Reply to Anonymous Referee #1

We would like to thank you for the review and your constructive comments which help to improve the manuscript. We have revised the manuscript accordingly and our point-by-point responses (in blue) to the specific comments (in red) are given. The modification made in the manuscript is presented in green. This document also includes a marked-up version of manuscript.

Best Regards,
Soheila Jafariserajehlou

Comments to the Author:

1. Line 97: If you apply your algorithm to AATSR, clouds can change significantly in a period of 2 days quoted here as the return time.

- ✓ Author's response: We agree with the statements with respect to the changes of clouds in 2 days. In fact, these changes can help us to separate clouds (highly variable) from cloud-free surface (which could be considered as relatively stable in 2 days). We modified the text and explained more about the advantage of a shorter revisit time to our algorithm:
- ✓ Modifications: Line 104-108: In addition to multiple imagery over the Arctic, the shorter time interval between satellites over-passes over the same scene provides images with less variability in the observed cloud-free areas which the algorithm looks for. For the two SLSTR, the revisit time is 0.9 days at the equator (Coppo et al., 2010) and this time becomes even shorter at higher latitudes due to orbital convergence.

2. Line 162: What about continuity in the data record? There is a ~4 year gap between the failure of AATSR and the launch of SLSTR.

- ✓ Author's response: To fill this gap, we plan to apply the algorithm to the AVHRR measurements which have observations at the required wavelengths in ASCIA.
- ✓ Modifications: lines 180-183: However, there is a ~4 years gap between the failure of AATSR and the launch of SLSTR. To fill this gap, we will apply ASCIA also to the Advanced Very High Resolution Radiometer (AVHRR) sensor carried by National Oceanic and Atmospheric Administration (NOAA).

3. Line 250: The 11 micron BT although closest to surface temperature is not the 'real surface temperature' as it includes atmospheric effects. To give an example, LST or SST is never equal to BT in the 11 micron channel. If it were, we wouldn't need to retrieve surface temperature.

- ✓ Author's response: We agree. Actually, we meant the measured BT in the 11 micron channel is regarded to be in good agreement with the real surface temperature especially over snow and ice (Istomina et al., 2010; Musial et al., 2014). We corrected the sentence as below:

- ✓ Modifications: Line 271-274: This is estimated from observations at 11 µm where absorption by water vapor and other trace gases is small; most objects in regions outside of the tropics can be treated as blackbodies and the measured BT considered being in good agreement with real surface temperature (Istomina et al., 2010; Musial et al., 2014).

4. Line 312: Are you including all measurements within the block in the PCC calculation? This isn't clear in the text. If you are then your PCC is based on ~25x25 pixels (x2 for two scenes?).

- ✓ Author's response: In this step we include all measurements (at 1.6 micron) within the block of 25x25 pixels in PCC calculation.
- ✓ Every PCC calculation is based on ~25x25 pixels of two images. But the algorithm repeats PCC calculations between every two images unless all measurements in the time series are included. For one block of ~25x25 pixels, PCC is calculated n times (n depends on the number of satellite overpass over the same scene). We explained this point at lines 330-332.

I would think in this scenario that most of the information is coming from the spatial variability rather than the temporal variability? If you are using a block average, this seems like too few observations to make a valid PCC calculation…

- ✓ In fact the information comes from spatial changes during the time. Therefore we believe both spatial and temporal variabilities contribute to this cloud detection algorithm. Since we do not have a single PCC to decide about one block of ~25x25 pixels, instead we have PCC values over the time span. So we would say temporal variability and we would say it's not a block average.
- ✓ Modifications: Line 337-339: After computing the first binary cloud flag at block level using the last measurement and one previous image, ASCIA keeps the result in a memory and repeats the procedure with second previous data. This procedure is iterated until the last measurement of the data series is involved.

5. Line 335: New ice is also dark and I think you would find it hard to distinguish from open water. It would be good to mention this here too.

- ✓ Author's response: We appreciate this hint from the reviewer and added explanation about new ice to this part as below:
- ✓ Modifications: Line 361-362: Sea-ice is distinguished from water on the basis of its higher brightness; one scene might be white enough to be considered as ice. However, melting or broken ice as well as new ice would not be labeled as ice.

6. Making comparisons with SYNOP using a 45-minute time window for validation could be problematic? Clouds can move across a scene within minutes?

- ✓ In preparation of the article, we performed a comprehensive review of previous works to define the optimal maximum time difference. However, statements in literature strongly vary: 10 min (Werkmeister et al., 2015), 15 min (Musial et al., 2014), 1 h (Dybbroe et al., 2005) and 4 h (Meerkötter et al., 2004), obviously

all for different kinds of meteorological conditions. The investigation and results in the previous publications indicate that temporal difference in validation of satellite retrievals against SYNOP may vary based on meteorological conditions. Therefore we also tried to check meteorological conditions in our study.

- ✓ First, we checked to see which fraction of our results could be affected by longer temporal difference? We found that in 30% of our validation scenarios time difference exceeds half an hour. We tried to answer your question in these 30%. Do clouds change in the interval of 45 minutes? Or clouds could be almost stable due to the unique Arctic environment compared to lower latitudes? To answer this question we have performed two investigations:

  1) We checked SYNOP data before and after validation time to see whether cloud changes could be observed during this time? Available to us were SYNOP data, i.e. cloud fraction every 3 hours over Svalbard and every 1 hour over Greenland. We found that only in 12% of our validation scenarios cloud fraction changed (±1 or ±2 oktas) within 2 hours (for Greenland) and – 6 hours (for Svalbard).

  2) Another perspective on this topic is to ask how fast cloud can travel during our validation time window? We have checked the wind speed over Svalbard to answer the question. How strong the wind needs to be to move clouds out of or into in our validation area which is defined as a circle with the radius of 20 km around each SYNOP station? If we assume that the cloud is in the middle of this circle, we need at least a wind speed of 7.5 m/s to move clouds out of or into this circle in 45 minutes. Based on information from the Norwegian Meteorological Institute which provides an average of hourly wind speed, we see that the average wind speed over or close to the selected SYNOP stations at the closest stations during satellite overpass time is usually very low. For example below 3m/s and only in one scenario, it exceeds 7.5m/s slightly.

- ✓ Second, choosing a smaller temporal difference like for instance 0.5h would limit the number of observations and introduce a sampling error. For example, by filtering the validation dataset with respect to 0.5h temporal difference, 30% of the validation data will be lost. Therefore, to have a trade-off between good statistics/sampling and representativeness, we decided to keep the temporal interval of 45 minutes.
- ✓ However, temporal difference between satellite and SYNOP measurements is one of several sources of uncertainty (different viewing perspective, different spatial footprint etc.) which affect validation results.
  We have updated the text to address this comment and explain the uncertainty which originates from time difference.

- ✓ Modification: Line 465-477: To define the optimal maximum temporal difference between SYNOP and satellite data, other comparable validation activities used different temporal intervals like 10 min (Werkmeister et al., 2015), 15 min (Musial et al., 2014), 1 h (Dybbroe et al., 2005) and 4 h (Meerkötter et al., 2004). The investigation and results in the previous publications indicate that temporal difference in validation of satellite retrievals against SYNOP depend on meteorological conditions. Allowing only a

small temporal difference between measurement datasets (here: SYNOP and ASCIA) ensures an optimal temporal overall but can introduce a significant sampling error due to the small number of scenes for validation (Bojanowski et al., 2014). According to Bojanowski et al. (2014) a temporal difference of 90 min to compare with SYNOP measurements at temporal resolution of 3 h minimizes the sampling error (Bojanowski et al., 2014). However, potential longer temporal difference will introduce an error which should be considered along other sources of uncertainty (different viewing perspective, different spatial footprint and etc.). In this study, the maximum allowed temporal difference between the ASCIA retrievals and SYNOP measurements is less than ±20 minutes in most cases and generally does not exceed ± 45 minutes.

- ✓ Line 558: Bojanowski, J., Stöckli, R., Tetzlaff, A., and Kunz, H.: The Impact of Time Difference between Satellite Overpass and Ground Observation on Cloud Cover Performance Statistics, Remote Sensing, 6, 12 866–12 884, https://doi.org/10.3390/rs61212866, 2014.
- ✓ Line 575: Dybbroe, A., Karlsson, K.-G., Thoss, A., NWCSAF AVHRR cloud detection and analysis using dynamic thresholds and radiative transfer modeling. Part II: Tuning and validation. J. Appl. Meteorol., 44, 55–71, 2005.
- ✓ Line 640: Meerkötter, R., König, C., Bissolli, P., Gesell, G., Mannstein, H., A 14-year European cloud climatology from NOAA/AVHRR data in comparison to surface observations. Geophys. Res. Lett., 31, doi:10.1029/2004GL020098, 2004.

7. *Please check the use of English throughout. The manuscript is readable but often the sentence structure isn't quite right and this makes it more difficult to understand. To give an example (line 222): o Manuscript: 'One major contributors of error in aerosol retrieval is misclassifying heavy aerosol loads with clouds.' o Correct English: 'One of the major contributors to error in aerosol retrievals is misclassification of heavy aerosol loads as cloud.'

- ✓ Author's response: Done.
- ✓ Modifications: Line 244-245: One of the major contributors to error in aerosol retrievals is misclassification of heavy aerosol loads as cloud.
- ✓ We also check the whole manuscript.

8. The structure of the text within the various sections of the manuscript could be improved/tightened up in places, particularly in the introduction.

- ✓ Done.

9. Figures 1-2, 5-13: It would be good to enlarge all of these figures so that they are more readable and cloud/surface features can be identified.

- ✓ Done; all figures are enlarged.

10. Figures 1-2, 5-13: It is far more intuitive for the reader if cloud is white and clear-sky is black. Colours should also match between algorithms. For example in one plot you have black=cloud, white=clear, and in another white=snow. This can become quite confusing for the reader to interpret.

- ✓ Done. Thanks for pointing out this issue. The figures (5-13) have been updated and they are more intuitive for the reader now. But in figures 1-2, we used different color schemes to highlight different features such as temperature (not only cloud information) observed at various wavelengths.

11. Table 4: The use of the terms 'correct' and 'incorrect' here assumes that there is no uncertainty in the SYNOP data. Is this really true? How are the oktas determined from this data? Does it involve human input? Is misclassification possible? I would be inclined to use the terms 'agreement' and 'disagreement' with the caveat that SYNOP isn't perfect made within the text.

- ✓ Done. We agree. The use of "correct" term is not suitable, as you mentioned, the measurements involve human input. Actually, we explained this uncertainty involved in these observations at lines 196-198. We changed the word correct and incorrect to agreement and disagreement.
- ✓ Modifications: line 714.

	Criteria	
Cloud data	within ±2 oktas	within ±1 okta
ASCIA vs. SYNOP	96 % agreement 4 % disagreement	83 % agreement 17 % disagreement
ESA vs. SYNOP	68 % agreement 32 % disagreement	50 % agreement 50 % disagreement

12. Figures with RGB: Are these actual RGB's or false colour images? This should be made clear and the channels used described in the figure captions.

- ✓ They are false colour images. We used 0.67, 0.87 and 0.55 µm channels. We added this information to figure captions.

13. Figure 16: Light blue error bars are difficult to see. Could you change the colour?

- ✓ Done.

14. There is some existing literature about cloud detection and surface classification over Polar Regions from within the SST CCI project: o Bulgin, C. E., Eastwood, S., Embury, O., Merchant, C. J. and Donlon, C. (2015) Sea surface temperature climate change initiative: alternative image classification algorithms for sea-ice affected oceans. Remote Sensing of Environment, 162. pp. 396-407. ISSN 0034-4257 doi: https://doi.org/10.1016/j.rse.2013.11.022 How does your algorithm compare to a Bayesian approach? This paper also shows the limitations of the SADIST mask that you mention in your paper.

- ✓ We read the above mentioned paper from Bulgin et al. (2015). The work aims to discriminating clear-sky ice-free ocean. Though these two algorithms (Bulgin et al. (2015) and this work) are different methods, there are some similarities in the general ideas of both algorithms. For example, we used similar

wavelengths (3.7, 1.6, 0.6 µm). Based on our investigation, the combination of the selected bands can improve the accuracy of snow/ice, cloud and ocean discrimination; similar conclusion can be found in Bulgin et al. (2015). The benefits can be increased when spectral and textural modification using the texture derived from 1.6 µm is employed. Additionally we also concluded that using temperature information alone, could lead to significant misclassification over Polar Regions as Bulgin et al. (2015) indicated.

- ✓ Both algorithms have their advantages and disadvantages. The advantageous of Bulgin et al. (2015) is that the algorithm has been optimized of the algorithm for night/twilight time. The advantageous of our algorithm is: it can be used over land covered by snow/ice along Open Ocean.
- ✓ We added the mentioned paper to our references as below:
- ✓ Modification: line 67-69: SADIST is known to misclassify ice, cloud and open ocean in Polar Regions. Bulgin et al. (2015) developed a Bayesian approach in ESA's Climate Change Initiative (CCI) project to overcome this limitation (Hollmann et al., 2013).
- ✓ Line 561-563: Bulgin, C. E., Eastwood, S., Embury, O., Merchant, C. J. and Donlon, C., Sea surface temperature climate change initiative: alternative image classification algorithms for sea-ice affected oceans. Remote Sensing of Environment, 162, pp. 396-407, 2015.
- ✓ Line 593-595: Hollmann, R.; Merchant, C.J.; Saunders, R.; Downy, C.; Buchwitz, M.; Cazenave, A.; Chuvieco, E.; Defourny, P.; de Leeuw, G.; Forsberg, R.; et al. The ESA climate change initiative Satellite Data Records for Essential Climate Variables. Bull. Am. Meteorol. Soc., 94, 1541–1552, 2013.

15. Have you given any consideration to classification at nighttime? I realise that in the context of aerosol retrievals this isn't relevant, but it is for other applications so would be good to include a sentence on this.

- ✓ We did not consider cloud identification at night time yet. As you already mentioned the aim of this work was masking clouds for aerosol retrieval. However, we may have such improvements in the next version of this algorithm. We modified the text as below:
- ✓ Modifications: Line 537-539: The objective of this study was to assess and validate the current version of ASCIA for daytime observations. For night time application, an adaption of ASCIA is planned to identify clouds during night.

Technical comments:

1. Line 32: Comma after 'though' doesn't make sense.

- ✓ Done.
- ✓ Modifications: Line 33: Though the attribution of the origins this phenomenon is …

2. Line 36: Comma after 'since' doesn't make sense. Please check throughout, as there are not many instances in English where it is appropriate to use a comma after the first word in a sentence, except with 'however' on occasion.

- ✓ Done.

- ✓ Modifications: Line 37: Since all developed cloud detection methods encounter many obstacles originating from the unique atmosphere and surface conditions in the Arctic (Curry et al., 1996).

3. Line 39: Snow/ice are also cold – this is the limiting factor in the thermal infrared.

- ✓ Done.
- ✓ Modifications: Line 40-42: For example, snow/ice are also like clouds cold, the lack of strong thermal contrast is a limitation in the retrieval of clouds in the thermal infrared (Rossow and Garder 1993; Curry et al., 1996).

- ✓ 4. Line 47: 'as' thin cloud?
- ✓ Done.
- ✓ Modifications: Line 49: To avoid the uncertainty included in AOT products due to significant misclassification of heavy aerosol load as thin cloud…

5. Line 151/152: The revisit times stated here are not consistent with those given in the introduction (0.9 days).

- ✓ Done.
- ✓ Modifications: Line 166-167: This yields global revisit times of 1.9 days at the equator with two satellites and 0.9 day with one satellite.

6. Line 277: Only in Polar Regions – deserts for example can be bright. Sunglint over water is also bright.

- ✓ Done.
- ✓ Modifications: Line 299: This is because land scenes in Polar region are dark in comparison to cloud and snow.

7. Line 650: Should be 'lower'.

- ✓ Done.
- ✓ Modifications: Line 722: (lower panels)

Reply to Anonymous Referee #2

We would like to thank you for the review and your constructive comments which help to improve the manuscript. We have revised the manuscript accordingly and our point-by-point responses (in blue) to the specific comments (in red) are given. The modification made in the manuscript is presented in green. This document also includes a marked-up version of manuscript.

Best Regards,
Soheila Jafariserajehlou

Comments to the Author:

Reviewer #2: Please check you English. The content of the paper really suffer from lack of structure, wrong grammar and misplaced commas. Also, you should not use an article before ASCIA, except when it is used as adjective. Example: "ASCIA retrieves clouds over Artic" or "The ASCIA retrieval over Artic". This is true for all acronyms and abbreviations (e.g. SLSTR). Since (which certainly does not require a comma) introduces a subordinate sentence, which cannot be separated from the main clause by a full stop.

- ✓ We have corrected all these points throughout the manuscript.

Reviewer #2: It is not clear to me whether or not your algorithm is applicable during winter. At line 357 you write that your targeted seasons are spring, summer and autumn and then you choose March, May and July. This is already confusing by itself. Later, at line 362 you write that ASCIA is not optimized for winter time. Could you please clarify this?

- ✓ We are sorry for this inaccurate statement. The algorithm is not optimized for night time retrievals (the information from the visible channels may not be valid during night) and we do not attempt to apply it during polar night. In fact, we only had access to 3 months (March, May and July) of SYNOP data which we used for validation. We understood the confusion arising from this statement and it would be better not to mention seasons here. We modified the text as below:
- ✓ Modification: Line 388- 395: Three months of data from March, May, and July have been acquired over Greenland and Svalbard to assess the performance of ASCIA in a wide range of solar zenith angles (60°-85°), surface and atmospheric conditions observed at high latitudes. In order to take various surface types in the Arctic into account, we selected case studies including, highly variable topography and fairly homogeneous snow cover, coast lines, land and ocean along snow and ice covered surface. The designed criteria for ASCIA are optimized for various regions over the Arctic observed under different solar illumination conditions. Polar night and transition seasons at low light conditions are excluded from our retrievals.

Reviewer #2: Fig. 5 The left panel over Greenland shows 2 rectangles in left part of the image. Could you please discuss where they arise from? Your algorithm shows promising results, but it is always worthwhile to discuss its limitations.

