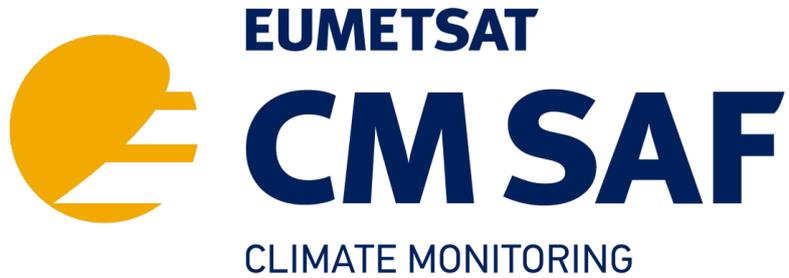


**EUMETSAT Satellite Application Facility on Climate Monitoring**



**Validation Report**

**CM SAF Cloud, Albedo, Radiation data record,  
AVHRR-based, Edition 3 (CLARA-A3) Surface Black-sky,  
White-sky and Blue-sky Albedo**

**DOI: [10.5676/EUM\\_SAF\\_CM/CLARA\\_AVHRR/V003](https://doi.org/10.5676/EUM_SAF_CM/CLARA_AVHRR/V003)**

	<b>TCDR</b>	<b>ICDR</b>
<b>Black-sky albedo from AVHRR-GAC</b>	<b>CM-11222</b>	<b>CM-6221</b>
<b>White-sky albedo from AVHRR-GAC</b>	<b>CM-11223</b>	<b>CM-6223</b>
<b>Blue-sky albedo from AVHRR-GAC</b>	<b>CM-11224</b>	<b>CM-6224</b>

Reference Number:  
Issue/Revision Index:  
Date:

SAF/CM/FMI/VAL/CLARA/SAL  
3.1  
06.02.2023

### Document Signature Table

	Name	Function	Signature	Date
<b>Author</b>	Aku Riihelä Emmihenna Jääskeläinen Viivi Kallio-Myers	CM SAF scientists (FMI)		06/02/2023
<b>Editor</b>	Marc Schröder	CM SAF Science Coordinator		06/02/2023
<b>Approval</b>	CM SAF Steering Group			
<b>Release</b>	Rainer Hollmann	CM SAF Project Manager		

### Distribution List

Internal Distribution	
Name	No. Copies
DWD / Archive	1
CM SAF Team	1

External Distribution		
Company	Name	No. Copies
Public		1

	<b>Validation Report Surface Albedo (SAL) CLARA Edition 3</b>	Doc. No: SAF/CM/FMI/VAL/CLARA/SAL Issue: 3.1 Date: 06.02.2023
---	---	---

## Document Change Record

Issue/ Revision	Date	DCN No.	Changed Pages/Paragraphs
2.2	06.02.2020	SAF/CM/DWD/VAL/GACSAL	Initial release for DRR3.13
2.3	07.05.2020	SAF/CM/DWD/VAL/GACSAL	Updated version after review
2.4	23.09.2020	SAF/CM/DWD/VAL/GACSAL	Layout revision
3.0	19.09.2022	SAF/CM/DWD/VAL/GACSAL	Update for CLARA-A3;  Initial version for joint DRR3.2/ORR
3.1	06.02.2023	SAF/CM/DWD/VAL/GACSAL	Updates following discussions at DRR3.2/ORR  Adapted ICDR processing

## Applicable Documents

Reference	Title	Code
AD 1	CM SAF Product Requirements Document (PRD)	SAF/CM/DWD/PRD/4.1

## Reference Documents

Reference	Title	Code
RD 1	Product User Manual  CM SAF Cloud, Albedo, Radiation data record, AVHRR-based,  Edition 3 (CLARA-A3) Surface Albedo	SAF/CM/FMI/PUM/GAC/SAL/3.1
RD 2	Algorithm Theoretical Basis Document  CM SAF Cloud, Albedo, Radiation data record, AVHRR-based,  Edition 3 (CLARA-A3) Surface Albedo	SAF/CM/FMI/ATBD/GAC/SAL/3.3

Reference	Title	Code
RD 3	Validation Report CM SAF Cloud, Albedo, Radiation data record, AVHRR-based, Edition 2 (CLARA-A2) Surface Albedo	SAF/CM/FMI/VAL/GAC/SAL/2.1

## Table of Contents

Executive Summary .....	12
1. The EUMETSAT SAF on Climate Monitoring .....	14
2. Introduction .....	16
3. Data and Methods .....	17
3.1. Reference observations – Baseline Surface Radiation Network.....	17
3.2. Reference observations – PROMICE Network.....	18
3.3. Reference observations – SHEBA and Tara ice camps .....	19
3.4. CLARA-A3 SAL/WAL/BAL surface albedo estimates .....	20
3.5. Validation strategy against reference observations .....	20
3.6. Intercomparison against MCD43D MODIS-based surface albedo data .....	22
3.7. Performance metrics and requirements .....	22
4. Stability analysis of surface albedo over quasi-stationary regions for CLARA-A3 SAL/WAL/BAL.....	24
5. BAL validation against BSRN in situ observations .....	30
5.1. Spatial representativeness of BSRN sites based on Google Earth Engine data.....	30
5.2. Validation metrics summary, performance against specifications.....	32
5.3. Validation result illustrations of the participating BSRN sites.....	36
5.4. Alert (ALE).....	36
5.5. Bondville (BON).....	37
5.6. Boulder (BOU).....	38
5.7. Cabauw (CAB).....	39
5.8. Desert Rock (DRA) .....	41
5.9. Southern Great Plains (E13).....	42
5.10. Fort Peck (FPE).....	43
5.11. Goodwin Creek (GCR).....	44
5.12. Georg von Neumayer (GVN).....	45
5.13. South Pole (SPO) .....	46

5.14. Sioux Falls (SXF).....	47
5.15. Syowa (SYO).....	48
5.16. Toravere (TOR).....	49
6. Validation of WAL and SAL against filtered diffuse/direct illumination BSRN observations .....	51
7. Validation against PROMICE in situ observations .....	54
7.1. BAL.....	54
7.2. WAL.....	58
7.3. SAL.....	60
8. Validation against Tara and SHEBA drifting ice camp albedo observations over the Arctic Ocean .....	65
9. Comparison against MCD43D51 edition 6.1 and CLARA-A2 SAL .....	68
10. Analysis of consistency between TCDR and its ICDR continuation .....	71
11. Discussion and conclusions .....	74
References.....	77
Glossary.....	79
Appendix A: Google Dynamic World-inferred land cover and climatological albedo for CLARA- A3 grid cells containing BSRN sites.....	81

## List of Tables

Table 1: BSRN stations used in validation .....	17
Table 2: Coordinates, elevations, and start dates of PROMICE sites. Sites with names in green indicate >90% snow/ice cover in containing CLARA NP grid cell.....	18
Table 3: Categorization levels for each performance metric of the CLARA-A3 albedo estimates.....	23
Table 4: Climatological surface albedos prescribed to Dynamic World land use classes for the representativeness analysis .....	30
Table 5: Overpass (level 2) & Monthly mean BAL performance metrics against BSRN observations.	33
Table 6: Pentad mean BAL performance metrics against BSRN observations .....	33
Table 7: Pearson correlations of BAL retrieval parameters against BAL relative bias (vs BSRN) at the overpass level.....	34
Table 8: WAL and SAL monthly and pentad mean bias (rMBE) validation results (at 0.25 degree resolution) against BSRN observations filtered for near-fully diffuse/direct illumination conditions. Metric is relative bias against reference observation [%]. .....	51
Table 9: WAL and SAL monthly and pentad mean validation results for precision (at 0.25 degree resolution) against BSRN observations filtered for near-fully diffuse/direct illumination conditions. Metric is bias-corrected root mean square deviation [bc-rmsd; unitless]. .....	52
Table 10: WAL MM performance metrics and compliance to requirements .....	52
Table 11: WAL PM performance metrics and compliance to requirements .....	53
Table 12: SAL MM performance metrics and compliance to requirements .....	53
Table 13: SAL PM performance metrics and compliance to requirements .....	53
Table 14: BAL MM performance metrics against PROMICE in situ reference observations .....	55
Table 15: BAL PM performance metrics against PROMICE in situ reference observations .....	57
Table 16: PROMICE validation metrics for monthly mean WAL .....	58
Table 17: PROMICE validation metrics for pentad mean WAL.....	59
Table 18: PROMICE validation metrics for monthly mean SAL .....	61
Table 19: PROMICE validation metrics for pentad mean SAL.....	64

## List of Figures

Figure 1: Locations of the PROMICE measurement sites on the Greenland Ice Sheet. From Fausto et al. (2021).....	18
Figure 2: The drift paths of the SHEBA and Tara expeditions on the Arctic Ocean. Tara drift illustrated for period March-September 2007. ....	19
Figure 3 : CLARA-A3 SAL (black-sky albedo) monthly mean retrievals over the designated analysis areas. Markers color coded by mean Sun Zenith Angle of valid observations during the month in	

question. Light blue envelope indicates  $\pm 5\%$  about the interannual mean (black horizontal line). Dark blue envelope indicates median standard deviation of ‘instantaneous’ SAL retrievals. Red region indicates the time period most affected by the atmospheric disturbances of the Mt. Pinatubo eruption in June 1991 (2 years)..... 24

Figure 4: As Figure 3, but for CLARA-A3 WAL (white-sky albedo). Note that standard deviation over snow/ice surfaces is not available for WAL, as the variable is based on statistical relationship between mean black/white-sky albedos. .... 25

Figure 5: As Figure 3, but for CLARA-A3 BAL (Blue-sky albedo). Note that standard deviation over snow/ice surfaces is not available for BAL due to the non-availability of std for WAL over snow and ice. .... 25

Figure 6: The numbers of clear-sky GAC-resolution AVHRR observations available for each monthly mean over each of the stability sites. Color of marker denotes mean SZA of each month. .... 26

Figure 7: Deviation from interannual mean BAL over the stability sites as a function of available clear-sky AVHRR observations. Marker color indicates mean SZA of each month of observations. .... 27

Figure 8: GrIS area's SAL retrievals by month. Thick blue line indicates the May-August mean. .... 28

Figure 9: CLARA-A2 SAL monthly mean estimates for the GrIS area between 1982-2015..... 28

Figure 10: Site representativeness analysis results from BSRN sites at monthly mean scale. Blue circle indicates expected BAL during June-August, based on land cover from Google Dynamic World and associated climatological surface albedos. Orange triangles indicate mean observed in situ albedo at each site. The length of the black line connecting the two indicates the dissimilarity between the two as an indicator of spatial un-representativeness. Finally, green cross with whiskers indicates the mean BAL from CLARA-A3 of the grid cell containing site, the whiskers showing the standard deviation of the June-August mean BAL retrievals during the evaluation period. .... 32

Figure 11: Distribution of performance metrics of monthly mean BAL over BSRN sites, inclusive of both representative and unrepresentative sites. Left: relative bias, center: precision, right: stability. Vertical colored lines indicate thresholds of performance levels..... 34

Figure 12: As Figure 11, but for pentad mean BAL..... 34

Figure 13: Left, BAL level 2 bias vs. BSRN observations as a function of 0.8 micron TOA reflectance of the AVHRR observation. Uncolored markers imply less than 20 observations in said bin. Right, relative BAL bias as a function of 0.8 micron TOA reflectance..... 35

Figure 14: As Figure 13, but against Sun Zenith Angle (SZA) of the observation time. .... 35

Figure 15: (Top) Retrieved CLARA-A3 monthly mean BAL (red circles), pentad mean BAL (black triangles), GAC-resolution (overpass) BAL (magenta x), in situ measured albedo (blue +). (Bottom) Relative retrieval error (bias) of the pentad, monthly mean and instantaneous satellite estimates, markers consistent with the top panel. .... 36

Figure 16: Validation results for Bondville (BON)..... 37

Figure 17: Validation results for Boulder (BOU)..... 38

Figure 18: Validation results for Cabauw (CAB)..... 39

Figure 19: Validation results for Desert Rock (DRA)..... 41

Figure 20: Validation results for Southern Great Plains (E13) ..... 42

Figure 21: Validation results for Fort Peck (FPE)..... 43

Figure 22: Validation results for Goodwin Creek (GCR) ..... 44

Figure 23: Validation results for Georg von Neumayer (GVN)..... 45

Figure 24: Validation results for South Pole (SPO)..... 46

Figure 25: Validation results for Sioux Falls (SXF) ..... 47

Figure 26: Validation results for Syowa (SYO)..... 48

Figure 27: Validation results for Toravere (TOR)..... 49

Figure 28: Top, relative bias (rMBE) of monthly mean CLARA-A3 BAL against PROMICE in situ observations. Each marker denotes one monthly mean at one PROMICE site, colored by site elevation. Red lines indicate +/- 25% rMBE thresholds. Magenta line indicates the mean rMBE for each month across all sites. Bottom, the overall mean rMBE during the evaluation period for the PROMICE sites, binned and shown as a function of site elevation (30 bins). Whiskers indicate std.dev. of rMBE in each bin. .... 54

Figure 29: Top, relative bias (rMBE) of pentad mean CLARA-A3 BAL against PROMICE in situ observations. Each marker denotes one pentad mean at one PROMICE site, colored by site elevation. Red lines indicate +/- 25% rMBE thresholds. Magenta line indicates the mean rMBE for each month across all sites. Bottom, the overall mean rMBE during the evaluation period for the PROMICE sites, binned and shown as a function of site elevation (30 bins). Whiskers indicate std.dev. of rMBE in each bin. .... 56

Figure 30: Precision (bc-rms) of BAL during June-August as a function of snow cover in CLARA grid cell (25 km EASE polar subset) containing PROMICE sites. Snow/ice fraction data from Google Earth Engine’s Dynamic World dataset. Color of markers indicates PROMICE site elevation. The red, magenta, and grey lines indicate the threshold, target, and optimum precision requirement levels, respectively..... 57

Figure 31: Top, relative bias (rMBE) of monthly mean CLARA-A3 WAL against PROMICE in situ observations. Each marker denotes one monthly mean at one PROMICE site, colored by site elevation. Red lines indicate +/- 25% rMBE thresholds. Magenta line indicates the mean rMBE for each month across all sites. Bottom, the overall mean rMBE during the evaluation period for the PROMICE sites, binned and shown as a function of site elevation (30 bins). Whiskers indicate std.dev. of rMBE in each bin. .... 59

Figure 32: Top, relative bias (rMBE) of pentad mean CLARA-A3 WAL against PROMICE in situ observations. Each marker denotes one pentad mean at one PROMICE site, colored by site elevation. Red lines indicate +/- 25% rMBE thresholds. Magenta line indicates the mean rMBE for each month across all sites. Bottom, the overall mean rMBE during the evaluation period for the PROMICE sites, binned and shown as a function of site elevation (30 bins). Whiskers indicate std.dev. of rMBE in each bin. .... 60

Figure 33: Top, relative bias (rMBE) of monthly mean CLARA-A3 SAL against PROMICE in situ observations. Each marker denotes one monthly mean at one PROMICE site, colored by site elevation. Red lines indicate +/- 25% rMBE thresholds. Magenta line indicates the mean rMBE for each month across all sites. Bottom, the overall mean rMBE during the evaluation period for the PROMICE sites,

	<b>Validation Report</b> <b>Surface Albedo (SAL)</b> <b>CLARA Edition 3</b>	Doc. No: SAF/CM/FMI/VAL/CLARA/SAL Issue: 3.1 Date: 06.02.2023
---	---	---

binned and shown as a function of site elevation (30 bins). Whiskers indicate std.dev. of rMBE in each bin. .... 62

Figure 34: Annual standard deviation of SAL and PROMICE-measured albedos under cloud-free conditions. All PROMICE sites considered. .... 63

Figure 35: Top, relative bias (rMBE) of pentad mean CLARA-A3 SWAL against PROMICE in situ observations. Each marker denotes one pentad mean at one PROMICE site, colored by site elevation. Red lines indicate +/- 25% rMBE thresholds. Magenta line indicates the mean rMBE for each month across all sites. Bottom, the overall mean rMBE during the evaluation period for the PROMICE sites, binned and shown as a function of site elevation (30 bins). Whiskers indicate std.dev. of rMBE in each bin. .... 64

Figure 36: Visualization of Tara drift during summer 2007 (left), with red/blue markers denoting its collocatable part, and visualization of the Tara albedo observations vs. SAL/BAL/WAL estimates (right). All data shown for conditions for which SZA<70 deg., Tara data visualized as hourly means. CLARA-A3 SAL/WAL/BAL estimates from the spatiotemporally matched monthly mean grid cells in which the Tara schooner drifted during the course of the summer. Dashed line indicates the relative availability of valid AVHRR observations vs. the summer maximum. Note that the CLARA estimates are at the 25 kilometer spatial resolution while Tara observations are point-like..... 65

Figure 37: As Figure 36, but for the pentad mean SAL/BAL/WAL data during summer 2007..... 66

Figure 38: Top: In situ albedo from SHEBA transect measurements (blue, 5 day average) versus CLARA-A3 albedo pentad mean products – BAL in orange line & markers, with the variability range between SAL and WAL shown with a beige envelope. Faint blue overlaid bar chart indicates the amount of valid GAC-resolution SAL retrievals at the ice camp’s grid cell during the Arctic summer 1998. Below: Relative bias (rMBE) of CLARA-A3 BAL vs. SHEBA in situ measurements. Red lines indicate the target accuracy of 25% relative. .... 67

Figure 39: CLARA-A3 SAL compared with MCD43D51, April-August 2015 mean albedo. Top left: MCD43D51 black-sky albedo. Top right: CLARA-A3 SAL black-sky albedo. Bottom left: Spatially resolved difference between CLARA-A3 SAL and MCD43D51. Bottom right: Zonal means of both albedo products over the joint coverage area (valid data in both records). For the comparison, both MCD43D51 and CLARA-A3 SAL have been normalized to SZA 60 degrees using the equation from Dickinson (1983). .... 68

Figure 40: Mean differences in black-sky albedo estimates between CLARA-A3 and MCD43D51 shown as a function of USGS land cover classes. Y-axis bars of the classes indicate standard deviation of differences over grid cells belonging to each class (converted from the GLOBCOVER2009 land cover dataset)..... 69

Figure 41: Comparison of black-sky albedo (SAL) from CLARA-A2 (top left) and CLARA-A3 (top right) data records. Annual means of year 2015. Bottom left: Difference in SAL between A3 and A2. Bottom right: Zonal means of SAL in CLARA-A2 and A3. Ocean surfaces normalized to SZA of 60 degrees in CLARA-A3 to improve comparability..... 70

Figure 42: Difference (TCDR minus ICDR) for black-sky albedo SAL in the monthly means of July-December 2020. .... 71

Figure 43: Difference (TCDR minus ICDR) for white-sky albedo WAL in the monthly means of July-December 2020. .... 72

	<p style="text-align: center;"><b>Validation Report Surface Albedo (SAL) CLARA Edition 3</b></p>	<p>Doc. No: SAF/CM/FMI/VAL/CLARA/SAL Issue: 3.1 Date: 06.02.2023</p>
---	--	--

Figure 44: Difference (TCDR minus ICDR) for blue-sky albedo BAL in the monthly means of July-December 2020. .... 72

Figure 45: Zonal means of TCDR-ICDR differences for black-sky surface albedo in the monthly means of July-December 2020. Blue color is for ICDR with Metop-C excluded (distributed version), orange for Metop-C included in the means. .... 73

Figure 46: Summary of monthly mean BAL retrieval bias against BSRN in situ observations. Color of markers indicates magnitude of bias (MBE), height of each (monthly) marker indicates the relative amount of observations available for the BAL retrieval, i.e. larger bars indicate more available observations. Data based on grid cell scale validation, i.e. point-to-pixel comparisons. .... 74

	<p style="text-align: center;"><b>Validation Report</b>  <b>Surface Albedo (SAL)</b>  <b>CLARA Edition 3</b></p>	<p>Doc. No: SAF/CM/FMI/VAL/CLARA/SAL  Issue: 3.1  Date: 06.02.2023</p>
---	--	--

## Executive Summary

This CM SAF report provides information on the validation of the CM SAF CLARA-A3 SAL data record, consistent of the climate data record (CDR), covering the period 1979-2020 and its consistent, operational extension with the Interim CDR (ICDR), derived from Advanced Very High Resolution Radiometer (AVHRR) observations on-board National Oceanic and Atmospheric Administration (NOAA) platforms and EUMETSAT’s Metop satellite constellation.

The shortwave surface albedo, the ratio of reflected solar flux to the incoming solar flux, is an important driver of the surface energy budget of the Earth. Variations and trends in surface albedo can influence near-surface air temperatures as well as the melt-freeze cycles of sea ice and snow cover. Accurate determination of surface albedo is particularly important in the Polar regions, where snow and ice dynamics largely govern the surface energy budget. The CLARA-A3 SAL data contains estimates of three different quantifications of surface albedo: surface albedo under unidirectional illumination (black-sky albedo), surface albedo under completely diffuse illumination (white-sky albedo), and surface albedo under the prevailing (ambient) illumination conditions valid for each time and place in the record (blue-sky albedo).

The validation of the CLARA-A3 SAL (Edition 3) follows the same procedure as the validation of CLARA-A2 SAL [RD 3]. We evaluate the product accuracy against reference ground observations of surface albedo from several sources: The Baseline Surface Radiation Network (BSRN), The Programme for the Monitoring of the Greenland Ice Sheet (PROMICE), and the Surface Heat Budget of the Arctic Ocean (SHEBA) and Tara-Arctic field campaigns on the Arctic sea ice. The comparability of the large-scale satellite observations against the point-like in situ observations is problematic (the “point-to-pixel problem”). To minimize the effect of this, we include in the report the mean performance of individual (GAC-resolution) SAL observations against the BSRN in situ measurements (N=144 950). We also include additional analysis on the likely impact of spatial representativeness issues on the observed bias. Finally, we compare the CLARA-A3 albedo estimates against widely used MODIS albedo data (MCD43) as well as the predecessor data record CLARA-A2 for the black-sky albedo. Also, we intercompare the full climate data record with its follow-up Interim Climate Data Record (ICDR) for a six-month period to ascertain coherence and continuity.

The product requirements and achieved (mean) performance for CLARA-A3 SAL Edition 3 CDR [AD 1] are summarized in the tables on the next page. For quantitative details on the performance against the various components of the validation, please see the full report below. the mean difference between TCDR and ICDR is small in the global scale.

Product	Accuracy requirement (relative bias)	Achieved performance level or level range
Black-sky surface albedo (SAL)	<5% optimum <15% target <20% threshold	Optimum / Target
White-sky surface albedo (WAL)	<5% optimum <25% target <50% threshold	Target
Blue-sky surface albedo (BAL)	<5% optimum <25% target <50% threshold	Optimum / Target

Product	Precision requirement (bc-rms)	Achieved performance level or level range
Black-sky surface albedo (SAL)	<0.1 / <0.05 target <0.15 / <0.1 threshold	Target
White-sky surface albedo (WAL)	<0.05 optimum <0.1 target <0.15 threshold	Target / Threshold <i>(some sites near ice sheet margins show non-threshold performance)</i>
Blue-sky surface albedo (BAL)	<0.05 optimum <0.1 target <0.15 threshold	Target

Product	Stability requirement (decadal trend in bias)	Achieved performance level or level range
Black-sky surface albedo (SAL)	<2% optimum <10% target <15% threshold	Optimum
White-sky surface albedo (WAL)	<2% optimum <15% target <20% threshold	Target
Blue-sky surface albedo (BAL)	<2% optimum <15% target <20% threshold	Optimum / Target

	<p style="text-align: center;"><b>Validation Report</b>  <b>Surface Albedo (SAL)</b>  <b>CLARA Edition 3</b></p>	<p>Doc. No: SAF/CM/FMI/VAL/CLARA/SAL  Issue: 3.1  Date: 06.02.2023</p>
---	--	--

## 1. The EUMETSAT SAF on Climate Monitoring

The importance of climate monitoring with satellites was recognized in 2000 by EUMETSAT Member States when they amended the EUMETSAT Convention to affirm that the EUMETSAT mandate is also to “contribute to the operational monitoring of the climate and the detection of global climatic changes”. Following this, EUMETSAT established within its Satellite Application Facility (SAF) network a dedicated centre, the SAF on Climate Monitoring (CM SAF, <http://www.cmsaf.eu>).

The consortium of CM SAF currently comprises the Deutscher Wetterdienst (DWD) as host institute, and the partners from the Royal Meteorological Institute of Belgium (RMIB), the Finnish Meteorological Institute (FMI), the Royal Meteorological Institute of the Netherlands (KNMI), the Swedish Meteorological and Hydrological Institute (SMHI), the Meteorological Service of Switzerland (MeteoSwiss), the Meteorological Service of the United Kingdom (UK MetOffice), and the Centre National de la Recherche Scientifique (CNRS). Since the beginning in 1999, the EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF) has developed and will continue to develop capabilities for a sustained generation and provision of Climate Data Records (CDR’s) derived from operational meteorological satellites.

In particular, the generation of long-term data records is pursued. The ultimate aim is to make the resulting data records suitable for the analysis of climate variability and potentially the detection of climate trends. CM SAF works in close collaboration with the EUMETSAT Central Facility and liaises with other satellite operators to advance the availability, quality and usability of Fundamental Climate Data Records (FCDRs) as defined by the Global Climate Observing System (GCOS). As a major task the CM SAF utilizes FCDRs to produce records of Essential Climate Variables (ECVs) as defined by GCOS. Thematically, the focus of CM SAF is on ECVs associated with the global energy and water cycle.

Another essential task of CM SAF is to produce data records that can serve applications related to the Global Framework of Climate Services initiated by the WMO World Climate Conference-3 in 2009. CM SAF is supporting climate services at national meteorological and hydrological services (NMHSs) with long-term data records but also with data records produced close to real time that can be used to prepare monthly/annual updates of the state of the climate. Both types of products together allow for a consistent description of mean values, anomalies, variability and potential trends for the chosen ECVs. CM SAF ECV data records also serve the improvement of climate models both at global and regional scale.

As an essential partner in the related international frameworks, in particular WMO SCOPE-CM (Sustained COordinated Processing of Environmental satellite data for Climate Monitoring), the CM SAF - together with the EUMETSAT Central Facility, assumes the role as main implementer of EUMETSAT’s commitments in support to global climate monitoring. This is achieved through:

- Application of highest standards and guidelines as lined out by GCOS for the satellite data processing,
- Processing of satellite data within a true international collaboration benefiting from developments at international level and pollinating the partnership with own ideas and standards,
- Intensive validation and improvement of the CM SAF climate data records,
- Taking a major role in data record assessments performed by research organisations such as WCRP (World Climate Research Program). This role provides the CM SAF

	<p style="text-align: center;"><b>Validation Report</b>  <b>Surface Albedo (SAL)</b>  <b>CLARA Edition 3</b></p>	<p>Doc. No: SAF/CM/FMI/VAL/CLARA/SAL  Issue: 3.1  Date: 06.02.2023</p>
---	--	--

with deep contacts to research organizations that form a substantial user group for the CM SAF CDRs,

- Maintaining and providing an operational and sustained infrastructure that can serve the community within the transition of mature CDR products from the research community into operational environments.

A catalogue of all available CM SAF products is accessible via the CM SAF webpage, <http://www.cmsaf.eu/>. Here, detailed information about product ordering, add-on tools, sample programs and documentation is provided.

	<p style="text-align: center;"><b>Validation Report</b>  <b>Surface Albedo (SAL)</b>  <b>CLARA Edition 3</b></p>	<p>Doc. No: SAF/CM/FMI/VAL/CLARA/SAL  Issue: 3.1  Date: 06.02.2023</p>
---	--	--

## 2. Introduction

Shortwave surface albedo is defined as the ability of a surface to reflect solar radiation, i.e. it is the ratio of the reflected shortwave solar flux to the incoming one. The albedo of natural surfaces varies from ~5-6% (water, concrete) up to 90% (fresh, small-grained snow). Determination of the Earth's surface energy budget is dependent on our ability to accurately and robustly monitor global surface albedo. Thus, climate change studies require information about surface albedo and its changes. This has been acknowledged by GCOS in through naming surface albedo as an Essential Climate Variable (ECV).

In the polar and boreal regions, seasonal snow cover and changes in polar sea ice cover cause considerable changes in surface albedo (and vice versa through feedback effects). Changes in polar surface albedo have global effects (Hudson et al., 2011), and changes in surface albedo of the Arctic sea ice are closely tied to its mass budget (Holland et al., 2010). Therefore, monitoring polar albedo is of highest importance.

This report provides information on the validation of the surface albedo estimates in the third edition of the CM SAF CLARA data record family. Now covering 1979-2020 on a global scale, the CLARA-A3 is based on data from every Advanced Very High Resolution Radiometer (AVHRR) satellite-based optical remote sensing instrument ever flown, from TIROS-N up to the MetOp satellite constellation.

In the CLARA-A3 data record, three surface albedo estimates are provided: black-sky albedo (directional-hemispherical reflectance), white-sky albedo (bidirectional reflectance under diffuse illumination), and blue-sky albedo (bidirectional reflectance under ambient illumination). The data are provided at pentad (5-day) and monthly mean temporal resolution, in a global spatial grid of 0.25 degree resolution (polar region subsets available at 25 km equal-area grid).

This investigation consists of the following components:

1. Analysis of temporal stability in retrieved surface albedo over a selection of sites with naturally (quasi-)stable surface properties.
2. Collocation and comparison (point-to-pixel) against Baseline Surface Radiation Network (BSRN) sites to address quality over vegetation and seasonal snow.
3. Collocation and comparison (point-to-pixel) against The Programme for Monitoring of the Greenland Ice Sheet (PROMICE) sites to address quality over glaciated surfaces.
4. Collocation and comparison (point-to-pixel) against observations of sea ice albedo from the Tara and SHEBA (Surface Heat Budget of the Arctic Ocean) campaigns.
5. Resampling and intercomparison with MODIS surface albedo dataset MCD43 to assess consistency with this widely used data record.
6. Intercomparison with the predecessor data record CLARA-A2 to assess consistency and effects of input and algorithm changes.
7. Analysis of coherence between the CLARA-A3 thematic climate data record (TCDR) and its follow-up ICDR data for a six-month overlap period.

We first provide a description of the reference data, followed with details of the processing strategy for each of the components of the validation. Validation results are described separately for each component; the report concludes with a brief discussion of the significance of the results.

### 3. Data and Methods

#### 3.1. Reference observations – Baseline Surface Radiation Network

As in the validations of the preceding CLARA surface albedo data records, the needs for quality-controlled and sufficiently long reference data lead us to choose the Baseline Surface Radiation Network (BSRN; Driemel et al., 2018) as the primary reference data source. The BSRN stations listed in Table 1 were used as reference. The choice of stations was done also for practical reasons; only a fraction of the BSRN stations worldwide record the reflected shortwave radiation necessary to calculate the surface albedo. Furthermore, to promote comparability in performance between the different CLARA albedo records, we maintain the site selection used in the original CLARA-A1 validation.

**Table 1:** BSRN stations used in validation

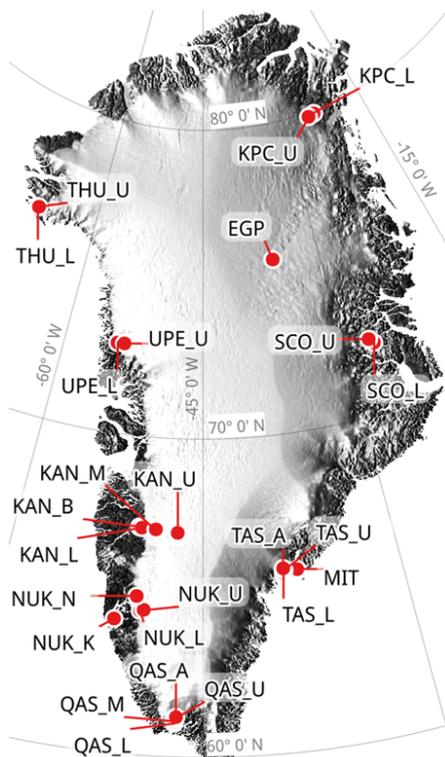
Station code	Name	Latitude (N)	Longitude (E)	Elevation (m)	Time period	Meas. height (m)
<b>ALE</b>	Alert	82.4900	-62.4200	127	2004-2013	2
<b>BON</b>	Bondville	40.0667	-88.3667	213	1995-2019	10
<b>BOU</b>	Boulder	40.0500	-105.0070	1577	1992-2015	300
<b>CAB</b>	Cabauw	51.9711	4.9267	0	2013-2019	2
<b>DRA</b>	Desert Rock	36.6260	-116.0180	1007	1998-2019	10
<b>E13</b>	Southern Great Plains	36.6050	-97.4850	318	1994-2018	10
<b>FPE</b>	Fort Peck	48.3167	-105.1000	634	1995-2019	10
<b>GCR</b>	Goodwin Creek	34.2547	-89.8729	98	1995-2019	10
<b>GVN</b>	Georg von Neumayer	-70.6500	-8.2500	42	1992-2019	2
<b>SPO</b>	South Pole	-89.9830	-24.7990	2800	1992-2017	2
<b>SXF</b>	Sioux Falls	43.7300	-96.6200	473	2003-2019	10
<b>SYO</b>	Syowa	-69.0050	39.5890	18	1998-2019	2
<b>TOR</b>	Toravere	58.2540	26.4620	70	1999-2019	2

Different BSRN sites exhibit different timeliness in the delivery of their measurements to the BSRN database. Therefore, the evaluation record lengths used here vary by evaluation site.

