Author responses are below the respective reviewer/editor comments in green text. A marked-up version of the manuscript with tracked changes is attached after the responses to the comments. We have significantly changed the manuscript (rewritten and restructured the discussion, expansion of methods section, changes to figures, …) and accordingly the manuscript with the marked-up changes is rather unpleasant to read. We hope our responses to the comments are comprehensive in giving an overview of the specific changes made and suggest looking at these in combination with the non-marked up version of revised manuscript.

Editor comments:

I also have a couple of additional comments on the submitted MS:
- Are the 246 spectra collected 246 averages of multiple spectra ("stacked")? What is typical in ground-based albedo studies?
- A better description of the measurement protocol (identified by reviewer #2) is essential, along with context, i.e., comparison against accepted best practices for such measurements.

The spectra we present in this study are not stacked. Most other studies that follow a similar measurement protocol as ours and measure reflectance of ice/snow with a portable spectroradiometer do not use stacked spectra as such, though spectra might be averaged after grouping by surface type (e.g. Naegeli et al., 2015, 2017; Malinka et al., 2016; Hendriksa et al, 2003). In contrast, Di Mauro et al. (2017) use stacked spectra, averaging over 15 scans each time. They do not comment on how this affects their results. Our test runs in the field indicated that the measured spectra are very consistent between "shots" provided the position of the instrument is not changed so we chose not to average over multiple spectra for each point. We cannot comment in detail on typical procedures in ground-based albedo studies that deal with very different kinds of surfaces (eg vegetation) but our understanding is that it varies depending on the specific questions that are being investigated. We have expanded the description of the measurement protocol and cite previous studies where a similar approach is used.

Reviewer 1

Reviewer comment:

This study by Hartl et al (2020) compares a detailed field survey of albedo on Jamtalferner with synchronous remote sensing derived albedo from Sentinel and Landsat images. The methods for both approaches to albedo determination are well explained. The comparison of the field albedo and remote sensing derived albedo is the key output of this paper and is well illustrated in Figures 7-9. The study provides a richer data set for understanding how Landsat or Sentinel images could be used and is simply interesting. The primary comments below are seeking more context: 1) On the value of detailed spatial and temporal albedo observations. 2) For connections with energy balances. I am not suggesting additional data or figures be presented, but instead additional reference to other work and how the data here fits with these.

Author response:

Thank you for your comments and suggestions. We have significantly expanded the discussion to include more depth on points 1 and 2, among other considerations.

We address the specific comments below. We have attempted to follow all suggestions within the constraints of the data available to us.

9: The first sentence reverses the cause and effect. "As Alpine glaciers recede, they are quickly becoming snow free in summer and, accordingly, spatial and temporal variations in ice albedo increasingly affect the melt regime. "

Instead I suggest, “As alpine glacier become snow free in summer, recession occurs, and further spatial and temporal variations in ice albedo increasingly accentuate the melt regime.“

Rephrased to: “As Alpine glaciers become snow free in summer, further spatial and temporal variations in ice albedo increasingly accentuate the melt regime and recession occurs.”

16: Finishing the sentence with fluid is confusing since that could be a surface type, “Spectra can roughly be grouped into dry ice, wet ice, and dirt/rocks, although transitions between types are fluid.” Maybe finish with, “although gradations between these groups occur”. Replace “fluid” with gradations throughout.

Changed as suggested and checked manuscript for use of “fluid”.

24: Explain that firn cover is lost when persistent loss of snow cover in the accumulation zone exposes the firn (Fischer, 2011).

Changes sentence to: “Glaciers in the Eastern Alps are losing mass rapidly, and due to persistent loss of snow cover exposing the underlying firn (Fischer, 2011), many have lost much of their firn cover.”

59: “. . .relatively recent times”, be more specific.

Replaced with “...throughout approximately the last decade”.

74: Azzoni et al. (2016) also found a significant impact from rain water.

Changed “as well as significant effects of melt water” to: “as well as significant effects of melt and rain water”

76: What is the basis for Brun et al (2015) stating importance of remote sensing in albedo assessment?

Brun et al (2015) point out that satellite products are critical for glaciological studies in data sparse regions such as the Himalayas, where their study sites are, as in situ data are often not available and glaciers may not be easily accessible. In their study they reconstruct annual mass balance from MODIS albedo data for two glaciers, validating this with in situ data. They suggest that this method can be applied to certain other glaciers in the HKH region from which no in situ mass balance or albedo data is available. This highlights that 1) remote sensing is often the only way of getting albedo data and 2) other important glaciological work can be carried out using remote sensing derived albedo data.

Changed text as follows to makes this clearer:

"Brun et al. (2015) highlight the importance of remote sensing data for monitoring of glacier albedo changes in remote regions..."

To

"Brun et al. (2015) highlight the importance of remote sensing data for monitoring of glacier albedo changes in remote regions where data collection on the ground is impossible or impractical..."

77: Resolution of Naegeli et al. (2015) aerial albedo observations?

From Naegeli et al. 2015: "The flying altitude of 4000 m above ground level (a.g.l.) in combination with an instantaneous field of view (FOV) of 0.0025° resulted in a surface projected pixel resolution of ~ 2 m."

Changed sentence to include resolution:

"Naegeli et al. (2015) use in situ spectrometer and airborne image spectroscopy data with a pixel resolution of approximately 2m to classify glacier surface types"

96: Is it worth observing that for degree day modelling changing albedo with time would alter parameters in the model.

Added the following sentence to the paragraph: "In addition, delineating temporal variability of reflectance properties is relevant to degree day modelling, as a changing albedo would alter parameters in the model."

109: Given the illustrations in Figure 2 leverage these with terminus retreat from 1990- 2017 and for the accumulation zone what is the mean AAR during this same period 1990-1999, 2000-2009 and 2010-2017?

The requested AAR values are as follows:

1990/91-99/00: 0.35

2000/01-09/10: 0.18

2010/11-17/18: 0.12

Jamtalferner has experienced a rapid loss of firn and AAR was 0/the glacier was essentially snow free in the hydrological seasons 2002/03, 2014/15, and 2016/17. AAR values are contained in the data sets downloadable at: https://doi.pangaea.de/10.1594/PANGAEA.818772

We removed this figure in the revised manuscript based on suggestions by Reviewer 2. We added the mean AAR values for the 1990/91-99/00 and 2010/11-17/18 periods to the text.

117: “Along each profile line spectra are gathered at equal intervals, with 14 profile lines containing 11 spectra spaced at 2(?)m and 2 profiles containing 40 spectra gathered at a higher resolution of 0.5(?) m.” // 132: “Google Earth Engine” // 161: “gradational” instead of “fluid”

Changed as suggested.

196: Profile 8 seems to have the least agreement in Figure 9 between field are remote sensing data, why?

We have attached photos of the profile at the end of this document to provide visual context. Profile 8 crosses a section of ice where the contrast between dark and bright areas is comparatively strong. The profile is roughly at a right angle to the flow direction and there are “stripes” of meltwater channels and/or dirt that cross the profile. The profile has a comparable number of individual spectra with reflectance values above and below the profile mean, i.e. it is not a dark profile with a few bright outliers (compare e.g. to P6 in Fig 8) or vice versa (e.g. P3), but alternates along the profile line. Agreement with the remote sensing data is decent for the darker spectra in P8 but the bright values are not captured.

While we cannot rule out that the lack of agreement between the field and remote sensing data is due to an unusually unfortunate/unrepresentative positioning of the field measurement points in the satellite pixels, this may be an instance where the diurnal melt cycle and the associated presence/absence of water on the surface exacerbates the contrast between the dark and bright sections of the profile. In the bright sections, the porous weathering crust and cryconite hole structures appear to be drained of water, while the depressions of the melt channels are noticeably wet. Cook et al. 2015 (https://doi.org/10.1002/hyp.10602) indicate the occurrence of “sudden drainage events” in the weathering crust on a day-to-day time scale and a diurnal cycle of the hydrology of the weathering crust driven by meteorological conditions (radiation, turbulent fluxes). The time of day of a satellite overpass would determine which stage of this cycle the satellite “sees” and consequently the satellite data would not capture this variability. A more definitive explanation would require further study and dedicated field experiments designed specifically to explore this aspect of reflectance variability – we hope to do this in the future.

We have added a version of the above commentary to the discussion (last paragraph of section 4.2).

204: Figure 8 has excellent potential for the direct spatial correlation of the Sentinel albedo to the point measurements. I think showing all the profiles prevents being able to visualize the relationship. I suggest focusing on a few of the same profiles that were a focus of Figure 5 and provide a range of conditions ie. P 3, 5, 8, and 11. Anzoni et al (2016) noted a future goal of generating an albedo map. Is that feasible for the area of the glacier shown in Figure 1?

Changed the figure so that only profiles 3, 5, 8, and 11 are shown. We rescaled the circles in order to give a visual representation of the horizontal uncertainty of the GPS coordinates.

Based on high resolution, close range digital images of the ice surface at Forni glacier, Anzoni et al (2016) develop a relationship between the area ratio of ice covered by fine debris and clean ice (d) such that albedo can be derived for a given area if d is known. To apply this method to the area of the glacier in Figure 1 would require an estimate of d, for which we would need close range imagery of the ice surface for the entire area. We could perhaps apply the method of Anzoni et al. to the photos we took of the ice surface at our sampling sites, but this would still result in albedo values only at our sampling sites without addressing the question of how representative these locations

are for the rest of the glacier area and how albedo might be interpolated between them. We hope that our work may eventually contribute to methods for producing high resolution albedo maps, but do not think making such a map is feasible with our current dataset and the method described by Anzoni et al. Naegeli et al. 2017 produce an albedo map based on a classification of different surface types in remote sensing imagery. This would be probably be the approach of choice for Jamtalferner given the currently available data for the site.

210: This is a key observation. What have other studies found in terms of the over/under-estimate transition?

We have not been able to find many other studies with explicit information on this issue specifically over glacier ice surfaces. Hendricks et al. (2004) state for measurements at Hintereisferner: “Except for ice, the glacier reflectances derived from the satellite image are large underestimations when comparing them to the spectrometer measurements. A maximum underestimation of 139 % was found for firn in band 4. New snow, with the highest reflectance of 86 % is predicted most accurately within a confidence interval of 15 - 18 %. The reflectance of ice seems to be highly variable with both under -and overestimations of up to 76 % and 31 % respectively.” This refers to Landsat ETM+ imagery acquired about 2 weeks before the corresponding field measurements. We have cited this in the revised manuscript with additional discussion of possible explanations for the location of the over/under-estimate transition. Further measurements specifically investigating this issue are needed to truly explain this effect.

226: The variation in energy balance as albedo/debris cover changes spatially and temporally was a focus of Nicholson and Benn (2006) provided a nice overview of this from Ghiacciaio del Belvedere. They observed for debris cover areas the dominant energy contribution varied from sensible heat to shortwave radiation due to decreased albedo and higher surface temperatures. They further found that for dry debris cover, sensible heat flux became negative as debris cover thickened, because of higher surface temperatures and that longwave radiation became negative even for thin debris cover.

We have added the following note in the discussion to reflect the findings on Nicholson and Benn: “Nicholson and Benn (2006) indicate that the surface albedo of ice with scattered debris can be simulated in a modelling approach be linearly varying between clean ice albedo values and values for debris, but this does not necessarily account for other types of surfaces and even the clean ice albedo can vary considerably, especially if liquid water is present.”

231: How significant is the time of day variation in albedo? How consistent would this variation be from day to day? Moller and Moller (2017) provide one measure of this in an examination of spatiotemporal variations of albedo across Svalbard glaciers, recognizing this is a larger scale model albedo product. Nicholson and Benn (2012) examining Ngozumpa Glacier identify surface albedo variation across an area of varied debris cover, as well as the changing diffusivity through the melt season. The surface temperature variation of this glacier in the Himalaya would be much different than in the Alps, yet the continuous record compiled does provide context to the degree of variation and the potential importance of ongoing point measurments. They observe the importance of distinguishing wet vs dry surfaces. Azzoni et al (2016) note the increased albedo due meltwater presence during the middle of the day to albedo, while rain led to increased albedo for several days.

We have added the following paragraph to the discussion to address these points:

"Cook et al. (2016) indicate the occurrence of "sudden drainage events" in the weathering crust on a day-to-day time scale and a diurnal cycle of the hydrology of the weathering crust driven by meteorological conditions (radiation, turbulent fluxes). The time of day of a satellite overpass would determine which stage of this cycle the satellite sees and consequently the satellite data would not capture this variability. In order to assess how much time of day of the overpass could systematically affect the representativeness of the satellite date for actual ground reflectance, it needs to be determined how significant and how consistent the diurnal cycle is. To do this, the driving processes must be identified, keeping in mind that these may be different for different types of glaciers and that different causes of short-term albedo change can overlap. E.g.: Azzoni et al. (2016) point out that meltwater increases albedo around midday in a daily cycle, while rain causes increased albedo for more than one day. A seasonal cycle of albedo has been demonstrated in previous observational studies and modelling efforts of broadband albedo, which also highlight the importance of continuous measurements (e.g. Hoinkes and Wendler, 1968; Nicholson and Benn, 2012; Möller and Möller, 2017)."

In order to quantitatively answer the questions posed at the beginning of this comment, we very much hope to expand our data collection at Jamtalferner and install instrumentation that would allow continuous point measurements.

249: Similarly, the question of how well the albedo variations need to be resolved to model or understand surface processes need to be acknowledged/discussed. One reason a relatively sparse ablation stake network can represent ablation during a melt season is that despite significant surface changes the spatial distribution of energy balance over time tends to balance. Your Figure 5 illustrates this that though albedo varies considerably along the Profile 3 and 11, and the profiles have been exposed ablation ice for some period, the ice surface is relatively even. Energy balance distribution across an ice surface in a small area responds to the variations in surface level, albedo and debris cover.

"Similarly, the question of how well the albedo variations need to be resolved to model or understand surface processes need to be acknowledged/discussed." - This is a valid and interesting point of discussion. We suggest that the answer to this question depends on the processes one is trying to understand and the scale at which they occur. In the context of glacier wide ablation monitoring via the direct glaciological method, resolving sub daily and sub meter variations is perhaps not exactly a very pressing need, but, in our opinion, still interesting. The area of the glacier shown in Figure 1 contains 9 ablation stakes, which we maintain as part of our mass balance monitoring program at Jamtalferner. We observe significant differences in the amount of melt that occurs at these stakes. Aspect, shading, and slope angle of course play a strong role in this, as does locally increasing debris cover. We hypothesize that darkening due to water on the glacier surface is more of a factor at some of our stakes than at others, depending e.g. on their position in relation to seasonally shifting meltwater channels. We would like to eventually achieve a clearer separation of the influence of these factors (especially the influence of water), their relative magnitudes, and possible changes over time. We think that small scale reflectance monitoring can contribute valuable insights in this context.

Additionally, data with high spatial and temporal resolution seems essential to improve understanding of micro-hydrological processes in the weathering crust and how these may affect a possible larger scale darkening of increasingly snow free glaciers, e.g. by favoring or impeding the growth of ice algae, or the collection/washing out of cryoconite.

We have added the following section to the discussion:

“4.3. Relevance of small-scale variability, way forward

The reflectance properties of ice are a central part of mass and energy balance modelling, usually in the form of a glacier wide broad band albedo, or using one value for ice in the ablation zone and one for snow covered areas. Resolving local albedo variations at a very small, sub-pixel scale is not required for regional or global studies, provided the albedo parametrization captures the conditions on the ground adequately for the region of interest. In their important 2015 study, Naegeli et al. find that Sentinel-2 and Landsat-8 reflectance data are within the suggested accuracy requirements for global climate modelling (±0.05, Henderson-Sellers and Wilson, 1983) over their study site, Glacier de la Plaine Morte in Switzerland. In the same study, they report a 10% difference in modelled mass balance when a spatially distributed albedo is used to force the model as opposed to a single, glacier wide albedo. Significantly larger differences occur in parts of the glacier where water is present on the surface or the ice surface contains a lot of light-absorbing impurities. While the glacier wide impact of a spatially distributed albedo on model results may be relatively small, this highlights that resolving local variability of reflectance properties and its causes is important for accurately predicting the future evolution of individual glaciers, especially in cases where the firn covered area is gone or greatly reduced and rapid melt is occurring. Only once the problem of different scales comparing point and spatially averaged data is solved, the relationship between albedo variability and mass balance point and averaged data can be tackled to calculate the effects on mass balance at glacier-wide or regional scale.

Aside from directly mass and energy balance related applications, reflectance data with high spatial and temporal resolution is essential to improve understanding of micro-hydrological processes in the weathering crust and how these may affect a possible larger scale darkening of increasingly snow free glaciers, e.g. by favoring or impeding the growth of ice algae, or the collection/washing out of cryoconite or other impurities. High resolution time series of spectral reflectance at representative locations in the ablation zone are needed to assess how changes in wetness and temperature, surface texture (cryoconite formation, roughness changes during the season), biotic productivity, deposition of sediment by melt water and rain affect albedo on a small spatial scale, throughout the day and over the course of the ablation season. Establishing measurement efforts aimed at generating such time series on glaciers with existing mass balance monitoring networks would be highly desirable.”

260: A significant source of uncertainty for what?

Surface reflectance and parameters that might be derived from it are key variables in glaciological modelling and uncertainty therein accordingly contributes to overall model uncertainty. This has implications for applications such as modelling runoff and catchment hydrology.

We have rephrased this to read “...source of uncertainty in modelling applications...”

271: Need a reference from a different region to emphasize this point.

We have removed the sentence this refers to as a part of the restructuring of the discussion and conclusion sections.

280: Did you sample spectra at any location over a period of time? If so, this helps relate the logistical challenge of temporal albedo monitoring.

We have not had opportunity to do that but hope we will in the future.

Profile 8, looking east:

Profile 8, looking west:

Reviewer 2

Reviewer comment:

In this paper, the authors present a comparison between spectral reflectance measurements of bare ice carried out in the ablation zone of the Jamtalferner glacier, Austria with concurrent Sentinel-2 and Landsat-8 acquisitions. In a first step, the spatial variability of the manually acquired surface albedo across the ablation zone of the glacier is presented, highlighting large differences in reflective properties from dry clean ice to surfaces covered in mineral and organic debris. Secondly, the paper focusses on comparing the field measurements with atmospherically-corrected satellite reflectance products to investigate whether physical processes related to deglaciation are fully captured by optical Earth Observation sensors. Results show that the differences observed between the ground-based and satellite measurements are not uniform depending on the wavelength, the sensor or surface type. The authors conclude by suggesting that further in-situ monitoring efforts are needed to be able to use satellite-derived reflectance for glacier change monitoring.

General assessment

The comparison of in-situ surface reflectance measurements with satellite-derived products is of great interest for anyone involved in space-borne observations of glaciers and more generally glacier surface processes monitoring,

and in that sense, the work here is timely and most welcome. I particularly commend the use of openly accessible world-wide available satellite data rather than higher-resolution commercial data, making the applications available to a wider audience. The article is overall well written, apart from a couple of minor approximations (see detailed comments). However, the manuscript presents two major shortcomings that leave the reader missing significant information (see General comments paragraph below).

In summary, this article would have merit for publication in The Cryosphere if the major points referred to below are addressed. Currently, the Methods and Discussion sections are insufficient.

Author response:

We thank the reviewer for their time and the detailed and constructive commentary. The points of criticism are valid and we will address them in a revised version of the manuscript, following the suggestions by both reviewers.

To remedy the main shortcomings of the Methods and Discussion sections as specified in this review, we have:

1) expanded the Methods section, particularly the description of the measurement protocol for the in situ data collection.

2) restructured and significantly expanded the discussion section to address the specific issues pointed out by the reviewer in the comments below.

We address further comments individually below.

General comments

The first deficiency mentioned in the paragraph above concerns the presentation of the Methods. The ground measurements of spectral reflectance presented in Section 2.2 (7 lines) are largely insufficient for a piece of work dedicated to comparing ground measurements to satellite products. Indeed, the section barely skims over the way measurements were collected and crucial information is lacking to clearly understand the comparisons made.

See response to specific comments below.

---

1. When were the measurements collected? No date or time of measurements is provided in the section describing ground measurements. The reader has to wait until Section 2.3 to understand that the measurements were acquired on 4th September 2019. Over what time period (start and end of acquisitions) was the data acquired? This is of significant importance for the comparison of the data, e.g. did the surface have time to change between the satellite overpass and the ground measurements?

Ground measurements were taken on 4th September 2019, between approximately 10 am and 3 pm local time. The Sentinel overpass occurred at 10:20 GMT on Sept. 4. The Landsat overpass occurred at 10:10 GMT on Sept. 3. We have specified this in the revised manuscript and begin the section describing the ground measurements by stating the date and time period of the data collection. The glacier surface is constantly changing to some extent, but weather conditions on Sept. 3 and 4. were very favorable and there was no change introduced by factors such as precipitation of deposition of impurities through wind during the time period between the acquisitions. We

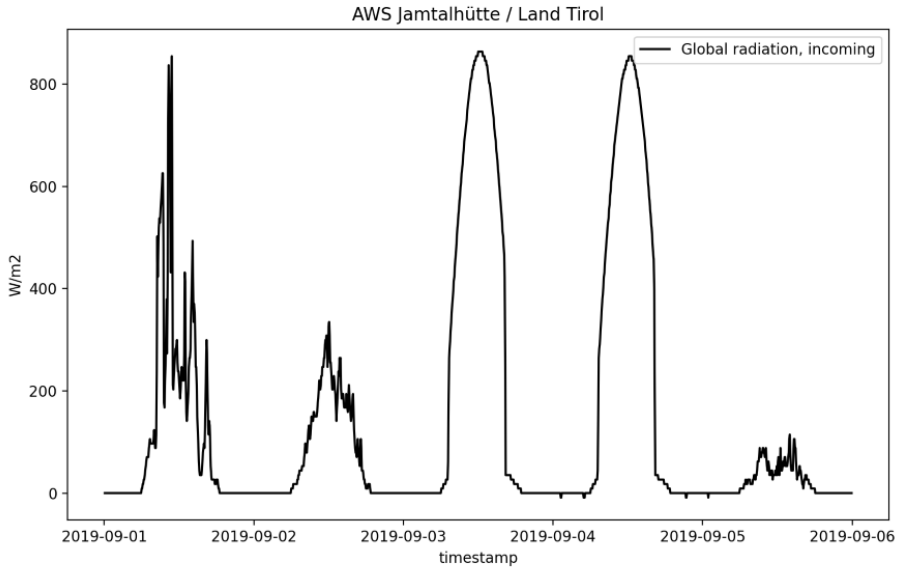
have added information on the weather situation in the Methods section and added further comments on this in the Discussion.

