#/sites>

is optimized to reduce the lag time between the acquisition and the distribution of the products. As an example, in April 2018, nearly 80% of the snow products were made available online 5 days or less after the date of the Sentinel-2 acquisition (Fig. 6). It is expected that this time lag should reach a median value of 2 days thanks to the increasing performance of MUSCATE.

Each snow product is distributed as a zipped file which contains raster and vector files. The file base name is the product ID,

5 i.e. `productId.zip`, where the `productId` is a character string which is constructed as follows:

`productId = Satellite_AcquisitionDate_L2B-SNOW_TileName_D_Version.`

For example, the eastern snow product in Fig. 5 was extracted from a file named:

`SENTINEL2A_20151130-105641-486_L2B-SNOW_T31TDH_D_V1-0.zip`,

which indicates that this product was generated from a Sentinel-2A image acquired on 30 November 2015 at 10:56:41.486

10 UTC and covers tile 31TDH. The version of the product is 1.0 (see above).

The size of each zipped product varies depending on the complexity of the snow and cloud mask, but generally ranges between 10 Mo and 100 Mo. After extracting the product the following files are created:

- `productId_SNW_R2.tif`: an 8-bit single band GeoTiff image which provides the following classification for each pixel:

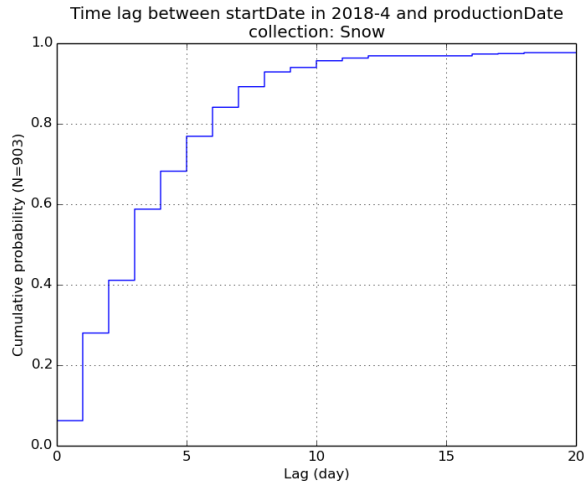


Figure 6. Cumulative probability of the time lags between the Sentinel-2 acquisition and the distribution date of the corresponding snow product, computed for all snow products with acquisition date in April 2018.

- 0: no-snow,
 - 100: snow,
 - 205: cloud including cloud shadow,
 - 254: no data.
- 5
- `productID_SNW_R2.{shp,shx,dbf,prj}`: a vector file in the ESRI Shapefile format, which contains a vectorized version of the `productID_SNW_R2.tif` using the polygon geometry.
 - `productID_CMP_R2.tif`: an 8-bit RGB GeoTiff image which shows the outlines of the snow mask (green lines) and cloud mask (magenta lines) on a false-color composite of the input L2A image (RGB image made with bands SWIR/Red/Green). This composition was chosen because RGB composites using the SWIR band image are useful to discriminate the snow cover from the snow-free areas and from the clouds (Vidot et al., 2017). Hence, this image mainly allows the user to visually check the consistency of the snow and cloud mask.
- 10
- `productID_MTD_ALL.xml`: a metadata file.
 - `productID_QKL_ALL.jpg`: a quicklook picture of the `productID_SNW_R2.tif` image made using this colormap (8-bit RGB code in parentheses): snow in cyan (0,255,255), cloud in white (255,255,255), no-snow in grey (119,119,119) and no-data in black (0,0,0) (see an example in **Fig. 7**).
- 15
- `DATA/productID_HIS_R2.txt` (since product version 1.4): a text file indicating the cloud, snow and no-snow fraction area by elevation bins (Sect. 2).

– MASKS/productID_EXS_R2.tif (in product version 1.0, it is located in the root of the zipped file): an 8-bit GeoTiff image for expert evaluation, where each bit indicates the value of an intermediate computation mask (Sect. 2):

- bit 1: snow (pass 1),
- bit 2: snow (pass 2),
- 5 – bit 3: clouds (pass 1),
- bit 4: clouds (pass 2),
- bit 5: clouds (initial all cloud),

For most applications, only files with the SNW suffix should be useful.

The shapefile and raster images are referenced in WGS-84 UTM system with the zone number given by the first two digits of the tile name. All raster files have a 20 m resolution. **Note that no-snow pixels can be any surface, including water surface. No data pixels are the pixels outside of the acquisition segment.**

4 Evaluation

In this section we present an evaluation of the Snow collection. We first present the methods developed to do this evaluation and then the results. **The evaluation was based only on Sentinel-2 products, because the production for Landsat-8 had just started at the time of writing and thus the Landsat-8 products represented a very minor fraction of the Theia Snow collection.**

4.1 Method

4.1.1 Comparison with in situ snow depth measurements

We collected all available snow depth measurements within tiles 31TGM, 31TGL, 31TGK, 32TLS, 32TLR, 32TLQ, 31TDH, 31TCH, 30TYN and 30TNX from the Météo-France database between 01-Sep-2017 and 31-Aug-2018. These 10 tiles cover most of the french Alps and Pyrenees (**Fig. 8**). We obtained 120 snow depth time series with at least one measurement. We gathered all available Sentinel-2 snow products for these tiles over the same period, i.e. a total of 1134 products. Then, we extracted the pixel values at the location of each snow measurement stations for all dates. When a station is located in two tiles we only kept the data from the first tile. We selected the snow depth measurements which were collected on the same day of a Theia snow product and discarded the measurements corresponding to a cloud detection in the Theia snow product. The snow depth measurements were converted to snow presence and absence using a threshold of $SD_0=0$ m. This eventually allowed us to compute a confusion matrix between a set of snow presence and absence data from in situ measurement and a set of simultaneous snow presence and absence data from the Theia Snow collection across the French Alps and Pyrenees. However, previous studies comparing MODIS binary snow products with ground measurements showed that a value of 0 m