### 3.2. Reference observations – PROMICE Network

The Programme for Monitoring of the Greenland Ice Sheet (PROMICE) operates 25 automated weather stations on the Greenland Ice Sheet (Fausto et al., 2021). Kipp & Zonen pyranometers provide radiative flux data so that deriving shortwave broadband surface albedo is possible for validation purposes. Their locations are illustrated in Figure 1. Measurement height at PROMICE sites is generally between 2 and 3 meters above the ice, but this may vary as the measurement booms move with ice flow and melt.

**Table 2:** Coordinates, elevations, and start dates of PROMICE sites. Sites with names in green indicate >90% snow/ice cover in containing CLARA NP grid cell.



**Figure 1:** Locations of the PROMICE measurement sites on the Greenland Ice Sheet. From Fausto et al. (2021).

Station	Lat	Lon	Elevation	Start
KPC_L	79,9108	24,0828	370	17/07/2008
KPC_U	79,8347	25,1662	870	17/07/2008
EGP	75,6247	35,9748	2660	01/05/2016
SCO_L	72,223	26,8182	460	21/07/2008
SCO_U	72,3933	27,2333	970	21/07/2008
MIT	65,6922	37,828	440	03/05/2009
TAS_L	65,6402	38,8987	250	23/08/2007
TAS_U	65,6978	38,8668	570	15/08/2007
TAS_A	65,779	38,8995	890	28/08/2013
QAS_L	61,0308	46,8493	280	24/08/2007
QAS_M	61,0998	46,833	630	11/08/2016
QAS_U	61,1753	46,8195	900	07/08/2008
QAS_A	61,243	46,7328	1000	20/08/2012
NUK_L	64,4822	49,5358	530	20/08/2007
NUK_U	64,5108	49,2692	1120	20/08/2007
NUK_K	64,1623	51,3587	710	28/07/2014
NUK_N	64,9452	49,885	920	25/07/2010
KAN_B	67,1252	50,1832	350	13/04/2011
KAN_L	67,0955	49,9513	670	01/09/2008
KAN_M	67,067	48,8355	1270	02/09/2008
KAN_U	67,0003	47,0253	1840	04/04/2009
UPE_L	72,8932	54,2955	220	17/08/2009
UPE_U	72,8878	53,5783	940	17/08/2009
THU_L	76,3998	68,2665	570	09/08/2010
THU_U	76,4197	68,1463	760	09/08/2010
CEN	77,1333	61,0333	1880	23/05/2017

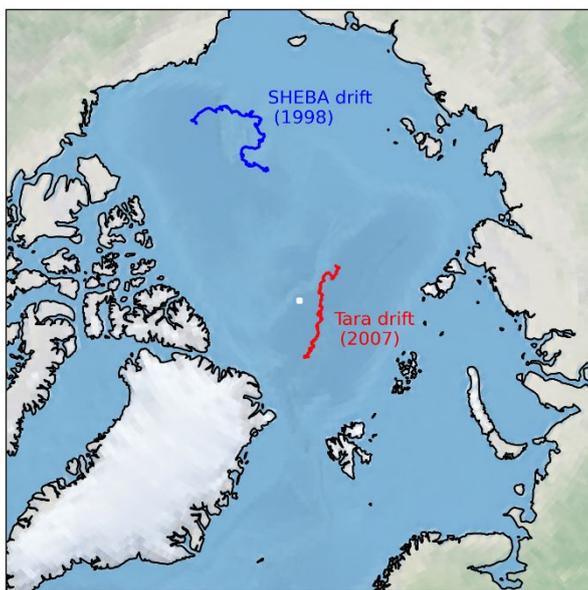
The accuracy of the albedo measurements has been (conservatively) reported as 10% (Fausto et al., 2021) with coverage factor assumed as  $k=1$ . The lengths of data records vary across the sites, but here all available sites have been included in the analysis to maximize spatial coverage across the ice sheet. It should be noted, though, that a number of the sites are located close to the ice sheet margins, where substantial spatial representativeness issues are expected to emerge at the 25 km CLARA-A3 grid cell scale due to much of the grid cell not

being covered by snow and ice. Therefore, we employed the Dynamic World dataset from Google Earth Engine (Brown et al., 2022) to classify land cover fractions in each grid cell containing PROMICE sites. The sites with <90% snow/ice cover in the CLARA grid cell are colored green in Table 2. The validation results are reported for both the full PROMICE group and the more representative set of sites with comprehensive snow/ice cover.

### 3.3. Reference observations – SHEBA and Tara ice camps

The best available in situ data for the validation of sea ice albedo estimates over the Arctic Ocean comes from two expeditions. The first is the SHEBA (Surface HEat Budget of the Arctic Ocean; Uttal et al., 2002) campaign, where a drifting ice camp was set up with the Canadian icebreaker *Des Groseilliers* between October 1997 - October 1998. A wide variety of atmospheric and sea ice measurements were undertaken during the period, including regular albedo transect measurements during the summer season (Perovich et al., 2002). Figure 2 visualizes the drift of the SHEBA ice camp.

These 200-m transects of surface albedo data serve as the reference data source for CLARA albedo validation from the SHEBA campaign. The transects provide enhanced spatial coverage relative to the point-like measurements available from stationary radiative flux measurement stations; on the other hand, the measurements then cannot be operated continuously but are available with a temporal spacing of a few days typically. Uncertainty is reported as 0.01, with the assumption of coverage factor  $k=1$ .



**Figure 2:** The drift paths of the SHEBA and Tara expeditions on the Arctic Ocean. Tara drift illustrated for period March-September 2007.

The second data source for sea ice albedo is the Tara expedition covering Arctic summer 2007. The French schooner *Tara* was frozen into the ice in September 2006, and eventually drifted free of the ice again in Fram strait in December 2007 after 15 months in the icepack. (Gascard et al., 2008). Continuous radiative flux and surface albedo measurements were performed at the ice camp between late March and September, providing a temporally well-resolved look into the surface albedo dynamic of sea ice during the summer melt (Vihma et

	<p style="text-align: center;"><b>Validation Report Surface Albedo (SAL) CLARA Edition 3</b></p>	<p>Doc. No: SAF/CM/FMI/VAL/CLARA/SAL Issue: 3.1 Date: 06.02.2023</p>
---	--	--

al., 2008), though with the restraint of being a point-like measurement. We have previously estimated the accuracy of the Tara sea ice albedo measurements to be 2-4% (Riihelä et al., 2010). Coverage factor for this uncertainty is assumed as  $k=1$ .

For Tara, measurement height for radiative fluxes is ~2 m above ice surface. For SHEBA, this is not reported, but the mobile nature of the transect measurements require a relatively low measurement height, assumed as ~1.5 m above ice surface.

### 3.4. CLARA-A3 SAL/WAL/BAL surface albedo estimates

All CLARA-A3 SAL/WAL/BAL estimates are not normalized any specific Sun Zenith Angle, but rather represent the mean solar illumination conditions of their time and place of record. As the SZA normalization of snow and ice in particular is challenging – because melting snow does not produce a symmetric diurnal variation in albedo – we elected not to normalize any albedo estimates to maintain internal consistency. As SZA data is provided for all CLARA-A3 albedo data records, users may define their normalization algorithms of choice for the data. The validation against in situ references accounts for the non-normalized nature of the albedo data. In the MCD43D intercomparison, first-order normalization to local noon conditions is carried out to promote comparability to the MODIS-based data record.

This report does not repeat the descriptions of the retrieval algorithm. A broad overview of the algorithm is available in the PUM [RD 1], and a much more detailed description is available in the ATBD [RD 2].

### 3.5. Validation strategy against reference observations

The wide variety of in situ data used in this validation implies that a similar variety of preprocessing choices and strategies is necessary to carry out the comparison against CLARA-A3 albedo estimates. Here, we outline the validation method per in situ data source, also outlining any QA-related preprocessing and filtering of the in situ observations.

#### **BSRN:**

The basic validation strategy is to compare the SAL/WAL/BAL estimates against in situ reference albedo measurements through spatiotemporal collocation. We acknowledge that this ‘point-to-pixel’ validation contains an inherent weakness in that the spatial representativeness of the in situ measurement may be limited against the areal mean observed by the satellite and its gridded end products. To compensate, the areal representativeness of each site is analyzed and discussed, and furthermore, the primary performance metrics of the analysis are not evaluated at the grid cell scale (0.25 degrees or 25 km), but rather at the ‘instantaneous’ GAC resolution of the AVHRR overpasses, enabled by the recording of albedo estimates at the validation sites during the data record computation.

To ensure comparability with SAL/WAL/BAL within the bounds of spatial representativeness, the in situ measurements (which occur under blue-sky conditions) are preprocessed to select, for each time period evaluated, only observations which fulfill the following conditions:

- SZA < 70 degrees
- Observed albedo is physically realistic, i.e.  $0 < \alpha < 1$

	<b>Validation Report Surface Albedo (SAL) CLARA Edition 3</b>	Doc. No: SAF/CM/FMI/VAL/CLARA/SAL Issue: 3.1 Date: 06.02.2023
---	---	---

- Observations occur in a 15-minute time slot centered on a valid clear-sky SAL observation at the validation site

The condition imply that the comparable in situ quantity to BAL is the temporal mean of clear-sky in situ albedo observations, as SAL is the temporal mean of clear-sky reflectance observations.

However, because BAL is defined as the mean blue-sky albedo over each time period given the estimated fractions of direct and diffuse illumination during all AVHRR overpasses (not only clear-sky), a secondary analysis is also carried out comparing monthly mean BAL against all valid BSRN in situ observations with SZA<70 degrees during each time period, irrespective of clear/cloudy sky conditions, with the results presented separately. The difference between these BAL evaluations could be understood as the difference between assessing BAL accuracy during clear-sky conditions vs. BAL accuracy during the full month/pentad, with consideration for the full variability in ambient cloudy/clear conditions.

Similarly, WAL is validated with in situ observations which are estimated to correspond to fully cloudy conditions with near-isotropic illumination at the surface level. The identification condition is that diffuse incoming irradiance is => 98% of total surface irradiance. A corresponding analysis is carried out for SAL, inversely selecting in situ observations with =>98% direct incoming solar radiation fraction.

We emphasize that the validation of SAL and WAL remains approximative even after these exclusions because validation at the GAC resolution level is not possible (no observations under cloudy skies). Therefore, the spatial representativeness issues inherent in the ~25 km resolution level of the end products are expected to become manifested in the results.

We report and illustrate validation results for each site in detail to promote comparability with the validation reports of previous CLARA editions, and also since BSRN observations represent the highest reference data quality due to their constant maintenance and quality monitoring.

## **PROMICE**

The PROMICE validation follows the ‘combined clear/cloudy sky’ BSRN approach due to the fact that clear-sky SAL observations are not recorded for PROMICE sites as they are for BSRN sites. While cloud-free periods could be estimated from global radiation records, the AVHRR observations in CLARA only cover a fraction of these. Therefore, all valid PROMICE albedo observations with SZA < 70 degrees are averaged over each time period in question to form the (point-like) reference to the BAL estimate. Precise overpass tracking for PROMICE sites is planned for future CLARA releases’ validation efforts.

As the PROMICE sites do not separately measure direct and diffuse illumination, validation of WAL is reliant on estimation of cloudy-sky periods from the in situ data (threshold >=99% cloud cover), which is admittedly uncertain. Similarly, SAL validation relies on estimation of clear-sky periods in the PROMICE record (cloud cover = 0).

Also, as noted earlier, because the PROMICE sites near the edges of the ice sheet will almost certainly suffer from poor spatial representativeness at the 25km CLARA-A3 resolution, the

	<p style="text-align: center;"><b>Validation Report</b>  <b>Surface Albedo (SAL)</b>  <b>CLARA Edition 3</b></p>	<p>Doc. No: SAF/CM/FMI/VAL/CLARA/SAL  Issue: 3.1  Date: 06.02.2023</p>
---	--	--

validation metrics are primarily presented with these sites excluded. For completeness and comparison, validation metrics are also provided for the entire PROMICE site catalogue.

### **Tara & SHEBA**

Due to the constant drift of the ice camps, SAL recording at overpass level is not available. Furthermore, the ice camp data are different, with SHEBA observations being spatially resolved but a snapshot in the temporal sense, and vice versa for the Tara observations. Therefore, the validation approach is simply to aggregate and average all valid observations over each BAL period, screening the Tara observations for SZA exceeding 70 degrees, and spatially averaging each of the 200-meter SHEBA transects.

### **3.6. Intercomparison against MCD43D MODIS-based surface albedo data**

Among the various MODIS-based surface albedo products, MCD43D51 to MCD43D53 offer daily shortwave broadband surface albedo for land surfaces at 30 arc-sec spatial resolution (~1000 m). Given their algorithmic maturity and broad scientific use base, they are suitable for a near-global intercomparison against CLARA-A3 albedo estimates. Due to the much coarser resolution of CLARA-A3 albedo products, the MCD data were resampled with drop-in-bucket averaging, i.e. all MCD grid cells belonging to each CLARA grid cell were determined and averaged. This operation is much more computationally demanding compared to the common (and here inaccurate) nearest-neighbor resampling. Considering also the I/O load required to download and process a multitude of MCD data (daily global data files at ~1km resolution), the intercomparison period was set to April-August 2015. For the chosen period, CLARA data was available from several AVHRR satellites, and the latest MODIS Collection 6.1 albedo data had been made available in time for this validation exercise.

However, as the MCD albedo products are provided as normalized to local solar noon SZA, the unnormalized CLARA albedo data need to be normalized before comparison. For vegetated land surfaces, the SZA normalization by Dickinson (1983) is a common and reasonably accurate choice. To implement the normalization at the 0.25 degree resolution level, three processing stages were required:

- 1) The dominant land use class in each CLARA grid cell was determined from the GLOBCOVER2009 data set
- 2) Average local noon SZA for each CLARA grid cell for the April-August period was calculated
- 3) CLARA-A3 albedo was normalized with the Dickinson equation

After resampling of the MCD data and SZA normalization of CLARA-A3, the differences could be examined over the full near-global (union) coverage of the datasets. Note that sea ice and Antarctica were not a part of the area where both data records provided retrievals.

### **3.7. Performance metrics and requirements**

The performance of CLARA-A3 SAL/WAL/BAL estimates is assessed through three indicators of quality: accuracy, precision, and stability (of bias), represented by the metrics (relative) bias, bias-corrected RMS error, and decadal trend in bias, respectively. For each metric, predetermined limits classify observed performance to categories of “threshold”, “target”, and “optimum”. Obtaining threshold-level (or better) performance for each metric is considered the

minimum for enabling public release. The target and threshold performance levels are defined for 90% of cases over flat terrain. Table 3 details the categorization levels for the performance of the metrics.

The metrics are defined as follows:

Bias = mean deviation or mean absolute deviation

Precision = bias-corrected root mean square difference [bc-rms]

Stability = decadal trend of bias (compared to a reference data record), i.e. decadal trend magnitude in bias over reference sites

**Table 3:** Categorization levels for each performance metric of the CLARA-A3 albedo estimates.

PRD Identifier	Name	Accuracy [bias] (threshold / target / optimum)	Precision [bc-rms] (threshold / target / optimum)	Stability [decadal] (threshold / target / optimum)
CM-11222	CLARA-A3 SAL	20% / 15% / 5% or 0.005 absolute	0.15 / 0.1 0.1 / 0.05 (pentads)	15% / 10% / 2%
CM-11223	CLARA-A3 SAW	50% / 25% / 5%	0.15 / 0.1 / 0.05	20% / 15% / 2%
CM-11224	CLARA-A3 SAB	50% / 25% / 5%	0.15 / 0.1 / 0.05	20% / 15% / 2%

Here, bias is calculated simply as the rMBE (relative mean bias error) – the relative difference between estimate and reference albedo. Precision is calculated with the standard RMS error equation with bias removal in squared space. For bias stability, we calculate decadal trends in bias over each validation site with the Theil-Sen trend estimator (Theil, 1950; Sen, 1968), which is by design outlier-resistant. The mean of the sites' decadal bias trends is the final metric for stability.

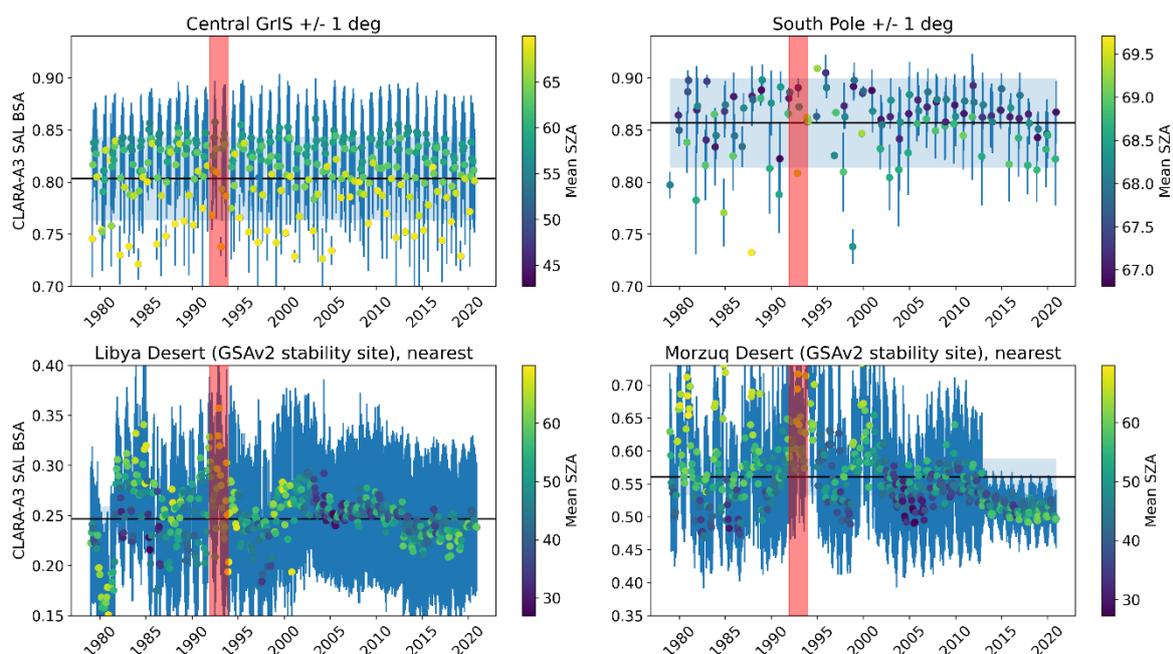
## 4. Stability analysis of surface albedo over quasi-stationary regions for CLARA-A3 SAL/WAL/BAL

Quantification of the stability in the surface albedo data record is necessary to assess the data record's capacity for e.g. trend studies. The quantitative aspect of stability assessed here relates to the decadal stability of biases in estimated surface albedo vs. reference in situ observations. However, in addition, a commonly accepted and adopted practice to assess surface albedo estimate stability is to examine the temporal evolution and variability over sites where natural albedo variability is expected to be minimal. These (quasi-)stationary sites are typically over inner parts of ice sheets where the snow surfaces do not experience seasonal melt, and over deserts with little or no vegetation.

Following this approach, and maintaining a similar practice from the evaluation of the CLARA-A2 SAL data record, we selected four sites to qualitatively assess CLARA-A3 SAL/WAL/BAL stability. These are:

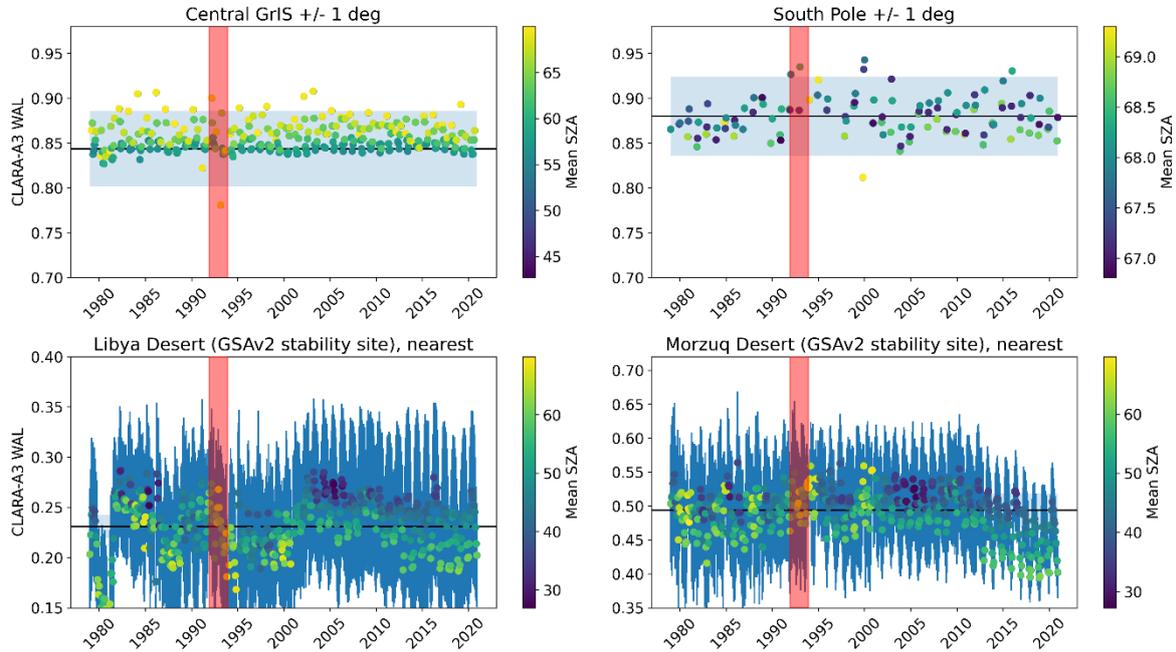
- 1) The central part of the Greenland Ice Sheet (GrIS), obtained as the mean of data from grid cells between 72-74 °N and 39-41 °W.
- 2) South Pole, obtained as the mean of data from grid cells between 88-90 °S and 0-2 °E.
- 3) Libya Desert, the nearest grid cell to 16.276 °N, 27.474 °E. This site is the same as used for stability analyses of the GSAv2 surface albedo data record of EUMETSAT.
- 4) Morzuq Desert, the nearest grid cell to 24.75 °N, 12.5 °E. This site is also used for stability analyses of the GSAv2 surface albedo data record of EUMETSAT.

The time series of CLARA-A3 SAL/WAL/BAL estimates over these areas are shown below in Figure 3 - Figure 5.

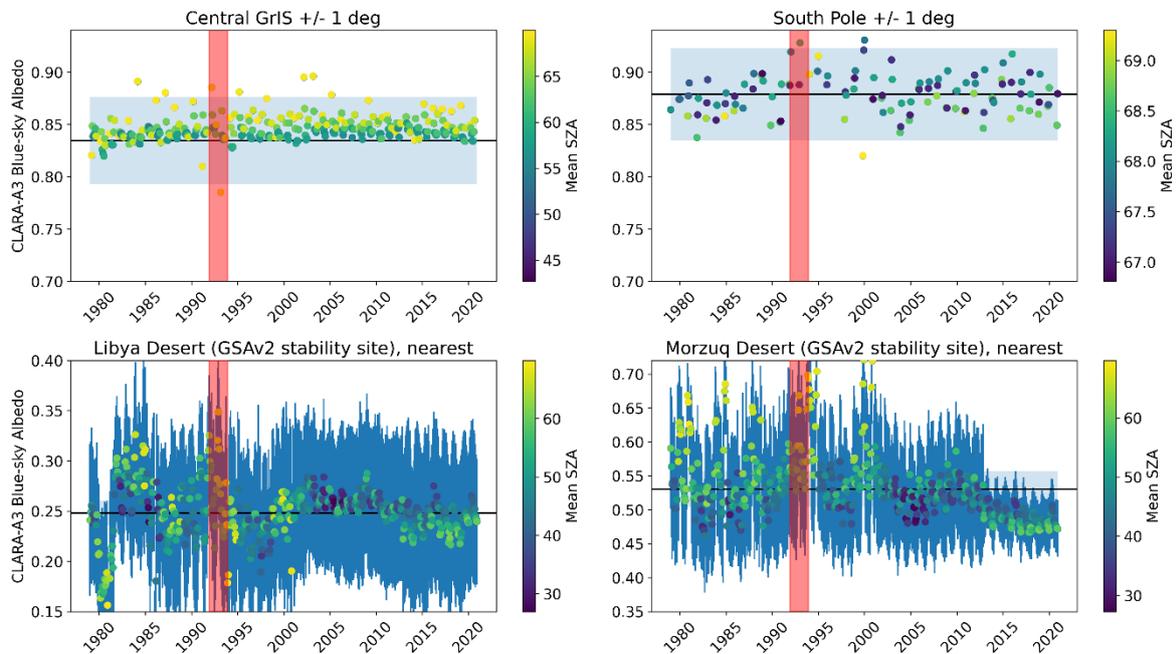


**Figure 3** : CLARA-A3 SAL (black-sky albedo) monthly mean retrievals over the designated analysis areas. Markers color coded by mean Sun Zenith Angle of valid observations during the month in question. Light blue envelope indicates  $\pm 5\%$  about the interannual mean (black horizontal line). Dark

blue envelope indicates median standard deviation of ‘instantaneous’ SAL retrievals. Red region indicates the time period most affected by the atmospheric disturbances of the Mt. Pinatubo eruption in June 1991 (2 years).



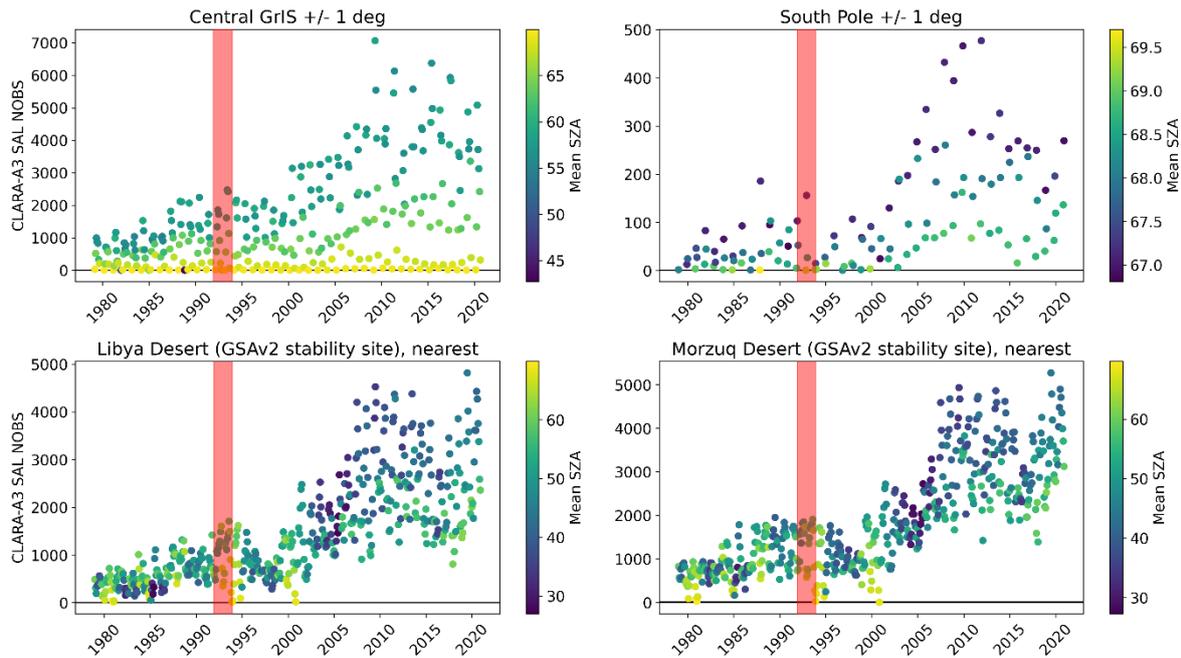
**Figure 4:** As Figure 3, but for CLARA-A3 WAL (white-sky albedo). Note that standard deviation over snow/ice surfaces is not available for WAL, as the variable is based on statistical relationship between mean black/white-sky albedos.



**Figure 5:** As Figure 3, but for CLARA-A3 BAL (Blue-sky albedo). Note that standard deviation over snow/ice surfaces is not available for BAL due to the non-availability of std for WAL over snow and ice.

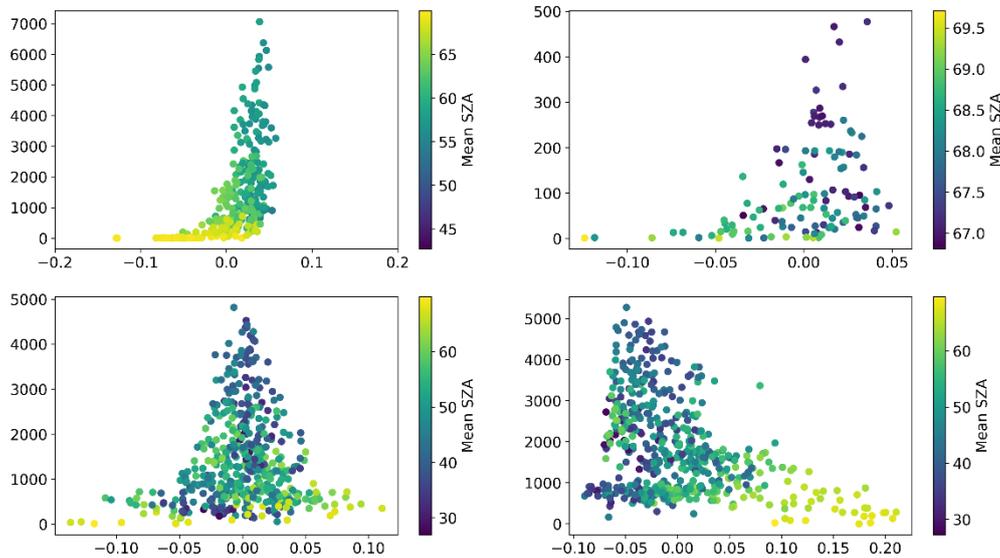
We first examine the desert sites which exhibit larger variability than the ice sheet areas. Some of the largest deviations are likely connected to atmospheric correction challenges related to

large-scale volcanic eruptions such as the Mt. Pinatubo in 1991 (red vertical bars in figures). However, it should be noted that the amount of available clear-sky AVHRR observations over these regions varies substantially across the CLARA coverage period as the AVHRR constellation changes. Figure 6 shows the number of clear-sky GAC-resolution observations in the monthly means over the stability sites. As is apparent, before the turn of the millennium the sampling rate was in general quite low as the AVHRR constellation consisted of only 1 or 2 satellites.



**Figure 6: The numbers of clear-sky GAC-resolution AVHRR observations available for each monthly mean over each of the stability sites. Color of marker denotes mean SZA of each month.**

A relevant follow-up question then is, is the variability in sampling related to the variability in retrieved albedo about the interannual mean? To explore this, Figure 7 shows the deviations from the interannual mean BAL over each of the stability sites. Here, the long “tails” of the BAL deviation distributions clearly show that observation density is a major factor in determining the magnitude of deviations, alongside variability in SZA.