---

2. There is no description of the environmental conditions during the acquisition, e.g cloud cover. Even a small amount of cloud cover, such as the presence of rapidly changing cirrus can introduce uncertainties of several percent in the measured reflectance.

The study site is free of cloud cover in both satellite images and the weather was sunny and dry on both days. Attached is a plot of incoming solar radiation from a weather station a short distance below the glacier showing the cloud free conditions of Sept. 3 and 4. We have included a description of the weather conditions in the revised methods section.

Figure: Incoming solar radiation at Jamtalhütte automatic weather station from September 1 2019 to September 5 2019. Data provided by the hydrology office of the state government of Tyrol, who operate this station.



---

3. The method for measuring the distance between the points on the profile is indicated, but how were the measurements geo-located in the field? Were there any GPS points acquired (especially as the authors refer to “GPS profile” in figure 1), with what uncertainty? The uncertainty in the positioning of the ground spectra may impact your point-to-pixel comparisons (to be addressed in the Discussion also).

GPS points were taken at the start and end point of each profile line, using a standard handheld GPS device. The horizontal uncertainty is < 3m. We have specified this in more detail in the expanded description of the in situ measurements and have added an approximation of the uncertainty in the point-to-pixel comparisons due to the GPS uncertainty in the results, which we comment on further in the discussion.

---

4. The measurement protocol is not described sufficiently, leaving the reader with a number of interrogations: how were the measurements carried out: was the ASD fibre optic handheld or placed on a device to reduce operator interference (Fig 3 in Wright et al. 2014, Kimes et al. 1983)? Did the authors use an optical lens on the fibre optic (if so, what field-of-view)? What height was the collector from the surface / spectral panel when performing the measurements? A description of how the measurements were performed is desired, or at the least, if the authors were following an existing protocol, a reference to the article is expected.

We have expanded the description of the measurements (Section 2.2). The fibre optic was handheld and used without an optical lens, at a distance of 35cm above the ground. This results in a circular field of view with a diameter of approx. 16cm. Our usage of the ASD device is similar to that of Naegeli et al. (2015, 2017) and Di Mauro et al. (2017), who carried out comparable measurements on glacier surfaces. We have cited these publications in this section.

---

5. The description of the processing of the raw ASD is missing. There are numerous steps to be carried out during the processing of data, including the application of instrument or spectral calibration files. In the current state, the description of the processing is too vague.

We used a feature of our instrument that saves the white reference measurement to the RAM of the instrument's computer. When this option is enabled, subsequent reflectance measurements are calculated with respect to the reference and the result of this calculation is saved to the output data file, such that there is no separate file for the reference and the output ASD files contain the calibrated values. We have added this information to the methods section.

---

6. The authors are not clear about the physical quantities measured. The title reads "Small scale variability of bare-ice albedo at Jamtalferner, Austria", and the author summarise the body of work on broadband and spectral albedo. However, in the methods, the field acquisitions are referred to as spectral reflectance and the (limited) description of the measurement protocol leads the author to believe that the authors are recording hemispherical–conical reflectance. The ground measurements are then compared to surface reflectance products derived from Sentinel-2 and Landsat-8. Particular care should be observed when describing remotely sensed quantities and I recommend that the authors verify inconsistencies throughout the paper. Very useful references in that sense are Schaepman-Strub et al., 2004, 2006 (besides an important corpus on the subject).

The reviewer's assessment here is correct. We have checked the manuscript for occurrences of these inconsistencies and define the quantities more clearly at the beginning of Section 1.2, referring to the works of Schaepman-Strub et al. Thank you for pointing these out, they are indeed very helpful.

The second shortfall mentioned in the overall remarks concerns the Discussion, that does not do justice to the paper. Indeed, in its current state, the section repeats the introduction and doesn't address the rich results obtained by the authors. The key points presented in the results are barely brushed past and the discussion on the limitations of the methods employed and possible explanations for the results obtained are missing. The paragraph starting P8, L247 would deserve (consequential) expanding in regard to the results obtained. By restructuring the

Discussion section, significant value could be brought to this otherwise valuable contribution to the observation of glacier ablation zones based on optical Remote Sensing.

We accept the reviewer's criticisms of the discussion and have rewritten and significantly expanded it. It now includes a section discussing the points relevant to the paragraph above, as well as other commentary that was previously lacking.

Specific comments

- P1, L14: in the Optical Remote Sensing community, ground reflectance is commonly referred to as Bottom-Of-Atmosphere (BOA) reflectance. I am not suggesting to replace the term, but maybe add a mention to BOA.

Changed sentence to include BOA in parentheses after 'ground reflectance'.

"...and are compared to the respective ground reflectance (Bottom-Of-Atmosphere) products"

- P1, L27: "The magnitude and [. . .] local production rates." > Although you go into further details later in the introduction, citations are missing here.

Added citations.

- P4, L106: Figure 2 and 3 seem irrelevant in the context of this paper that focusses on the comparison of ground and satellite acquisitions of reflectance and not the evolution of the surface properties over time. I suggest their removal, as they cloud the overall message. Rather, the satellite images (used in the study), of the glacier tongue with the profiles overlaid would be a nice addition to the paper.

Removed Figures 2 and 3 and added a new Figure 2 with two panels showing the satellite images.

- Section 2.3: Table 3 would benefit being completed with additional information on the Sentinel-2 and Landsat-8 acquisitions, such as acquisition time or the angular information (solar and viewing angles). A column with the corresponding ground measurement information would be a plus.

We have added additional columns as suggested to Table 3 (satellite acquisitions) and to Table 2 (in situ measurements).

- P5, L126: The acquisition time of Sentinel-2 is not specified: yet this information is important to investigate the differences between the measurements from both sensors.

We have specified this in the text and added it to Table 3 as suggested.

- P5, L139: Did the authors consider integrating the spectral measurements using the (available at least for Sentinel-2) spectral response of each band? Do the authors think that the difference with the average would be negligible or not?

We extracted the associated measured in-situ reflectances for each spectral range per band per sensor. Thus, the averages of the in-situ measurements can be directly compared with the reflectances per spectral band. For the comparison of Landsat 8 vs Sentinel-2 mean reflectances it should be noted that the BOA reflectances are used as they are provided by NASA and ESA, respectively, i.e. products are prepared with different radiative transfer models and different parameterizations of the atmospheric conditions. For a proper usage of the spectral response function, the L1C data should be processed using the same atmospheric correction approach and parameters. Although this is another very interesting topic, it is out of the scope of our study with the main objectives (i) analysing the spatial variability of the reflectance on a glacier's ablation zone and (ii) comparing the commonly used satellite L2A products with in-situ measurements.

- P6, L175: This is an interesting find. Have the authors considered the difference in viewing/solar geometries between the two acquisitions? The strong anisotropy of the ice could partly explain the differences (see the previous comment). Basic simulations of ice reflectance (using e.g. Malinka et al. 2016) could help investigate this point. To be clear, this is not expected from the authors, but a point that could be worth thinking about for future studies. Another factor that could influence the differences observed could be the different atmospheric corrections schemes used (a reference in the Discussion would be of value).

We have added comments on anisotropy and solar angles, as well as the issue of atmospheric correction schemes in the discussion. These factors likely contribute to the differences between Landsat and Sentinel, but without targeted further analysis and data collection it is not possible to quantify the contribution of each factor. We have also added a citation of the interesting Malinka et al. (2016) paper. Modelling reflectance properties is indeed beyond the scope of this study but would be very interesting in the future, if we can expand our monitoring situation at Jamtalferner as we hope. We believe that any modelling would have to be tuned carefully for the kind of ice surface one is dealing with, especially for very heterogenous surfaces like we have at our site. Malinka's work is based on in situ data measured on sea ice, which appears to be significantly more uniform in terms of texture and reflective properties. Their case of dark and wet sea ice still appears brighter than the majority of our spectra.

- P6, L183: This suggests that for surfaces with strong sub-pixel variability the resolution of the images is essential for an accurate description of the surface. The representativeness of field sampling when comparing in situ measurements to satellite images is of particular interest in the snow and ice community. Did the authors consider investigating the sensitivity to resolution by degrading the 10m bands to 30 then 60 meters?

We have added a comparison of differences between in situ and satellite data for the original pixel sizes and pixels resized to 30 and 60 m in the results section.

- P7, L200: Very interesting find, which links to the question of the representativeness of the in-situ sampling. It would be nice to see this point further discussed in the Discussion section.

We agree that representativeness of the in situ sampling is an important issue and discuss this in a newly added section 4.2.

- P7, L206: Again, this key result deserves some discussion.

This is now addressed in the expanded discussion section (section 4.2)

- P8, L222-226: the observation is repeated from the introduction.

Removed as a part of the rewrite of the discussion section.

- P8, L228: This paragraph should be placed in the context of the results of this study and is overall too vague.

Removed as a part of the rewrite of the discussion section

- P8, L234: Again, the paragraph reads like an introduction and doesn't have a place in the discussion.

Removed.

- P8, L244: Some lines of reflection in the context of the authors' study, such as discussing the anisotropy of ice in line with the differences in overpass geometries would be most welcome here.

We have added a discussion of these issues in Section 4.1

- Figure 4: is the highlighting of the maximum and minimum spectra necessary? A single emphasised black spectrum of the mean and the others in light grey could be clearer (if the authors agree).

Changed figure as suggested.

- Figure 6: in the printed manuscript, the tape measure is unreadable in the photos. Adding a small simple scale bar int the pictures would help grasp the scale of the images. This is an interesting figure showing the important variability of reflectance across the glacier.

Added scale bars to the figures.

- Figure 7: the caption is unclear and the reader has to read Section 3.2 several times to understand the figure. The term “ground measurements” for satellite images (P20, L419) is confusing. I would suggest revising the caption to clearly state what the blue and orange bars represent.

Rephrased the figure caption and the associated part of the text in order to improve clarity.

- Table 1: why are the PROMICE network measurements not referenced (Fausto and van As 2019)? They have been used for satellite calibration also.

Added references to PROMICE to the revised table and in the text.

Technical corrections

- P1, L12: exits > exist. // - P1, L16: at dark spectra > for dark spectra - // P1, L 25: “so that darker bare ice is exposed” > I suggest specifying "in Summer” to be more precise. // - P2, L33: ”gap of knowledge” > “knowledge gap”

Changed as suggested.

- P2, L39: “comparatively high resolution” > Comparatively to what? Please be more specific. Sentinel-2 and Landsat-8 could be referred to as “medium resolution sensors”.

The comparative statement was meant mainly in reference to the resolution of MODIS, but this was poorly phrased.

Changed:

“2) Compare commonly used, comparatively high resolution satellite-derived reflectance products with in situ measurements, highlighting areas in which further study is required if ongoing processes related to deglaciation are to be fully captured by satellite data.”

To:

"2) Compare reflectance products derived from Landsat 8 and Sentinel 2 data with in situ measurements, highlighting areas in which further study is required if ongoing processes related to deglaciation are to be fully captured by satellite data."

- P2, L59: "in relatively recent times" > Please be more specific.

Replaced with "...throughout approximately the last decade".

- P3, L86: "different kinds of remote sensing" > this phrasing is a little vague, could you clarify?

Changed:

"....albedo products derived from different kinds of remote sensing data..."

To:

"...albedo products derived from airborne imaging spectroscopy (APEX) and Landsat and Sentinel data..."

- P4, L122: "specdal" > "spectral"

Thisshould be "SpecDal" and refers to a python package we used to process the data. Rephrased to make this clearer and added a citation of the documentation for the package.
https://specdal.readthedocs.io/en/latest/index.html

- Figure 9: please specify the wavelength of band 3.

Added wavelength in the figure caption.

- Table 2: is lacking the first column header

Added missing column header.

- Table 1, 2 and 3: I am guessing that the authors will format the tables correctly in the next iteration? They are currently unpleasant to read.

We have reformatted the tables.

References Reviewer

Fausto, R.S. and van As, D., (2019). Programme for monitoring of the Greenland ice sheet (PROMICE): Automatic weather station data. Version: v03, Dataset published via Geological Survey of Denmark and Greenland. DOI: https://doi.org/10.22008/promice/data/aws

Kimes, D. S., J. A. Kirchner, and W. Wayne Newcomb. "Spectral radiance errors in remote sensing ground studies due to nearby objects." Applied optics 22.1 (1983): 8-10.

Malinka, Aleksey, et al. "Reflective properties of white sea ice and snow." Cryosphere 10.6 (2016).

Schaepman-Strub, G., et al. "About the importance of the definition of reflectance quantities-results of case studies." Proceedings of the XXth ISPRS Congress. 2004.

Schaepman-Strub, Gabriela, et al. "Reflectance quantities in optical remote sensingâA˘TDefinitions and case studies." Remote sensing of environment 103.1 (2006): 27-42.

Wright, Patrick, et al. "Comparing MODIS daily snow albedo to spectral albedo field measurements in Central Greenland." Remote Sensing of Environment 140 (2014): 118-129.

Author references:

Naegeli, K., Damm, A., Huss, M., Wulf, H., Schaepman, M., & Hoelzle, M.: Cross-Comparison of albedo products for glacier surfaces derived from airborne and satellite (Sentinel-2 and Landsat 8) optical data, Remote Sensing, 9(2), 110, 2017. Naegeli, K., Huss, M., & Hoelzle, M.: Change detection of bare-ice albedo in the Swiss Alps, The Cryosphere, 13(1), 397- 412, 2019.

Di Mauro, B., Baccolo, G., Garzonio, R., Giardino, C., Massabò, D., Piazzalunga, A., Rossini, M., and Colombo, R.: Impact of impurities and cryoconite on the optical properties of the Morteratsch Glacier (Swiss Alps), The Cryosphere, 11(6), 2393, 2017.

Small scale spatial variability of bare-ice ~~albedo~~reflectance at Jamtalferner, Austria

Lea Hartl (1), Lucia Felbauer (1), Gabriele Schwaizer (2), Andrea Fischer (1)

1) Institute for Interdisciplinary Mountain Reasearch, Austrian Academy of Sciences, Technikerstraße, 21a, ICT, 6020 Innsbruck, Austria
2) ENVEO GmbH, Fürstenweg 176, 6020 Innsbruck, Austria

Abstract

As Alpine glaciers ~~recede, they are quickly becoming~~become snow free in summer ~~and, accordingly,~~, further spatial and temporal variations in ice albedo increasingly ~~affect~~accentuate the melt regime~~.~~ and recession occurs. To ~~accurately model~~include this feedback mechanism in models of future ~~developments, such as~~ deglaciation ~~patterns~~, it is important to understand the processes governing broadband and spectral albedo at a local scale. However, little in situ ~~data~~reflectance data has been measured in the ablation zones of ~~ice albedo exits.~~mountain glaciers. As a contribution to this knowledge gap, we present spectral reflectance data (Hemispherical-Conical-Reflectance-Factor) from 325 to 1075 nm collected along several profile lines in the ablation zone of Jamtalferner, Austria. Measurements were timed to closely coincide with a Sentinel-~~-~~2 and Landsat-~~-~~8 overpass and are compared to the respective ground reflectance (Bottom-Of-Atmosphere) products. The brightest spectra have a maximum reflectance of up to 0.7 and consist of clean, dry ice. In contrast, reflectance does not exceed 0.2 ~~at~~for dark spectra where liquid water and/or fine-~~-~~grained debris are present. Spectra can roughly be grouped into dry ice, wet ice, and dirt/rocks, although ~~transitions~~gradations between ~~types are fluid~~these groups occur. Neither satellite captures the full range of in situ reflectance values. The difference between ground and satellite data is not uniform across satellite bands, between Landsat and Sentinel, and to some extent between ice surface types (underestimation of reflectance for bright surfaces, overestimation for dark surfaces). We wish to highlight the need for further, systematic measurements of in situ spectral ~~albedo, its~~reflectance properties, their variability in time and space, and in- depth analysis of time-synchronous satellite data.

1. Introduction

1.1 General context and aims

Under ongoing climate change, mountain glaciers are retreating at unprecedented rates (Zemp et al, 2015, 2019). Glaciers in the Eastern Alps are losing mass rapidly, and due to persistent loss of snow cover exposing the underlying firn (Fischer, 2011), many have ~~also~~ lost much of their firn cover~~, so that~~. An increasing amount of darker bare ice is exposed~~. This~~ in Summer and at some glacier tongues, darkening of the ice has been observed (Klok et al., 2003). These feedback mechanisms in turn ~~increases~~increase the amount of energy absorbed and ~~accelerates~~accelerate melt (e.g. Paul et al., 2005; Box et al., 2012; Naegeli et al., 2017 & 2019). The ~~magnitude and variability of albedo~~reflective properties of glacier ice ~~is~~are affected by e.g. the absence or presence and amount of dust, pollen, debris, cryoconite, supraglacial water, and biota including local production rates~~.~~ (Dumont et al., 2009; Gabbi et al. 2015; Azzoni et al., 2016). Variability is understood to be high, but few measurements and models exist. In a glaciological context, the spatial and temporal variability of ice albedo is understudied compared to snow albedo.

We present spectroradiometric data on the spatial variability of bare-~~-~~ice ~~albedo~~reflectance at the tongue of Jamtalferner, Austria, aiming to contribute to closing the ~~gap of~~ knowledge gap in bare ice variability as an important feedback mechanism in glacier mass loss. Specifically, we aim to:

1) Provide a first-order quantitative assessment of spatial variability of surface reflectance in the ablation area of the rapidly melting Jamtalferner, quantifying possible ranges of spectral reflectance and qualitatively summarizing different surface types.
2) Compare commonly used~~, comparatively high resolution satellite derived~~ reflectance products derived from Landsat-8 and Sentinel-2 data with in situ measurements, highlighting areas in which further study is required if ongoing processes related to deglaciation are to be fully captured by satellite data.

1.2 In situ and remote sensing-~~-~~based change detection of surface reflectance properties of glacier ice

In the following section we summarize previous studies on this topic. For clarity, we begin with a note on terminology: Following the definitions and guidelines detailed in Schaepman-Strub et al., (2004, 2006) and Nicodemus et al. (1977), we use the term “albedo” for bihemispherical reflectance (BHR), including cases where this parameter is approximately measured with an albedometer. In situ measurements with field spectrometers – such as they were carried out for this study – generally represent Hemispherical-Conical-Reflectance-Factors (HCRF). For exact specifications of what is represented by satellite derived surface reflectance products we refer to the documentation of the respective products as this differs between sensors and product suites.

While it is generally understood that albedo is a major driving factor for the energy balance and radiative regime of glaciers, few studies discuss ice albedo and its variability at the local level. Early investigations of ice albedo were carried out by Sauberer in 1938. Building on this work, Sauberer and Dirmhirn (1951) showed that albedo is highly variable in time and space and strongly affects the radiation balance. They reported mean values of 0.37 for clean ice and 0.13 for dirty ice at Sonnblick glacier (Austria), a pronounced diurnal cycle of albedo related to refreezing of the surface, and influence of wind transported fine mineral dust. In another study based on measurements at Sonnblick, they highlighted that the collection of mineral dust in cryoconite holes affects albedo, as does liquid water, and showed a diurnal reduction of albedo of about 0.2 under clear sky conditions, which they attribute to melt-freeze cycles on the ice surface. (Sauberer and Dirmhirn, 1952). Jaffé (1960) also pointed out the importance of cryoconite and air content in the upper most ice layer for the radiative properties. Dirmhirn and Trojer (1955) presented a histogram-like curve of the frequency of different ice albedo values measured on the tongue of Hintereisferner (Austria): Broadband ice albedo ranges from <0.1 to about 0.58, with a frequency maximum at 0.28. Similar to the results from Sonnblick, melt-related diurnal albedo variations were also found at Hintereisferner. In a detailed study of the radiation balance at Hintereisferner, Hoinkes and Wendler (1968) showed the importance of summer snow falls onfor albedo, as well as seasonal changes in ice albedo, and their significant contribution to ablation.

The spectral reflectance of bare ice areas of Alpine glaciers, how it changes over time, and the associated driving processes at the glacier surface have become of increasing scientific interest in relatively recent times, alongsideConsidering the growing dominance of bare ice areas both compared to overall glacier area and in terms of glacier-wide mass- and energy balance, the sensitivity of the latter parameters to changing reflectance properties has become of increasing interest throughout approximately the last decade. Using a combination of mass balance data from multiple Swiss glaciers and the Landsat-8 surface reflectance product, Naegeli and Huss (2015) show that mass balance decreases on average by 0.14 m w.e. a−1 per 0.1 albedo decrease. In order to better delineate associated driving processes at the glacier surface, it is important to assess reflectance properties not only as broadband albedo at the scale of a glacier, but at a high spectral and spatial resolution. A number of studies attribute recent darkening of European glaciers to increased accumulation of mineral dust (e.g. Oerlemans et al., 2009, Azzoni et al., 2016) and black carbon (e.g. Painter et al., 2013, Gabbi et al., 2015). Similar findings have been reported from the Himalayas (e.g. Ming et al. 2012, 2015; Qu et al. 2014) and the Greenland ice sheet (Dumont et al., 2009). Some discussion remains as to whether the observed darkening is primarily due to the increase of bare ice areas compared to overall glacier area, or whether there is a darkening of the bare ice areas as such, and if so, whether bare ice areas are darkening due to local processes or large scale systemic change (e.g. Box et al., 2012; Alexander et al., 2014; Naegeli, 2019).