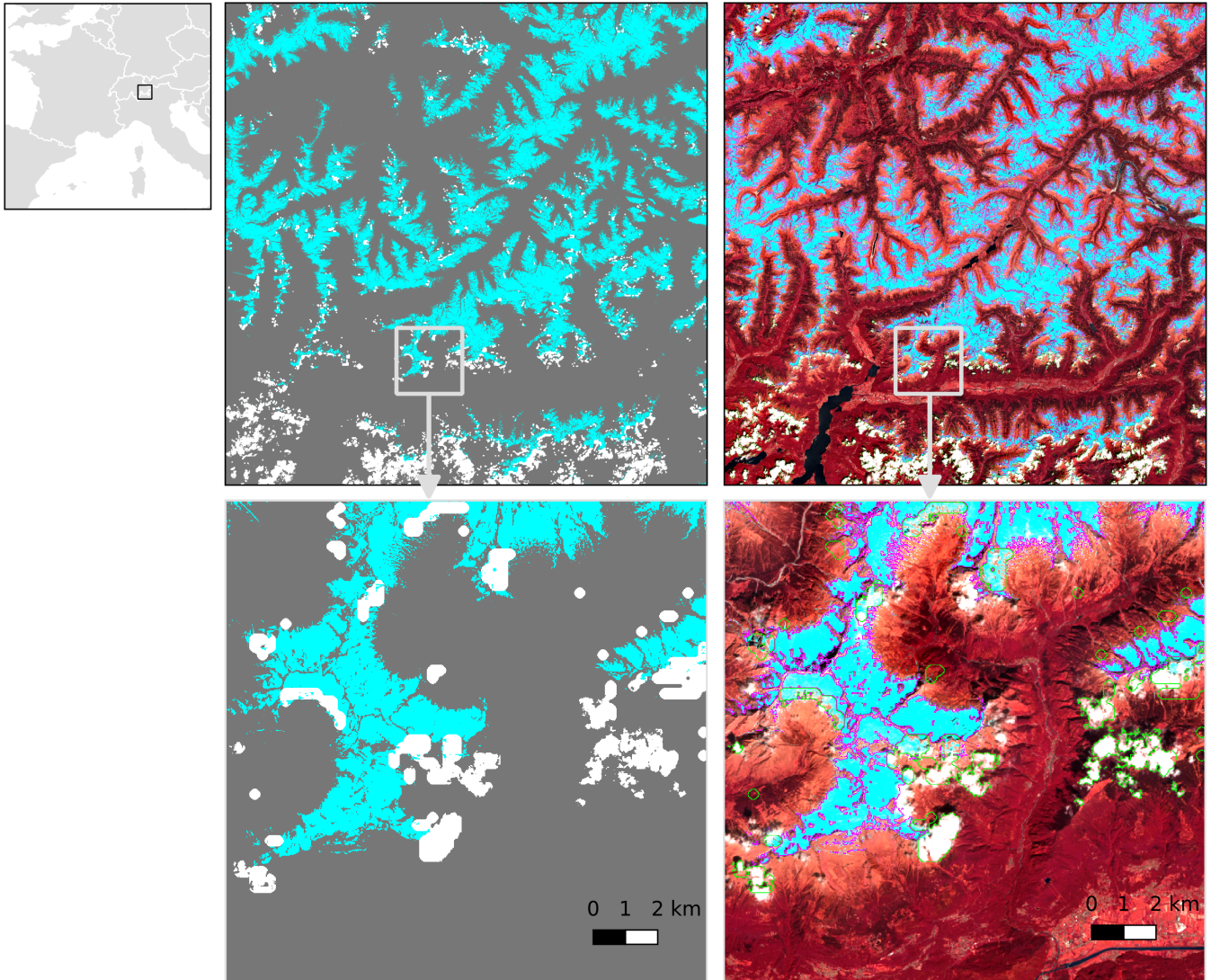


Figure 7. Example of a snow map (SNW file) and its corresponding color composite (CMP file). In the snow map the snow is represented in cyan, no-snow in grey and cloud in white. In the RGB composite, the snow cover areas are delineated in magenta while the clouds are delineated in green (Theia snow product of tile 32TNS on 27 May 2017).

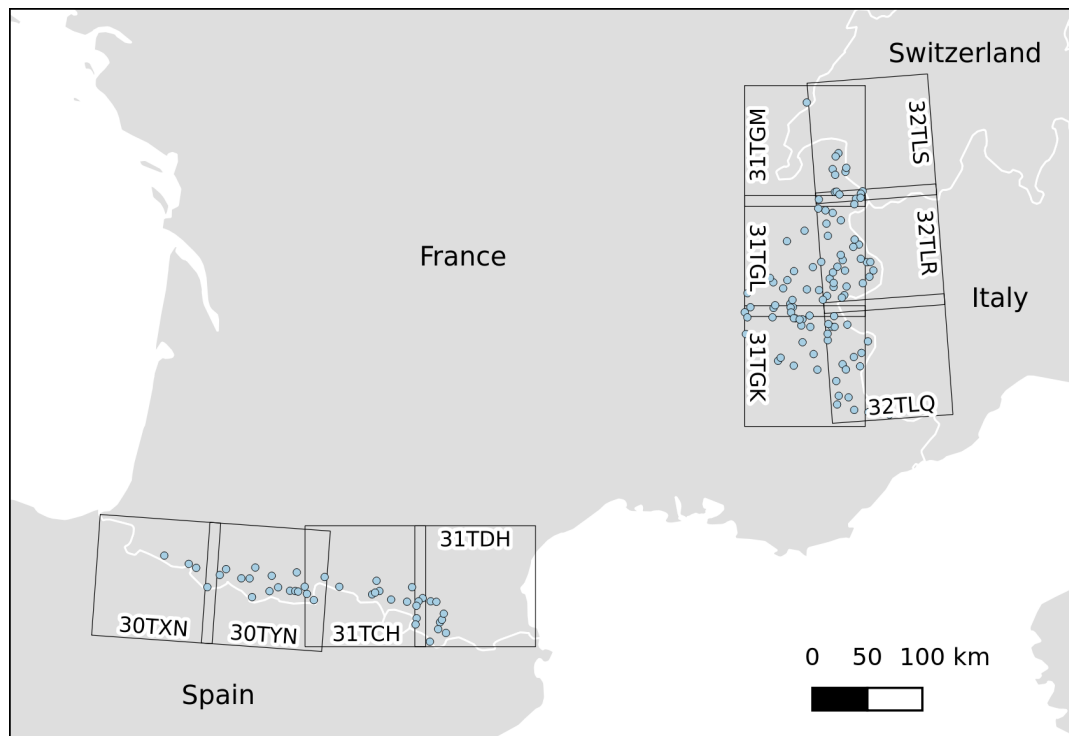


Figure 8. Map showing the location of the stations in the *réseau nivo-météorologique* used for this study (source: Météo-France) and the corresponding Sentinel-2 tiles.

may not be optimal (Klein and Barnett, 2003; Gascoin et al., 2015), therefore the sensitivity of the results to SD_0 was tested by recomputing the confusion matrix for 1 cm increments of SD_0 from 0 to 1 m.

4.1.2 Comparison with snow maps of higher spatial resolution

We used SPOT-6 and SPOT-7 images with a resolution of 1.5 m in panchromatic and 6 m in multispectral (blue, green, red, near-infrared) to evaluate the accuracy of the snow detection in Theia snow products. For this comparison, the LIS processor was run with its current operational configuration (i.e. the configuration used to routinely produce the Theia snow products since December 2017 Tab. 1). SPOT-6 and SPOT-7 sensors acquire images with a large radiometric depth coded in 16 bits, which enables to identify surfaces features in alpine regions with dark shaded slopes and bright snow surfaces. We searched in the catalogue of the Kalideos database for available SPOT images that could match a cloud-free (or nearly cloud-free) Theia snow map over the French Alps region (<https://alpes.kalideos.fr>). We identified six pairs of images in 2016 and 2017 with a maximum time lag of 6 days (Tab. 4, Fig. 9). The SPOT images were obtained as orthorectified products from the Kalideos database. The SPOT images were used to generate reference snow maps by a pixel-based supervised classification. For each SPOT image, we manually drew about fifteen polygons of homogeneous of snow and no-snow surfaces using the panchromatic

Table 4. Pairs of Sentinel-2 and SPOT-6/7 products used for the evaluation of the snow detection accuracy (see also **Fig. 9**).

Sentinel-2 product		Reference product	
Tile	Date	Sensor	Date
31TGK	2016-08-13	SPOT-6	2016-08-08
32TLS	2016-10-12	SPOT-7	2016-10-12
31TGK	2016-12-11	SPOT-7	2016-12-17
31TGL	2016-12-01	SPOT-7	2016-12-03
31TGK	2017-03-11	SPOT-6	2017-03-11
31TGL	2017-03-11	SPOT-6	2017-03-11

image. The **few** cloud pixels in the SPOT images were also manually delineated with a large buffer to restrict the classification to strictly cloud-free areas. Color composites made from the multispectral images were also used as a visual support to help the snow and cloud classification. These polygons were then used to extract training samples from the SPOT products. The samples were taken from both the panchromatic and the multispectral image. We tested a number of classification algorithms by splitting the polygons in calibration and validation datasets. The detail of this analysis is not shown here, but it allowed us to conclude that the random forest classifier was the best choice for our purpose, given its good accuracy and its numerical efficiency (Bouchet, 2018).

In addition, we generated the snow maps from the same Sentinel-2 level 1 products using ESA’s **Sen2Cor** version 2.5 (Louis et al., 2016). The **Sen2Cor** output snow masks are expected to be nearly identical to the ones included in the distributed L2A products by ESA since the same processor is used. We have generated ourselves the L2A products with **Sen2Cor** for this study, because the L2A products were not yet available when me made this analysis (ESA started operational production of Sentinel-2 data at level 2A in 2017). The Theia and **Sen2Cor** snow maps were compared to the reference snow maps from SPOT images using standard metrics derived from the confusion matrix (accuracy, f1-score, kappa, false detection rate and false negative rate).

4.1.3 Evaluation metrics

From the confusion matrices of Sect. 4.1.1 and 4.1.2, we derived the following metrics: accuracy (the proportion of the total number of predictions that were correct), false positive rate (FPR, i.e. the proportion of no-snow pixels that were incorrectly classified as snow), false negative rate (FNR, i.e. the proportion of snow pixels that were incorrectly classified as no-snow), F1-score (harmonic average of the precision and recall) and kappa coefficient (Cohen, 1960).