- ✓ We see these rectangles on the middle and right panels over Greenland:

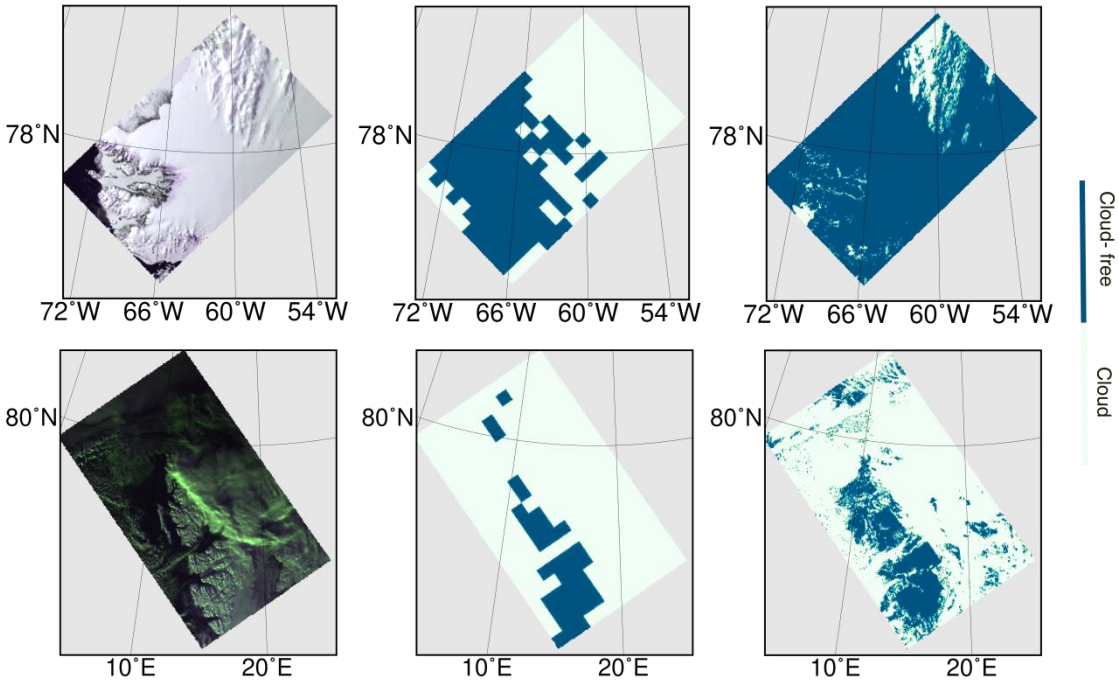
  Middle panel: In the middle panel, these 2 rectangles are indicators of the blocks with low Pearson Correlation Coefficient (PCC smaller than 0.4). As we can see from the corresponding RGB image, the locations of these two rectangles are over/near open water. We already discussed this shortcoming of the PCC analysis at lines 322-325 (in the original version of this manuscript): "Small PCC values may be caused by rapid surface change, high aerosol load or lack of recognizable spatial pattern, which is often the case over homogeneous snow covered surface". Therefore the lack of geo-physical patterns and structure could be a potential reason in this case. We also mentioned at lines 318-319 in the original version of manuscript "The combination of these two constraints is necessary because neither PCC analysis nor the reflectance part of 3.7 μ m is adequate on its own for accurate cloud detection".

  Right panel: However, using the reflectance of 3.7 μ m compensates the limitation of the PCC criterion as we explained at line 393-396. As we can see in the final results of this image (in Fig. 12 in the results section) these two rectangles over open water disappeared. Therefore the right panel in Fig. 5 (in original version of the manuscript) should not have these rectangles as well.

  We thank the referee for this information. We realized that we have used erroneously an outdated version of this picture (which was not created using the right running version of the algorithm).

  We have updated the right panel in Fig. 5 (with an improved color scheme (comment from the reviewer #1)). Please see below.

- ✓ Modification: Line 721 Fig. 5.



Reviewer #2: L398 You say the computing time is higher. How higher? Please give an estimate.

- ✓ The run-time of one scene is 30 minutes on a state-of-the-art computer system.

Reviewer #2: L428 I agree with the other reviewer, 45 minute time difference seem to me quite large for validation purposes. Maybe you should introduce a filtering?

- ✓ In preparation of the article, we performed a comprehensive review of previous works to define the optimal maximum time difference. However, statements in literature strongly vary: 10 min (Werkmeister et al., 2015), 15 min (Musial et al., 2014), 1 h (Dybbroe et al., 2005) and 4 h (Meerkötter et al., 2004), obviously all for different kinds of meteorological conditions. The investigation and results in the previous publications indicate that temporal difference in validation of satellite retrievals against SYNOP may vary based on meteorological conditions. Therefore we also tried to check meteorological conditions in our study.
- ✓ First, we checked to see which fraction of our results could be affected by longer temporal difference? We found that in 30% of our validation scenarios time difference exceeds half an hour. We tried to answer your question in these 30%. Do clouds change in the interval of 45 minutes? Or clouds could be almost stable due to the unique Arctic environment compared to lower latitudes? To answer this question we have performed 2 investigations:

  1) We checked SYNOP data before and after validation time to see whether cloud changes could be observed during this time? Available to us were SYNOP data, i.e. cloud fraction every 3 hours over Svalbard and every 1 hour over Greenland. We found that only in 12% of our validation scenarios cloud fraction changed (±1 or ±2 oktas) within 2 hours (for Greenland) and – 6 hours (for Svalbard).

  2) Another perspective on this topic is to ask how fast cloud can travel during our validation time window? We have checked the wind speed over Svalbard to answer the question. How strong the wind needs to be to move clouds out of or into in our validation area which is defined as a circle with the radius of 20 km around each SYNOP station? If we assume that the cloud is in the middle of this circle, we need at least a wind speed of 7.5 m/s to move clouds out of or into this circle in 45 minutes. Based on information from the Norwegian Meteorological Institute which provides an average of hourly wind speed, we see that the average wind speed over or close to the selected SYNOP stations at the closest stations during satellite overpass time is usually very low. For example below 3m/s and only in one scenario, it exceeds 7.5m/s slightly.

- ✓ Second, choosing a smaller temporal difference like for instance 0.5h would limit the number of observations and introduce a sampling error. For example, by filtering the validation dataset with respect to 0.5h temporal difference, 30% of the validation data will be lost. Therefore, to have a trade-off between good statistics/sampling and representativeness, we decided to keep the temporal interval of 45 minutes.
- ✓ However, temporal difference between satellite and SYNOP measurements is one of several sources of uncertainty (different viewing perspective, different spatial footprint etc.) which affect validation results.

We have updated the text to address this comment and explain the uncertainty which originates from time difference.

- ✓ Modification: Line 465-477: To define the optimal maximum temporal difference between SYNOP and satellite data, other comparable validation activities used different temporal intervals like 10 min (Werkmeister et al., 2015), 15 min (Musial et al., 2014), 1 h (Dybbroe et al., 2005) and 4 h (Meerkötter et al., 2004). The investigation and results in the previous publications indicate that temporal difference in validation of satellite retrievals against SYNOP depend on meteorological conditions. Allowing only a small temporal difference between measurement datasets (here: SYNOP and ASCIA) ensures an optimal temporal overall but can introduce a significant sampling error due to the small number of scenes for validation (Bojanowski et al., 2014). According to Bojanowski et al. (2014) a temporal difference of 90 min to compare with SYNOP measurements at temporal resolution of 3 h minimizes the sampling error (Bojanowski et al., 2014). However, potential longer temporal difference will introduce an error which should be considered along other sources of uncertainty (different viewing perspective, different spatial footprint and etc.). In this study, the maximum allowed temporal difference between the ASCIA retrievals and SYNOP measurements is less than ±20 minutes in most cases and generally does not exceed ± 45 minutes.
- ✓ Line 558: Bojanowski, J., Stöckli, R., Tetzlaff, A., and Kunz, H.: The Impact of Time Difference between Satellite Overpass and Ground Observation on Cloud Cover Performance Statistics, Remote Sensing, 6, 12 866–12 884, https://doi.org/10.3390/rs61212866, 2014.
- ✓ Line 575: Dybbroe, A., Karlsson, K.-G., Thoss, A., NWCSAF AVHRR cloud detection and analysis using dynamic thresholds and radiative transfer modeling. Part II: Tuning and validation. J. Appl. Meteorol., 44, 55–71, 2005.
- ✓ Line 640: Meerkötter, R., König, C., Bissolli, P., Gesell, G., Mannstein, H., A 14-year European cloud climatology from NOAA/AVHRR data in comparison to surface observations. Geophys. Res. Lett., 31, doi:10.1029/2004GL020098, 2004.

Reviewer #2: Sections Results and Validation could be compressed in one section, as even when presenting the results you do some qualitative validation against other cloud products.

- ✓ Done.

Reviewer #2: L454 Could you please show part of the evaluation against AERONET? AS the latter is a well-known reference for every reader, the validation against it deserves more than 2 lines of text. Also, which version are you using? And why L1.5 instead of L2.0?

- ✓ We agree and modified the text as below. Unfortunately it was not possible to perform further statistical analysis (like we did for SYNOP) in validation against AERONET because the comparison is pixel-based and does not include a circular area around the station where we could estimate cloud fraction.

  The reason for selecting level 1.5 instead of level 2.0 is that, level 1.5 data are cloud screened but level 2.0 data are quality assured. This means if we use level 1.5 data and check whether we have aerosol

measurements or not (having aerosol measurements means cloud-free condition) we could have information of cloudiness. But, in level 2.0 data, missing aerosol data could also be due to low quality of measurements.

- ✓ Modification: Line: 504-508: We also validated ASCIA cloud identification results with AERONET level 1.5 measurements, which are cloud screened. The procedure for this validation takes place in 2 steps: (1) covering AERONET observed AOT to a cloud flag (AOT is provided in AERONET only in cloud-free conditions); (2) Validation of ASCIA with AERONET cloud flag. In 86.1 % of 36 studied scenes over Svalbard, both ASCIA and AERONET confirm the presence of clouds.

- ✓ Reviewer #2: Technical comments L151 The SLSTR revisit time is 1.9 day at the equator with one satellite and 0.9 day with two satellites, not single/dual view.
- ✓ Done.
- ✓ Modification: lines 166-167: This yields global revisit times of 1.9 days at the equator with two satellites and 0.9 day with one satellite.

Reviewer #2: Table 3. The title of the second column should be something like "Test"

- ✓ Done.
- ✓ Modification: line 712.

Surface Type	Test	Description
Water	$R_{0.87} < 11\%$ & NDSI≥0.4	MODIS snow and ice mapping ATBD (Hall et al., 2001)
Sea-ice	$R_{0.87} > 11\%$ & NDSI≥0.4	
Land	$R_{3.7}$>0.04 & $R_{0.66} < 0.2$ || NDSI<0.4	Allen et al.,1999
Snow	$R_{3.7}$≤0.04	Allen et al., 1999

Reviewer #2: Figures 1-2, fig. from 5 to 13 and fig. 16 should indeed be larger.

- ✓ Done.

Reviewer #2: L339-347 please simplify these lines. Throughout your manuscript sentences are often too long, but here they really affect the readability. Simplify the lines here and maybe add more information on the caption of the figures (for example the exact coordinates of the corners)

- ✓ Done.
- ✓ Modification: line 365-370: A representative example of the block level (25×25 km^2) and scene level (1×1 km^2) results of ASCIA applied to AATSR observations is shown in Fig. 5. This example was selected to show the performance of ASCIA in presence of different surface conditions: 1) one scene is over a combination of fairly homogenous snow cover, land, ocean, sea-ice and cloud scene at north-west of

Greenland taken on the 18 May 2008; 2) another example is over a surface with highly variable topography over Svalbard with relatively higher solar zenith angle (>80°) on the 1 March 2008.

- ✓ Line 707-711: Figure 5. Examples of the results of ASCIA on AATSR observations on the scenes over Greenland (upper panels) between (75°N, 48°W), (75°N, 75°W), (81°N, 48°W), (81°N, 75°W), taken on the 18 May 2008 and Svalbard (lower panels) within (75°N, 4°E), (75°N, 32°E), (81°N, 4°E), (81°N, 32°E) (lower panels), on the 1 March 2008, Left panels: RGB images, middle panels: Cloud detection at block level (25×25 km^2), right panels: cloud detection at scene level.

Reviewer #2: L16 reflection at 3.7 um

- ✓ Done.
- ✓ Modification: line 16: Subsequently, the reflection at 3.7 μm is used for accurate cloud identification at…

Reviewer #2: L18 e,g. e.g.

- ✓ Done.
- ✓ Modification: line 17: The ASCIA data product has been validated by comparison with independent observations e.g. surface synoptic observations (SYNOP)

Reviewer #2: L32 Though the attribution of the origins of this phenomenon

- ✓ Done.
- ✓ Modification: line 33: Though the attribution of the origins of this phenomenon is controversially discussed.

Reviewer #2: L76 the aim is

- ✓ Done.
- ✓ Modification: line 82: … where the aim is to simultaneously retrieve aerosol and surface properties.

Reviewer #2: L104 it is also planned to apply it to the observations acquired by SLSTR

- ✓ Done.
- ✓ Modification: line 115-116: It is also planned to apply it to the observations acquired by SLSTR onboard Sentinel-3A and Sentinel-3B launched in 2016 and 2018 respectively which provide continuity of AATSR observations.

Reviewer #2: L123 In the upper right

- ✓ Done.
- ✓ Modification: line137: For example, in the upper right panel…

Reviewer #2: L131 Each scene is

- ✓ Done.
- ✓ Modification: line 145: Each scene is imaged twice.

Reviewer #2: L142 These algorithms are typically not optimized

- ✓ Done. We deleted this line.

Reviewer #2: L168 For example, they are almost absent in the central parts

- ✓ Done.
- ✓ Modification: line 188-189: For example, there is almost no observation in the central parts of the Arctic Circle as is shown in Fig. 3.

- ✓ Reviewer #2: L187 AERONET is . . .
- ✓ Done.
- ✓ Modification: line 208: AERONET is a network of approximately 700 ground-based sun photometers established by National Aeronautics and Space Administration (NASA) and PHOtométrie pour le Traitement Opérationnel de Normalisation Satellitaire (PHOTONS).

Reviewer #2: L206 Then the PCC can be written as

- ✓ Done.
- ✓ Modification: line 226-228: To describe the computational procedure developed, we assume x, y to be two random variables, then the PCC can be written as a function of the covariance…

- ✓ Reviewer #2: L207 function of the covariance
- ✓ Done.
- ✓ Modification: line 228: … then the PCC can be written as a function of the covariance…

- ✓ Reviewer #2: L241 no new line
- ✓ Modification: line 264.
- ✓ Done.

Reviewer #2: L254 Here you should compact everything in one sentence

- ✓ Done.
- ✓ Modification: line 278-281: where $R_{3.7}$ is the reflectance i.e. the ratio of scattered radiance to incident solar radiance; L measured radiance at 3.7 μm, $B_{3.7}(T_{11})$ the Planck function radiance (the contribution from thermal emission at 3.7 μm) T_{11} measurements at 11 μm; $F_{3.7,0}$ the solar constant at 3.7 μm and μ_0 the cosine of solar zenith angle.

Reviewer #2: L285 AATSR provides

- ✓ Done.
- ✓ Modification: line 309: AATSR provides more data over higher latitudes, …

Reviewer #2: L310 found that a PCC of 06

- ✓ Done.
- ✓ Modification: line 334-336: …we defined a lower threshold for PCC of 0.4 over the Arctic region and found that a PCC of 0.6 is appropriate for middle latitudes based on a number of statistical analyses.

Reviewer #2: L319 ASCIA starts looking for remaining small cloud scenes within a block, i.e. scenes . . . (R3.7 >0.04)

- ✓ Done.
- ✓ Modification: line 346: ASCIA starts looking for remaining small cloud scenes within a block, i.e. scenes with $R_{3.7}$ …

Reviewer #2: L333 it is important to note that one scene, . . .

- ✓ Done. We changed this sentence.
- ✓ Modification: line 359: Although characterized as land, a scene may include soil, different types of vegetation cover or even melting snow.

Reviewer #2: L334 The latter mix with soil and becomes

- ✓ Done.
- ✓ Modification line 360: The latter mix with soil and becomes dark enough to be filtered out from the snow class. Sea-ice is distinguished from water on the basis of its greater brightness…

Reviewer #2: L339 a representative example

- ✓ Done.
- ✓ Modification: line 365: A representative example of the block level (25×25 km^2) and scene level (1×1 km^2) results of ASCIA…

Reviewer #2: L373 cloud free scene which ISTO failed to detect but are correctly labeled by ASCIA.

- ✓ Done.
- ✓ Modification: line 406: The reddish scenes show cloud free cases, which ISTO fails to detect, but are correctly labeled by ASCIA as cloud free.

Reviewer #2: L393 Both the ESA and ISTO

- ✓ Done.
- ✓ Modification line 426: Both the ESA and ISTO cloud products showed good results for this case with the exception of thin cloud scenes which are falsely labeled as clear snow by ISTO.

Reviewer #2: L447 would be expected from SYNOP

- ✓ Done.
- ✓ Modification: line 497: As discussed earlier an error of ±1 to ±2 okta would be expected as the accepted accuracy range from SYNOP cloud cover values due to man-made nature of its observation and viewing conditions.

A cloud identification algorithm over the Arctic for use with AATSR/SLSTR measurements

Soheila Jafariserajehlou[1], Linlu Mei[1], Marco Vountas[1], Vladimir Rozanov[1], John P. Burrows[1], Rainer Hollmann[2]

[1]Institute of Environmental Physics, University of Bremen, Otto-Hahn-Allee 1, Bremen, 28359, Germany
[2]DWD – Deutscher Wetterdienst, Frankfurter Straße 135, 63067 Offenbach, Germany

Correspondence to: Soheila Jafariserajehlou (jafari@iup.physik.uni-bremen.de)

Abstract. The accurate identification of the presence of cloud in the ground scenes observed by remote sensing satellites is an end in itself. The~~Our~~ lack of knowledge of cloud at high latitudes increases the error and uncertainty in the evaluation and assessment of the changing impact of aerosol and cloud in a warming climate. A prerequisite for the accurate retrieval of Aerosol Optical Thickness, AOT, is the knowledge of the presence of cloud in a ground scene.

In this study observations of the up welling radiance in the visible (VIS), near infrared (NIR), shortwave infrared (SWIR), and the thermal infrared (TIR), coupled with solar extraterrestrial irradiance are used to determine the reflectance. We have developed a new cloud identification algorithm for application to the reflectance observations of Advanced Along-Track Scanning Radiometer (AATSR) on European Space Agency (ESA)-Envisat and Sea and Land Surface Temperature Radiometer (SLSTR) onboard the ESA Copernicus Sentinel-3A and -3B. The resultant AATSR/SLSTR Cloud Identification Algorithm (ASCIA) ~~developed~~ addresses the requirements for the study AOT at high latitudes and utilizes time-series measurements. It is assumed that cloud free surfaces have unchanged or little changed patterns for a given sampling period, whereas cloudy or partly cloudy scenes show much higher variability in space and time. In this method, the Pearson Correlation Coefficient (PCC) parameter is used to measure the 'stability' of the atmosphere-surface system observed by satellites. The cloud free surface is classified by analyzing the PCC values at the block scale 25×25 km^2. Subsequently, the reflection at~~of~~ 3.7 µm is used for accurate cloud identification at ~~the~~ scene level: having areas of ~~e~~either 1×1 km^2 or 0.5×0.5 km^2. The ASCIA data product has been validated by comparison with independent observations e.~~,~~g. s~~S~~urface synoptic observations (SYNOP), the data from AErosol RObotic NETwork (AERONET) and the following satellite-products ~~from~~ i) The ESA standard cloud product from AATSR L2 nadir cloud flag~~,~~; ii) The product from a ~~one~~ method based on clear-snow spectral shape developed at IUP Bremen (Istomina et al., 2010), which we call, ISTO~~,~~; iii) the Moderate Resolution Imaging Spectroradiometer (MODIS) products. In comparison to ground based SYNOP measurements, we achieved a promising agreement better than 95 % and 83 % within ±2 and ±1 okta respectively. In general, ASCIA shows an improved performance in comparison to other algorithms applied to AATSR measurements for the~~cloud~~ identification of clouds in a ground scene observed at high latitudes.

1 Introduction

The large trends in warming over the Arctic in recent decades, has received much attention from the global and regional climate change research community (Wendisch et al., 2017; Cohen et al., 2014). ~~This study is part of research activities to meet the scientific objectives of Collaborative Research Centers, CRC/Transregio 172 "ArctiC Amplification: Climate Relevant Atmospheric and SurfaCe Processes, and Feedback Mechanisms (AC)³ (Wendisch et al., 2017).~~ A number of studies using global observations and climate models confirm this phenomenon, called Arctic Amplification and provide evidence that its impacts extends ~~grows~~ beyond the Arctic (Kim et al., 2017; Cohen et al., 2014). Though~~,~~ the attribution of the origins of this phenomenon is controversially discussed (Serreze et al., 2011; Pithan et al., 2014), cloud cover is well-known to play a role in the Arctic surface-atmosphere radiation balance (Kellogg et al., 1975; Curry et al., 1996). The accurate identification of Arctic clouds in the ground scenes of remote sensing measurements made from space is therefore of intrinsic importance. However, cloud identification and screening over the Arctic is a challenging task. Since~~,~~ all developed cloud detection methods encounter many obstacles originating from the unique atmosphere and surface conditions in the Arctic (Curry et al., 1996). The Arctic clouds are mostly optically thin and low with no remarkable contrast in the commonly used visible or thermal or microwave measurements to the underlying surface covered with highly reflecting snow and ice. For example, snow/ice are also like clouds cold, the lack of strong thermal contrast is a limitation in the retrieval of clouds in the thermal infrared ~~-~~(Rossow and Garder 1993; Curry et al., 1996).