Different methodological approaches have been used to address specific changes in the surface characteristics of the ablation zone as they relate to changes in albedoreflectance properties and energy absorption across the electromagnetic spectrum: Using both hyperspectral satellite data and in situ HCRF measurements, Di Mauro et al. (2017) find that the presence of elemental and organic carbon leads to darkening of the ablation zone at Vadret da Morteratsch glacier (Switzerland) and discuss potential anthropogenic contributions. Azzoni et al. (2016) use semi-automatic image analysis techniques on photos of the ice surface at Forni glacier (Italy) to quantify the amount of fine debris present on the surface and its effect on the albedo. They find an overall darkening due to increasing dust, as well as significant effects of melt and rain water.

Brun et al. (2015) highlight the importance of remote sensing data for monitoring of glacier albedo changes in remote regions and compare MODIS data with in situ radiation measurements. Naegeli et al. (2015) use in situ spectrometer and airborne image spectroscopy data with a pixel resolution of approximately 2m to classify glacier surfacessurface types and map spectral albedo on Glacier de la Plaine Morte in Switzerland. Additionally, they highlight the difference in scale between albedo variability at the ice surface and the pixel resolution of satellite data and the need for detailed case studies combining ground truth data and remote sensing techniques to bridge this gap. In situ data is also essential for model verification, as shown e.g. by Malinka et al. (2016), who use reflectance spectra (HCRF) gathered on sea ice to validate modelled reflectance parameters.

In order to scale assessments of ice albedo from the local to a regional or global level, satellite-derived data are indispensable. Earlier in the satellite era, several studies carried out comparisons of albedo data measured on the ground and surface reflectance derived from Landsat-5 Thematic Mapper scenes, finding considerable differences between in situ and satellite data especially in the ablation area (e.g. Hall et al., 1989 & 1990; Koelemeijer et al., 1993; Winther, 1993; Knap et al., 1999). These works are mostly based on albedo data from a single location, such as an automatic weather station (AWS), and it was often not possible to carry out ground measurements so that they coincided with the satellite overpasses. More recently, Brun et al. (2015) highlight the importance of remote sensing data for monitoring of glacier albedo changes in remote regions where data collection on the ground is impossible or impractical and compare MODIS data with in situ radiation measurements. Albedo measurements from AWS sites on the Greenland ice sheet –

associated with the PROMICE and GC-Net monitoring networks - have been used to improve gridded albedo products based on MODIS data, showing the importance of using ground truth in conjunction with satellite data (Box et al., 2013; van As et al., 2017). Narrow-to-broadband conversions remain a challenge in this regard and commonly used conversions are typically designed for use with Landsat-5 or 7, rather than Landsat-8 or Sentinel-2, which increases the uncertainties inherently associated with any narrow-to-broadband conversion (Gardner et al, 2010; Naegeli et al., 2017). In addition, studies assessing the potential effects of anisotropy on satellite-derived surface reflectance data are sparse and the magnitude of associated uncertainties is hard to quantify (Naegeli et al., 2015 & 2017).

Naegeli et al. (2019) quantify trends in bare ice albedo for 39 Swiss glaciers using Landsat surface reflectance data products for a 17-year period. While they do not find a clear, wide spread darkening trend of bare ice surfaces throughout the entirety of their data set, they note significant negative trends at the local level, most notably for certain terminus areas. A detailed comparison of different albedo products derived from ~~different kinds of remote sensing data (~~airborne imaging spectroscopy (APEX) and Landsat~~, and~~ Sentinel~~, APEX)~~ data by Naegeli et al. (2017) further highlights the gap between albedo variability on the ground and ~~it's~~its representation in remote sensing data of varying resolution. A recent study by Di Mauro et al. (2020) uses in situ HCFR data and DNA analysis to show that ice algae affect albedo on a Swiss glacier.

Despite the growing body of work on this topic (see Table 1), reflectance properties – spectral as well as broadband~~–~~ ~~remains~~, local as well as regional, short time as well as seasonal - remain understudied compared to other parameters routinely recorded at Jamtalferner and other long-term glaciological monitoring sites. However, surface changes and associated changes of the spectral characteristics in the ablation area (e.g. due to debris cover, supraglacial meltwater, deposition of impurities) are expected to play a significant role in determining the future development of these glaciers. Incorporating relevant parameters into monitoring efforts is highly desirable. The accuracy of direct measurements of mass balance depends on the representation of all surface types in the stake network, and the correct attribution of unmeasured areas to measured stake ablation. Accordingly, a better understanding of how surface ~~albedo~~ types differ in terms of their reflective properties is required to maintain the stake network on a rapidly changing glacier. To this end, it is important to understand whether satellite-derived data can provide a basis for defining surface ~~albedo~~ classes to be covered by stakes, or whether it does not allow for the retrieval of the full bandwidth of ~~albedo~~reflectance variability relevant to the ice melt rate. In addition, delineating the temporal variability of reflectance properties is relevant to degree day modelling, as a changing albedo would alter parameters in the model.

Table 1: ~~Overview on~~ Measurements of bare ice ~~albedo measurements~~reflectance properties on mountain glaciers: Overview.

Glacier	Albedo type	Temporal resolution	Spatial resolution	Reference
Hintereisferner, AT	Total	Multiple days	Multiple points on different surface types	Dirmhirn and Trojer, 1955.
Hintereisferner, AT	Total	Multiple times on one day	2 points	Jaffé, 1960.
Northern China (glacier not specified)	Spectral	Not specified	Different surfaces	Zeng et al., 1984.
Forbindels, Greenland	Spectral	One measurement campaign	Regular grid of points around multiple study sites	Hall et al., 1990.
Hintereisferner, AT	Spectral	7 days during ablation season	Points along a profile	Van de Wal et al., 1992.
Austre Brøggerbreen, Midre Lovénbreen, Svalbard	Spectral, total shortwave	Multiple days during ablation season	1 point	Winther, 1993.
Haut Glacier d'Arolla, CH	Total	One measurement campaign	Multiple points	Knap et al., 1999.
Chhota Shigri, Mera Glaciers, Nepal	Total shortwave	Continuous AWS measurements	AWS location	Brun et al., 2015.
Forni Glacier, IT	Total	Multiple measurements during multiple years	Multiple points	Azzoni et al., 2016.
Glacier de la Plaine Morte, CH	Spectral	One measurement campaign	Multiple points	Naegeli et al., 2015.
Findelen, CH	Total	Continuous AWS measurements	AWS location	Naegeli et al., 2017.
Morteratsch, CH	Spectral	One measurement campaign	Multiple points	Di Mauro et al., 2017.
Jamtal, AT	Spectral	One measurement campaign	Multiple points	This study

Glacier	Albedo type	Temporal resolution	Spatial resolution	Reference
Hintereisferner, AT	Total	Multiple days	Multiple points on different surface types	Dirmhirn and Trojer, 1955.
Hintereisferner, AT	Total	Multiple times on one day	2 points	Jaffé, 1960.
Northern China (glacier not specified)	Spectral	Not specified	Different surfaces	Zeng et al., 1984.
Forbindels, Greenland	Spectral	One measurement campaign	Regular grid of points around multiple study sites	Hall et al., 1990.
Hintereisferner, AT	Spectral	7 days during ablation season	Points along a profile	Van de Wal et al., 1992.

Austre Brøggerbreen, Midre Lovénbreen, Svalbard	Spectral, total shortwave	Multiple days during ablation season	1 point	Winther, 1993.
Morteratsch, CH	Narrow band (Landsat TM bands 2 and 4)	One measurement campaign	Multiple points	Greuell and de Wildt, 1999.
Haut Glacier d'Arolla, CH	Total	One measurement campaign	Multiple points	Knap et al., 1999.
Hintereisferner, AT	Spectral	One measurement campaign	Multiple points	Hendriksa et al., 2003
Morteratsch, CH	Total	Continuous AWS measurements	Multiple AWS locations	Klok et al., 2003
Chhota Shigri, Mera Glaciers, Nepal	Total shortwave	Continuous AWS measurements	AWS location	Brun et al., 2015.
Forni Glacier, IT	Total	Multiple measurements during multiple years	Multiple points	Azzoni et al., 2016.
Glacier de la Plaine Morte, CH	Spectral	One measurement campaign	Multiple points	Naegeli et al., 2015.
Findelen, CH	Total	Continuous AWS measurements	AWS location	Naegeli et al., 2017.
Morteratsch, CH	Spectral	One measurement campaign	Multiple points	Di Mauro et al., 2017; Di Mauro et al, 2020
Greenland ice sheet	Total	Continuous AWS measurements	Multiple AWS locations	van As et al., 2017; Box et al., 2013
De Geerfonna and Elfenbeinbreen, Svalbard	Total	Continuous AWS measurements	1 AWS on each glacier	Möller and Möller, 2017
Jamtal, AT	Spectral	One measurement campaign	Multiple points	This study

155

2. Data, Methods, and Study Site

2.1. Study site – glaciological background

Jamtalferner was chosen for this study as it has the smallest end-of-season snow cover amongst the glaciers with long term mass balance monitoring in Austria. Jamtalferner is located in the Silvretta mountain range, which intersects the border between Austria and Switzerland. Jamtalferner is the largest glacier on the Austrian side of Silvretta (Fig. 1, size in 1970: 4.115km², size in 2015: 2.818km²). The history of scientific research at the site goes back as far as 1892, when length change measurements were first carried out, and a wealth of cartographic, geodetic, and glaciological data are available (Fischer et al., 2019). Orthophotos and cartographic analysis show that debris cover at the glacier terminus and in the lower elevation zones has increased (debris covered percentage of total area: 1.7% in 1970, 24.1% in 2015~~,~~ ~~Fig. 2 and 3~~), while firn cover is decreasing (firn covered area in 1970: 75%, in 2015: 13%, ~~Fig. 2 and 3~~mean accumulation area ratio (AAR) 1990/91-99/00: 0.35, mean AAR 2010-2017/18: 0.12, Fischer et al., in review).

Mass balance measurements via the direct glaciological method began in 1988/1989. In recent years, increasing mass loss was recorded across all elevation zones (~~Fig. 3~~ Fischer et al., 2016; Fischer et al., in review). The lowest elevation zones are dominant in terms of total ablation and thus net balance. Melt in the lowest altitudes has been increasing during the last two decades of negative mass balances and the variability of surface albedo at and near the glacier terminus affects melt over the full duration of the ablation season.

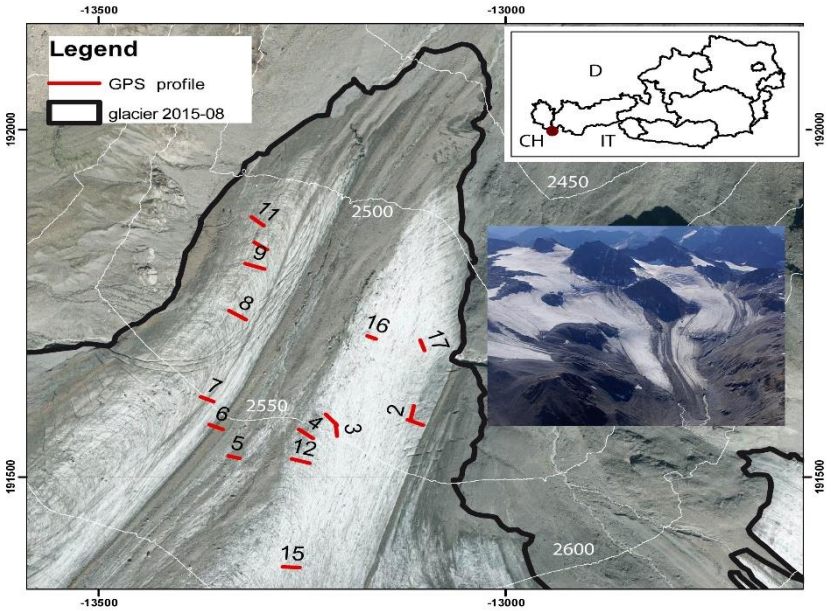
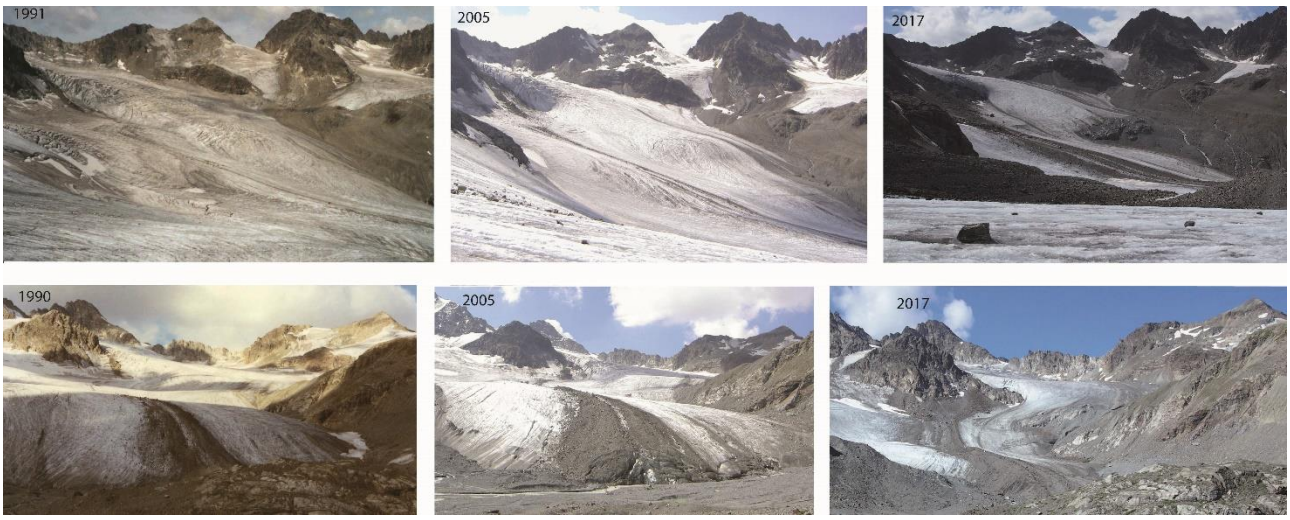
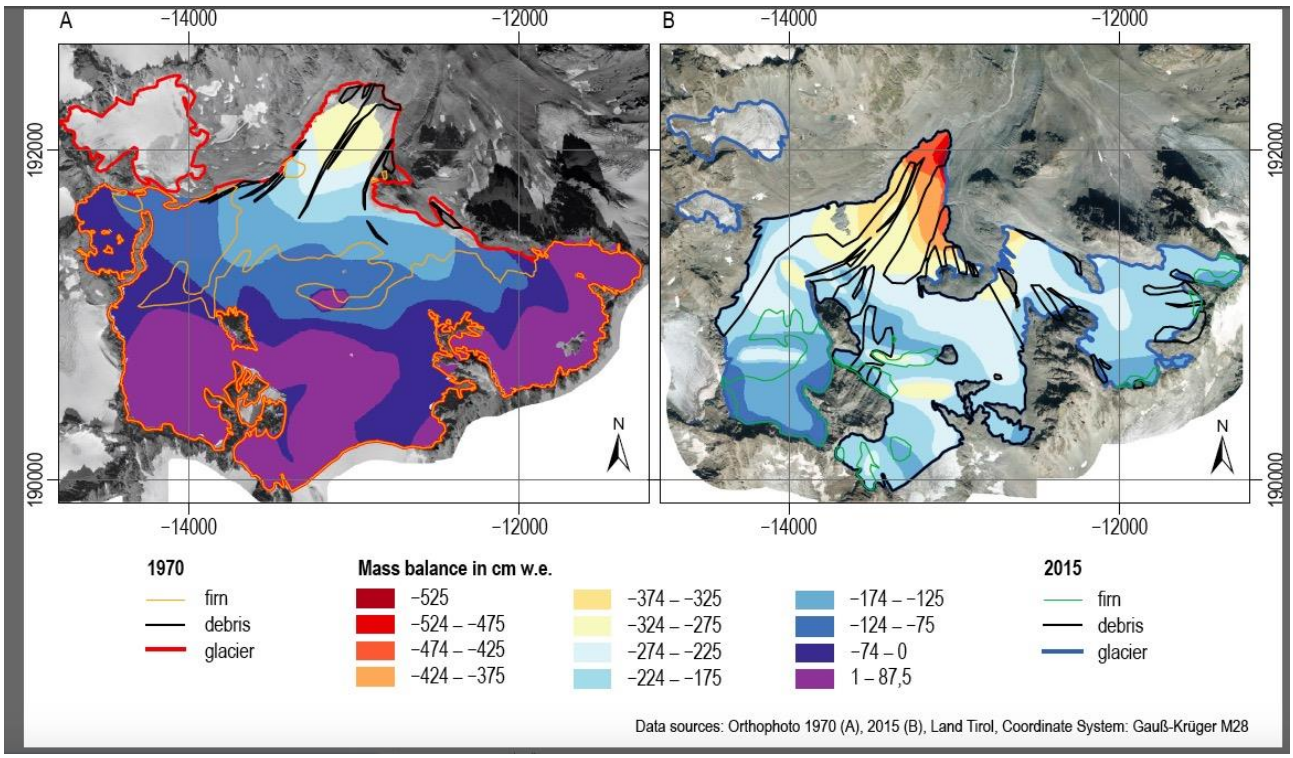


Figure 1: Tongue of Jamtalferner glacier (Orthophoto, August 2015, Source: Tyrolean Government/ TIRIS) with profile lines of spectroadiometer measurements indicated in red. Insert: ~~Areal~~Aerial photograph of Jamtalferner, 20.09.2018 (Photo: Andrea Fischer).



~~Figure~~

2~~: The upper part of the ablation zone of Jamtalferner (top panel) and the terminus (bottom panel) in the early 1990s (Photos: Gerhard Markl), 2005 and 2017 (Photos: Andrea Fischer).~~



~~Figure 3: Outlines indicate summer firn coverage, bare ice areas, and debris cover in 1970 (left) and 2015 (right). Coloured areas show mass balance in cm water equivalent. in 1988/89 (first year of mass balance measurements at Jamtalferner) and the extremely negative mass balance year 2014/15. Background imagery: Orthophotos 1970 (A), 2015 (B), Tyrolean government.~~

.2~~.2 Ground~~ **In situ measurements of spectral reflectance**

The field campaign was carried out on September 4th, 2019. This date was selected for two reasons: Favourable weather conditions and temporal proximity to overpasses of both Sentinel-2 (on the same day) and Landsat-8 (on September 3rd). With a large area of high pressure over western and central Europe, the weather at the study site was sunny and dry on throughout Sept. 3rd and 4th. Using ~~a~~an ASD Field Spec Handheld 2 spectroradiometer~~,~~ (ASD Inc., 2010), a total of

246 reflectance spectra (HCFR) was collected, with 12 spectra measured at point locations and 234 spectra measured along 16 profile lines. ~~14 profile lines contain 11 spectra.~~Profiles were measured along a 20m measuring tape in such a way that individual spectra were gathered at equal intervals~~ between the start and end point of the profile, along a 20m measuring tape.~~, with 14 profile lines containing 11 spectra spaced at 2 m. 2 profiles contain 40 spectra – these were also gathered at equal intervals but with a higher resolution. Measurements began at 08:28 GMT (10:28 local time) and ended at 13:43 GMT. The coordinates of the start and end points of each profile line, as well as any spectra measured outside of the lines, were recorded with a Garmin etrex VISTA HCx, a standard handheld GPS device, which also recorded the time of day. The horizontal accuracy of the GPS coordinates is better than 3 m as per the internal accuracy assessment of the GPS device. The timestamps of the GPS points for the start and end points of the profiles were used to compute solar elevation and azimuth. For each profile, the mean solar elevation and azimuth between the respective start and end points is given in Table 2. Measurements were taken 35 cm above ground from nadir with a bare fibre optic. Test measurements in the field showed high consistency between multiple measurements at the same point, so that we chose to use single measurements at each location rather than average over multiple measurements. The instrument was handheld and not mounted on a stand to minimize shading. This measurement set up is similar to that of previous studies (Naegeli, 2015; Di Mauro et al., 2017) and yields a circular field of view (FOV) with a radius of approximately 7.8 cm for flat ground. The instrument operates between 325 and 1075 nm with an accuracy of ±1 nm and a resolution of <3 nm at 700 nm. ~~For calibration, a~~We used a feature of the instrument that allows the user to save the white reference measurement to the RAM of the built-in computer. HCFR is computed for subsequent target reflectance measurements based on the saved reference. This is saved to the output file, eliminating the need to calibrate the target measurements to the white reference in post-processing. A new SRT-- 99-020 Spectralon (serial number 99AA08-0918-1593) manufactured by Lab Sphere was used~~. Initial processing~~ for the measurement of the ~~raw~~white reference. The ASD data files ~~was carried out~~ were imported into a python script for further analysis using the ~~specdal~~ Python ~~package.~~module SpecDal (Lee, 2017) to read the ASD format. Further data analysis was carried out using numerous other Python (Van Rossum and Drake., 2009) packages, mainly NumPy (van der Walt et al., 2011), pandas (McKinney, 2010), Matplotlib (Hunter, 2007), Rasterio (Gillies et al., 2013), GeoPandas (GeoPandas developers, 2019), rasterstats (Perry, 2015), and PyEphem (Rhodes, 2020).