4.1.4 Visual verification

The evaluation methods presented above under-sample the actual resolution of the Theia snow collection both in space (only 120 points of comparisons in Sect. 4.1.1) and time (only 6 dates of comparisons in Sect. 4.1.2). Given that we do not have a

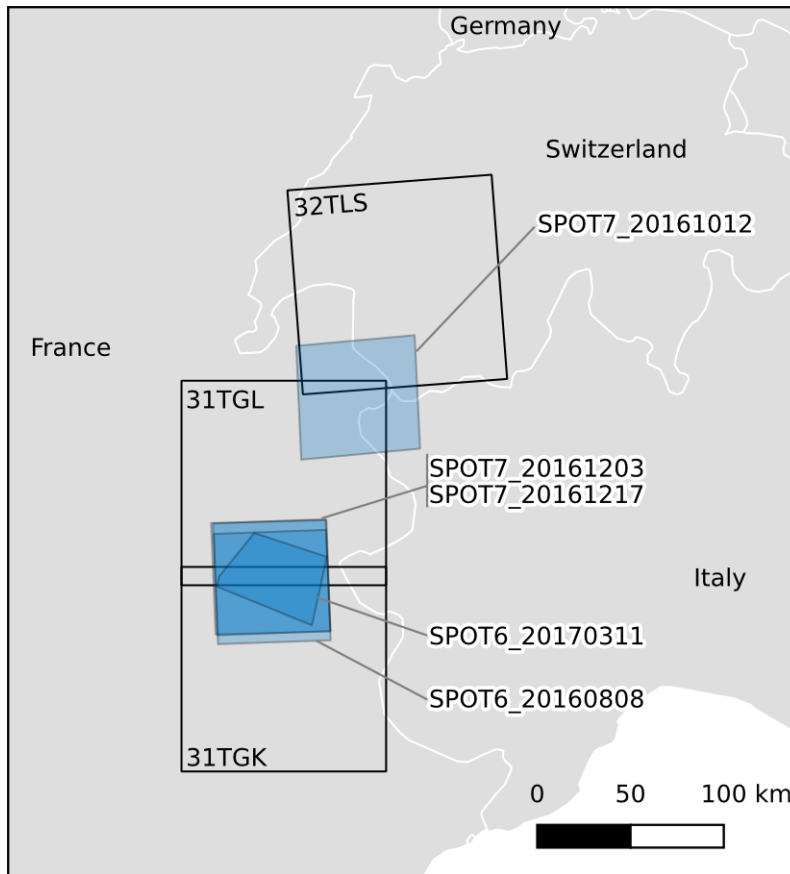


Figure 9. Location of the five SPOT-6/7 products used for the evaluation of the snow detection accuracy and the corresponding Sentinel-2 tiles (see also Tab. 4).

more extensive validation dataset, we used a time series of 64 Theia snow products from 06 July 2015 to 02 July 2017 over tile 31TCH to control the consistency of the snow and cloud masks based on the visual inspection of the false color composites. This approach is efficient to evaluate the frequency of gross errors at the tile scale, i.e. large patches of false snow or false no-snow detection (Vidot et al., 2017).

5 4.2 Results

4.2.1 Comparison with in situ snow depth measurements

The confusion matrix between the Theia snow products and in situ snow depth measurements is given in Table 5. We find 1414 pairs of data for the study period (i.e. there are 1414 individual snow depth measurements for which a cloud-free retrieval can be found in the Theia Snow collection on the same day). From this confusion matrix, the accuracy (proportion of correct classifications) is 94% and the kappa coefficient is 0.83, which indicates an excellent agreement according to Fleiss et al. (2013).

Table 5. Confusion matrix between the Theia snow products and in situ snow depth data (SD).

In situ snow depth	Theia snow products	
	no snow	snow
SD=0	276	8
SD>0	76	1054

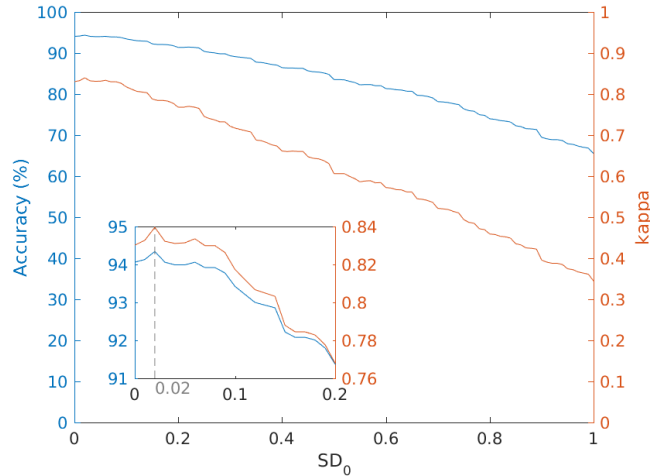


Figure 10. Sensitivity of the agreement between in situ and Theia snow product to SD_0 in m (threshold to convert the measured snow depth to snow presence or absence). The inset shows a zoom in the region 0-0.2 m.

The false positive rate (2.8%) is lower than the false negative rate (6.7%), which means that the Theia snow product is more likely to underestimate than to overestimate the snow detection at the station locations. The false negative rate decreases if SD_0 is set to higher value, however the false positive rate also increases in such a way that an optimum is reached at $SD_0 = 2$ cm (Fig. 10).

- 5 The comparisons of each individual time series before matching the data by date is shown in Appendix B. These plots illustrate that the high revisit time of Sentinel-2 enables to well capture the seasonal cycle of the snow cover. Even intermediate melt out events at lower elevation stations can be identified over the course of the snow season.

4.2.2 Comparison with snow maps of higher spatial resolution

- 10 Table 6 shows that the MAJA-LIS workflow provides a better detection of the snow cover than the **Sen2Cor** processor in all the studied cases. Although the improvement in the accuracy coefficient is low in some cases (31TGL/2017-03-11, 31TGK/2016-08-13) or even slightly negative in one case (32TLS/2016-10-12), the F1 score and kappa coefficient of the Theia snow products are always greater than or equal to those of **Sen2Cor**. The differences are significant for four products among the six tested. The

Table 6. Results of the validation of Theia (LIS 1.2.1) and ESA (**Sen2Cor** 2.5) snow products using SPOT-6/7 reference snow maps.

Sentinel-2 product		Accuracy		F1		Kappa		FPR		FNR	
Tile	Date	Sen2Cor	LIS	Sen2Cor	LIS	Sen2Cor	LIS	Sen2Cor	LIS	Sen2Cor	LIS
31TGK	2016-12-11	0.83	0.92	0.73	0.91	0.61	0.84	0.41	0.04	0.05	0.14
31TGL	2017-03-11	0.89	0.90	0.92	0.93	0.74	0.76	0.09	0.02	0.06	0.11
31TGL	2016-12-01	0.84	0.95	0.78	0.95	0.66	0.90	0.35	0.05	0.01	0.06
31TGK	2017-03-11	0.85	0.89	0.89	0.92	0.65	0.72	0.15	0.01	0.07	0.13
31TGK	2016-08-13	0.98	0.99	0.35	0.77	0.34	0.77	0.78	0.24	0.22	0.22
32TLS	2016-10-12	0.98	0.97	0.63	0.64	0.62	0.62	0.49	0.08	0.17	0.51

lower accuracy in the case of the the 32TLS/2016-10-12 product is due to a higher false negative rate. The false positive rates in the Theia snow products are generally lower than the false negative rates, which is consistent with the previous evaluation using in situ data. The spatial analysis of the errors indicates that the false positives (i.e. over-detection of snow pixels) in the Theia products are mostly located near the snow cover edges (Fig. A1). This suggests that these errors might not be due to a false detection of snow (e.g. confusion with a lake or a cloud surface) but are rather due to the resolution discrepancy with the reference data. By contrast, the spatial comparison of the **Sen2Cor** products with the reference data shows large patches of false negative (omission of snow pixels) in the snow cover area, which are typically due the omission of snow pixels in shaded areas (Fig. A2).