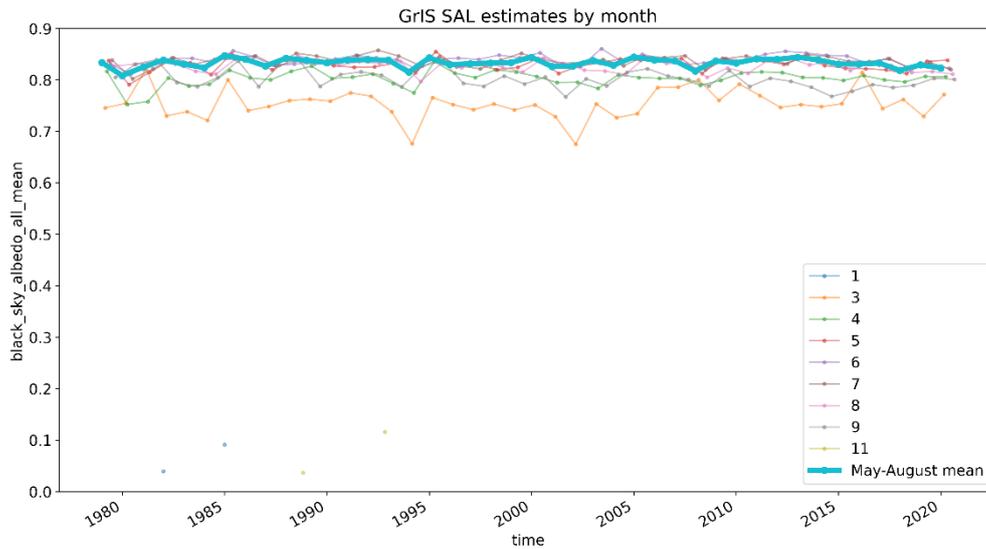


**Figure 7:** Deviation from interannual mean BAL over the stability sites as a function of available clear-sky AVHRR observations. Marker color indicates mean SZA of each month of observations.

Furthermore, it is important to note the apparent decrease in retrieved albedo at the Morzuq desert site from 2012 onwards. This cannot be connected simply to observation density, as the latest years in the CLARA record have the best coverage. Also, it is noteworthy that the phenomenon is not replicated over the other desert site in Libya, suggesting that the effect is localized rather than general algorithm issue. Given that the aerosol background in CLARA switches to a climatology after 2015, it is possible that a part of the seemingly fixed underestimation vs. interannual mean at Morzuq is related to the aerosol loading fixing to an anomalous level for the last 5 years of data. Further study is required to ascertain the causes of this behaviour, but since the anomaly is limited to one of four stability sites, we conclude that the overall stability of the dataset is not placed in question.

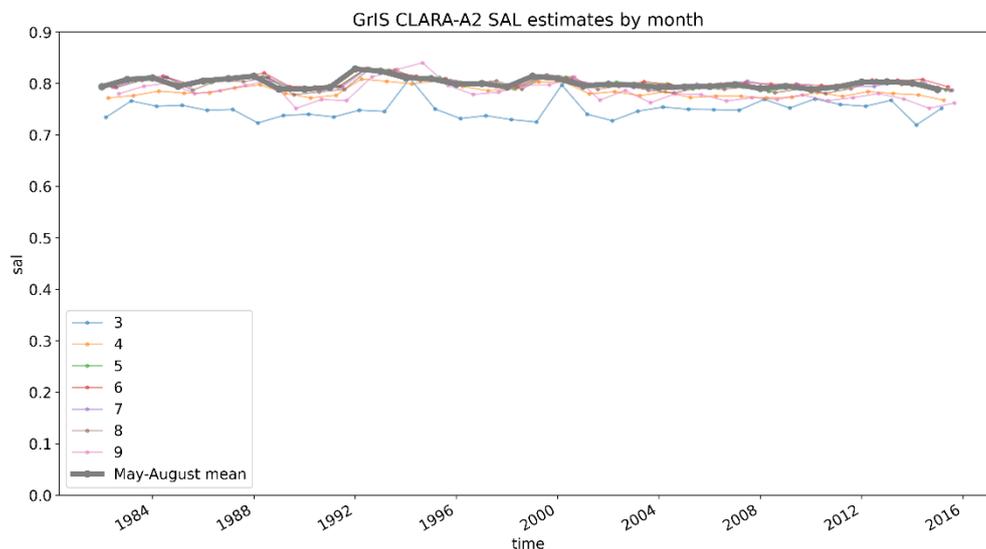
We then examine the ice sheet areas. Over these sites, where surface melt has been rare during the four decades of coverage, all retrieved albedo variables appear stable. In particular WAL, which by definition should not vary with solar elevation, and BAL exhibit variance which fits within the illustrated 5% envelope apart from some outliers. No trends are apparent over GrIS or South Pole.

In the black-sky albedo estimates (SAL), it is however apparent that some retrievals seem biased low relative to the mass of the estimates. To examine more closely, the SAL monthly mean data is displayed below month by month in Figure 8.



**Figure 8:** GrIS area's SAL retrievals by month. Thick blue line indicates the May-August mean.

From the data, it is apparent that the low-biased estimates generally occur in March. This behaviour is consistent with CLARA-A2 SAL, shown below for the same area in Figure 9. The likeliest causes for the apparent underestimation are challenges in obtaining an accurate atmospheric correction over the very bright snow surfaces during very low solar elevation conditions.



**Figure 9:** CLARA-A2 SAL monthly mean estimates for the GrIS area between 1982-2015

Another feature to note are the singular very low SAL estimates in CLARA-A3 over the area in the 1980s and 1990s. They occur during polar winter months and should therefore not happen at all. Their cause has been identified as incorrect solar zenith angle calculations sometimes occurring in PPS due to corrupted raw AVHRR orbital data; these corrupted SAL data affect only a few grid cells poleward of the true SZA cutoff line and are thus easily identifiable. The issue will be noted in the known issues & limitations-section of the PUM.

	<b>Validation Report Surface Albedo (SAL) CLARA Edition 3</b>	Doc. No: SAF/CM/FMI/VAL/CLARA/SAL Issue: 3.1 Date: 06.02.2023
---	---	---

The period of reliable observations, May-August, exhibit similar mean black-sky albedo over the area in CLARA-A2 ( $0.80 \pm 0.01$ ) and CLARA-A3 ( $0.83 \pm 0.01$ ), as expected. The slight increase in mean albedo is likely due to improved cloud detection, leading to fewer clouds misclassified as clear (typically causing an underestimation over bright snow) and secondarily to the revised AVHRR FCDR calibration. The increased mean albedo in CLARA-A3 over the central part of GrIS is now well in line with expected long-term mean albedo of the ice sheet's surface ( $\sim 0.84$ , see e.g. Konzelmann and Ohmura, 1995).

## 5. BAL validation against BSRN in situ observations

In this section, we first present a quantified analysis of representativeness effects on BSRN validation results in section 6. Sites with unacceptably poor representativeness of the surrounding area are flagged for exclusion from the calculation of the summary metrics (section 5.2), although their validation results are presented for completeness.

After the summary, we present the collated metrics and results of the BAL validation against BSRN reference data as well as result visualizations for each BSRN site. All visualizations are made in a common template to promote comparability. Then a brief discussion of results is presented alongside site-specific performance metrics (bias, precision, stability) over the site.

For each site-specific visualization, the top panel shows the retrieved and measured albedos (displayed in range 0-100%), and the bottom panel shows the relative retrieval errors (biases) of the level 3 pentad and monthly mean records, as well as the ‘instantaneous’ level 2 albedo estimates from each clear-sky overpass. The EUMETSAT Surface Albedo Validation Sites (SAVS) database (Loew et al., 2016) is frequently cited for information about the sites’ spatiotemporal representativeness.

### 5.1. Spatial representativeness of BSRN sites based on Google Earth Engine data

The representativity of an in situ albedo measurement against the surrounding areal mean observed by a satellite is a long-standing problem in validation of satellite retrievals. To analyze the issue’s impact on the CLARA-A3 surface albedo retrievals, we employed the Google Earth Engine, and in particular its Dynamic World dataset (Brown et al., 2022), to investigate the variability of high-resolution land cover in the areas surrounding the BSRN sites participating in this study.

Specifically, we extracted the median Dynamic World land cover data at 30 m spatial resolution for June-August 2018 with the 0.25 degree BAL grid cells containing BSRN sites. Using climatological surface albedos valid for midsummer of the Dynamic World land cover classes, we then constructed ‘expected’ climatological BAL values for these grid cells. Comparing these to the actual retrieved BAL during this period (June-August means) and the measured in-situ blue-sky albedo, we can assess the impact of spatial representativeness on the retrieval bias at each BSRN site.

**Table 4:** Climatological surface albedos prescribed to Dynamic World land use classes for the representativeness analysis

Dynamic World land use class	Prescribed surface albedo	Source & comments
Forest	0.14	He et al. (2014)
Grassland	0.2	He et al. (2014)
Flooded vegetation	0.16	He et al. (2014)
Cropland	0.19	He et al. (2014)
Shrubs	0.2	He et al. (2014)
Built area	0.15	Trlica et al. (2017)
Bare ground	0.25	He et al. (2014)

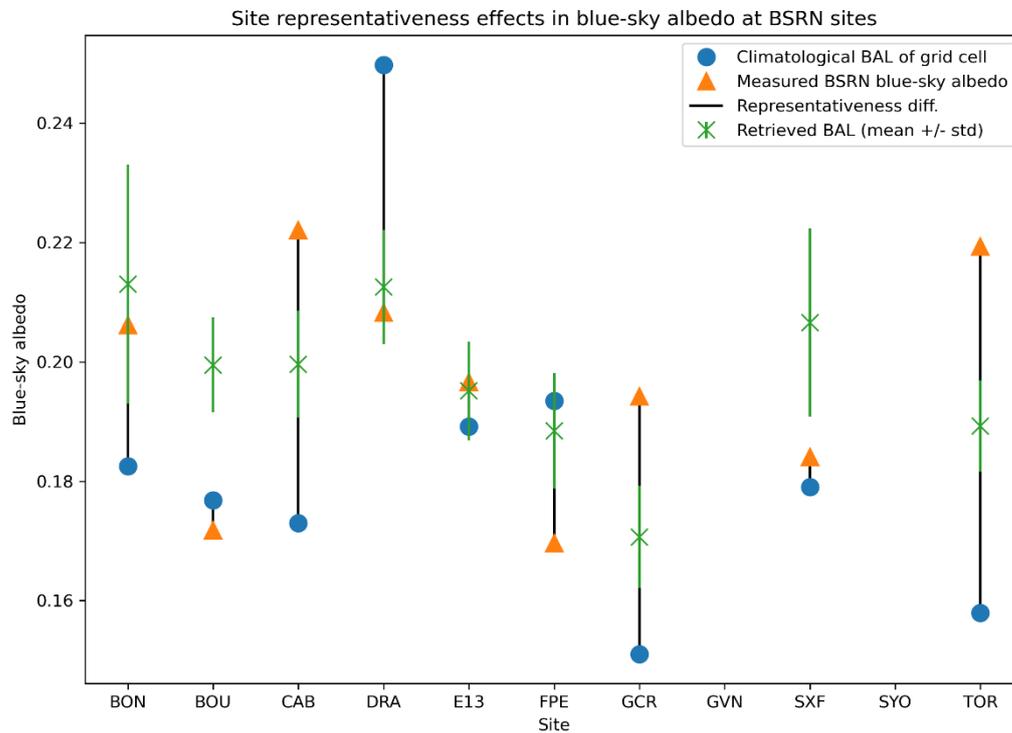
	<p style="text-align: center;"><b>Validation Report</b>  <b>Surface Albedo (SAL)</b>  <b>CLARA Edition 3</b></p>	<p>Doc. No: SAF/CM/FMI/VAL/CLARA/SAL  Issue: 3.1  Date: 06.02.2023</p>
---	--	--

The results of the analysis are shown in Figure 10. The black lines connecting the expected climatological BAL of the grid cell (blue) with the in situ measured albedo at some minutely small patch within (orange) indicate the spatial dissimilarity of each grid cell in question. We generally expect the actually retrieved BAL (green) to fall somewhere on the black line, as perfect agreement is not expected with neither the climatology nor the in situ measurement. The prescribed climatological albedo ignores potentially significant albedo variations within each class, and the climatological albedo further assumes the Dynamic World land use to be fully correct. Also, the phenological variability in surface albedo is not captured, although the analysis here is only focused on midsummer period; thus also seasonal snow cover is not an issue. Some sites may also measure larger footprints (Table 1).

As an example of the uncertainty in the analysis, we note that the expected BAL at Desert Rock (DRA) is markedly larger than either the retrieved BAL or the in situ measurement. The cause is that Dynamic World classifies nearly the entire grid cell as 'bare ground', whose prescribed albedo is set higher than the other classes. However, according to the SAVS database ([https://savs.eumetsat.int/html/DRA\\_report.html](https://savs.eumetsat.int/html/DRA_report.html)), the region is classified dominantly as shrubland. This land cover class change would bring the climatological BAL at DRA down to ~0.2, in close agreement with the in situ measurement and the multi-year BAL retrievals.

We note good agreement with representativeness expectations at Cabauw (CAB), Southern Great Plains (E13), Fort Peck (FPE), Goodwin Creek (GCR), and Toravere (TOR). However, at Bondville (BON), Boulder (BOU), and Sioux Falls (SXF), the retrieved BAL overestimates compared to the expected range. No extractable valid data was available for GVN and SYO during the evaluation. We noted some differences in land cover around BOU and SXF from Dynamic World when compared to the ESA Land Cover CCI-based estimates in the SAVS database, particularly regarding the cropland/grassland delineation. However, as the climatological albedos of these classes are very similar, this cannot alone explain the larger than expected BAL retrievals. At this point, no clear reason for the different behaviour at these two sites has emerged; followup studies using e.g. Google Earth Engine to map surface reflectance patterns directly from Landsat or Sentinel data may shed more light on the issue. The Dynamic World classification maps and inferred climatological albedo maps are available in Appendix A of this report.

The question of which of the sites are not representative enough of their surrounding area is somewhat subjective. TOR, GCR, DRA and CAB show the largest differences between expected areal mean albedo and the actual measurement. However, as discussed, at DRA this difference is likely artificial due to land cover classification ambiguity. A similar situation likely exists at CAB due to likely larger than reported grass coverage combined with uncertainty about the true albedo variance between cropland and grassland, the dominant land covers around CAB. Thus, GCR and TOR are likely the most unrepresentative sites due to considerable land cover heterogeneity and abundance of forest cover around the measurement site (BSRN measurements usually taking place over a grass patch). We thus exclude GCR and TOR from the summary metrics calculation. Furthermore, because ALE and GVN feature highly variable snow and sea ice cover in their grid cells, we also exclude these sites from the summary metrics, although the analysis is retained and presented below.



**Figure 10:** Site representativeness analysis results from BSRN sites at monthly mean scale. Blue circle indicates expected BAL during June-August, based on land cover from Google Dynamic World and associated climatological surface albedos. Orange triangles indicate mean observed in situ albedo at each site. The length of the black line connecting the two indicates the dissimilarity between the two as an indicator of spatial un-representativeness. Finally, green cross with whiskers indicates the mean BAL from CLARA-A3 of the grid cell containing site, the whiskers showing the standard deviation of the June-August mean BAL retrievals during the evaluation period.

## 5.2. Validation metrics summary, performance against specifications

Table 5 and Table 6 display the time-averaged performance indicators and their metrics of BAL against all BSRN in situ reference data considered representative. The obtained mean metrics are then compared against the performance targets in Table 3 to ascertain the requirement level fulfilled (optimum, target, threshold, or none). BAL performance clearly fills the target requirement for all evaluated metrics. For the ‘instantaneous’ level 2 BAL data and the pentad means, the mean bias fills the optimum requirement level. For comparison purposes, performance levels are reported for BAL both as the mean of site-specific performance and against all valid BSRN data – this choice has relevance for some metrics.

As discussed earlier, the metrics were also evaluated for BAL MM for the case where all valid in situ albedo observations taken when  $SZA < 70$  degrees form the reference (vs observations coinciding with clear-sky AVHRR overpasses). The corresponding bias was -4.12% and precision was 0.059 (mean of precisions of sites; precision over all data was 0.07), matching the overpass-centered metrics and confirming that the validation results are valid irrespective of applied temporal filtering.

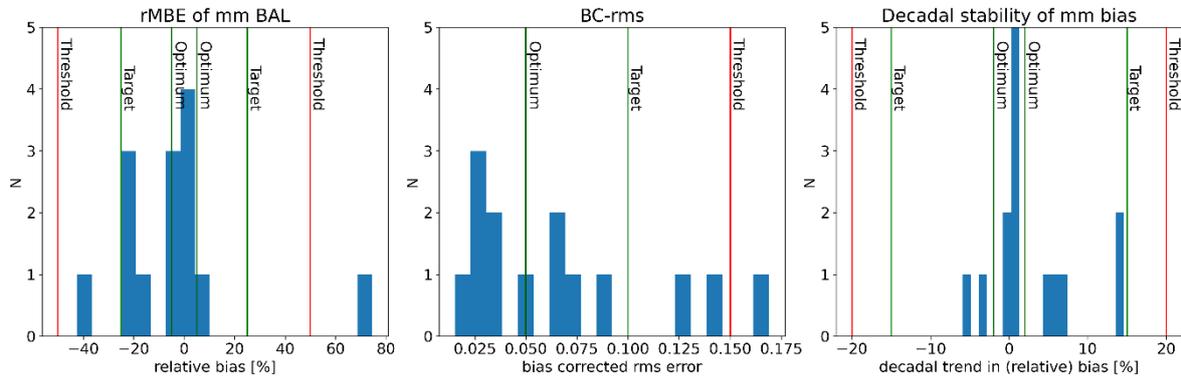
**Table 5:** Overpass (level 2) & Monthly mean BAL performance metrics against BSRN observations.

Variable	Time resolution	Indicator (Metric)	Value	Requirement level fulfilled
BAL	Overpass	Bias (GAC level2 rMBE)	0.03% (mean of sites) 0.69% (all data)	Optimum
BAL	MM	Bias (level 3 rMBE)	-4.35% (mean of sites) -4.12% (all data)	Optimum / Target
BAL	MM	Precision (bc-rms)	0.057 (mean of sites) 0.07 (all data)	Target
BAL	MM	Stability (bias dec. trend)	2.30% (mean of sites) -0.31% (all data)	Target / Optimum

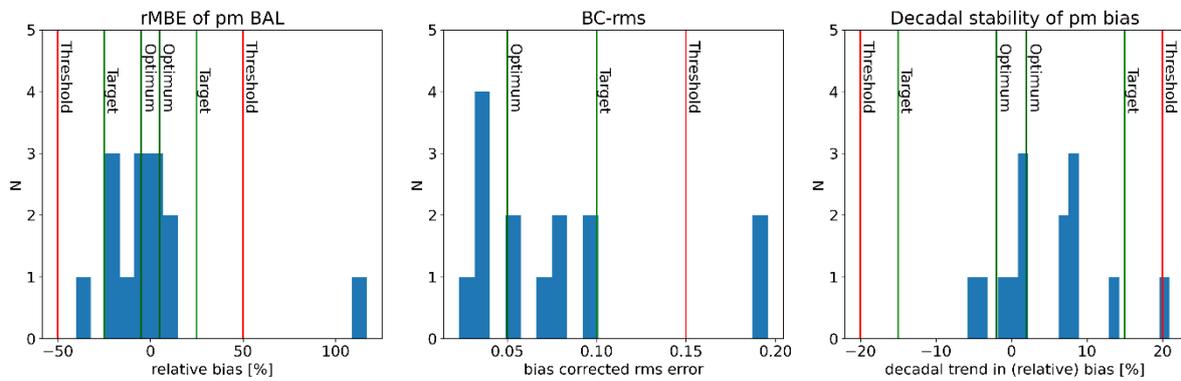
**Table 6:** Pentad mean BAL performance metrics against BSRN observations

Variable	Time resolution	Indicator (Metric)	Value	Requirement level fulfilled
BAL	PM	Bias (level 3 rMBE)	-2.63% (mean of sites) -2.12% (all data)	Optimum
BAL	PM	Precision (bc-rms)	0.063 (mean of sites) 0.074 (all data)	Target
BAL	PM	Stability (bias dec. trend)	3.38% (mean of sites) 0.26% (all data)	Target / Optimum

The performance metric distribution across the participating BSRN sites is illustrated in Figures Figure 11 & Figure 12. Nearly all the individual sites fulfill the target requirement levels of the three metrics.



**Figure 11:** Distribution of performance metrics of monthly mean BAL over BSRN sites, inclusive of both representative and unrepresentative sites. Left: relative bias, center: precision, right: stability. Vertical colored lines indicate thresholds of performance levels.



**Figure 12:** As Figure 11, but for pentad mean BAL.

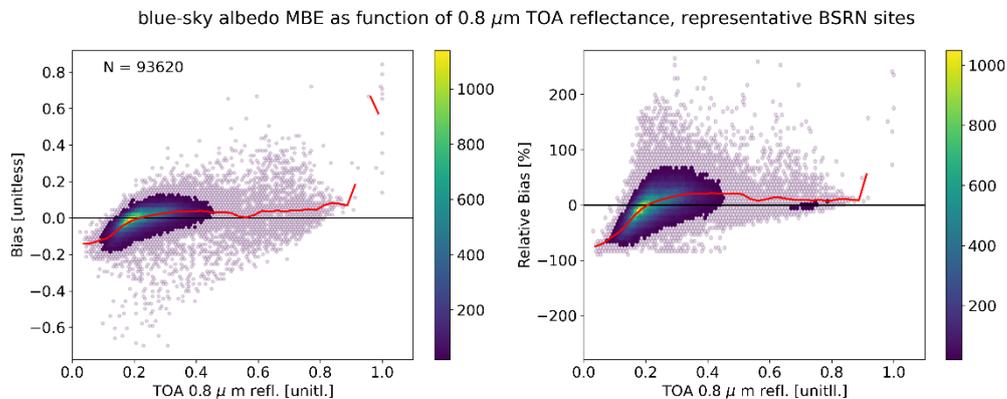
For a more in-depth look into the level 2 BAL performance, we first calculated the Pearson linear correlations of relative blue-sky albedo bias vs. BSRN reference against the input parameters of the retrieval, such as AOD or the illumination-viewing geometry, in order to identify the parameters most relevant for retrieval bias. Table 7 shows the top 5 of these parameters. Note that the retrieved black- and white-sky albedos are excluded, as they would naturally correlate very well with BAL bias.

**Table 7:** Pearson correlations of BAL retrieval parameters against BAL relative bias (vs BSRN) at the overpass level.

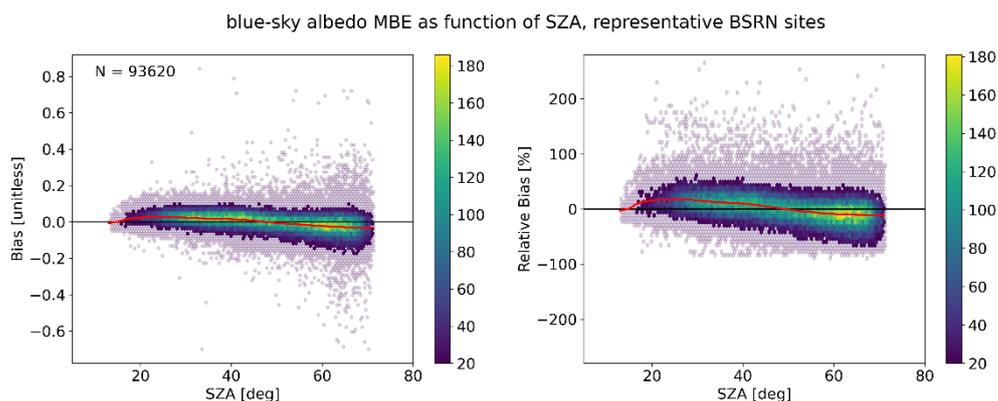
Retrieval parameter	Pearson correlation vs. BAL relative bias
0.8 micron TOA reflectance	0.42
SZA	-0.35
Cloud probability (in GAC pixel)	0.26
0.6 micron BOA reflectance	0.23
Surface pressure	-0.20

The binned distributions of retrieved albedo bias at representative vegetated and seasonal snow sites for AVHRR overpass – station observation (15 minute window) pairs (N = 93620

pairs) are shown in Figures Figure 13 and Figure 14 as functions of 0.8 micron TOA reflectance and SZA, respectively.



**Figure 13:** Left, BAL level 2 bias vs. BSRN observations as a function of 0.8 micron TOA reflectance of the AVHRR observation. Uncolored markers imply less than 20 observations in said bin. Right, relative BAL bias as a function of 0.8 micron TOA reflectance.



**Figure 14:** As Figure 13, but against Sun Zenith Angle (SZA) of the observation time.

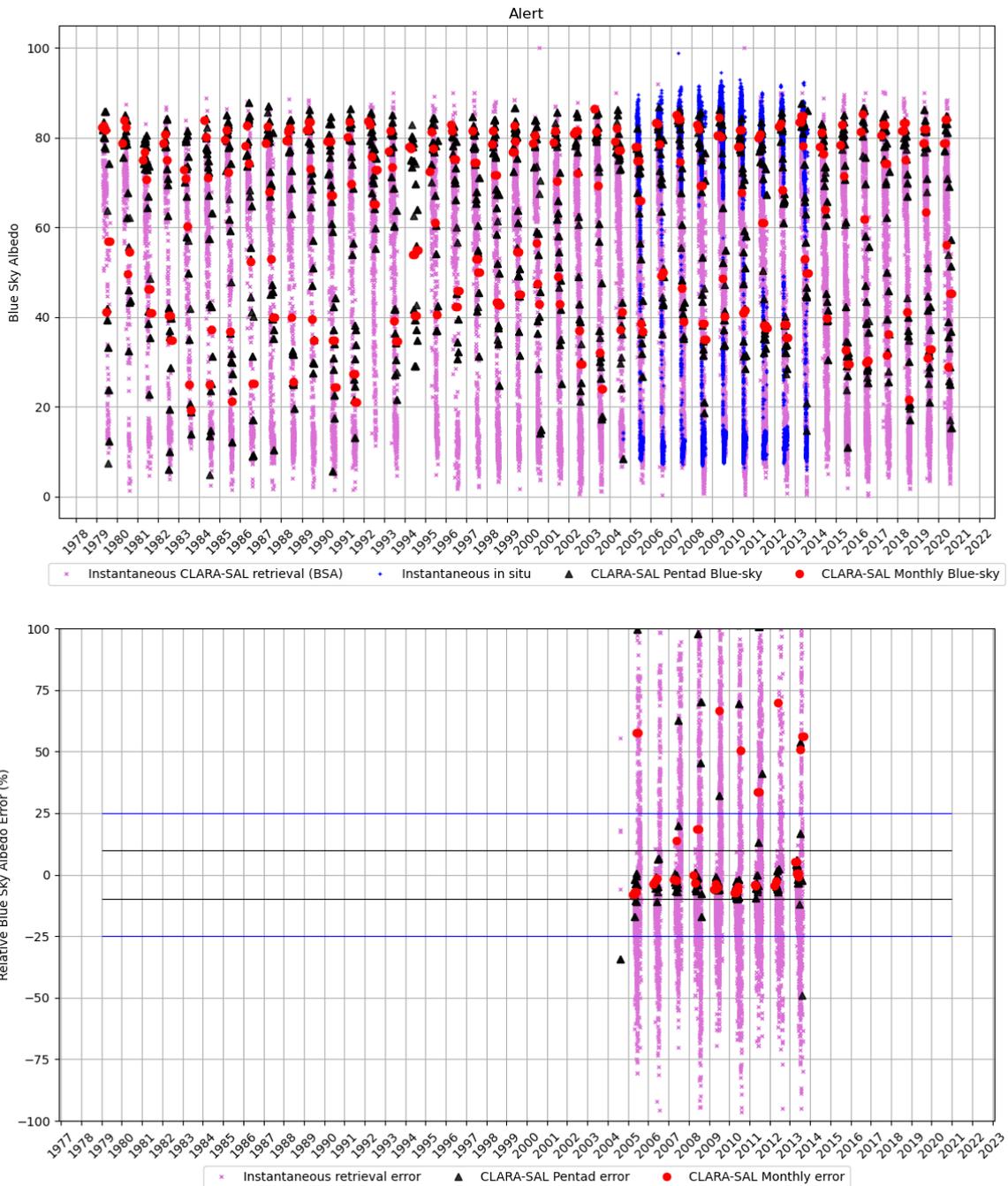
The distribution of bias as function of NIR band TOA reflectance shows no significant trend; the uptick at very high TOA reflectance may rather reflect the small number of samples in the BSRN dataset, as few of the representative data are over naturally brightest snow surfaces. A more thorough evaluation of retrieval performance on glaciated snow and ice against PROMICE station observations on the Greenland Ice Sheet is available in section 7.

The evolution of bias as a function of SZA suggests that BAL somewhat overestimates albedo with high solar elevation (low SZA) and underestimates with low solar elevation (high SZA). Notably, though, the range of potential retrieval biases increases strongly once we pass SZA of 60 degrees. This is consistent with our a priori information that the atmospheric correction becomes less robust at high SZA due to increasing path lengths through the atmosphere.

Note that the large negative biases exist solely over snow and ice surfaces, whose in situ albedo is typically large ( $>0.7$ ). For these cases, even large biases of  $-0.3$  to  $-0.4$  result in relative biases of  $-50\%$  to  $-70\%$ , which explains why the expanding “tail” of the biases is not present in the relative bias distribution (Figure 14, right). Note that the more consistently snow/ice covered sites (ALE, SYO, SPO) are not considered here.

### 5.3. Validation result illustrations of the participating BSRN sites

#### 5.4. Alert (ALE)

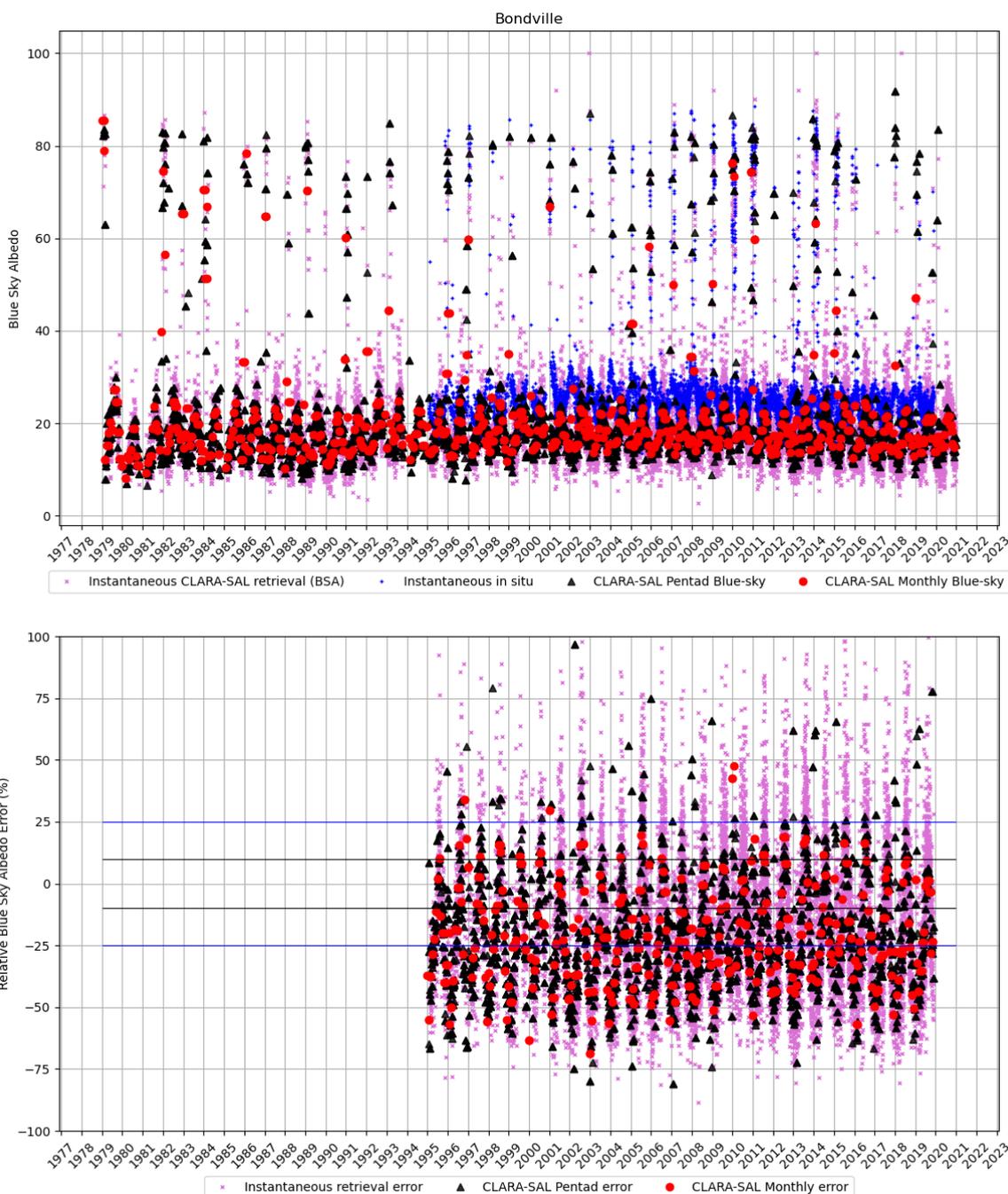


**Figure 15:** (Top) Retrieved CLARA-A3 monthly mean BAL (red circles), pentad mean BAL (black triangles), GAC-resolution (overpass) BAL (magenta x), in situ measured albedo (blue +). (Bottom) Relative retrieval error (bias) of the pentad, monthly mean and instantaneous satellite estimates, markers consistent with the top panel.