In addition to the importance of clouds to Arctic Amplification, errors in the identification of cloud within a ground~~in~~ scene are also one of the major sources of error in retrievals of a variety of data products for both satellite and ground-based measurements at high latitude. For example~~instance~~, the interference of cloud contamination in the Aerosol Optical Thickness (AOT) retrieved by passive satellite remote sensing is a well-known issue (Shi et al., 2014; Várnai and Marshak, 2015; Christensen et al., 2017; Arola et al., 2017). This limits the reliability and usefulness of the AOT products in the assessment of the direct/ indirect impact of aerosols in Earth's energy balance in particular over the Arctic. To avoid the uncertainty included in AOT products due to significant misclassification of heavy aerosol load as ~~by~~ thin clouds (which have similar reflectance properties) the development of an adequate cloud identification algorithm is a prerequisite (Martins et al., 2002; Remer et al., 2012; Wind et al., 2016; Mei et al., 2017a, 2017b; Christensen et al., 2017).

One recent approach to detect cloud-free snow and ice ~~for aerosol retrieval~~ over high latitudes used the spectral shape of clear snow, ISTO (Istomina et al., 2010). The latter analyses the spectral behavior of each ground scene and identifies clear snow or ice scenes from Advanced Along-Track Scanning Radiometer (AATSR) measurements. Thresholds of the reflectance were empirically determined in seven spectral channels from the VIS to TIR. Defining a reliable threshold which can guarantee a successful separation of cloud and cloud-free regions for the wide range of atmospheric conditions and surface types is a challenging task. This is because of the similarity between spectral reflectance of cloud and snow-ice (Lyapustin et al., 2008). In spite of progress made by this approach, adequate discrimination of thin cloud above ice or snow is an inherent limitation of such threshold based techniques.

The European Space Agency (ESA) standard cloud product from AATSR is another example of an existing cloud data product over the Arctic. This operational cloud mask is called the Synthesis of ATSR Data Into Sea-Surface Temperature (SADIST) and is based on the latitudinal thresholds for various cloud types (Ghent et al., 2017). ~~The~~

SADIST was initially developed for cloud screening over the ocean (Zavody et al., 2000). Birks et al. (2007) modified this method to apply it over land. Later, Kolmonen et al. (2013) reported that the cloud flags included in AATSR product are noticeably restricted and using this cloud product results in aerosol episodes not being observed. SADIST is known to misclassify ice, cloud and open ocean in Polar Regions. Bulgin et al. (2015) developed a Bayesian approach in ESA's Climate Change Initiative (CCI) project to overcome this limitation (Hollmann et al., 2013). Sobrino et al. (2016) reviewed different cloud clearing methods including the AATSR operational cloud mask in the framework of Synergistic Use of The Sentinel Missions For Estimating And Monitoring Land Surface Temperature (SEN4LST) project. Theyand highlighted the potential uncertainty in different versions of this product, which result in these errors being propagated in subsequent data products. For example, the AATSR operational cloud mask falsely detects cloud in ~ 16 % of the observations. This is attributed to the flagging of land features (such as rivers) incorrectly as cloud (see Sobrino et al., 2013).

To avoid the uncertainty arising from the similarity of spectral characteristics of snow, ice and clouds, we decided to develop an algorithm based on a different strategy, namely the use of time series measurements. The use of abrupt changes of TOA reflectance in time with the aim of cloud identification has been reported previously (Gómez-Chova et al., 2017; Lyapustin et al., 2008). An early example of this idea was proposed for low to middle latitudes by Rossow and Garder (1993) in the International Satellite Cloud Climatology Project (ISCCP). This method later evolved as a part of MultiAngle Implementation of Atmospheric Correction (MAIAC) algorithm (Lyapusitn et al., 2008)., which MAIAC is mainly designed for use with observations over land (low to middle latitudes), where the aim is to simultaneously retrieve aerosol and surface properties. However, it has also been utilized by another study to identify snow grain size over Greenland (Lyapustin et al., 2009). Though, further optimization for the Arctic region is required and reported, a better performance in comparison to Moderate Resolution Imaging Spectroradiometer (MODIS) cloud mask is reported by Lyapustin et al. (2009).

The central assumption used in these algorithms for cloud identification, is that clear-sky reflectance is different to that of clouds, which exhibit, in comparison, a high variation as a function of time (Lyapustin et al., 2008; Gómez-Chova et al., 2017). Knowledge of cloud-free scenes within a given time period, is achieved from knowledge of the variability of the measured TOA reflectance. Covariance analysis is used to estimate the spatial coherence. This has a long history in remote sensing studies using time series measurements (Leese et al., 1970; Lyapustin et al., 2008). The covariance computation assumes changes in the textural patterns of the observed scene, which originate from natural and man-made features such as topography, lakes or urban areas (Lyapustin et al., 2008). The use of the covariance analysis, which accounts for geometrical structures, minimizes issues originating from illumination variation and results in the same algorithm being applicable over both dark and bright surfaces (Lyapustin et al., 2008). For these reasons we decided to use Pearson Correlation Coefficient (PCC) as a function of covariance value for cloud detection over the Arctic. However, Lyapustin et al. (2008) reported that in spite of relative good performance, the covariance itself is not alone adequate for cloud identification in the case of homogenous surfaces or thin clouds. Therefore, we decided to use a combination of a PCC analysis and the reflectance of solar radiation at 3.7 μm. The latter utilizes the contrast between cloud and underlying surface making it possible to distinguish cloud-free snow and ice.

Another argument in favor of the use of time series analysis is the availability of multiple images by the AATSR and Sea and Land Surface Temperature Radiometer (SLSTR) sensor over the Arctic. For AATSR the revisit time of 3-4 days over mid-latitudes (Kolmonen et al., 2016) with more frequent at higher latitude which increase to 2 days over the Arctic (Soliman et al., 2012; Mei et al., 2013). In addition to multiple imagery over the Arctic, the shorter time interval between satellites over-passes over the same scene provides images with less variability in the observed cloud-free areas which the algorithm looks for. For the two SLSTR, the revisit time ~~it~~ is 0.9 days at the equator (Coppo et al., 2010) and this time becomes even shorter at ~~with these values increasing at~~ higher latitudes due to orbital convergence.

The AATSR/SLSTR Cloud Identification Algorithm (ASCIA) has been developed as a part of research activities to meet the scientific objectives of Collaborative Research Centers, CRC/Transregio 172 "ArctiC Amplification: Climate Relevant Atmospheric and SurfaCe Processes, and Feedback Mechanisms (AC)³ project (Wendisch et al., 2017). ~~for use in the (AC)³ project (Wendisch et al., 2017)~~. The project aims to identify, investigate and evaluate parameters and feedback mechanisms which contribute to Arctic A~~a~~mplification (Wendisch et al., 2017). Consequently, a long-term data record of AOT and cloud is required. It is planned to use ~~the~~ ASCIA to identify cloud free scenes for AOT retrieval. It is also planned to ~~be~~ apply it~~ied~~ to the observations acquired by ~~the~~ SLSTR onboard Sentinel-3A and Sentinel-3B launched in 2016 and 2018 respectively which provide continuity of AATSR observations.

A full description of ~~the~~ this new cloud identification and its application to AATSR data is presented in the following sections of this manuscript. First, a brief data description is presented in Sect. 2. The theory and methodology, used in our new ASCIA, are discussed in detail in Sect. 3 and 4. We evaluated the performance of ~~the~~ ASCIA by comparison of the cloud identification with i)~~that of~~ the ESA standard cloud product for AATSR level2 nadir cloud flag; ii) the data obtained by applying ISTO to AATSR data; ~~while ASCIA is also applied to AATSR nadir observations, those obtained by applying ISTO,~~ iii) the MODIS cloud mask;~~,~~ iv) the Surface synoptic observations (SYNOP);~~ and~~ vi) the AErosol RObotic NETwork (AERONET). The results of the comparisons with these five different ~~source~~ source of cloud data are reported in Sect. 5~~6~~. A discussion and set of conclusions, drawn from the study, are presented in Sect. 6~~7~~.

2 Instruments and Data

2.1 AATSR data

The AATSR sensor flown on board polar orbiting Envisat was primarily designed for measuring Sea Surface Temperature (SST) with accuracy higher than 0.3 k. ~~after~~ As the successor of ATSR-1 and ATSR-2 on European Remote Sensing-1, ERS-1 and ERS-2, ~~(http://envisat.esa.int/handbooks/aatsr/CNTR.html). The~~ AATSR delivered data from March 2002 until Envisat failed in 2012 (http://envisat.esa.int/handbooks/aatsr/CNTR.html). The unique design of spectral coverage of AATSR enabled this sensor to measure reflected and emitted radiances in the VIS, (0.55 μm, 0.66 μm), NIR (0.87 μm, 1.6 μm) and three TIR channels (3.7 μm, 10.85 μm, 12.00 μm) with spatial resolution of 1×1 km^2 at nadir view and swath wide of 512 km. In Fig. 1 one example of the AATSR image over Svalbard is shown. It comprises three different wavelengths to highlight the different information, which are

retrieved~~one can gain~~ from the wide spectral coverage of this instrument. For example, in the upper right panel of~~in~~ Fig. 1 the large change ~~drop of~~ in reflectance over snow/ice created a notable contrast between the cloud and the underlying surface at this wavelength as compared~~in comparison~~ to that found from the VIS channels used in the R(0.66 μm) G(0.87 μm) B(0.55 μm) image. A similar significant difference~~separation~~ of snow/ice and cloud is observed in the reflectance at 3.7 μm shown in the lower left panel in Fig. 1. However, at the longer wavelength of 11 μm thin cloud patterns appear in the south-western scenes close to and above Svalbard, which have small signatures in the shorter wavelength. Combining the information from the different channels in an appropriate way enables the presence of cloud in the ground scenes to be accurately identified.

The conical imaging geometry of AATSR yields the dual viewing capability of this sensor. Each scene is~~was~~ imaged twice. The first measurement of the ground scenes is in the forward direction at a viewing angle of 55°. The second occurs 150 sec later at a near-nadir viewing angle. This capability is a design feature of AATSR to deliver an optimal and accurate atmospheric correction and thereby invert an accurate surface reflectance. The two views theoretically yield independent information about atmosphere and the surface to be retrieved. (http://envisat.esa.int/handbooks/aatsr/CNTR.html). The dual view approach intrinsically provides more information than the single view for the study of surfaces with complex reflectance characteristics, such as snow and ice (Istomina, 2012).

~~Examples of AOT algorithms applied to AATSR data are as follows: the AATSR Dual View algorithm (ADV) which was initially proposed by Veefkind et al. (1999) and AATSR single view algorithm (ASV) by Veefkind et al. (1998), the Swansea University (SU) algorithm (North et al., 1999) and Oxford RAL Aerosol and Cloud retrieval (ORAC) algorithm (Thomas et al., 2009). These algorithms typically not optimized for the retrieval of AOT at high latitudes. As the first task in delivering an algorithm, which delivers AOT at high latitude, t~~The ~~new~~ ASCIA has been applied to AATSR measurements to identify cloud and cloud free ground scenes. ~~has been applied to AATSR measurements~~.

2.2 SLSTR data

~~The~~ SLSTR on-board Sentinel-3A was launched on the 16th of February in 2016 as the successor to~~of~~ AATSR ~~series~~ to provide the continuity of long term SST measurements. The Sentinel-3B satellite, which contains an identical payload, was also launched by a Rockot/Breeze-KM launch vehicle from the Plesetsk Cosmodrome in northern Russia, on the 25th of April 2018. The design of the SLSTR instrument has some significant improvements with respect to ATSR (Coppo et al., 2010). For example, the swath width of single view and dual view was increased from 500 km to 1420 km and 750 km respectively. This yields global revisit times of 1.9 days at the equator with two satellites ~~for the dual view~~and 0.9 ~~1~~day with one satellite~~for the single view~~. There are measurements of two additional channels in the SWIR, at the wavelengths of 1.37 μm and 2.25 μm, which are used to provide more accurate cloud, cirrus and aerosol information and used to correct for atmospheric radiative transfer effects in the determination of surface reflectance (Coppo et al., 2010). The Fig. 2 upper right panel shows the use of the new 1.37 μm measurements to detect thin cirrus clouds, which are only weakly identified in reflectance at 3.7 μm shown in Fig. 2. The current design of ASCIA does not yet include 1.37 μm measurements. This is because~~As~~ the radiance

and TOA reflectance at this wavelength are not measured by AATSR. ~~and because of~~ In addition, SLSTR data at this wavelength currently have unresolved calibration issues. ~~in SLSTR data~~, ~~the current design of ASCIA does not yet include 1.37 μm measurements. In addition, water vapor absorption above and within clouds is considered as an obstacle in using this channel for cirrus detection (Meyer et al., 2010)~~. Nevertheless, the use of the measurements at this wavelength in thin cirrus detection should improve the performance of ASCIA in the future. SLSTR also has a higher spatial resolution of 0.5×0.5 km^2 in the VIS and SWIR measurements and two channels dedicated to fire detection (Coppo et al., 2010). The use of the observations from SLSTR and AATSR enables a long-term time series of clouds and aerosol parameters including AOT over the Arctic to be derived. However, there is a ~4 years gap between the failure of AATSR and the launch of SLSTR. To fill this gap, we will apply ASCIA also to the Advanced Very High Resolution Radiometer (AVHRR) sensor carried by National Oceanic and Atmospheric Administration (NOAA).

2.3 Data used in the cloud identification comparison studies

2.3.1 SYNOP

The SYNOP have been provided by World Meteorological Organization (WMO) for~~with~~ the purpose of mapping ~~large scale~~ weather information around the world. However, the availability of the data is limited in the Arctic ~~studies~~ due to the coverage of SYNOP stations in this region. For example, there~~re are almost~~is almost no observation ~~absent~~ in the central parts of the Arctic Circle as is shown in Fig. 3. The SYNOP measurements, which are made by an observer or, an ~~-~~automated device~~fixed stations~~ are available in a standardized layout of numerical code which is called FM-12 by WMO (1995). The SYNOP reports include a variety of meteorological parameters such as temperature, barometric pressure, visibility etc. as well as cloud amount which are observed at synoptic hours simultaneously throughout the globe. We used SYNOP cloud fraction, which have a temporal resolution of 1-3 hours, to evaluate the performance of our new developed ASCIA over the Arctic region.

The use of SYNOP measurements to validate a cloud identification algorithm, or for that matter the cloud predicted by a climate model, the fact that the SYNOP cloud fraction is reported using the~~in~~ okta scale, has to be appropriately taken into account. ~~which ranges from 0 (completely clear sky) to 8 (completely obscured by clouds) has to be appropriately taken into account.~~ Converting discrete okta values, which ranges from 0 (completely clear sky) to 8 (completely obscured by clouds) to continuous percentage ones has been done in different ways by climatologists. A common assumption is that 1 okta equals 12.5 % of cloud coverage (Boers et al., 2010; Kotarba, 2009). For use in this study it was necessary to make an estimate of the error or uncertainty in the okta in measurements. It is assumed that the man-made nature of cloudiness okta estimation have errors of ±1 okta and even larger values of ±2 okta for~~in~~ the non 0 or 8 okta situations (Boers et al., 2010; Werkmeister et al., 2015). Boers et al. (2010) suggested defining a larger range of 18.75 % for 1 okta instead of commonly used value of 12.5 %. We used this approach and defined percentage of cloud values for each okta, which are given in Table 1. More details about validation procedure are provided in Sect. 6.

2.3.2 AERONET

The AERONET is a network of approximately 700 ground-based sun photometers established by National Aeronautics and Space Administration (NASA) and PHOtométrie pour le Traitement Opérationnel de Normalisation Satellitaire (PHOTONS). This globally distributed network aims to provide long-term and continuous measurements of AOT, inversion products and perceptible water in diverse aerosol regimes (Holben et al., 1998). The high temporal resolution of 15 minutes for these data, expected low accuracy of ~ 0.01 to 0.021 (Eck et al., 1999) as well as readily accessible public domain database provides a suitable dataset for aerosol research and characterization.

AERONET data are categorized and available in 3 levels: Level 1.0 (unscreened), Level 1.5 (cloud screened and quality controlled) and level 2.0 (quality assured). The data used in this work are selected from Level 1.5 to validate cloud identification results from newly developed ASCIA. More details of validation procedure are discussed in Sect. 6.

3 Theoretical background

3.1 Pearson Correlation Coefficient (PCC)

The PCC was proposed by Karl Pearson (1896) and is used in this study as an indicator of the correlation between sequential AATSR measurements. The PCC is also known as the Pearson Product-Moment Correlation Coefficient (PPMCC). It is a standard dimensionless statistical parameter commonly used to measure the strength and direction of the linear association between a pair of variables (Benesty et al., 2009). This parameter has extensively been used in many studies which pursue pattern analysis and recognition.

Our use of the PCC analysis is to separates the surface reflectance at a given viewing angle, which is stable over short time periods, from the cloud reflectance, which is highly variable over a short time period. To describe the computational procedure developed here, welet assume x, y to be as two random variables, then the PCC can be written as a function of the covariance of x and y which is normalized by square root of their variances (Rodgers et al., 1988; Benesty et al., 2009):

$$PCC = \frac{COV(x,y)}{\sigma_x \sigma_y}, \quad (1)$$

where COV(x,y) is the covariance of variables and σ is the root-mean-square variations of each random variables (Rodgers et al., 1988; Benesty et al., 2009):

$$COV(x,y) = \frac{1}{N^2}\sum\sum(x_i - \bar{x})(y_i - \bar{y}) \ and \ \sigma_x^2 = \frac{1}{N^2}\sum_{i=1}^{N}(x_i - \bar{x})^2, \quad (2)$$

$$PCC = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{(\sum(x_i - \bar{x})^2 \ \sum(y_i - \bar{y})^2)^{1/2}}, \quad (3)$$

where $\bar{x}$ and $\bar{y}$ are the mean value of x and y variables respectively. The correlation coefficient parameter has values between -1 and +1 (Rodgers et al., 1988). The PCC values were prepared in this study. The association between the two variables is stronger if the absolute value is closer to 1, whereas if two variables are independent or in another

word “uncorrelated” PCC value will become 0 (Benesty et al., 2009). As a consequence of the above the PCC values computed between several data pairs for ground scenes of the same area at different times provide an indication of whether the scene is cloud covered or free of clouds.

For this aim, the use of all seven channels (0.55 µm, 0.66 µm, 0.87 µm, 1.6 µm, 3.7 µm, 11 and 12 µm) was investigated. The visible channels (0.55 µm, 0.66 µm) on their own are not optimal to separate cloud free form cloudy scenes, in particular for thin clouds. The SWIR and TIR such as 1.6 µm and beyond, where liquid water and ice absorb provide useful information. There is a large reduction of reflectance between clear snow/ice as compared to clouds between 0.87 µm and 1.6 µm (Kokhanovsky, 2006). Our routine takes advantage of this contrast through the PCC calculation. ~~One major contributors of error in aerosol retrieval is misclassifying heavy aerosol loads with clouds~~ One of the major contributors to error in aerosol retrievals is misclassification of heavy aerosol loads as cloud. Using 1.6 µm reflectance which is less affected by aerosols than visible wavelengths addresses in part this issue (Lyapustin et al., 2008).