4.2.3 Visual verification

We found 4 products among 64 with significant patches of false positive pixels (i.e. pixels falsely classified as snow). These false snow areas are exclusively located in cloud areas (Fig. 11). False negative pixels (i.e. pixels falsely classified as **snow-free**) can be found by looking at higher resolution along the snow cover edges, but we did not find large areas of false negative pixels.

5 Discussion

The evaluation of the Snow collection using both in situ and remote sensing reference datasets indicates **an** excellent accuracy of the snow detection, albeit with a tendency to snow under-detection. The omission of snow pixels can be due to the presence of trees or shadows. In the case of the in situ comparison, part of the errors may be explained by the uncertainty in the geolocalization of the Sentinel-2 data, which is close to 11 m at 95% for both satellites according to the latest Sentinel-2 L1C data quality report (MPC Team, 2018). In addition, in situ measurements may not **be** representative of the snow conditions within the Sentinel-2 pixel. However, the sensitivity analysis of the accuracy and kappa coefficient to the SD_0 value suggests that the discrepancy between the scale of the in situ observations and the scale of the pixel observation is not significant since

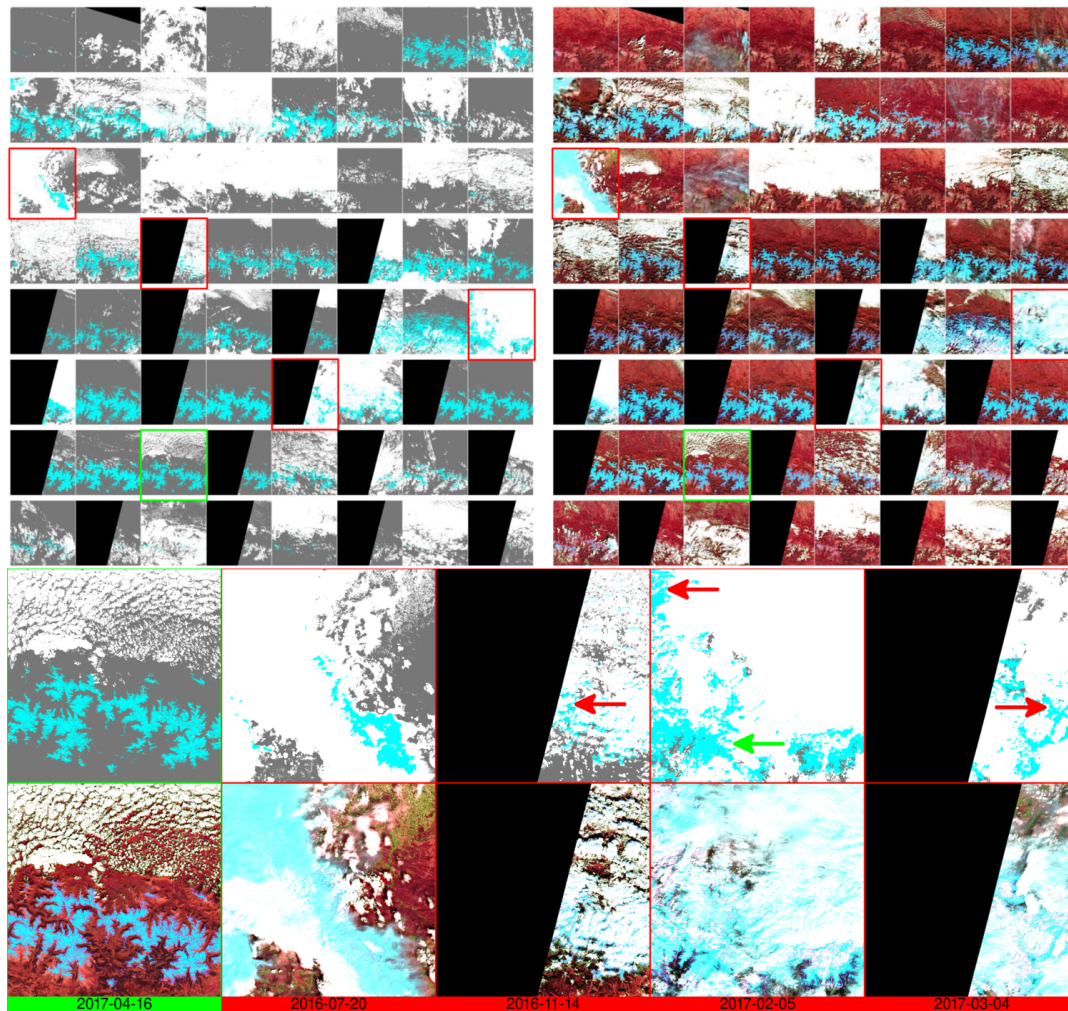


Figure 11. Full time series of Theia snow products from 06 July 2015 to 02 July 2017 over tile 31TCH. Four products with visually-evident false snow detections are framed in red and shown at higher resolution. Red arrows indicate the location of the false snow patches, while the green arrow indicates a true snow patch. The product of 16 April 2017 is shown as a reference to help the visualization of the errors.

the optimal SD_0 is close to 0 m. In contrast, a similar analysis with lower resolution snow products from MODIS found an optimal SD_0 of 0.15 m **Gascoin et al. (2015)**. This illustrates that high resolution snow products from Sentinel-2 enable to partly resolve the typical scale issue between in situ and remote sensing products Blöschl (1999).

5 The visual inspection of the products reveals that the main issue in some Theia snow products is rather the occurrence of false positive due to the confusion with cloud surfaces. This problem can be due to two main reasons:

- Cold clouds: cold clouds have a spectral signature that is close to the snow cover since they contain ice crystals (high NDSI). If these clouds are not accurately detected by the level 2A processor then the LIS algorithm will also classify them as snow after pass 1 (Sect 2). As a result the snowline elevation z_s is not well estimated, which can generate erroneous snow patches within these clouds even at low elevation (e.g. product of 20 July 2017 in Fig. 11).
- 10 – Shaded clouds: the three-dimensional structure of the clouds can form shadows within the cloud cover area (e.g. product of 04 March 2017 in Fig. 11). These shaded cloud areas have a lower reflectance in the visible wavelengths, and therefore can be considered as “dark clouds” in the LIS algorithm, provided that the shaded cloud cover area is large enough to significantly reduce the reflectance in the red band even after the down-sampling step (Sect 2).

Among 64 products, we find 4 cases with significant false snow cover detection in this tile (Fig 11). Although it is not shown
15 here, we have made this exercise over other tiles in the Alps and Atlas mountains and we estimate that this proportion is similar in these areas. Regarding this issue we mainly rely on our visual judgment using color composite to assess the ability of the algorithm to discriminate the cloud and the snow cover. However, to further quantify the algorithm performance and eventually optimize its parameters, we would need a database of classified imagery with labelled regions of cloud and snow cover in many different cases.

20 **Both Sen2Cor and LIS use the NDSI to detect the snow cover, however the algorithms differ in many aspects. The better performances of LIS in Sect. 4.2.2 can be related to the mono-date approach for atmospheric correction and cloud detection in Sen2Cor, which can cause snow/cloud confusion in the L2A product, while LIS uses L2A products from the more accurate multi-temporal MAJA processor Baetens et al. (2019). In addition, a unique feature of the LIS algorithm is the use of the snow line elevation concept to improve the robustness of the snow detection in mountain regions. In**
25 **addition, LIS was designed to retrieve snow below thin clouds and to reclassify snow pixels which are frequently found near the snow cover edges in L2A products. Apart from the MAJA processing, however, the effect of these algorithmic differences could be less evident in flat areas.**

6 Conclusions

The Theia Snow collection is a free collection of snow products which indicate the presence or absence of snow on the land
30 based on Sentinel-2 (20 m) and Landsat-8 (30 m) data. Most of the snow products are derived from Sentinel-2. **However, in late September 2018, Landsat-8 snow products started being routinely added to the Theia Snow collection. Previous Landsat-8 data were also reprocessed back to March 2017. At the time of writing (15 March 2019), for example in tile**

T31TCH in the Pyrenees, 20.5% of all available products were derived from Landsat-8 data (63 among 244). In addition, the processing is done to facilitate the combination of Landsat-8 and Sentinel-2 snow products since the data format and the tiling scheme are the same (only the pixel size differs). Although the co-registration of Sentinel-2 and Landsat-8 data can be still problematic Storey et al. (2016), the combination of "analysis ready" Landsat-8 and Sentinel-2 is important to increase the number of cloud-free observations.