The Alert site in the Canadian High Arctic is characterized by a very strong seasonal cycle in measured in-situ albedo resulting from the melt of local snow cover. However, at the ~25 km

scale of CLARA grid cells, the coastal location of the site means that half of the grid cell is water-covered, with seasonally variable sea ice cover. The implication is that the spatial representativeness of the site is poor, as also noted in the analysis of the EUMETSAT SAVS validation site database: [https://savs.eumetsat.int/html/ALE\\_report.html](https://savs.eumetsat.int/html/ALE_report.html)

### 5.5. Bondville (BON)

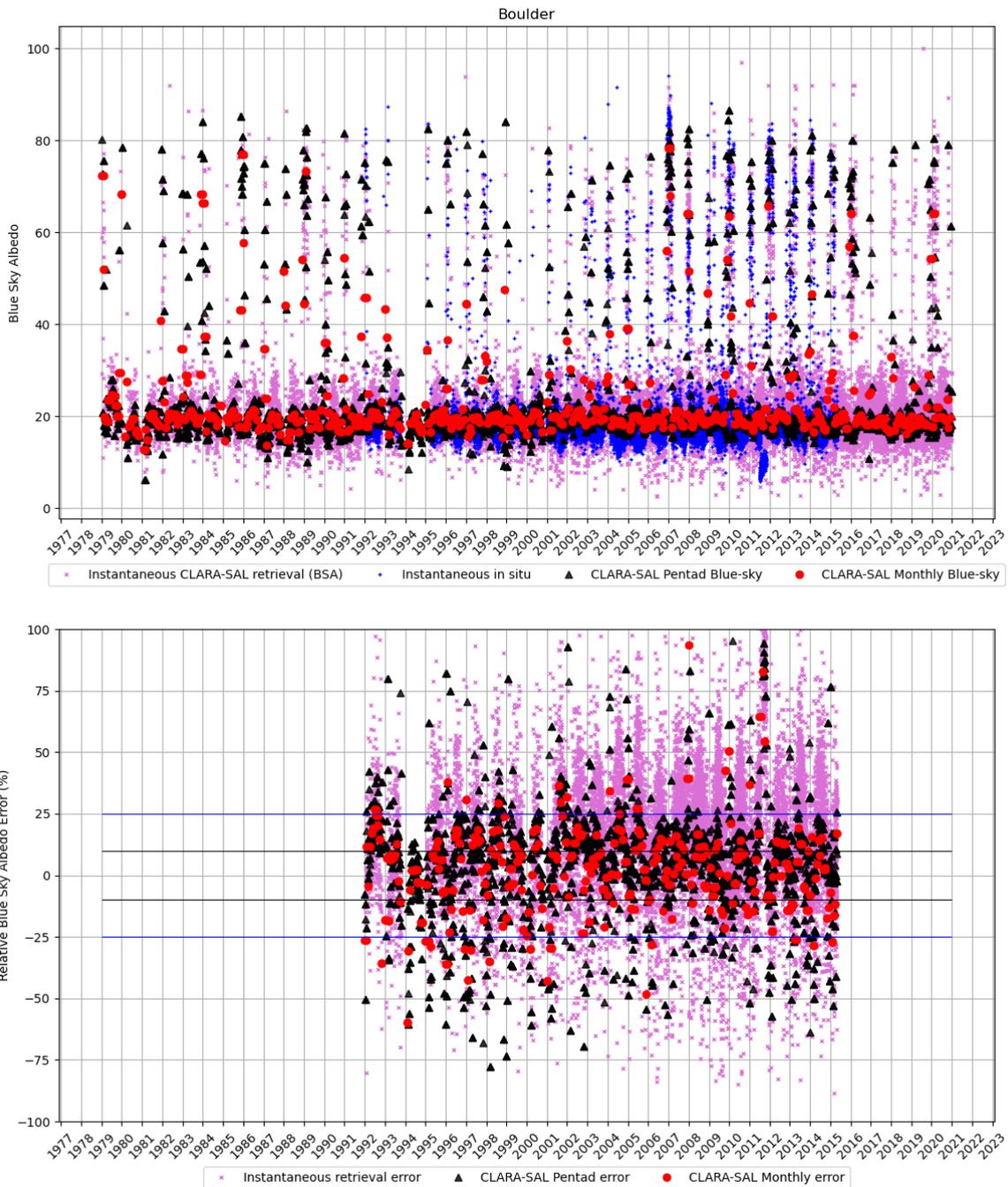


**Figure 16:** Validation results for Bondville (BON).

Bondville is located in the middle of relatively homogeneous croplands ([https://savs.eumetsat.int/html/BON\\_report.html](https://savs.eumetsat.int/html/BON_report.html)), implying relatively good spatial

representativeness of the in situ measurements. Only the nearby town of Champaign, IL, disturbs the cropland cover and may thus contribute to the increase in underestimation for pentad/monthly mean BAL vs. the overpass-level estimates. However, long-term mean bias of all CLARA estimates still remains within the target specification (-11% to -20%).

### 5.6. Boulder (BOU)

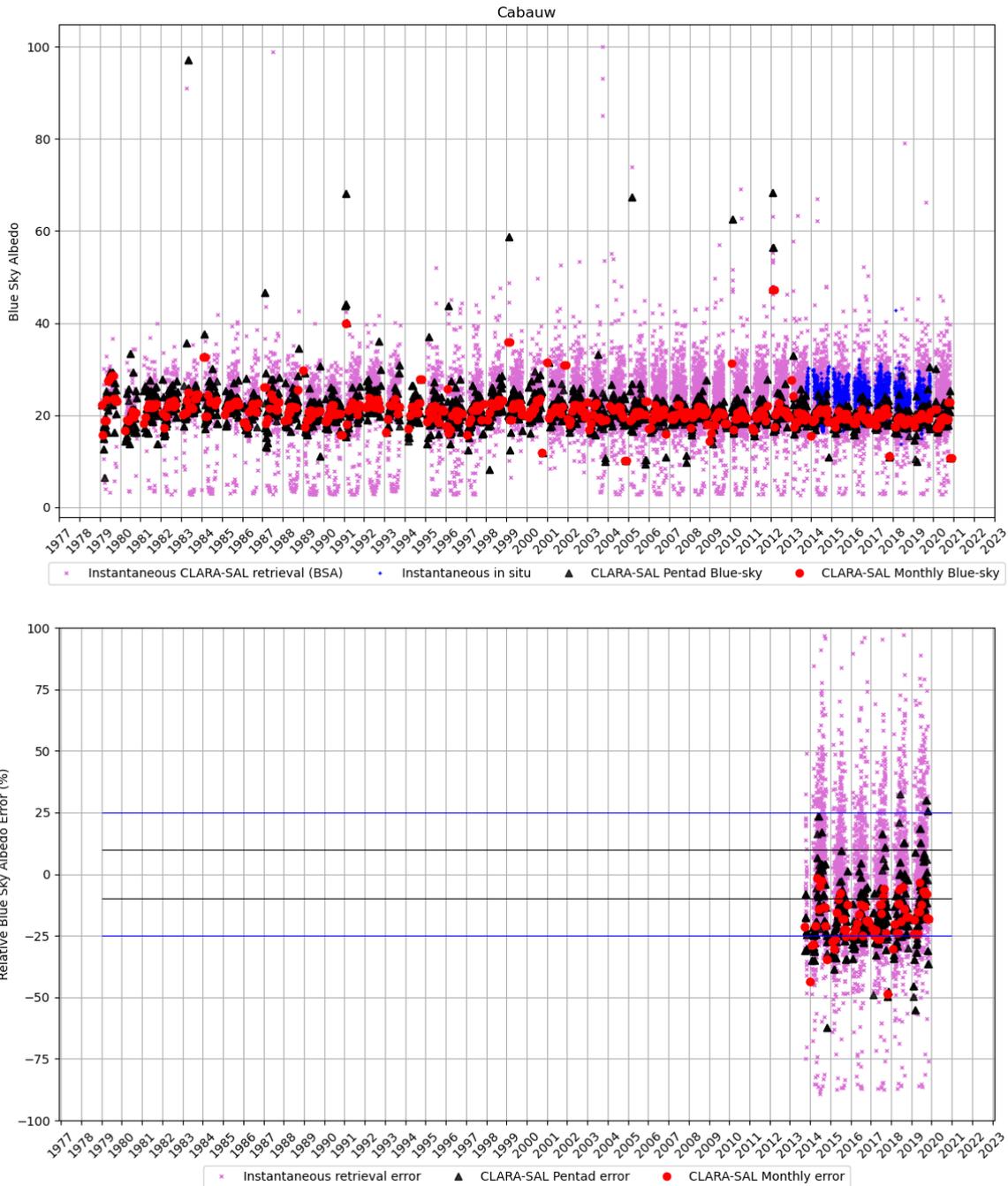


**Figure 17:** Validation results for Boulder (BOU)

The site at Boulder is predominantly mid-latitude grasslands, implying a good spatial representativeness and a modest seasonal cycle in albedo apart from snowfall occurrences.

The metropolitan area of Denver to the south is nearby, but the center of the CLARA grid cell containing the site is to the NW, meaning that only the urban area of the small town of Longmont may contaminate the grid cell. However, as the retrieved CLARA albedos are consistent and stable, it appears that contamination is a minor issue.

### 5.7. Cabauw (CAB)



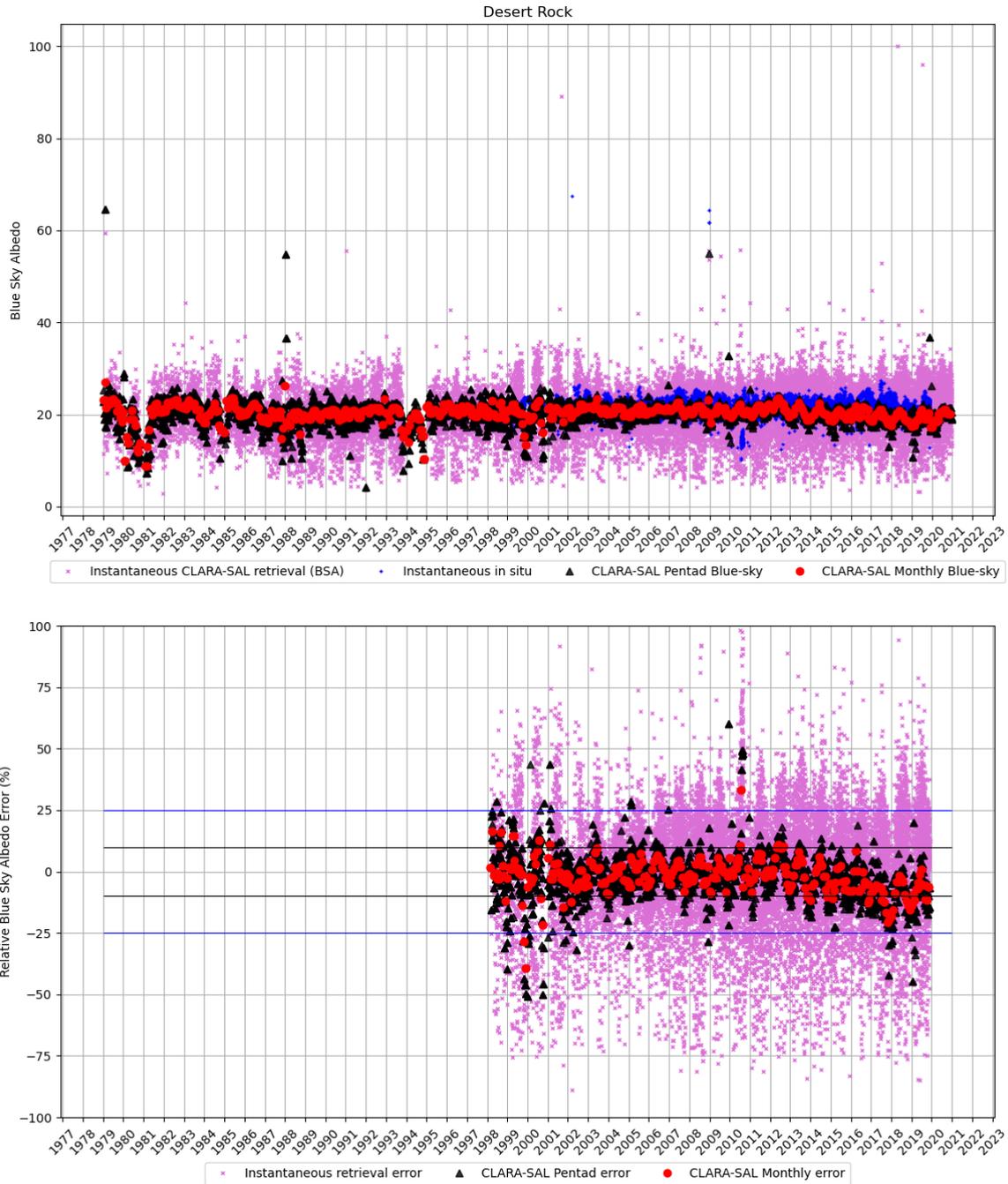
**Figure 18:** Validation results for Cabauw (CAB)

The Cabauw site is surrounded mainly by grassland, with a narrow river and urban areas rounding out the land cover according to the SAVS database. Occasional snowfall events

	<p style="text-align: center;"><b>Validation Report</b>  <b>Surface Albedo (SAL)</b>  <b>CLARA Edition 3</b></p>	<p>Doc. No: SAF/CM/FMI/VAL/CLARA/SAL  Issue: 3.1  Date: 06.02.2023</p>
---	--	--

during autumn-spring are possible but relatively rare. The level 3 BAL underestimates the in situ measurements somewhat, but the overpass-level retrievals are clearly in fairly good agreement.

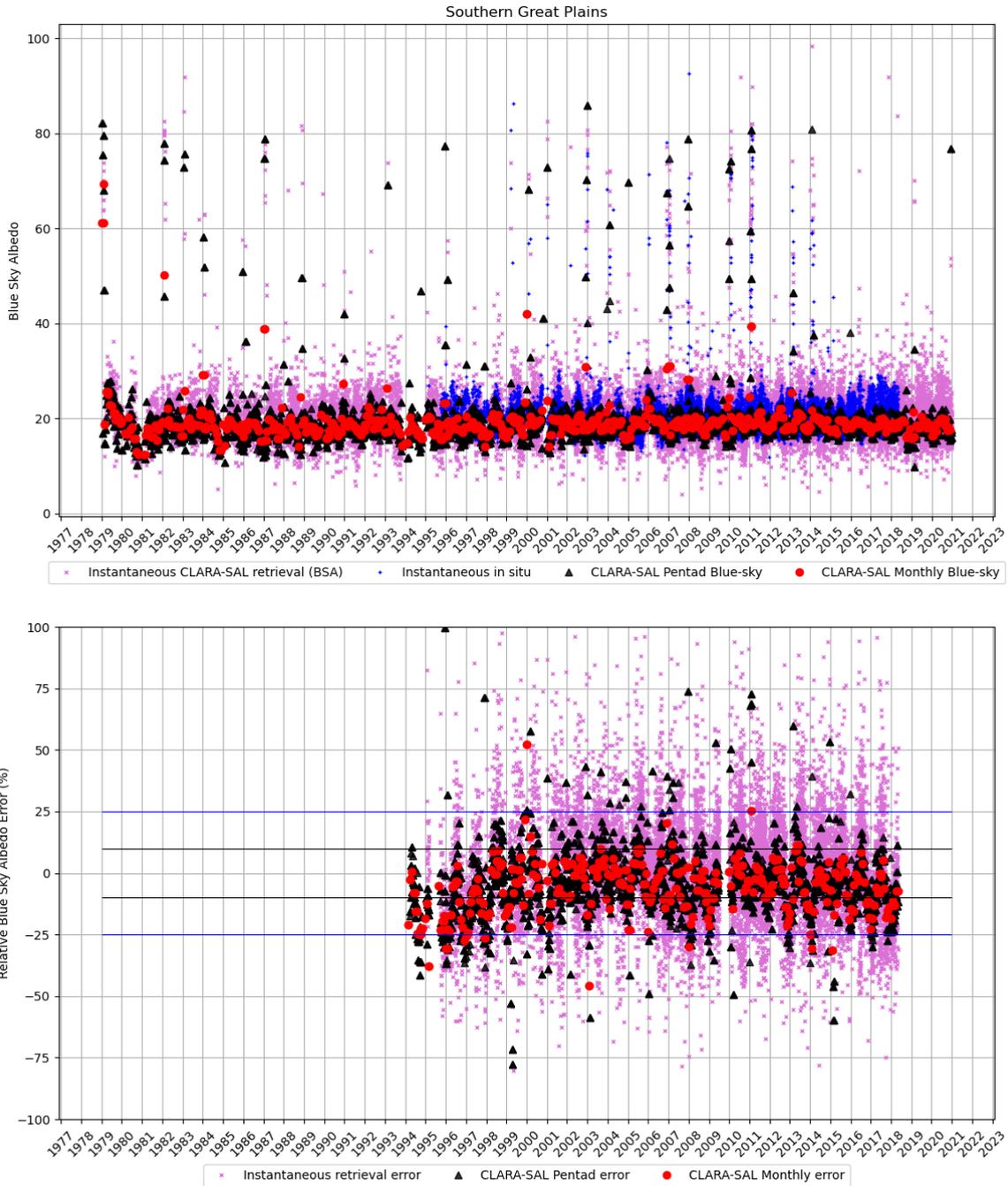
## 5.8. Desert Rock (DRA)



**Figure 19:** Validation results for Desert Rock (DRA)

Desert Rock, located in the dry and barren region of the American Southwest, is a spatially homogeneous and stable site ([https://savs.eumetsat.int/html/DRA\\_report.html](https://savs.eumetsat.int/html/DRA_report.html)). Consistently, BAL retrievals are stable over the site, although the overpass-level retrievals do exhibit variability related to illumination geometry changes (the retrievals are not normalized to a specific Sun Zenith Angle). Mean bias over the evaluation period is very small due to mutually canceling (low) under- and overestimations. Retrieval stability over the full 1979-2020 CLARA coverage is also very good.

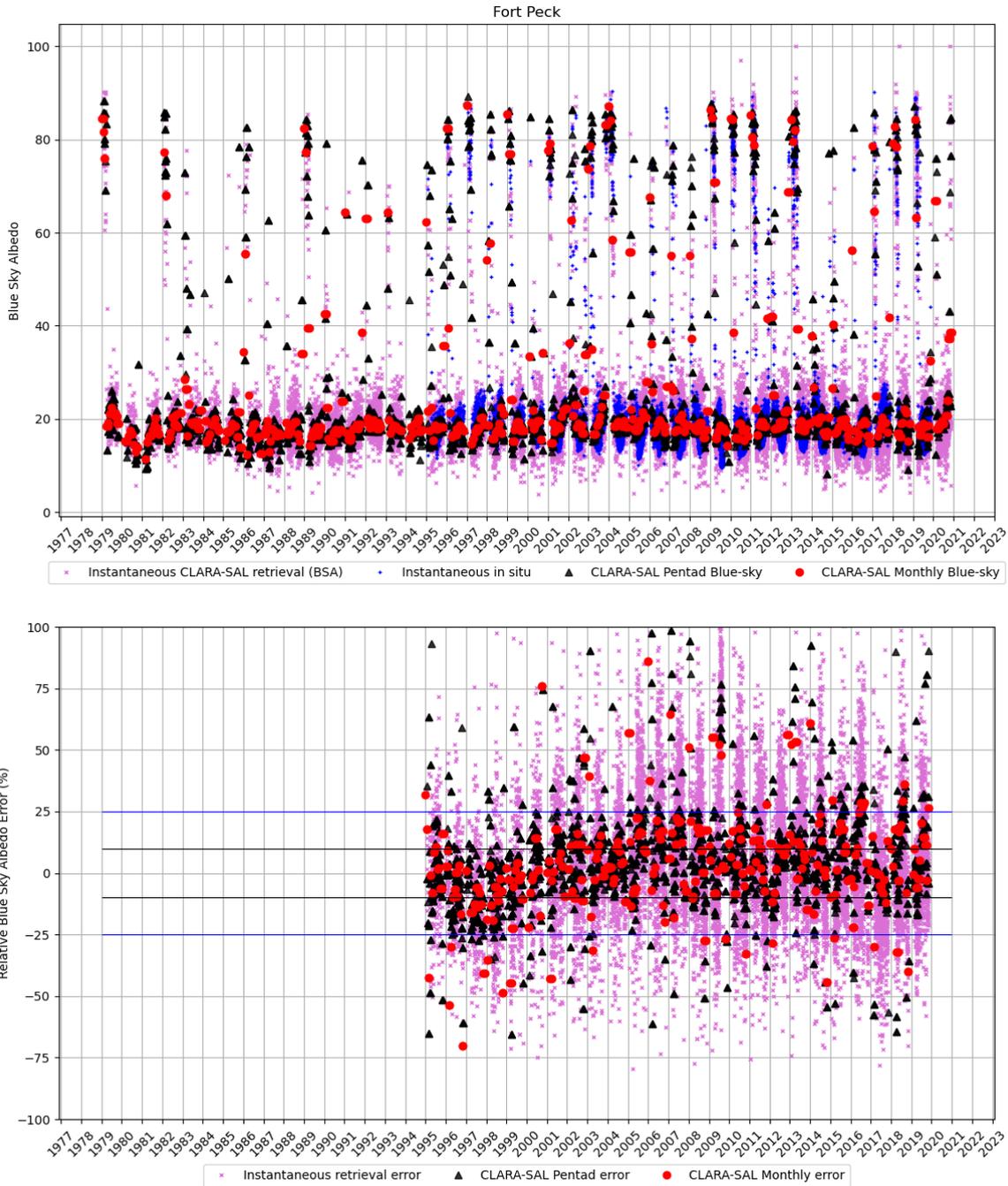
### 5.9. Southern Great Plains (E13)



**Figure 20:** Validation results for Southern Great Plains (E13)

Like Desert Rock, the Southern Great Plains BSRN site is often cited as a good example of a spatially representative validation site even at coarse 10+ km resolutions ([https://savs.eumetsat.int/html/E13\\_report.html](https://savs.eumetsat.int/html/E13_report.html)). The region is composed of grassland and pastures, implying relatively low seasonal albedo cycles, again apart from snowfall events. Here, BAL retrievals match the in situ observations well, and display very good temporal stability, with possible exceptions in the 1979—1981 timeframe of the earliest AVHRR observations with least amounts of available sampling.

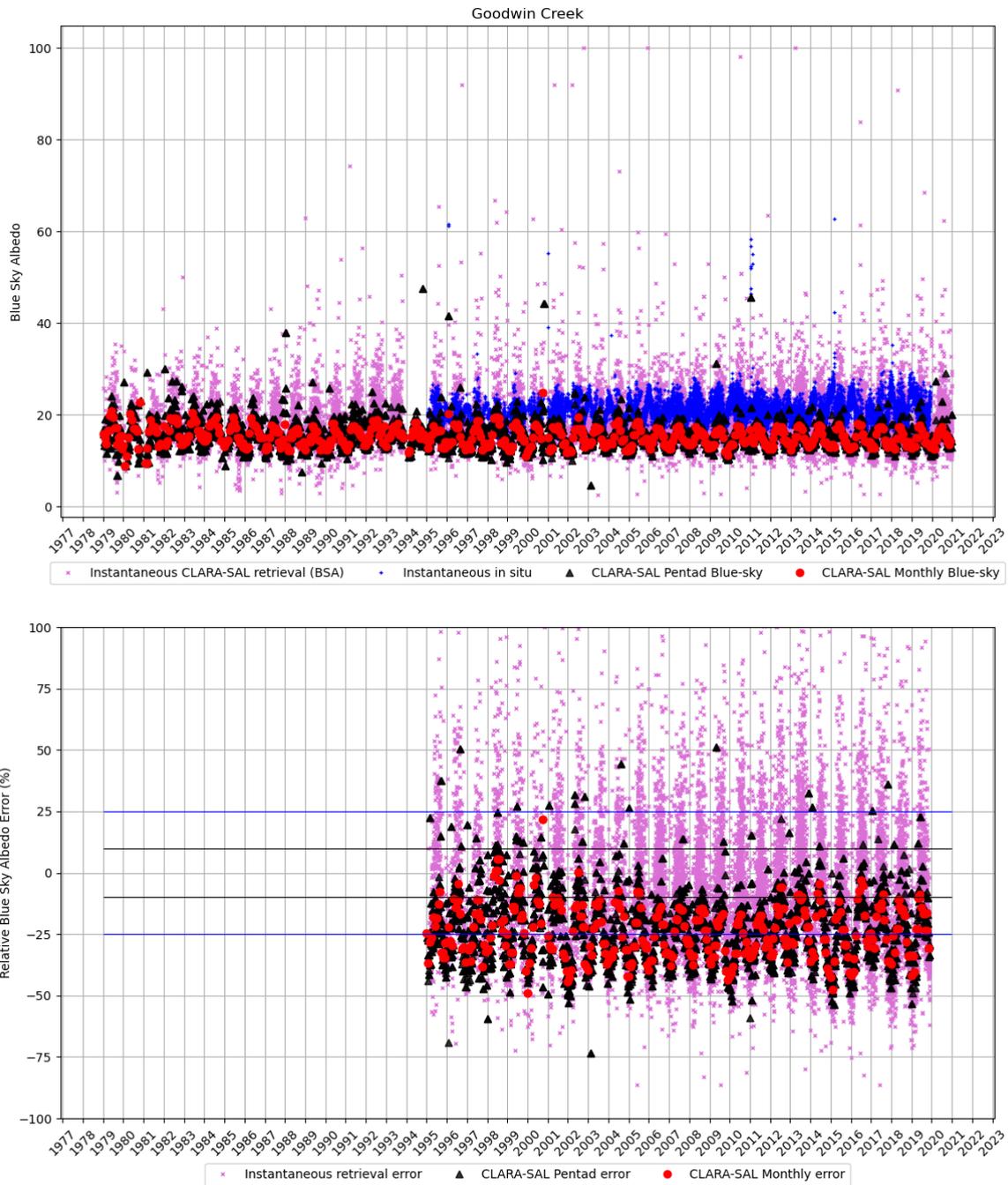
### 5.10. Fort Peck (FPE)



**Figure 21:** Validation results for Fort Peck (FPE)

Fort Peck, located in Montana, is predominantly grassland, although with substantial tree and shrub cover at coarser scales. Snow cover is common during winter and vegetation senescence alters land surface albedo markedly, combining into a considerably large seasonal cycle in total surface albedo. BAL generally tracks this variability well, although discrepancies in snow cover in the point-to-pixel comparison can produce large biases in single pentad or monthly mean products.

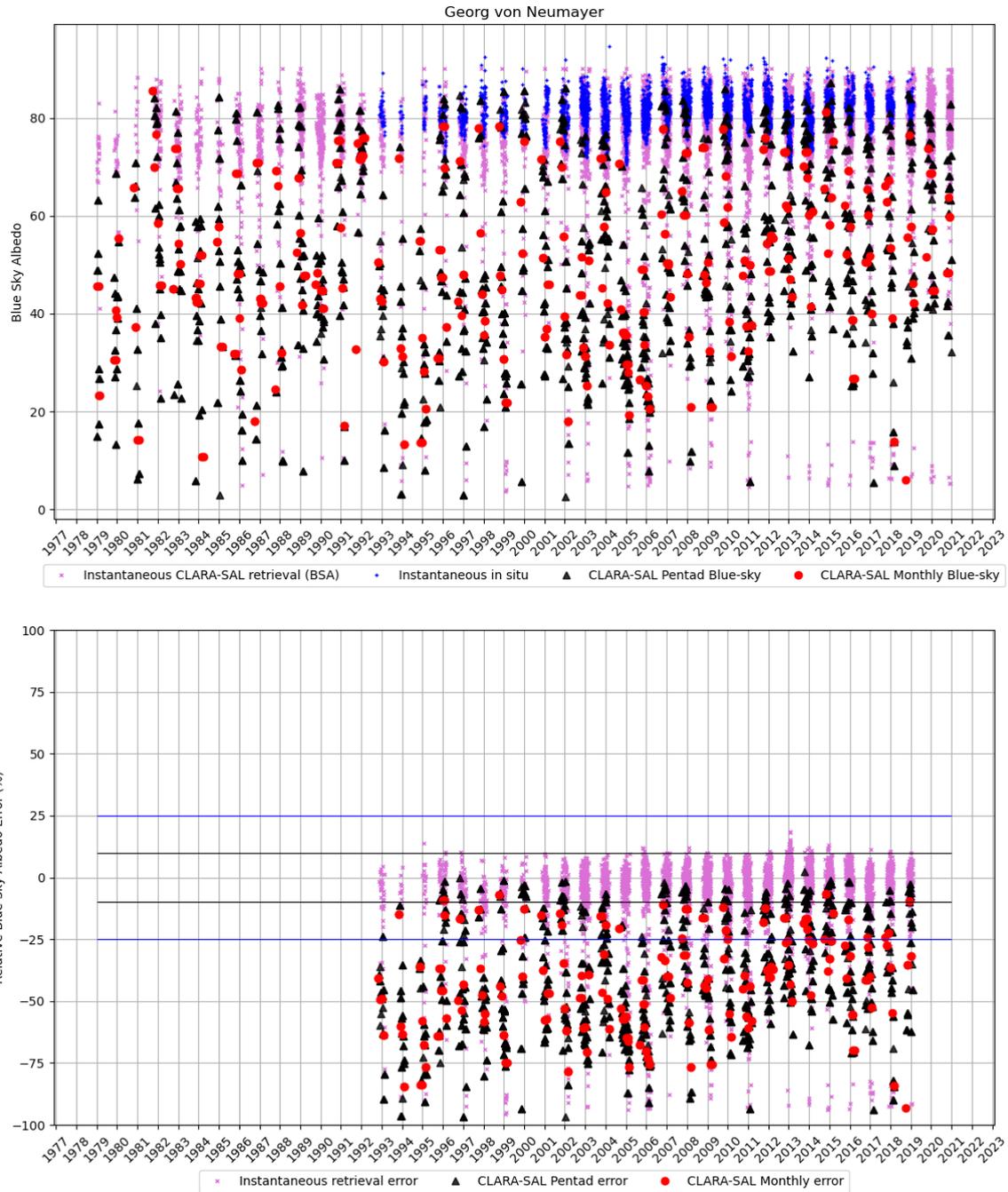
### 5.11. Goodwin Creek (GCR)



**Figure 22:** Validation results for Goodwin Creek (GCR)

Goodwin Creek in Mississippi is located among a mix of forest, cropland, and grassland land cover areas at the 20 km scale ([https://savs.eumetsat.int/html/GCR\\_report.html](https://savs.eumetsat.int/html/GCR_report.html)). Here, BAL pentad and monthly mean estimates generally underestimate the in situ measured albedo, but the GAC-resolution (5-10 km) retrievals provide a better match, implying that the representativeness issues at coarse resolution produce the majority of the bias instead of retrieval algorithm-based issues. Again, the BAL retrievals (and their bias) display negligible trends over the CLARA (validation) coverage period.