2.3 Satellite data

We compare the ~~ground~~in situ measurements with surface reflectance products derived from a Landsat-~~-~~8 Operational Land Imager (OLI) scene acquired on September 3rd, 2019 (10:10~~:32Z~~ GMT), the day before the ~~ground measurements~~field campaign, and a Sentinel 2A scene acquired on September 4th~~,~~ (10:20 GMT), the same day as the ~~ground measurements.~~field campaign. Both scenes are cloud free over the study area~~.~~ (Figure 2). Details on the atmospheric correction algorithm used to generate the Landsat-~~-~~8 OLI level-2 surface reflectance data product from top of atmosphere (TOA) reflectance can be found in Vermote et al. (2016~~).~~) and in the product guide of the algorithm used to derive surface reflectance (USGS, 2020). Details on the equivalent Sentinel-~~-~~2 product – the Level-2A bottom of atmosphere reflectance – are given in Main-Knorn et al. (2017) and Richter and Schläpfer (2011). For the sake of readability, we refer to the Landsat-~~-~~8 OLI level-2 surface reflectance as "Landsat" data in the following, and to the Sentinel~~-2A~~-2 level-2A surface reflectance as "Sentinel" data. The Landsat and Sentinel surface reflectance raster data used in this study were acquired using ~~google earth engine~~Google Earth Engine (Gorelick et al., 2017).

The wavelength range of the ~~spectral reflectance~~spectroradiometric measurements carried out on the ground overlaps with bands 1-5 of the Landsat data and bands 1-9 and 8A of the Sentinel data, respectively. Only spectral ranges covered by these bands are considered for this study. The wavelengths and resolution of the individual bands, as well as the relevant viewing and solar angles are given in Table 3. For each ground measurement point, band values were extracted from the satellite scenes at the overlaying pixel.

In order to compare the satellite values with ground data, we compute mean values for the subsets of the spectral reflectance curves measured on the ground that correspond to the Landsat and Sentinel bands, respectively. Data are then grouped into profile lines and/or different bands, the Pearson correlation is computed for ground- and corresponding satellite data, and further comparisons are carried out using standard statistical metrics.

To assess the influence of the spatial resolution of the satellite data on results, band 3 imagery was resampled (cubic interpolation) from the original 10 m resolution to 30 m and 60 m for Sentinel and from 30 m to 60 m for Landsat, respectively. To account for the potential effects of the uncertainty in the GPS coordinates, we created a circular buffer with a radius of 3m around each in situ measurement point. For each buffer, the corresponding satellite value is

computed as the median of the values of all pixels the buffer overlaps with.

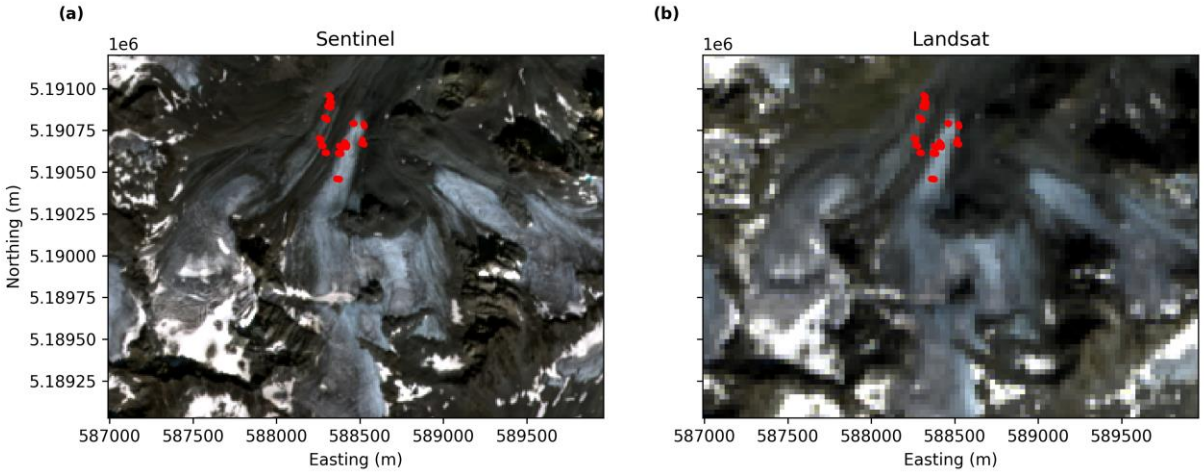
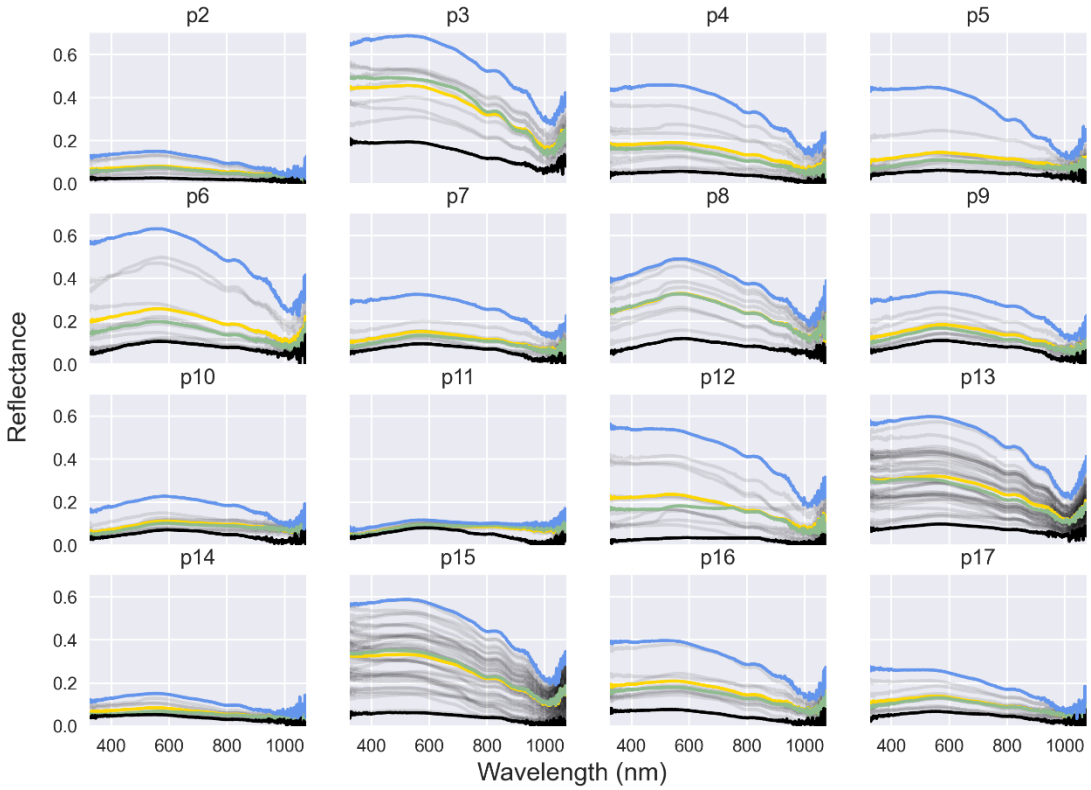


Figure 2: Jamtalferner as seen in the Sentinel (a) and Landsat (b) scenes used in this study. The images shown here are composites of bands 2, 3, and 4 of each satellite's L2A surface reflectance product displayed at a resolution of 10 (Sentinel) and 30 (Landsat) m/pixel, respectively. Profiles where reflectance spectra were collected are marked in red. Coordinate reference system: EPSG: 32632.

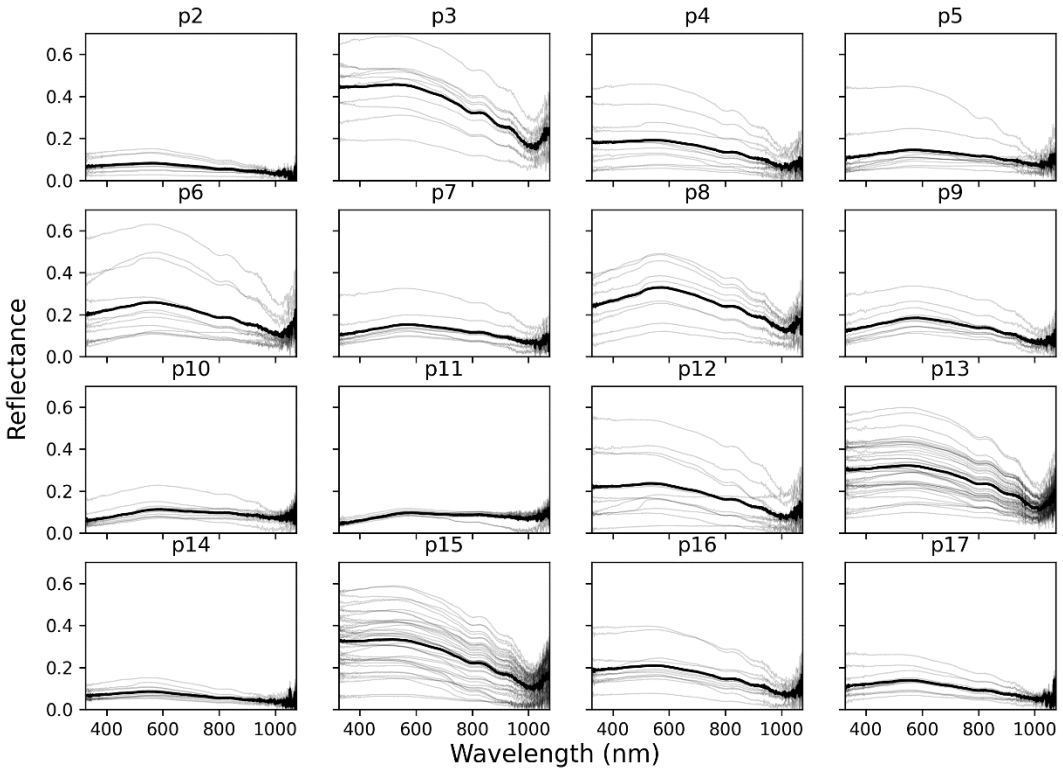
3. Results

3.1 Surface measurements

The in situ measurements exhibit extreme differences in ~~surface albedo~~HCFR depending on the characteristics of the surface. Figure 43 shows the spectra grouped into profiles, ~~as well as~~with the mean~~, median, maximum, and minimum~~ spectral ~~reflectance per~~HCFR highlighted for each profile. P3 is the "brightest" profile, with the highest maximum (up to 0.7) and minimum (up to 0.2) values of all profiles. Profiles 2, 11, and 14 are the darkest profiles and all of their respective spectral reflectance remain below 0.2 at all measured wavelengths. Figure ~~5~~4 shows the ice surface along profile lines 3 (brightest) and 11 (darkest) for a visual comparison. In P3, the surface is mainly comprised of clean, dry ice. In P11, the ice surface is wet and impurities (rocks, fine grained debris) are present. The profile line crosses several small melt water channels with running water.


p2
p3
p4
p5
p6
p7
p8
p9
p10
p11
p12
p13
p14
p15
p16
p17
Reflectance
Wavelength (nm)
0.0
0.2
0.4
0.6
400
600
800
1000

278
279


p2
p3
p4
p5
p6
p7
p8
p9
p10
p11
p12
p13
p14
p15
p16
p17
Reflectance
Wavelength (nm)
0.0
0.2
0.4
0.6
400
600
800
1000

280

Figure 43: Each subplot shows the spectra along a profile line. ~~The~~The bold black lines highlight the mean~~, median, maximum, and minimum~~ spectral reflectance (HCRF) in each profile ~~are highlighted in yellow, green, blue, black, respectively~~.

Figure ~~5~~4: Photos of the ice surface along profile 3 (left) and profile 11 (right), at the time of measurement (~~Photo~~Photos: Andrea Fischer)

Table 2 contains a qualitative description of the ice surface along each profile line, the length of the line, the number of spectra per line, and the number of Landsat and Sentinel band 3 pixels that each line crosses~~.~~, as well as the mean solar elevation and azimuth angles for the profile. The maximum number of pixels per line is 5 for Sentinel and 3 for Landsat, respectively. All lines cross at least 2 pixels for Sentinel, while 3 lines fall into a single Landsat pixel. See Fig. 1 for the location of each profile on the glacier.

Table 2: Description of the surface characteristics along each profile line, as well as number of spectra collected along the line and number of pixels intersected by the line in band 3 of the Sentinel and Landsat scenes, respectively.

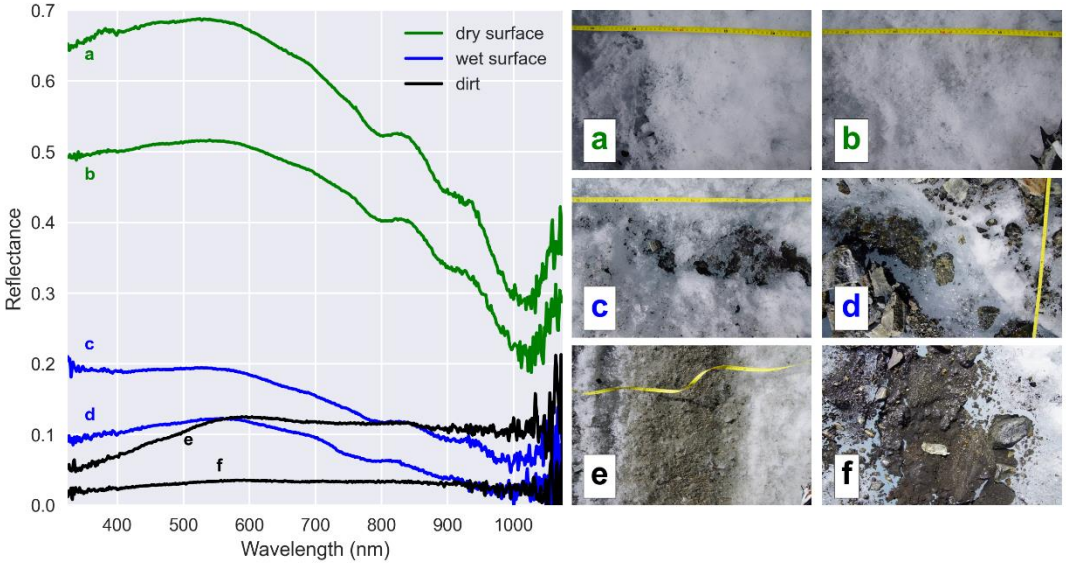
	Qualitative description	Spectra	Sentinel B3 pixels	Landsat B3 pixels
P2	Relatively smooth, uniform ice surface, slightly wet.	11	3	2
P3	Mostly dry surface, clean cryoconite.	11	4	1
P4	Mostly dry ice surface, some dirt, some rocks/debris on ice surface where profile approaches moraine.	11	4	2
P5	Significant debris cover along profile. Where ice is exposed, ice surface is wet. Profile crosses meltwater channels with running water.	11	3	1
P6	Wet ice surface with dust/dirt transitions to cleaner, brighter ice.	11	4	1
P7	Grey-ish ice surface with meltwater channels and fine grained debris/small rocks.	11	2	2
P8	Similar to P7, fewer rocks.	11	4	2
P9	Wet ice surface with mixture of relatively clean cryoconite and more dusty areas.	11	3	2
P10	Wet ice surface with several small melt water channels. Mostly dirty, grey ice.	11	3	2
P11	Wet ice surface with several small meltwater channels. Very dirty ice with scatteredsmall rocks.	11	4	2
P12	Relatively clean, bright ice interspersed with larger meltwater ponds/channels, which contain dirt and small rocks.	11	4	3
P13	Clean cryoconite with some darker patches.	40	5	2
P14	Wet ice surface with fine grained dirt in relatively uniform cryoconite.	11	4	2
P15	Uneven ice surface, mostly clean, dry ice.	40	3	2
P16	Mixture of wet and dry ice surface and fine grained dirt.	11	3	2
P17	Mostly wet ice surface, fine grained dirt with some cleaner patches.	11	2	2

Profile Nr.	Qualitative description	Mean solar elevation, azimuth in degrees	Spectra	Sentinel B3 pixels	Landsat B3 pixels
P2	Relatively smooth, uniform ice surface, slightly wet.	24.69, 106.70	11	3	2
P3	Mostly dry surface, clean cryoconite.	26.43, 108.92	11	4	1
P4	Mostly dry ice surface, some dirt, some rocks/debris on ice surface where profile approaches moraine.	28.64, 111.87	11	4	2
P5	Significant debris cover along profile. Where ice is exposed, ice surface is wet. Profile crosses meltwater channels with running water.	31.34, 115.72	11	3	1
P6	Wet ice surface with dust/dirt transitions to cleaner, brighter ice.	34.45, 120.61	11	4	1
P7	Grey-ish ice surface with meltwater channels and fine-grained debris/small rocks.	36.20, 123.57	11	2	2
P8	Similar to P7, fewer rocks.	38.05, 126.99	11	4	2
P9	Wet ice surface with mixture of relatively clean cryoconite and more dusty areas.	39.40, 129.68	11	3	2
P10	Wet ice surface with several small melt water channels. Mostly dirty, grey ice.	40.71, 132.51	11	3	2
P11	Wet ice surface with several small meltwater channels. Very dirty ice with scattered small rocks.	42.08, 135.75	11	4	2
P12	Relatively clean, bright ice interspersed with larger meltwater ponds/channels, which contain dirt and small rocks.	47.61, 153.83	11	4	3
P13	Clean cryoconite with some darker patches.	48.63, 159.14	40	5	2
P14	Wet ice surface with fine grained dirt in relatively uniform cryoconite.	49.80, 168.30	11	4	2
P15	Uneven ice surface, mostly clean, dry ice.	50.30, 179.29	40	3	2
P16	Mixture of wet and dry ice surface and fine-grained dirt.	49.43, 194.99	11	3	2
P17	Mostly wet ice surface, fine grained dirt with some cleaner patches.	48.33, 202.34	11	2	2

The spectral reflectance curves of the individual spectra as well as of the profile lines indicate high spatial variation of surface types and associated reflective properties. The spectral signatures of the individual spectra can roughly be grouped into dry ice, wet ice, and dirt/rocks. (We use the word “dirt” to describe all types of mineral or organic materials and fine-grained debris that may collect on the glacier surface.) However, transitions between these types are ~~fluid~~gradational and in practice these categories cannot always be clearly separated - both dry and wet ice might be clean or dirty, dirt might be wet or dry.

The reflectance curves for clean ice exhibit the typical shape frequently found in literature (Zeng et al., 1984), with highest reflectance values (up to 0.69) in the lower third of our wavelength range and declining values for wavelengths

greater than approximately 580 nm. The spectral reflectance curves of wet ice surfaces follow roughly the same shape as for dry ice but are strongly dampened in amplitude with reflectance values typically not exceeding 0.2. In contrast, the reflectance curve of dirty surfaces remains at uniformly low values throughout our wavelength range in some cases and exhibits an increase between 325 and approximately 550 nm before flattening out in other cases. Reflectance values have similar magnitudes as for wet ice. Example reflectance curves of these surface types are given in Fig. ~~6~~5.



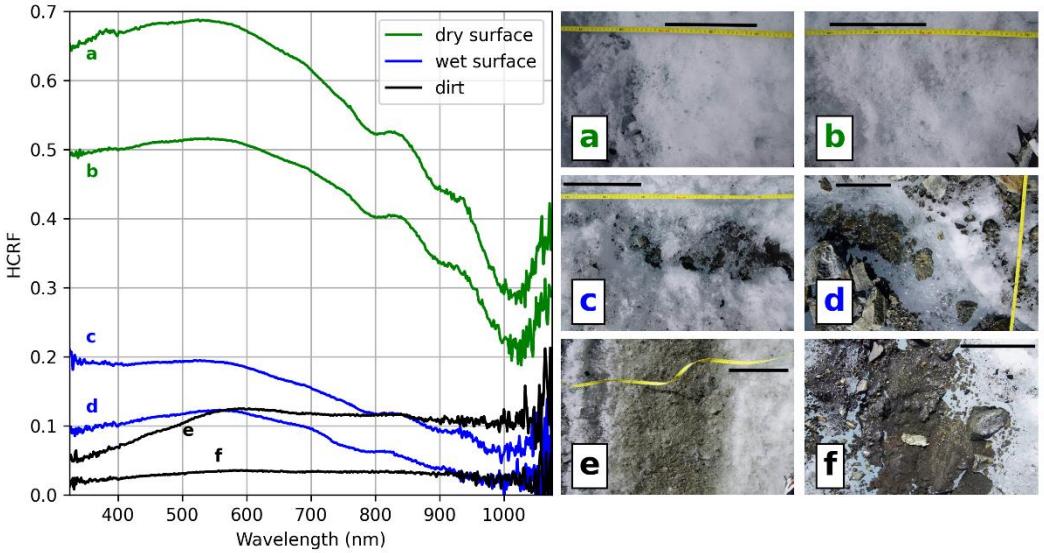
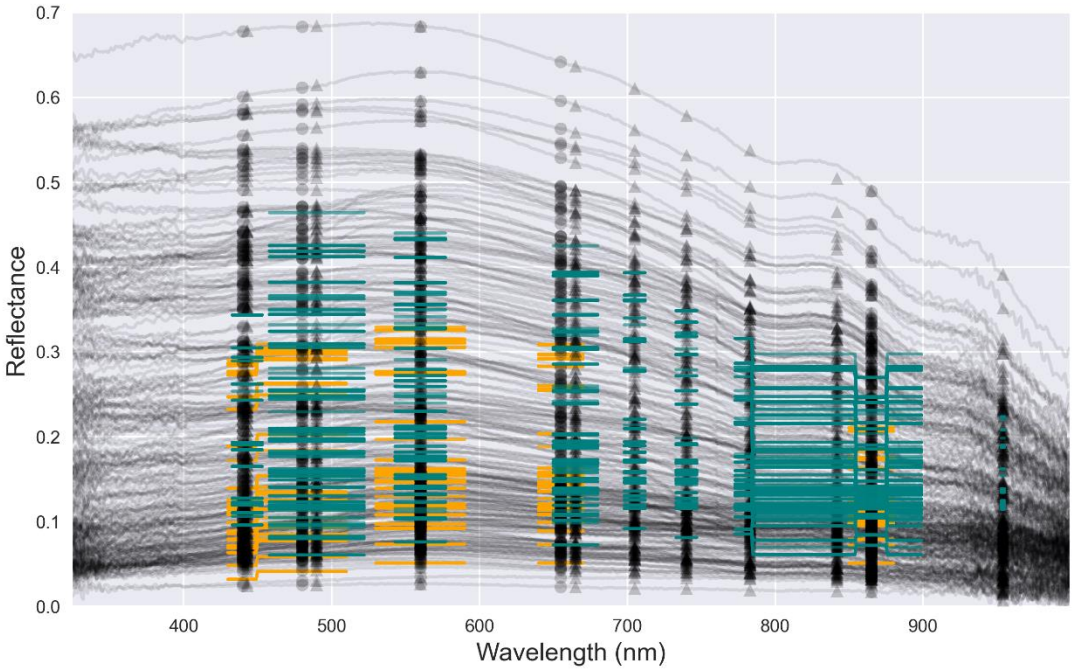


Figure ~~6~~5: Spectra of different kinds of ice surface types encountered in the ablation zone of Jamtalferner. The photos on the right show the ice surface at the sampling sites of the respective spectra. The black bar in each photo represents approximately 20 cm, to provide a sense of scale. The spectra shown in this figure are part of the following profile lines: a, b, c – p3; d – p4; e – p6; f – p12.