The evaluation presented in this article indicate that the Theia Snow product allows an accurate detection of the snow cover (and of the absence of snow). However, it remains a preliminary assessment that should be extended in the future. For example, the evaluation was based only on Sentinel-2 products, **since the integration of Landsat-8 data started during the writing of this manuscript and should be extended to Landsat-8 products. However, the processing is done with the same algorithm, which is based on the calculation of the NDSI from slope-corrected surface reflectance (level 2A data).** In addition, the evaluation was limited to the French Alps and French Pyrenees, while the collection covers other mountain regions. Last, the spatial evaluation using higher resolution remote sensing data was focused on the accuracy of the snow detection using very high resolution clear-sky images, while, as discussed above, the main difficulty is not to detect the snow when the sky is cloud-free, but rather to avoid false snow detection within clouds. The reduction of the false positive rate is the main challenge for the future developments in the LIS algorithm. In the meantime, we also welcome feedback from other users.

The Theia Snow collection is based on optical observations therefore it is not adapted to the detection of the snow cover in dense forest areas where the ground is obstructed by the canopy (Xin et al., 2012). This may typically occur in evergreen conifer forests of alpine regions (e.g. Alps, Pyrenees). We have noted that the snow detection can be successful even in alpine forests, but we lack quantitative data to quantify its accuracy. Therefore it is recommended to use a land cover map to exclude these forest areas from the analysis. For the tiles in France, this can be done using the Theia land cover map (Inglada et al., 2017), which is distributed at the same spatial resolution. In the Pyrenees, the mean altitude of zero-degree isotherm in winter is close to the treeline elevation at 1600 m, therefore the impact of the forest cover is limited (Gascoin et al., 2015). The LIS algorithm may also fail to detect the snow cover in steep shaded slopes if the solar elevation is very low (typically below 20°). This can occur in mid-latitude areas in winter. In this case the slope correction in the L2A product is generally not applied. This is indicated in the L2A mask but not in the snow product. A future release of the Snow collection should include this information. However, the visual inspection of many images in regions of complex topography gives us the impression that the snow detection algorithm performs already well in shaded slopes. We are planning to further investigate this issue using alternative approaches (Sirguey et al., 2009).

In spite of all these limitations, the Theia snow products have already been successfully used for the evaluation of the MODIS snow products (Masson et al., 2018) and the assimilation of snow cover area data into a snowpack model in the High Atlas (Baba et al., 2018). Given that the snow cover is a key driver of many natural processes in mountains regions, we envision various potential applications of the Theia snow products, including the modelling of the distribution of the permafrost in mountain regions, the validation and calibration of hydrological models in snow-dominated catchments, and the spatial modelling of biodiversity and productivity of ecosystems in mountain regions.

7 Code and data availability

The Theia snow collection can be accessed and cited using this Digital Object Identifier: DOI:10.24400/329360/F7Q52MNK.

The let-it-snow source code is available under the GNU Affero General Public License v3.0 in this repository: <http://tully.ups-tlse.fr/grizonnet/let-it-snow>.

Appendix A: Spatial comparison between Theia and Sen2Cor snow products with SPOT-6/7 reference data

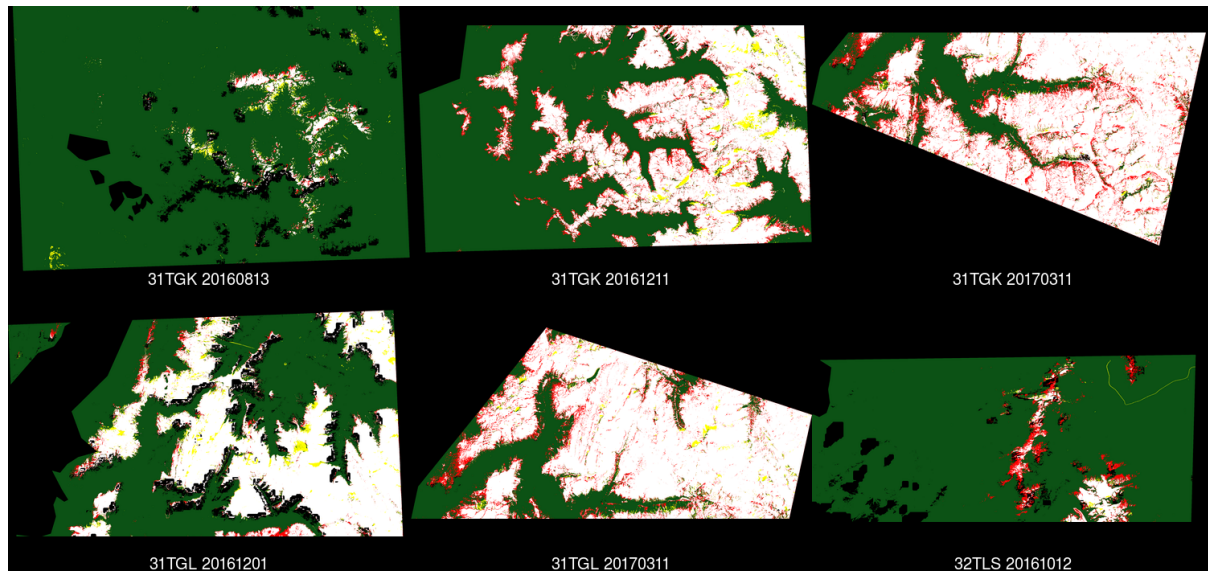


Figure A1. Theia snow products vs. SPOT-6/7 snow maps (green: true negative, white: true positive, red: false positive, yellow: false negative).

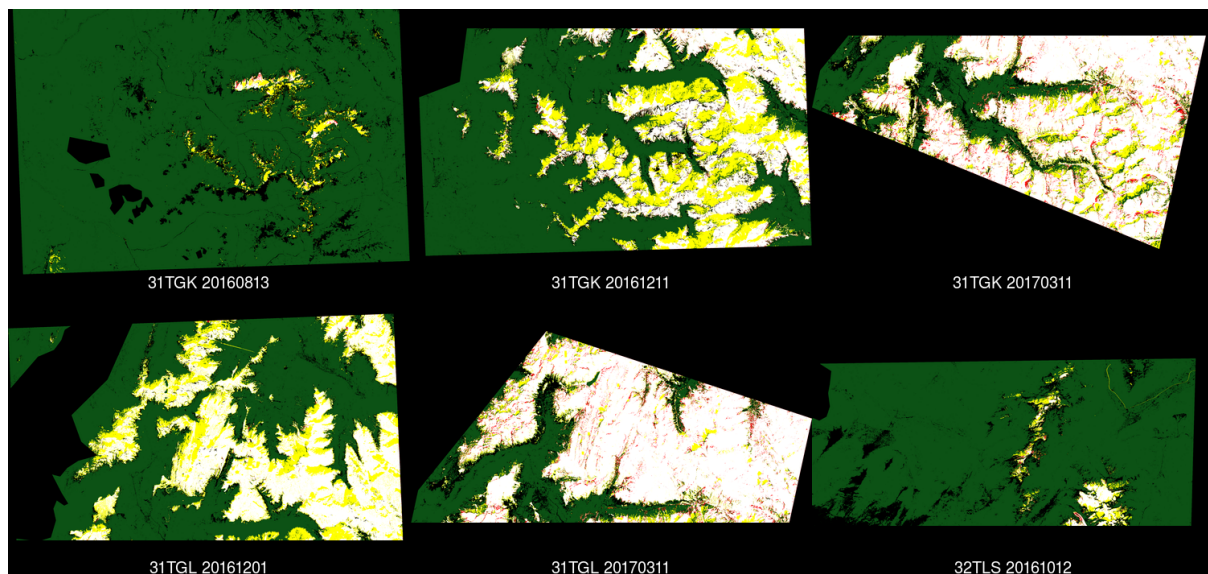


Figure A2. Sen2Cor snow product vs. SPOT-6/7 snow maps (green: true negative, white: true positive, red: false positive, yellow: false negative).

Appendix B: Time series of Theia snow products and snow depth records at 120 snow observation stations from 01-Sep-2017 to 31-Aug-2018.