## 5.12. Georg von Neumayer (GVN)

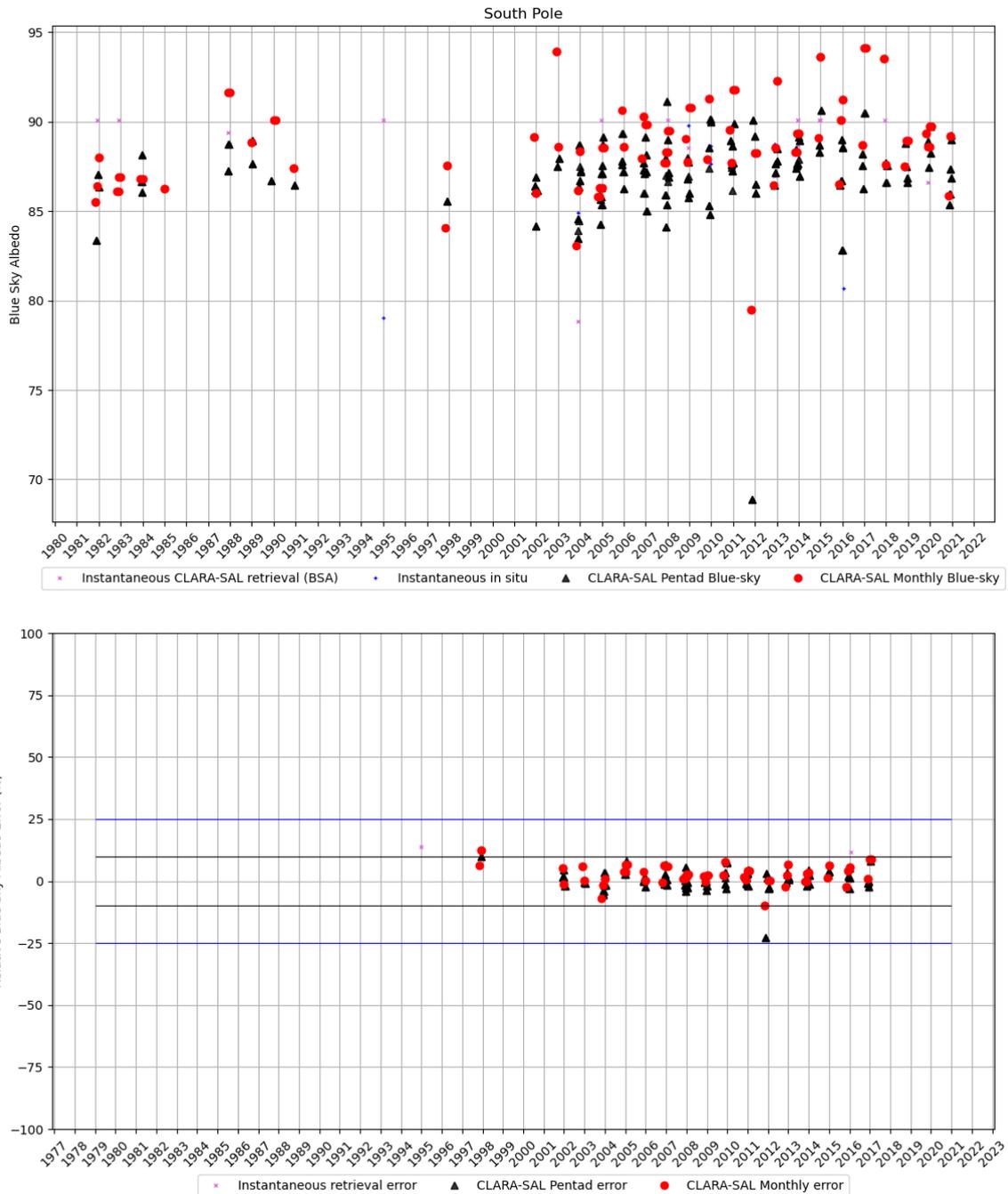


**Figure 23:** Validation results for Georg von Neumayer (GVN)

Neumayer is our first Antarctic site; located on the ice shelf off the coast of Queen Maud Land, the in situ measurements cover only snow with limited or no annual surface melt, resulting in a consistently high albedo. Against this background, the GAC-resolution CLARA retrievals are in good agreement (mean bias -6%), whereas both pentad and monthly means show considerable underestimation (-40%). Again, this resolution dependency shows that the root cause is in the point-to-pixel comparison rather than in the algorithm itself. GVN is also the

only site in the evaluation where some measurable trend in bias exists, although its significance is hard to decipher due to the large month-to-month variability in retrieval errors.

### 5.13. South Pole (SPO)

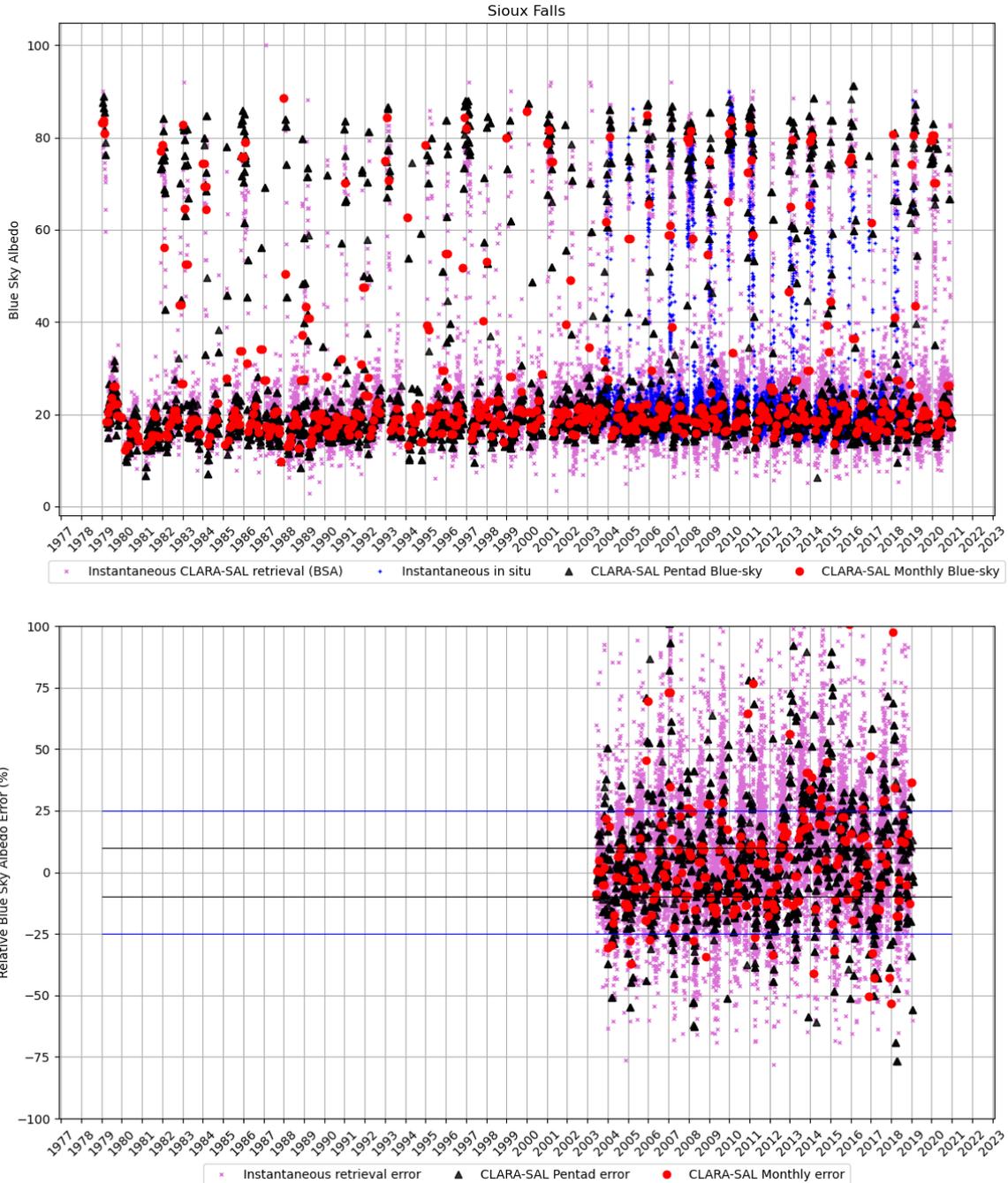


**Figure 24:** Validation results for South Pole (SPO)

The South Pole site understandably has very challenging solar geometry conditions, with very few comparable BAL-insitu retrieval/observation pairs available. Within the confines of this limited data, the retrievals compare favourably against the observations. Being at the center

of the the ice sheet, SPO has naturally a very stable, dry, and bright snow cover with very little variability in surface albedo.

### 5.14. Sioux Falls (SXF)

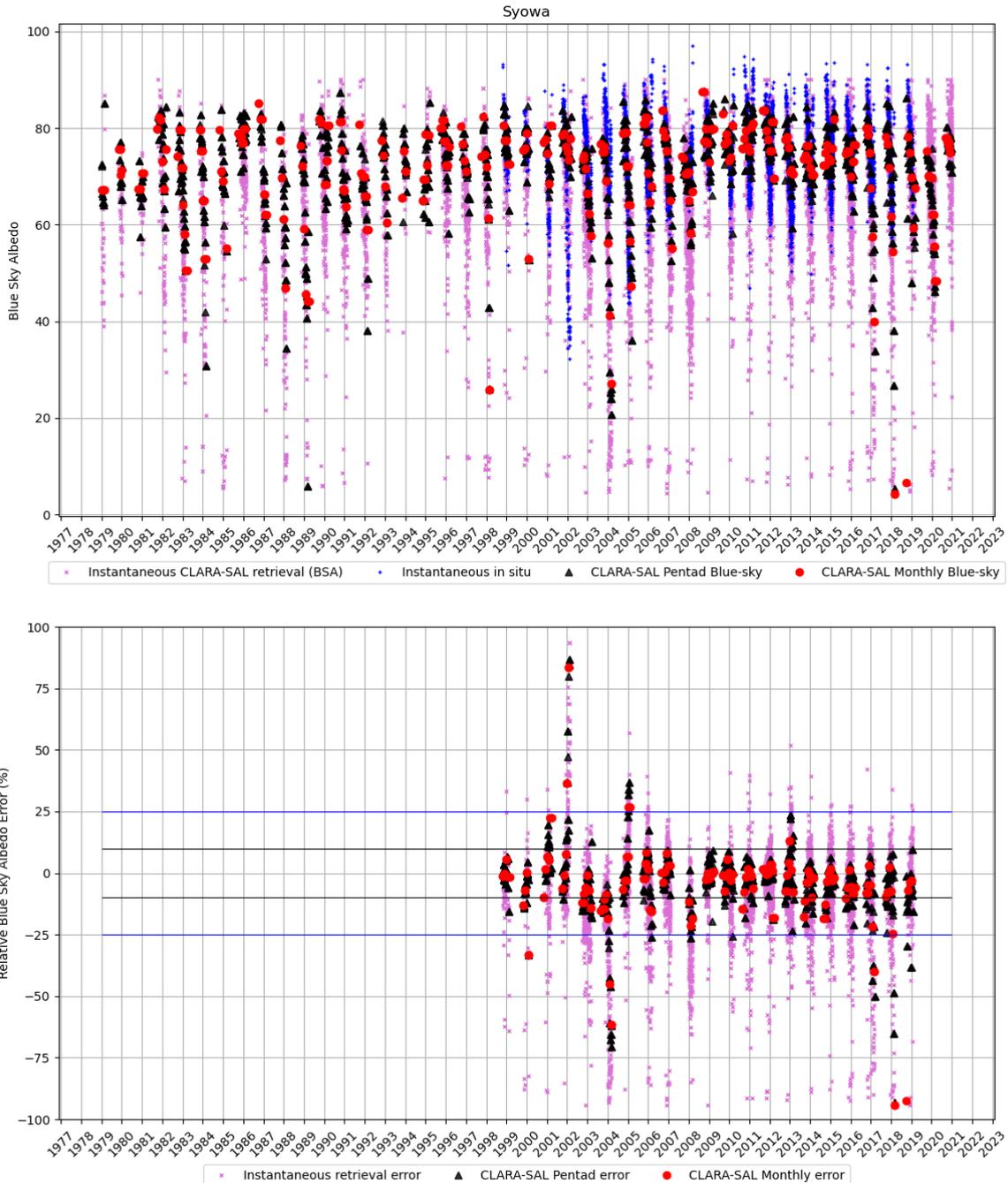


**Figure 25:** Validation results for Sioux Falls (SXF)

The Sioux Falls site in South Dakota is dominated by croplands (over 80% of surrounding land cover; [https://savs.eumetsat.int/html/SXF\\_report.html](https://savs.eumetsat.int/html/SXF_report.html)), implying good spatial representativeness, though frequent snowfalls in winter and the agricultural impacts on surface albedo do enhance variability at both local and grid cell scales. Despite this, BAL estimates at

all resolution levels generally track the mean albedo well, with larger retrieval errors mainly occurring during the snow-covered spring where local snow cover may differ from the areal mean.

### 5.15. Syowa (SYO)



**Figure 26:** Validation results for Syowa (SYO)

Syowa (Shōwa) station is a Japanese research outpost located on East Ongul Island, Queen Maud Land, Antarctica. Although the location is coastal, sea ice coverage is typically consistent, resulting in few issues in the point-to-pixel comparison as the sea ice is usually

snow-covered with limited surface melt during the austral summer. During the most recent years of coverage, though, anomalously low sea ice coverage has caused a decrease in grid cell-resolution retrievals (see 2017 and 2019 in Figure 26). This was also noted during the validation of the CLARA-A2.1 data record (DOI: 10.5676/EUM\_SAF\_CM/CLARA\_AVHRR/V002\_01), where the sea ice coverage anomalies were also verified from high-resolution satellite imagery.

### 5.16. Toravere (TOR)

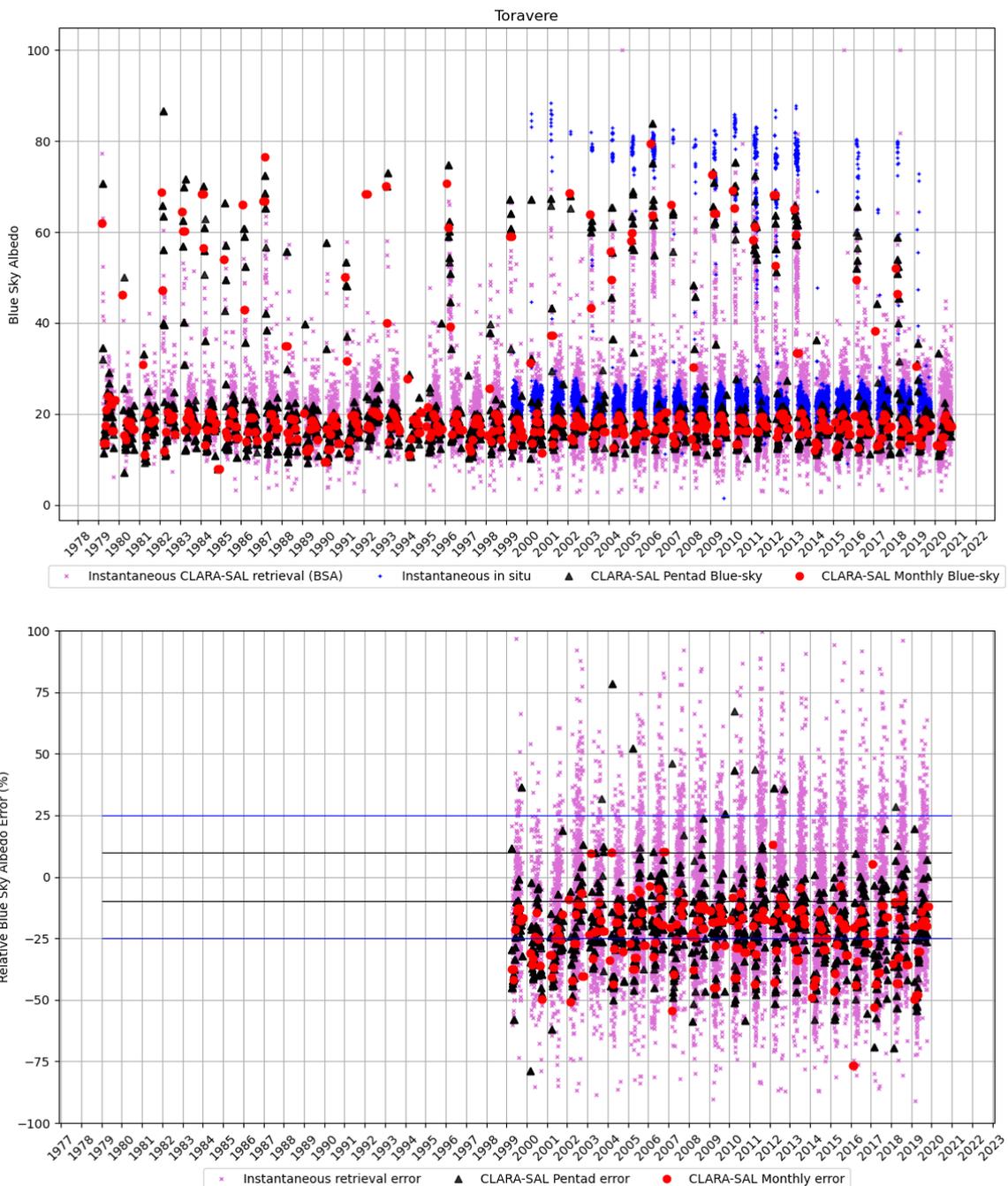


Figure 27: Validation results for Toravere (TOR)

	<p style="text-align: center;"><b>Validation Report</b>  <b>Surface Albedo (SAL)</b>  <b>CLARA Edition 3</b></p>	<p>Doc. No: SAF/CM/FMI/VAL/CLARA/SAL  Issue: 3.1  Date: 06.02.2023</p>
---	--	--

At Toravere, the in situ measurement takes place at a grassland patch, while the surrounding area is generally composed of a forest-cropland mix. The typical albedo for European grasslands (0.2 – 0.24; Sütterlin et al., 2016) is well manifested in the in situ measurements, whereas the CLARA retrievals at the 0.25 degree resolution clearly underestimate this. Again, we note that the GAC-resolution retrievals are more closely aligned with the in situ record, suggesting that representativeness rather than an algorithmic tendency for underestimation causes the majority of the bias, although spatial representativeness of the 5-9 km GAC resolution retrievals is still at best partial against the point-like nature of the in situ measurement. However, the bias at both resolution levels is stable in time throughout the 20-year evaluation period.

## 6. Validation of WAL and SAL against filtered diffuse/direct illumination BSRN observations

### Bias

The validation results for bias in WAL and SAL against BSRN sites are collected and shown in Table 8. The bottom row indicates the overall mean bias, where the spatially poorly representative sites are excluded (row above that displays bias with their inclusion)

**Table 8:** WAL and SAL monthly and pentad mean bias (rMBE) validation results (at 0.25 degree resolution) against BSRN observations filtered for near-fully diffuse/direct illumination conditions. Metric is relative bias against reference observation [%].

Station	WAL_MM	WAL_PM	SAL_MM	SAL_PM
ALE	101,9	153,7	48,5	79,1
BON	-23,3	-22,5	-23,7	-26,1
BOU	-10,9	-3,5	12,0	12,0
DRA	-5,3	-7,9	-2,5	-3,4
FPE	-2,2	1,5	0,7	3,3
GVN	-39,8	-37,3	-42,4	-35,9
SXF	5,0	5,6	2,8	1,0
SPO	-0,4	-2,8	2,6	0,8
E13	-10,2	-7,7	-8,2	-8,2
TOR	-26,8	-24,0	-28,3	-28,6
SYO	-1,2	-0,8	-9,6	-8,5
GCR	-25,9	-25,1	-27,8	-30,5
Mean	-11,4	-9,3	-10,2	-10,2
Mean of representative sites	-8,9	-7,3	-4,6	-5,3

Relative to the requirement for bias, both WAL and SAL fulfill the ‘target’ accuracy (<25% for WAL, <15% for SAL) excepting some (unrepresentative) sites, notably Alert (ALE), and Georg von Neumayer (GVN). Both of these are coastal or near-coastal sites with variable snow/ice cover and low ocean albedo contributions in the level3 WAL and SAL data. ALE furthermore exhibits very low albedo during summer (~0.1), indicating snow-free tundra conditions, whereas the 0.25 degree SAL/WAL grid cell evidently contains either snow or sea ice contributions, as the corresponding satellite-based estimates do not fall below 0.4 – implying a 300% bias for any months or pentads thus affected.

### Precision

The results for bias-corrected root mean square deviation are collected in Table 9. Apart from ALE and GVN, all sites fulfill the criteria for precision performance. The mean precision across all evaluated BSRN sites fulfills the ‘target’ level criterion (<0.1) for both pentad and monthly mean SAL and WAL estimates.

The reasonably good precision metrics obtained here are also fully in line with the temporally stable BAL retrievals evaluated in the preceding section, and further demonstrate the good

radiometric stability of the underlying AVHRR radiance FCDR, as well as the benefits obtained from the use of probabilistic cloud masking to enhance the robustness of processable AVHRR observations.

**Table 9:** WAL and SAL monthly and pentad mean validation results for precision (at 0.25 degree resolution) against BSRN observations filtered for near-fully diffuse/direct illumination conditions. Metric is bias-corrected root mean square deviation [bc-rmsd; unitless].

Station	WAL_MM	WAL_PM	SAL_MM	SAL_PM
ALE	0,235	0,327	0,120	0,138
BON	0,095	0,099	0,091	0,091
BOU	0,118	0,125	0,043	0,049
DRA	0,023	0,034	0,014	0,021
FPE	0,068	0,055	0,039	0,033
GVN	0,296	0,340	0,310	0,331
SXF	0,112	0,105	0,052	0,052
SPO	0,028	0,034	0,045	0,045
E13	0,051	0,050	0,029	0,033
TOR	0,095	0,074	0,167	0,123
SYO	0,069	0,061	0,138	0,121
GCR	0,044	0,054	0,048	0,050
<b>Mean</b>	<b>0,083</b>	<b>0,085</b>	<b>0,067</b>	<b>0,065</b>
<b>Mean of repr. sites</b>	<b>0,07</b>	<b>0,073</b>	<b>0,048</b>	<b>0,048</b>

### Stability

In accord with the BAL visualizations by site, with the possible exception of GVN there are no significant trends in bias during the evaluation period for either white- or black-sky albedo estimates. Stability metrics fulfill target levels for WAL, and optimum levels for SAL. Tables Table 10 to Table 13 summarize the results including stability (for the representative BSRN sites).

**Table 10:** WAL MM performance metrics and compliance to requirements

Variable	Time resolution	Indicator (Metric)	Value	Requirement level fulfilled
WAL	MM	Bias (level3 rMBE)	-8.9%	Target
WAL	MM	Precision (bc-rms)	0.07	Target
WAL	MM	Stability (bias dec. trend)	2.76 %	Target

**Table 11:** WAL PM performance metrics and compliance to requirements

Variable	Time resolution	Indicator (Metric)	Value	Requirement level fulfilled
WAL	PM	Bias (level3 rMBE)	-7.3%	Target
WAL	PM	Precision (bc-rms)	0.073	Target
WAL	PM	Stability (bias dec. trend)	3.49%	Target

**Table 12:** SAL MM performance metrics and compliance to requirements

Variable	Time resolution	Indicator (Metric)	Value	Requirement level fulfilled
SAL	MM	Bias (level3 rMBE)	-4.6%	Optimum
SAL	MM	Precision (bc-rms)	0.048	Target
SAL	MM	Stability (bias dec. trend)	-1.01%	Optimum

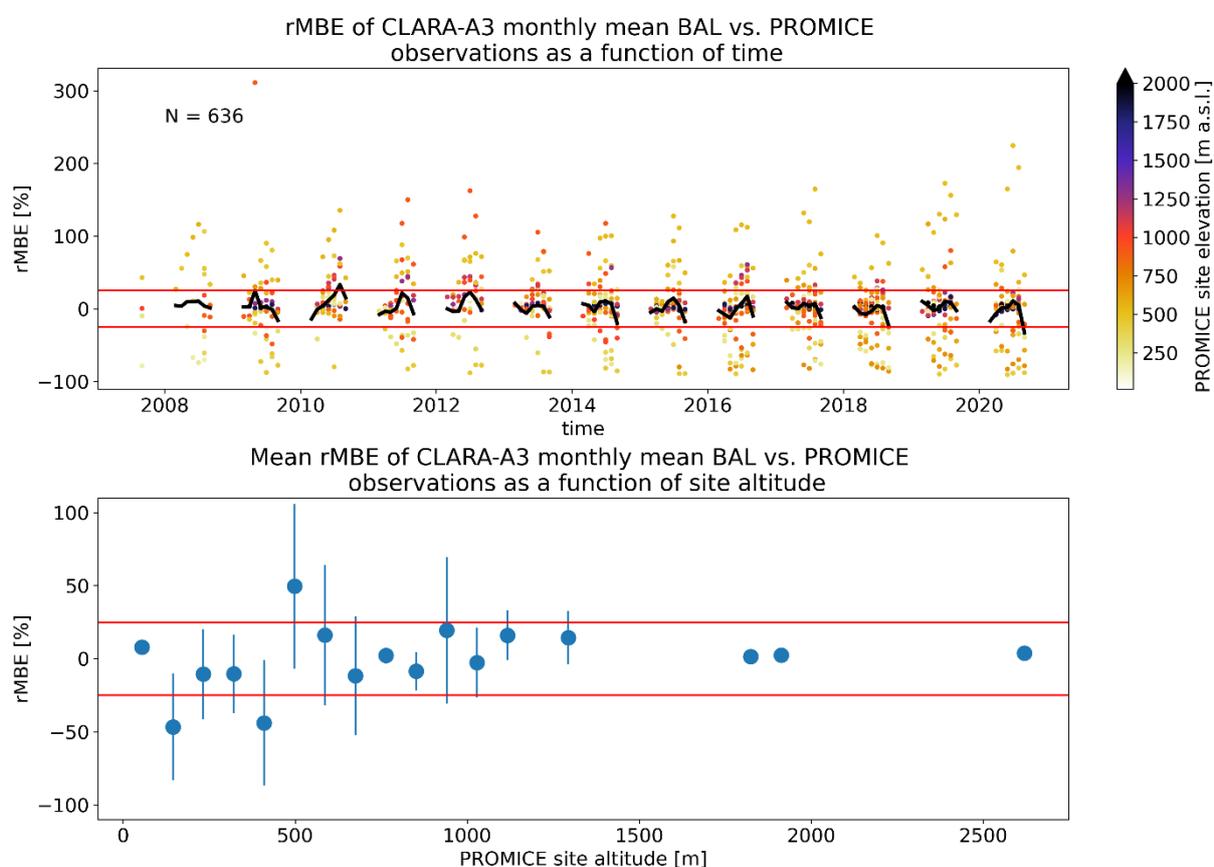
**Table 13:** SAL PM performance metrics and compliance to requirements

Variable	Time resolution	Indicator (Metric)	Value	Requirement level fulfilled
SAL	PM	Bias (level3 rMBE)	-5.3%	Target
SAL	PM	Precision (bc-rms)	0.048	Target
SAL	PM	Stability (bias dec. trend)	3.11%	Target

## 7. Validation against PROMICE in situ observations

### 7.1. BAL

The validation results of monthly mean BAL against PROMICE observations are shown in Figure 28. As is readily apparent, the retrievals are stable over higher elevations on the ice sheet (the accumulation zone), where the snow cover experiences little surface melt and slow metamorphism during the observable summer period. At low elevations near to the ice sheet margins, there is markedly larger variability in the bias. This is expected given that the melting season over the ablation zone induces considerable surface changes, ranging from large-scale formation of melt ponds and meltwater canals to exposure of impurities following the melt of topmost snow layers (Knap & Oerlemans, 1996).



**Figure 28:** Top, relative bias (rMBE) of monthly mean CLARA-A3 BAL against PROMICE in situ observations. Each marker denotes one monthly mean at one PROMICE site, colored by site elevation. Red lines indicate +/- 25% rMBE thresholds. Magenta line indicates the mean rMBE for each month across all sites. Bottom, the overall mean rMBE during the evaluation period for the PROMICE sites, binned and shown as a function of site elevation (30 bins). Whiskers indicate std.dev. of rMBE in each bin.

The collated performance metrics of MM BAL vs PROMICE (Table 14) quantify these findings, showing low mean bias and excellent stability (of said bias), but conversely low precision due to the challenges in tracking in situ surface albedo across the heterogeneous surface conditions of the ice sheet's ablation zone.

We note that precision in the full PROMICE group does not fulfill any of the stated performance requirements, but we argue that the inclusion of the sites near the ice sheet margins (with incomplete surrounding snow/ice cover) presents an unfairly challenging environment for the coarse-resolution BAL retrievals. As a reminder, because the PROMICE sites are not tracked for overpass data in the same way as BSRN sites are, the validation here is based on point-to-pixel comparisons at the 25 kilometer North Polar subset of CLARA-A3 BAL.

**Table 14:** BAL MM performance metrics against PROMICE in situ reference observations

Variable	Time resolution	Indicator (Metric)	Value	Requirement level fulfilled
BAL	MM	Bias (level3 rMBE)	2.73%	Optimum
BAL	MM	Bias of sites with full snow/ice cover in CLARA grid cell	3.55%	Optimum
BAL	MM	Precision (bc-rms)	0.19	None
BAL	MM	Precision of sites with full snow/ice cover in CLARA grid cell	0.12	Threshold
BAL	MM	Stability (bias dec. trend)	-0.28%	Optimum
BAL	MM	Stability of sites with full snow/ice cover in CLARA grid cell	-1.01%	Optimum

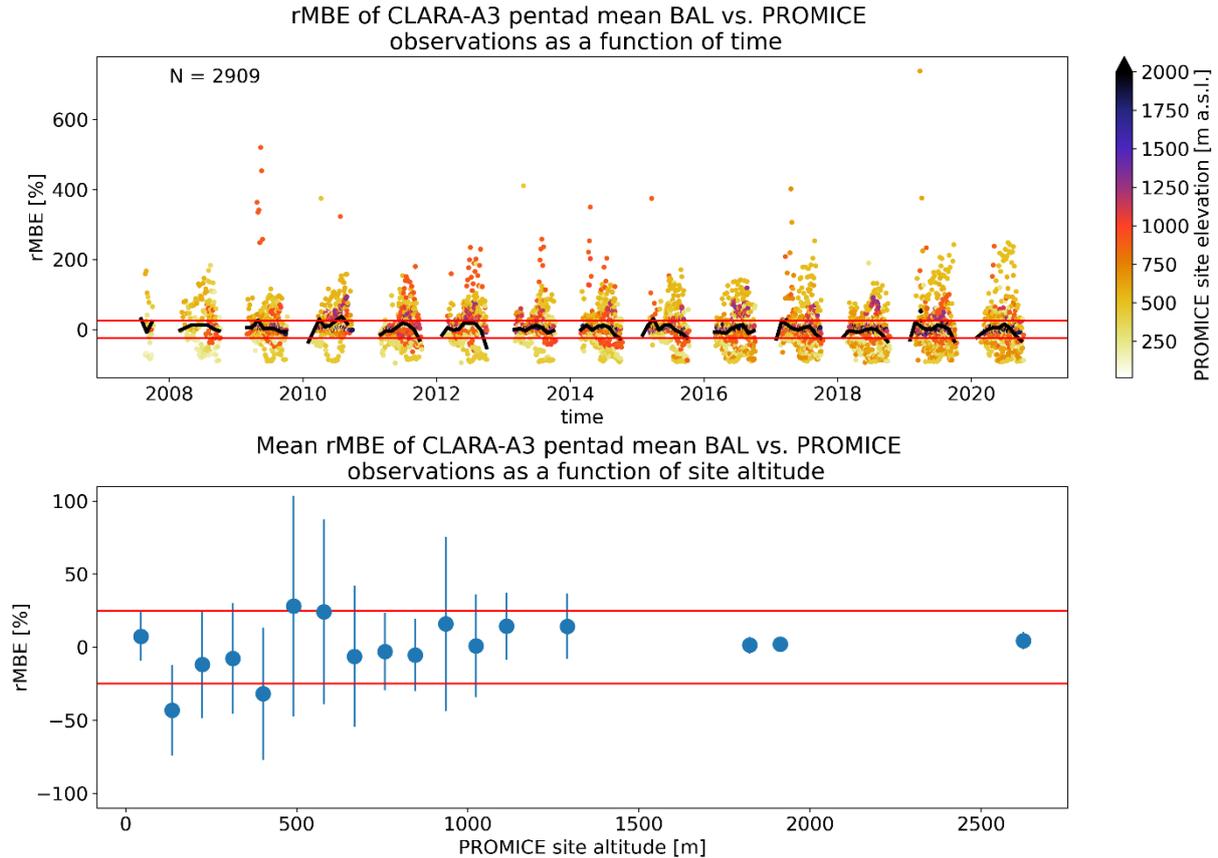
For the pentad means, the validation results (Figure 29) naturally show similarity with the monthly mean results, but with increased scatter implying a lower precision. This impression is confirmed in the overall BAL PM performance metrics (Table 15). The mutually compensating under- and overestimations from sites near the ice sheet margins result in low mean bias with a negligible trend, but the metric for precision shows their effect.