3.2 Comparison with satellite data

Figure ~~7~~6 shows all measured spectral reflectance curves, as well as the Sentinel and Landsat values in the bands that overlap the wavelength range of the ground measurements. ~~The satellite~~ Reflectance values ~~in the figure are the values~~were extracted from the ~~Sentinel and Landsat pixels, respectively, for~~satellite imagery at the coordinates of each

~~ground measurement~~sampling point and overlayed onto the plots of the in situ spectra as coloured bars. Naturally, neither satellite captures the full range of reflectance values measured on the ground. In all overlapping bands of Sentinel and Landsat, the Sentinel values are higher, in the sense that the maximum values of the Sentinel data are closer to the maximum values measured on the ground, while the minimum Landsat data are closer to the minimum values measured on the ground.



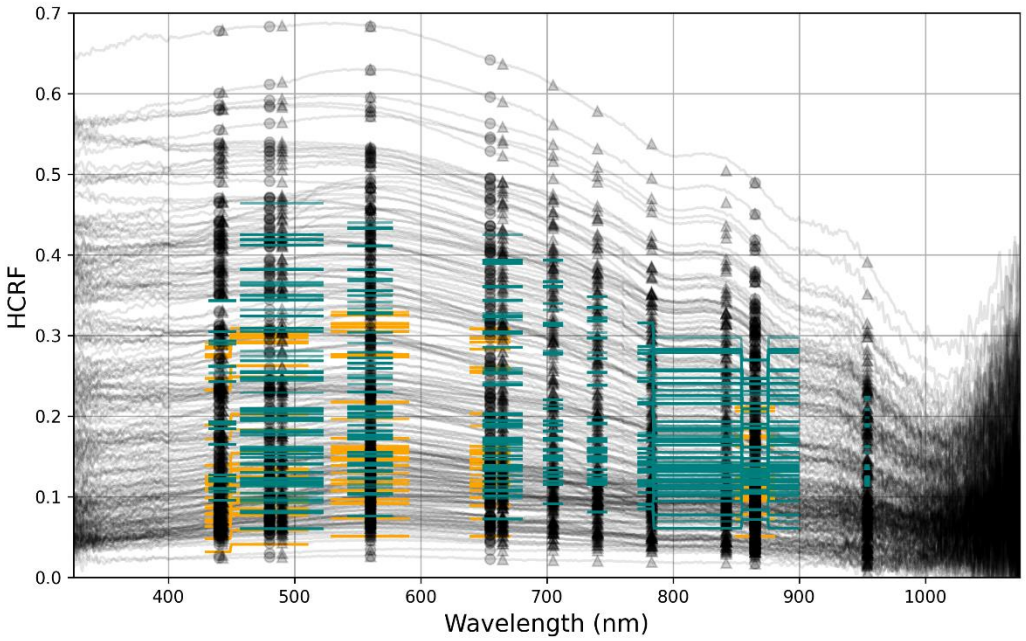


Figure ~~7: Spectra~~6: The spectra measured ~~on the ground~~in situ are plotted in black. ~~Orange and blue lines represent the mean reflectance of all ground measurements in the wavelength range of the Landsat and Sentinel bands listed in Table~~

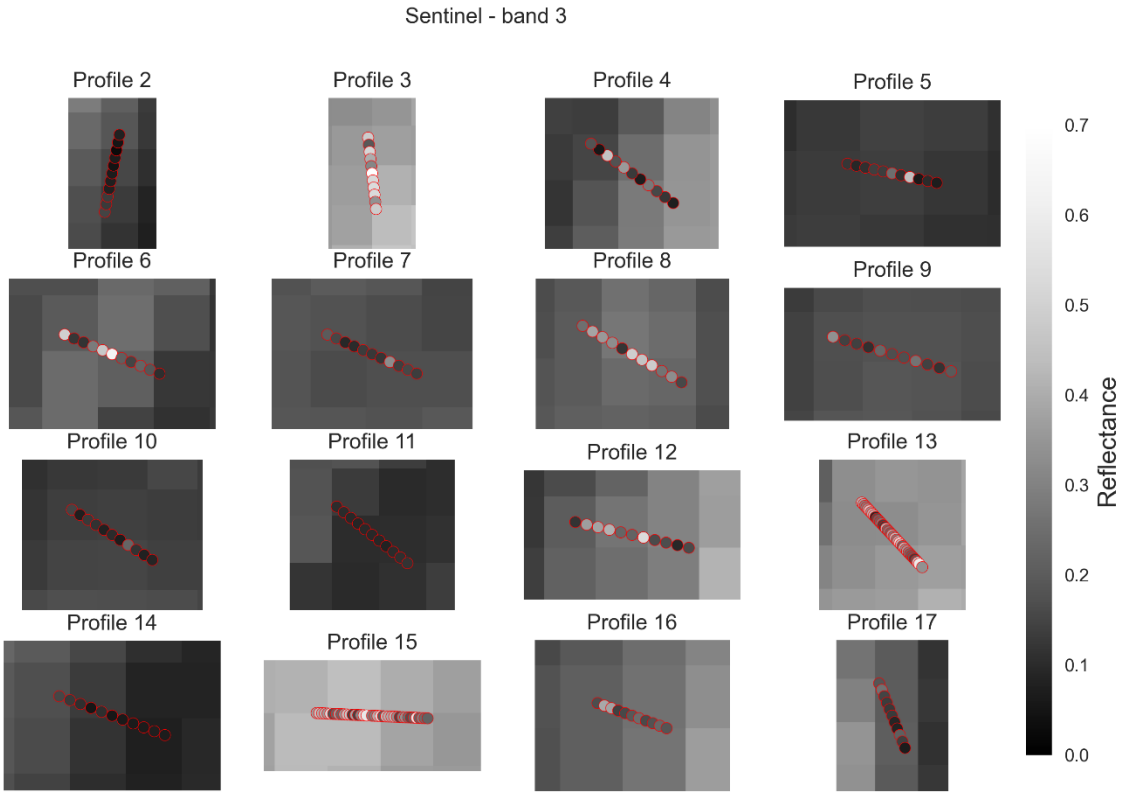
~~3.~~ Black circles indicate the central wavelengths of the Landsat bands, black triangles those of the Sentinel bands (see Table 3). Orange and blue lines represent the wavelength range of the respective Landsat and Sentinel bands along the horizontal axis and the satellite derived reflectance at the sampling points of each spectrum on the vertical axis.

Comparing the mean of the ~~spectral reflectances~~HCRF spectra measured on the ground for each satellite band with the associated satellite values yields a Pearson correlation coefficient ranging from 0.53 (band 5) to 0.62 (band 1) for the Landsat bands and 0.3 (band 9) to 0.65 (band 2) for Sentinel. Table 3 lists the correlation coefficients, as well as the wavelength range and resolution of each band. The two lower resolution Sentinel bands (band 1, band 9 – 60m resolution) have notably lower correlation coefficients than the higher resolution bands. The Sentinel and Landsat data at the ~~ground~~in situ measurement points are strongly correlated with each other in the bands where both satellites overlap, with r=0.69 in band 1 and r>0.8 for bands 2, 3, 4, and 5.

For a visual comparison of the location of the profile lines and the range of measured values in the profiles in relation to the satellite pixel boundaries and pixel band values, see Fig. ~~8~~7 for Sentinel (band 3 selected as an example) and supplementary material for an analogous figure of the Landsat data~~.~~).

Table 3: Band names and respective wavelength range and resolution for Landsat and Sentinel as used in this study. Pearson correlation given for mean band values of ground-measurements and associated satellite data. For Landsat, the solar zenith and azimuth angles given in the surface reflectance image are listed. The view zenith angle is hardcoded to 0 in the Land Surface Reflectance Code (LaSRC_1.3.0) for the Landsat surface reflectance product, as per the LaSRC documentation (USGS, 2020). For Sentinel, the incidence angles refer to the mean viewing zenith and azimuth angles for each band. The solar angles are the averages for all bands.

Band	Range (nm)	Resolution (m)	Pearson Corr.
	Landsat		
1 (Coastal/Aerosol)	430-450	30	0.62
2 (Blue)	450-510	30	0.61
3 (Green)	530-590	30	0.58
4 (Red)	640-670	30	0.57
5 (NIR)	850-880	30	0.53
	Sentinel		
1(Coastal/Aerosol)	433-453	60	0.46
2 (Blue)	457.5-522.5	10	0.65
3 (Green)	542.5-577-5	10	0.63
4 (Red)	650-680	10	0.61
5 (Vegetation Red Edge)	697.5-712.5	20	0.57
6 (Vegetation Red Edge)	732.5-747.5	20	0.56
7 (Vegetation Red Edge)	773-793	20	0.55
8 (NIR)	784.5-899.5	10	0.56
8A (NIR narrow band)	855-875	20	0.53
9 (Water vapour)	953-955	60	0.3



371
372

	Landsat	Sensing time: 2019-09-03 10:10 GMT					
Band	**Range (nm)**	**Resolution (m)**	**Pearson Corr.**	**View zenith angle**	**View azimuth angle**	**Solar zenith angle**	**Solar azimuth angle**
1 (Coastal/Aerosol)	430-450	30	0.62	0	-	42.63	153.57
2 (Blue)	450-510	30	0.61				
3 (Green)	530-590	30	0.58				
4 (Red)	640-670	30	0.57				
5 (NIR)	850-880	30	0.53				
	Sentinel	Sensing time:2019-09-04 10:20 GMT		**Mean incidence zenith angle**	**Mean incidence azimuth angle**	**Mean solar zenith angle**	**Mean solar azimuth angle**
1 (Coastal/Aerosol)	433-453	60	0.46	3.13	193.02	40.83	159.93
2 (Blue)	457.5-522.5	10	0.65	2.48	198.51		
3 (Green)	542.5-577.5	10	0.63	2.59	196.22		
4 (Red)	650-680	10	0.61	2.72	194.92		
5 (Vegetation Red Edge)	697.5-712.5	20	0.57	2.79	194.43		

6 (Vegetation Red Edge)	732.5-747.5	20	0.56	2.87	193.84		
7 (Vegetation Red Edge)	773-793	20	0.55	2.95	193.53		
8 (NIR)	784.5-899.5	10	0.56	2.54	197.22		
8A (NIR narrow band)	855-875	20	0.53	3.04	193.30		
9 (Water vapour)	953-955	60	0.3	3.22	192.89		

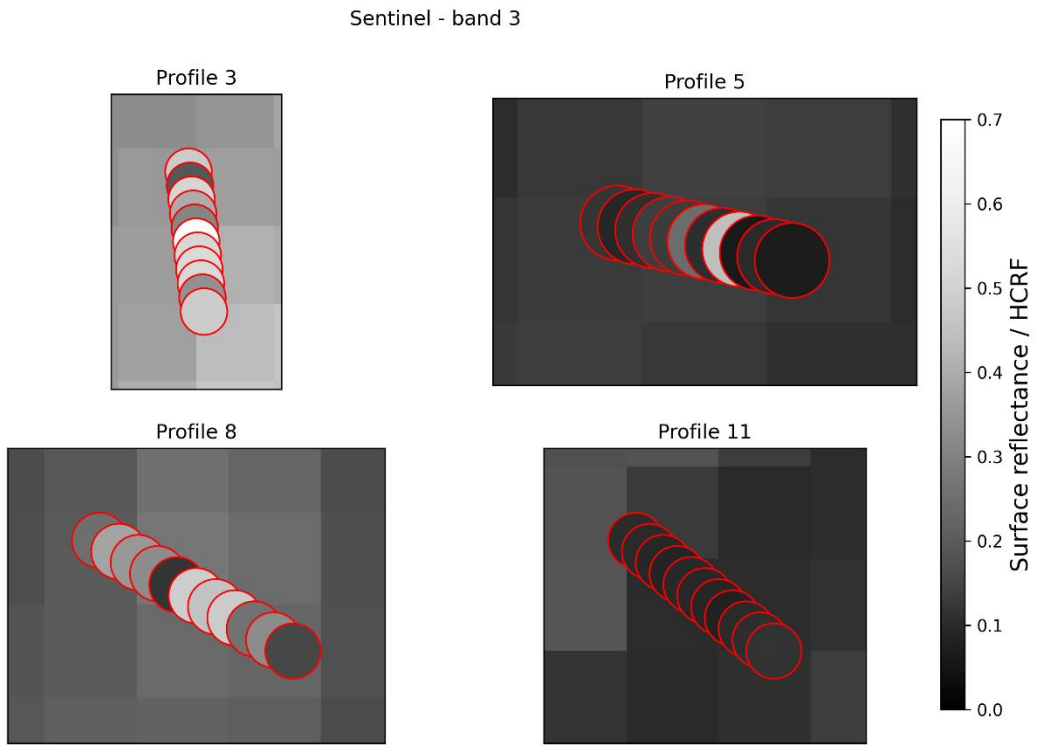
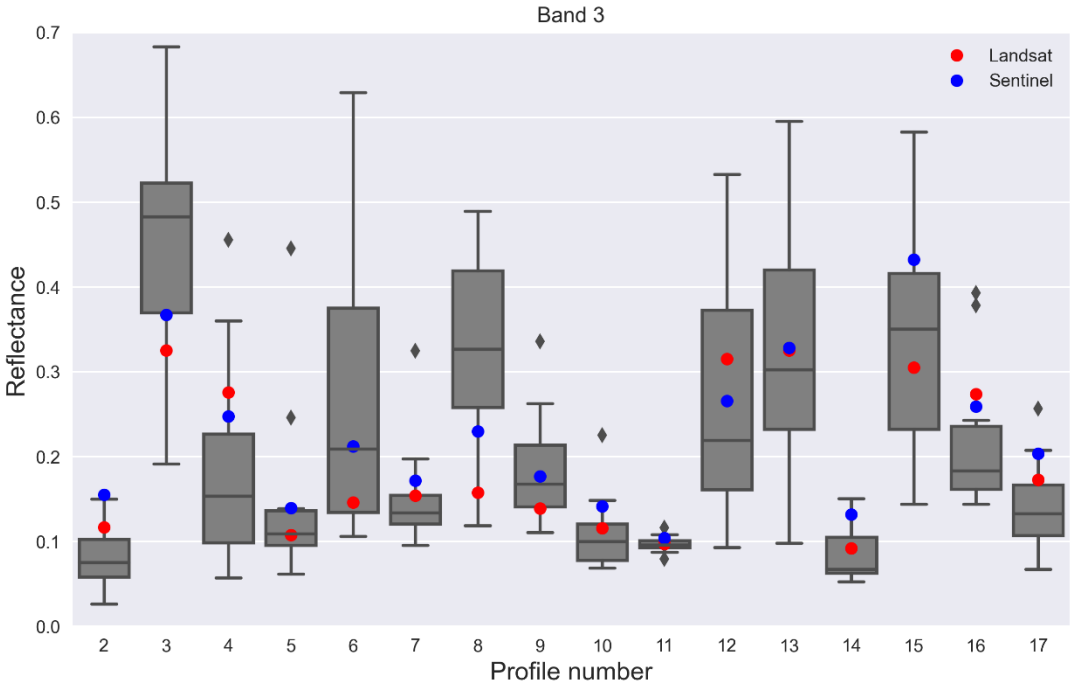


Figure ~~8~~7: The spectra comprising the profile lines are plotted over the corresponding satellite pixels for selected profiles. The colour bar is the same for the background raster and the circles indicating the sampling sites of the spectra and represents the Sentinel band 3 pixel value and the mean reflectance in the Sentinel band 3 wavelength range of each spectrum, respectively. The pixel size of the raster is 10m². The GPS coordinates of the sampling sites are centred in the circles. The circle radius is set to 3m to represent the horizontal uncertainty of the GPS points.

The spread of ~~ground reflectance~~in situ HCRF values per profile is generally lower for profiles that are darker overall, and greater for brighter profiles, although not in all cases (Fig.~~4~~3, Fig. ~~9~~8). In the Sentinel band 3 wavelength range, profile 3 is brightest with a median reflectance of 0.48 and spread of 0.49. Profile 6 (median ~~reflectance~~ in Sentinel band 3 range: 0.21) has the largest spread of ~~reflectances~~HCRF (0.52). Broadly speaking, profiles with a high median ~~reflectance~~HCRF tend to include individual measurement points that are both very bright and very dark, while darker profiles are more uniformly dark. Profile 6 in particular transitions between surface types and contains wet/dirty spectra as well as dry ice spectra (see Table 2). Figure ~~9~~8 shows boxplots of the ground measurements (band 3 mean) for all profiles to exemplify this and indicates where the Landsat and Sentinel values fall compared to the spread of values in each profile.



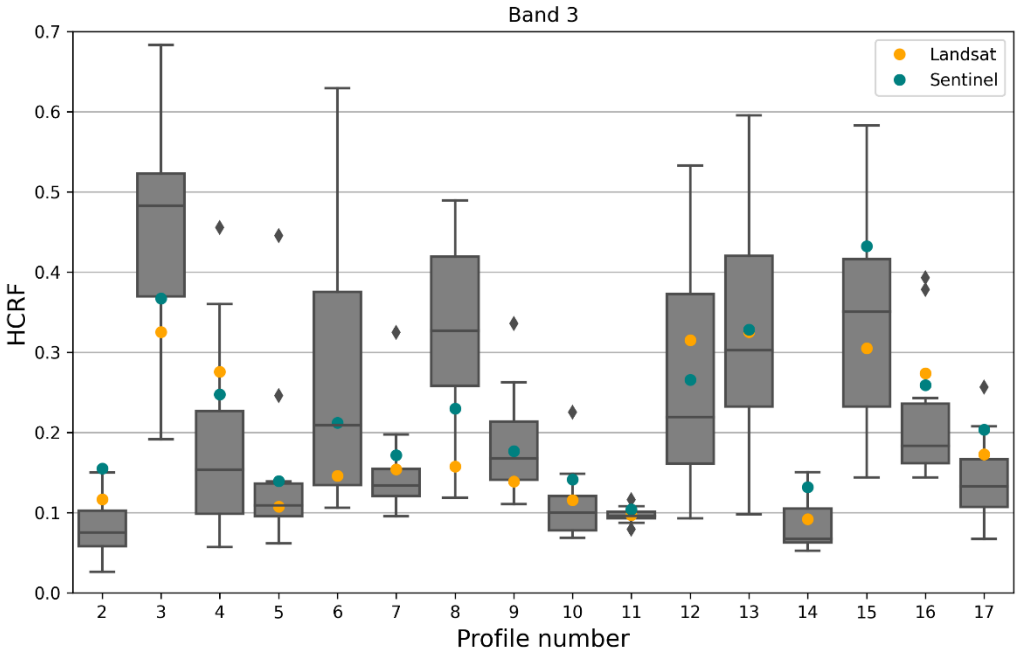


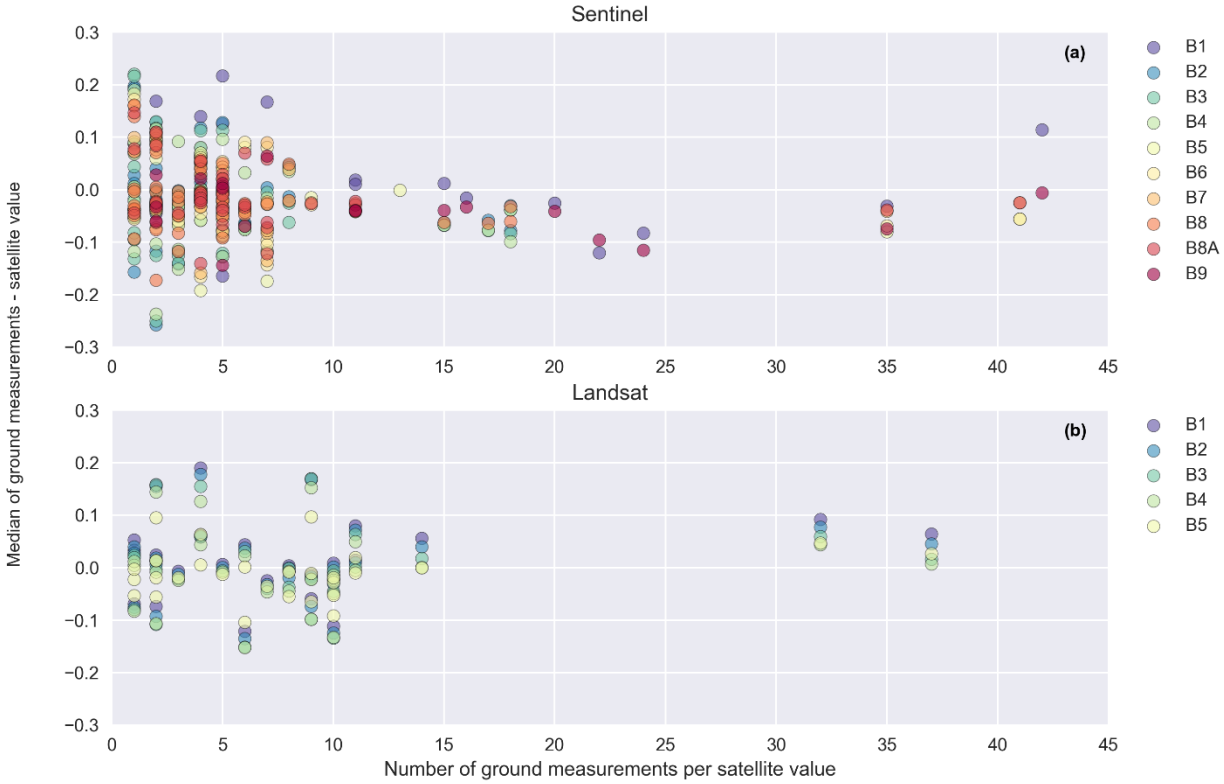
Figure ~~9~~8: Spread of the ~~band 3 (~~Sentinel band 3 (wavelength range: 542.5-577.5 nm) mean values of the measured spectra, grouped by profile. ~~Red~~Orange and blue circles show corresponding mean pixel values of data extracted from Landsat and Sentinel pixels at the sampling sites of the spectra, respectively.