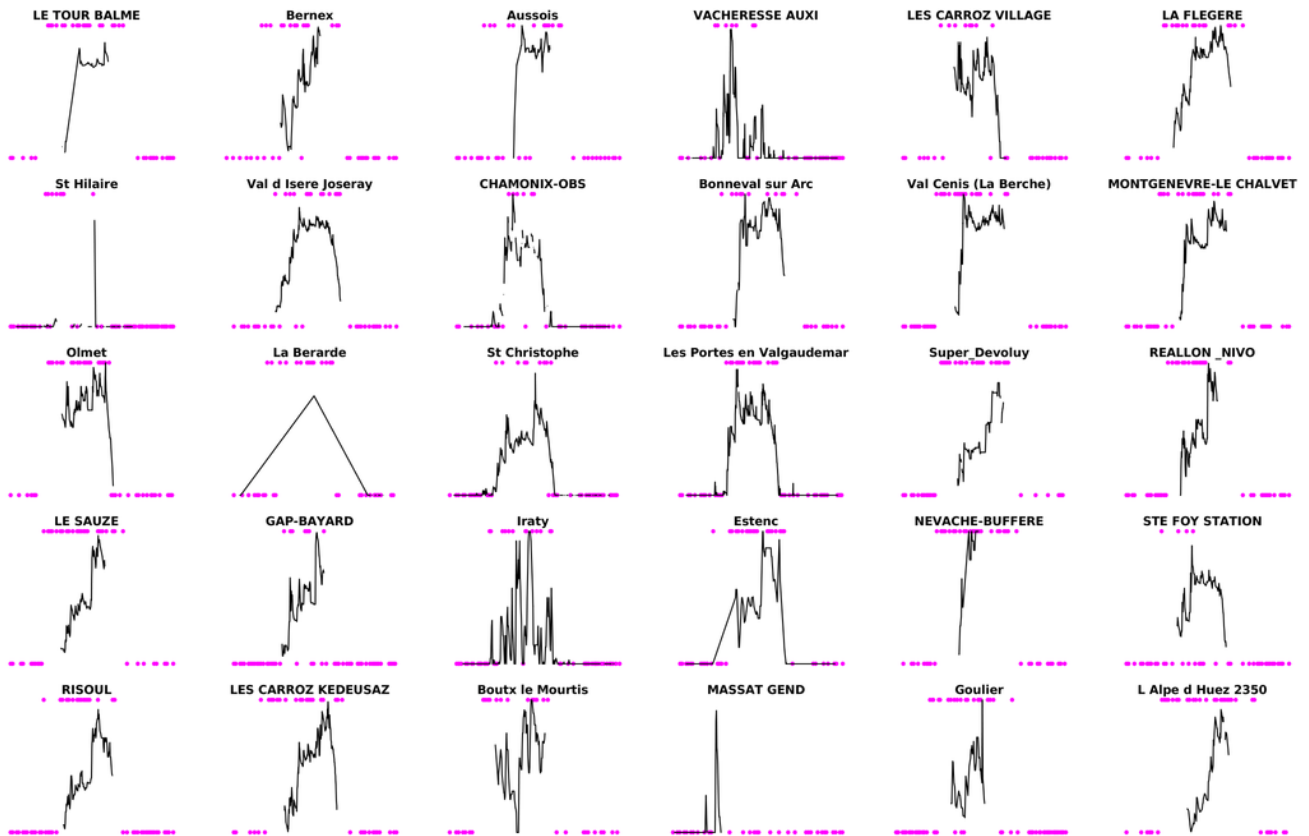


Figure B1. Magenta dots: snow presence (top) or absence (bottom) in the Theia snow collection. Black line: snow depth time series.

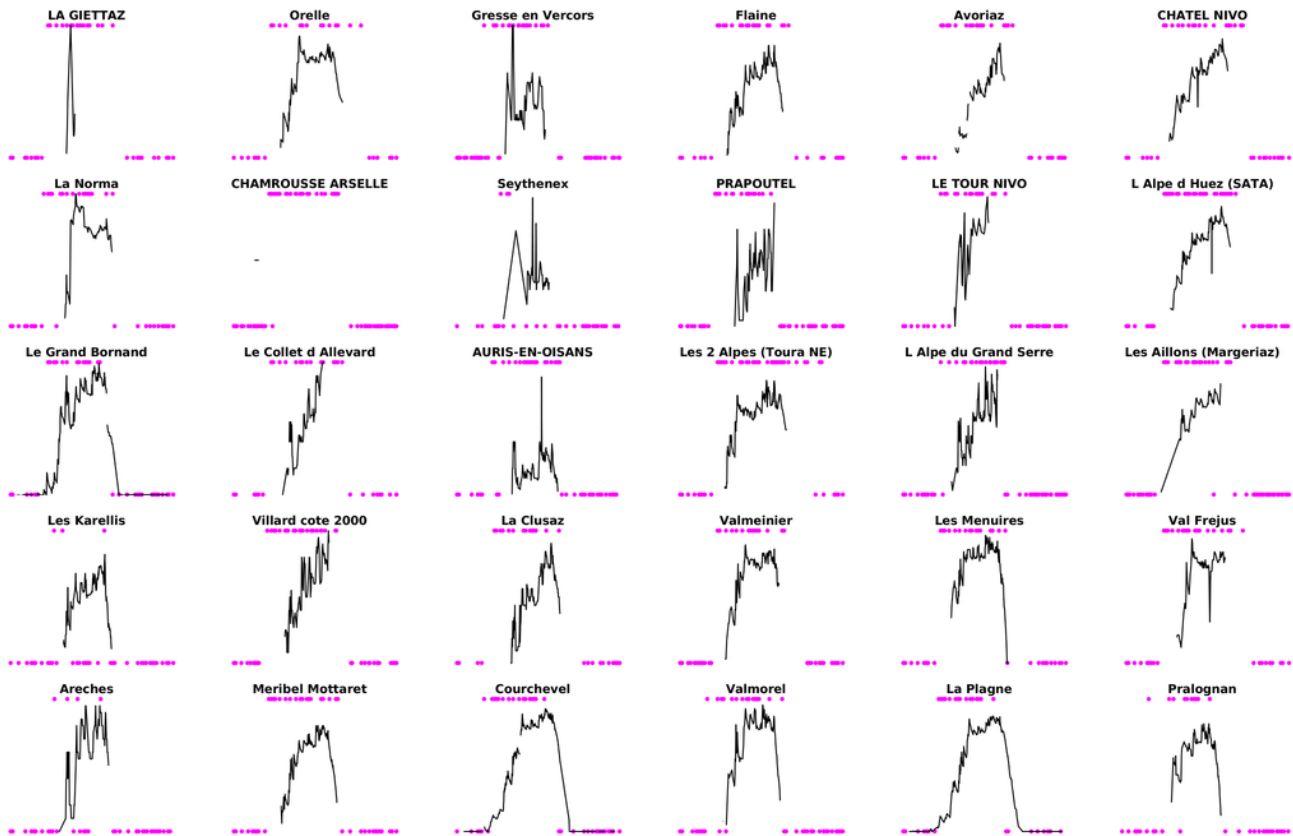


Figure B2. Magenta dots: snow presence (top) or absence (bottom) in the Theia snow collection. Black line: snow depth time series.