The question then is, is this substantial scatter in bias an indication of algorithm deficiency, or an unfair comparison in the point-to-pixel validation procedure? To investigate, we again obtained land cover fractions of the North Polar-subset CLARA grid cells containing PROMICE sites from Google's Dynamic World dataset. Figure 30 shows the June-August site-specific precision as a function of mean snow cover fraction in the CLARA grid cell for the same months. Obviously and as expected, precision improves with better coverage of ice sheet in the level 3 averages.

Further, in grid cells with total snow cover, it is apparent that precision against station observations does still display marked variability. However, it should be noted that the Dynamic World data does not delineate between e.g. wet snow, dry snow, bare ice patches, or melt ponds, meaning that the surface albedo variability around the sites is not fully characterized by snow cover fraction alone. Site elevation is often a useful proxy for the intensity of melt

season, and evidently here as well precision is better higher up the ice sheet where melt effects are more subdued.

This evidence indicates that the non-fulfillment of the precision threshold for the full PROMICE evaluation is indeed a result of poor representativeness. As the metric is calculated in squared space, sites with high bias dominate the all-data mean. Calculated on a per-site basis, the pentad BAL precision is outside the threshold requirement for only 5 sites out of 26 (Figure 30).

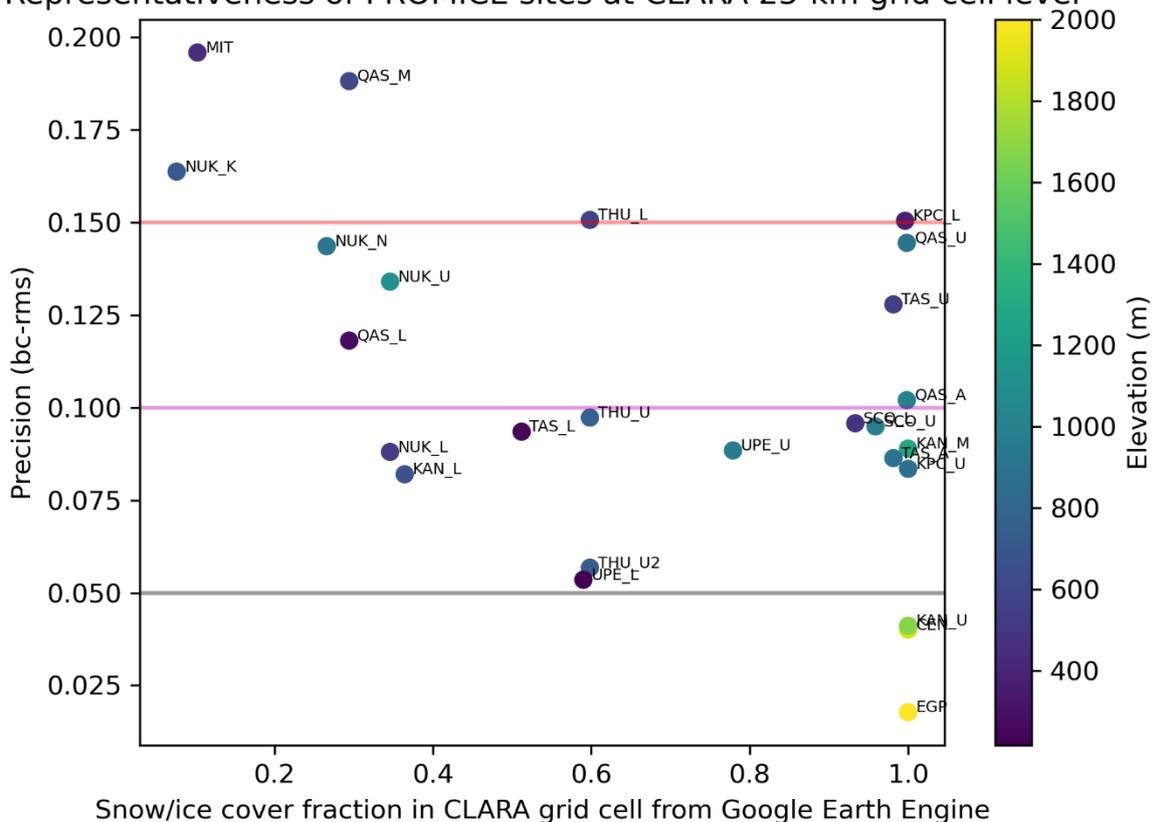


**Figure 29:** Top, relative bias (rMBE) of pentad mean CLARA-A3 BAL against PROMICE in situ observations. Each marker denotes one pentad mean at one PROMICE site, colored by site elevation. Red lines indicate +/- 25% rMBE thresholds. Magenta line indicates the mean rMBE for each month across all sites. Bottom, the overall mean rMBE during the evaluation period for the PROMICE sites, binned and shown as a function of site elevation (30 bins). Whiskers indicate std.dev. of rMBE in each bin.

**Table 15:** BAL PM performance metrics against PROMICE in situ reference observations

Variable	Time resolution	Indicator (Metric)	Value	Requirement level fulfilled
BAL	PM	Bias (level3 rMBE)	4.41%	Optimum
BAL	PM	Bias of sites with full snow/ice cover in CLARA grid cell	-1.47%	Optimum
BAL	PM	Precision (bc-rms)	0.21	None
BAL	PM	Precision of sites with full snow/ice cover in CLARA grid cell	0.13	Threshold
BAL	PM	Stability (bias dec. trend)	-0.32%	Optimum
BAL	PM	Stability of sites with full snow/ice cover in CLARA grid cell	-1.04%	Optimum

Representativeness of PROMICE sites at CLARA 25-km grid cell level



**Figure 30:** Precision (bc-rms) of BAL during June-August as a function of snow cover in CLARA grid cell (25 km EASE polar subset) containing PROMICE sites. Snow/ice fraction data from Google Earth Engine's Dynamic World dataset. Color of markers indicates PROMICE site elevation. The red,

magenta, and grey lines indicate the threshold, target, and optimum precision requirement levels, respectively.

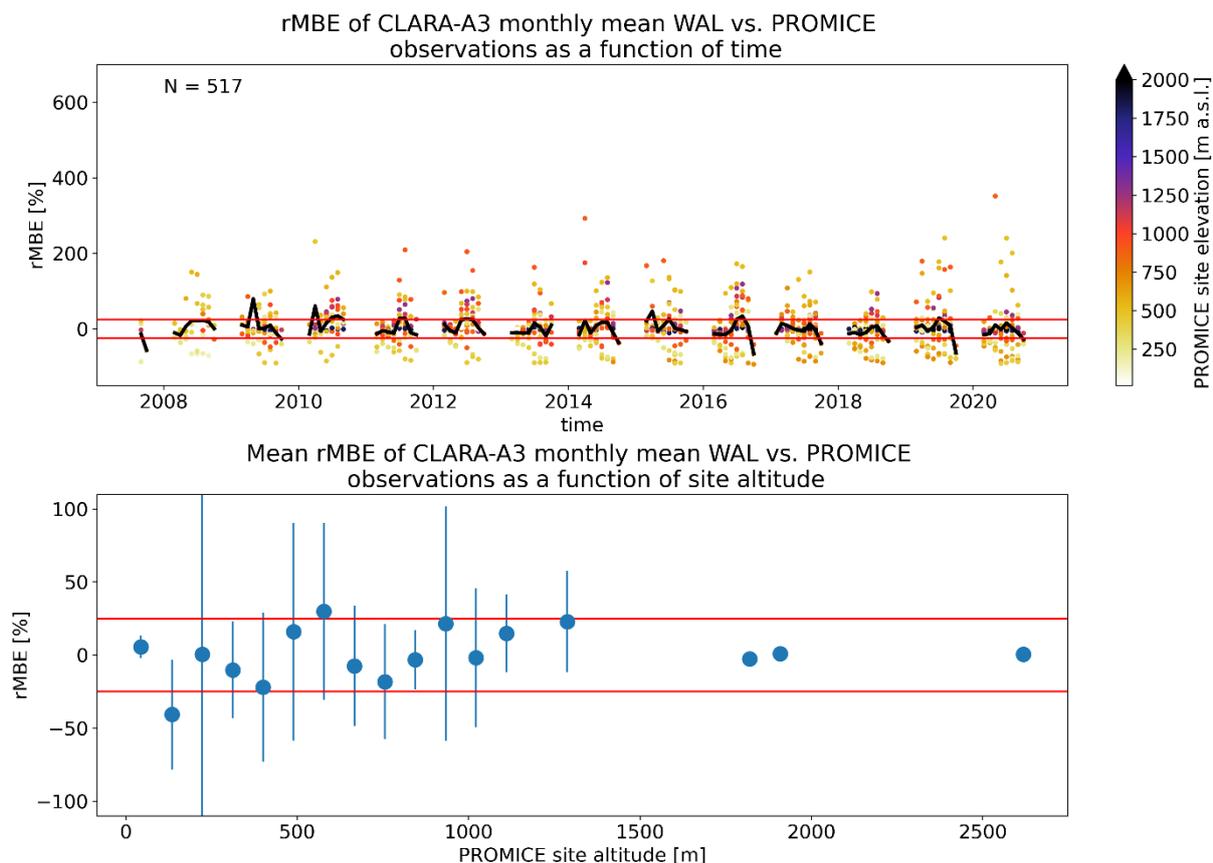
## 7.2. WAL

In the comparison against the black-sky albedo estimate SAL, the PROMICE reference data is filtered to contain only those hourly observations where measured cloud cover is >0.99. This is admittedly an imprecise method to attempt to obtain reference albedo under fully diffuse illumination conditions (as even full cloud cover does not guarantee lack of directionality in incoming irradiance, especially under optically thin clouds), yet it is the only available means for a quantitative evaluation of the white-sky WAL estimate over the PROMICE sites.

The monthly mean validation metrics (Table 16) display low bias, consistently with BAL. However, in terms of precision WAL exhibits more scatter, which is consistent with expectations given that WAL over snow and ice is derived from SAL observations using a statistical relationship. Therefore, spatial representativeness issues in SAL will translate into similar issues in WAL (and BAL). In contrast, decadal stability in bias remains very good, confirmed with a visual inspection of the rMBE across time (Figure 31).

**Table 16:** PROMICE validation metrics for monthly mean WAL

Variable	Time resolution	Indicator (Metric)	Value	Requirement level fulfilled
WAL	MM	Bias (level3 rMBE)	4.62%	Target
WAL	MM	Bias of sites with full snow/ice cover in CLARA grid cell	4.98%	Optimum
WAL	MM	Precision (bc-rms)	0.22	None
WAL	MM	Precision of sites with full snow/ice cover in CLARA grid cell	0.17	None
WAL	MM	Stability (bias dec. trend)	-0.40%	Optimum
WAL	MM	Stability of sites with full snow/ice cover in CLARA grid cell	-0.36%	Optimum



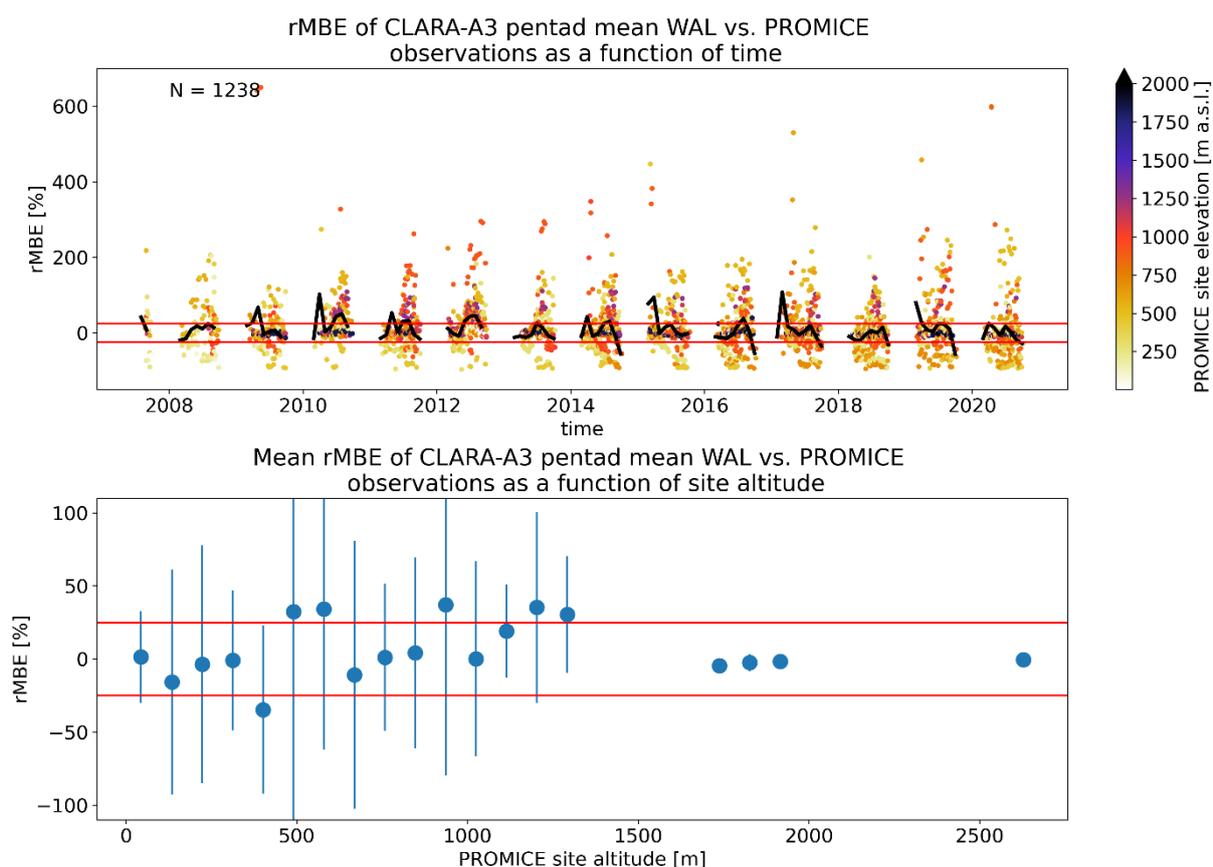
**Figure 31:** Top, relative bias (rMBE) of monthly mean CLARA-A3 WAL against PROMICE in situ observations. Each marker denotes one monthly mean at one PROMICE site, colored by site elevation. Red lines indicate  $\pm 25\%$  rMBE thresholds. Magenta line indicates the mean rMBE for each month across all sites. Bottom, the overall mean rMBE during the evaluation period for the PROMICE sites, binned and shown as a function of site elevation (30 bins). Whiskers indicate std.dev. of rMBE in each bin.

The results at the pentad mean level (Table 17 & Figure 32) broadly reflect the same characteristics. Of note, the bias over the full PROMICE set of sites is higher due to a small number of very high rMBE over individual sites and pentads during the evaluation period. The bias over the subset of sites with continuous surrounding snow/ice cover remains comparable to the monthly means. Precision is again very low due to the influence of the high scatter. Stability is very good and comparable to the monthly mean retrievals, as expected.

**Table 17:** PROMICE validation metrics for pentad mean WAL

Variable	Time resolution	Indicator (Metric)	Value	Requirement level fulfilled
WAL	PM	Bias (level3 rMBE)	12.02%	Target
WAL	PM	Bias of sites with full snow/ice cover in CLARA grid cell	4.76%	Optimum
WAL	PM	Precision (bc-rms)	0.25	None
WAL	PM	Precision of sites with full snow/ice	0.26	None

		cover in CLARA grid cell		
WAL	PM	Stability (bias dec. trend)	-0.81%	Optimum
WAL	PM	Stability of sites with full snow/ice cover in CLARA grid cell	-1.04%	Optimum



**Figure 32:** Top, relative bias (rMBE) of pentad mean CLARA-A3 WAL against PROMICE in situ observations. Each marker denotes one pentad mean at one PROMICE site, colored by site elevation. Red lines indicate +/- 25% rMBE thresholds. Magenta line indicates the mean rMBE for each month across all sites. Bottom, the overall mean rMBE during the evaluation period for the PROMICE sites, binned and shown as a function of site elevation (30 bins). Whiskers indicate std.dev. of rMBE in each bin.

### 7.3. SAL

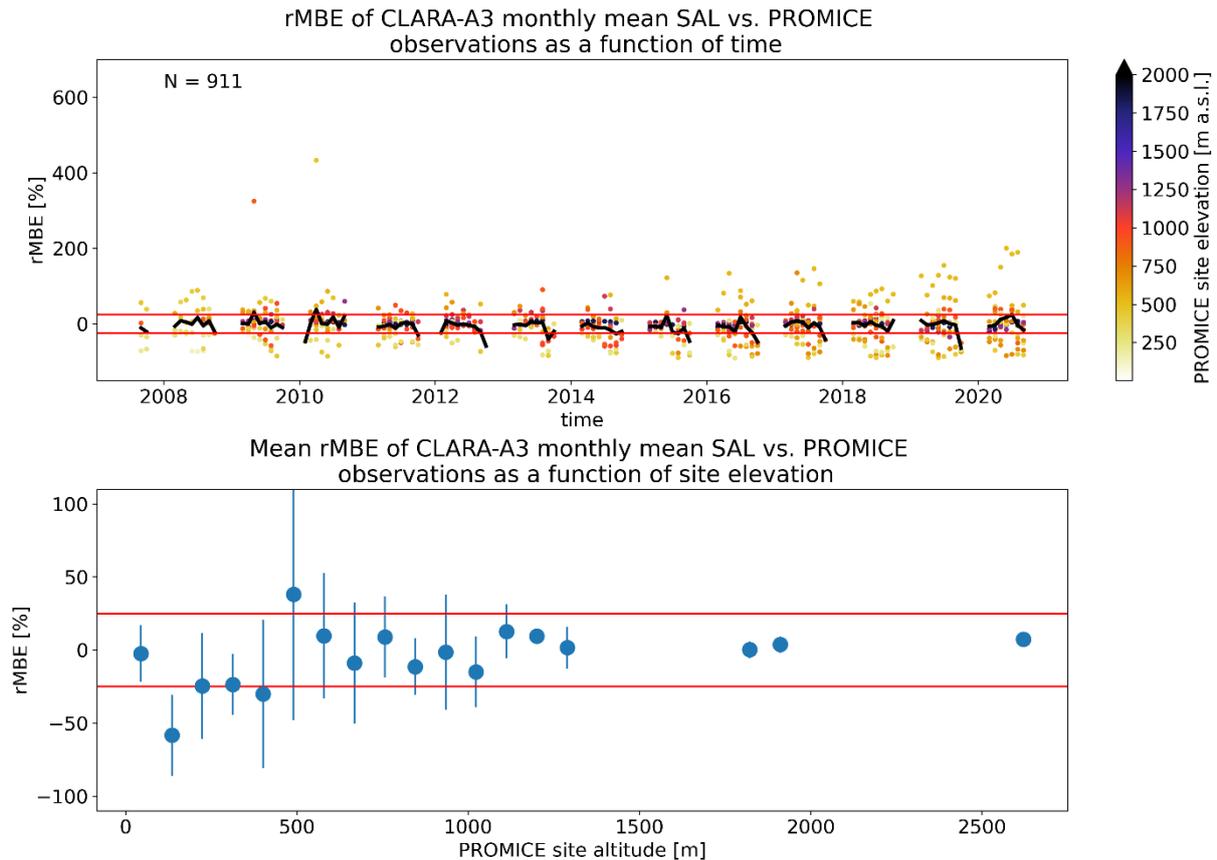
In the comparison against the black-sky albedo estimate SAL, the PROMICE reference data is filtered to contain only those hourly observations where measured cloud cover equals zero. The incoming solar radiation is not, strictly speaking, unidirectional even under fully cloud-free conditions. However, when SZA is close to 60 degrees, atmospheric influences tend to compensate for each other so that black- and blue-sky albedos align closely under cloud-free

conditions (Manninen et al., 2012). Given that the mean SZA over the ice sheet during summer is close to this condition (mean in evaluation set is 59.2 degrees), we elect to simply use the cloud-free in situ albedo observations as a proxy for black-sky albedo.

The results at monthly mean scale (Table 18, Figure 33) are naturally very similar to the blue- and white-sky validation results already discussed. Bias is very low and stable through the years, but precision remains low. However, as discussed earlier, this follows from poor representativeness in the point-to-pixel comparison. Furthermore, if we were to assess the overall precision as the mean of site-specific precision instead of precision from the full MBE data, the precision for monthly mean SAL would be reduced from 0.2 to 0.09, with the new result fulfilling the target requirement.

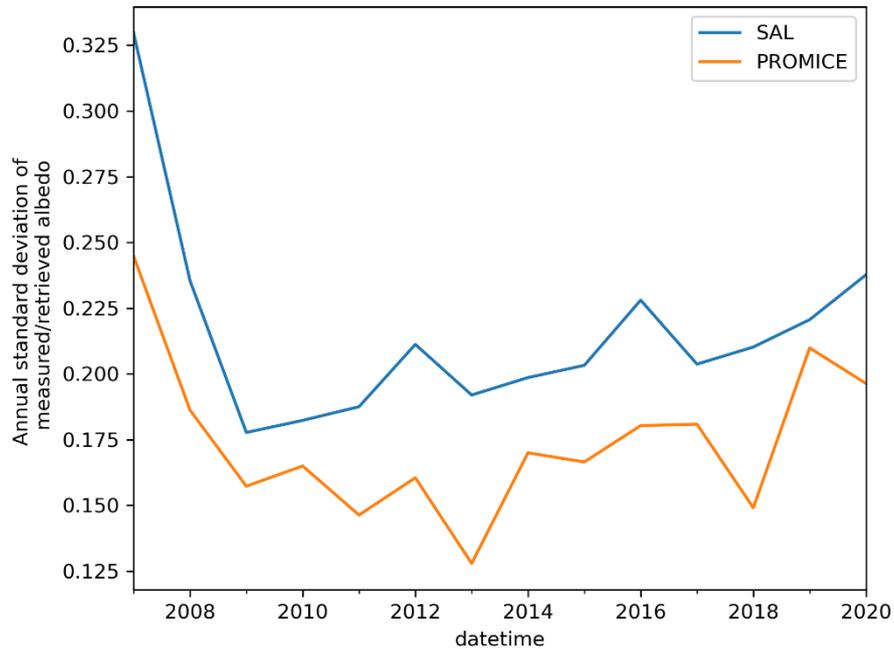
**Table 18:** PROMICE validation metrics for monthly mean SAL

Variable	Time resolution	Indicator (Metric)	Value	Requirement level fulfilled
SAL	MM	Bias (level3 rMBE)	-2.75%	Optimum
SAL	MM	Bias of sites with full snow/ice cover in CLARA grid cell	-5.20%	Target
SAL	MM	Precision (bc-rms)	0.2	None
SAL	MM	Precision of sites with full snow/ice cover in CLARA grid cell	0.13	Threshold
SAL	MM	Stability (bias dec. trend)	-0.3%	Optimum
SAL	MM	Stability of sites with full snow/ice cover in CLARA grid cell	-0.8%	Optimum



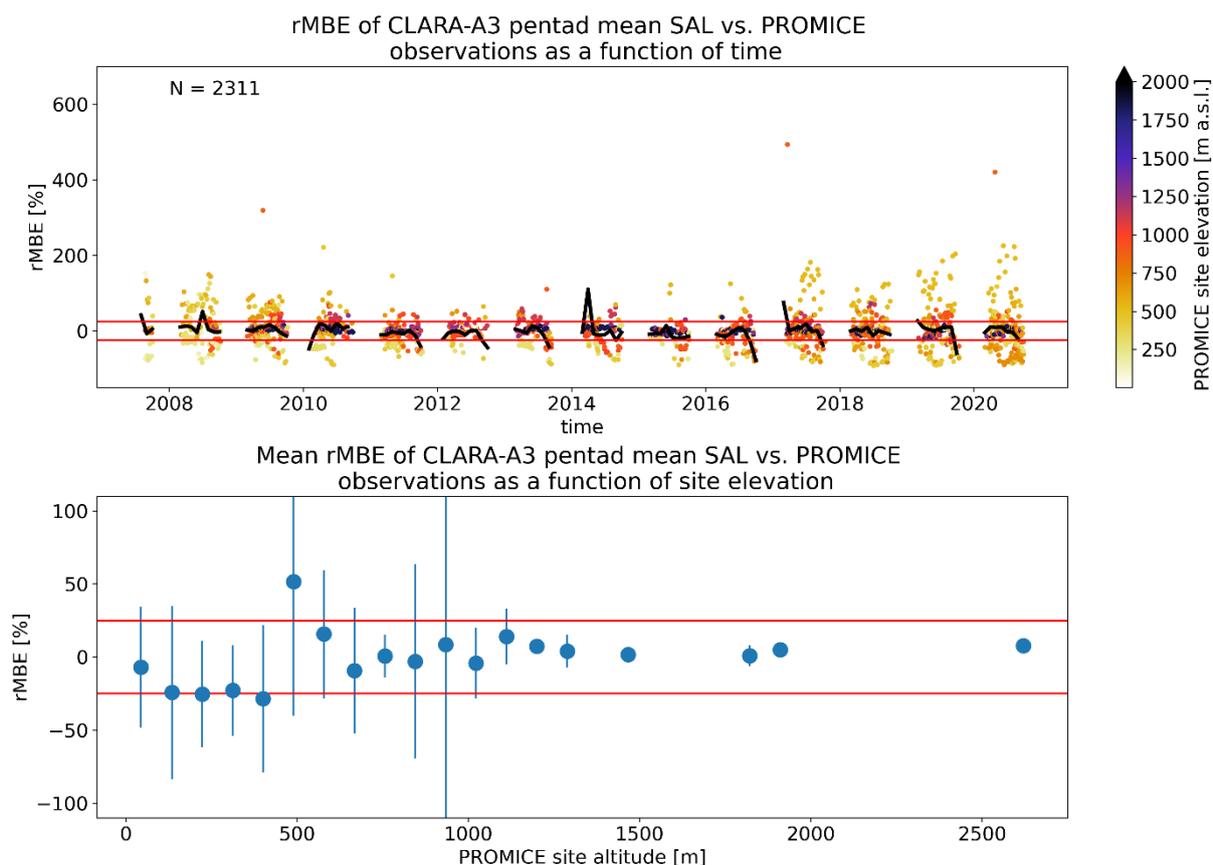
**Figure 33:** Top, relative bias (rMBE) of monthly mean CLARA-A3 SAL against PROMICE in situ observations. Each marker denotes one monthly mean at one PROMICE site, colored by site elevation. Red lines indicate  $\pm 25\%$  rMBE thresholds. Magenta line indicates the mean rMBE for each month across all sites. Bottom, the overall mean rMBE during the evaluation period for the PROMICE sites, binned and shown as a function of site elevation (30 bins). Whiskers indicate std.dev. of rMBE in each bin.

The pentad mean SAL validation results (Figure 35) are naturally similar to those of BAL and WAL. An interesting feature is the emerging increase in rMBE scatter during 2017-2020. Examination of the estimated and measured albedos over PROMICE sites show that the standard deviations of annually averaged SAL and PROMICE data are both exhibiting increasing trends since 2010 (Figure 34). The reason may be twofold: first, the amplifying surface melt of the ice sheet (e.g. Riihelä et al., 2019) broadens the variability range of its surface albedo towards the lower end, and second, the gradually increasing number of sites in the PROMICE network implies a gradually broadening set of measured/retrieved surface conditions, both contributing towards increasing standard deviations. The apparent variability increase in the surface observations implies that any algorithmic deficiencies would not fully explain this apparent performance decrease.



**Figure 34:** Annual standard deviation of SAL and PROMICE-measured albedos under cloud-free conditions. All PROMICE sites considered.

The quantified mean metrics (Table 19) show, as for the other albedo estimates, very good performance in terms of bias and stability, but substantial scatter in bias again leads to very large precision. Note that the precision threshold for pentad mean SAL is set to a stricter value than for the other estimates, leading to non-fulfillment of the threshold performance even for the set of sites with full snow/ice cover in the CLARA grid cell.



**Figure 35:** Top, relative bias (rMBE) of pentad mean CLARA-A3 SWAL against PROMICE in situ observations. Each marker denotes one pentad mean at one PROMICE site, colored by site elevation. Red lines indicate  $\pm 25\%$  rMBE thresholds. Magenta line indicates the mean rMBE for each month across all sites. Bottom, the overall mean rMBE during the evaluation period for the PROMICE sites, binned and shown as a function of site elevation (30 bins). Whiskers indicate std.dev. of rMBE in each bin.

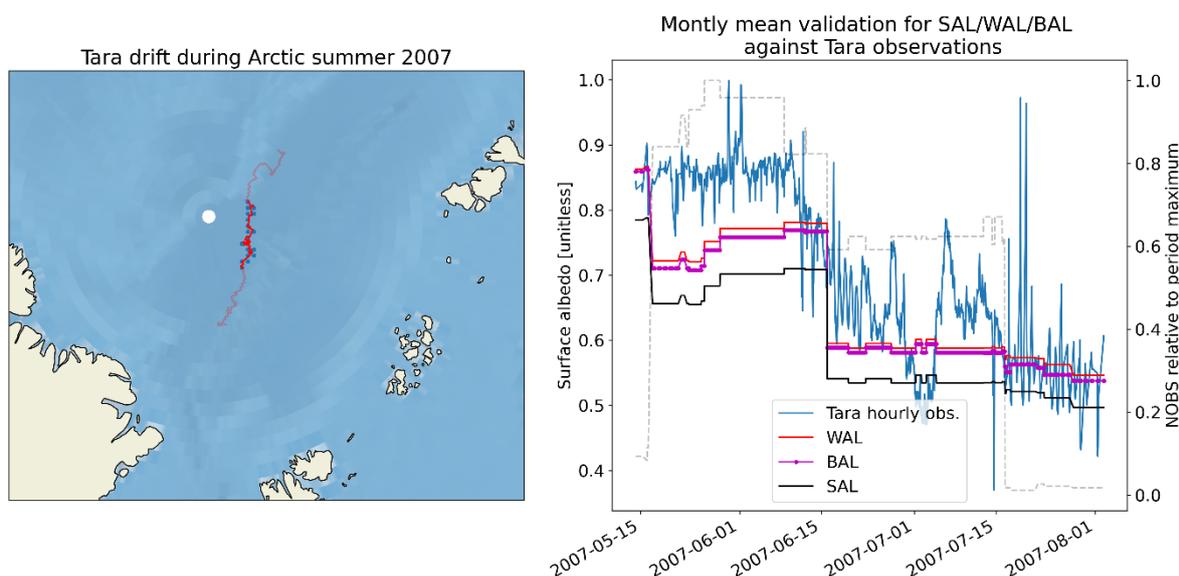
**Table 19:** PROMICE validation metrics for pentad mean SAL

Variable	Time resolution	Indicator (Metric)	Value	Requirement level fulfilled
SAL	PM	Bias (level3 rMBE)	2.07%	Optimum
SAL	PM	Bias of sites with full snow/ice cover in CLARA grid cell	-0.72%	Optimum
SAL	PM	Precision (bc-rms)	0.2	None
SAL	PM	Precision of sites with full snow/ice cover in CLARA grid cell	0.13	None
SAL	PM	Stability (bias dec. trend)	-0.7%	Optimum
SAL	PM	Stability of sites with full snow/ice cover in CLARA grid cell	-3.4%	Optimum

## 8. Validation against Tara and SHEBA drifting ice camp albedo observations over the Arctic Ocean

We first examine the monthly mean CLARA albedo estimates against Tara ice camp observations in Figure 36. The drift of the ice camp carried it over high-latitude multiyear sea ice with snow cover during the early part of the summer (May-early June), implying very high observed albedos of 0.8 – 0.9. The CLARA estimates generally underestimate the period albedo, although it is also clear that significant variability exists between CLARA grid cell retrievals in the region, as shown by the substantial ‘step changes’ apparent when the ice camp drifts from one CLARA grid cell to another. Note that because of the Sun Zenith Angle cutoff in SAL processing, no comparable albedo estimates are available in August.