A second question in PCC analysis (after wavelength selection) is definition of the optimal size of the block of ground scene for PCC calculation. In early version of current algorithm, we set up 10×10 km^2 as the block size. Since~~,~~ aerosol retrieval would be carry out with the same spatial resolution. However, our investigations and previous studies show that 10×10 km^2 is not sufficient to capture surface patterns. Thus, blocks of 25×25 km^2 area as proposed in previous studies (Lyapustin et al., 2008) were used. The implementation of PCC analysis as used in this study is discussed in more detail in Sect. 4.

3.2 Reflectance of 3.7 µm thermal infrared channel

The reflectance part of TIR Channels at 3.7 µm and 3.9 µm have been used in different studies to determine cloud properties such as cloud effective radius and thermodynamic phase of the cloud or to discriminate cloud and snow/ice covered surface (Meirink et al., 2016; Klüser et al., 2015; Musial et al., 2014; Khlopenkov, et al., 2007; Pavolonis et al., 2005; Rosenfeld et al., 2004; Spangenberg et al., 2001; Allen et al., 1990). The reason for the wide application of this channel in cloud identification methods is the difference in Single Scattering Albedo (SSA) at this band compared to shorter VIS and INR wavelengths, which in turn results from the significant sensitivity of SSA to thermodynamic phase and particle size of clouds (Platnick et al., 2008). For example, the scattering of liquid clouds, having small droplets, is relatively larger than absorption and the ratio of NIR/VIS reflectance approaches 1. But ~~while~~ in the case of large liquid droplets or ice particles, the absorption increases and this ratio is closer to zero (Platnick et al., 2008).

~~—~~In addition, cloud-free snow reflects at a relatively weak level in comparison to clouds at 3.7 µm channel (Derrien et al., 1993; Platnick et al., 2008). Therefore, the contrast due to different physical properties and radiance of snow/ice and cloud at 3.7 µm makes the use of this channel advantageous for the identification of clouds. During daytime, the measured Brightness Temperature (BT) at 3.7 µm is determined from the upwelling radiation which comprises both reflected or scattered solar radiation and the thermal emission from the surface (Musial et al., 2014). To use TOA reflectance at 3.7 µm, procedures are needed to account for and subtract the emission portion of measured BT at 3.7 µm wavelength (Allen et al., 1990). To achieve this goal independent information about the

surface TIR is needed. This is estimated from observations at 11 μm where absorption by water vapor and other trace gases is ~~very~~ small~~;~~, most objects ~~phenomena~~ in regions outside of the tropics can be treated~~behave~~ as blackbodies and the measured BT considered being in good agreement with ~~as the~~ real surface temperature (Istomina et al., 2010; Musial et al., 2014).

To do that, we use the method described in Meirink et al. (2016) and Musial et al. (2014), where the reflectance of 3.7 μm can be written as:

$$R_{3.7} = \frac{L_{3.7} - B_{3.7}(T_{11})}{\mu_0 F_{3.7,0} - B_{3.7}(T_{11})}, \quad (4)$$

where $R_{3.7}$ is the reflectance i.e. the ratio of scattered radiance to incident solar radiance; L ~~is~~ measured radiance at 3.7 μm~~,~~ ~~T~~ $B_{3.7}(T_{11})$ the Planck function radiance (the contribution from thermal emission at 3.7 μm) ~~is the Planck function radiance $B_{3.7}(T_{11})$ estimated from the temperature value obtained from the~~ T_{11} measurements at 11 μm; $F_{3.7,0}$ ~~is~~ the solar constant at 3.7 μm and ~~which is weighted by~~ μ_0 ~~as~~ the cosine of solar zenith angle.

Theoretical reflectance values at 3.7 μm band, computed by Allen et al. (1990) have been compared to satellite measurements at the same channel from Advanced Very High ReSolution Radiometer (AVHRR). The results of this work are summarized in Table 2. According to this study, the reflectance of liquid clouds primarily depends on droplet size and solar zenith angle, whereas for ice clouds, ice particle shape and size distribution are of great importance together with Cloud Optical Thickness (COT) and sun-satellite geometry. The observed reflectance is reported in a range of 0.08 to 0.36 for liquid clouds and 0.02 to 0.27 for ice clouds (Allen et al., 1990). Arking and Childs (1985) calculated 3.7 μm reflectance for ice clouds which varies between 0.01 to 0.30 for the COT of 0.1 to 100 and ice crystal effective radius of 2 μm to 32 μm, solar zenith angle of 60°. Spangenberg et al. (2001) reported a typical value of 0.04 to 0.4 for clouds. In the case of snow covered surface 3.7 μm reflectance is dependent on many factors including snow grain size, solar zenith angle, liquid water content, snow impurities and etc. Considering the snow grain size of 50 μm to 200 μm, with a solar zenith angle of 40° to 80°, the modeled values for snow reflectance varies between 0.005 and 0.025 at 3.7 μm (Allen et al., 1990). However, a range of 0.02 to 0.04 is observed from the satellite measurements over the same wavelength for snow cover. This difference between model calculations and measurements is explained by snow impurities (Allen et al., 1990). For land areas, the 3.7 μm reflectance is impacted by soil type, vegetation type, coverage and moisture content. An average value of 0.15 is derived for clear sky land scenes at 3.7 μm (Allen et al., 1990). In order to use the remarkable contrast between snow cover and clouds at 3.7 μm channel, two main issues have to be taken into account: 1) the interference between snow and ice-cloud values; 2) the interference between cloud and land reflectance. The latter is easily solved by using information from visible channels with 3.7 μm reflectance. This is because land scenes in Polar region are dark in comparison to cloud and snow. The first issue, discriminating ice clouds from snow is a challenging task. To detect ice clouds, we combined 3.7 μm reflectance with PCC analysis. A full description of this new method is given in Sect. 4.

4 Methodology

The ASCIA implementation is initiated by preparing a time series of data. A time span of one month for the ground scene was selected. Hagolle et al. (2015) indicated that in Sentinel-2 measurements with revisit time of 5 days, most of the given scenes would be observed cloud-free at least once a month. Consequently, we also assume that every scene of AATSR measurements, which have a higher revisit time of 3 days, will be cloud-free at least once a month.

Depending on the latitude and the time of year the number of downloaded data varies from 10 to 50 or more over the same scene. AATSR provides more data over higher latitudes, which increase in spring and summer ~~time~~ due to longer polar days and solar illumination. The AATSR L1b data are already provided as gridded and calibrated 1×1 km^2 scenes., ~~which~~ These include geo-location information interpolated from the tie point scenes which are equally distributed across a single AATSR image (http://envisat.esa.int/handbooks/aatsr/CNTR.html). Consequently~~Therefore~~, there is no necessity to re-grid them for the geo-referencing step. ~~which~~ This is considered as an advantage, because it ~~to~~ preserves the original reflectance value of each scene for following steps. However, the time series data are acquired by the satellite from different viewing geometries. To compute PCC values over the same areas from different days, ~~the~~ ASCIA selects~~looks for~~ the closest similar scenes using geo-location information provided in the data. The closest distance is often found to be within 0.006 degree and increases to 0.01 degree in the worst case and thus is considered to be of negligible significance. After this procedure to select the observations~~finding the same blocks over different dates and building blocks~~, ~~the~~ ASCIA comprises two main parts to identify the presence of cloud: i) a PCC analysis at 1.6 μm; ii) the use of thresholds for~~Applying thresholds on~~ reflectance of 3.7 μm channel.

In the first step, a PCC analysis for a block of ground scenes (25×25 km^2) is used to identify cloud and cloud-free blocks, which are assumed to have low and high PCC values respectively. The output of this step is a binary flag at the block level. This serves as input for the second step to produce at ground scene level (1×1 km^2 or 0.5×0.5 km^2 depending on spatial resolution of instrument) cloud identification, by using the knowledge of the reflectance of solar radiation at 3.7 μm channel. The combination of these two constraints is necessary because neither PCC analysis nor reflectance part of 3.7 μm channel is adequate on its own for accurate cloud detection. A high PCC value cannot guarantee the clearness of the whole block of scenes (Lyapustin et al., 2008) because some ground scenes may still contain clouds, which are not enough in number to decrease significantly the PCC value. This case occurs frequently over small or semitransparent clouds where the textural pattern of surface is still observable through the clouds (Lyapustin et al., 2008). Small PCC values may be caused by ~~a~~ rapid surface change, ~~or~~ high aerosol load or ~~the~~ lack of recognizable spatial pattern, which is often the case over homogenous snow covered surface (Lyapustin et al., 2008). A PCC value of 0.63 is suggested by Lyapustin et al. (2008) to separate cloud-free blocks over middle latitudes. Considering less surface patterns in a large area of the Arctic compared to lower latitudes, and our PCC analysis over both middle and high latitudes, we defined a lower threshold for~~of~~ PCC of 0.4 over the Arctic region and found that a~~the~~ PCC of 0.6 is appropriate for middle latitudes based on a~~large~~ number of statistical analyse~~is~~.

After computing the first binary cloud flag at block level using the last measurement and one previous image, ~~the~~ ASCIA keeps the result in a memory and repeats the procedure with second previous data. This procedure is iterated ~~and so on,~~ until the last measurement of the data series is involved ~~in PCC analysis~~. The final binary blocks are

imported through the second step to identify cloudy scenes based on thresholds defined for blocks with low and high PCC value differently. We note~~However, we would like to underline~~ that, the snow/ice reflectance at 3.7 μm channel (~0.005-0.025) has interference with those of ice clouds (0.01-0.3) at this wavelength. To avoid the uncertainty arising from this problem we defined the PCC analysis as a decision point of ASCIA requiring further optimized analysis:

(i) For the high PCC ≥ 0.4, the whole block is considered to be cloud free and then ~~the~~ ASCIA starts looking for remaining small cloud scenes within a block~~,~~ i.e. scenes with $R_{3.7}$ larger than the maximum value observed over snow at 3.7 μm: $R_{3.7} > 0.04$, (Allen et al., 1990).

(ii) For PCC < 0.4, the block is assumed to be cloudy; ASCIA removes all scenes within the block and only keeps scenes which satisfy the $R_{3.7} < 0.015$ test. This threshold is equal or lower than the lowest observation of ice cloud reflectance at 3.7 μm (Allen et al., 1990).

In our method, PCC analysis constrains the procedure and the strict decision is only made within low PCC blocks. The loss of some clear scenes in low PCC blocks is an unavoidable side effect of using this strict criteria in particular over land scenes, having low PCC and high 3.7 μm reflectance values. However, ~~the~~ ASCIA detects the presence of thin cirrus cases with a relatively high confidence level. A schematic flowchart of ~~the~~ ASCIA is shown in Fig. 4, with the use of the two main constraints being highlighted. In addition to picking out clear scenes, a simple land classification procedure is undertaken in this step of ~~the~~ ASCIA. Snow/ice scenes are identified with low 3.7 μm reflection whereas land scenes with high reflection are classified with the aid of the darkness test over visible channels. The corresponding thresholds for land classification scheme are described in the Table 3.

~~It is important to note that if one scene, a~~Although characterized as land, a scene may include soil, different types of vegetation cover or even melting snow. The latter mix with soil and becomes dark enough to be filtered out from the snow class. Sea-ice is distinguished from water on the basis of its greater~~higher~~ brightness; one scene might be white enough to be considered as ice. However, melting or broken ice as well as new ice would not be labeled as ice. Snow over sea-ice is not distinguished from pure sea-ice and both of them are labeled as sea-ice. This also means that ice over land is also marked as snow as well as pure snow.

A representative example~~s~~ of the block level (25×25 km^2~~.~~) and scene level (1×1 km^2~~.~~) results of ~~the~~ ASCIA applied to~~on~~ AATSR observations is shown in Fig. 5. This example was selected to show the performance of ASCIA in presence of different surface conditions: 1) one scene is over a combination of fairly homogenous snow cover, land, ocean, sea-ice and cloud scene at north-west of Greenland taken on the 18 May 2008; 2) another example is over a surface with highly variable topography over Svalbard with relatively higher solar zenith angle (>80°) on the 1 March 2008. ~~on the scenes within the region over northwest of Greenland in spring time enclosed in the coordinates for four corners (75°N, 48°W), (75°N, 75°W), (81°N, 48°W), (81°N, 75°W) taken on the 18 May 2008 are shown in Fig. 5. This example selected to show the performance of ASCIA over a combination of fairly homogenous snow cover, land, ocean, sea ice and cloud.~~ As we discussed earlier, the ambiguity of the PCC analysis over homogeneous surfaces on the right and left sides of the AATSR scene in middle panel of Fig. 5, is ~~entirely~~ compensated in the right panel by using additional information from 3.7 μm channel. ~~Another example over a surface with highly variable topography in March with relatively higher solar zenith angle (>80°) is selected over~~

~~Svalbard enclosed in the coordinates for four corners (75°N, 4°E), (75°N, 32°E), (81°N, 4°E), (81°N, 32°E) taken on the 1 March 2008.~~

5 Results and validation

5.1 The comparison of ASCIA products with products from other algorithms using space-borne observation

In this study, we applied our recently developed ASCIA to identify cloud in the scenes using AATSR L1b (TOA reflectance) and SLSTR L1b gridded data. The input file to the process chain is one scene of AATSR L1b product the output comprises 5 classes of surface types including snow/ice, sea ice, water, cloud and land. The procedure of surface classification is explained in Sect. 4. The location and time of selected case studies are used to show that the identification of cloud by our new ASCIA is adequate. ~~In this regard, t~~The AATSR data are selected from several years starting from 2006, during strong Arctic haze episode, which originated predominantly from agricultural fires burning in Eastern Europe. The event has been reported in previously (Law et al., 2007). A second episode in 2008 is also considered for which validation data are available from SYNOP stations. Three~~One~~ months of data from ~~the targeted seasons spring, summer and autumn vis.~~ March, May, and July ~~respectively~~ have been acquired over Greenland and Svalbard to assess the performance of ~~the~~ ASCIA in a wide range of solar zenith angles (60°-85°), surface and atmospheric conditions observed at high latitudes. In order to take ~~take into account~~ various surface types in the Arctic into account, we selected case studies including, highly variable topography and fairly homogeneous snow cover, coast lines, land and ocean along snow and ice covered surface. The designed criteria for ~~the~~ ASCIA are optimized for ~~an over~~ various regions over~~of~~ the Arctic observed under~~in~~ different solar illumination conditions. Polar night and transition seasons at low light conditions are excluded from our retrievals. ~~with the exception of the dark winter period.~~ The results obtained are compared with i) the AATSR L2 nadir cloud flag and ii) those results obtained with ISTO (Istomina et al., 2010) and iii) MODIS.

As we discussed in Sect. 1, misclassification~~fying~~ of thin cirrus cloud with clear snow is reported to be~~as~~ an unresolved problem of ISTO approach. Two representative scenarios of this problem are illustrated in Fig. 6 and Fig. 7 over Greenland and Svalbard respectively in which thin cloud is detected as clear snow by the ISTO method whereas ASCIA confirmed the presence of cloud. ~~In fact, o~~Over ~~such~~ a homogenous surface such as~~like~~ Greenland, the second step of ~~the~~ ASCIA is decisive~~plays the main rule~~. ~~Since, t~~The lack of structural patterns on surface lead to low PCC values in the first step and consequently overestimation of cloudy scenes. However, the reflection part of 3.7 µm could help to label and bring back clear homogenous surface as cloud free snow in second step. The right panel in Fig. 6 and 7 shows the difference between the result of ASCIA and ISTO. In this panel, the dark blue scenes show clouds, which are not detected by ISTO but are by~~while~~ ~~the~~ ASCIA could identify them sufficiently. The~~On the other hand,~~ reddish scenes show cloud free cases, ~~scenes~~ which ISTO fails~~ed~~ to detect, ~~them~~ but are correctly labeled by ~~the~~ ASCIA ~~labeled them~~ as cloud free. ~~As we can see, i~~ In addition to the edge of clouds which are difficult to detect ~~specially~~ over snow and ice, there are ~~we have~~ a significant ~~remarkable~~ number of undetected cloud scenes in ISTO results which are identified successfully by ~~the~~ ASCIA. However, for the rest of these two scenes, the~~both of~~ two algorithms show good~~a promising~~ agreement.

The ESA cloud product from L2 data, ~~shows a significant~~ overestimate~~sion of~~ cloud~~,s~~ which leads to a loss of ~~missing~~ clear snow and ice scenes. The tendency of this product to flag clear scenes as cloud is also visible in Fig. 6, 7. The results in Fig. 8 show undetected clouds as another problem of AATSR level 2 cloud product, which happens frequently at~~in~~ high solar zenith angle~~winter time~~. To have a better understanding of this misclassification, we validated the AATSR L2 nadir cloud flag against SYNOP measurements and results are described in Sect. 6.

Poor~~The lack of good~~ performance for~~in~~ cases ~~winter time~~ over the Arctic with high solar zenith is observed in all of the results using ISTO method. Figure 8 is an example over Svalbard in March 2008. Over such a highly variable surface type, such as~~like~~ Svalbard, the reflection at~~part of~~ 3.7 µm can~~could~~ approach the high~~est~~ values such as 0.035, which is similar to that from cloud reflection. In this ~~difficult~~ case, PCC analysis is of great importance to keep cloud free snow scenes from the strict criteria of second step in particular in cases~~winter time~~ with higher solar zenith angle. ~~The~~ ASCIA in high PCC block covers~~accounts~~ a wide range of solar zenith angle (40-80 degree) and results in the reflectance of snow/ice being defined as between 0.02 and 0.04 at 3.7 µm channel. On the right panel of Fig. 8, a relatively ~~one can see the~~ large number of red scenes, which are falsely detected as cloud by~~in~~ the ISTO method, are observed.~~.~~

Figure 10 shows one example of a haze event over Svalbard on 3rd of May, 2006. Both the ~~of~~ ESA and ISTO cloud products showed~~had~~ good results~~performance~~ for this case with the exception of~~ the undetected~~ thin cloud scenes which are falsely labeled~~detected~~ as clear snow by~~ the~~ ISTO. ~~In fact, t~~The appropriate design and application of PCC analysis over 1.6 µm enables cloud to be discriminated from heavy aerosol load. However, aerosol load over cloud could not be separated from cloudy scenes.

The only season, in which all three approaches detected clouds with similar success, was summer in July as shown in Fig. 9. Although ASCIA shows an overall better performance in particular for thin clouds, the required computational time for cloud detection and surface classification is higher than two other methods.

~~In addition, w~~We also compared our results with those from the MODIS cloud identification algorithm used for masking cloudy~~ed~~ scenes. As an example, Fig. 11 shows the AATSR scene over Svalbard on 14th July 2008, where a large part of sea-ice is covered with thin clouds which have a small signature in visible channels. The middle panel shows the MODIS cloud mask for the same area. Although there is a small time difference of 15 minutes between MODIS and AATSR overpasses, we see that scenes identified as~~ with~~ cloudy by ASCIA correspond well with those of MODIS.

Figure 12 shows another example over northwest of Greenland on 18th ~~-~~May 2008. The thin and broken clouds are well detected over the snow cover by ASCIA, as well as the clouds over the southern part of the scene, which is covered with snow and ocean. As we can see from the comparison between ASCIA and MODIS cloud scene identification, cloudy scenes in the northern part of scene are not captured by MODIS product but the presence of clouds is seen in the RGB image in left panel. We observed other cases with similar differences especially for the case of thin and broken clouds. There are two potential sources of these differences, 1) time differences, which are 10 minutes in this case, or 2) an inadequate~~ proper~~ performance of the MODIS cloud mask over bright surfaces covered by snow and ice.