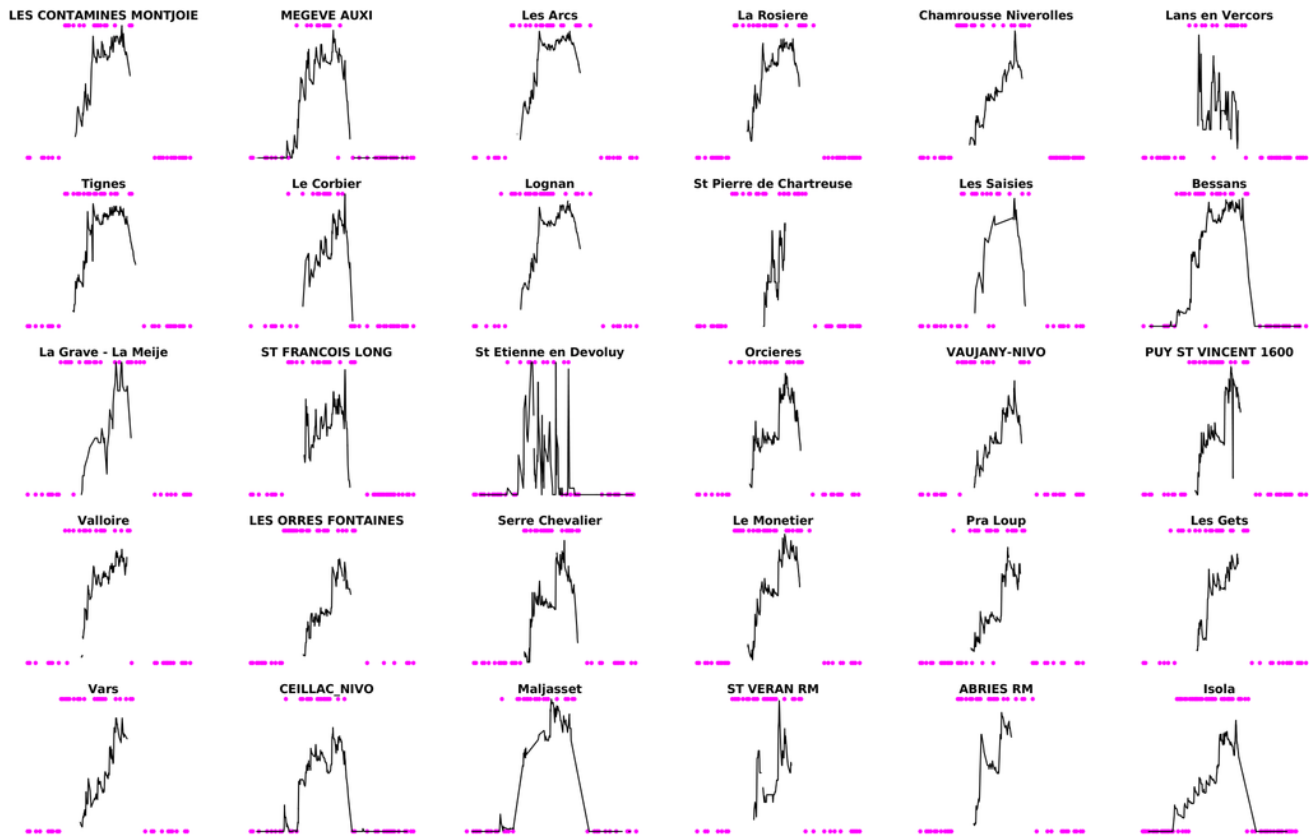


Figure B3. Magenta dots: snow presence (top) or absence (bottom) in the Theia snow collection. Black line: snow depth time series.

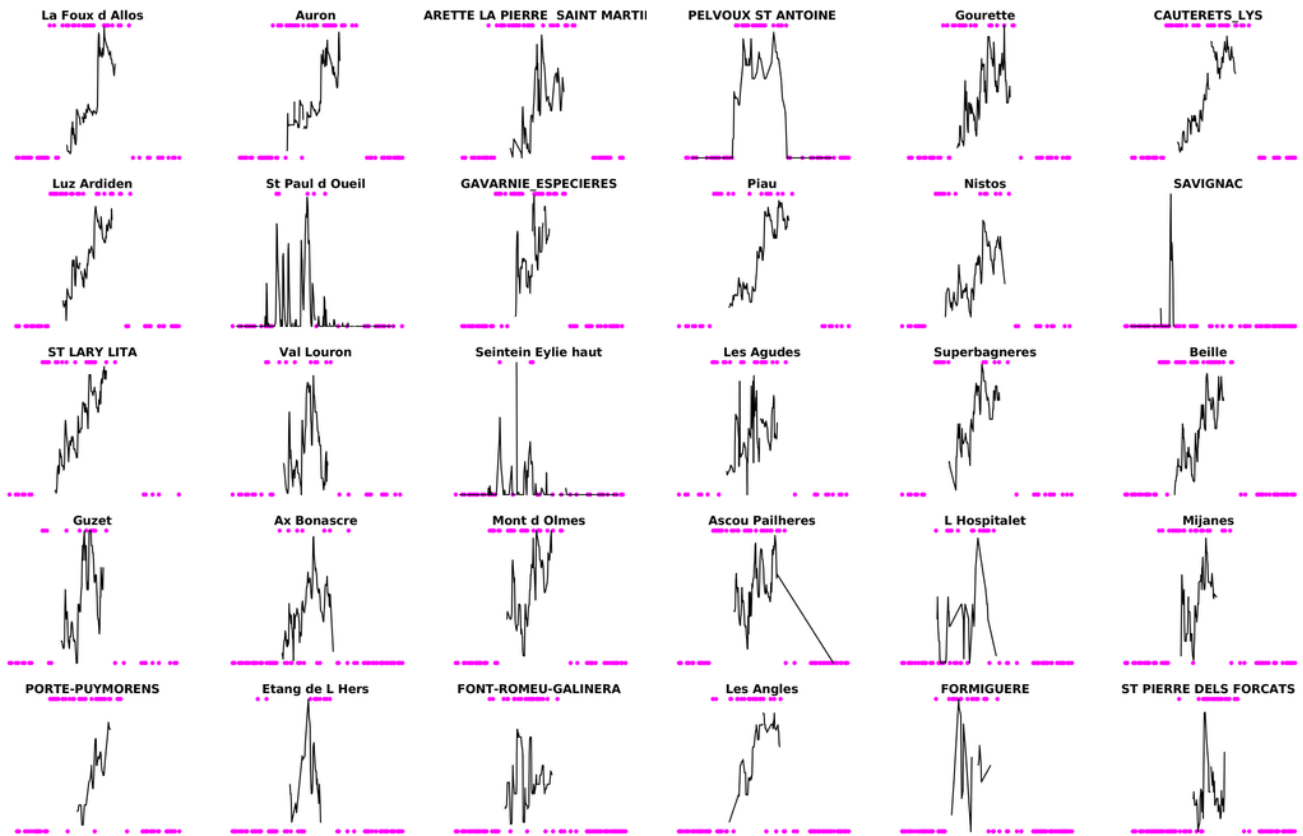


Figure B4. Magenta dots: snow presence (top) or absence (bottom) in the Theia snow collection. Black line: snow depth time series.

Author contributions. SG and OH defined the LIS algorithm. MG and GS worked on the implementation and developments of the LIS processor. MG, MB and SG worked on the validation of the snow products. SG wrote the manuscript with inputs from all the co-authors.

Competing interests. The authors declare that no competing interests are present.

Acknowledgements. The development and production of the Theia Snow collection is supported by the CNES. We thank Tristan Klempka
5 for his contribution to earlier developments of the LIS processor. We thank Marc Leroy, Arnaud Sellé and Philippe Pacholczyk for the coordination of the Theia project, Joelle Donadieu and Céline l’Helguen for the transfer to the production at MUSCATE. **This manuscript was greatly improved thanks to the constructive comments of Gaia Piazzì, Javier Herrero, Jeff Dozier and an anonymous referee.**