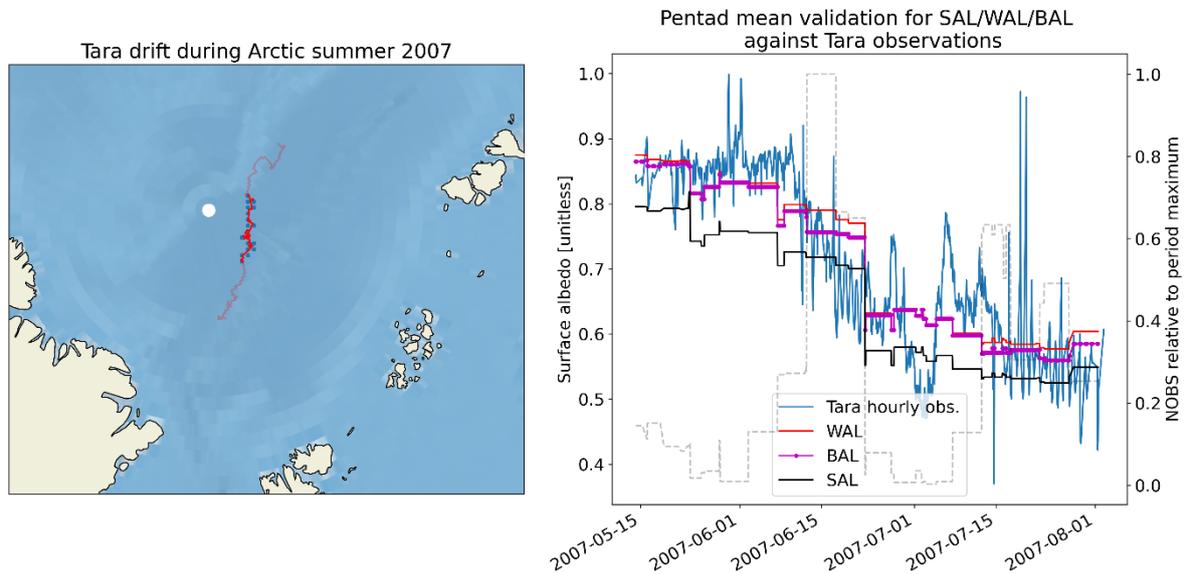
Melt onset of sea ice occurs at the ice camp during the first half of June, with the measured albedo beginning to decrease as the snow cover wets and melt ponds start to form. The monthly mean albedo does not have the temporal resolution to capture this transition phase, which instead appears as a step change in late June. However, it is notable that the mean level of surface albedo during the melt season is captured correctly in July, except during the albedo ‘rebound’ in early July which is a typical result of small-scale drainage of melt ponds (Perovich et al., 2002) and thus likely not of sufficiently broad scale to be capturable by the CLARA albedo estimates.



**Figure 36:** Visualization of Tara drift during summer 2007 (left), with red/blue markers denoting its collocatable part, and visualization of the Tara albedo observations vs. SAL/BAL/WAL estimates (right). All data shown for conditions for which  $SZA < 70$  deg., Tara data visualized as hourly means. CLARA-A3 SAL/WAL/BAL estimates from the spatiotemporally matched monthly mean grid cells in which the Tara schooner drifted during the course of the summer. Dashed line indicates the relative availability of valid AVHRR observations vs. the summer maximum. Note that the CLARA estimates are at the 25 kilometer spatial resolution while Tara observations are point-like.

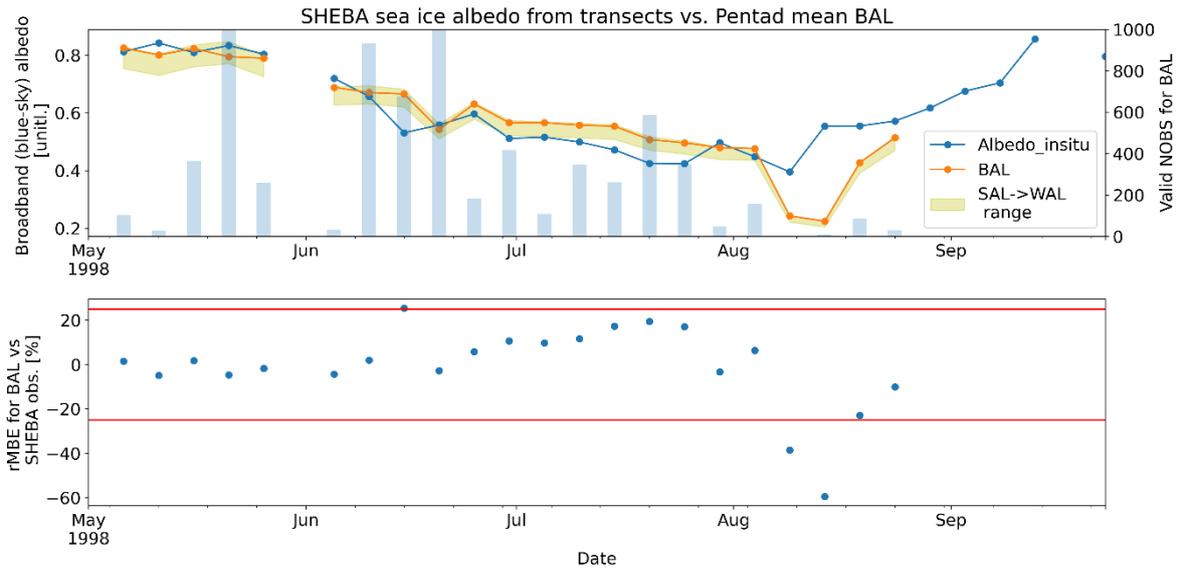
Conversely, and as expected, the validation of pentad mean CLARA albedo estimates against Tara observations (Figure 37) shows that the five-day means are better at capturing the albedo dynamic during melt onset and thereafter. The directly comparable albedo estimate (blue-sky

albedo BAL) is close to the white-sky (WAL) albedo because of the persistently high cloud cover over the Arctic Ocean during summer. Note also that the WAL estimate is constrained during processing to not exceed SAL by more than 10% over sea ice.



**Figure 37:** As Figure 36, but for the pentad mean SAL/BAL/WAL data during summer 2007.

The SHEBA ice camp observations are spatially more comprehensive (transects), but available only sparsely in time. We therefore limit the comparison to pentad means, whose validation is shown in Figure 38. Here, the blue-sky BAL estimates track the evolution of sea ice albedo well, with bias against the SHEBA transects within  $\pm 10\text{-}15\%$  for most pentads, excepting one 10-day period in early August. This increase in bias has been apparent in comparison against SHEBA since CLARA-A1 (Riihelä et al., 2013) and therefore is highly likely a combined result of an observation-independent lapse in spatial representativeness and very low sampling (i.e. high cloud cover) rather than any algorithm or cloud masking effect.



**Figure 38:** Top: In situ albedo from SHEBA transect measurements (blue, 5 day average) versus CLARA-A3 albedo pentad mean products – BAL in orange line & markers, with the variability range between SAL and WAL shown with a beige envelope. Faint blue overlaid bar chart indicates the amount of valid GAC-resolution SAL retrievals at the ice camp’s grid cell during the Arctic summer 1998. Below: Relative bias (rMBE) of CLARA-A3 BAL vs. SHEBA in situ measurements. Red lines indicate the target accuracy of 25% relative.

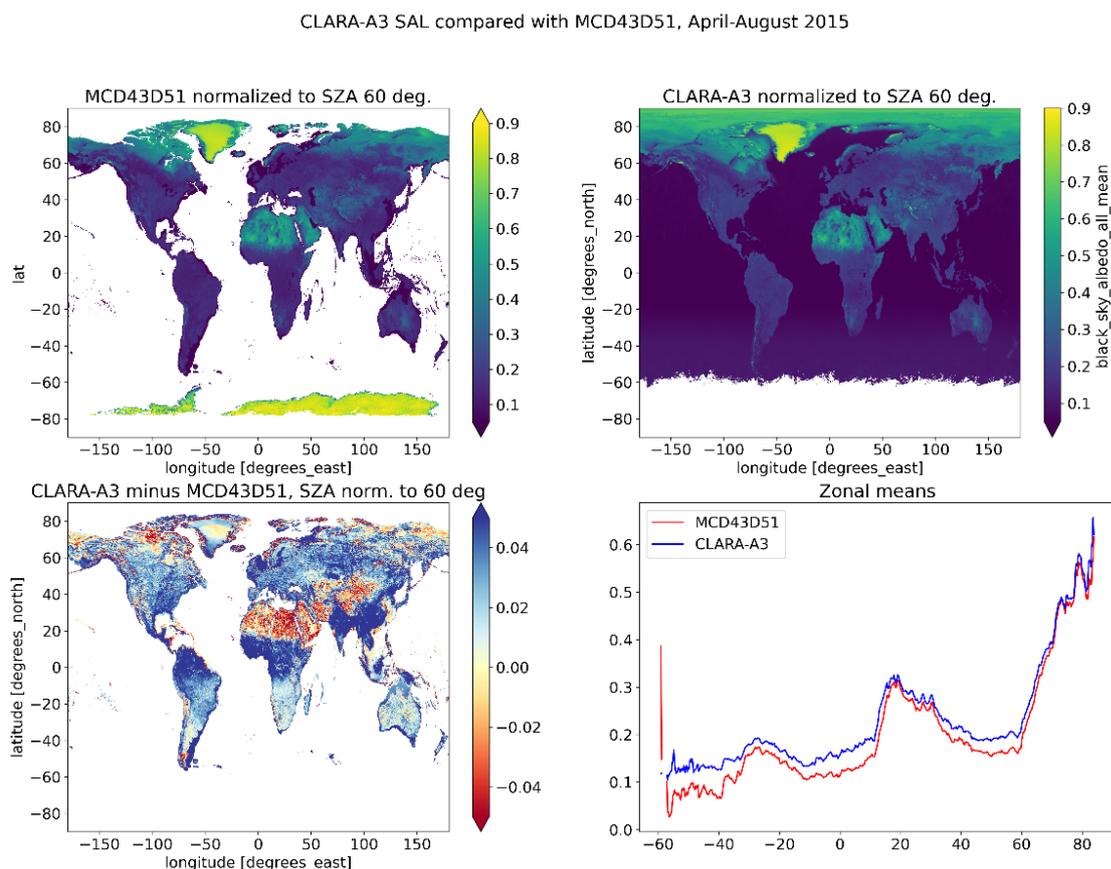
For the sea ice albedo validation, we only report the mean bias (rMBE) as the performance metric. The single-year durations of the expeditions exclude any stability analyses, and for precision it is very unlikely that calculable bc-rms error estimates would be a ‘true’ reflection of precision over a constantly moving validation target whose surface is in a near-constant state of flux during the evaluation period. Also, as diffuse and direct illumination measurements are not available, we only report the mean bias against pentad and monthly mean BAL.

Campaign	Monthly mean rMBE	Pentad mean rMBE
Tara	-7.3%	0.3%
SHEBA	--	-1.1%

## 9. Comparison against MCD43D51 edition 6.1 and CLARA-A2 SAL

In addition to point-to-pixel comparisons against reference data, we have sought to assess CLARA-A3 performance relative to other benchmark data records of surface albedo. Of these, the MODIS-based surface albedo estimates are perhaps the most widely used and very thoroughly validated (Schaaf et al., 2002; Cescatti et al., 2012, Wang et al., 2018). Also, to evaluate the consistency and evolution of CLARA albedo retrievals, we compare the A3 edition against the predecessor data record CLARA-A2 (Karlsson et al., 2017).

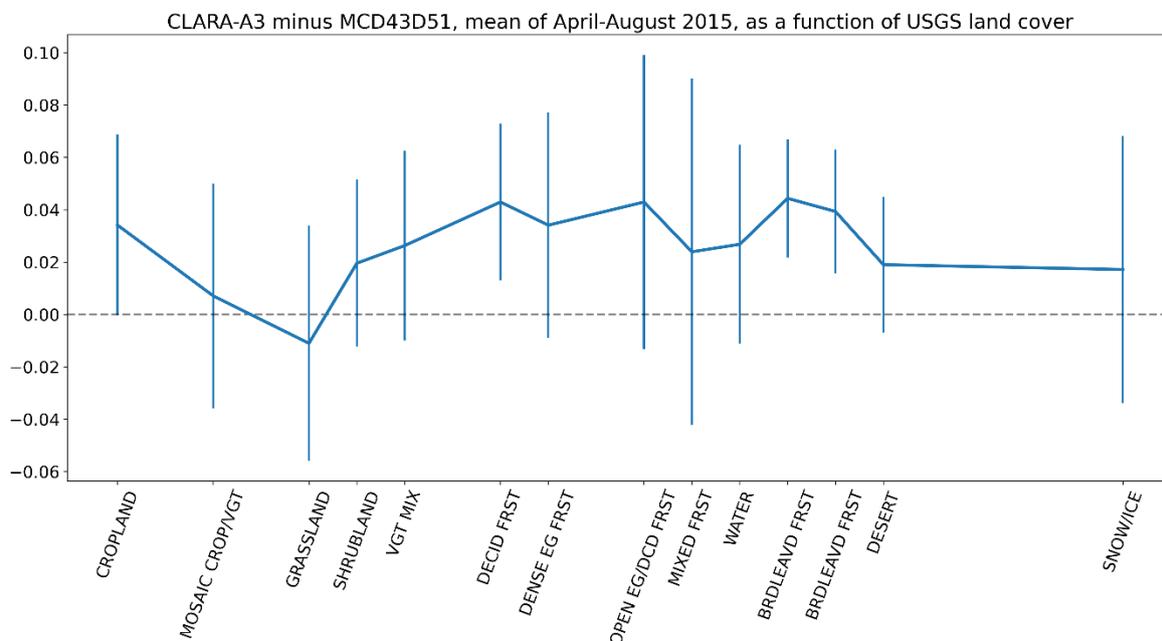
The comparison between MODIS black-sky albedo (MCD43D51) and CLARA-A3 SAL for the April-August 2015 mean is shown in Figure 39. The top row shows both MCD and CLARA mean albedos with a consistent scale – note that MCD has by design no coverage over oceans and sea ice, and its SZA cutoff angle is higher, resulting in more coverage over Antarctica. For this comparison, both MCD and CLARA black-sky albedo has been normalized to correspond with SZA of 60 degrees (Dickinson, 1983). The higher-resolution MCD albedo has been resampled to CLARA grid using bucket resampling.



**Figure 39:** CLARA-A3 SAL compared with MCD43D51, April-August 2015 mean albedo. Top left: MCD43D51 black-sky albedo. Top right: CLARA-A3 SAL black-sky albedo. Bottom left: Spatially resolved difference between CLARA-A3 SAL and MCD43D51. Bottom right: Zonal means of both albedo products over the joint coverage area (valid data in both records). For the comparison, both MCD43D51 and CLARA-A3 SAL have been normalized to SZA 60 degrees using the equation from Dickinson (1983).

The bottom row in Figure 39 shows the difference (CLARA-MCD43) per grid cell (left) and the zonal means of the data records (right). Both records track the zonal means similarly, although with clear differences over deserts and the tropical (rain) forest zone. Conversely, the records appear to agree very well over snow surfaces (Greenland) and the mid-latitude vegetated land surfaces of both hemispheres.

In an effort to further quantify and understand the differences, we classified the differences by land cover of each grid cell using the GLOBCOVER2009 LU classification data. The results are shown in Figure 40. The thick line illustrates the mean difference across land cover classes, whereas the whiskers show the standard deviation of albedo differences for each class. Clearly, cropland, grassland and vegetation mix classes agree within the standard deviation bounds, whereas CLARA retrievals over most forest classes are higher than MCD43. Overall, CLARA-retrieved black-sky albedos are on average  $\sim 0.02$  higher than MCD43 during the evaluation period, though variability for any individual grid cell may be substantially higher.



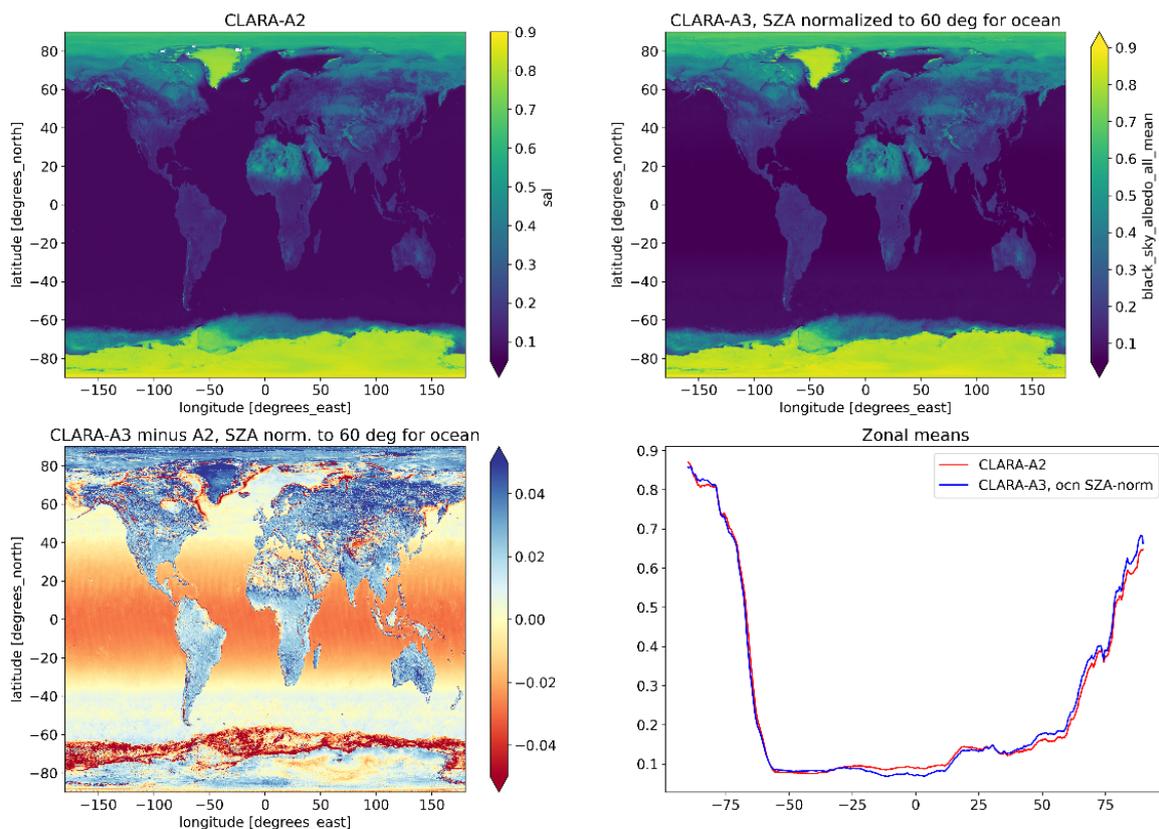
**Figure 40:** Mean differences in black-sky albedo estimates between CLARA-A3 and MCD43D51 shown as a function of USGS land cover classes. Y-axis bars of the classes indicate standard deviation of differences over grid cells belonging to each class (converted from the GLOBCOVER2009 land cover dataset).

In the predecessor record CLARA-A2, ocean albedo was normalized to a fixed SZA of 60 degrees, whereas CLARA-A3 is unnormalized for all parts of the record. To enhance the intercomparison, we performed a first-order normalization to the ocean albedo of CLARA-A3 (SAL) following Dickinson (1983) and based on the mean SZA of valid clear-sky observations in each CLARA-A3. Please note that this procedure should not be expected to result in a perfect replication.

Indeed, while the intercomparison of annual mean black-sky albedo shows generally good agreement between A2 and A3 (Figure 41) over land and snow/ice, residual latitude-dependent differences are apparent in ocean albedo. Interestingly, while the sea ice albedo

over the Arctic Ocean is now somewhat higher in A3 compared to A2 (and better aligned with in situ measurements as seen before), the Antarctic sea ice appears somewhat dimmer than in the predecessor. Given that no algorithmic difference exists for the treatment of sea ice over the two polar cryospheres, it is most likely that changes in cloud detection are the source of the change. Given the lack of in situ measurements of Antarctic sea ice albedo, it is not possible to ascertain if the change is beneficial or not.

2015



**Figure 41:** Comparison of black-sky albedo (SAL) from CLARA-A2 (top left) and CLARA-A3 (top right) data records. Annual means of year 2015. Bottom left: Difference in SAL between A3 and A2. Bottom right: Zonal means of SAL in CLARA-A2 and A3. Ocean surfaces normalized to SZA of 60 degrees in CLARA-A3 to improve comparability.

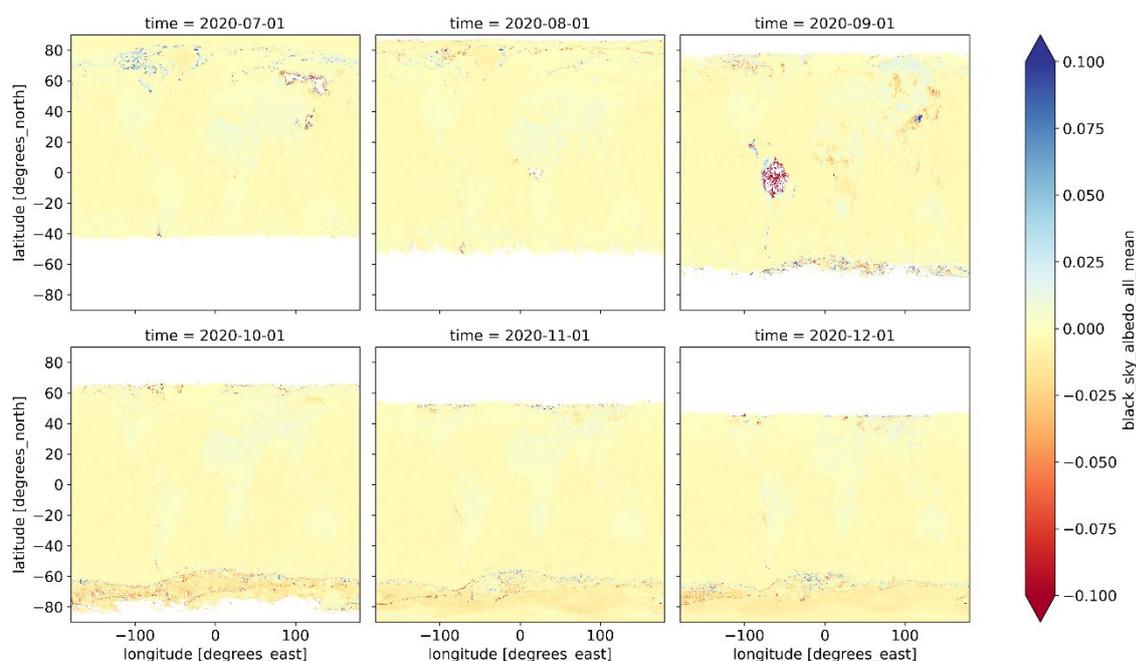
## 10. Analysis of consistency between TCDR and its ICDR continuation

The CLARA-A3 Thematic Climate Data Record (TCDR) covers 1979-2020. The time series is continued through an Interim Climate Data Record (ICDR), delivered continuously with consistent processing methods but with inputs adapted to use continuously available data streams (e.g. replacing ERA5 with ERA5T) or with climatologies in case of inputs where continuously updated data is not available.

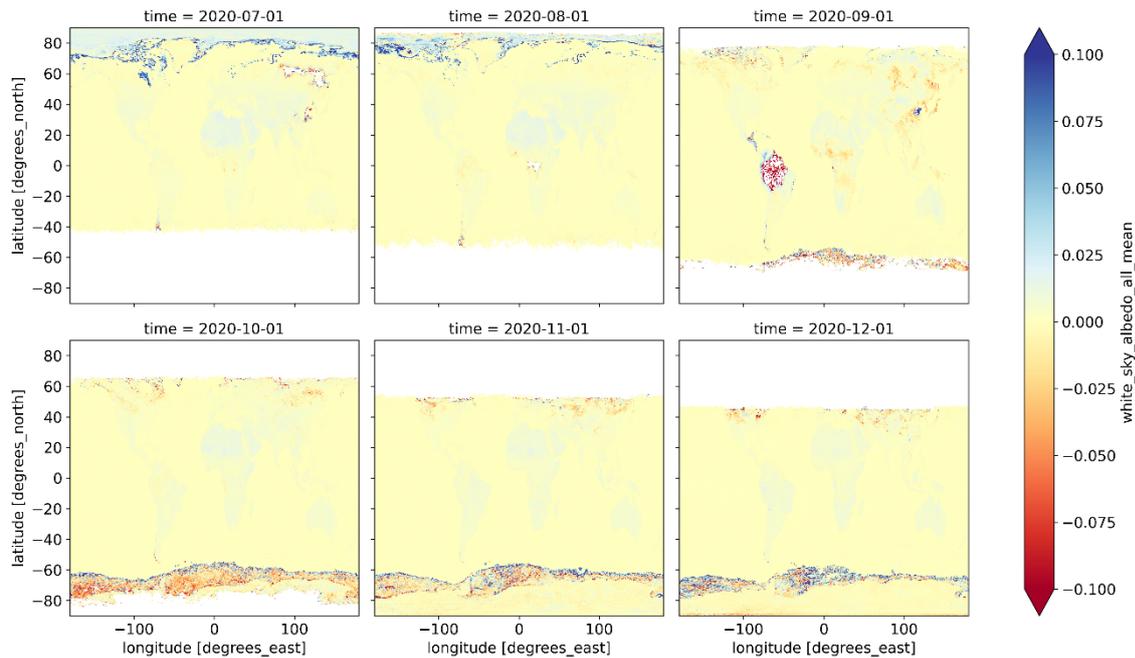
In order to assess consistency between TCDR and ICDR, a six-month period (July-December 2020) was produced both with TCDR and ICDR methods. Here, we assess the monthly means of global gridded surface albedo estimates; as the data are in equal grids, simple differencing (TCDR minus ICDR) is used as the analysis method.

During ICDR preparation it was discovered that the calibration of visible channels for Metop-C is too uncertain to warrant its inclusion in the ICDR processing. The reason appears to be that presently the available Metop-C time series is too short for an effective (vicarious) calibration to remove time decay effects from the data. As a result, Metop-C is not included in the ICDR processing until such time as a robust calibration may be carried out, likely in 2023 or later. Further details are available in e.g. CLARA-A3 Cloud Products PUM.

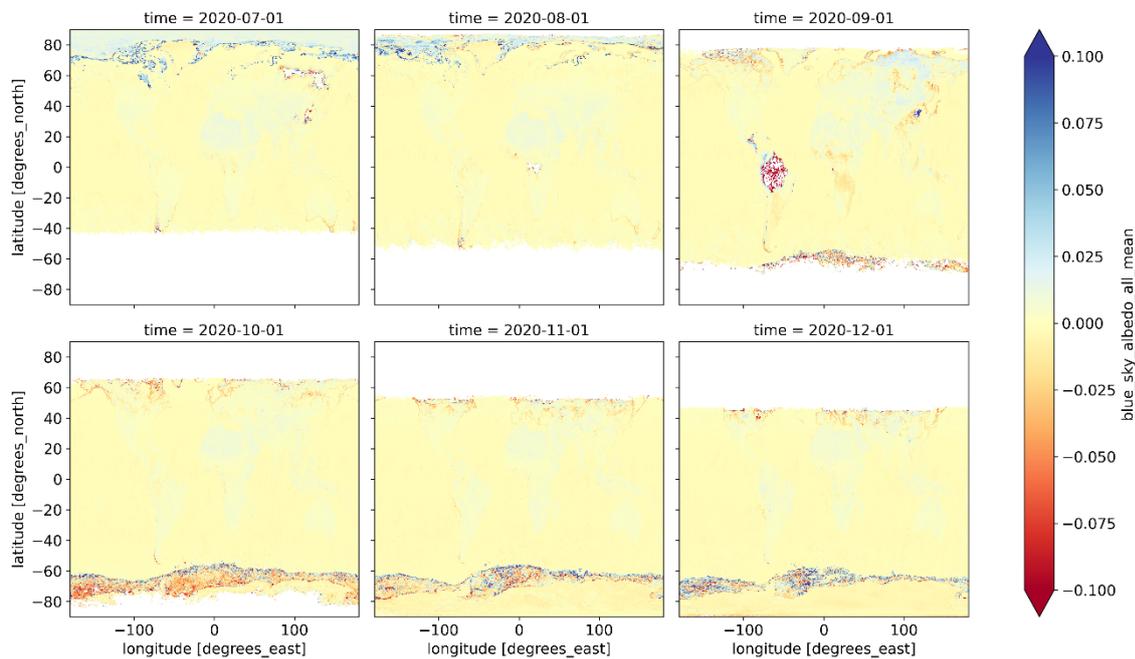
Figure 42 to Figure 44 illustrate the differences on the global grid for the surface albedo estimates per month. Differences over oceans and non-snow land surfaces are generally negligible, though occasional larger differences occur near cutoff conditions for aerosol loading or illumination geometry (i.e. low Sun). These conditions are challenging for the atmospheric correction and sensitive to even minor changes in e.g. atmospheric composition input. Note that the AOD used for both TCDR and ICDR is a climatology for 2020; thus differences must originate elsewhere.



**Figure 42:** Difference (TCDR minus ICDR) for black-sky albedo SAL in the monthly means of July-December 2020.



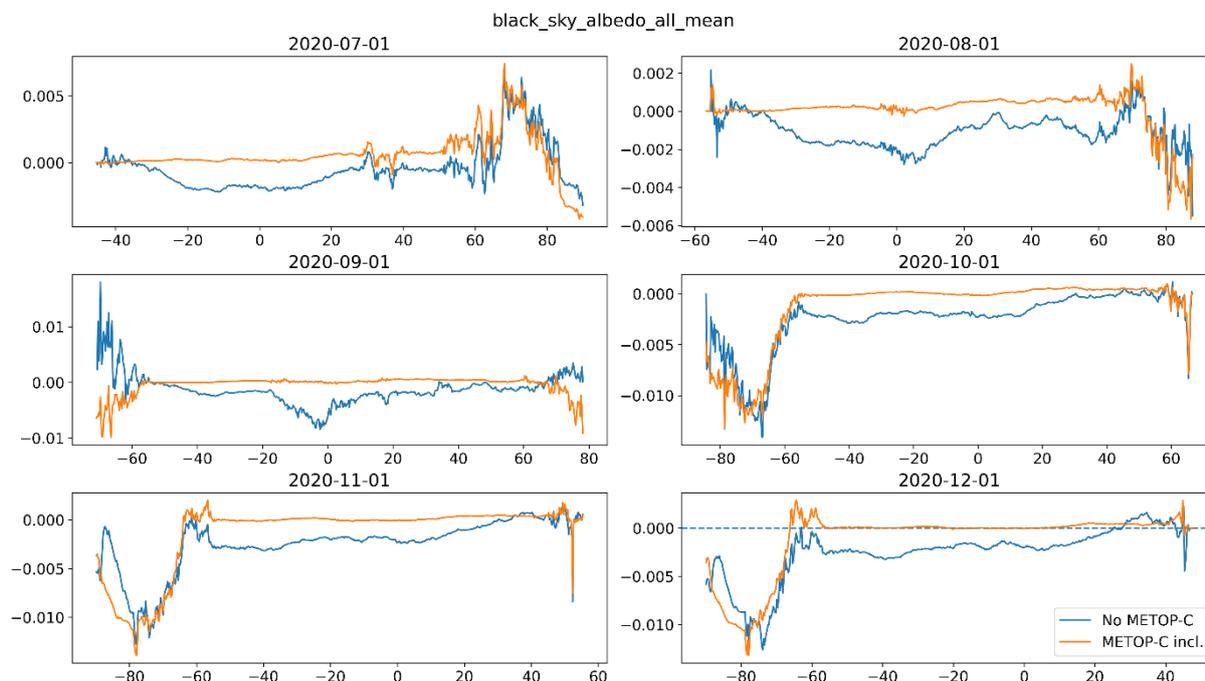
**Figure 43:** Difference (TCDR minus ICDR) for white-sky albedo WAL in the monthly means of July-December 2020.