Due to the loss of Envisat and thus AATSR data in 2012, and the need for long time series of consistent data, we tested ASCIA on the AATSR successor SLSTR as well. Figure 13 shows some results over Svalbard on the 18th of April 2017. Due to the smaller swath width of AATSR compared to SLSTR, ~~the~~ ASCIA is not applied to the full coverage of SLSTR and the selected scene is cropped to have a~~the~~ similar coverage of 500×500 km^2. In spite of some unresolved calibration issues in this sensor, the higher spatial resolution in SLSTR clearly helps to improve cloud identification in first step, because the PCC analysis is more sensitive to smaller changes in 0.5×0.5 km^2 scenes, as compared to 1×1 km^2. Moreover, the shorter revisit time of the Sentinel-3 satellite provides more acquired images over the same scene. This results in a larger number of reference images, compared to those from Envisat. Overall these effects result in ~~the ASCIA~~ an expected improved performance of ASCIA when applied to~~application on~~ SLSTR data ~~being improved to the performance with~~, as compared to when applied to AATSR. ~~However, t~~ The comparison of MODIS and ASCIA results indicates that ASCIA detected more cloudy scenes than the MODIS algorithm in agreement with the above.

5.2 The comparison to ground-based measurements: SYNOP and AERONET

~~**6 Validation**~~

In this section, we present a quantitative validation of our ASCIA results by making comparisons with simultaneous ground-based SYNOP and AERONET measurements. The ESA standard cloud product is also compared with these validation data sets. The difference in spatial and temporal resolution of the new cloud identification datasets and the data sets used to validate this dataset has to be taken into account. To define the optimal maximum temporal difference between SYNOP and satellite data, other comparable validation activities used different temporal intervals like 10 min (Werkmeister et al., 2015), 15 min (Musial et al., 2014), 1 h (Dybbroe et al., 2005) and 4 h (Meerkötter et al., 2004). The investigation and results in the previous publications indicate that temporal difference in validation of satellite retrievals against SYNOP depend on meteorological conditions. Allowing only a small temporal difference between measurement datasets (here: SYNOP and ASCIA) ensures an optimal temporal overall but can introduce a significant sampling error due to the small number of scenes for validation (Bojanowski et al., 2014). According to Bojanowski et al. (2014) a temporal difference of 90 min to compare with SYNOP measurements at temporal resolution of 3 h minimizes the sampling error (Bojanowski et al., 2014). However, potential longer temporal difference will introduce an error which should be considered along other sources of uncertainty (different viewing perspective, different spatial footprint and etc.). In this study, the maximum allowed temporal difference between the ASCIA retrievals and SYNOP measurements is less than ±20 minutes in most cases and generally does not exceed ± 45 minutes. ~~The difference in the time of satellite and SYNOP measurements is small being below ±20 minutes in most cases and generally does not exceed ± 45 minutes.~~ To compare surface measurement from SYNOP hemispheric view with the cloud identification at a spatial resolution 1×1 km^2 resolution satellite measurement, we calculated cloudiness as the percentage of cloudy scenes within a window of 20×20 km^2 around each SYNOP station. This is a similar distance to that used in previous studies to validate satellite based cloud identification SYNOP or similar surface measurements (Kotarba, 2017; Werkmeister et al., 2015; Minnis et

al., 2003). The cloud detection data product was then compared to the three selected months (March, May and July) of SYNOP observations. These result in~~comprise~~ 100 measurements over Svalbard and Greenland.

In Fig. 14 we present the relation between the calculated Cloud Fractional Cover (CFC) from ASCIA and SYNOP measurements and density plot of occurrences of the CFC by ASCIA as a function of SYNOP following the idea of Werkmeister et al. (2015). The~~We find that these~~ two data sets have a correlation coefficient of R=0.92. In 31 % of scenarios, ASCIA estimates 1 okta more than SYNOP while in 14 % of match-ups SYNOP shows higher CFC of 1 okta. Figure 14 also reveals that most of ±1 okta differences occur when either SYNOP or ASCIA estimates 7 or 8 oktas which could be due to definition of 8 oktas (100 % CFC) and conversion of continuous percentage to okta (Werkmeister et al., 2015). For instance, CFC of 99.9 % is considered as 7 oktas by using Table 1, but~~while~~ the CFC difference is only 0.1 % with 8 oktas. The underestimation of CFC by SYNOP is also indicated~~confirmed~~ in the histogram of difference between ASCIA-SYNOP in Fig. 15. This underestimation~~ which~~ was confirmed~~indicated~~ by~~in~~ previous studies as well (Kotarba, 2009; Werkmeister et al., 2015). We also indicate the higher accuracy of ASCIA for cloud detection compared to the ESA cloud product. The results of the validation are summarized in Table 4. The cloud cover reported from SYNOP agrees in~~has an overall agreement of~~ 96 % (within ±2 okta) and 83 % (within ±1 okta) of the observations with the cloud identification data from ASCIA. As~~ we~~ discussed earlier an error of ±1 to ±2 okta would be expected as the accepted accuracy range from~~form~~ SYNOP cloud cover values due to man-made nature of its observation and viewing conditions (Boers et al., 2010; Werkmeister et al., 2015). In comparison, ~~T~~the ESA cloud product agrees 68 % (within ±2 okta) and 50 % (within ±1 okta) with SYNOP CFCs. The larger differences of SYNOP and ESA cloud product are also indicated in Fig. 16 where the CFC values in percentage are shown for ASCIA, ESA and SYNOP for validation scenarios. The blue error bars, indicate the range of okta values for each SYNOP as explained in~~according to~~ Table 1.

We also validated ASCIA cloud identification results with AERONET level 1.5 measurements, which~~.~~ are cloud screened. The procedure for this validation takes place in 2 steps: (1) covering AERONET observed AOT to a cloud flag (AOT is provided in AERONET only in cloud-free conditions); (2) Validation of ASCIA with AERONET cloud flag. ~~-~~In 86.1 % of 36 studied~~s~~ scenes over Svalbard, both ASCIA and AERONET confirm the presence of clouds.

6~~7~~ Conclusions

A new cloud detection algorithm, called ASCIA, for use at high altitudes above bright surfaces has been developed to generate stand-alone products and for subsequent use in the retrieval of AOT over the Arctic. ASCIA has been developed for use with the~~ uses~~ data from the European instruments ~~using~~ AATSR on the ESA Envisat (2002 to 2012) and SLSTR on the ESA Sentinel 3A or 3B. ~~The~~ ASCIA employs initially a time series analysis of PCC to identify cloud presence, the stability and cloud-free conditions at the block scale of scenes (25×25 km^2). It then uses the 3.7 µm solar reflectance to discriminate cloud presence at the spatial resolution of the scene~~ level~~, which is 1x1 km^2 or 0.5×0.5 km^2 for AATSR and SLSTR measurements respectively. The PCC parameter analysis of a block of data is independent to a first approximation of threshold settings, which often lead to misclassification of cloud and snow due to the similarity of their spectral characteristics and thus the thresholds. The brightness temperature

measurements from 3.7 µm channel provide the information to convert a block level resolution (25×25 km^2) to a scene level resolution (1×1 km^2 or 0.5×0.5 km^2) cloud identification. ASCIA thereby exploits the contrast in reflectance between snow/ice and cloud at 3.7 µm wavelength.

The results of applying the new developed ASCIA are compared and validated against 5 existing products and methods over the Arctic: 1) SYNOP measurements, 2) AERONET measurements, 3) one of existing methods based on spectral shape of clear snow 4) AATSR L2 nadir cloud flag, 5) MODIS cloud product. The validation ~~is~~ resulted in overall agreement of 96 % (within ±2 oktas) and 83 % (within ±1 okta) between SYNOP and ASCIA. The comparison of the ASCIA and ISTO methods shows the improved~~a better~~ performance of ASCIA in extreme situations, such as high solar zenith angle conditions.

The validation results indicate that the current ESA AATSR L2 nadir cloud flag often falsely identify clouds over snow/ice, ~~with the~~ except ~~ion of~~ during summer. The comparison between the ESA AATSR L2 cloud product and SYNOP measurements resulted in agreement of 68 % (within ±2 oktas) and 50 % (within ±1 okta). The overall better performance of ASCIA has also been demonstrated~~shown~~ when it is applied to ~~by~~ the SLSTR data. Nevertheless ~~However~~, ~~more investigation and~~an optimization is ~~are~~ needed for the detection of cloud over land (soil, vegetation etc.) for~~in~~ the PCC blocks having~~with~~ lower values. This is because ~~Since,~~ the strict performance of ~~the~~ ASCIA in cloudy blocks results in scenes of clear land (without snow cover) being identified as cloud due to high reflectance of land scenes at 3.7 µm channel. We also note that~~Additionally,~~ sub-scene cloud detection has not been studied with the current version of ASCIA. The use of reflectance in the 1.37 µm channel will be tested in the future to improve thin cirrus detection in ASCIA. The objective of this study was to assess and validate the current version of ASCIA for daytime observations. For night time application, an adaption of ASCIA is planned to identify clouds during night.

Acknowledgements: We gratefully acknowledge the support by the Collaborative Research Centres, CRC/Transregio 172 "ArctiC Amplification: Climate Relevant Atmospheric and SurfaCe Processes, and Feedback Mechanisms (AC)³. This work has been funded in part by the DFG CRC 172 and the State and University of Bremen.

References

Allen, R. C., Durkee, P. A., and Wash, C. H.: Snow/cloud discrimination with multispectral satellite measurements, J. Appl. Meteor., 29, 994–1004, 1990.

Arking, A. and Childs, J. D.: Retrieval of cloud cover parameters from multispectral satellites images, J. Climate Appl. Meteor., 24, 322-333, 1985.

Arola, A., Eck, T. F., Kokkola, H., Pitkanen, M. R. A., Romakkaniemi, S.: Assessment of cloud-related fine-mode AOD enhancement based on AERONET SDA product, Atmos. Chem. Phys., 17, 5991–6001, 2017.

Benesty, J., Chen, J., Huang, Y., Cohen, I.: Noise Reduction in Speech Processing, Springer Topics in Signal Processing 2, Springer-Verlag Berlin Heidelberg, doi: 10.1007/978-3-642-00296-0_5, 2009.

Birks, A. R.: Improvements to the AATSR IPF relating to Land Surface Temperature Retrieval and Cloud Clearing over Land, AATSR Technical Note, Rutherford Appleton Laboratory, Chilton, UK, 2007.

Boers R., de Haij, M.J., Wauben, W. M. F., Baltink, H. K., van Ulft, L. H., Savenije, M., Long, C. N.: Optimized fractional cloudiness determination from five ground-based remote sensing techniques, J. Geophys. Res., 115, D24116, 2010.

Bojanowski, J., Stöckli, R., Tetzlaff, A., and Kunz, H.: The Impact of Time Difference between Satellite Overpass and Ground Observation on Cloud Cover Performance Statistics, Remote Sensing, 6, 12 866–12 884, https://doi.org/10.3390/rs61212866, 2014.

Bulgin, C. E., Eastwood, S., Embury, O., Merchant, C. J. and Donlon, C., Sea surface temperature climate change initiative: alternative image classification algorithms for sea-ice affected oceans. Remote Sens. Environ, 162, pp. 396-407, 2015.

Christensen, M. W., Neubauer, D., Poulsen, C. A., Thomas, G. E., McGarragh, G. R., Povey, A. C., Proud, S. R., Grainger, R. G.: Unveiling aerosol–cloud interactions – Part 1: Cloud contamination in satellite products enhances the aerosol indirect forcing estimate, Atmos. Chem. Phys., 17, 13151–13164, 2017.

Cohen, J., Screen, J. A., Furtado, J., C., et al.: Recent Arctic amplification and extreme mid-latitude weather, Nat. Geosci., 7(9), 627–637, 2014.

Coppo, P., Ricciarelli, B., Brandani, F., Delderfield, J., Ferlet, M., Mutlow, C., Munro, G., Nightingale, T., Smith, D., Bianchi, S., Nicol, P., Kirschstein, S., Hennig, T., Engel, W., Frerick, J., Nieke, J., SLSTR: a high accuracy dual scan temperature radiometer for sea and land surface monitoring from space, J. Mod. Optic., 57(18), 1815-1830, 2010.

Curry, J. A., Rossow, W. B., Randall, D., Schramm, J. L.: Overview of Arctic cloud and radiation characteristics. J. Climate, 9, 1731–1764, 1996.

Dybbroe, A.; Karlsson, K.-G.; Thoss, A. NWCSAF AVHRR cloud detection and analysis using dynamic thresholds and radiative transfer modeling. Part II: Tuning and validation. J. Appl. Meteorol., 44, 55–71, 2005.

Derrien, M., Farki, B., Harang, L., Pochic, D., Sairouni, A., LeGldau, H., Noyalet, A.: Automatic Cloud Detection Applied to NOAA-11/AVHRR Imagery, Remote Sens. Environ., 46, 246-267, 1993.

Eck, T. F., B. N. Holben, J. S. Reid, O. Dubovik, A. Smirnov, N. T. O Neill, I. Slutsker, and S. Kinne, Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols, J. Geophys. Res., 104(1), 333-349, 1999.

Ghent, D. J., Corlett, G. K., Göttsche,F.-M., & Remedios, J. J.: Globalland surface temperature from the Along-Track Scanning Radiometers, J. Geophys. Res., 122, 12167–12193, 2017.

Gómez-Chova, L., Amorós-López, J., Mateo-García, G., Muñoz-Marí, J., Camps-Valls, G.: Cloud masking and removal in remote sensing image time series, J. Appl. Remote Sens., 11(1), 015005, 2017.

Hagolle, O., Sylvander, S., Huc, M., Claverie, M., Clesse, D., Dechoz, C., Lonjou, V., Poulain, V.: SPOT4 (Take5): A simulation of Sentinel-2 time series on 45 large sites. Remote Sens., 7, 12242–12264, 2015.

Hall, D. K., Riggs, G., Salomonson, V. V.: Algorithm Theoretical Basis Document (ATBD) for the MODIS snow and sea-ice mapping algorithms, September 2001.

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., Smirnov, A.: AERONET - a federated instrument network and data archive for aerosol characterization Remote Sens. Environ., 66, 1-16, 1998.

Hollmann, R.; Merchant, C.J.; Saunders, R.; Downy, C.; Buchwitz, M.; Cazenave, A.; Chuvieco, E.; Defourny, P.; de Leeuw, G.; Forsberg, R.; et al. The ESA climate change initiative Satellite Data Records for Essential Climate Variables. Bull. Am. Meteorol. Soc., 94, 1541–1552, 2013.

Istomina, L., Hoyningen-Huene, W., Kokhanovsky, A.A., Burrows, J.P.: The detection of cloud-free snow-covered areas using AATSR measurements, Atmos. Chem. Phys., 9, 1279-1288, 2010.

Istomina, L.: Retrieval of aerosol optical thickness over snow and ice surfaces in the Arctic using Advanced Along Track Scanning Radiometer, Ph.D. thesis, University of Bremen, Germany, 2012.

Kellogg, W. W.: Climatic feedback mechanisms involving the polar regions, Climate of the Arctic, Weller, G., and Bowling, S. A., (Eds.), Geophysical Institute, University of Alaska, Fairbanks, AK, 111–116, 1975.

Key, J. and Barry, R. G.: Cloud cover analysis with Arctic AVHRR data. 1. Cloud detection, J. Geophys. Res., 94, D15, 18521– 18535, 1989.

Kim B-M, Hong J-Y, Jun S-Y, Zhang X, Kwon H, Kim S-J, Kim J-H, Kim S-W and Kim H-K: Major cause of unprecedented Arctic warming in January 2016: critical role of an Atlantic windstorm Sci. Rep., 7, 40051, doi:10.1038/srep40051, 2017.

Khlopenkov, K., Trishchenko, A. SPARC: New cloud, snow, and cloud shadow detection scheme for historical 1-km AVHHR data over Canada, J. Atmos. Ocean. Technol., 24, 322–343, 2007.

Klüser, L., Killius, N., Gesell, G.: APOLLO_NG – a probabilistic interpretation of the APOLLO legacy for AVHRR heritage channels, Atmos. Meas. Tech., 8, 4155–4170, 2015.

Kokhanovsky, A. A.: Cloud Optics, Eds.: Mysak, L. A., Hamilton, K., Publ. Springer, 2006.

Kolmonen, P., Sundström, A. -M., Sogacheva, L., Rodriguez, E., Virtanen, T. H., & de Leeuw, G.: Uncertainty characterization of AOD for the AATSR dual and single view retrieval algorithms, Atmos. Meas. Tech. Discuss., 6, 4039–4075, 2013.

Kolmonen, P., Sogacheva, L., Virtanen, T. H., de Leeuw, G., Kulmala, M.: The ADV/ASV AATSR aerosol retrieval algorithm: current status and presentation of a full-mission AOD dataset, Int. J. Digit. Earth, 9, 545-561, 2016.

Kotarba, A. Z.: A comparison of MODIS-derived cloud amount with visual surface observations, J. Atmos. Res., 92, 522–530, doi:10.1016/j.atmosres.2009.02.001, 2009.

Kotarba A. Z.: Inconsistency of surface-based (SYNOP) and satellite-based (MODIS) cloud amount estimations due to the interpretation of cloud detection results. Int. J. Climatol. 37(11):4092–4104, 2017.

Law, K. S., Stohl, A.: Arctic Air pollution: origins and impacts, J. Science, 315 (5818), 1537-1540, doi: 10.1126/science.1137695, 2007.

Leese, J. A., Novak, C. S., Taylor, V. R.: The determination of cloud pattern motions from Geosynchronous Satellite Image Data, J. Pattern Recognition, 2, 279-292, 1970.

Lyapustin, A., Wang, Y., and Frey, R.: An automated cloud mask algorithm based on time series of MODIS measurements, J. Geophys. Res., 113, D16207, doi:10.1029/2007JD009641, 2008.

Lyapustin, A., Tedesco, M., Wang, Y., Aoki, T., Hori, M., Kokhanovsky, A.: Retrieval of snow grain size over Greenland from MODIS Remote Sens. Environ., 113 (9), 1976-198, 2009.

Martins, J.V., Tanre, D., Remer, L., Kaufman, Y.J., Mattoo, S., Levy, R.: MODIS Cloud screening for remote sensing of aerosols over oceans using spatial variability. Geophys. Res. Lett. 29 (12), 1619, 2002.

Mei, L., Xue, Y., Kokhanovsky, A. A., von Hoyningen-Huene, W., Istomina, L., de Leeuw, G., Burrows, J. P., Guang, J., Jing, Y., Aerosol optical depth retrieval over snow using AATSR data, International Journal of Remote Sensing, 34, 5030-5041, 2013.

Mei, L., Rozanov, V. V., Vountas, M., Burrows, J. P., Levy, R. C., and Lotz, W. A.: Retrieval of aerosol optical properties using MERIS observations: algorithm and some first results, Remote Sens. Environ., 197, 125–141, 2017a.

Mei, L. L., Vountas, M., Gómez-Chova, L., Rozanov, V., Jäger, M., Lotz, W., Burrows, J. P., and Hollmann, R.: A Cloud masking algorithm for the XBAER aerosol retrieval using MERIS data, Remote Sens. Environ., 197, 141–160, 2017b.

Meerkötter, R., König, C., Bissolli, P., Gesell, G., Mannstein, H., A 14-year European cloud climatology from NOAA/AVHRR data in comparison to surface observations. Geophys. Res. Lett., 31, doi:10.1029/2004GL020098, 2004.