When binning ~~ground~~in situ measurements by the associated satellite value/pixel and taking the median or mean of the binned values, the difference between the median/mean ~~ground~~in situ value and the satellite value tends to decrease with increasing number of ~~ground~~in situ measurements mapped to unique satellite values. This is to be expected, as each satellite value represents an integration of the emission characteristics over the area contained in the pixel.

However, for our data, this relationship is not obviously linear and differs between Sentinel and Landsat, as well as between different bands (Figure ~~10~~9).

Comparing ~~ground~~in situ and satellite values for individual ~~ground~~in situ measurement points, it is apparent that both satellites tend to overestimate the reflectance values of dark ground surfaces, and underestimate the reflectance of bright surfaces, in all bands (Figure ~~11~~10). The shift from over- to underestimation appears linear and has a similar increase rate in all bands. The zero crossings of the regression lines, i.e. the ground reflectance values for which ground measurements and satellite values match, fall between 0.15 (band 5) and 0.21 (band 1) for Landsat and 0.17 (band 9) and 0.27 (band 3) for Sentinel.

Figure ~~12~~11 shows histograms of the mean reflectance in band 3 of Landsat and Sentinel, respectively, compared with associated in situ values, as well as density plots ~~over~~of the satellite derived surface reflectance over all pixels in the study area. The mean is highest in the in situ measurements and lowest in Landsat images. Both Sentinel and Landsat fail to capture ~~reflectances~~HCRF values below 0.05 and above 0.45. A second peak in frequency evident from the in situ measurements at a reflectance of 0.4 is not represented in the remote sensing data.



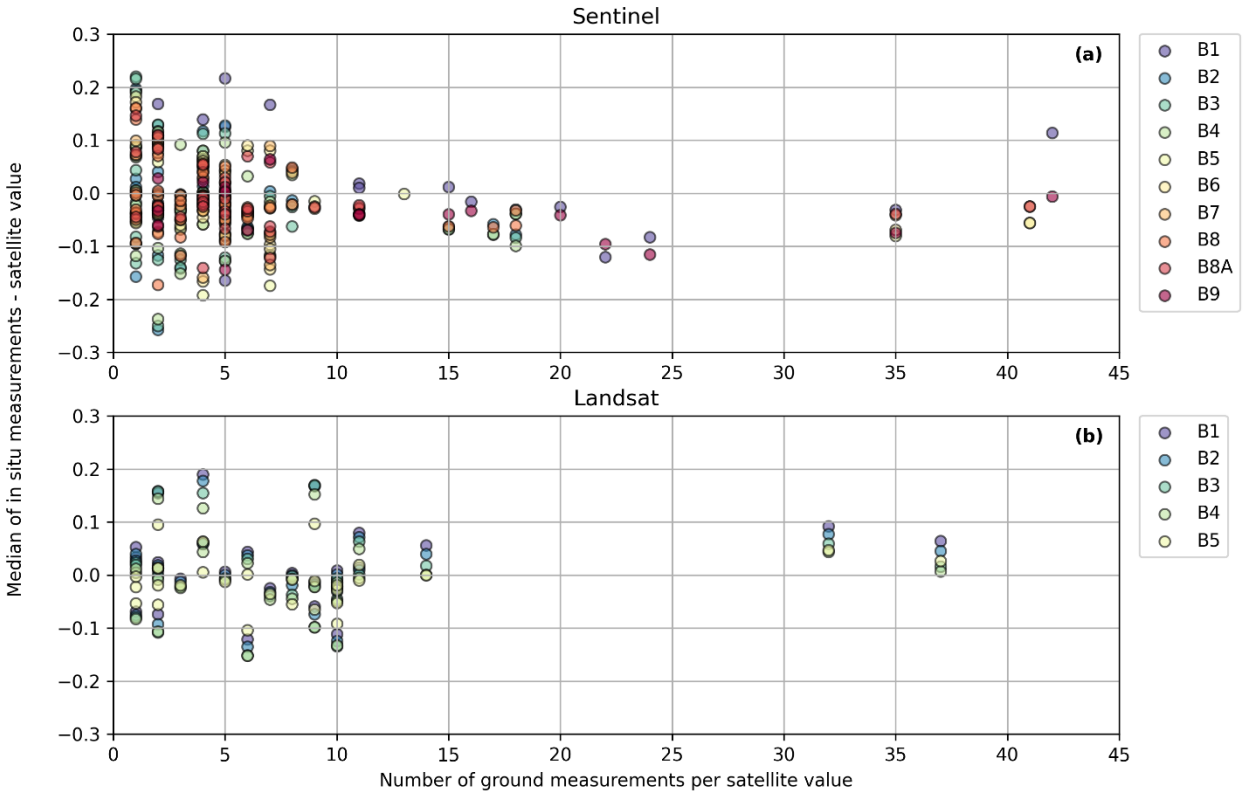
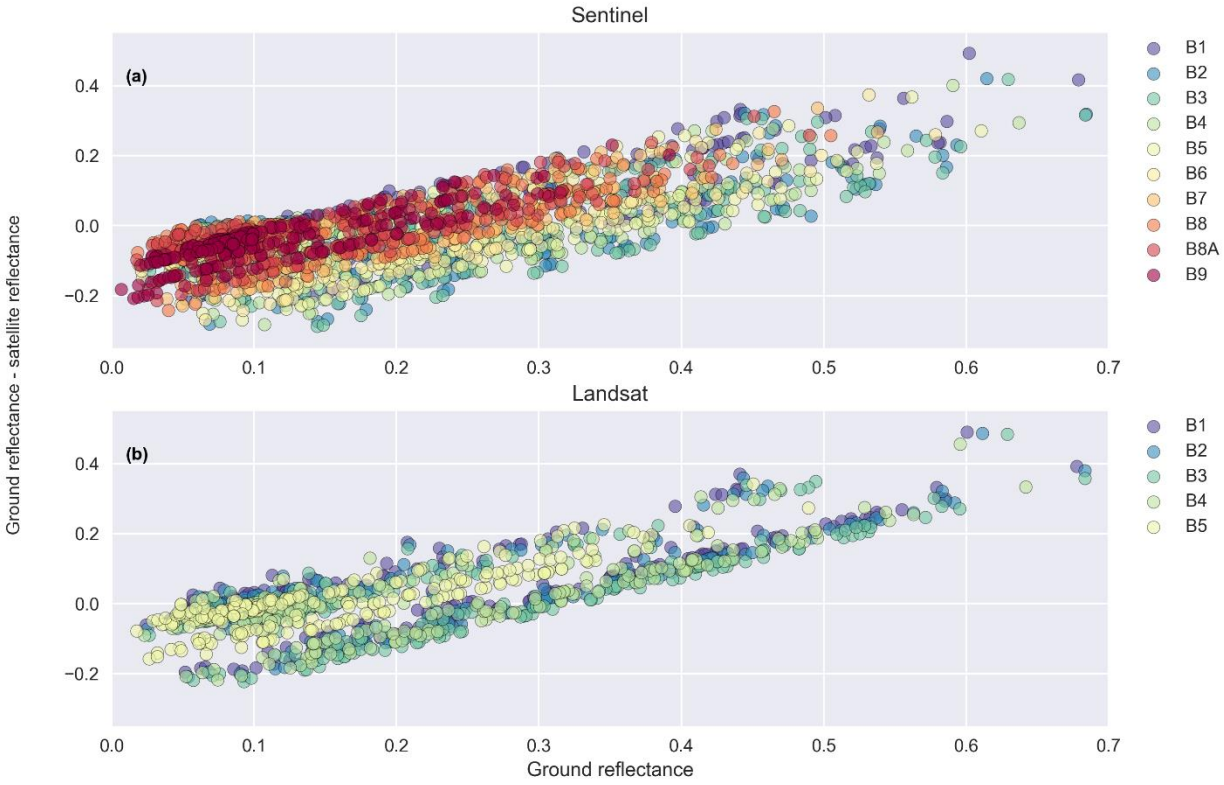
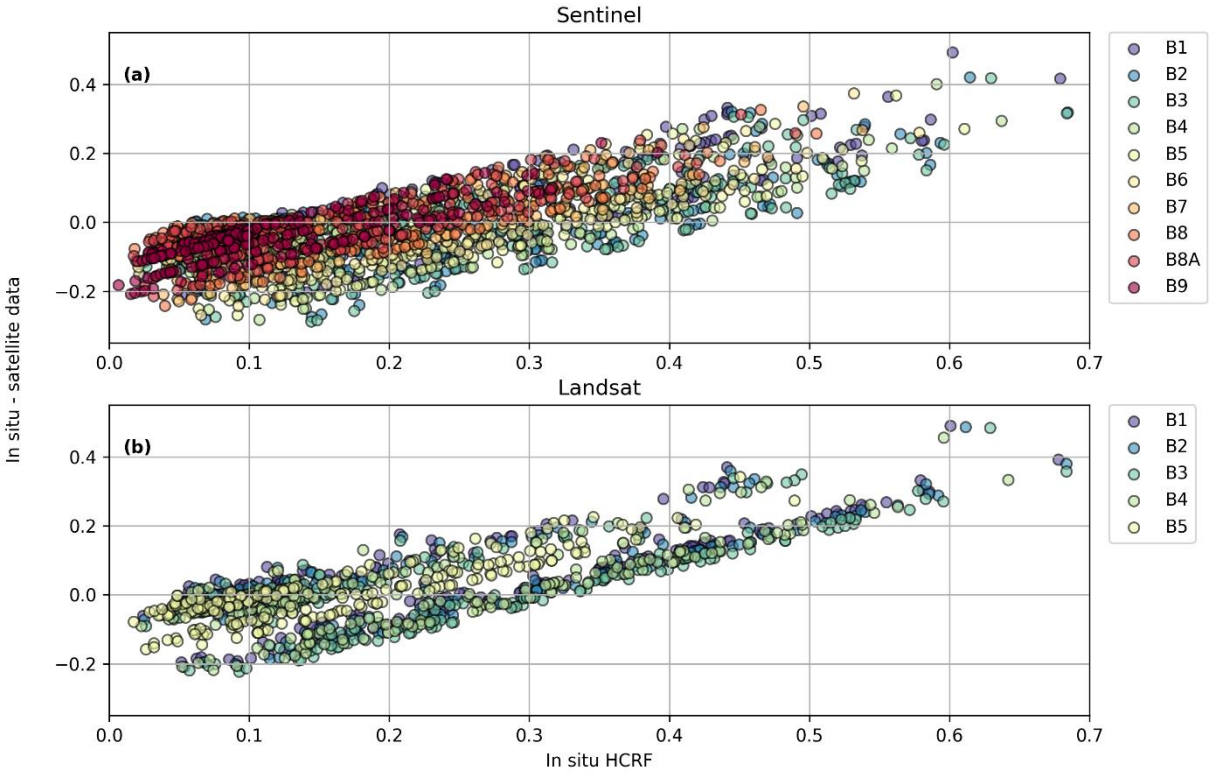


Figure ~~10~~9: The number of ground measurements per unique satellite value (x-axis) is plotted against the difference between the median of these ground measurements in the respective wavelength band and the corresponding satellite value (y-axis). i.e. values that are positive in the vertical axis represent cases where ground reflectance is higher than satellite derived reflectance, whereas negative values represent the opposite. Different colours represent the different satellite bands, as indicated by the legend next to the plots.

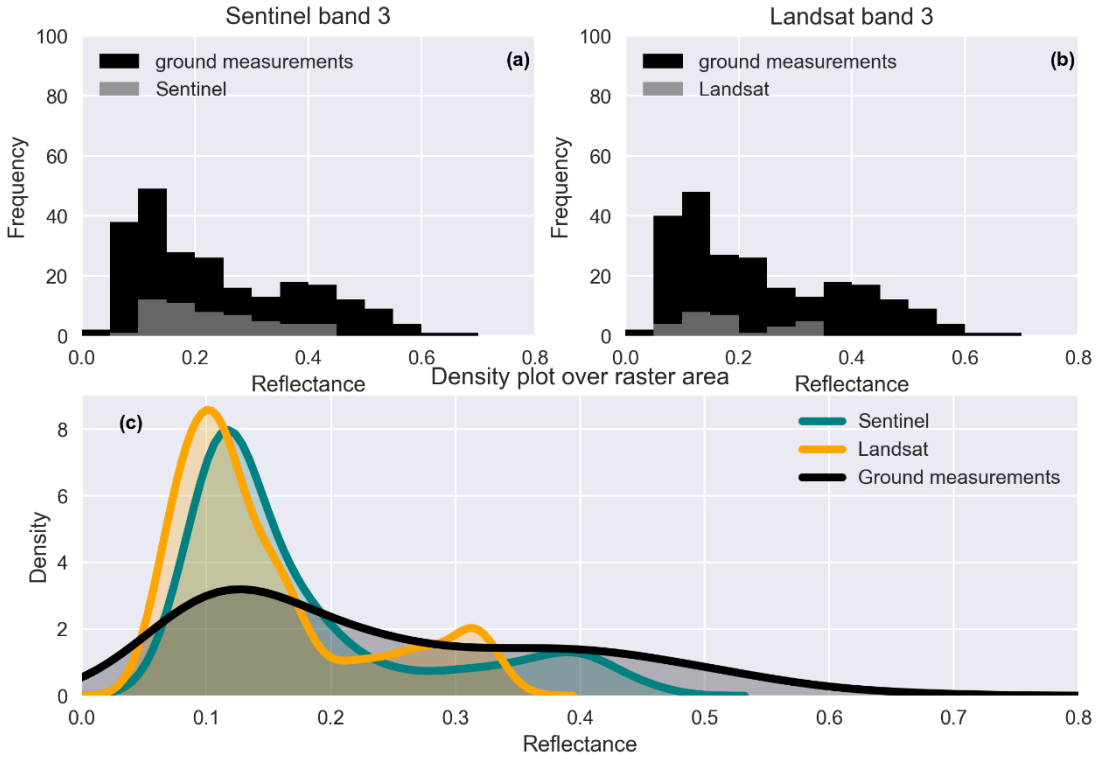

Sentinel
(a)
B1
B2
B3
B4
B5
B6
B7
B8
B8A
B9
Landsat
(b)
B1
B2
B3
B4
B5
Ground reflectance - satellite reflectance
Ground reflectance
0.4
0.2
0.0
−0.2
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7

433
434


Sentinel
(a)
B1
B2
B3
B4
B5
B6
B7
B8
B8A
B9
Landsat
(b)
B1
B2
B3
B4
B5
In situ - satellite data
In situ HCRF
0.4
0.2
0.0
−0.2
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7

435

Figure ~~11~~10: Same data as in Fig. 10, but showing individual sampling points without grouping by common satellite pixels.



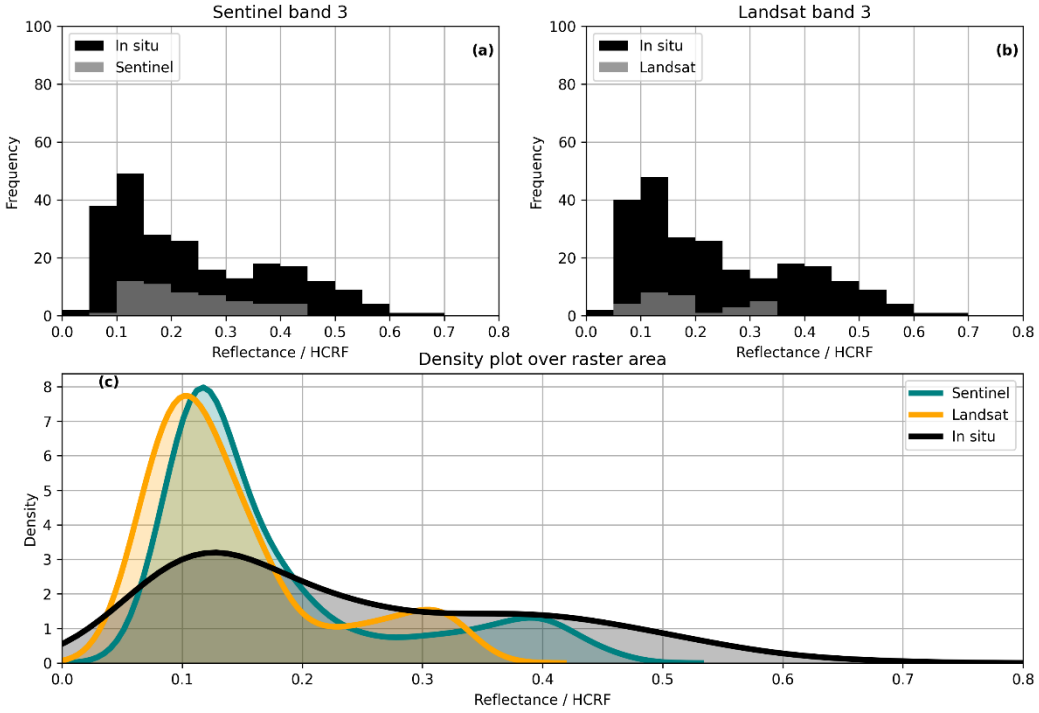


Figure ~~12~~11: The histograms in the top panels (a, b) show the frequency of occurrence of the band 3 mean values of the ground measurements per reflectance bin. Bin width: 0.05. Overlaid in grey are the histograms of the corresponding satellite pixel values. The bottom panel (c) shows density plots of the Sentinel and Landsat band 3 surface reflectance rasters over the study area (smallest possible rectangle containing all ground measurements), with the density of the ~~ground reflectance~~in situ HCRF for comparison.

To conclude the results, a note on the sensitivity of data and results to the spatial resolution of the satellite data and the accuracy of the geolocation of the in situ data: To assess the possible effects of the GPS accuracy or lack thereof, we compare the differences between in situ and satellite values presented previously to the differences that result when a buffer corresponding to the GPS uncertainty is created around each in situ measurement point. For the Sentinel data in the original 10 m resolution of band 3, the maximum number of pixels that any buffer touches is 4, the mean is 2.6, and most buffered in situ measurement points overlap with 2 pixels. For the 30 m Landsat data in band 3, the maximum number of pixels touched is also 4 while the mean is 1.5 and most in situ points are fully within only one pixel. Table 4 gives the standard deviation of differences between the in situ HCFR and the satellite data in different resolutions, grouped by the number pixels the buffered measurement points overlap with, to show how variability of results shifts depending on the buffer and the raster resolution. Changes caused by introducing the buffer are small in all groups. As expected, standard deviation increases with decreasing resolution of the satellite pixels due to the loss of detail in the satellite data. Figure 12 gives an overview of the ungrouped dataset with and without the buffer and at different raster resolutions.

Table 4: Comparison of in situ and satellite data by the standard deviation (SD) of the difference between in situ HCFR and satellite surface reflectance. Values are grouped by number of pixels that buffered in situ measurements overlap.