**Figure 44:** Difference (TCDR minus ICDR) for blue-sky albedo BAL in the monthly means of July-December 2020.

Apart from slight differences over mid-latitude land surfaces differences naturally follow from the exclusion of Metop-C. Furthermore, some systematic, though mostly minor, differences are apparent for snow and sea ice surfaces. An analysis of differences in zonal mean black-sky albedo (Figure 45) following the Metop-C exclusion suggests that calibration differences over (bright) polar regions are the likeliest reason. Without Metop-C, the difference to TCDR is typically smaller over high latitudes, suggesting that a) Metop-C exclusion is warranted from

the consistency viewpoint, and b) ICDR consistency in general is sensitive to the quality of the vicarious calibration and its time dependency. It is recommended and foreseen that the ICDR calibration is updated in the future as more data from the Metop series allows for a more robust calibration.



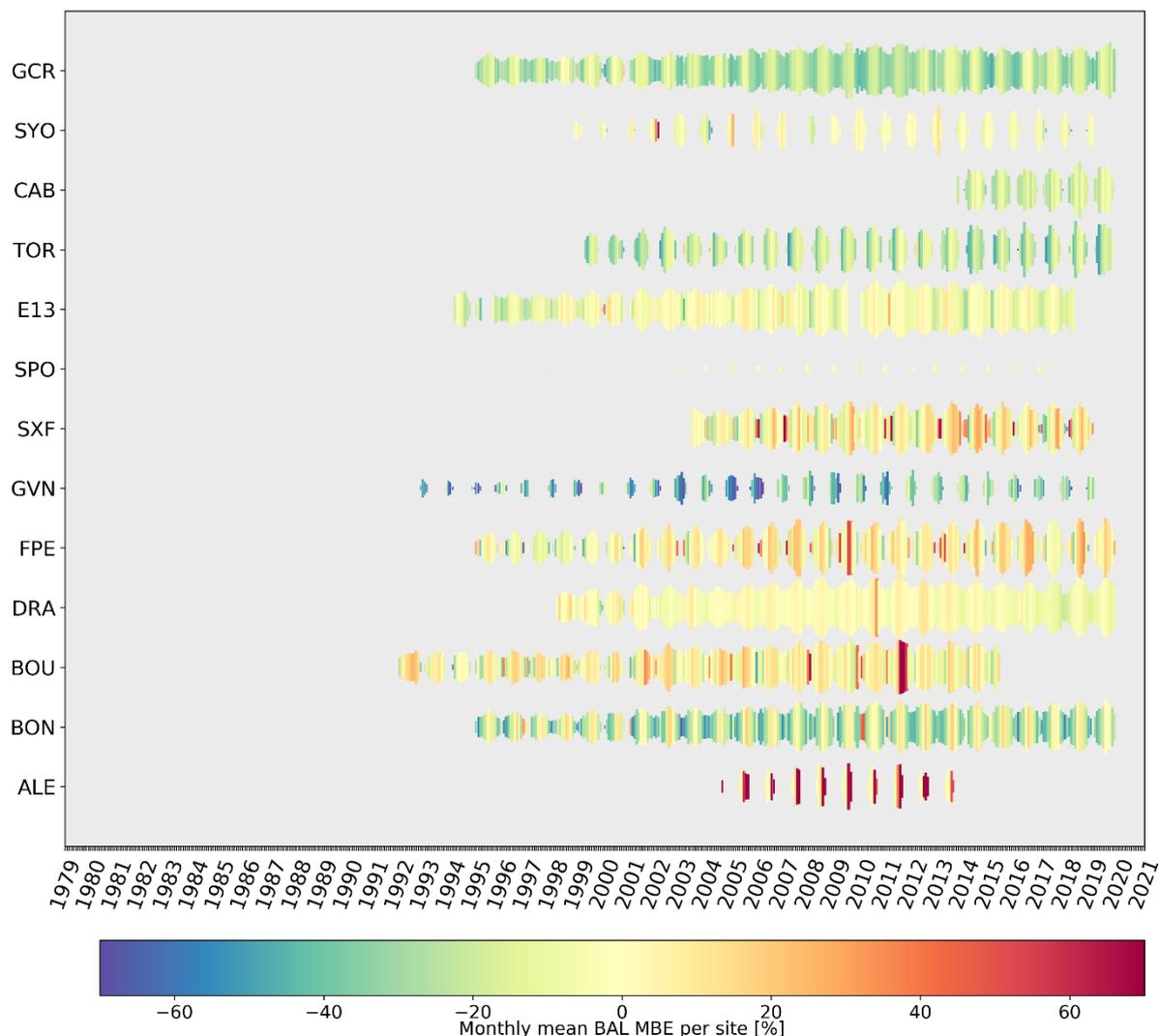
**Figure 45:** Zonal means of TCDR-ICDR differences for black-sky surface albedo in the monthly means of July-December 2020. Blue color is for ICDR with Metop-C excluded (distributed version), orange for Metop-C included in the means.

To summarize, the mean difference between TCDR and ICDR is small in the global scale ( $<0.002$  for SAL/WAL/BAL, albedo estimates in range 0..1), but the standard deviation of all differences is non-negligible for WAL and BAL (stdev of  $\sim 0.013$ ), resulting in a full width at half maximum (FWHM) of  $2.355 * 0.013 \sim 0.03$  (difference distributions appear approximately gaussian). Differences in this range are therefore to be expected to be common over snow and ice surfaces (with larger differences possible but less likely). Given an aged snow surface with an albedo of  $\sim 0.7$ , for example, the additional average uncertainty would be on the order of 4-5%.

## 11. Discussion and conclusions

Here, we reported on a multifaceted effort to validate the latest (third) edition of the surface albedo data records of the CM SAF CLARA product family, now covering 1979-2020. The validation comprised of point-to-pixel comparisons of ground truth measurements of surface albedo versus their satellite-based estimates, as well as intercomparisons with the well-established MODIS surface albedo data as well as the predecessor data record CLARA-A2. Efforts were also made to establish the stability of the data record through examination of retrievals over (quasi-) stationary surface albedo targets, and to examine the spatial representativeness of in situ albedo measurements against their surrounding satellite-observed area.

The principal long-term in situ data source is the Baseline Surface Radiation Network (BSRN). A summary of the observed retrieval biases for blue-sky albedo is shown in Figure 46.



**Figure 46:** Summary of monthly mean BAL retrieval bias against BSRN in situ observations. Color of markers indicates magnitude of bias (MBE), height of each (monthly) marker indicates the relative amount of observations available for the BAL retrieval, i.e. larger bars indicate more available observations. Data based on grid cell scale validation, i.e. point-to-pixel comparisons.

	<b>Validation Report</b> <b>Surface Albedo (SAL)</b> <b>CLARA Edition 3</b>	Doc. No: SAF/CM/FMI/VAL/CLARA/SAL Issue: 3.1 Date: 06.02.2023
---	---	---

The general tendency is that of slight underestimation (-4 to -9% on average; section 5.2), but with substantial variability across sites. This variability is generally consistent with the expected deviations resulting from differing representativeness of the point-like surface measurement for its surrounding area (section 11). This finding is also supported by the lower bias found when examining the individual AVHRR overpass (at GAC resolution of ~5-9 km) retrievals against collocated in situ measurements.

In terms of precision, it is clear that legacy sensors such as AVHRR cannot match the retrieval-to-retrieval precision of more modern instruments even under optimal conditions despite the good results seen during the retrieval stability assessment done here over stable snow/ice surfaces (section 4). Indeed, precision estimates against both BSRN and PROMICE in situ observations display substantial variability even after attempts to account for spatial representativeness effects. A part of this result is due to the nature of the bias-corrected rmse as a metric; large retrieval errors, which occur consistently at a few PROMICE sites in particular, will drive the metric, even though the vast majority of sites by themselves display relatively good precision (Figure 30).

While some of scatter in retrieval error is likely unavoidable, the results illustrate that a part of the future development efforts of CLARA surface albedo data records should be directed towards further improvement of the atmospheric correction of satellite retrievals, examining both methods. through updating the currently used SMAC algorithm (Rahman and Dedieu, 1994), and through updates of the atmospheric composition inputs. Particularly the 0.8 micron band treatment is found to be a significant source of error (section 5.2). Development of overpass-level bidirectional reflectance correction methods for snow and ice is another potential development direction, but that comes with the caveat that sea ice and melting snow are particularly difficult to characterize as surface variability may be great even over small distances.

The temporal stability in mean retrieval errors is very good. Considering that the evaluation here spans 1992-2020 with 39 participating in situ sites (plus two measurement campaigns on Arctic sea ice), the lack of observable trends in bias suggests that the radiance intercalibration of the AVHRR imagery performed as a basis for the CLARA-A3 has been successful.

Comparison with MODIS-based albedo estimates and the predecessor CLARA-A2 suggested good agreement with both (section 9); against MODIS, CLARA-A3 black-sky albedos appear somewhat brighter over forests, in particular over tropics. This is likely connected to the differing atmospheric correction routines and to the different atmospheric compositions used as their input. However, consistency over mid-latitude crop/grassland as well as snow/ice surfaces is good. Against the predecessor CLARA-A2, we note that the tendency for CLARA-A3 is to retrieve slightly higher albedo. As principal algorithms (for black-sky albedo) have not changed, it is likely that the change is related primary to improved cloud detection/screening, and potentially to changes in the intercalibrated AVHRR radiances between A2 and A3. The updated atmospheric composition data source (now ERA5) may also play a role through changes in e.g. atmospheric water vapour content.

Regarding the usage suggestions for the CLARA-A3 surface albedo, we note that quality over the cryospheres remains solid (good stability, biases of <10%) and usage for cryospheric climate studies is warranted; however, care should be taken when using the data over rapidly

	<p style="text-align: center;"><b>Validation Report</b>  <b>Surface Albedo (SAL)</b>  <b>CLARA Edition 3</b></p>	<p>Doc. No: SAF/CM/FMI/VAL/CLARA/SAL  Issue: 3.1  Date: 06.02.2023</p>
---	--	--

varying surfaces, or over partial snow/ice cover near the edges of cryospheric coverage. Screening the data with the available QA data fields (number of observations, stdev, skewness, kurtosis) is important to ensure sufficient quality, and we urge all users to note and make use of this data accordingly. In summary, the evaluation here suggests that the CLARA-A3 surface albedo estimates fulfill their stated performance goals and that no severe data quality problems were discovered. It should be noted, however, that the ICDR extension is presently processed without Metop-C because of calibration issues and that recalibration of the AVHRR imagers in the ICDR is very likely needed in 2023 or later to increase the robustness of the ICDR component and to re-introduce Metop-C.

## References

- Brown, C. F., Brumby, S. P., Guzder-Williams, B., Birch, T., Hyde, S. B., Mazzariello, J., ... & Tait, A. M. (2022). Dynamic World, Near real-time global 10 m land use land cover mapping. *Scientific Data*, 9(1), 1-17.
- Cescatti, A., Marcolla, B., Vannan, S. K. S., Pan, J. Y., Román, M. O., Yang, X., ... & Schaaf, C. B. (2012). Intercomparison of MODIS albedo retrievals and in situ measurements across the global FLUXNET network. *Remote sensing of environment*, 121, 323-334.
- Dickinson, R. E. (1983). Land surface processes and climate—Surface albedos and energy balance. In *Advances in geophysics* (Vol. 25, pp. 305-353). Elsevier.
- Driemel, A., Augustine, J., Behrens, K., Colle, S., Cox, C., Cuevas-Agulló, E., ... & König-Langlo, G. (2018). Baseline Surface Radiation Network (BSRN): structure and data description (1992–2017). *Earth System Science Data*, 10(3), 1491-1501.
- Fausto, R. S., van As, D., Mankoff, K. D., Vandecrux, B., Citterio, M., Ahlstrøm, A. P., ... & Box, J. E. (2021). Programme for Monitoring of the Greenland Ice Sheet (PROMICE) automatic weather station data. *Earth System Science Data*, 13(8), 3819-3845.
- Gascard, J.-C., et al. (2008), Exploring Arctic Transpolar Drift During Dramatic Sea Ice Retreat, *Eos Trans. AGU*, 89(3), 21, doi:10.1029/2008EO030001.
- Holland, M., Serreze, M., and Stroeve, J. (2010). The sea ice mass budget of the arctic and its future change as simulated by coupled climate models. *Climate Dynamics*, 34(2):185–200.
- Hudson, S. R. (2011). Estimating the global radiative impact of the sea ice–albedo feedback in the Arctic, *Journal of Geophysical Research*, 116, D16102, doi:10.1029/2011JD015804.
- Karlsson, K. G., Anttila, K., Trentmann, J., Stengel, M., Meirink, J. F., Devasthale, A., ... & Werscheck, M. (2017). CLARA-A2: CM SAF cLoud, Albedo and surface RAdiation dataset from AVHRR data—Edition 2, Satellite Application Facility on Climate Monitoring, doi: 10.5676 (Vol. 2, p. 2017). EUM\_SAF\_CM/CLARA\_AVHRR.
- Knap, W. H., & Oerlemans, J. (1996). The surface albedo of the Greenland ice sheet: satellite-derived and in situ measurements in the Søndre Strømfjord area during the 1991 melt season. *Journal of Glaciology*, 42(141), 364-374.
- Konzelmann, T., & Ohmura, A. (1995). Radiative fluxes and their impact on the energy balance of the Greenland ice sheet. *Journal of Glaciology*, 41(139), 490-502.
- Loew, A., Bennartz, R., Fell, F., Lattanzio, A., Doutriaux-Boucher, M., & Schulz, J. (2016). A database of global reference sites to support validation of satellite surface albedo datasets (SAVS 1.0). *Earth System Science Data*, 8(2), 425-438.
- Manninen, T., Riihelä, A., and de Leeuw, G. (2012): Atmospheric effect on the ground-based measurements of broadband surface albedo, *Atmos. Meas. Tech. Discuss.*, 5, 385-409, doi:10.5194/amtd-5-385-2012

Perovich, D. K., Grenfell, T. C., Light, B., & Hobbs, P. V. (2002). Seasonal evolution of the albedo of multiyear Arctic sea ice. *Journal of Geophysical Research: Oceans*, 107(C10), SHE-20.

Rahman, H. and Dedieu, G. (1994): SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum. *International Journal of Remote Sensing*, 15, 123-143.

Riihelä, A., Laine, V., Manninen, T., Palo, T., Vihma, T. (2010). Validation of the CM SAF surface broadband albedo product: Comparisons with in situ observations over Greenland and the ice-covered Arctic Ocean. *Remote Sensing of Environment*, 114 (11), 2779-2790.

Riihelä, A., Manninen, T., Laine, V., Andersson, K and Kaspar, F. (2013). CLARA-SAL: a global 28 yr timeseries of Earth's black-sky surface albedo. *Atmospheric Chemistry and Physics*, 13, 3743-3762.

Riihelä, A., King, M. D., & Anttila, K. (2019). The surface albedo of the Greenland Ice Sheet between 1982 and 2015 from the CLARA-A2 dataset and its relationship to the ice sheet's surface mass balance. *The Cryosphere*, 13(10), 2597-2614.

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., ... & Roy, D. (2002). First operational BRDF, albedo nadir reflectance products from MODIS. *Remote sensing of Environment*, 83(1-2), 135-148.

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American statistical association*, 63(324), 1379-1389.

Sütterlin, M., Stöckli, R., Schaaf, C. B., & Wunderle, S. (2016). Albedo climatology for European land surfaces retrieved from AVHRR data (1990–2014) and its spatial and temporal analysis from green-up to vegetation senescence. *Journal of Geophysical Research: Atmospheres*, 121(14), 8156-8171.

Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. *Indagationes mathematicae*, 12(85), 173.

Uttal, T., Curry, J. A., McPhee, M. G., Perovich, D. K., Moritz, R. E., Maslanik, J. A., ... & Grenfeld, T. C. (2002). Surface heat budget of the Arctic Ocean. *Bulletin of the American Meteorological Society*, 83(2), 255-276.

Vihma, T., Jaagus, J., Jakobson, E., and Palo, T. (2008). Meteorological conditions in the Arctic Ocean in spring and summer 2007 as recorded on the drifting ice station Tara, *Geophysical Research Letters*, 35, L18706, doi:10.1029/2008GL034681.

Wang, Z., Schaaf, C. B., Sun, Q., Shuai, Y., & Román, M. O. (2018). Capturing rapid land surface dynamics with Collection V006 MODIS BRDF/NBAR/Albedo (MCD43) products. *Remote sensing of environment*, 207, 50-64.

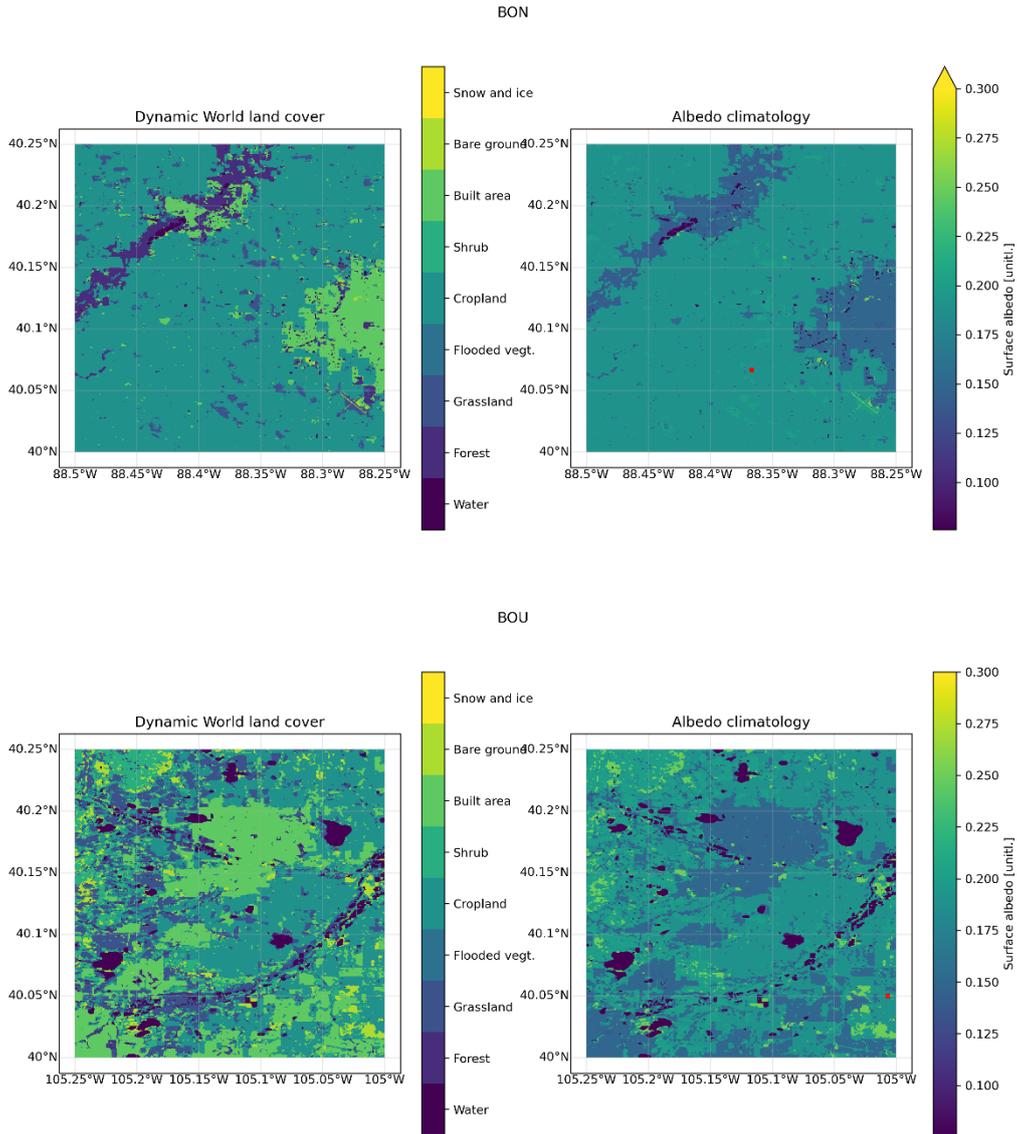
## Glossary

AOD	Aerosol Optical Depth
AVHRR	Advanced Very High Resolution Radiometer (NOAA)
BAL	Surface Albedo (blue-sky)
BB	Broadband
BRDF	Bidirectional Reflectance Distribution Function
BSRN	Baseline Surface Radiation Network
CCI	Climate Change Initiative
CDOP	Continuous Development and Operations Phase
CLARA-A2 SAL CMSAF	cLouds, Albedo and Radiation – Surface Albedo
CM SAF	Satellite Application Facility on Climate Monitoring
DWD	Deutscher Wetterdienst
ECMWF	European Center for Medium-Range Weather Forecasts
ECV	Essential Climate Variable
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FMI	Finnish Meteorological Institute
GCOS	Global Climate Observing System
KNMI	Koninklijk Nederlands Meteorologisch Instituut (Royal Netherlands Meteorological Institute)
LUC	Land Use Classification
LUT	Look-Up Table
MM	Monthly Mean
MODIS	Moderate Resolution Imaging Spectroradiometer
NH	Northern Hemisphere
NOAA	National Oceanic and Atmospheric Administration
NTB (C)	Narrow-to-Broadband (Conversion)
PM	Pentad (5-day) Mean
PPS	Polar Platform System

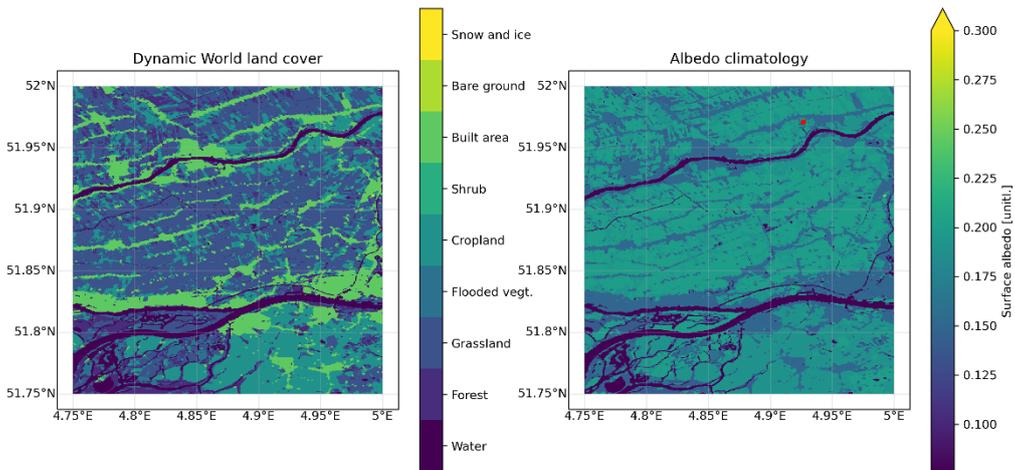
PROMICE	Programme for the Monitoring of the Greenland Ice Sheet
RMIB	Royal Meteorological Institute of Belgium
SAF	Satellite Application Facility
SAL	Surface Albedo (black-sky)
SHEBA	Surface Heat Budget of the Arctic Ocean
SMAC	Simplified method for the atmospheric correction of satellite measurements in the solar spectrum
SMHI	Swedish Meteorological and Hydrological Institute
SZA	Sun Zenith Angle
TOA	Top of Atmosphere
USGS	United States Geological Survey
VZA	Viewing Zenith Angle
WAL	Surface albedo (white-sky)

## Appendix A: Google Dynamic World-inferred land cover and climatological albedo for CLARA-A3 grid cells containing BSRN sites

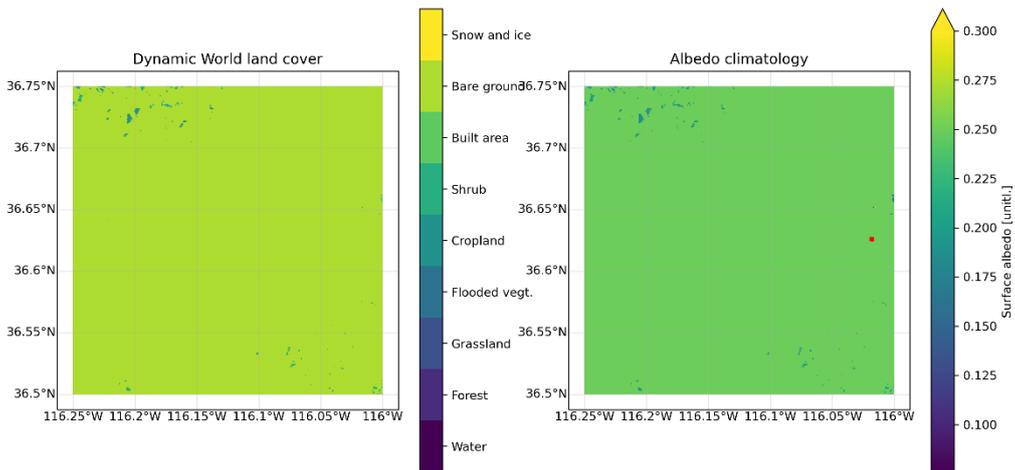
In all images, left subplot shows the Dynamic World data and the right subplot shows the resulting climatological summer albedo map. Location of the BSRN site in question is marked with a red X on the right subplot.



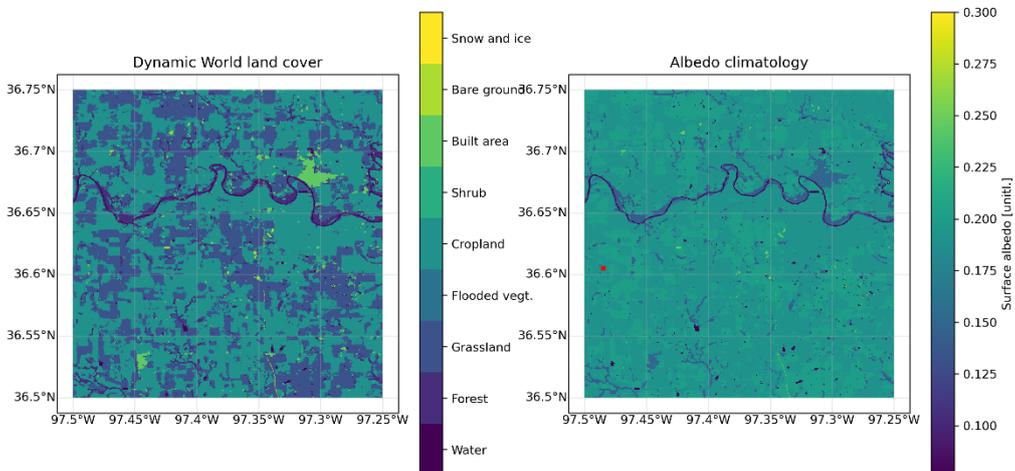
CAB



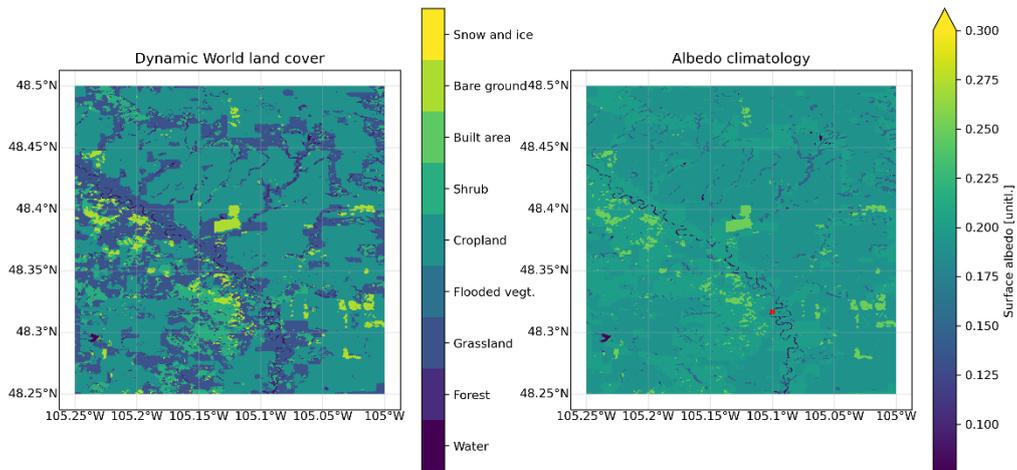
DRA



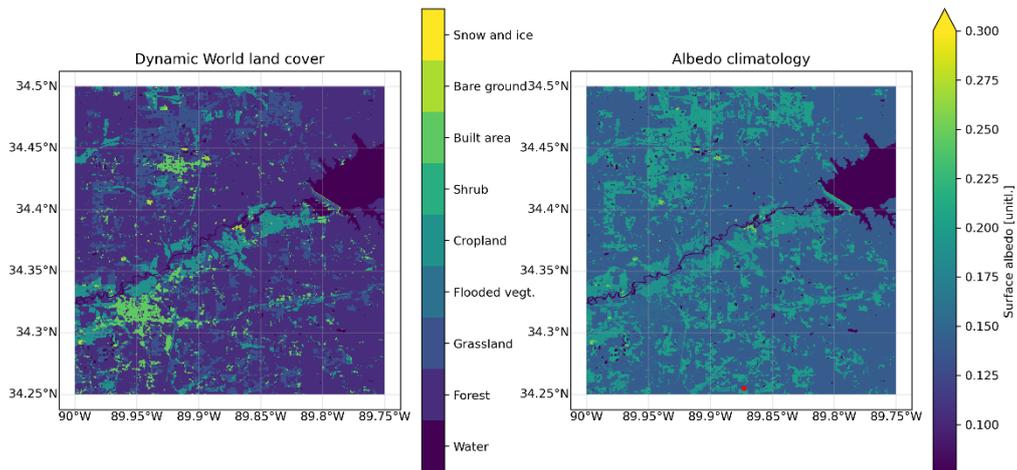
E13



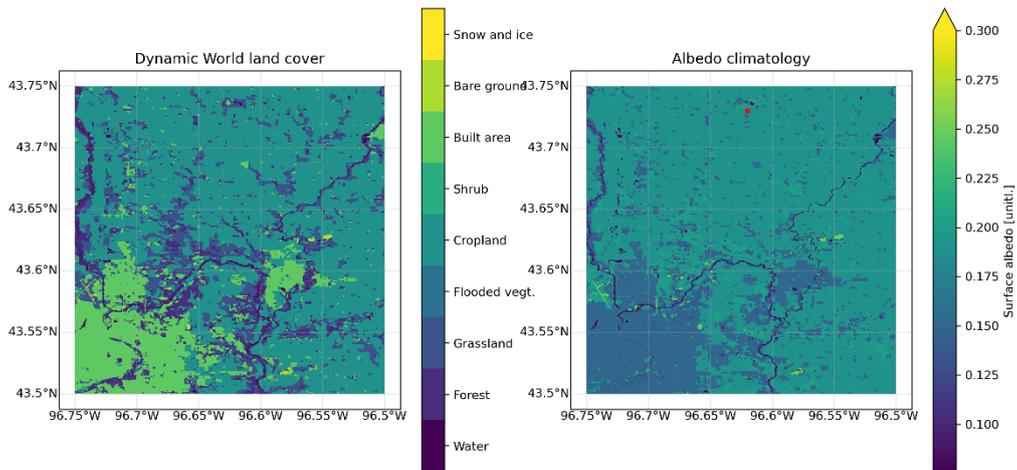
FPE



GCR



SXF



TOR