~~Meyer, K., Platnick, S.: Utilizing the MODIS 1.38 μm channel for cirrus cloud optical thickness retrievals: Algorithm and retrieval uncertainties, J. Geophys. Res., 115, D24209, 2010.~~

Minnis, P., Spangenberg, D.A., Chakrapani, V.: Distribution and validation of cloud cover derived from AVHRR data over the Arctic Ocean during the SHEBA year. Proc. 13th ARM Science Team Meeting, Broomfield, Colorado, March 31–April 4, 2003.

Musial, J. P., Husler, F., Sutterlin, M., Neuhaus, C., Wunderle, S.: Dyatime low Stratiform cloud detection on AVHRR Imagery, Remote Sens., 6, 5124-5150, doi: 10.3390/rs6065124, 2014.

North, P. R. J., Briggs, S. A., Plummer, S. E., & Settle, J. J.: Retrieval of land surface bidirectional reflectance and aerosol opacity from ATSR-2 multi-angle imagery, IEEE Trans. Geosci. Remote Sens. 37, 526–537, 1999.

Pavolonis, M. J., Heidinger, A. K., and Uttal, T.: Daytime global cloud typing from AVHRR and VIIRS: Algorithm description, validation, and comparison, J. Appl. Meteorol., 44, 804-826, 2005.

Pearson, K., VII. Mathematical contributions to the theory of evolution.—III. Regression, heredity, and panmixia, Philos. T. Roy. Soc. A., 187, 253-318, 1896.

Pithan, F., Mauritsen, T.: Arctic amplification dominated by temperature feedbacks in contemporary climate models, Nat. Geosci., 7, 181–184, 2014.

Platnick, S., Fontenla, J. M.: Model calculations of solar spectral irradiance in the 3.7 band for Earth remote sensing application, Am. Meteorol. Soc., 47, 124-134, 2008.

Remer, L. A., Mattoo, S., Levy, R. C., Heidinger, A., Pierce, R. B., Chin, M.: Retrieving aerosol in a cloudy environment: aerosol product availability as a function of spatial resolution, Atmos. Meas. Tech., 5, 1823–1840, 2012.

Rodgers, J. L., Nicewander, W. A.: Thirteen Ways to Look at the Correlation Coefficient, The American Statistician, 42 (1), 59-66, 1988.

Rosenfeld, D., Cattani, E., Melani, S., Levizzani, V.: Considerations on daylight operation of 1.6 versus 3.7 mm channel on NOAA and METOP satellites, American Meteorologica Society, 85, 873-881, 2004.

Rossow, W.B., and Garder, L. C.: Validation of ISCCP cloud detections. J. Climate, 6, 2370-2393, 1993.

Serreze M. C., Barry R. G.: Processes and impacts of Arctic amplification: a research synthesis, Global and Planetary Change, 77, 85–96, 2011.

Meirink, J., F., van Zadelhoff, G. J., Algorithm Theoretical Basis Document SEVIRI Cloud Physical Products, CLAAS Edition 2, Date: 10.06.2016, Issue: 2.2. 2016.

Shi, Y., Zhang, J., Reid, J. S., Liu, B., and Hyer, E. J.: Critical evaluation of cloud contamination in the MISR aerosol products using MODIS cloud mask products, Atmos. Meas. Tech., 7, 1791– 1801, doi:10.5194/amt-7-1791-2014, 2014.

Sobrino, J. A., Jiménez-Muñoz, J. C., Sòria, G., Ruescas, A. B., Danne, O., Brockmann, C., Ghent, D., Remedios, J., North, P., Merchant, C., Berger, M., Mathieu, P. P., Göttsche, F. M.: Synergistic use of MERIS and AATSR as a proxy for estimating Land Surface Temperature from Sentinel-3 data, Remote Sens. Environ., 179, 149–161, 2016.

Sobrino, J. A., Jiménez-Muñoz, J. C., Barres, G. S., Julien, Y.: Synergistic Use of The Sentinel Missions For Estimating And Monitoring Land Surface Temperature (SEN4LST FINAL REPORT), ESA technical report, doi: 10.13140/RG.2.2.34693.35049, 2013.

Soliman, A., Duguay, C., Saunders, W., Hachem, S.: Pan-Arctic land surface temperature from MODIS and AATSR: Product development and intercomparison, Remote. Sens-Basel, 4, 3833-3856, 2012.

Spangenberg, D. A., Chakrapani, V., Doelling, D. R., Minnis, P., and Arduini, R. F.: Development of an automated Arctic cloud mask using clear-sky satellite observations taken over the SHEBA and ARM NSA sites, Proc. 6th Conf. on Polar Meteor. and Oceanography, San Diego, CA, 14–18 May 2001, 246–249, 2001.

Thomas, G. E., Carboni, E., Sayer, A. M., Poulsen, C. A., Siddans, R., and Grainger, R. G.: Oxford-RAL Aerosol and Cloud (ORAC): aerosol retrievals from satellite radiometers, in: Satellite Aerosol Remote Sensing Over Land, edited by: Kokhanovsky, A. A. and de Leeuw, G., Springer, Berlin, Germany, 193–225, 4042, 4051, 2009.

Várnai, T., Marshak, A.: Effect of Cloud Fraction on Near-Cloud Aerosol Behavior in the MODIS Atmospheric Correction Ocean Color Product, Remote Sens., 7, 5283–5299, 2015.

~~Veefkind, J. P.,&de Leeuw, G.: A new algorithm to determine the spectral aerosol optical depth from satellite radiometer measurements., J. Aerosol Sci. 29, 1237–1248, 1998.~~\

~~Veefkind, J. P., de Leeuw, G. D., and Durkee, P. A.: Retrieval of aerosol optical depth over land using two-angle view satellite radiometry during TARFOX, Geophys. Res. Lett., 25, 3135–3138, 1999.~~\

Werkmeister, A., Lockhoff, M., Schrempf, M., Tohsing, K., Liley, B., and Seckmeyer, G.: Comparing satellite- to ground-based automated and manual cloud coverage observations – a case study, Atmos. Meas. Tech., 8, 2001-2015, 2015.

Wendisch, M., M. Brückner, J. P. Burrows, S. Crewell, K. Dethloff, K. Ebell, Ch. Lüpkes, A. Macke, J. Notholt, J. Quaas, A. Rinke, and I. Tegen,: Understanding causes and effects of rapid warming in the Arctic. Eos, 98, doi: 10.1029/2017EO064803, 2017.

Wind, G., da Silva, A. M., Norris, P. M., Platnick, S., Mattoo, S., Levy, R. C.: Multi-sensor cloud and aerosol retrieval simulator and remote sensing from model parameters – Part 2: Aerosols, Geosci. Model Dev., 9, 2377–2389, 2016.

WMO: Manual on Codes. Part A – Alphanumeric Codes. Secretariat of the World Meteorological Organization: Geneva, Switzerland, 1995.

Zavody, A. M., Mutlow, C. T., and Llewellyn-Jones, D. T.: Cloud Clearing over the Ocean in the Processing of Data from the Along-Track Scanning Radiometer (ATSR), Journal of Atmospheric and Oceanic Technology, 17, 595-615, 2000.

Table 1. Calculation of cloudiness in percentage for corresponding okta values

Percentage of Cloud	Okta
0	0
$0 < \% < 18.75$	1
$18.75 \leq \% < 31.25$	2
$31.25 \leq \% < 43.75$	3
$43.75 \leq \% < 56.25$	4
$56.25 \leq \% < 68.75$	5
$68.75 \leq \% < 81.25$	6
$81.25 \leq \% < 100$	7
100	8

711 **Table 2.** Simulated and observed reflectance values at 3.7 μm (Allen et al., 1990)

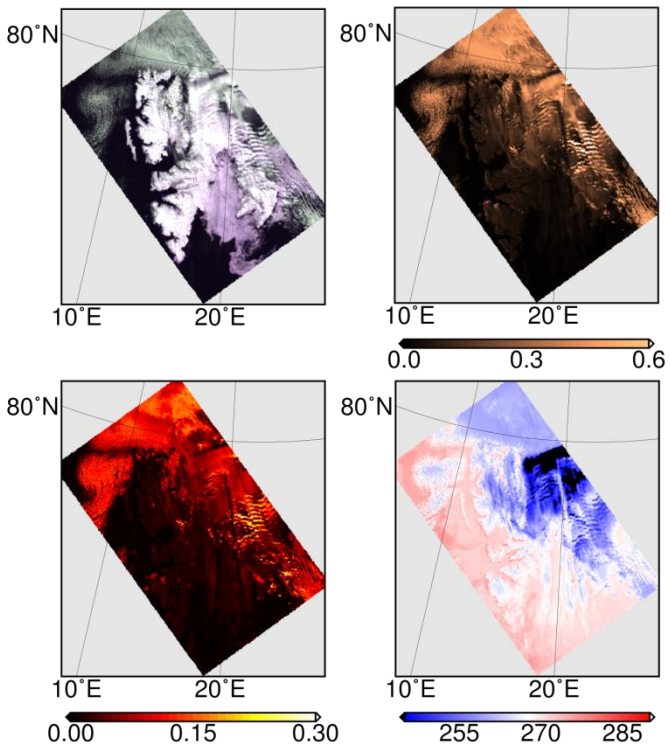
Surface/cloud Type	Simulation of 3.7 μm Reflectance	Observation of 3.7 μm Reflectance
Ice cloud	0.01-0.3	0.02-0.27
Liquid cloud	0.1-0.45	0.08-0.36
Clear land	~0.15	0.03-0.1
Snow cover	0.005-0.025	0.02- 0.04

712 **Table 3.** Land classification criteria in cloud-free scene.

Surface Type	Test~~Simulation of 3.7 μm Reflectance~~	Description
Water	$R_{0.87} < 11\%$ & NDSI≥0.4	MODIS snow and ice mapping ATBD (Hall et al., 2001)
Sea-ice	$R_{0.87} > 11\%$ & NDSI≥0.4	
Land	$R_{3.7}$>0.04 & $R_{0.66} < 0.2$ || NDSI<0.4	Allen et al.,1999
Snow	$R_{3.7}$≤0.04	Allen et al., 1999

713 **Table 4.** A summary of the comparison of ASCIA and ESA cloud product with SYNOP measurements used to validate these
714 products.

	Criteria	
Cloud data	within ±2 oktas	within ±1 okta
ASCIA vs. SYNOP	96 % ~~correct~~agreement 4 % ~~incorrect~~disagreement	83 % ~~correct~~agreement 17 % ~~incorrect~~disagreement
ESA vs. SYNOP	68 % ~~correct~~agreement 32 % ~~incorrect~~disagreement	50 % ~~correct~~agreement 50 % ~~incorrect~~disagreement


80˚N
10˚E
20˚E
80˚N
10˚E
20˚E
0.0
0.3
0.6
80˚N
10˚E
20˚E
0.00
0.15
0.30
80˚N
10˚E
20˚E
255
270
285

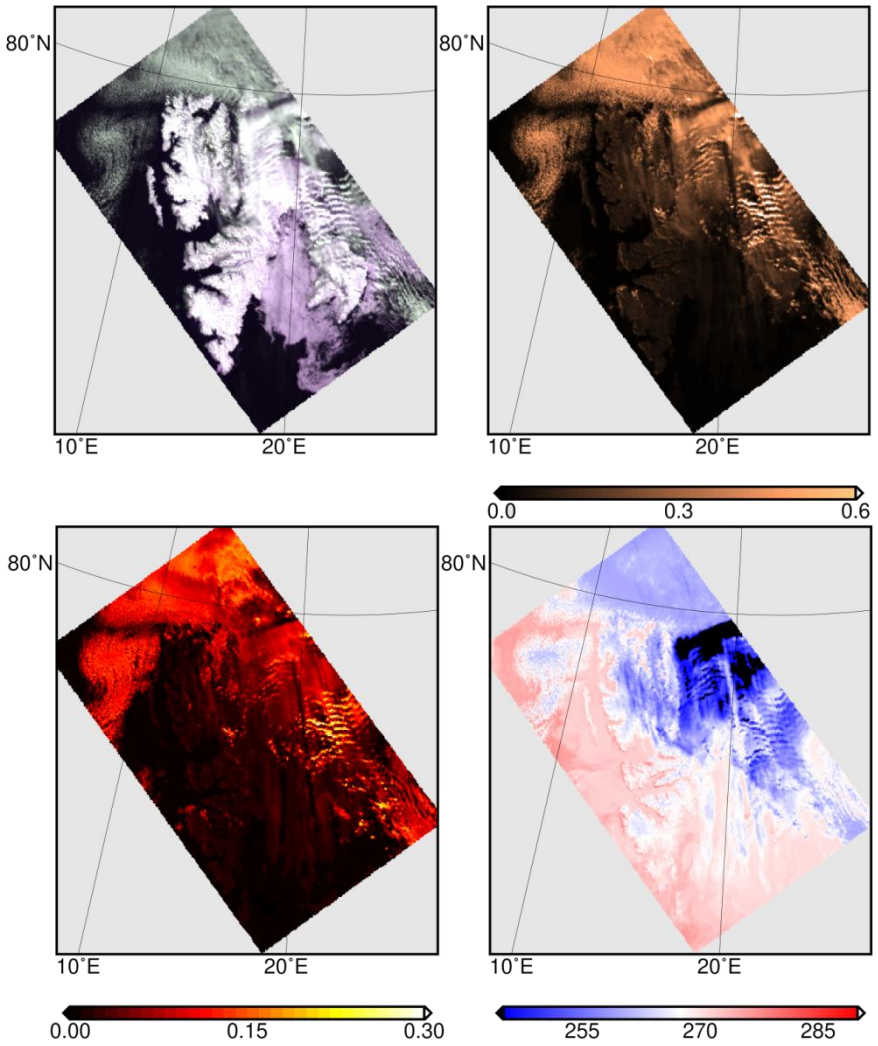
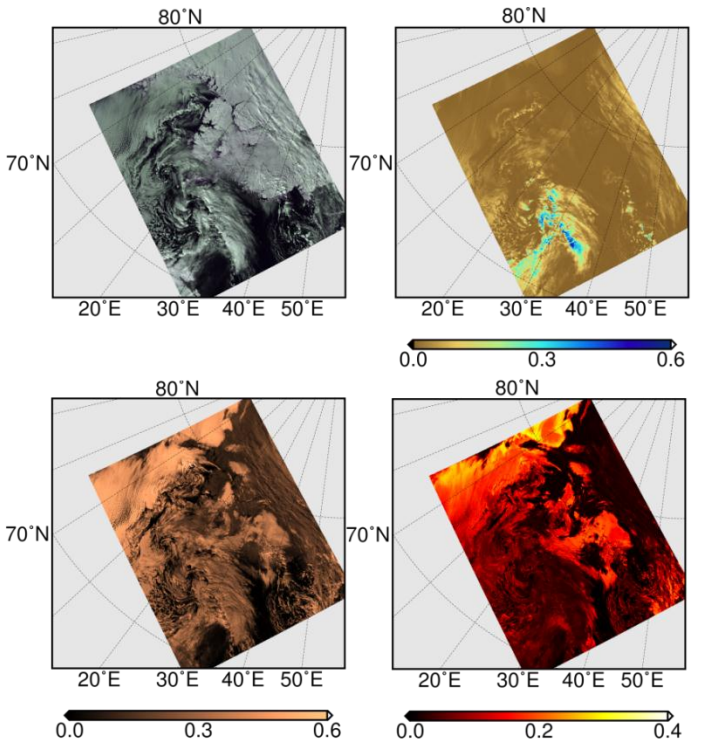


Figure 1. Upper left: the false colour RGB image of AATSR (using 0.67, 0.87 and 0.55 µm channels) over Svalbard, 10 May 2006, upper right: 1.6 µm reflectance, lower left: 3.7 µm reflectance, lower right: 11 µm brightness temperature.