References

- Baba, M. W., Gascoïn, S., and Hanich, L.: Assimilation of Sentinel-2 Data into a Snowpack Model in the High Atlas of Morocco, *Remote Sensing*, 10, <https://doi.org/10.3390/rs10121982>, 2018.
- Baetens, L., Desjardins, C., and Hagolle, O.: Validation of Copernicus Sentinel-2 Cloud Masks Obtained from MAJA, Sen2Cor, and FMask Processors Using Reference Cloud Masks Generated with a Supervised Active Learning Procedure, *Remote Sensing*, 11, 433, <https://doi.org/10.3390/rs11040433>, 2019.
- Blöschl, G.: Scaling issues in snow hydrology, *Hydrological processes*, 13, 2149–2175, 1999.
- Bouchet, M.: Validation et amélioration des produits Surfaces Enneigées Sentinel-2, <https://doi.org/10.5281/zenodo.1446460>, 2018.
- Cohen, J.: A coefficient of agreement for nominal scales, *Educational and psychological measurement*, 20, 37–46, 1960.
- 10 Commissariat général au développement durable: Plan d'applications satellitaires 2018, <https://www.ecologique-solidaire.gouv.fr/sites/default/files/Plan%20d%E2%80%99applications%20satellitaires%202018.pdf>, 2018.
- Dozier, J.: Spectral signature of alpine snow cover from the Landsat Thematic Mapper, *Remote sensing of environment*, 28, 9–22, 1989.
- Drusch, M., Bello, U. D., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., and Bargellini, P.: Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services, *Remote Sensing of Environment*, 120, 25 – 36, <https://doi.org/10.1016/j.rse.2011.11.026>, the Sentinel Missions - New Opportunities for Science, 2012.
- 15 Fleiss, J. L., Levin, B., and Paik, M. C.: *Statistical methods for rates and proportions*, John Wiley & Sons, 2013.
- Frei, A., Tedesco, M., Lee, S., Foster, J., Hall, D., Kelly, R., and Robinson, D. A.: A review of global satellite-derived snow products, *Advances in Space Research*, 50, 1007–1029, <https://doi.org/10.1016/j.asr.2011.12.021>, 2012.
- 20 Gao, B.-C., Goetz, A. F., and Wiscombe, W. J.: Cirrus cloud detection from airborne imaging spectrometer data using the 1.38 μm water vapor band, *Geophysical Research Letters*, 20, 301–304, 1993.
- Gascoïn, S., Hagolle, O., Huc, M., Jarlan, L., Dejoux, J.-F., Szczypta, C., Marti, R., and Sánchez, R.: A snow cover climatology for the Pyrenees from MODIS snow products, *Hydrology and Earth System Sciences*, 19, 2337–2351, <https://doi.org/10.5194/hess-19-2337-2015>, 2015.
- 25 Gascoïn, S., Grizonnet, M., Klempka, T., and Salgues, G.: Algorithm theoretical basis documentation for an operational snow cover product from Sentinel-2 and Landsat-8 data (Let-it-snow), <https://doi.org/10.5281/zenodo.1414452>, 2018.
- GDAL/OGR contributors: GDAL/OGR Geospatial Data Abstraction software Library, Open Source Geospatial Foundation, <http://gdal.org>, 2018.
- Grizonnet, M., Michel, J., Poughon, V., Inglada, J., Savinaud, M., and Cresson, R.: Orfeo ToolBox: open source processing of remote sensing images, *Open Geospatial Data, Software and Standards*, 2, 15, <https://doi.org/10.1186/s40965-017-0031-6>, 2017.
- 30 Hagolle, O., Huc, M., Pascual, D. V., and Dedieu, G.: A multi-temporal method for cloud detection, applied to FORMOSAT-2, VEN μ S, LANDSAT and SENTINEL-2 images, *Remote Sensing of Environment*, 114, 1747–1755, 2010.
- Hagolle, O., Huc, M., Desjardins, C., Auer, S., and Richter, R.: MAJA Algorithm Theoretical Basis Document, <https://doi.org/10.5281/zenodo.1209633>, 2017.
- 35 Hall, D., Riggs, G., Salomonson, V., DiGirolamo, N., and Bayr, K.: MODIS snow-cover products, *Remote sensing of Environment*, 83, 181–194, 2002.

- Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., and Rodes, I.: Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series, *Remote Sensing*, 9, <https://doi.org/10.3390/rs9010095>, 2017.
- Jarvis, A., Reuter, H. I., Nelson, A., and Guevara, E.: Hole-filled SRTM for the globe Version 4, 2008.
- Klein, A. and Barnett, A.: Validation of daily MODIS snow cover maps of the Upper Rio Grande River Basin for the 2000-2001 snow year, *Remote Sensing of Environment*, 86, 162–176, [https://doi.org/10.1016/S0034-4257\(03\)00097-X](https://doi.org/10.1016/S0034-4257(03)00097-X), 2003.
- 5 Krajčí, P., Holko, L., and Parajka, J.: Variability of snow line elevation, snow cover area and depletion in the main Slovak basins in winters 2001–2014, *Journal of Hydrology and Hydromechanics*, 64, 12–22, 2016.
- Li, J. and Roy, D. P.: A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring, *Remote Sensing*, 9, 902, <https://doi.org/10.3390/rs9090902>, 2017.
- 10 Louis, J., Debaecker, V., Pflug, B., Main-Korn, M., Bieniarz, J., Mueller-Wilm, U., Cadau, E., and Gascon, F.: Sentinel-2 Sen2Cor: L2A Processor for Users, in: *Living Planet Symposium*, vol. 740, p. 91, 2016.
- Malnes, E., Buanes, A., Nagler, T., Bippus, G., Gustafsson, D., Schiller, C., Metsämäki, S., Pulliainen, J., Luojus, K., Larsen, H. E., Solberg, R., Diamandi, A., and Wiesmann, A.: User requirements for the snow and land ice services – CryoLand, *The Cryosphere*, 9, 1191–1202, <https://doi.org/10.5194/tc-9-1191-2015>, 2015.
- 15 Masson, T., Dumont, M., Mura, M., Sirguey, P., Gascoin, S., Dedieu, J.-P., and Chanussot, J.: An Assessment of Existing Methodologies to Retrieve Snow Cover Fraction from MODIS Data, *Remote Sensing*, 10, 619, <https://doi.org/10.3390/rs10040619>, 2018.
- Metsämäki, S., Pulliainen, J., Salminen, M., Luojus, K., Wiesmann, A., Solberg, R., Böttcher, K., Hiltunen, M., and Ripper, E.: Introduction to GlobSnow Snow Extent products with considerations for accuracy assessment, *Remote Sensing of Environment*, 156, 96–108, 2015.
- MPC Team: Sentinel-2 L1C Data Quality Report, Tech. Rep. 31, European Space Agency, 2018.
- 20 Painter, T. H., Rittger, K., McKenzie, C., Slaughter, P., Davis, R. E., and Dozier, J.: Retrieval of subpixel snow covered area, grain size, and albedo from MODIS, *Remote Sensing of Environment*, 113, 868–879, 2009.
- Parajka, J. and Blöschl, G.: Spatio-temporal combination of MODIS images–potential for snow cover mapping, *Water Resour. Res.*, 44, W03406, <https://doi.org/10.1029/2007WR006204>, 2008.
- Ramsay, B. H.: The interactive multisensor snow and ice mapping system, *Hydrological Processes*, 12, 1537–1546, 1998.
- 25 Sirguey, P., Mathieu, R., Arnaud, Y., Khan, M., and Chanussot, J.: Improving MODIS Spatial Resolution for Snow Mapping Using Wavelet Fusion and ARSIS Concept, *Geoscience and Remote Sensing Letters, IEEE*, 5, 78–82, <https://doi.org/10.1109/LGRS.2007.908884>, 2008.
- Sirguey, P., Mathieu, R., and Arnaud, Y.: Subpixel monitoring of the seasonal snow cover with MODIS at 250 m spatial resolution in the Southern Alps of New Zealand: Methodology and accuracy assessment, *Remote Sensing of Environment*, 113, 160–181, 2009.
- Storey, J., Roy, D. P., Masek, J., Gascon, F., Dwyer, J., and Choate, M.: A note on the temporary misregistration of Landsat-8 Operational Land Imager (OLI) and Sentinel-2 Multi Spectral Instrument (MSI) imagery, *Remote Sensing of Environment*, 186, 121 – 122, <https://doi.org/10.1016/j.rse.2016.08.025>, 2016.
- U.S. Geological Survey Earth Resources Observation And Science Center: Landsat OLI Level-2 Surface Reflectance (SR) Science Product, <https://doi.org/10.5066/f78s4mzj>, 2014.
- Vidot, J., Bellec, B., Dumont, M., and Brunel, P.: A daytime VIIRS RGB pseudo composite for snow detection, *Remote Sensing of Environment*, 196, 134 – 139, <https://doi.org/10.1016/j.rse.2017.04.028>, 2017.
- 35 Xin, Q., Woodcock, C. E., Liu, J., Tan, B., Melloh, R. A., and Davis, R. E.: View angle effects on MODIS snow mapping in forests, *Remote Sensing of Environment*, 118, 50 – 59, <https://doi.org/10.1016/j.rse.2011.10.029>, 2012.

Zhu, Z., Wang, S., and Woodcock, C. E.: Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images, *Remote Sensing of Environment*, 159, 269–277, 2015.