	Sentinel					Landsat			
Nr. of overlapping pixels	Nr. of points	SD, no buffer, 10m	SD, buffer, 10m	SD, buffer, 30m	SD, buffer, 60m	Nr. of points	SD, no buffer, 30m	SD, buffer, 30m	SD, buffer, 60m
1	25	0.098	0.098	0.108	0.129	134	0.129	0.129	0.134
2	124	0.119	0.118	0.120	0.124	94	0.106	0.107	0.103

3	9	0.057	0.065	0.069	0.099	1	-	-	-
4	76	0.122	0.121	0.127	0.136	5	0.082	0.074	0.083

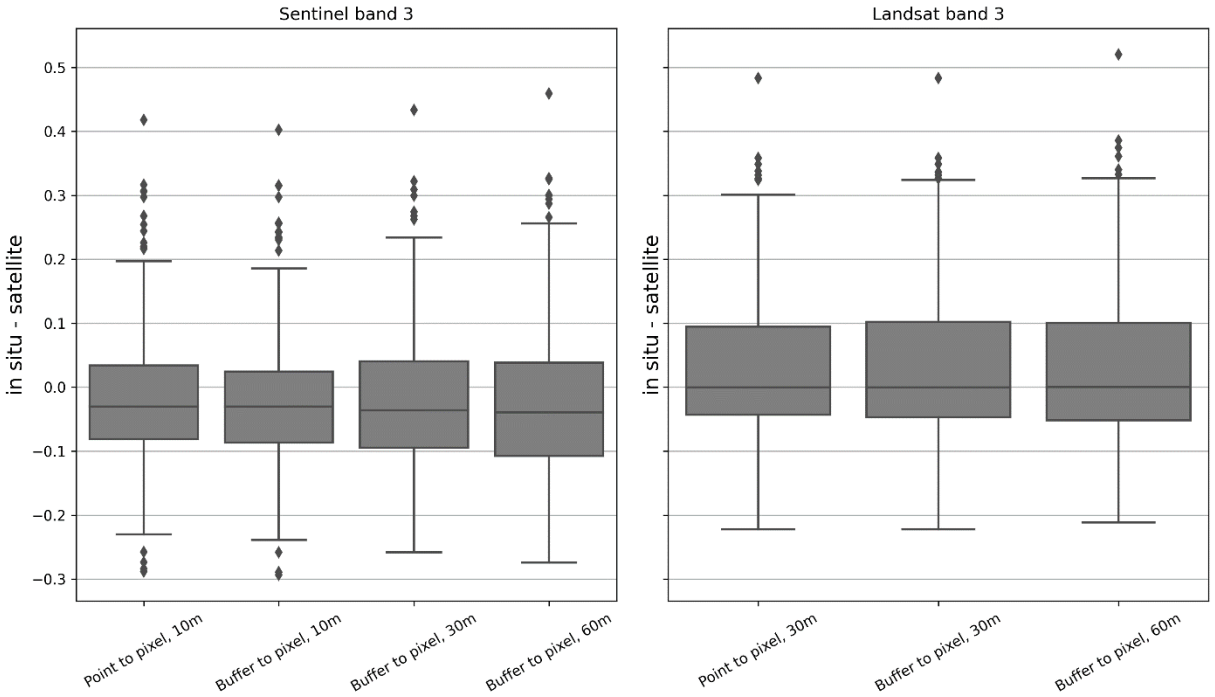


Figure 12: For the respective Sentinel and Landsat band 3 wavelength range, the difference between the in situ HCRF and satellite surface reflectance product is on the vertical axis. Point to pixel refers to the data as presented in previous figures. Buffer to pixel refers to data generated using a buffer around the in situ measurement points to account for GPS accuracy. For Sentinel, the original 10m resolution data was resampled to 30 and 60m. For Landsat, the original 30m resolution data was resampled to 60m.

4. Discussion

There are a number of complexities associated both with measuring reflectance properties on the ground and with any comparison between different products and data sets. Perhaps more than anything else, our results highlight the need for further in situ measurements and targeted data collection campaigns designed specifically to address some of the uncertainties detailed in the following.

4.1. Reflectance anisotropy and changing solar and atmospheric conditions

Ice is an anisotropic material and previous studies have shown that for glacier surfaces, anisotropy increases with decreasing albedo and depends on wavelength and solar zenith angle (Greuell and de Wildt, 1999; Klok et al., 2003; Naegeli et al., 2015). In order to truly quantify the effects of anisotropy in in situ spectroradiometric measurements, the bidirectional reflectance distribution function (BRDF) must be obtained – ideally for each measurement point. The BRDF cannot be measured directly but is approximated, e.g. by interpolating between multi-angular spectroradiometer measurements (Naegeli et al., 2015), or with modelling approaches (Malinka et al., 2016). While multi-angular HCRF measurements allow for the estimation of the BRDF, they are intrinsically dependent on the atmospheric conditions (cloud cover) at any given time, as well as on the topography and structure of the surface. Naegeli et al. (2015, 2017) use this approach to develop anisotropy correction factors for different glacier surface types in order to account for the typical underestimation of albedo in observations from nadir in remote sensing data. They find a difference between corrected and uncorrected albedo values of up to 11% for dirty ice in airborne imaging spectroscopy data. Nonetheless, the application of constant correction factors for clustered surface types is a simplification that obscures both the gradational nature of surface classification and the complexity of accounting for the effects of varying surface roughness on effective illumination angles. We consider a quantitative assessment of anisotropy beyond the scope of our

study and hope to tackle this issue in detail in future work. We assume that our in situ data as well as the satellite products underestimate the quantities they measure (HCRF and surface reflectance as per the respective documentation of the satellite products) due to the nadir or near-nadir observational angle, in particular for dark surfaces, and that uncertainties caused by anisotropy are likely to be in the range found by Naegeli et al. (2017). The local variability of reflectance properties of glacier ice is comprised of the spectral, as well as spatial and temporal variability of reflectance anisotropy, which require a combination of targeted, continuous measurements and modelling that accounts for the surface roughness of different glacier surface types to truly delineate.

The weather on September 3 (Landsat overpass) and September 4, 2019, (Sentinel overpass, in situ measurements) was very favourable. There was no cloud cover at the study site during either of the satellite overpasses and for the duration of the field measurements and we consider any changes in atmospheric conditions to be negligible. While the illumination angles naturally change over the course of the day and accordingly changed during the in situ measurements (Table 2), very low solar elevation angles were avoided. In their study on parametrizing BRDFs for glacier ice and Landsat TM, Greuell and de Wildt (1999) show that the spectrally integrated albedo of dark ice changes with the solar zenith angle and is particularly low for low zenith angles. Accordingly, we acknowledge that the changing solar angles are a source of uncertainty in our data and the comparison with the satellite derived reflectances, but we consider this uncertainty relatively small since measurements were carried out within a few hours before and after the satellite overpasses, avoiding very low solar elevation angles. Greuell and de Wildt (1999) also point out that the drop in albedo for low zenith angle is related to the presence of meltwater at later times of day (lower zenith angles), which highlights the difficulty of isolating one variable (zenith angle) in a complex system with multiple variables that change over time (surface processes like meltwater affecting reflectance properties).

The Landsat and Sentinel surface reflectance products both incorporate an atmospheric correction applied to TOA reflectance in the generation of the BOA product (Vermote et al., 2016; Main-Knorn et al., 2017). This introduces some uncertainty into the comparison with in situ data since the correction methods differ. Nonetheless, we believe that assessing how in situ data compare to the frequently used surface reflectance products of the Landsat-8 and Sentinel-2 suites is a necessary first step in being able to determine whether custom atmospheric corrections would improve results and if such improvements would be large enough to outweigh the added complexity and computational cost. We suggest that the answer to this question depends on the application and the spatial scale of the intended analysis. Again, this is beyond the scope of the presented study and is a point that needs to be specifically addressed in future work. We suggest that case studies at individual, well-studied glaciers can serve as an ideal testing ground for such issues, and will help to determine whether custom atmospheric corrections should be applied and are feasible on a regional or even global scale in satellite-based studies of ablation area reflectance properties.

4.2. Implications of in situ and satellite comparison

The results presented in section 3.1. highlight the large spatial variability of HCRF and different surface types encountered in the ablation area, both of which are in line with findings from other studies (Naegeli et al., 2015, 2017; Di Mauro et al., 2017). Section 3.2., the comparison of the in situ data with satellite values, arguably presents greater challenges in terms of interpretation and implications of the results.

In summary, there are three key findings which we believe may be important for further studies and for delineating the relationship between in situ and satellite derived reflectance:

- Sentinel surface reflectance values tend to be closer to the higher end of HCFR values measured in situ, while Landsat tends to be closer to the in situ minimum.
- The difference between in situ data and satellite data tends to decrease when there are more in situ data points per pixel, but not always and not in a clearly linear way.
- The reflectance of dark surfaces tends to be overestimated in the satellite products, while the reflectance of bright surfaces tends to be underestimated.

Explaining the above points in full requires targeted investigations specifically addressing the contributing factors and uncertainties, which – with our current data set – we can only provide a qualitative overview of:

As mentioned previously, different atmospheric corrections are used for the Sentinel and Landsat surface reflectance products. This may contribute to systematic differences in how surface reflectance is represented under differing lighting conditions and in different spectral ranges. Efforts to harmonize the Landsat and Sentinel surface reflectance data sets have great potential for minimizing this problem for applications where data from both satellites is used (Claverie et al., 2018).

Another issue that deserves more detailed attention is the narrow/spectral to broad band conversion required for comparing satellite reflectance in individual bands with the in situ data of the same wavelength range. We intentionally do not compute a shortwave broadband albedo from the satellite band values or the spectral in situ data to avoid introducing a further source of uncertainty. Instead, we limit ourselves to averaging over the band wavelength range in order to keep the comparison as straightforward as possible, but acknowledge that a glacier wide broad band albedo is a key parameter for many regional or global modelling applications.

The standard atmospherically corrected BOA reflectance products from satellite data are provided without correcting for the BRDF. The BRDF, describing the change of the reflectance with different observation and incidence geometries, can have a significant impact on the satellite-based reflectance as well as on the in situ data, leading to inherent challenges when comparing satellite based BOA reflectance with in situ reflectance measurements (Schaepman-Strub et al., 2006). Correcting Landsat and Sentinel surface reflectance with MODIS or VIIRS BRDF products to produce surface albedo has been shown to be a viable approach in some cases (Shuai et al., 2011; Li et al., 2018), but the coarse resolution of MODIS and VIIRS data is unlikely to capture the small-scale anisotropy effects of different glacier surface types. This would therefore be of limited use for our purposes. Optimizing methods for computing surface albedo from the L2A products, as well as from the in situ HCRF, requires further study and customized solutions accounting for local topographic effects and the spectral characteristics of the surfaces. We assume that for our case uncertainties due to the intrinsic difference between HCRF and satellite derived HDRF are small compared to other sources of uncertainty: The influence of local topography as a source of indirect radiation is not represented in the satellite derived values and the microstructure of the ice surface may locally affect in situ values on a scale that not visible to the satellite, but could be very significant for in situ measurements (e.g. small ice ridges or similar features acting as reflectors and/or scattering light into the FOV of the instrument).

Hendricks at al. (2004) state for spectroradiometric measurements at Hintereisferner compared to Landsat ETM+ imagery acquired about 2 weeks before the field measurements: “The reflectance of ice seems to be highly variable with both under -and overestimations of up to 76 % and 31 % respectively.” This corresponds well with our finding that both under- and overestimation occur frequently for both satellites. The factors mentioned above may partly explain the location of the shift from under- to overestimation (Fig. 9), but –again- targeted measurement campaigns are needed to truly quantify this.

The influence of very local backscattering could play a role in the seeming inconsistencies in the dependency of the difference between in situ and satellite data the on number of in situ measurement points per pixel (Fig. 10), but this also ties in with questions regarding the positional accuracy of the in situ measurement points and the satellite data, and the spatial representativity of point measurements for a larger area:

Our comparison of in situ and satellite data is based on the assumption that we know where both are located in a common coordinate reference system to a sufficient degree of accuracy. The accuracy of the position of the GPS points at the start and end points of the measurement profiles is approximately 3m. Sentinel-2 orthorectification is based on the PlanetDEM 90 digital evelation model (DEM), which incorporates the SRTM DEM in areas where SRTM is available, such as Austria (Kääb et al. 2016). The geometric accuracy of the Sentinel data hence depends on the accuracy of the underlying DEM, which is subject to a number of uncertainties particularly over mountainous terrain. Vertical inaccuracies – which propagate into horizontal inaccuracies - increase over glacier surfaces, especially in areas with large changes in surface elevation, as the DEM can only provide a snapshot of conditions for a moment in time and quickly becomes outdated in rapidly changing environments. Pandžic et al (2016) determine an average offset in the Sentinel-2 data for Austria of about 6m compared to a high-resolution regional DEM. The performance requirement of Landsat-8 OLI for geometric terrain corrected accuracy is specified as 12m (Storey et al, 2014). Kääb et al. (2016) find cross-track offsets of 20-30 m over glacier termini in the Swiss Alps when comparing Landsat-8 and Sentinel-2 scenes acquired on September 8, 2015. Accordingly, uncertainties regarding the GPS points of the in situ measurements as delineated in our sensitivity analysis (Table 4, Figure 12) can be considered relatively small compared to the those related to the orthorectification of the satellite data. Comparisons between in situ point data and pixel values from the satellite products must be interpreted keeping positional uncertainties in mind.

Decreasing the pixel resolution and averaging over multiple in situ measurement points can serve as an approach to reduce the influence of geometric errors. However, any sort of averaging procedure must also be assessed in terms of spatial representativeness of the point measurements for a greater area and, conversely, the down sampled satellite data for small scale surface processes. What can be considered representative will always be a question of scale and application. The glacier surface at the study site is locally very heterogenous and hence prone to representativeness errors (Wu et al., 2019). We selected the location of the in situ profile lines so that they cover what we consider to be the

typical surface features and types of a given section of the ablation zone and argue that our 20 m long profile lines with equidistant measurements at least every two meters capture any variations that are likely to influence the corresponding pixel values of the satellite data. Naturally, the less overlap there is between the profile lines and any given satellite pixel, the more likely it is that the in situ point data happen to capture something that differs strongly from what the satellite sees.

The different surface types identified at Jamtalferner (Fig. ~~6~~5) and their ~~spectral reflectances~~reflectance spectra are comparable to types of surfaces identified in Switzerland at Morteratsch and Glacier de la Plaine Morte by Di Mauro et al. (2017) and Naegeli et al. (2015), respectively, supporting the use of a classification scheme based on differentiating between a) clean and dirty ice surfaces and b) the presence or absence of liquid water on the ice surface. ~~It is generally understood that both "dirt" (organic or inorganic impurities) and liquid water reduce the reflectance of ice (e.g. Hall, 2012) and early studies showed diurnal cycles and high spatial variability of broadband albedo (Sauberer and Dirmhirn, 1951 & 1951; Dirmhirn and Trojer, 1955). However, understanding of how and why spectral ice albedo changes in time and space and how this affects the amount of energy available for melt remains incomplete~~Classifying the surface characteristics into discrete types can help to ensure representativeness e.g. by quantifying how much of a given area subsection relevant to the comparison with remote sensing data is comprised of which type and then sampling accordingly. However, surface types are not always discrete in practice. Nicholson and Benn (2006) indicate that the surface albedo of ice with scattered debris can be simulated in a modelling approach be linearly varying between clean ice albedo values and values for debris, but this does not necessarily account for other types of surfaces and even the clean ice albedo can vary considerably, especially if liquid water is present. Additionally, classification by type of any kind cannot address the issue of temporal representativeness unless the temporal variability of different surface types is first determined.

Profile 8 shows particularly poor agreement with the corresponding satellite data and may be an example where temporal variability plays a role: The profile crosses a section of ice where the contrast between dark and bright areas is comparatively strong. The profile line is roughly at a right angle to the flow direction of the glacier and "stripes" of meltwater channels and/or dirt cross the line. The profile has a comparable number of individual spectra with reflectance values above and below the profile mean, i.e. it is not a dark profile with a few bright outliers (compare e.g. to P6 in Fig. 7) or vice versa (e.g. P3), but alternates along the profile line. Agreement with the remote sensing data is decent for the darker spectra in P8 but the bright values are not captured. While we cannot rule out that the lack of agreement between the field and remote sensing data is due to an unusually unfortunate/unrepresentative positioning of the field measurement points in the satellite pixels, this may be an instance where the diurnal melt cycle and the associated presence/absence of water on the surface exacerbates the contrast between the dark and bright sections of the profile. In the bright sections, the porous weathering crust and cryconite hole structures appear to be drained of water, while the depressions of the melt channels are noticeably wet. Cook et al. (2016) indicate the occurrence of "sudden drainage events" in the weathering crust on a day-to-day time scale and a diurnal cycle of the hydrology of the weathering crust driven by meteorological conditions (radiation, turbulent fluxes). The time of day of a satellite overpass would determine which stage of this cycle the satellite sees and consequently the satellite data would not capture this variability. In order to assess how much the time of day of the overpass could systematically affect the representativeness of the satellite date for actual ground reflectance, it needs to be determined how significant and how consistent the diurnal cycle is. To do this, the driving processes must be identified, keeping in mind that these may be different for different types of glaciers and that different causes of short-term albedo change can overlap. E.g.: Azzoni et al. (2016) point out that meltwater increases albedo around midday in a daily cycle, while rain causes increased albedo for up to 4 days after the precipitation event. A seasonal cycle of albedo has been demonstrated in previous observational studies and modelling efforts of broadband albedo, highlighting the importance of continuous measurements (e.g. Hoinkes and Wendler, 1968; Nicholson and Benn, 2012; Möller and Möller, 2017).

4.3. Relevance of small-scale variability

The reflectance properties of ice are a central part of mass and energy balance modelling, usually in the form of a glacier wide broadband albedo, or using one value for ice in the ablation zone and one for snow covered areas. Resolving local albedo variations at a very small, sub-pixel scale is not required for regional or global studies, provided the albedo parametrization captures the conditions on the ground adequately for the region of interest. In their important 2015 study, Naegeli et al. find that Sentinel-2 and Landsat-8 reflectance data are within the suggested accuracy requirements for global climate modelling (±0.05, Henderson-Sellers and Wilson, 1983) over their study site, Glacier de la Plaine Morte in Switzerland. In the same study, they report a 10% difference in modelled mass balance when a spatially distributed albedo is used to force the model as opposed to a single, glacier wide albedo. Significantly larger

differences occur in parts of the glacier where water is present on the surface or the ice surface contains a lot of light-absorbing impurities. While the glacier-wide impact of a spatially distributed albedo on model results may be relatively small, this highlights that resolving local variability of reflectance properties and its causes is important for accurately predicting the future evolution of individual glaciers, especially in cases where the firn covered area is gone or greatly reduced and rapid melt is occurring. Only once the problem of different scales comparing point and spatially averaged data is solved, the relationship between albedo variability and mass balance point and averaged data can be tackled to calculate the effects on mass balance at glacier-wide or regional scale.

Aside from directly mass and energy balance related applications, reflectance data with high spatial and temporal resolution is essential to improve understanding of micro-hydrological processes in the weathering crust and how these may affect a possible larger scale darkening of increasingly snow free glaciers, e.g. by favoring or impeding the growth of ice algae, or the collection/washing out of cryoconite or other impurities. High resolution time series of spectral reflectance at representative locations in the ablation zone are needed to assess how changes in wetness and temperature, surface texture (cryoconite formation, roughness changes during the season), biotic productivity ~~and erosion and~~, deposition of sediment by melt water and rain affect ~~albedo~~reflectance properties on a small spatial scale, throughout the day and over the course of the ablation season. Establishing measurement efforts aimed at generating such time series on glaciers with existing mass balance monitoring networks would be highly desirable~~.~~ in order to better link small scale surface processes with mass and energy balance modelling.

5. Conclusion & Outlook

~~In order to scale assessments of ice albedo from the local to a regional or global level, satellite-derived data are indispensable. Earlier in the satellite era, several studies carried out comparisons of albedo data measured on the ground and surface reflectance derived from Landsat 5 Thematic Mapper scenes, finding considerable differences between in situ and satellite data especially in the ablation area (e.g. Hall et al., 1989 & 1990; Koelemeijer et al., 1993; Winther, 1993; Knap et al., 1999). These works are mostly based on albedo data from a single location, such as a weather station, and it was often not possible to carry out ground measurements so that they coincided with the satellite overpasses.~~

~~In more recent studies, an increasing focus is given to narrow-to-broadband conversions (Gardner et al, 2010; Naegeli et al., 2017). However, commonly used conversions are typically designed for use with Landsat 5 or 7, rather than Landsat 8 or Sentinel 2, which increases the challenges and uncertainties inherently associated with any narrow-to-broadband conversion. In addition, studies assessing the potential effects of anisotropy on satellite-derived surface reflectance data are sparse and the magnitude of associated uncertainties is hard to quantify (Naegeli et al., 2015 & 2017).~~

In comparing our in situ measurements with readily available L2A satellite products, we chose an "as simple as possible" approach to gain a general understanding of where sources of uncertainties are. ~~The albedo~~We found that the difference between in situ and satellite data is not uniform across satellite bands, between Landsat and Sentinel, and to some extent between surface types. Reflectance variability on the ground is not fully represented in the satellite data, which raises questions as to how well surface processes at rapidly changing glaciers such as Jamtalferner can be resolved with satellite data. ~~Surface reflectance products might be improved by developing dedicated atmospheric correction algorithms and quantifying the influence of anisotropy and different narrow-to-broadband conversions. Systematic collection of ground truth data will be fundamental to assessing the potential range and variations of uncertainty in satellite-derived glacier ice albedo and potentially reducing this uncertainty.~~

~~In addition to albedo,~~The reflectance properties of ice, along with other feedback mechanisms such as changing topography and glacier geometry, ~~also~~ significantly impact the rate of glacial retreat, contributing to the non-linear characteristics of glacier change and the high variability of defining parameters such as mass-balance or area change even among neighbouring glaciers subject to common climatic drivers (Charalampidis et al., 2018). Understanding these feedback mechanisms and associated processes is key to successfully predicting future glacier changes across spatial and temporal scales. Ice albedo will remain a significant source of uncertainty in modelling applications as long as the processes governing temporal and spatial variability are not fully understood.