80˚N
70˚N
20˚E
30˚E
40˚E
50˚E
0.0
0.3
0.6
0.2
0.4

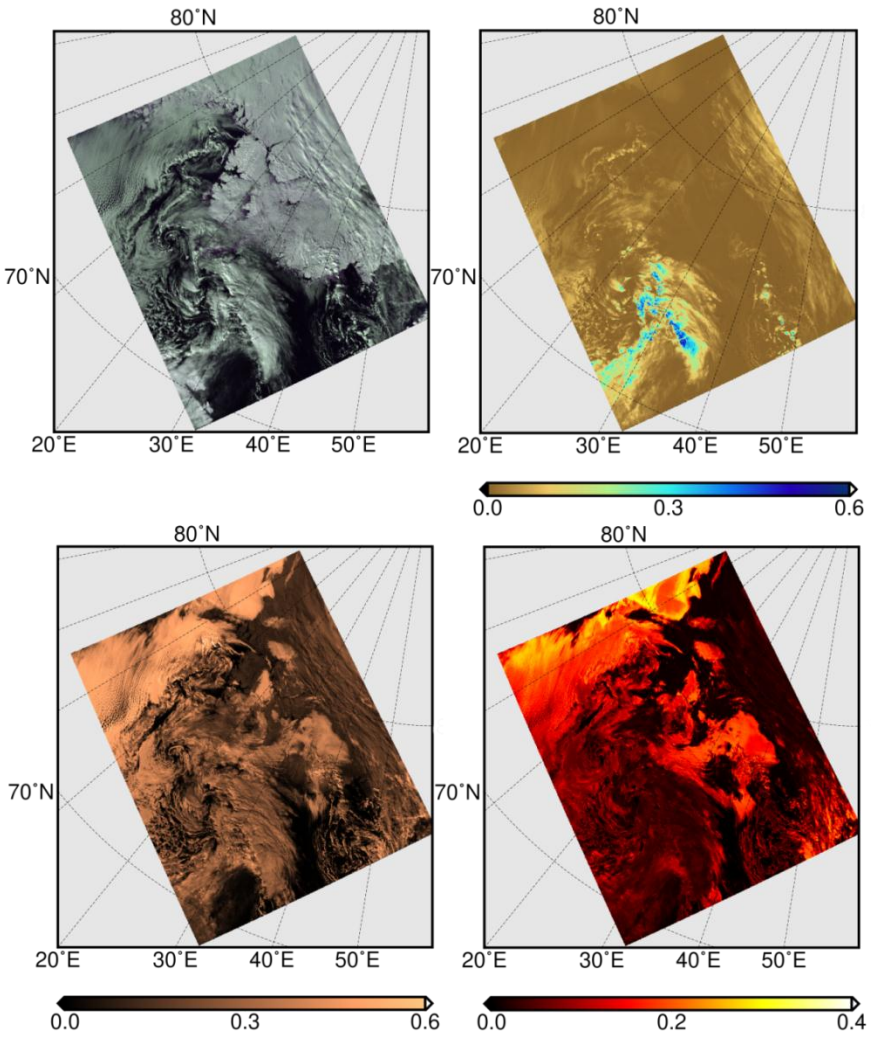
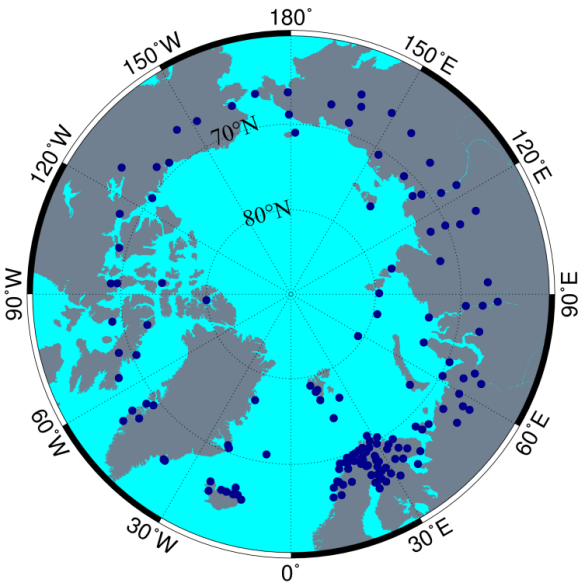
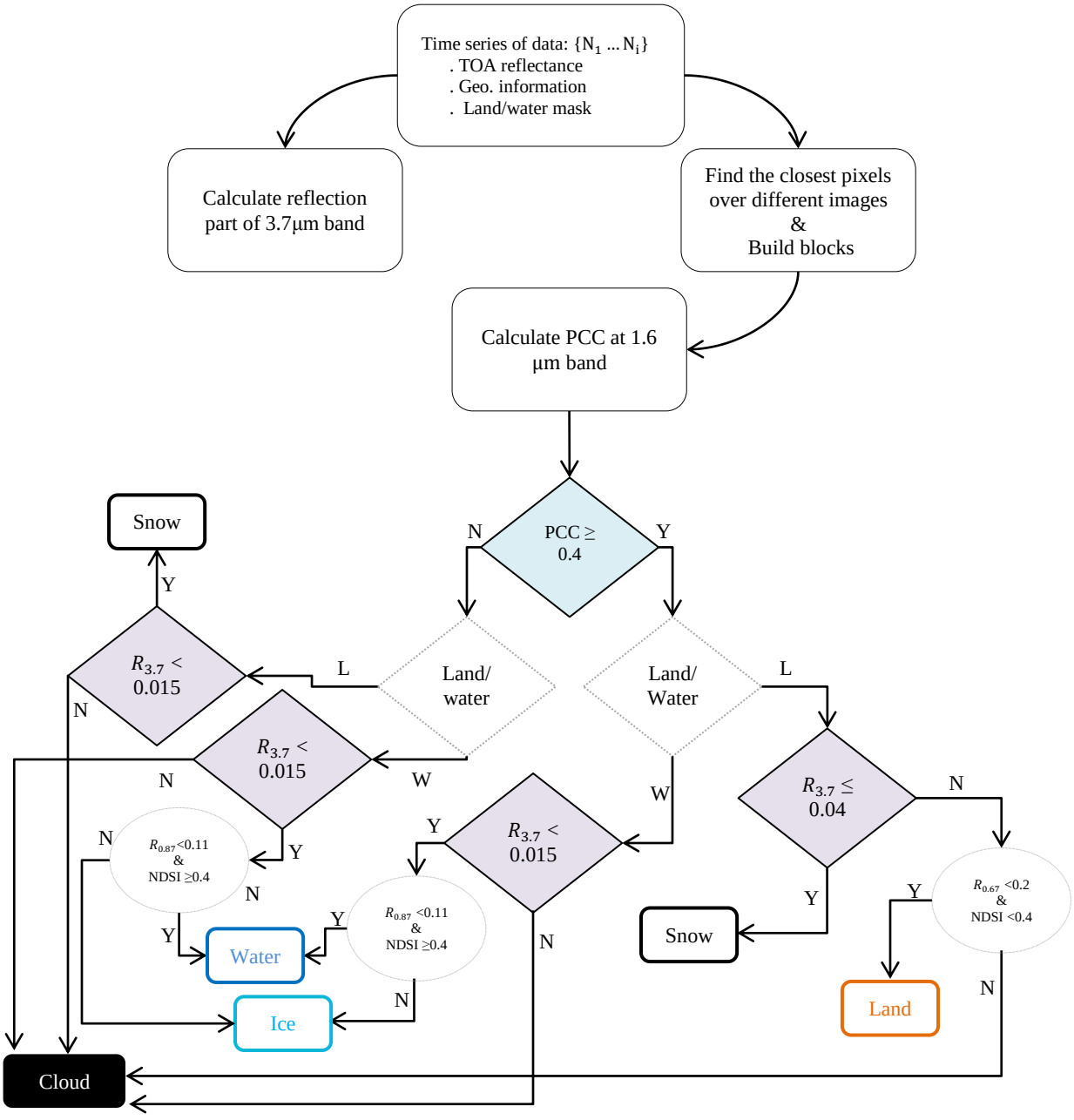


Figure 2. Upper left: the RGB false colour image (using 0.67, 0.87 and 0.55 µm channels) of SLSTR over Svalbard, 18 April 2017, upper right: 1.37 µm reflectance, lower left: 1.6 µm reflectance, lower right: 3.7 µm reflectance.



719 **Figure 3.** SYNOP network coverage over the Arctic, the dark blue points indicate the location of SYNOP stations.



720 **Figure 4.** The schematic flowchart of ASCIA.

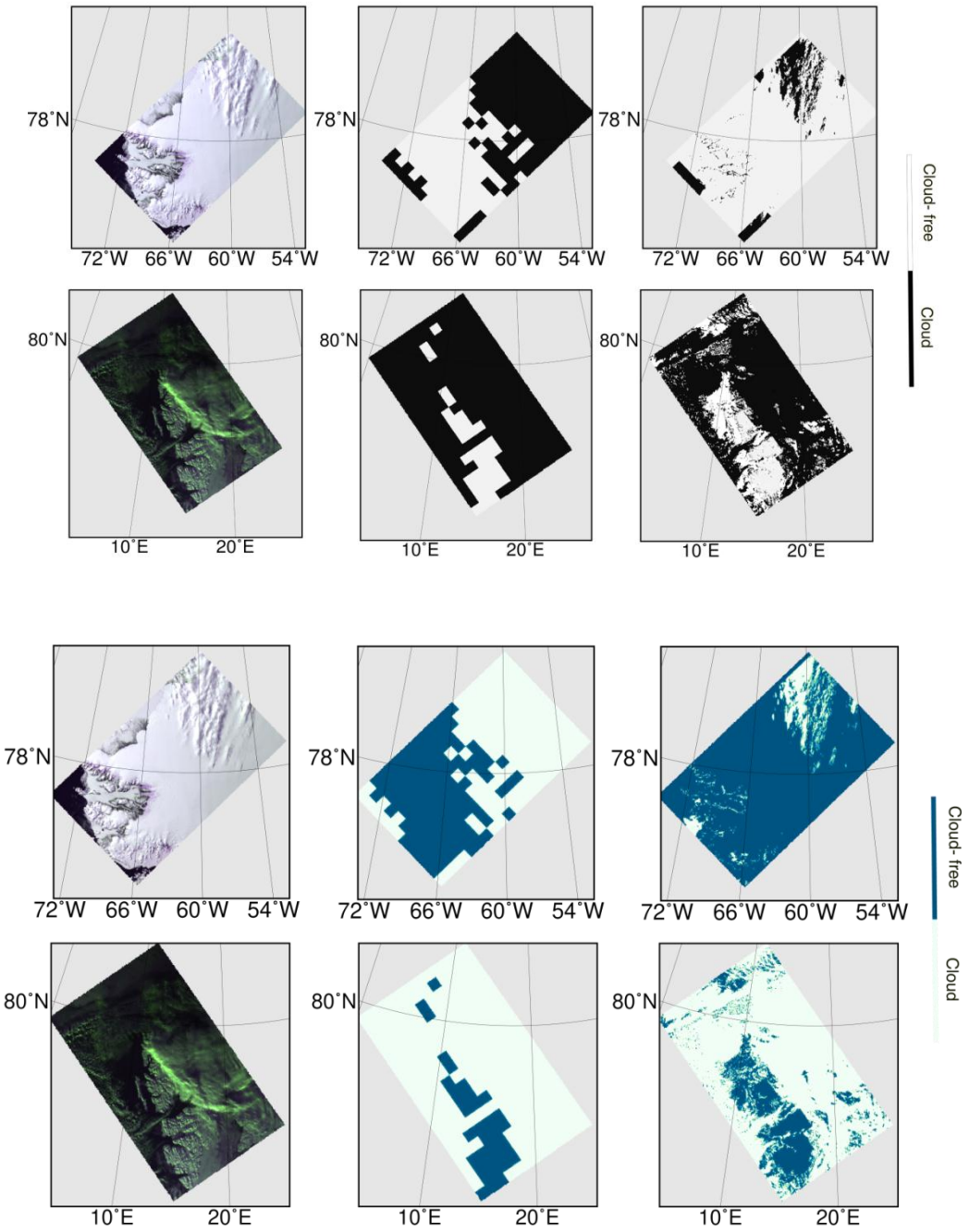
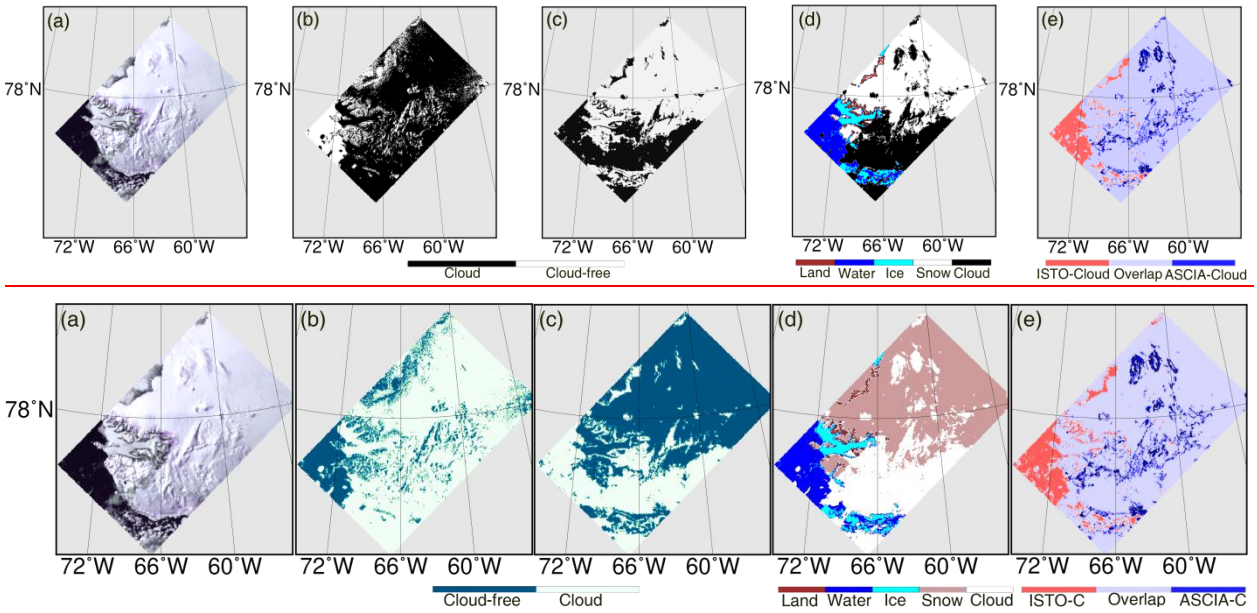
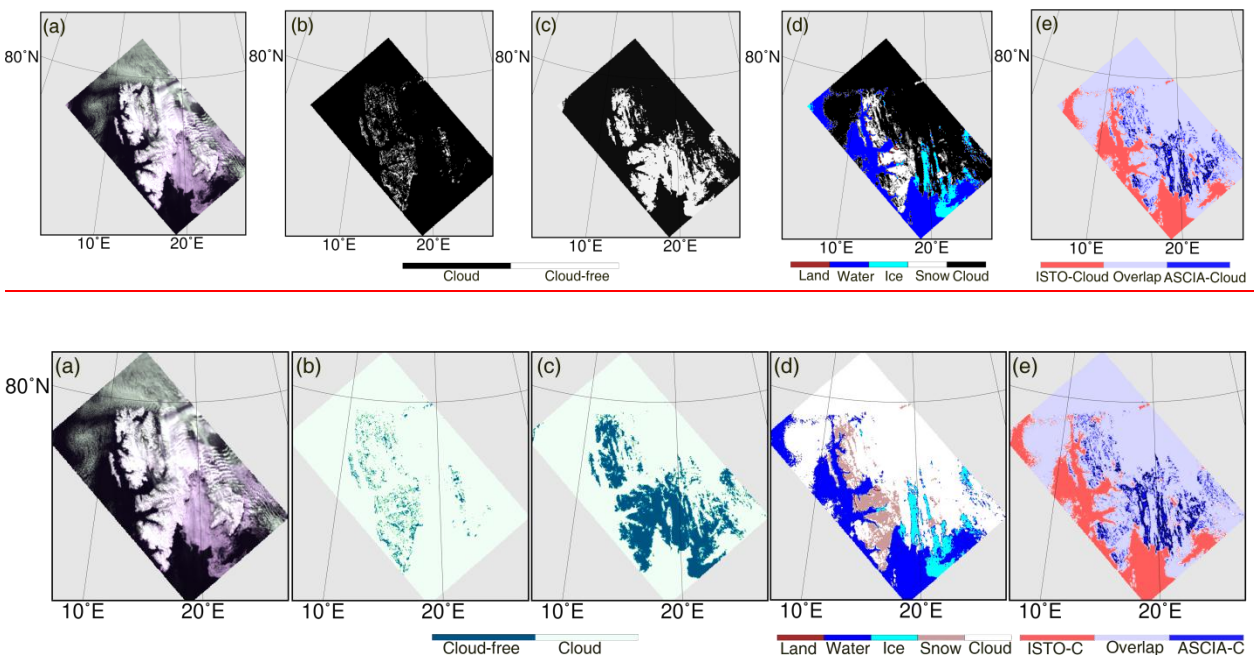


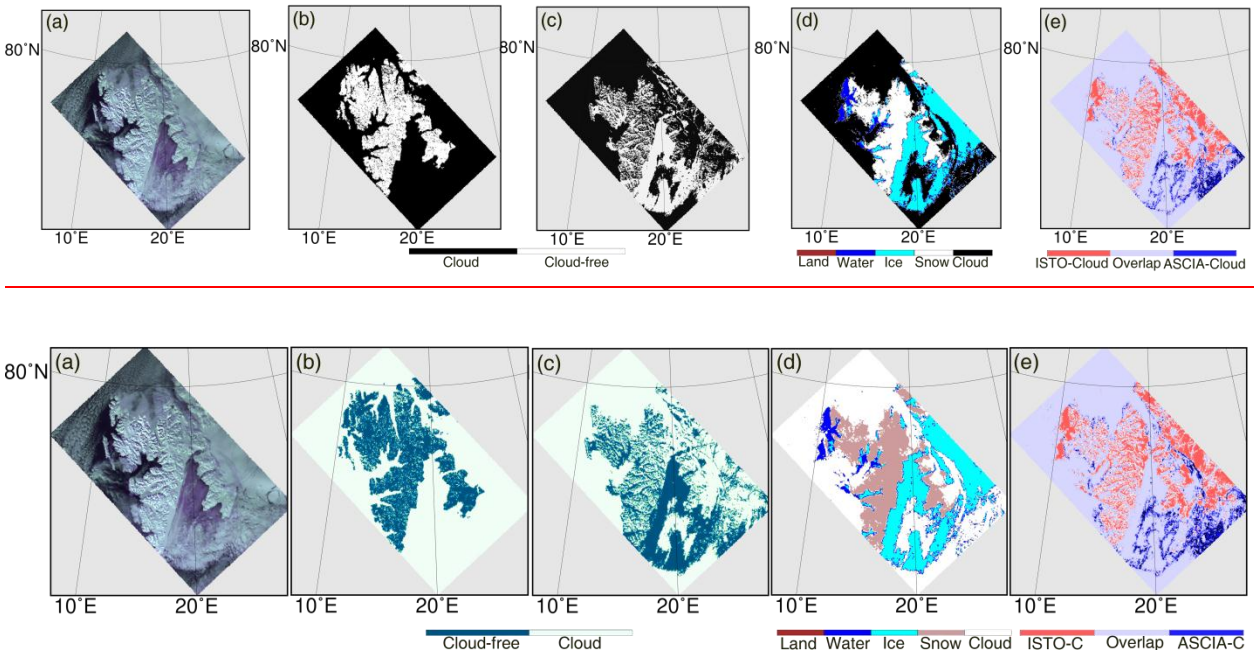
Figure 5. Examples of the results of ASCIA on AATSR observations on the scenes over Greenland (upper panels) between (75°N, 48°W), (75°N, 75°W), (81°N, 48°W), (81°N, 75°W), taken on the 18 May 2008 and Svalbard (lower panels), within (75°N, 4°E), (75°N, 32°E), (81°N, 4°E), (81°N, 32°E) ~~(Lowe panels), taken on the 18 May 2008 and~~on the 1 March 2008 ~~respectively~~, Left panels: RGB false colour images (using 0.67, 0.87 and 0.55 μm channels), middle panels: Cloud detection at block level (25×25 km^2), right panels: cloud detection at scene level.



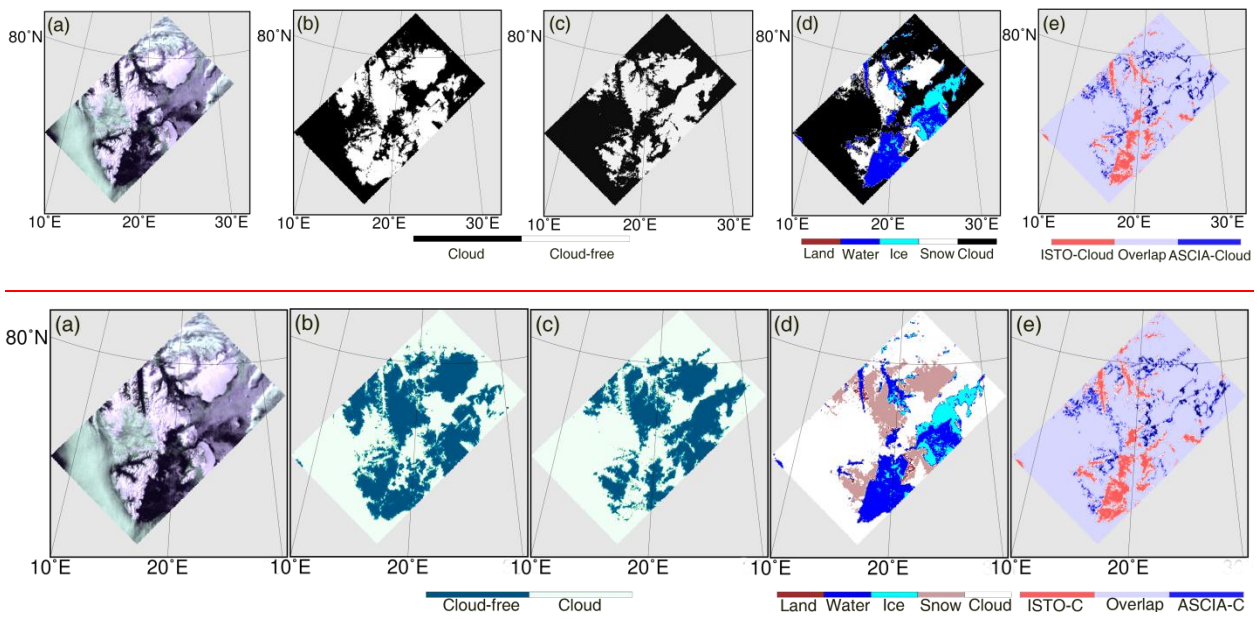
726 **Figure 6.** (a) The RGB false colour image (using 0.67, 0.87 and 0.55 µm channels) of AATSR over northern Greenland, 24 May
727 2008, (b) Nadir cloud flag from AATSR L2 product, (c) cloud detection based on spectral shape of clear snow, (d) cloud
728 detection of ASCIA, (e) difference between ISTO and ASCIA



729 **Figure 7.** (a) The RGB false colour image (using 0.67, 0.87 and 0.55 µm channels) of AATSR over Svalbard, 10 May 2006, (b)
730 Nadir cloud flag from AATSR L2 product, (c) cloud detection based on spectral shape of clear snow, (d) cloud detection of
731 ASCIA, (e) difference between ISTO and ASCIA.



732 **Figure 8.** (a) The RGB false colour image (using 0.67, 0.87 and 0.55 µm channels) of AATSR over Svalbard, 18 March 2008, (b) Nadir
733 cloud flag from AATSR L2 product, (c) cloud detection based on spectral shape of clear snow, (d) cloud detection of ASCIA, (e) difference
734 between ISTO and ASCIA.



735 **Figure 9.** (a) The RGB false colour image (using 0.67, 0.87 and 0.55 µm channels) of AATSR over Svalbard, 6 July 2008, (b) Nadir cloud
736 flag from AATSR L2 product, (c) cloud detection based on spectral shape of clear snow, (d) cloud detection of ASCIA, (e) difference between
737 ISTO and ASCIA.

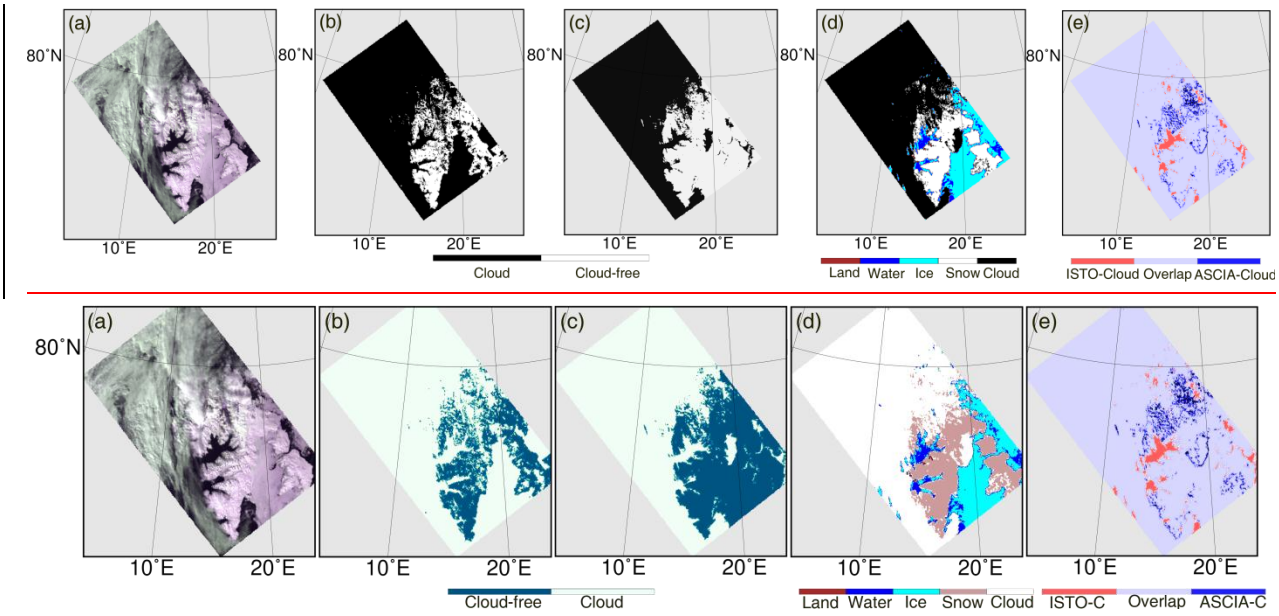


Figure 10. (a) The RGB false colour image (using 0.67, 0.87 and 0.55 µm channels) of AATSR over Svalbard, 3 May 2006, (b) Nadir cloud flag from AATSR L2 product, (c) cloud detection based on spectral shape of clear snow, (d) cloud detection of ASCIA, (e) difference between ISTO and ASCIA.

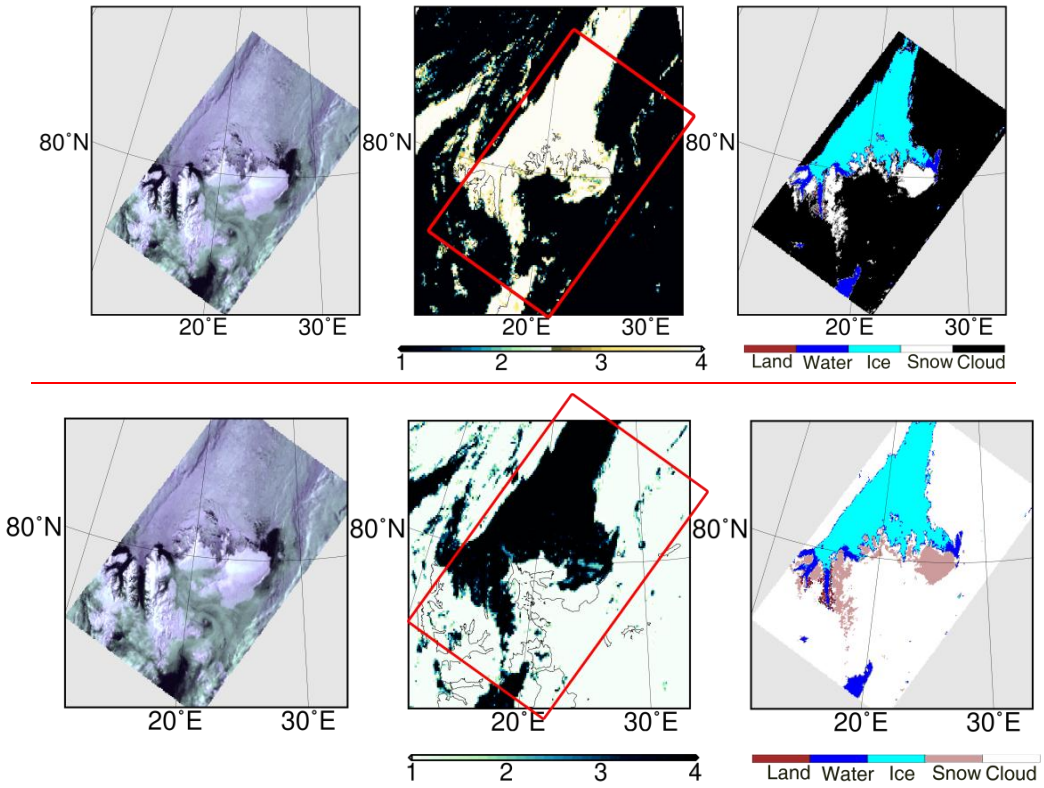


Figure 11. Left panel: RGB false colour image (using 0.67, 0.87 and 0.55 μm channels) of AATSR over Svalbard, 14 July 2008, 16h 40min 45s, middle panel MODIS cloud mask algorithm retrieved data: 1- cloudy, 2- probably cloudy, 3- probably clear, 4- clear, (red rectangle shows the coverage of AATSR) for 16h 25 min, right panel: the results for the cloud detection of ASCIA.

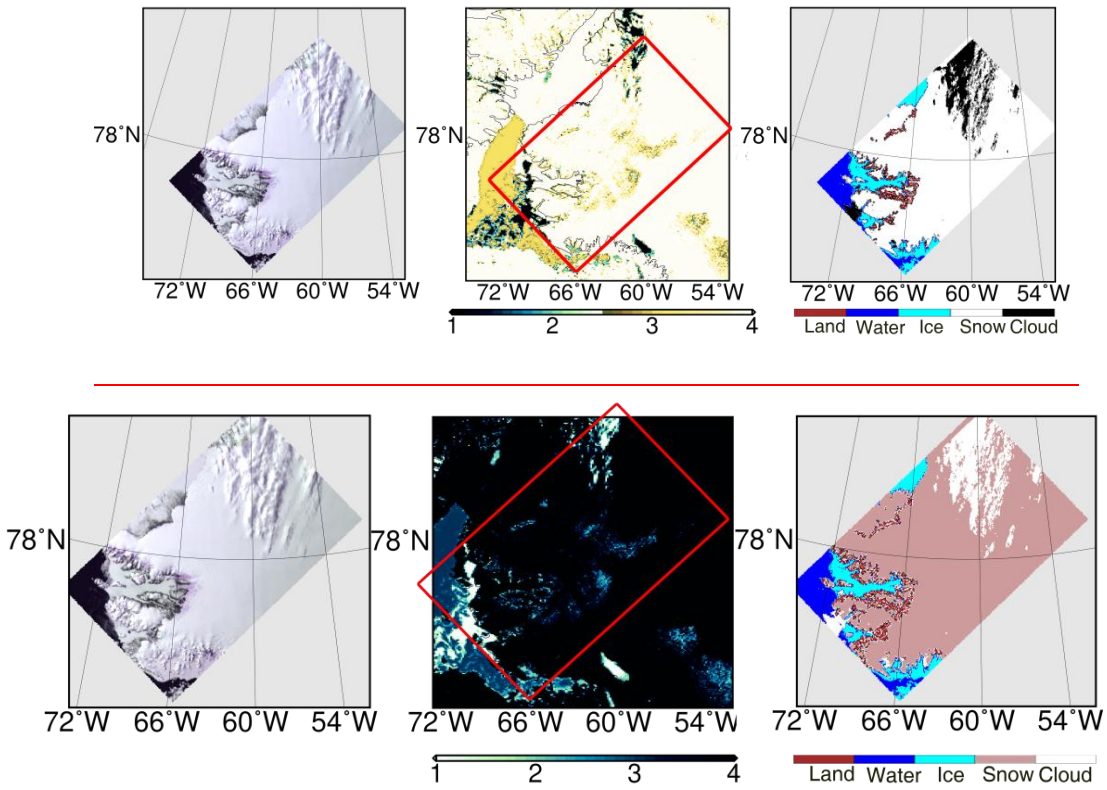


Figure 12. Left panel: RGB false colour image (using 0.67, 0.87 and 0.55 µm channels) of AATSR over Greenland, 18 May 2008, 23h 13min 38s, middle panel: MODIS cloud mask: 1- cloudy, 2- probably cloudy, 3- probably clear, 4- clear, (red rectangle shows the coverage of AATSR) for 23h 5min, right panel: Cloud detection of ASCIA.

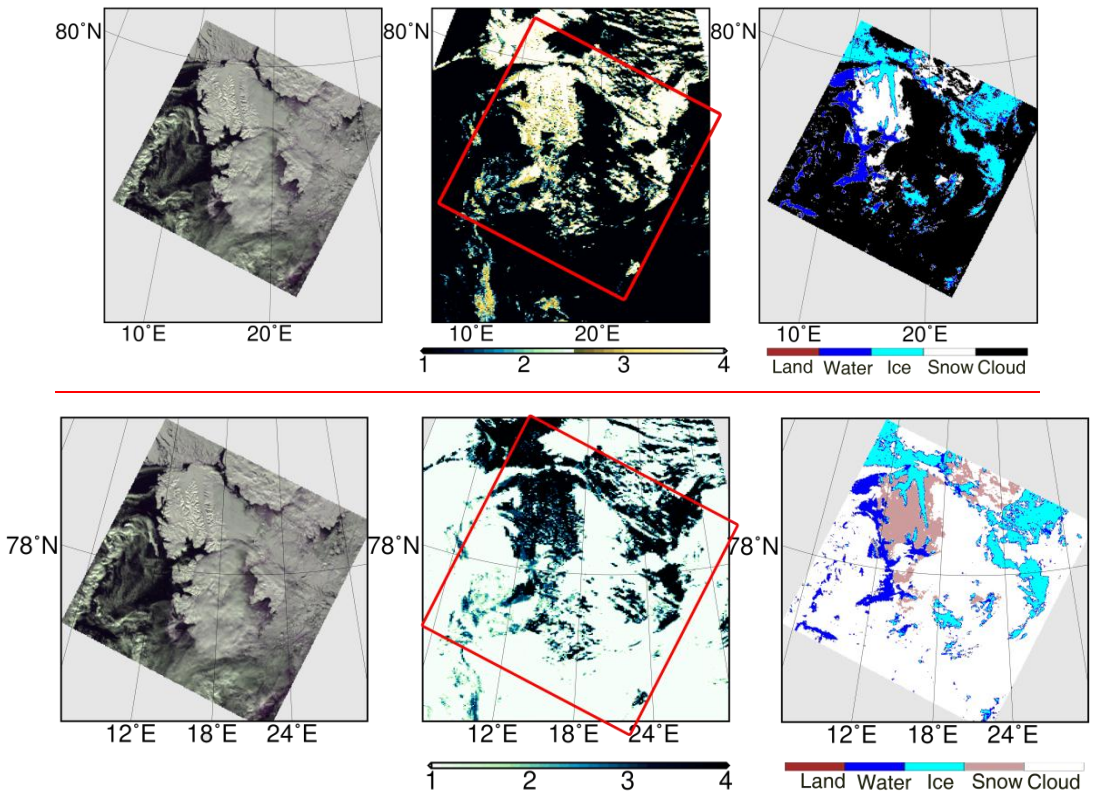
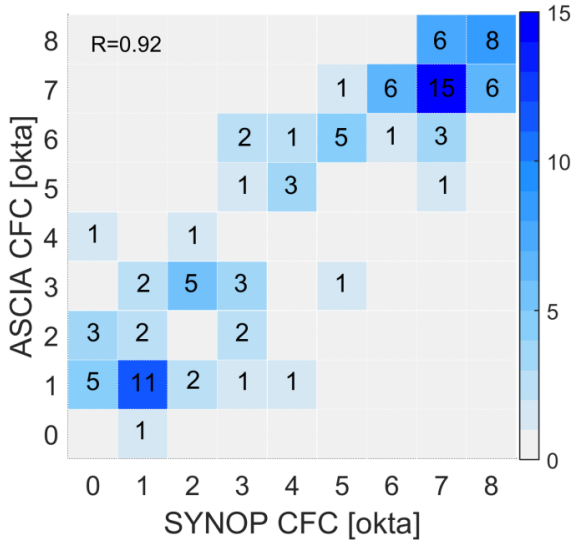
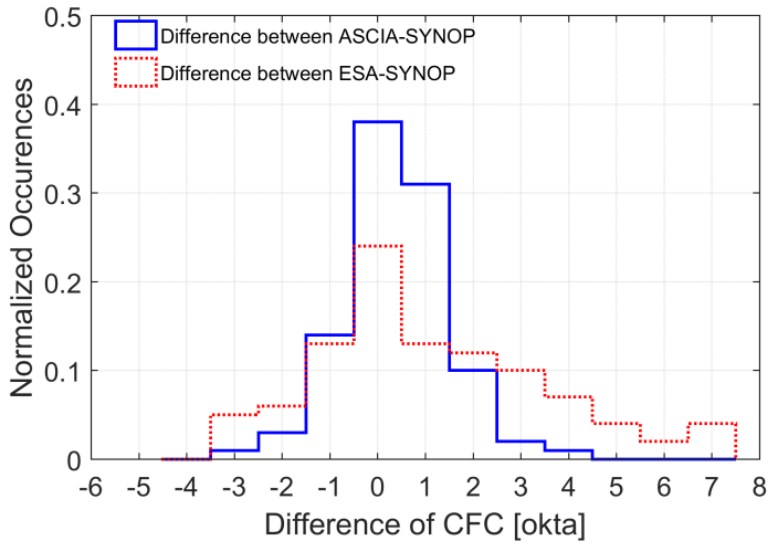


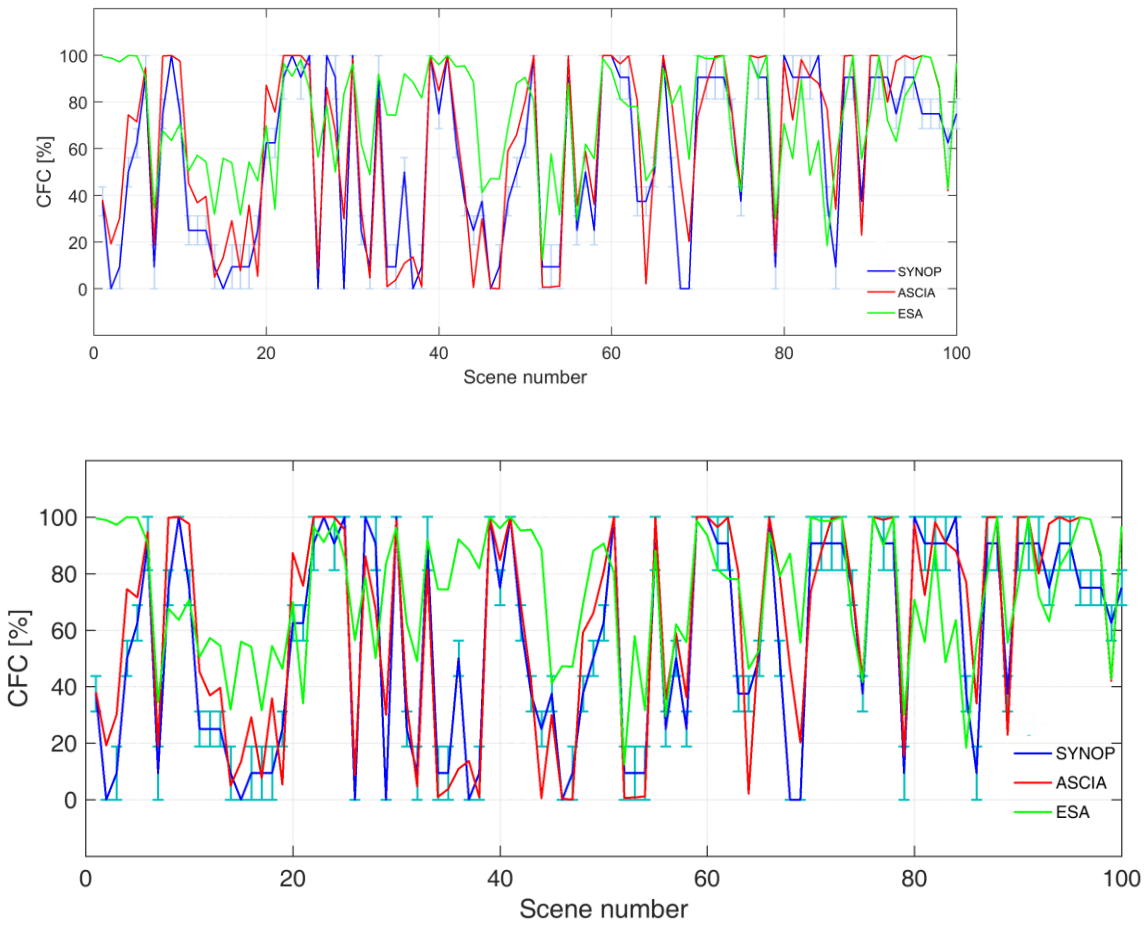
Figure 13. Left panel: The RGB false colour image (using 0.67, 0.87 and 0.55 µm channels) of SLSTR over Svalbard, 18 April 2017, 10hr 15min 6s, Middle panel: MODIS cloud mask: 1- cloudy, 2- probably cloudy, 3- probably clear, 4- clear, (red rectangle shows the coverage of AATSR) for 11h 30m, right panel: Cloud detection of ASCIA.



752 **Figure 14.** Density plot of occurrences of the CFC by ASCIA as a function of SYNOP.



753 **Figure 15.** Histogram of CFC differences (blue: ASCIA minus SYNOP; red: ESA cloud product minus SYNOP).



754 **Figure 16.** CFC in percent by ASCIA (red), SYNOP (blue) and ESA Cloud Product (green) for 100 scenarios of March, May and
755 July 2008 over Svalbard and Greenland. Light blue error bars show the range of percentage values for each okta from SYNOP
756 measurements.