~~5. Conclusion & Outlook~~

~~Our comparatively simplistic statistical comparison of Landsat and Sentinel L2A products with in situ data serves to exemplify that the difference between ground and satellite data is not uniform across satellite bands, between Landsat and Sentinel, and to some extent between ice surface types (underestimation of reflectance for bright surfaces, overestimation for dark surfaces). Assessing the reasons for these differences – with the eventual goal of improving satellite-derived surface reflectance of glacier ice – requires 1) further, systematic measurements of in situ spectral albedo and 2) in depth analysis of time-synchronous satellite data.~~

~~Given that increasing debris cover is already observed to be an major unknown during glacier disintegration and even deglaciation (e.g. Fischer et al., 2016; Fischer et al., in review), we urgently need to improve our process understanding and our data basis regarding albedo changes on glaciers.~~ Quantifying spatial and temporal variability of spectral reflectance and delineating the main causes of this variability for individual glaciers will improve modelling capabilities of glacier evolution and catchment hydrology. Satellite-derived reflectance products are a key component of tackling similar questions on the regional and global level. However, ground truth data from representative sites is essential in order to understand uncertainties associated with satellite albedo and surface reflectance products and potentially improve them for specific contexts.

Moving forward, an expansion of the monitoring network at Jamtalferner and, ideally, other glaciers, by continuous reflectance measurements in the ablation zone at a fixed location is needed, as well as "snap-shot" measurements of spectral, multi-angular reflectance at multiple strategic points in regular intervals. Combining analysis of spectral reflectance data from in situ and remote sensing sources with the wealth of contextual information available at Jamtalferner and other established monitoring sites has the potential to greatly improve our understanding of the complex interplay of surface changes, glacier dynamics, and mass- and energy balance.

Author Contribution

L. Felbauer and A. Fischer collected the in situ data. Subsequent data curation was carried out by L. Felbauer and L. Hartl. G. Schwaizer conceptualized the comparison of in situ and satellite derived data. L. Hartl developed the code for data analysis and visualizations, and wrote the manuscript with contributions from all co-authors.

Competing interests

The authors have no competing interests to declare.

Data availability

The spectral reflectance data ~~was submitted to the data repository pangaea.de and the DOI of the data set will be cited in the final version of this publication once it becomes available.~~can be downloaded at:
https://doi.pangaea.de/10.1594/PANGAEA.915932
And interactively explored in a web-app at: http://spectralalbedo.mountainresearch.at/

Acknowledgements

We are very grateful to Gottlieb Lorenz and the entire team at the Jamtal Hütte for providing an excellent base for field work at Jamtalferner and invaluable support over the years. We sincerely thank M. Pelto and the second, anonymous reviewer for their helpful comments!

References

Alexander, P. M., Tedesco, M., Fettweis, X., Van De Wal, R., Smeets, C. J. P. P., and Van Den Broeke, M. R.: Assessing spatio-temporal variability and trends in modelled and measured Greenland Ice Sheet albedo (2000-2013), The Cryosphere, 8(6), 2293-2312, 2014.

ASD Inc.: FieldSpec® HandHeld2™ Spectroradiometer User Manual. https://www.malvernpanalytical.com/en/support/product-support/asd-range/fieldspec-range/handheld-2-hand-held-vnir-spectroradiometer#manuals, Accessed: Sep. 22, 2020.

Azzoni, R. S., Senese, A., Zerboni, A., Maugeri, M., Smiraglia, C., and Diolaiuti, G. A.: Estimating ice albedo from fine debris cover quantified by a semi-automatic method: the case study of Forni Glacier, Italian Alps, The Cryosphere, 10, 665-679, 2016.

Box, J. E., Fettweis, X., Stroeve, J. C., Tedesco, M., Hall, D. K., and Steffen, K.: Greenland ice sheet albedo feedback: thermodynamics and atmospheric drivers, The Cryosphere, 6(4), 821-839, 2012.

Box J. E., van As D., and Steffen, K. Greenland, Canadian and Icelandic land-ice albedo grids (2000–2016). Geological Survey of Denmark and Greenland Bulletin, 38, 53-56, 2017.

Brun, F., Dumont, M., Wagnon, P., Berthier, E., Azam, M. F., Shea, J. M., Sirguey, P., Rabatel, A., and Ramanathan, A.: Seasonal changes in surface albedo of Himalayan glaciers from MODIS data and links with the annual mass balance, The Cryosphere, 9(1), 341-355, 2015.

Charalampidis, C., Fischer, A., Kuhn, M., Lambrecht, A., Mayer, C., Thomaidis, K., and Weber, M.: Mass-budget anomalies and geometry signals of three Austrian glaciers, Frontiers in Earth Science, 6, p. 218, 2018.

Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.C., Skakun, S.V. and Justice, C.: The Harmonized Landsat and Sentinel-2 surface reflectance data set. Remote sensing of environment, 219, pp.145-161, 2018.

Cook, J.M., Hodson, A.J. and Irvine-Fynn, T.D.: Supraglacial weathering crust dynamics inferred from cryoconite hole hydrology. Hydrological Processes, 30(3), pp.433-446, 2016.

Dirmhirn, I. and Trojer, E.: Albedountersuchungen auf dem Hintereisferner, Archiv für Meteorologie, Geophysik und Bioklimatologie, Serie B, 6(4), pp.400-416, 1955.

Dumont, M., Brun, E., Picard, G., Michou, M., Libois, Q., Petit, J. R., Geyer, S., and Josse, B.: Contribution of light-absorbing impurities in snow to Greenland's darkening since 2009, Nature Geoscience, 7(7), 509, 2014.

Fischer, A.: Comparison of direct and geodetic mass balances on a multi-annual time scale. The Cryosphere, 5, pp.107-124, 2011.

Fischer, A., Helfricht, K., Wiesenegger, H., Hartl, L., Seiser, B., and Stocker-Waldhuber, M.: Chapter 9 - What Future for Mountain Glaciers? Insights and Implications From Long-Term Monitoring in the Austrian Alps, in: Developments in Earth Surface Processes, edited by: Greenwood, G., B. and Shroder, J. F., Elsevier, 21, 325-382, 2016.

Fischer, A., Markl, G., and Kuhn, M.: Glacier mass balances and elevation zones of Jamtalferner, Silvretta, Austria, 1988/1989 to 2016/2017, Institut für Interdisziplinäre Gebirgsforschung der Österreichischen Akademie der Wissenschaften, Innsbruck, PANGAEA, https://doi.org/10.1594/PANGAEA.818772, 2016.

Fischer, A., Fickert, T., Schwaizer, G., Patzelt, G. and Groß, G.: Vegetation dynamics in Alpine glacier forelands tackled from space, Scientific reports, 9(1), pp.1-13, 2019.

Fischer, A., Seiser, B., -and Stocker-Waldhuber, M.: Capturing deglaciation in the Austrian Silvretta: Methods and Results, in review.

Gabbi, J., Huss, M., Bauder, A., Cao, F., & Schwikowski, M.: The impact of Saharan dust and black carbon on albedo and long-term mass balance of an Alpine glacier, The Cryosphere, 9(4), 1385-1400, 2015.

Gardner, A. S. and Sharp, M. J.: A review of snow and ice albedo and the development of a new physically based broadband albedo parameterization, Journal of Geophysical Research: Earth Surface, 115(F1), 2010.

GeoPandas developers: GeoPandas 0.8.0. URL: https://geopandas.org/, 2013-2019. Accessed: 2020.

Gillies, S., & others. (n.d.). Rasterio: Geospatial raster i/o for Python programmers. Mapbox, 2013. Retrieved from https://github.com/mapbox/rasterio, 2020.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: Google Earth Engine: Planetary-scale geospatial analysis for everyone, Remote Sensing of Environment, 202, pp.18-27, 2017.

Greuell, W. and de Wildt, M.D.R.: Anisotropic reflection by melting glacier ice: Measurements and parametrizations in

Landsat TM bands 2 and 4. Remote Sensing of Environment, 70(3), pp.265-277, 1999.

Hall, D. K., Chang, A. T. C., Foster, J. L., Benson, C. S., and Kovalick, W. M.: Comparison of in situ and Landsat derived reflectance of Alaskan glaciers, Remote Sensing of Environment, 28, 23-31, 1989.

Hall, D. K., Bindschadler, R. A., Foster, J. L., Chang, A. T. C., and Siddalingaiah, H.: Comparison of in situ and satellite-derived reflectances of Forbindels Glacier, Greenland, Remote Sensing, 11(3), 493-504, 1990.

Hall, D.: Remote sensing of ice and snow, Springer Netherlands, 2012.

Henderson-Sellers, A. and Wilson, M.F.: Surface albedo data for climatic modeling. Reviews of Geophysics, 21(8), pp.1743-1778, 1983.

Hendriksa J, Pellikkaa P, Peltoniemib J. Estimation of anisotropic radiance from a glacier surface-Ground based spectrometer measurements and satellite-derived reflectances. InProceedings, 30th International Symposium on Remote Sensing of Environment: Information for Risk Management and Sustainable Development, November 10-14, Honolulu, Hawaii, 2003.

Hoinkes, H., and Wendler, G.: Der Anteil der Strahlung an der Ablation von Hintereis-und Kesselwandferner (Ötztaler Alpen, Tirol) im Sommer 1958, Archiv für Meteorologie, Geophysik und Bioklimatologie, Serie B, 16(2-3), 195-236, 1968.

Hunter, J.D.: Matplotlib: A 2D Graphics Environment, Computing in Science & Engineering, 9, 90-95, 2007.

Jaffé, A.: Über Strahlungseigenschaften des Gletschereises, Archiv für Meteorologie, Geophysik und Bioklimatologie, Serie B, 10(3), pp.376-395, 1960.

Kääb, A., Winsvold, S.H., Altena, B., Nuth, C., Nagler, T. and Wuite, J.: Glacier remote sensing using Sentinel-2. part I: Radiometric and geometric performance, and application to ice velocity. Remote Sensing, 8(7), p.598, 2016.

Klok, E.L., Greuell, W. and Oerlemans, J.:Temporal and spatial variation of the surface albedo of Morteratschgletscher, Switzerland, as derived from 12 Landsat images. Journal of Glaciology, 49(167), pp.491-502, 2003.

Knap, W. H., Brock, B. W., Oerlemans, J., and Willis, I. C.: Comparison of Landsat TM-derived and ground-based albedos of Haut Glacier d'Arolla, Switzerland, International Journal of Remote Sensing, 20(17), 3293-3310, 1999.

Koelemeijer, R., Oerlemans, J., and Tjemkes, S.: Surface reflectance of Hintereisferner, Austria, from Landsat 5 TM imagery, Annals of Glaciology, 17, 17-22, 1993~~,~~.

Kuhn, M.: The response of the equilibrium line altitude to climate fluctuations: theory and observations, in: Glacier fluctuations and climatic change, 407-417, Springer, Dordrecht, 1989.

Lee, Y.: SpecDAL Reference. https://specdal.readthedocs.io/en/latest/, 2017. Accessed: September 2019.

Li, Z., Erb, A., Sun, Q., Liu, Y., Shuai, Y., Wang, Z., Boucher, P. and Schaaf, C.: Preliminary assessment of 20-m surface albedo retrievals from sentinel-2A surface reflectance and MODIS/VIIRS surface anisotropy measures. Remote sensing of environment, 217, pp.352-365, 2018.

Malinka, A., Zege, E., Heygster, G. and Istomina, L.: Reflective properties of white sea ice and snow. The Cryosphere, 10, 2541–2557, 2016.

Mayer, B. and Kylling, A.: Technical note: The libRadtran software package for radiative transfer calculations - description and examples of use, Atmos. Chem. Phys., 5, 1855-1877, https://doi.org/10.5194/acp-5-1855-2005, 2005.

~~Hall, D. K., Chang, A. T. C., Foster, J. L., Benson, C. S., and Kovalick, W. M.: Comparison of in situ and Landsat derived reflectance of Alaskan glaciers, Remote Sensing of Environment, 28, 23-31, 1989.~~
~~Hall, D. K., Bindschadler, R. A., Foster, J. L., Chang, A. T. C., and Siddalingaiah, H.: Comparison of in situ and satellite-derived reflectances of Forbindels Glacier, Greenland, Remote Sensing, 11(3), 493-504, 1990.~~

~~Hall, D.: Remote sensing of ice and snow, Springer Netherlands, 2012.~~

~~Hoinkes, H., and Wendler, G.: Der Anteil der Strahlung an der Ablation von Hintereis-und Kesselwandferner (Ötztaler Alpen, Tirol) im Sommer 1958, Archiv für Meteorologie, Geophysik und Bioklimatologie, Serie B, 16(2-3), 195-236, 1968.~~

Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., and Gascon, F.: Sen2Cor for Sentinel-2, in: Image and Signal Processing for Remote Sensing XXIII, International Society for Optics and Photonics, Warsaw, Poland, 11-13 September 2017, Vol. 10427, p. 1042704, 2017.

McKinney W.: Data structures for statistical computing in python. In: Proceedings of the 9th Python in Science Conference 2010 Jun 28 (Vol. 445, pp. 51-56).

Nicholson, L. and Benn, D.I.: Properties of natural supraglacial debris in relation to modelling sub-debris ice ablation. Earth Surface Processes and Landforms, 38(5), pp. 490-501, 2012.

Di Mauro, B., Baccolo, G., Garzonio, R., Giardino, C., Massabò, D., Piazzalunga, A., Rossini, M., and Colombo, R.: Impact of impurities and cryoconite on the optical properties of the Morteratsch Glacier (Swiss Alps), The Cryosphere, 11(6), 2393, 2017.

Di Mauro, B., Garzonio, R., Baccolo, G., Franzetti, A., Pittino, F., Leoni, B., Remias, D., Colombo, R. and Rossini, M.: Glacier algae foster ice-albedo feedback in the European Alps. Scientific reports, 10(1), pp.1-9, 2020.

Ming, J., Du, Z., Xiao, C., Xu, X., and Zhang, D.: Darkening of the mid-Himalaya glaciers since 2000 and the potential causes, Environmental Research Letters, 7(1), 014021, 2012.

Ming, J., Wang, Y., Du, Z., Zhang, T., Guo, W., Xiao, C., Xu, X., Ding, M., Zhang, D., and Yang, W.: Widespread albedo decreasing and induced melting of Himalayan snow and ice in the early 21st century, PLoS One, 10(6), e0126235, 2015.

Moller, M. and Moller, R.: Modeling glacier-surface albedo across Svalbard for the 1979–2015 period: The HiRSvaC500-a data set. J. Adv. Model. Earth Syst., 9, 404– 422, 2017.

Naegeli, K., Damm, A., Huss, M., Schaepman, M., ~~&~~and Hoelzle, M.: Imaging spectroscopy to assess the composition of ice surface materials and their impact on glacier mass balance, Remote Sensing of Environment, 168, 388-402, 2015.

Naegeli, K. and Huss, M.: Mass balance sensitivity of mountain glaciers to changes in bare-ice albedo. Annals of Glaciology, No. 75, The Cryosphere in a Changing Climate, 2017.

Naegeli, K., Damm, A., Huss, M., Wulf, H., Schaepman, M., ~~&~~and Hoelzle, M.: Cross-Comparison of albedo products for glacier surfaces derived from airborne and satellite (Sentinel-2 and Landsat 8) optical data, Remote Sensing, 9(2), 110, 2017.

Naegeli, K., Huss, M~~., &~~. and Hoelzle, M.: Change detection of bare-ice albedo in the Swiss Alps, The Cryosphere, 13(1), 397-412, 2019.

Nicodemus, F.E., Richmond, J.C., Hsia, J.J., Ginsberg, I.W. and Limperis, T.: Geometrical considerations and nomenclature for reflectance, Vol. 160. Washington, DC: US Department of Commerce, National Bureau of Standards, 1977.

Oerlemans, J.: Glaciers and climate change. CRC Press, 2001.

Oerlemans, J., R. H. Giesen, and M. R. Van den Broeke: Retreating alpine glaciers: increased melt rates due to accumulation of dust (Vadret da Morteratsch, Switzerland), Journal of Glaciology, 55, no. 192,- 729-736, 2009.

Painter, T. H., Flanner, M. G., Kaser, G., Marzeion, B., VanCuren, R. A., and Abdalati, W.: End of the Little Ice Age in the Alps forced by industrial black carbon, Proceedings of the National Academy of Sciences, 110(38), 15216-15221, 2013.

Pandzić, M., Mihajlović, D., Pandzić, J. and Pfeifer, N.: Assessment of the geometric quality of sentinel-2 data. In XXIII ISPRS Congress, Commission I, Vol. 41, No. B1, 489-494. International Society for Photogrammetry and Remote Sensing, 2016.

Paul, F., Machguth, H., and Kääb, A.: On the impact of glacier albedo under conditions of extreme glacier melt: the summer of 2003 in the Alps, EARSeL eProceedings, 4(2), 139-149, 2005.

Qu, B., Ming, J., Kang, S. C., Zhang, G. S., Li, Y. W., Li, C. D., Zhao, S.Y., Ji, Z. M., and Cao, J. J.: The decreasing albedo of the Zhadang glacier on western Nyainqentanglha and the role of light-absorbing impurities, Atmospheric Chemistry and Physics, 14(20), 11117-11128, 2014.

Richter, R. and Schläpfer, D.: Atmospheric/Topographic Correction for Satellite Imagery: ATCOR-2/3 UserGuide", DLR IB 565-01/11, Wessling, Germany, 2011.

Rhodes, B.: PyEphem. URL: https://rhodesmill.org/pyephem/toc.html, 2020.

Sauberer, F.: Versuche über spektrale Messungen der Strahlungseigenschaften von Schnee und Eis mit Photoelementen, Meteorol. Z., 55, 250, 1938.

Sauberer, F. and Dirmhirn, I.: Untersuchungen über die Strahlungsverhältnisse auf den Alpengletschern, Archiv Met. Geoph. Biokl. Ser. B.3, 256, 1951.

Sauberer, F. And Dirmhirn, I.: Der Strahlungshaushalt horizontaler Gletscherflächen auf dem Hohen Sonnblick, Geogr. Ann., 34, 261, 1952.

Schaepman-Strub, G., T. Painter, S. Huber, S. Dangel, M. E. Schaepman, J. Martonchik, and F. Berendse.: About the importance of the definition of reflectance quantities-results of case studies. In: Proceedings of the XXth ISPRS Congress, 361-366, 2004.

Schaepman-Strub, G., Schaepman, M. E., Painter, T. H., Dangel, S., and Martonchik, J. V.: Reflectance quantities in optical remote sensing—Definitions and case studies. Remote sensing of environment, 103(1), 27-42, 2006.

Shuai, Y., Masek, J.G., Gao, F. and Schaaf, C.B.: An algorithm for the retrieval of 30-m snow-free albedo from Landsat surface reflectance and MODIS BRDF. Remote Sensing of Environment, 115(9), 2204-2216, 2011.

Storey, J., Choate, M. and Lee, K.: Landsat 8 Operational Land Imager on-orbit geometric calibration and performance. Remote sensing, 6(11), 11127-11152, 2014.

U.S. Geological Survey: Landsat 8 Collection 1 (C1) Land Surface Reflectance Code (LaSRC) Product Guide Version 3.0, 2020. https://www.usgs.gov/media/files/landsat-8-collection-1-land-surface-reflectance-code-product-guide Accessed: September 17, 2020.

van As, D., Fausto, R. S., Colgan, W.T., and Box, J.E.: Darkening of the Greenland ice sheet due to the melt albedo feedback observed at PROMICE weather stations. Geological Survey of Denmark and Greenland (GEUS) Bulletin 28:69-72, 2013.

Van de Wal, R.S.W., Oerlemans, J., and Van der Hage, J.C.: A study of ablation variations on the tongue of Hintereisferner, Austrian Alps, Journal of Glaciology, 38(130), pp.319-324, 1992.

Van der Walt, S., Colbert, C., and Varoquaux, G.: The NumPy Array: A Structure for Efficient Numerical Computation, Computing in Science & Engineering, 13, 22-30, 2011.

Van Rossum, G., Drake, F.L.: Python 3 Reference Manual. Scotts Valley, CA: CreateSpace; 2009.

Vermote, E., Justice, C., Claverie, M., and Franch, B.: Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product, Remote Sensing of Environment, 185, 46-56, 2016.

Winther, J. G.: Landsat TM derived and in situ summer reflectance of glaciers in Svalbard, Polar Research,12(1), 37-55, 1993.

Wu, X., Wen, J., Xiao, Q., You, D., Lin, X., Wu, S. and Zhong, S.: Impacts and Contributors of Representativeness Errors of In Situ Albedo Measurements for the Validation of Remote Sensing Products. IEEE Transactions on Geoscience and Remote Sensing, 57(12), 9740-9755, 2019

Zemp, M., Frey, H., Gärtner-Roer, I., Nussbaumer, S.U., Hoelzle, M., Paul, F., Haeberli, W., Denzinger, F., Ahlstrøm, A.P., Anderson, B., and Bajracharya, S.: Historically unprecedented global glacier decline in the early 21st century, Journal of Glaciology, 61(228), ~~pp.~~745-762, 2015.

Zemp, M., Huss, M., Thibert, E., Eckert, N., McNabb, R., Huber, J., Barandun, M., Machguth, H., Nussbaumer, S.U., Gärtner-Roer, I. and Thomson, L.: Global glacier mass changes and their contributions to sea-level rise from 1961 to 2016, Nature, 568(7752), pp.382-386, 2019.

Zeng, Q., Cao, M., Feng, X., Liang, F., Chen, X., and Sheng, W.: A study of spectral reflection characteristics for snow, ice and water in the north of China, Hydrological applications of remote sensing and remote data transmission, 145, pp.451-462, 1984.